



جامعة التنمية البشرية
UNIVERSITY OF HUMAN DEVELOPMENT

p-ISSN 2521-4209
e-ISSN 2521-4217

UHD Journal of Science and Technology

A Scientific periodical issued by University of Human Development

Vol.6 No.(1) June 2022

2022

2722

www.jst.uhd.edu.iq



UHD Journal of Science and Technology

A periodic scientific journal issued by University of Human Development

Editorial Board

Professor Dr. Mariwan Ahmed Rasheed.....	Executive publisher
Assistant Professor Dr. Aso Mohammad Darwesh.....	Editor-in-Chief
Professor Dr. Muzhir Shaban Al-Ani.....	Member
Assistant Professor Dr. Raed Ibraheem Hamed.....	Member
Professor Dr. Salih Ahmed Hama.....	Member
Dr. Nurouldeen Nasih Qader.....	Member

Technical

Mr. Hawkar Omar Majeed.....	Head of Technical
-----------------------------	-------------------

Advisory Board

Professor Dr. Khalid Al-Quradaghi.....	Qatar
Professor Dr. Sufyan Taih Faraj Aljanabi.....	Iraq
Professor Dr. Salah Ismaeel Yahya.....	Kurdistan
Professor Dr. Sattar B. Sadkhan.....	Iraq
Professor Dr. Amir Masoud Rahmani	Kurdistan
Professor Dr. Muhammad Abulaish.....	India
Professor Dr. Parham Moradi	Iran

Introduction

UHD Journal of Science and Technology (UHDJST) is a semi-annual journal published by the University of Human Development, Sulaymaniyah, Kurdistan Region, Iraq. UHDJST member of ROAD, e-ISSN: 2521-4217, p-ISSN: 2521-4209 and a member of Crossref, DOI: 10.21928/issn.2521-4217. UHDJST publishes original research in all areas of Science, Engineering, and Technology. UHDJST is a Peer-Reviewed Open Access journal with Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0). UHDJST provides immediate, worldwide, barrier-free access to the full text of research articles without requiring a subscription to the journal, and has article processing charge (APC). UHDJST applies the highest standards to everything it does and adopts APA citation/referencing style. UHDJST Section Policy includes three types of publications: Articles, Review Articles, and Letters.

By publishing with us, your research will get the coverage and attention it deserves. Open access and continuous online publication mean your work will be published swiftly, ready to be accessed by anyone, anywhere, at any time. Article Level Metrics allow you to follow the conversations your work has started.

UHDJST publishes works from extensive fields including, but not limited to:

- Pure Science
- Applied Science
- Medicine
- Engineering
- Technology

Scope and Focus

UHD Journal of Science and Technology (UHDJST) publishes original research in all areas of Science and Engineering. UHDJST is a semi-annual journal published by the University of Human Development, Sulaymaniyah, Kurdistan Region, Iraq. We believe that if your research is scientifically valid and technically sound then it deserves to be published and made accessible to the research community. UHDJST aims to provide a service to the international scientific community enhancing swap space to share, promote and disseminate the academic scientific production from research applied to Science, Engineering, and Technology.

SEARCHING FOR PLAGIARISM

We use plagiarism detection: detection; According to Oxford online dictionary, Plagiarism means: *The practice of taking someone else's work or ideas and passing them off as one's own.*

Section Policies

No.	Title	Peer Reviewed	Indexed	Open Submission
1	Articles: This is the main type of publication that UHDJST will produce	✓	✓	✓
2	Review Articles: Critical, constructive analysis of the literature in a specific field through summary, classification, analysis, comparison.	✓	✓	✓
3	Letters: Short reports of original research focused on an outstanding finding whose importance means that it will be of interest to scientists in other fields.	✓	✓	✓

PEER REVIEW POLICIES

At UHDJST we are committed to prompt quality scientific work with local and global impacts. To maintain a high-quality publication, all submissions undergo a rigorous review process. Characteristics of the peer review process are as follows:

- The journal peer review process is a "double-blind peer review".
- Simultaneous submissions of the same manuscript to different journals will not be tolerated.
- Manuscripts with contents outside the scope will not be considered for review.
- Papers will be refereed by at least 2 experts as suggested by the editorial board.
- In addition, Editors will have the option of seeking additional reviews when needed. Authors will be informed when Editors decide further review is required.
- All publication decisions are made by the journal's Editors-in-Chief on the basis of the referees' reports. Authors of papers that are not accepted are notified promptly.
- All submitted manuscripts are treated as confidential documents. We expect our Board of Reviewing Editors, Associate Editors and reviewers to treat manuscripts as confidential material as well.
- Editors, Associate Editors, and reviewers involved in the review process should disclose conflicts of interest resulting from direct competitive, collaborative, or other relationships with any of the authors, and remove oneself from cases in which such conflicts preclude an objective evaluation. Privileged information or ideas that are obtained through peer review must not be used for competitive gain.
- Our peer review process is confidential and the identities of reviewers cannot be revealed.

Note: UHDJST is a member of CrossRef and CrossRef services, e.g., CrossCheck. All manuscripts submitted will be checked for plagiarism (copying text or results from other sources) and self-plagiarism (duplicating substantial parts of authors' own published work without giving the appropriate references) using the CrossCheck database. Plagiarism is not tolerated.

For more information about CrossCheck/iThenticate, please visit

<http://www.crossref.org/crosscheck.html>.

OPEN ACCESS POLICY

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge. Open Access (OA) stands for unrestricted access and unrestricted reuse which means making research publications freely available online. It access ensures that your work reaches the widest possible audience and that your fellow researchers can use and share it easily. The mission of the UHDJST is to improve the culture of scientific publications by supporting bright minds in science and public engagement.

UHDJST's open access articles are published under a Creative Commons Attribution CC-BY-NC-ND 4.0 license. This license lets you retain copyright and others may not use the material for commercial purposes. Commercial use is one primarily intended for commercial advantage or monetary compensation. If others remix, transform or build upon the material, they may not distribute the modified material. The main output of research, in general, is new ideas and knowledge, which the UHDJST peer-review policy allows publishing as high-quality, peer-reviewed research articles. The UHDJST believes that maximizing the distribution of these publications - by providing free, online access - is the most effective way of ensuring that the research we fund can be accessed, read and built upon. In turn, this will foster a richer research culture and cultivate good research ethics as well. The UHDJST, therefore, supports unrestricted access to the published materials on its main website as a fundamental part of its mission and a global academic community benefit to be encouraged wherever possible.

Specifically:

- The University of Human Development supports the principles and objectives of Open Access and Open Science
- UHDJST expects authors of research papers, and manuscripts to maximize the opportunities to make their results available for free access on its final peer-reviewed paper
- All manuscript will be made open access online soon after final stage peer-review finalized.
- This policy will be effective from 17th May 2017 and will be reviewed during the first year of operation.
- Open Access route is available at <http://journals.uhd.edu.iq/index.php/uhdjst> for publishing and archiving all accepted papers,
- Specific details of how authors of research articles are required to comply with this policy can be found in the Guide to Authors.

ARCHIVING

This journal utilizes the LOCKSS and CLOCKSS systems to create a distributed archiving system among participating libraries and permits those libraries to create permanent archives of the journal for purposes of preservation and restoration.

LOCKSS: Open Journal Systems supports the LOCKSS (Lots of Copies Keep Stuff Safe) system to ensure a secure and permanent archive for the journal. LOCKSS is open source software developed at Stanford University Library that enables libraries to preserve selected web journals by regularly polling registered journal websites for newly published content and archiving it. Each archive is continually validated against other library caches, and if the content is found to be corrupted or lost, the other caches or the journal is used to restore it.

CLOCKSS: Open Journal Systems also supports the CLOCKSS (Controlled Lots of Copies Keep Stuff Safe) system to ensure a secure and permanent archive for the journal. CLOCKSS is based upon the open-source LOCKSS software developed at Stanford University Library that enables libraries to preserve selected web journals by regularly polling registered journal websites for newly published content and archiving it. Each archive is continually validated against other library caches, and if the content is found to be corrupted or lost, the other caches or the journal is used to restore it.

PUBLICATION ETHICS

Publication Ethics and Publication Malpractice Statement

The publication of an article in the peer-reviewed journal UHJST is to support the standard and respected knowledge transfer network. Our publication ethics and publication malpractice statement is mainly based on the Code of Conduct and Best-Practice Guidelines for Journal Editors (Committee on Publication Ethics, 2011) that includes;

- General duties and responsibilities of editors.
- Relations with readers.
- Relations with the authors.
- Relations with editors.
- Relations with editorial board members.
- Relations with journal owners and publishers.
- Editorial and peer review processes.
- Protecting individual data.
- Encouraging ethical research (e.g. research involving humans or animals).
- Dealing with possible misconduct.
- Ensuring the integrity of the academic record.
- Intellectual property.
- Encouraging debate.
- Complaints.
- Conflicts of interest.

ANIMAL RESEARCHES

- For research conducted on regulated animals (which includes all live vertebrates and/or higher invertebrates), appropriate approval must have been obtained according to either international or local laws and regulations. Before conducting the research, approval must have been obtained from the relevant body (in most cases an Institutional Review Board, or Ethics Committee). The authors must provide an ethics statement as part of their Methods section detailing full information as to their approval (including the name of the granting organization, and the approval reference numbers). If an approval reference number is not provided, written approval must be provided as a confidential supplemental information file. Research on non-human primates is subject to specific guidelines from the Weather all (2006) report (The Use of Non-Human Primates in Research).
- For research conducted on non-regulated animals, a statement should be made as to why ethical approval was not required.
- Experimental animals should have been handled according to the highest standards dictated by the author's institution.
- We strongly encourage all authors to comply with the '*Animal Research: Reporting In Vivo Experiments*' (ARRIVE) guidelines, developed by NC3Rs.
- Articles should be specific in descriptions of the organism(s) used in the study. The description should indicate strain names when known.

ARTICLE PROCESSING CHARGES

UHDJST is an Open Access Journal (OAJ) and has article processing charges (APCs). The published articles can be downloaded freely without a barrier of admission.

Address

University of Human Development, Sulaymaniyah-Kurdistan Region/Iraq
PO Box: Sulaymaniyah 6/0778

Contact

Principal Contact

Dr. Aso Darwesh

Editor-in-Chief

University of Human Development –
Sulaymaniyah, Iraq

Phone: +964 770 148 5879

Email: jst@uhd.edu.iq

Support Contact

UHD Technical Support

Phone: +964 770 247 3391

Email: jst@uhd.edu.iq

Contents

No.	Author Name	Title	Pages
1	Kaziwa Ahmad Kaka Alla Salih Ahmed Hama	Seroprevalence and Molecular Detection of Influenza A Virus (H1N1) in Sulaimani Governorate-Iraq	1-6
2	Mohammed Ibrahim Khalil	Dissimilarity in the Nuclear Ribosomal DNA Internal Transcribed Spacer Regions of <i>Haplophyllum</i> spp. Founded in Ashdagh Mountain, Sangaw, Kurdistan of Iraq	7-11
3	Bayan Omar Mohammed Hamsa D. Majeed Siti Zaiton Mohd Hashim Muzhir Shaban Al-Ani	New Feature-level Algorithm for a Face-fingerprint Integral Multi-biometrics Identification System	12-20
4	Shakhawan Hares Wady	Classification of Acute Lymphoblastic Leukemia through the Fusion of Local Descriptors	21-33
5	Safa Abdul Wahab Hameed Alaa Badeea Ali	An Intelligent and Precise Method Used for Detecting Gestational Diabetes in the Early Stages	34-42
6	Tara Qadir Kaka Muhammad Hawar Othman Sharif Mazen Ismaeel Ghareb	Eye Tracking Technique for Controlling Computer Game Objects	43-51
7	Nigar M. Shafiq Surameery	Modified Advanced Encryption Standard for Boost Image Encryption	52-59
8	Goran Saman Nariman Hamsa D. Majeed	Adaptive Filter based on Absolute Average Error Adaptive Algorithm for Modeling System	60-69
9	Mustafa N. Rashad Dana L. Hussein Haval D. Abdalkarim Ribwar R. Azeez	Future IoT Software in Healthcare Also Exploring IoT Industry Application	70-75
10	Sardar M. Weli Sabiha M. Salih Abdullah A. Hama Ary B. Faiq Fatimah M. Ali	Determination of Potassium Bromate in Bread Brands in Sulaimani City, Kurdistan-Iraq	76-79

Seroprevalence and Molecular Detection of Influenza A Virus (H1N1) in Sulaimani Governorate-Iraq



Kaziwa Ahmad Kaka Alla¹, Salih Ahmed Hama^{1,2}

¹Department of Biology, College of Science, University of Sulaimani, Kurdistan Region, Sulaymaniyah, Iraq, ²Department of Medical Laboratory Science, College of Health Sciences, University of Human Development, Kurdistan Region, Sulaymaniyah, Iraq

ABSTRACT

Influenza A (H1N1) virus is now rapidly scattering across the world. Early detection is one of the most effective measures to stop the further spread of the virus. The current study was aimed to detect influenza A (H1N1) serologically and by polymerase chain reaction (PCR) techniques. From September 2020 to June 2021, three hundred nasopharyngeal swabs and blood samples were collected from Hiwa and Shahid Tahir Hospitals in Sulaimani city. Obtained results revealed that 23.3% of the tested patients were seropositive anti-IgG for Influenza A, while 13.3% showed anti-IgM seropositive results although 10% of the tested cases were with both anti-IgG and anti-IgM seropositive results. Gender, residency, and flu symptoms showed no significant relations with seropositive results ($p < 0.05$) whereas valuable relations were found between seropositive observations and smoking, the previous history of chronic diseases as well as employment status ($p < 0.05$). It was concluded that hematologic investigations (CBC) were not dependable if H1N1 diagnosis and detection. Only 1% of the tested samples showed positive results for influenza A (H1N1) RNA using reverse transcription-PCR.

Index Terms: Influenza A, H1N1, Anti-IgG, Anti-IgM, Reverse transcription-polymerase chain reaction, ssRNA

1. INTRODUCTION

Influenza viruses which are enveloped ssRNA viruses can cause annual epidemics and pandemics with serious consequences for public health and the global economy, assessed with 1 billion cases, including 3-5 million severe cases, and 290 000-650 000 influenza-related respiratory deaths worldwide [1]. Influenza A virus (IAV) is due to the family Orthomyxoviridae which possess a segmented, single-stranded, negative-sense RNA genome. This family consists of five genera: Influenzavirus A, B, and C, Togavirus [2]. The virus

is with a pleomorphic morphology, characterized by spherical, elongated, or filamentous particles [3]. In 2009 a pandemic influenza infection was caused by a subtype known as swine flu (H1N1) virus with genes that originate from human and avian influenza virus [4]. Humans can be infected with H1N1, H1N2, or H3N2 through direct contact with infected animals or contaminated surroundings. The pandemic strain contains genes from four different flu viruses including two swine strains, one human strain, and one avian [5]. Enveloped viruses have a matrix that interacted with the viral glycoproteins and nucleocapsid that can play an essential role in the gathering of the viral proteins and budding of the progeny virions [6]. Novel re-assorted influenza H1N1 virus produced by reassortment between the viral genome segments and it was behind the pandemic H1N1 in 2009 [7]. During the past 100 years, five pandemic influenza outbreaks have occurred Spanish flu (H1N1) in 1918, Asian flu (H2N2) in 1957, Hong Kong flu (H3N2) in 1968, Russian flu (H1N1) 1977, and swine flu

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp1-6

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Copyright © 2022 Alla and Hama This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: kaziwaahmad91@gmail.com

Received: 14-08-2021

Accepted: 28-12-2021

Published: 02-01-2022

(H1N1) in 2009. In particular, the 1918 influenza pandemic affected almost 30% of the global population and is believed to have killed over 50 million people [8]. Multiple one-step real-time reverse transcription-polymerase chain reaction (RT-PCR) assays can simultaneously detect and discriminate Flu A subtypes with dependable sensitivity and specificity, which is required for the early clinical diagnosis and viral surveillance of patients with Flu A infection [9].

Serological techniques commonly can be depended on for detection of Influenza A infections through anti-influenza Immunoglobulin G (IgG_ and IgM detection by ELISA technique, especially IgG and IgM against hemagglutinin [10]. The aims of the current study are; serologic detection of Influenza anti-IgG and anti-IgM and molecular detection of Influenza RNA using RT-PCR.

2. MATERIALS AND METHODS

2.1. Study Population

The study's population included people visiting Hiwa and Shahid Tahir Hospitals in Sulaimani city from September 2020 to June 2021, difficulties were found during sample collection due to the negative view of the patients. All tested patients were suffered from flue signs and symptoms, including fever, chills, cough, muscle or body aches, runny or stuffy nose, sneeze, headaches, fatigue, sore throat, and sweating. The sample size was 300 patients included 163 males and 137 females.

2.2. Sample Collection

From each tested patient nasopharyngeal swabs were collected as well as 5 mL fresh venous blood was taken aseptically and divided into two parts; one for serum preparation and the rest for hematologic investigations. The collected samples were stored according to their uses as following; the blood samples were stored in 4°C till hematological investigations were done. The serum samples were divided into two parts; one for serology and stored in -20°C, while the other part of the separated serum was stored in -80°C (for molecular tests).

2.3. Anti-Influenza Virus Antibody Detection by ELISA

Indirect-ELISA method was depended to detect anti-Influenza virus A antibody IgG and IgM using a special ELISA kit (CUSABIO/Whan-China, Elab-Science/Korea, NovaLisa®/Germany). The microtiter plate wells were precoated with recombinant influenza antigens. All preserved sera samples were transferred to room temperature for about 30 min. 100 µL of each diluted sample, standard, and blank were added to the desired wells for IgM (for IgG 200 µL of diluted sample was added). The plate was incubated for 30 min but (an hour for IgM

according to supplied company instructions) at 37°C in shading light. The process of washing and aspirating of each well with 350 µL washing buffer was done five times for IgM using ELISA washer (for IgG 300 µL of washing buffer was used four times as directed by the supplied company). About 100 µL of HRP conjugate was added to each well except the blank and incubated for 30 min at 37°C in shading light. The process of washing was repeated and aspirated five times for IgM and four times for IgG. To each well, 50 µL of the substrate reagent A and 50 µL substrate reagent B was added and mixed, then incubated for 15 min at 37°C in shading light for both IgG and IgM. For each well, 50 µL of stop solution was added and the Optical Density was measured at 450 nm and 620 nm for IgG and IgM.

2.4. Viral RNA Extraction and Amplification

The RNA extraction was performed according to the manufacturer's protocol included in AddPrep Viral Nucleic Acid Extraction Kit (Add Bio-Tech, Korea). AddPrep Viral Nucleic Acid Extraction Kit (Add Bio-Tech/Korea) buffer system provides the effective binding condition of RNA to the microfiber-silica-based membrane through the mix with lysis and binding buffers, and then the impurities on the membrane are washed away by two different washing buffers. Starting with a 200 µL of swab sample to 1.2 tubes and followed spin column purification with final elute of 150 µL RNA. Extracted viral nucleic acid was stored at -80°C until the day of examination.

2.5. PCR Reaction

A total volume of Master Mix AddScript RT-PCR Nuclease-free (D.W), Forwarding primer, Reverse primer, and nasopharyngeal swab fluid/standard/Negative/Positive Control was prepared as directed by the supplied company. The process of PCR programming for detecting IVA nucleic acid was performed starting with the reverse transcription step, denaturation, renaturation, annealing, elongation, and the data were collected.

Step	Temperature (°C)	Duration	Cycle
cDNA synthesis	50	30 min	1
initial denaturation	95	10 min	
Denaturation,	95	15–30 s	
Annealing,	55–65	15–30 s	
Extention	72	1 min	
and final extension	72	5 min	35

Item	Volume
Nuclease-free (D.W)	5 µl
Forwarding primer	1 µl
Reverse primer	1 µl
Nasopharyngeal swab fluid/Standard/Negative/	3 µl
Positive Control	10 µl
2x Master mix AddScript RT-PCR	
Total volume	20 µl

2.6. Primers and Probes

Subtype	Oligo	SEQ
Influenza A (H1N1) pdm09	H1F1	5' - AGCAAAAGCAGGGGAAAATAAAAGC - 3' (25mer)
	H1R1264	5' - CCTACTGCTGTGAACTGTGTATTC - 3' (24mer)
	H1F848	5' - GCAATGCAAAGAAATGCTGGATCTG - 3' (25mer)
	H1RUc	5' - ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT - 3' (35mer)
	N1F1	5' - AGCAAAAGCAGGAGTTTAAATG - 3' (23mer)
	N1R1099	5' - CCTATCCAAACACCATTGCCGTAT - 3' (24mer)
	N1F401	5' - GGAATGCAGAACCTTCTTCTTGAC - 3' (24mer)
	NARUc	5' - ATATGGTCTCGTATTAGTAGAAACAAGGAGTTTTTT - 3' (36mer)
	NAFUc	5' - TATTGGTCTCAGGGAGCAAAAGCAGGAGT - 3' (29mer)
	MF1	5' - AGCAAAAGCAGGTAGATATTGAAAGA - 3' (26mer)
	MR1027	5' - AGTAGAAACAAGGTAGTTTTTACTC - 3' (26mer)
	NAFUc	5' - TATTGGTCTCAGGGAGCAAAAGCAGGAGT - 3' (29mer)
	H3N2R109	5' - TCATTTCCATCATCRAAGGCCCA - 3' (23mer)
Influenza A (H3N2)	N2F387	5' - CATGCGATCCTGACAAGTGTTATC - 3' (24mer)
	NARUc	5' - ATATGGTCTCGTATTAGTAGAAACAAGGAGTTTTTT - 3' (36mer)
	H3A1F6	5' - AAGCAGGGGATAATTCTATTAACC - 3' (24mer)
	H3A1R1	5' - GTCTATCATTCCCTCCCAACCATT - 3' (24mer)
	H3A1F3	5' - GTCTATCATTCCCTCCCAACCATT - 3' (24mer)
	HARUc	5' - ATATCGTCTCGTATTAGTAGTAGAAACAAGGGTGTTTT - 3' (35mer)

3. RESULTS

Both sexes were included in the current study, out of 300 participants (163 males and 137 females), (71, 23.7%) showed seropositive results for anti-H1N1 IgG, respectively, considering the gender (Table 1). Seropositive observations considering anti-H1N1 IgM showed lower positive results comparing to anti-H1N1 IgG. It was noticed that (40, 13.33%) cases were seropositive for anti-H1N1 IgM among males and females (Table 1). It was noticed that some tested cases were seropositive for both anti-H1N1 IgG and IgM at the same time (10%) (Fig. 1).

The percentage of seropositive results among males was relatively higher (56.3%) than among females (43.7%), although there were significant differences considering the gender regarding anti-H1N1 IgG ($P < 0.05$) (Fig. 2). As in the case of IgM results, the seropositive results were higher among males (55%) when compared with females (45%). Statistical analysis showed significant differences between males and females considering anti-H1N1 IgM ($P < 0.05$) (Fig. 2). From this ratio, 55% were among males and 45% among females. Furthermore, the percentage of IgG and IgM among males (11.7%) was higher to compare females (7.8%) (Fig. 2).

The PCR positive result was among seropositive males (0.67%) only, while the seropositive females showed negative PCR results (Fig. 3).

When the relationships of certain risk factors were evaluated on the seropositive observations, it appeared that gender has significant effects on the H1N1 seropositive results considering H1N1 anti-IgG, anti-IgM ($P < 0.05$) (Table 1). Moreover, As mentioned in the methodology, some of the cases were symptomatic others were asymptomatic, so depending on the presence of flu syndrome, it appeared that the occurrence of flu symptoms have significant relations with the obtained seropositive results ($P < 0.05$) which indicates that the symptoms are dependable in H1N1 diagnosis (Table 1). Studying the effects of residency indicated that it has no significant effects on the percentage of seropositive results ($P > 0.05$) (Table 1). In Addition to these factors, the effects of smoking also were evaluated, it was noticed that smoking has significant effects on the results ($P < 0.05$), so smoking can be considered as a risk factor for H1N1 infections (Table 1). Similarly, both previous history of chronic diseases and employment can be strongly related with observations recorded in the current study ($P < 0.05$) for both factors (Table 1).

Depending on the complete blood count (CBC) picture done for all studied cases, it was concluded that no valuable changes were seen between seropositive cases and negative ones ($P > 0.05$). As well as comparison of the calculated hematologic parameters with the normal ranges from reference textbooks clarified that no significant abnormal (elevation and decline) of these parameters were recorded although slight changes or elevations in some parameters were seen, but were non-significant ($P > 0.05$) (Table 2).

TABLE 1: Evaluation of relations between some risk factors and H1N1 seropositive results

Variables		Anti-IgG positive (No, %)	Anti-IgM positive (No, %)	Anti-IgG, IgM positive (No, %)	P-value
Gender	Males	40 (13.33)	22 (07.33)	18 (06.00)	P<0.05
	Females	31 (10.33)	18 (06.00)	12 (04.00)	
Flu symptoms	Yes	37 (12.33)	23 (07.67)	20 (06.67)	P>0.05
	No	34 (11.33)	17 (05.67)	10 (03.33)	
Residency	Urban	33 (11.00)	19 (06.33)	12 (04.00)	P<0.05
	Rural	38 (12.67)	21 (07.00)	18 (06.00)	
Smoking	Smoker	33 (11.00)	29 (09.67)	21 (07.00)	P<0.05
	Non-smoker	18 (6.000)	11 (03.67)	09 (03.00)	
Chronic diseases	Yes	40 (13.33)	22 (07.33)	18 (06.00)	
	No	31 (10.33)	18 (06.00)	12 (04.00)	
Employed	Yes	47 (15.67)	31 (10.33)	22 (07.33)	
	No	24 (08.00)	09 (03.00)	08 (02.67)	

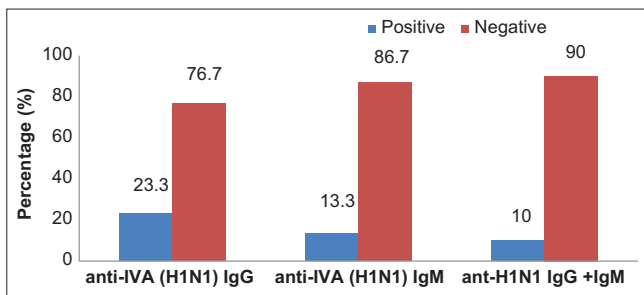


Fig. 1. Seropositive results of IVA (H1N1) among tested patients.

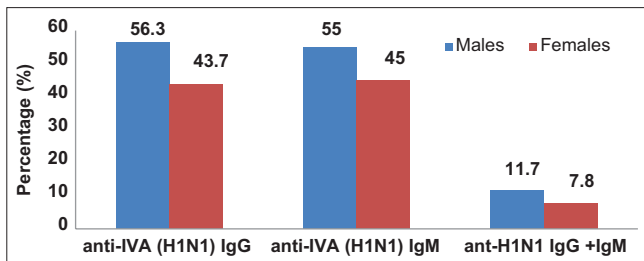


Fig. 2. Seropositive results of ant-H1N1 IgG and IgM among males and females.

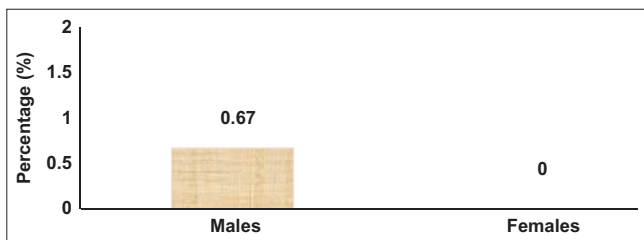


Fig. 3. Positive PCR results among males and females.

4. DISCUSSIONS

Life-threatening infection by influenza A virus stays behind health complaints and death worldwide. It was estimated around the world that seasonal influenza can cause about

3–5 million cases of severe illness, and about 290,000–650,000 respiratory deaths worldwide each year [11]. Certain factors may explain the low percentage rates of RT-PCR results from the current study; among them, the limited numbers of the samples, technical errors as well as high sensitivity of the viral RNA for degradation by enzymes and environmental factors, since most of the analyzed samples were previously collected and preserved in the specified hospitals. The relatively high seropositivity rates (23.3%) of INF-A (H1N1) virus infection among the studied cases in the current study can be explained, especially since the vast majority of patients were with a previous history of flu infection. They were suspected of having an influenza virus infection. On the other hand, most of the studied cases were from Cancer treatment centers and suffering from immunologic complaints, and were at high risk for different infections including influenza. Several studies and investigators reported a higher prevalence of influenza A virus infections than our observation. In a previous study, it was reported that the prevalence of influenza A virus seropositivity (anti-IgG and anti-IgM) was relatively higher than the current results [12]-[14]. Whereas the current results were in agreement with conclusions reported by other investigators [15]. It was reported that some factors were significantly effective on the seropositivity of influenza A (H1N1), which was parallel with observations recorded by a study done in the American Society of Clinical Oncology, who found that occupation, immunocompetency, previous history of chronic diseases, smoking, showed significant effects on respiratory viral infections especially influenza A virus [16]. The current observations were relatively similar and agreed with results reported by the Iranian research groups who reported in 2019 [17]. Moreover, our conclusions nearby with results reported in a study done in Switzerland [18]. Moreover, other investigators reported a relatively higher prevalence of influenza A viral infections and transmissions [19].

TABLE 2: Hematologic parameter evaluation of seropositive cases with normal ranges

Hematologic parameters	Units	H1N1 seropositive (Mean±SD)	Normal range	P-value
WBC total	10 ⁹	7.4±2.2	4–11	P>0.05
Hemoglobin	g/dl	11.92±1.71	11.5–15.5	
RBC	10 ¹²	2.72±0.88	3.8–4.8	
Platelets	10 ⁹	186.61±59.24	150–450	
Lymphocytes	%	28.8±7.67	2–45	
Granulocytes	%	56.48±12.73	40–80	
MID	%	7.64±1.92	2–10	

Some factors may be behind the high prevalence rates of H1N1 anti-IgG and anti-IgM seropositivity among males in the current study including the cultural behavior where males mostly enter into the crowded areas without following standard protection protocols, as well as smoking are more common among males in comparison to females. These observations were in agreement with results reported by epidemiological studies conducted in different areas among different groups and populations [20]. Preparation planning surveyed by the response to the first influenza pandemic of the 21st Century delivered a unique opportunity for construction and applying a global system of surveillance to chance both global and national needs [21]-[23]. The current work found a limited number of pandemic influenza A (H1N1) among the tested cases although the vast majority of them were within the flu-like syndrome. This may be due to the other pandemic viral infection by SARS-CoV 2 which is known as COVID-19. There are mixes between symptoms for both cases that may confuse the physicians and researchers in their discussions and more other laboratory investigations are necessary to be followed. Reports achieved by other workers support this conclusion and explanation [24]. This opinion opens a gate for a fact which is essential to recognize the co-infections by way of some individual can be treated with antibiotics and antivirals [25]. The current study revealed that CBC may not help identify influenza A virus (H1N1), which was parallel to other conclusions reported by others [26], although other investigators reported that the possibility of high monocytosis and lymphopenia could be considered as a good indicator [27].

5. CONCLUSIONS

It was concluded that the percentage rates of anti-IgG and anti-IgM seropositivity for influenza A (H1N1) viral infections was relatively in an accepted range in Sulaymani Governorate. Smoking, previous history of chronic diseases, and the

employment status of the tested cases showed to be among the significant risk factors for influenza A viral infections, especially H1N1. It was concluded that hematologic tests and parameters are not dependable in H1N1 diagnosis. Limited numbers of the studied cases showed positive results for RT-PCR comparing to the serologic investigations.

REFERENCES

- [1] N. Takeshi. "Native morphology of influenza virions". *Frontiers in Microbiology*, vol. 2, p. 269, 2012.
- [2] S. V. Bourmakina and A. García-Sastre. "Reverse genetics studies on the filamentous morphology of influenza A virus". *Journal of General Virology*, vol. 84, no. 3, pp. 517-527, 2003.
- [3] B. Szewczyk, K. Bieńkowska-Szewczyk and E. Król. "Introduction to molecular biology of influenza A viruses". *Acta Biochimica Polonica*, vol. 61, no. 3, pp. 397-401, 2014.
- [4] D. B. Smith, E. R. Gaunt, P. Digard, K. Templeton and P. Simmonds. "Detection of influenza C virus but not influenza D virus in Scottish respiratory samples". *Journal of Clinical Virology*, vol. 74, pp. 50-53, 2016.
- [5] C. Brockwell-Staats, R. G. Webster and R. J. Webby. "Diversity of influenza viruses in swine and the emergence of a novel human pandemic influenza A (H1N1)". *Influenza and Other Respiratory Viruses*, vol. 3, no. 5, pp. 207-213, 2009.
- [6] K. Wu, J. Liu, R. Saha, D. Su, V. D. Krishna, M. C. J. Cheeran and J. P. Wang. "Magnetic particle spectroscopy for detection of influenza A virus subtype H1N1". *ACS Applied Materials and Interfaces*, vol. 12, no. 12, pp. 13686-13697, 2020.
- [7] Z. Yu, K. Cheng, H. He and J. Wu. "A novel reassortant influenza A (H1N1) virus infection in swine in Shandong Province, Eastern China". *Transboundary and Emerging Diseases*, vol. 67, no. 1, pp. 450-454, 2020.
- [8] J. A. Pulit-Penalosa, C. Pappas, J. A. Belser, X. Sun, N. Brock, H. Zeng, T. M. Tumpey and T. R. Maines. "Comparative in vitro and in vivo analysis of H1N1 and H1N2 variant influenza viruses isolated from humans between 2011 and 2016". *Journal of Virology*, vol. 92, no. 22, p. e01444-18, 2018.
- [9] P. J. Campbell, S. Danzy, C. S. Kyriakis, M. J. Deymier, A. C. Lowen and J. Steel. "The M segment of the 2009 pandemic influenza virus confers increased neuraminidase activity, filamentous morphology, and efficient contact transmissibility to A/Puerto Rico/8/1934-based reassortant viruses". *Journal of Virology*, vol. 88, no. 7, p. 3802, 2014.
- [10] C. W. Potter. "A history of influenza". *Journal of Applied Microbiology*, vol. 91, no. 4, pp. 572-579, 2001.
- [11] S. Davis. "The different types of flu explained-seasonal influenza, swine flu, and avian flu". *SA Pharmacists Assistant*, vol. 19, no. 2, pp. 10-11, 2019.
- [12] P. J. Gavin and R. B. Jr. Thomson. "Review of rapid diagnostic tests for influenza". *Clinical and Applied Immunology Reviews*, vol. 4, no. 3, pp. 151-172, 2004.
- [13] M. Petric, L. Comanor and C. A. Petti. "Role of the laboratory in the diagnosis of influenza during seasonal epidemics and potential pandemics". *The Journal of Infectious Diseases*, vol. 194, no. Suppl 2, pp. S98-S110, 2006.
- [14] S. Dellière, M. Salmona, M. Minier, A. Gabassi, A. Alanio, J. Le Goff, C. Delaugerre, M. L. Chaix and Saint-Louis CORE

- (COVid REsearch) Group. "Evaluation of the COVID-19 IgG/IgM rapid test from orient gene biotech". *Journal of Clinical Microbiology*, vol. 58, no. 8, p. e01233-20, 2020.
- [15] M. Von Lilienfeld-Toal, A. Berger, M. Christopeit, M. Hentrich, C. P. Heussel, J. Kalkreuth, M. Klein, M. Kochanek, O. Penack, E. Hauf, C. Rieger, G. Silling, M. Vehreschild, T. Weber, H. H. Wolf, N. Lehnert, E. Schalk and K. Mayer. "Community-acquired respiratory virus infections in cancer patients guideline on diagnosis and management by the infectious diseases working party of the German society for haematology and medical oncology". *European Journal of Cancer*, vol. 67, pp. 200-212, 2016.
- [16] R. El Ramahi and A. Freifeld. "Epidemiology, diagnosis, treatment, and prevention of influenza infection in oncology patients". *Journal of Oncology Practice*, vol. 15, no. 4, pp. 177-184, 2019.
- [17] V. Rahmanian, M. Shakeri, H. Shakeri, A. S. Jahromi, A. Bahonar and A. Madani. "Epidemiology of influenza in patients with acute lower respiratory tract infection in South of Iran (2015-2016)". *Acta Facultatis Medicae Naissensis*, vol. 36, no. 1, pp. 27-37, 2019.
- [18] L. P. Hariri, C. M. North, A. R. Shih, R. A. Israel, J. H. Maley, J. A. Villalba, V. Vinarsky, J. Rubin, D. A. Okin, A. Sclafani, J. W. Alladina, J. W. Griffith, M. A. Gillette, Y. Raz, C. J. Richards, A. K. Wong, A. Ly, Y. P. Hung, R. R. Chivukula, C. R. Petri, T. F. Calhoun, L. N. Brenner, K. A. Hibbert, B. D. Medoff, C. C. Hardin, J. R. Stone and M. Mino-Kenudson. "Lung histopathology in coronavirus disease 2019 as compared with severe acute respiratory syndrome and H1N1 influenza". *Chest*, vol. 159, no. 1, pp. 73-84, 2020.
- [19] E. Kenah, D. L. Chao, L. Matrajt, M. E. Halloran and I. M. Jr. Longini. "The global transmission and control of influenza". *PLoS One*, vol. 6, no. 5, p. e19515, 2011.
- [20] V. M. Konala, S. Adapa, V. Gayam, S. Naramala, S. R. Daggubati, C. B. Kammari and A. Chenna. "Co-infection with influenza A and COVID-19". *European Journal of Case Reports in Internal Medicine*, vol. 7, no. 5, p. 001656, 2020.
- [21] S. Briand, A. Mounts and M. Chamberland. "Challenges of global surveillance during an influenza pandemic". *Public Health*, Vol. 125, no. 5, pp. 247-256, 2011.
- [22] B. N. Archer, C. Cohen, D. Naidoo, J. Thomas, C. Makunga, L. Blumberg, M. Venter, G. Timothy, A. Puren, J. McAnerney, A. Cengimbo and B. Schoub. "Interim report on pandemic H1N1 influenza virus infections in South Africa, April to October 2009: Epidemiology and factors associated with fatal cases". *Eurosurveillance*, vol. 14, no. 42, pp. 19369, 2009.
- [23] D. Miyazawa. "Why obesity, hypertension, diabetes, and ethnicities are common risk factors for COVID-19 and H1N1 influenza infections". *Journal of Medical Virology*, vol. 93, no. 1, pp. 127-128, 2021.
- [24] Z. A. Memish, A. M. Assiri, R. Hussain, I. Alomar and G. Stephens. "Detection of respiratory viruses among pilgrims in Saudi Arabia during the time of a declared influenza A (H1N1) pandemic". *Journal of Travel Medicine*, vol. 19, no. 1, pp. 15-21, 2012.
- [25] Z. Shimoni, J. Glick and P. Froom. "Clinical utility of the full blood count in identifying patients with pandemic influenza A (H1N1)". *The Journal of Infection*, vol. 66, no. 6, pp. 545-547, 2013.
- [26] O. Coşkun, I. Y. Avci, K. Sener, H. Yaman, R. Ogur, H. Bodur and C. P. Eyigün. "Relative lymphopenia and monocytosis may be considered as a surrogate marker of pandemic influenza A (H1N1)". *Journal of Clinical Virology*, vol. 47, no. 4, pp. 388-389, 2010.
- [27] Y. Egawa, S. Ohfuji, W. Fukushima, Y. Yamazaki, T. Morioka, M. Emoto, K. Maeda, M. Inaba and Y. Hirota. "Immunogenicity of influenza A (H1N1) pdm09 vaccine in patients with diabetes mellitus: With special reference to age, body mass index, and HbA1c". *Human Vaccines and Immunotherapeutics*, vol. 10, no. 5, pp. 1187-1194, 2014.

Dissimilarity in the Nuclear Ribosomal DNA Internal Transcribed Spacer Regions of *Haplophyllum* spp. Founded in Ashdagh Mountain, Sangaw, Kurdistan of Iraq



Mohammed Ibrahim Khalil

Department of Animal Production, College of Agricultural Engineering Sciences, University of Garmian, Kalar, Sulaymaniyah, Kurdistan Region, Iraq

ABSTRACT

Studying and understanding the changes in the molecular sequences of plants are important to better identification and classifying the species. Internal transcribed spacer (ITS) regions are considered an informative genetic sequence to find out the variation among species of the same genus. In this study, ITS regions of *Haplophyllum* collected in Ashdagh Mountain were amplified, analyzed, and compared with reference species gathered from gene bank. The results showed that the collected species was closely related to *Haplophyllum blanchei* and *Haplophyllum tuberculatum*. In addition, there were differences in the number of the base pairs in the ITS1 region between *Haplophyllum* sp. and *H. blanchei* and *H. tuberculatum*. The transition to transversion ratio between *Haplophyllum* sp. and *H. tuberculatum* was lower (= 1) than with other species. The results reveal that the studied plant could be a new species in Iraq. To the best of our knowledge, this study is the first molecular taxonomic study done to identify *Haplophyllum* species in Ashdagh mountain in Sangaw/Suleimani.

Index Terms: ITS regions, Base pairs, Transversion, Molecular, Sangaw/Suleimani

1. INTRODUCTION

The genus *Haplophyllum* is considered one of the most common and rich genera in Rutaceae family [1,2]. The family includes 69 species of *Haplophyllum*. This genus is widely distributed throughout tropical, subtropical, and temperate regions [3,4]. Studies showed that some species of this genus contain flavonoid, lignans, alkaloids, and glycosides [5-7].

These phytochemical properties of the species made it an important medicinal plant [8]. The life form of the species

is characterized as perennial herbs, growing on rocky hills, slopes, sandy soil, stony mountains, or steppes [2,3].

Many morphological studies have been conducted on the species of this genus and some new species have been identified recently. For example, new species (*Haplophyllum ermenekense*) has been identified as new species in Turkey [9]. However, few species have been identified or studied through molecular approaches [2]. Nuclear ribosomal internal transcribed spacer (ITS) regions are more powerful than other plastid regions in identification of intraspecific variability among species of the same genus [10-12]. ITS regions are characterized as ITS1 and ITS2. In addition, the high evolution rates of these two regions make them more useful for studying and tracking the diversification levels among closely related species [13]. Many new plant species have been identified and recorded based on ITS regions [14]. Sangaw area with unique

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp7-11
E-ISSN: 2521-4217
P-ISSN: 2521-4209

Copyright © 2022 Khalil. This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Mohammed Ibrahim Khalil, College of Agricultural Engineering Sciences, University of Garmian, Kalar, KRG.
E-mail: mohammed.ibrahim@garmian.edu.krd

Received: 19-08-2021

Accepted: 24-01-2022

Published: 01-02-2022

geological and ecological characteristics (abiotic factors) would have a massive influence on the *Haplophyllum* species that would be assessed at the molecular level. We predict that the ITS regions will be different in compare with other available reference species in GenBank. The rationale for our prediction is because the surveyed has its unique geological and ecological characteristics (abiotic factors) and it is possible that the abiotic factors have significant influence on the composition of ITS sequences. Therefore, the aim of this study is to examine the ITS regions of *Haplophyllum* collected in Ashdagh mountain in Sangaw. Sangaw is a district located in the south east of Sulaimani Province. The aim of this study is to examine the ITS regions of *Haplophyllum* collected in Ashdagh mountain in Sangaw (a district located in the south east of Sulaimani Province).

2. MATERIALS AND METHODS

2.1. Sample Collection

The leaves samples (ten leaves) were collected from the unknown plant sample founded on Ashadah Mountain in Sangaw (Latitude: 35° 12' 49", longitude: 45° 33' 47") in February 5, 2021 (Fig. 1). The collected samples placed inside zipper plastic bags within ice box. The zipper bags were transferred immediately to lab and stored in deep freezer to prepare it for molecular analysis.

2.2. DNA Extraction and Polymerase Chain Reaction (PCR)

AddPrep Genomic DNA Extraction Kit was used to extract the total genomic DNA. Two universal nuclear regions named ITS 1 and 2 (ITS1 and ITS2) primers were used to identify the unknown species based on the phylogenetic analysis. The primers for ITS1 are ITS1F (ITS1 forward) (5'- GGAAGKARAAGTCGTAACAAGG -3') and ITS1R (ITS1 reverse) (5'- GCGTTCAAAGAYTCGATGRITTC-3') and primers of the ITS2 gene are ITS2F (ITS2 forward) (5'- CAWCGATGAAGAACGYAGC-3') and ITS2R (ITS2 reverse) (5'-RGTTTCITTTTCCITCCGCTTA-3') [15]. In the standard PCR, each reaction contained (MyTaq™ HS Mix-Bioline, USA) master mix (10 µL) 10 pmol of each primer and 20 ng template DNA in a final volume of 20µL. PCR was performed using a three-step cycling protocol: initial denaturation (95°C/5 min) 1 cycle; (denaturation (95°C/30 s), annealing (57°C/30 sec), (extension 72°C/30 s) 40 cycles); and final extension (72°C/5) 1 cycle (Bio-Rad C1000 Thermal Cycler, USA). The PCR products were analyzed in 1% agarose gel (Only 1 g of agarose was dissolved in 100 mL (1X TAE buffer) TAE: Tris-acetate EDTA) stained with Ethidium bromide (0.07% ready to use). PCR products were purified and sequenced, the sequencing reactions were investigated for both strands of all two purified PCR products using HiSeq4000 (Illumina, San Diego, USA) of Macrogen Inc., Korea.

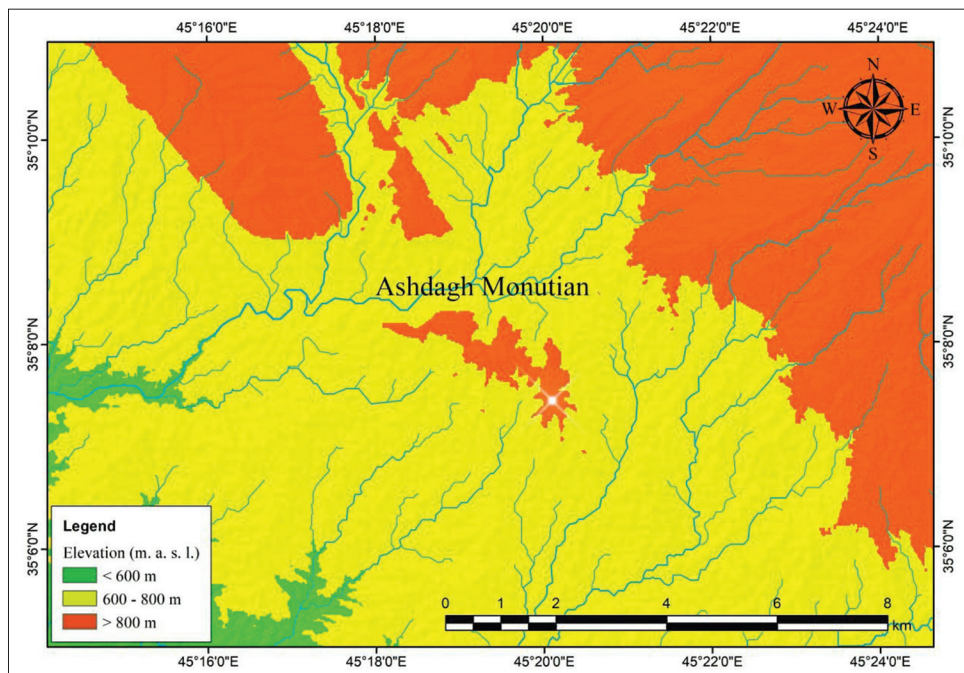


Fig. 1. Location map of study area. The white dot indicates the collection site.

The forward and reverse sequences of ITS1 and ITS2 of plant sample were combined with each other using Bioedit software v.7. [16] Blast software in National Center for Biotechnology Information (NCBI) database was used to search the identity of our sequence. Then, the species with high percentage identity (Table 1) were downloaded from GenBank (NCBI) and aligned using ClustalW embedded in Bioedit.

2.3. Sequence Analyses and Phylogenetic Tree

The pairwise distance (i.e. divergence) among *Haplophyllum* species was estimated using MegaX. Tamura 3-parameter (T92 + G) was picked as a best model to estimate the distance. Furthermore, pairwise deletion was used to treat the gaps among the sequences. PAUP 4.0 [17] was used to estimate transition and transversion rates (Table 2). The phylogenetic tree (Fig. 2) was reconstructed using Maximum likelihood (ML). MegaX was used to generate the trees and T92 with gamma distribution (+G) was picked as best substitution model. To test the confidence of the clades, bootstrap (BS) method was performed with 100 replicates.

3. RESULTS AND DISCUSSION

In comparison with reference species, the unknown sample showed that the plant genus is belong to *Haplophyllum* and

the percent identity of *Haplophyllum* sp. (samples) was 93.4% with *Haplophyllum tuberculatum* and 92.34% with *Haplophyllum blanchei* (Table 1).

The analysis of the sequences (Table 2) showed that the proportion of the difference (i.e., pairwise distance) among *Haplophyllum* sp., *H. tuberculatum*, and *H. blanchei* is 2% and it is less than others. In addition, the ratio of base pairs transition to transversion between *Haplophyllum* sp. and *H. tuberculatum* is 1% which means that the rate to transition (CT = 4) is equal to transversion (AT = 1, CG = 2, and GT = 1). On the other hand, *H. blanchei* had high transition to transversion rate (ratio = 3) in comparison with other species. The alignment of the sequences illustrates a clear difference in ITS1 region in *Haplophyllum* sp.

In compare with other species sequences (Fig. 3). For example, there is a clear difference in nucleotide composition among *Haplophyllum* sp. and *H. tuberculatum* and *H. blanchei* in position 14 to 98 (Fig. 2). *Haplophyllum* sp. has 84 and 15 more nucleotides than *H. tuberculatum* and *H. blanchei*, respectively. The length of sequence in *Haplophyllum* sp. was 589 base pairs which is longer than other species.

The ML tree strongly support the clade of *Haplophyllum* sp. and *H. tuberculatum* and *H. blanchei* (BS = 100) (Fig. 2). Based on the

TABLE 1: The percentage identification (Per. Ident) of the plant sample (*Haplophyllum* sp.) with available reference species in GenBank

No.	Scientific Name	Max Score	Total Score	Query Cover	Per. ident	Acc. Len	Accession
1	<i>Haplophyllum tuberculatum</i>	739	883	73%	93.4	635	KF805116.1
2	<i>Haplophyllum blanchei</i>	730	868	74%	92.32	637	AY484571.1
3	<i>Haplophyllum linifolium</i>	656	656	62%	90.64	628	AY484572.1
4	<i>Haplophyllum coronatum</i>	647	647	62%	90.25	630	AY484573.1
5	<i>Haplophyllum bastetanum</i>	645	645	62%	90.25	627	AY484576.1
6	<i>Haplophyllum rosmarinifolium</i>	638	638	62%	90.06	626	AY484574.1
7	<i>Haplophyllum suaveolens</i>	558	558	62%	86.77	625	AY484575.1
8	<i>Ruta montana</i>	283	283	50%	80.09	612	AY484577.1

TABLE 2: Sequence analysis of *Haplophyllum* sp. and other reference species from GenBank database (i.e., NCBI; National Center for Biotechnology Information)

No	Taxa	Ti		Tv				Ident				Prop diff	Ti/Tv ratio	Total	
		AG	CT	AC	AT	CG	GT	AA	CC	GG	TT				
	<i>Haplophyllum</i> sp. vs. :														
1	<i>Haplophyllum tuberculatum</i>	0	4	0	1	2	1	87	176	152	79	0.02	1	502	
2	<i>Haplophyllum blanchei</i>	2	7	0	0	3	0	98	205	173	84	0.02	3	572	
3	<i>Haplophyllum linifolium</i>	4	8	1	1	4	3	80	164	144	77	0.04	1.33	486	
4	<i>Haplophyllum coronatum</i>	5	9	4	2	3	2	82	162	143	76	0.05	1.27	488	
5	<i>Haplophyllum bastetanum</i>	4	9	1	1	4	3	79	164	144	76	0.05	1.44	485	
6	<i>Haplophyllum rosmarinifolium</i>	5	8	1	1	4	3	79	163	143	77	0.05	1.44	484	
7	<i>Haplophyllum suaveolens</i>	9	14	3	2	5	3	77	164	140	78	0.07	1.77	495	
8	<i>R. montana</i>	5	16	5	3	6	4	67	114	110	63	0.1	1.17	393	

The table include transition (ti) and transversion (tv) rates, rates of identical base pairs (Ident), proportion of difference (Prop diff), transition to transversion ratio (ti/tv ratio), and total length of the sequences (total)

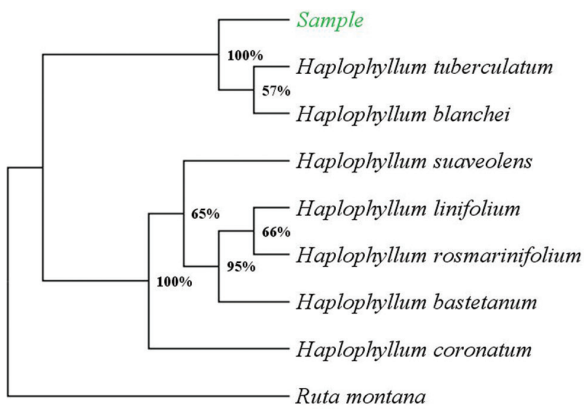


Fig. 2. Maximum Likelihood (ML) tree. The highlighted “sample” is the founded *Haplophyllum* sp. Numbers represent bootstrap values.

phylogenetic tree, *Haplophyllum* sp. is sister to subclade of *H. tuberculatum* and *H. blanchei*. ITS 1 and ITS2 are widely used in plant phylogenetic studies due to their ability to teasing out the intra- and interspecific variation among species [10]. In general, the base substitutions in the ITS region among *Haplophyllum* species are more than other species in angiosperm [2]. The high substitution rates among *Haplophyllum* species increase in their evolution rates [2,18,19]. The estimated transition/transversion bias (*R*) is 1.59. Substitution pattern and rates were estimated under the Kimura (1980) 2-parameter model [20]. The nucleotide frequencies are A = 25.00%, T/U = 25.00%, C = 25.00%, and G = 25.00%. For estimating ML values, a tree topology was automatically computed. The maximum Log likelihood for this computation was -1322.866.

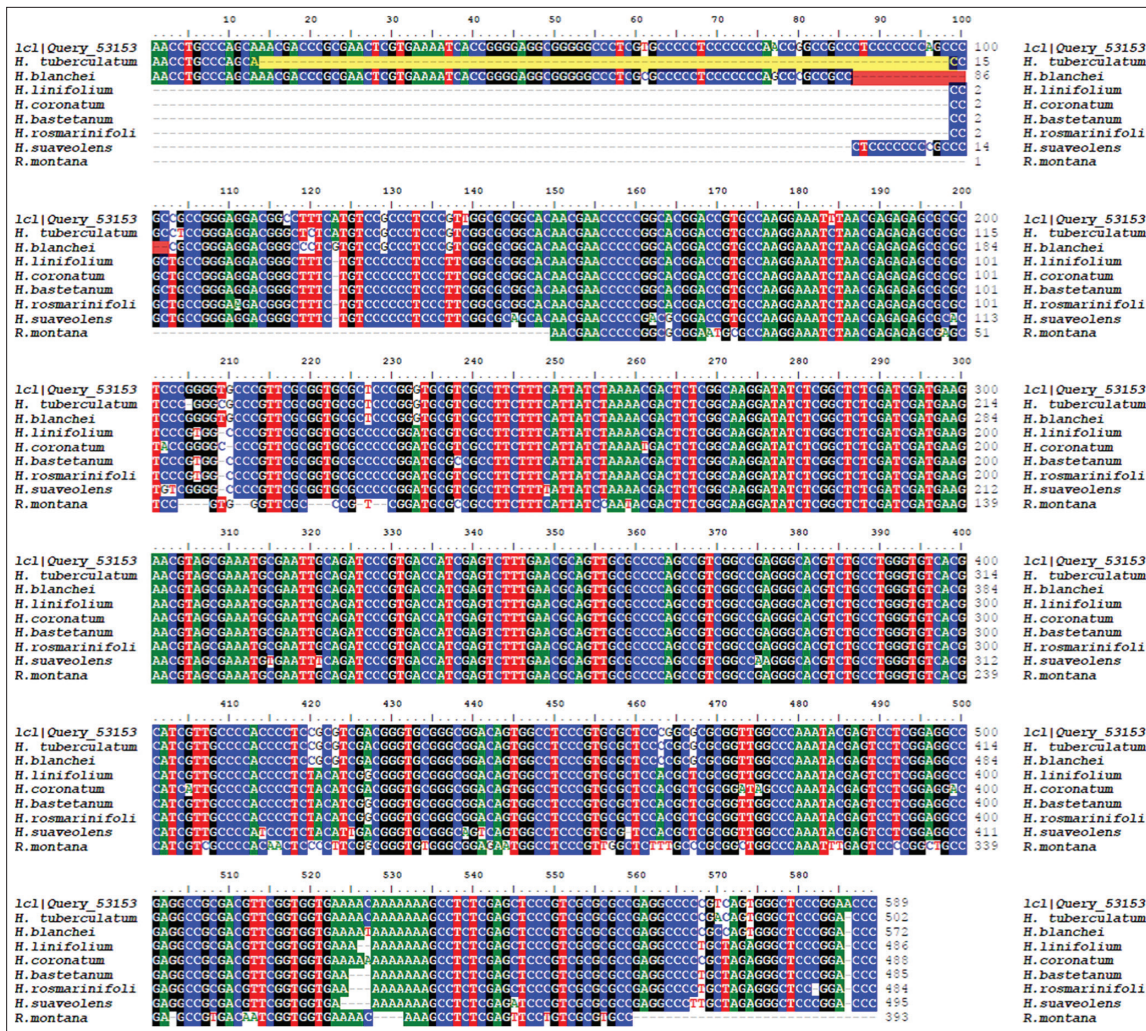


Fig. 3. Clustal W multiple sequence alignment of ITS1 and ITS2 regions. The sample species is indicated as “Icl|Query_53153” with 589 bp. The highlighted gaps with yellow and dark orange colors indicate the base pair differences among our sample species and *Haplophyllum tuberculatum* and *Haplophyllum blanchei*, respectively.

The analysis involved nine nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. There were a total of 589 positions in the final dataset. Evolutionary analyses were conducted in MEGA X [21]

Furthermore, *H. tuberculatum* and *H. blanchei* are most abundant species in Kurdistan [22]. The phylogenetic and sequence analyses reveal that the founded species is closely related to the *H. tuberculatum* and *H. blanchei* and distantly related with others. The differences in the number of base pairs in the ITS1 region between *Haplophyllum* sp. and *H. tuberculatum* and *H. blanchei* are the main cause of the phylogenetic placement of the species in the clade. Based on the sequence information for the ITS regions in the GenBank, the finding species could be considered as a new species in Iraq and have not been submitted before.

4. CONCLUSION

The existed dissimilarities in the ITS1 sequence and phylogenetic position of *Haplophyllum* sp. in compare with other reference species reveals that the founded species is a different species. In addition, this study showed the importance of using molecular studies to see the variability among closely related species.

REFERENCES

- [1] S. Mohammad and K. Ahmad Reza. "A new species of *Haplophyllum* (Rutaceae) from SW Iran". *Willdenowia*, vol. 35, no. 2, pp. 293-298, 2005.
- [2] F. B. Navarro, V. N. Suárez-Santiago and G. Blanca. "A new species of *Haplophyllum* A. Juss. (Rutaceae) from the Iberian Peninsula: Evidence from morphological, karyological and molecular analyses". *Annals of Botany*, vol. 94, no. 4, pp. 571-582, 2004.
- [3] C. C. Townsend. "Taxonomic revision of the genus *Haplophyllum* (Rutaceae)". Bentham-Moxon Trustees, Kew, Surrey, 1986.
- [4] M. Fay. Flowering plant families of the world. In: V. H. Heywood, R. K. Brummitt, A. Culham and O. Seberg. "Curtis's Botanical Magazine". Vol. 24. Firefly Books Ltd., Canada, pp. 198-200, 2007.
- [5] S. S. Nazrullaev, I. A. Bessonova and K. S. Akhmedkhodzhaeva. "Estrogenic activity as a function of chemical structure in *Haplophyllum* quinoline alkaloids". *Chemistry of Natural Compounds*, vol. 37, no. 6, pp. 551-555, 2001.
- [6] M. S. Ali, M. K. Pervez, M. Saleem and R. B. Tareen. "Haplophytin-A and B: The alkaloidal constituents of *Haplophyllum acutifolium*". *Phytochemistry*, vol. 57, no. 8, pp. 1277-1280, 2001.
- [7] M. Mohammadhosseini, A. Venditti, C. Frezza, M. Serafini, A. Bianco and B. Mahdavi. "The Genus *Haplophyllum* Juss.: Phytochemistry and bioactivities-a review". *Molecules*, vol. 26, no. 15, p. 4664, 2021.
- [8] A. B. Mahmoud, O. Danton, M. Kaiser, S. Han, A. Moreno, S. Abd Algaffar, S. Khalid, W. K. Oh, M. Hamburger and P. Mäser. "Lignans, amides, and saponins from *Haplophyllum tuberculatum* and their antiprotozoal activity". *Molecules*, vol. 25, no. 12, p. 2825, 2020.
- [9] D. Ulukuş and O. Tugay. "*Haplophyllum ermenekense* (Rutaceae), a new species from Turkey". *Phytokeys*, vol. 111, pp. 119-131, 2018.
- [10] B. Fazeli-Nasab, R. Z. Sayyed, M. Farsi, S. Ansari and H. A. El-Enshasy. "Genetic assessment of the internal transcribed spacer region (ITS1.2) in *Mangifera indica* L. landraces". *Physiol Mol Biol Plants*, vol. 26, no. 1, pp. 107-117, 2020.
- [11] X. Li, Y. Yang, R. J. Henry, M. Rossetto, Y. Wang and S. Chen. "Plant DNA barcoding: From gene to genome". *Biological Reviews*, vol. 90, no. 1, pp. 157-166, 2015.
- [12] China Plant BOL Group, D. Z. Li, L. M. Gao, H. T. Li, H. Wang, X. J. Ge, J. Q. Liu, Z. D. Chen, S. L. Zhou, S. L. Chen, J. B. Yang, C. X. Fu, C. X. Zeng, H. F. Yan, Y. J. Zhu, Y. S. Sun, S. Y. Chen, L. Zhao, K. Wang, T. Yang and G. W. Duan. "Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants". *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 49, pp. 19641-19646, 2011.
- [13] S. R. Downie and D. S. Katz-Downie. "A molecular phylogeny of apiaceae subfamily Apioideae: Evidence from nuclear ribosomal DNA internal transcribed spacer sequences". *American Journal of Botany*, vol. 83, no. 2, pp. 234-251, 1996.
- [14] P. Mishra, A. Kumar, V. Rodrigues, A. K. Shukla and V. Sundaresan. "Feasibility of nuclear ribosomal region ITS1 over ITS2 in barcoding taxonomically challenging genera of subtribe Cassiinae (Fabaceae)". *PeerJ*, vol. 4, pp. e2638-e2638, 2016.
- [15] T. Cheng, C. Xu, L. Lei, C. Li, Y. Zhang and S. Zhou. "Barcoding the kingdom plantae: New PCR primers for ITS regions of plants with improved universality and specificity". *Molecular Ecology Resources*, vol. 16, pp. 138-149, 2015.
- [16] T. A. Hall. "BioEdit: A user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT". *Nucleic Acids Symposium Series*, vol. 41, pp. 95-98, 1999. Available from: <https://www.ci.nii.ac.jp/naid/10029727236/en> [Last accessed on 2021 Aug 15].
- [17] D. L. Swofford. "PAUP: Phylogenetic Analysis Using Parsimony". Mac Version 3. 1. 1. (Computer Program and Manual), 1993. Available from: <https://www.ci.nii.ac.jp/naid/10003335753/en> [Last accessed on 2021 Aug 13].
- [18] Y. Suh, K. Heo and C. W. Park. "Phylogenetic relationships of maples (*Acer* L.; *Aceraceae*) implied by nuclear ribosomal ITS sequences". *Journal of Plant Research*, vol. 113, no. 2, pp. 193-202, 2000.
- [19] B. G. Baldwin, M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell and J. D. Michael. "The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny". *Annals of the Missouri Botanical Garden*, vol. 82, pp. 247-277, 1995. Available from: <http://www.search.ebscohost.com/login.aspx?direct=true&db=edspsc&AN=A1995QY12500007&site=eds-live> [Last accessed on 2022 Jan 10].
- [20] M. Kimura. "A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences". *Journal of Molecular Evolution*, vol. 16, no. 2, pp. 111-120, 1980.
- [21] S. Kumar, G. Stecher, M. Li, C. Knyaz and K. Tamura. "MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms". *Molecular Biology and Evolution*, vol. 35, pp. 1547-1549, 2018.
- [22] G. Salvo, S. Manafzadeh, F. Ghahremaninejad, K. Tojibaev, L. Zeltner and E. Conti. "Phytogeny, morphology, and biogeography of *Haplophyllum* (Rutaceae), a species-rich genus of the Irano-Turanian floristic region". *Taxon*, vol. 60, no. 2, pp. 513-527, 2011. Available from: <http://www.jstor.org/stable/41317146> [Last accessed on 2021 Aug 14].

New Feature-level Algorithm for a Face-fingerprint Integral Multi-biometrics Identification System



Bayan Omar Mohammed¹, Hamsa D. Majeed², Siti Zaiton Mohd Hashim³,
Muzhir Shaban Al-Ani⁴

^{1,2,4}Department of Information Technology, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq, ³Department of Data Science, Universiti Malaysia Kelantan (UMK), Taman Bendahara, 16100 Pengkalan Chepa, Kelantan

ABSTRACT

This article delves into the power of multi-biometric fusion for individual identification. A new feature-level algorithm is proposed that is the Dis-Eigen algorithm. Here, a feature-fusion framework is proposed for attaining better accuracy when identifying individuals for multiple biometrics. The framework, therefore, underpins the new multi-biometric system as it guides multi-biometric fusion applications at the feature phase for identifying individuals. In this regard, the Face-fingerprints of 20 individuals represented by 160 images were used in this framework. Experimental resultants of the proposed approach show 93.70 % identification rate with feature-level fusion multi-biometric individual identification.

Index Terms: Multi-model Biometric, Dis-Eigen algorithm, Identification, Aspect United Moment Invariant

1. INTRODUCTION

Technologies of biometric are utilized as a source in multiple applications that outline the interests of identification and authorization, such attentiveness demands a high level of privacy and security. Compared with traditional security, biometrics is considered an important type of security since it provides unique features through identifying biometrics characteristics [1].

A large number of security-based researchers used biometrics systems to improve the performance of their approaches. The biometric specifications are classified into two main classes: Physiological and behavioral, body parts included under fixed human characteristics and classified to be physiological

characteristics for instance iris, fingerprint, face, DNA, and retina, also classified as passive biometrics, while gait, voice, and handwritten signature classified as active biometrics as it represented by skills or functions performed by an individual and that make them belong to behavioral characteristics. Either way, those characteristics led to high authentication and verification for security [2].

Biometric security can be achieved through two kinds of categories: uni-model and multi-model [3]. Online list of biometric characteristics (active or passive) is used as a feature in the uni-model, this model has a low-security level against the multi-model which uses two or more of those characteristics that achieve a higher security level.

The main biometric characteristics for personal verification are obtained from face-fingerprint characteristics, the main aim of this paper is to improve human identification through the multi-model biometric process through merging between two biometric features, face-fingerprint, via these features, the system can compare, detect and identify the candidate within the constructed dataset. Both of these characteristics require

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp12-20

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Copyright © 2022 Mohammed *et al.* This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Bayan.omar@uhd.edu.iq, hamsa.al-rubaie@uhd.edu.iq, sitizaiton@umk.edu.my, muzhir.al-ani@uhd.edu.iq

Received: 22-11-2021

Accepted: 03-02-2022

Published: 11-02-2022

the existence of reference biometric data samples taken from different volunteers of different ages to be compared against the respective biometrical data of every person enrolled in a database or against a single reference template of particular enrolled individuals for identity confirmation of that person respectively [4].

The identification system accuracy is determined by its success comparison relying on the uniqueness of people's biometric characteristics, i.e, two persons can never have the same features [1]-[4].

2. RELATED WORK

According to the importance of using biometric systems in different applications and implementations, this field attracted countless researchers to propose their approaches using different biometrics characteristics. This section presents the most recent published approaches in this field.

In fingerprint recognition orientation, several techniques are recommended for accuracy enhancement of the recognition, those techniques were different upon various criteria, certain of these proposals areas of interest was in the preprocessing stage of fingerprint images [5], implemented an approach for the features extraction of both right and left-hand thumbs using many levels of two-dimensional discrete wavelet transform (2D-DWT), while another approach [6] used Discrete Cosine Transform (DCT) technique for extraction features [7], [8] presented techniques for fingerprint enhancement through localizing and recognizing the minutiae for minutiae extraction relying on the optimal thinning operation that took place in the preprocessing stage of fingerprint image. Prasad *et al.* [9] proposed a system consisting of many stages starting from data gathering which includes fingerprint images belonging to many different people then pre-processing those depending on their characteristics, finally, the algorithm is used for the purpose of recognition of fingerprint. Another area of interest was in proposing and developing algorithms for extracted feature stage from the process of recognition of that fingerprint, [10] proposed multi-biometric fusion for identical twins at the feature-level with Dis-Mean algorithm, Aspect United Moment Invariant (AUMI) used to define the individual biometric fingerprint characteristic. The extracted features regard the twin handwriting fingerprint for both word and shape. Furthermore, [11] proposed a fusion algorithm using the Mean-Discrete feature for identical twins fingerprint detection, The main method of the presented

algorithm requires the person's class labeling and multi-model biometric features to uni-modal biometric features conversion. As in multi-modal biometrics, the individuality represented by Mohammed and Hashim [12] using (AUMI) for global feature extractions to serve as a means to identical twins fingerprint detection, the procedure of Individuality representation measures the AUMI capacity for the individuality of the main of twin handwriting-fingerprint. A modified algorithm proposed by Mohammed [13] for Individuality Representation by employing the Mean-Discrete Algorithm, the vector of feature carries the features which generalized the global features owned by individuals. The developed model generalized the features in an earlier stage of classification.

In face recognition, many approaches have been proposed over the years that compound pattern recognition with computer vision and image processing, those algorithms are implemented in many forms to accomplish high-efficiency recognition. Al-Shayea *et al.* [14] present algorithm to specify the recognition rate of the PCA algorithm before and after applying DWT. The outcome was that applying DWT increases the recognition rate with minimization in the feature matrix size. Paper [15] proposed an approach based on wavelet-curvelet for facial features extraction, the used technique aims to reduce mathematical computational image analysis by reducing the dimensionality. The Nearest Mean Classifier (NMC) is employed for recognition at a high rate. An efficient approach is proposed [16] for face recognition concentrated on the efficiency of the image to calculate the required features from the image that contained the face. The outcome result gives a fine performance of the face recognized from that system. Qeethara Al-Shayea [17] proposed a system for information measurement between the main features of different face parts like angles and distances. An algorithm has been constructed with all the digital data compared with massive face images databases. Histogram equalization technique is used for extraction feature for recognition. Research [18] has a main objective of designing an efficient face recognition approach by generating a matrix with a significant rank by applying the technique of singular value decomposition. A relative study is made by Prasad *et al.* [19] scientific survey generated to study various methods and techniques with all the sufferings and benefits of these approaches. Al Ani and Al Waisy [20] presented an approach using a kernel machine for face detection from different views, the proposed algorithm shows powerful ability to effectively multi-view face detection. Another proposal presented by Nejr's and Al-Ani [21] for face classification through a structured approach that is implemented using

dissimilar levels in Two-Dimensional Discrete Fourier Transform (2D-DWT). While [22] proposed a technique to track faces and detect happiness parameters. The system is performed using a Raspberry Pi device, high-resolution camera, and high-definition screen. The study is compatible to apply in real-time. The main goal of this proposal is for studying the happiness level that the students have in the class and raise that level during the lecture.

Other researchers shared our interest in integrating face-fingerprint recognition in one system. Sriram J and Jacob J [23] proposed a voting system EVM where all the data of voters are digitally recorded. Through the proposed project, the faces of the voters are recognized first by a camera then the fingerprint is used for giving authentication through the data that is already stored in the database. Another multimodal system was proposed [24] for identification using face-fingerprint features. The calculations are taking place in the work first for each feature individually rather than when they are combined together. Divyakant and Meva [25] proposed a biometric system for performance assessment of face-fingerprint biometric traits collected from 30 humans. The recognition stage performance is measured through (False Acceptance Rate) FAR and (False Rejection Rate) FRR calculations. A comb filter approach is proposed systems [26] regarding face-fingerprint recognition through an encryption process to present cancelable patterns then make a comparison with any other encrypted biometric that have been randomly attacked. In article [27], firstly in the part of novel step, a system of smart cards accessing is presented by face authentication first then by fingerprint concurrently. The verification should be for both mutual authentications, after that the transaction is approvable [12].

3. THE MULTI-BIOMETRIC IDENTIFICATION SYSTEM APPROACH PROPOSAL

With regards to getting a high accuracy identification, the classifier input is rich with the major features extraction of individuals. In this work, the features are extracted for individual face-fingerprint. After extraction, the classification takes place so as to improve biometric identification. In this module, the feature sets of the defined processed biometrics data are extracted. Feature extraction is crucial for resolving raw data into simple, clear, and comprehensible data that would be able for matching learning. This phase is crucial in almost all systems of pattern recognition. The extraction of global features for both face-fingerprints including the extraction of macro features and method of minutiae-based extraction.

The Aspect United Moment Invariant (AUMI) is extracted to attain the individual features from the images of individual face-fingerprint. The extracted features are all kept in the Storage of Invariant Feature Vector.

The proposed system for multimodal biometrics identification that employs two biometrics modalities is presented in Fig. 1. The adopted phase following feature extraction comprises the suggested fusion feature level.

At the level of feature fusion, signals that come from different biometrics channels are processed first. Meanwhile, feature vectors are separately extracted then go through a specific field algorithm of fusion known as the Dis-Eigen algorithm. These feature vectors are composite to generate a combined feature vector prior to being employed for similarity measurement and classification process.

4. EXTRACTION OF THE FEATURE

The feature sets are extracted from the defined processed biometrics data in this work using Aspect United Moment Invariant (AUMI) [28]. It is powerful in capturing the biometrics' individual global characteristics of Physiognomy and style of the fingerprint and for global features continuously and separately into an individual representation. The scorce of Aspect United Moment Invariant Structure is below and represented in Fig. 2.

Entirely There are eight AUMI features reported below:

$$\theta_1 = \frac{\sqrt{\phi_2}}{\phi_1} \tag{1}$$

$$\theta_2 = \frac{\phi_6}{\phi_1\phi_4} \tag{2}$$

$$\theta_3 = \frac{\sqrt{\phi_5}}{\phi_4} \tag{3}$$

$$\theta_4 = \frac{\phi_3}{\phi_2\phi_4} \tag{4}$$

$$\theta_5 = \frac{\phi_1\phi_6}{\phi_1\phi_3} \tag{5}$$

$$\theta_6 = \frac{(\phi_1 + \sqrt{\phi_2})\phi_2}{\phi_6} \tag{6}$$

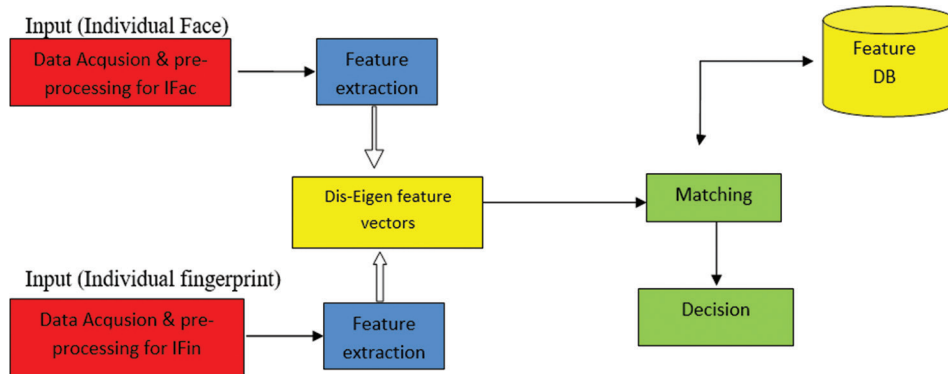


Fig. 1. Proposed multi-biometric system.

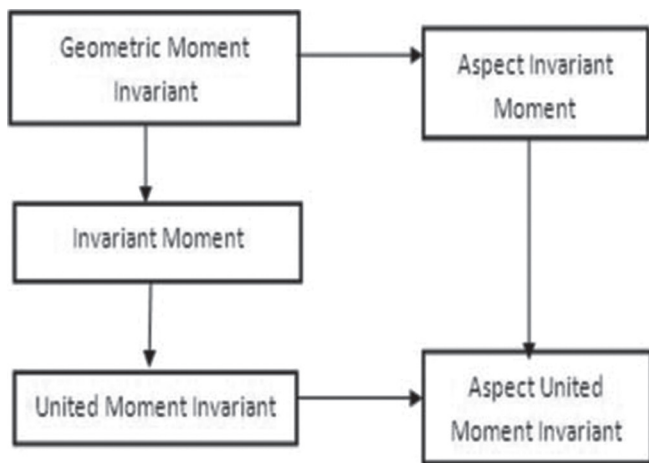


Fig. 2. Aspect united moment invariant structure.

$$\theta_7 = \frac{\varnothing_1 \varnothing_5}{\varnothing_3 \varnothing_6} \tag{7}$$

$$\theta_8 = \frac{\varnothing_3 + \varnothing_4}{\sqrt{\varnothing_3}} \tag{8}$$

Where φ_i represents Hu's moment invariants. Due to the large values of φ_i , the natural logarithm is applied thus giving us: for $i = 1$ to 7 ; $\theta_i \leftarrow \log_{10} \varphi_i$.

A set of global features are generated from the extracted features. For the purpose of improving identification performance, these features are gathered individually. Since the extracted features are in the multi-representations zone, then it has been used in combined form. Such combined features are termed as Dis-Eigen feature vectors in the Uni-representation zone, which is employed after the process of feature extraction [12], [28], [29].

5. THE PROPOSED DISCRETIZED-EIGEN (DIS-EIGEN) ALGORITHM IN FEATURE LEVEL FUSION

This work attempts to design a more effective multimodal biometrics identification system introduces by introducing the new proposed Dis-Eigen feature-based fusion as Fig. 3 illustrates with the capacity in generating distinctive features of numerous modalities of individuals, where *fac* and *fin* represent the face and fingerprint features consecutively. First, an improved AUMI is used as global in the extraction of features obtained from the individual face-fingerprint shape and style. Then, the features-based fusion is examined in terms of its generalization. Further, to achieve better classification accuracy, the Dis-Eigen feature-based fusion algorithm was used.

At the start of the process, Dis-Eigen feature-based fusion processes the raw biometrics input images using the suitable technique of processing. This results in images of standard size and quality. Then, the feature extraction algorithm is used to extract these images to generate discriminatory information which can provide a distinction between the identities. It is not easy to come up with a method that could comprehensively capture the discriminatory information from both raw input images among numerous identities or modalities and resolve problems in biometrics analysis. This study proposes the use of an enhanced combined

Feature vector fusion and uni-representation are known as the Dis-Eigen feature for multimodal biometrics identification. Also, Dis-Eigen replaces the feature fusion as a feature transformation agent to provide better feature representation from numerous modalities.

In the Dis-Eigen process, multi representation features from multi- biometrics are converted to uni- and systematic

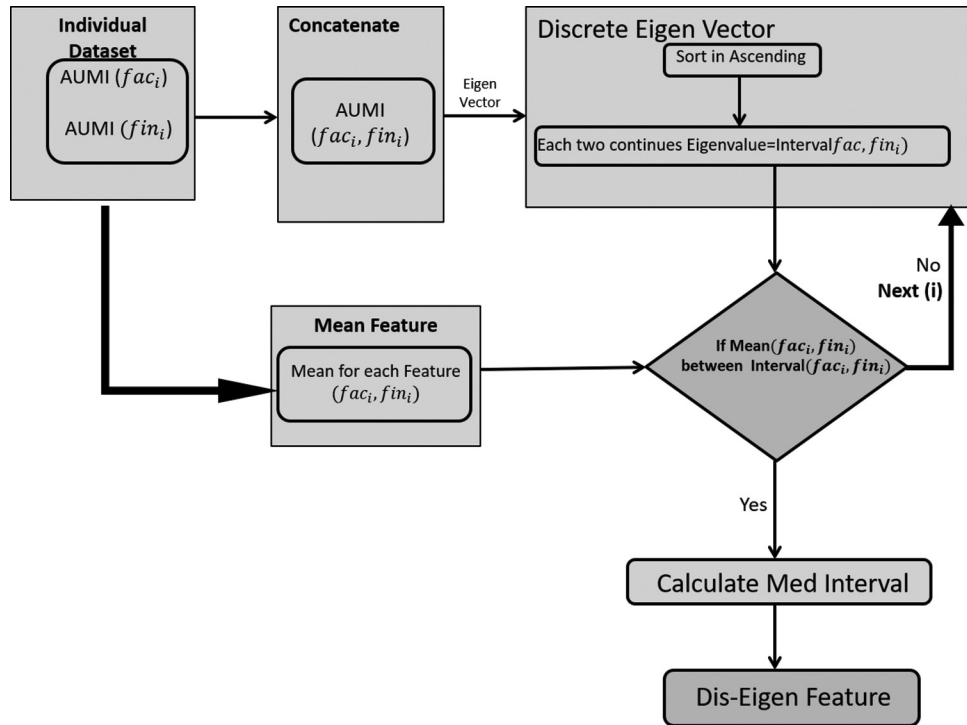


Fig. 3. The proposed dis-eigen feature-based fusion.

representation features for reduction of the complexity (dimension) of the feature vector. After concatenating an individual face-fingerprint the sorted Eigenvector in ascending order for a person in the aim of determining discrete intervals in the Dis-Eigen line. The dis-Eigen line is a line of an invariant feature vector that starts from the minimum Eigenvalue and ends with the maximum Eigenvalue from the Eigen feature vector for that person. The interval number is fit equally to the capacity of Eigenvector -1.

The eight features columns of AUMI have been applied in this study. The eigenvalue is the divider of intervals in the Dis-Eigen line. Each mean Face-Fingerprint feature (FF) vector that falls within the same interval will have the same Med Interval value. Med Interval (MI) for each interval is the average of an interval that is calculated using the formula as in (9):

$$MI = \frac{ev_i - ev_{i+1}}{2} \tag{9}$$

Med Interval value for interval one to seven represents the invariant feature

vector that falls within if $FF \geq ev_i$ and if $FF \leq ev_{i+1}$. On the other hand, compute the mean features that come from the individual face (fac_{ij}) and individual fingerprint (fin_{ij}) invariant

feature vector for a person in a twin. The computation and reduction for each face-fingerprint feature are expressed as below:

$$FF_{ij} = (fac_{ij} + fin_{ij}) / 2 \tag{10}$$

Where:

fac_{ij} : face features for an individual.

fin_{ij} : fingerprint features for an individual.

Eight features are created in this study and this is the column number for the features number of the applied AUMI for multi-biometrics. These features are called the Dis-Eigen feature-based fusion vector.

Tables 1-3 below exemplify the transformation of individual multimodal biometrics feature vector into Dis-Eigen feature-based fusion vector:

Tables 1 comprise eight columns representing the eight columns of invariant feature vectors within the AUMI then concatenated. These data are further applied in the Dis-Eigen process.

While Table 2 represents the Dis-Eigen features vector composed of the generalized features of individual features for an individual. Meantime Table 3 presents Dis-Eigen

TABLE 1: Real Data for face and fingerprint for individual number 1



Image	F1	F2	F3	F4	F5	F6	F7	F8
	1.0075	2.6937	1.7301	0.3341	2.9618	3.3513	1.2401	5.7571
	1.0075	5.2967	1.7303	0.334	5.8222	1.7048	630.6227	5.758
	1.0054	5.7202	1.7308	0.3338	6.3065	1.5772	583.5766	5.7626
	1.0075	2.6937	1.7301	0.3341	2.9618	3.3513	1.2401	5.7571
	1.0178	0.1602	1.7237	0.3363	0.0175	56.1404	2.0994	5.7059
	1.0142	0.1418	1.7242	0.3361	0.0156	63.2681	2.3703	5.71
	1.0305	0.0901	1.7238	0.3363	0.0096	101.1057	3.7337	5.7062
	1.0178	0.1021	1.7239	0.3362	0.011	88.8608	3.2916	5.7073

TABLE 2: Dis-Eigen face and fingerprint for individual number1

F1	F2	F3	F4	F5	F6	F7	F8
1.2312	1.2312	1.2312	1.2312	1.2312	57.6571	1.2312	1.2312
1.2312	1.2312	1.2312	1.2312	1.2312	57.6571	0	1.2312
1.2312	1.2312	1.2312	1.2312	1.2312	57.6571	0	1.2312
1.2312	1.2312	1.2312	1.2312	1.2312	57.6571	1.2312	1.2312

TABLE 3: Example of dis-eigen feature for individuals

F1	F2	F3	F4	F5	F6	F7	F8	Individual
3.4057	3.4057	3.4057	0.437	3.4057	47.802	0	3.4057	P10
3.4057	3.4057	3.4057	0.437	3.4057	47.802	0	3.4057	P10
3.4057	3.4057	3.4057	0.437	3.4057	47.802	3.4057	3.4057	P10
3.4057	3.4057	3.4057	0.437	3.4057	47.802	3.4057	3.4057	P10
0.5237	4.5258	4.5258	0.5237	4.5258	54.7769	0	4.5258	P11
0.5237	4.5258	4.5258	0.5237	4.5258	54.7769	0	4.5258	P11
0.5237	4.5258	4.5258	0.5237	4.5258	54.7769	4.5258	4.5258	P11
0.5237	4.5258	4.5258	0.5237	4.5258	54.7769	4.5258	4.5258	P11
1.2422	5.3632	1.2422	1.2422	5.3632	58.042	0	5.3632	P12
1.2422	5.3632	1.2422	1.2422	5.3632	58.042	0	5.3632	P12
1.2422	5.3632	1.2422	1.2422	5.3632	58.042	58.042	5.3632	P12
1.2422	5.3632	1.2422	1.2422	5.3632	58.042	0	5.3632	P12
1.8191	6.8858	1.8191	1.8191	6.8858	64.3154	0	6.8858	P13
1.8191	6.8858	1.8191	1.8191	1.8191	64.3154	0	6.8858	P13
1.8191	1.8191	1.8191	1.8191	1.8191	64.3154	0	6.8858	P13
1.8191	6.8858	1.8191	1.8191	6.8858	64.3154	0	6.8858	P13
1.8032	1.8032	1.8032	1.8032	1.8032	52.7556	0	5.5007	P14
1.8032	1.8032	1.8032	1.8032	1.8032	52.7556	1.8032	5.5007	P14
1.8032	5.5007	1.8032	1.8032	5.5007	52.7556	5.5007	5.5007	P14
1.8032	1.8032	1.8032	1.8032	1.8032	52.7556	1.8032	5.5007	P14
1.7582	3.1658	1.7582	0.6345	3.1658	55.8628	0	55.8628	P15
1.7582	55.8628	1.7582	0.6345	55.8628	55.8628	0	55.8628	P15
1.7582	0.6345	1.7582	0.6345	0.6345	55.8628	3.1658	55.8628	P15
1.7582	3.1658	1.7582	0.6345	3.1658	55.8628	0	55.8628	P15
1.3383	3.4765	1.3383	1.3383	3.4765	51.3403	0	51.3403	P16
1.3383	3.4765	1.3383	1.3383	3.4765	51.3403	0	51.3403	P16
1.3383	1.3383	1.3383	1.3383	1.3383	51.3403	1.3383	51.3403	P16
1.3383	3.4765	1.3383	1.3383	3.4765	51.3403	0	51.3403	P16

features for both face and fingerprint for individuals (10–16).

6. EXPERIMENT AND RESULTS

Precise results of identification process generated by Dis-Eigen algorithm within the focal point of multi-representation analysis for an individual’s face-fingerprint. An enhancement level of an individual’s face-fingerprint for Dis-Eigen feature-based fusion data utilization is proven in this work.

In this study, the proposed Dis-Eigen algorithm is Dis-Eigen feature-based fusion. Datasets with different numbers of both test and other train are used in this work, precisely two kinds of examples are established, The first adopt individual datasets with a split percentage of 60% as training and 40% for testing. The other has a split percentage of 80% as training and 20% as testing. The implementation of training is demonstrated using NaiveBayes, RandomForest, RandomForest, and J48

along with folds cross-validation of eight (8) and ten (10). In this experiment, the data sets comprise 160 data which are broken down into two categories with 20 individuals. As presented in Tables 4-7 along with Figs. 4 and 5.

Tables 4 and 5 show the performance of Dis-Eigen, face, fingerprint, and Concatenate features for the eight classifiers for the two experimental analysis setup. In average, the performance of Dis-Eigen feature for all classifiers has

TABLE 4: Provide the accuracy for classification process for Dis-Eigen feature-based fusion with all Methods for Split Percentage of 60% Training and 40% of Testing

Methods	NaiveBayes	Random forest	Random tree	J48
<i>Dis-Eigen</i>	84.37	100	84.37	65.62
Face	6.25	21.78	18.75	15.62
Fingerprint	18.75	18.75	21.75	12.5
Concatenate	10.93	18.75	10.93	14.06

TABLE 5: Provide the accuracy for classification process for Dis-Eigen feature-based fusion with all Methods for Split Percentage of 80% Training and 20% of Testing

Methods	NaiveBayes	RandomForest	Random Tree	J48
<i>Dis-Eigen</i>	87.5	100	100	100
Face	25	31.25	25	18.75
Fingerprint	12.5	18.75	12.5	6.25
Concatenate	15.62	18.75	12.5	15.62

TABLE 6: Provide the accuracy for Dis-Eigen feature-based fusion with all Methods for eight (8) folds Cross-Validation

Methods	NaiveBayes	RandomForest	Random Tree	J48
<i>Dis-Eigen</i>	88.75	100	98.75	100
Face	13.75	23.75	17.5	22.5
Fingerprint	32.5	23.75	27.5	27.5
Concatenate	5	22.5	22.5	25.62

TABLE 7: Provide the accuracy for Dis-Eigen feature-based fusion with all Methods for ten (10) folds Cross-Validation

Methods	NaiveBayes	RandomForest	Random Tree	J48
<i>Dis-Eigen</i>	91.25	100	100	98.75
Face	16.25	21.25	17.5	25
Fingerprint	33.75	25	22.5	28.75
Concatenate	4.37	25.62	20	20.62

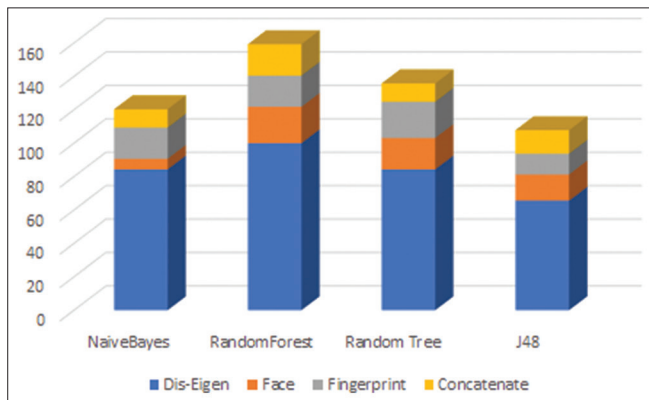


Fig. 4. Percentage of 60% training and 40% of testing.

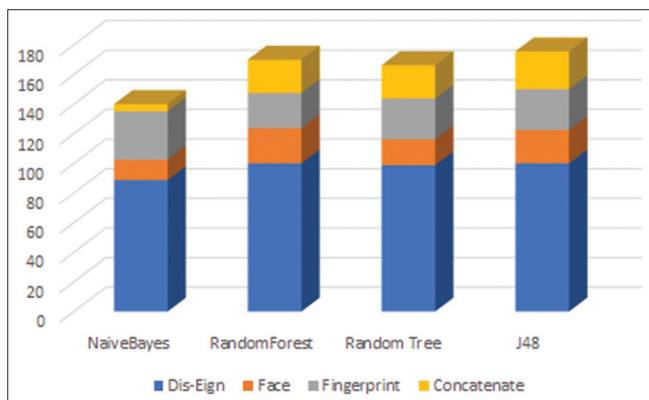


Fig. 5. All methods for eight (8) folds Cross.

succeeded as the heights performance with 90.23% accuracy rate in average. This is followed by 20.3 % for face rule, 14.64% for Concatenate, and 15.21 % for Fingerprint.

The Dis-Eigen algorithm has presented the best performance accuracy of 97.18% in average in Tables 6 and 7 for all classifiers with eight and ten fold cross validation environment setup. Though, six fusion and non-fusion algorithm have achieved quite a lower average performance of 19.68% for Face rule, 27.65% for Fingerprint and 18.20% for Concatenate. This is a very poor performance in comparison toward the performance of Dis-Eigen features. This has shown that Dis-Eigen algorithm significantly increased their classification performance.

As referred, Tables 4-7 and Figs. 4 and 5 presented the overall results of the various methods, it is a noticeable sign that the Dis-Eigen feature-based fusion has the optimum accuracy than face or fingerprint individually besides the Concatenate data. The preferable outcome achieves the applied improvement to the features which are individually represented through Dis-Eigen feature-based fusion.

7. CONCLUSION

The Dis-Eigen feature-based fusion algorithm has been proposed in this work as an attempt for multi-model biometric system improvement for individuality in face-fingerprint identification.

The proposed algorithm converts the multi-representations of individual features into a uni-representation with the technique of the Dis-Eigen algorithm. Generalized features of an individual have been presented significantly. The new approach has been evaluated and beard comparison with the conventional one with regard to similarity measurement. According to this, an individuals' face- fingerprints were identified. Resultant scrutinization has been made. The outcome feature from the Dis-Eigen feature application is systematic and more informative as experimental resultants of the proposed approach show 93.70 % identification rate with feature-level fusion multi-biometric individual identification. Furthermore, a particular improvement in system performance accuracy is achieved.

REFERENCES

- [1] M. Al-Ani and K. Al-Baset. "Efficient Watermarking based on Robust Biometric Features." *IRACST-Engineering Science and Technology: An International Journal*, vol. 3, pp. 529-534, 2013.
- [2] M. S. Al-Ani and M. A. Rajab. "Biometrics hand geometry using discrete cosine transform (DCT)." *Science and Technology*, vol. 3, no. 4, pp. 112-117, 2013.
- [3] M. Al-Ani and S. Nejrs. "Efficient biometric iris recognition based on iris localization approach." *UHD Journal of Science and Technology*, vol. 3, pp. 24-32, 2020.
- [4] Z. A. Kakarash, D. F. Abd, M. Al-Ani, G. A. Omar and K. Mohammed. "Biometric Iris Recognition Approach Based on Filtering Techniques." 2019.
- [5] M. S. Al-Ani, T. N. Muhamad, H. A. Muhamad and A. A. Nuri. "Effective Fingerprint Recognition Approach Based on Double Fingerprint Thumb." In: 2017 International Conference on Current Research in Computer Science and Information Technology (ICCRIT), 2017.
- [6] M. ShabanAl-Ani and W. M. Al-Aloosi. "Biometrics fingerprint recognition using discrete cosine transform (DCT)." *International Journal of Computer Applications*vol, 69, no. 6, pp. 44-48, 2013.
- [7] O. H. A. Al-Ani. "Human identification based on thinning minutiae of fingerprint." *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 17, pp. 5918-5929, 2018.
- [8] M. S. Al-Ani. "A novel thinning algorithm for fingerprint recognition." *International Journal of Engineering Sciences*, vol. 2, no. 2, pp. 43-48, 2013.
- [9] R. S. Prasad, M. S. Al-Ani and S. M. Nejres. "An efficient approach for fingerprint recognition." *International Journal of Engineering Innovation and Research*, vol. 4, no. 2, pp. 307-313, 2015.
- [10] B. O. Mohammed and S. M. Shamsuddin. "Feature Level Fusion for Multi-Biometric with Identical Twins." In: 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE), 2018.
- [11] B. O. Mohammed. "Fusion method with mean-discrete algorithm in feature level for identical twins identification." *UHD Journal of Science and Technology*, vol. 4, no. 2, pp. 141-150, 2020.
- [12] B. O. Mohammed and Z. M. Hashim. "Individuality representation using multimodal biometrics with aspect uniated moment invariant for identical twins." *Journal of Theoretical and Applied Information Technology*, vol. 98, no. 12, pp. 2148-2157, 2020.
- [13] B. O. Mohammed. "Mean-discrete algorithm for individuality representation." *Journal of Al-Qadisiyah for Computer Science and Mathematics*, vol. 13, no. 1, pp. 1-10, 2021.
- [14] Q. K. Al-Shayea, M. S. Al-Ani and M. S. A. Teamah. "The effect of image compression on face recognition algorithms." *International Journal of Computer and Network Security*, vol. 2, no. 8, pp. 56-60, 2010.
- [15] M. S. Al Ani. "Face Recognition Approach Based on Wavelet-Curvelet Technique." *Signal Image Process*, vol. 3, no. 2, pp. 21-31, 2012.
- [16] R. S. Prasad, M. S. Al-Ani and S. M. Nejres. "An efficient approach for human face recognition." *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 5, no. 9, pp. 133-136, 2015.
- [17] M. A. A. Qeethara Al-Shayea. "Biometric face recognition based on enhanced histogram approach." *International Journal of Communication Networks and Information Security*, vol. 10, no. 1, pp. 148-154, 2018.
- [18] O. H. Ahmed, J. Lu, Q. Xu and M. S. Al-Ani. "Face recognition based rank reduction SVD approach." *The ISC International Journal of Information Security*, vol. 11, no. 3, pp. 39-50, 2019.
- [19] R. S. Prasad, M. S. Al-Ani and S. M. Nejres. "Human identification

- via face recognition: Comparative study." *IOSR Journal of Computer Engineering*, vol. 19, no. 3, pp. 17-22, 2017.
- [20] M. S. Al Ani and A. S. Al Waisy. "Multi-view face detection based on kernel principal component analysis and kernel support vector techniques." *International Journal on Soft Computing*, vol. 2, no. 2, pp. 1-13, 2011.
- [21] S. M. Nejrj and M. S. Al-Ani. "Face image classification based on feature extraction." *Solid State Technology*, vol. 63, no. 6, pp. 13515-13526, 2020.
- [22] M. S. Al-Ani. "Happiness measurement through classroom based on face tracking." *UHD Journal of Science and Technology*, vol. 3, no. 1, pp. 9-18, 2019.
- [23] Sriram J and Jacob J. "Smart EVM based on face and fingerprint recognition." *IJRASET*, vol. 8, no. 8, pp. 1606-1610, 2020.
- [24] M. Szymkowski and K. Saeed. "A Multimodal Face and Fingerprint Recognition Biometrics System." In: *Computer Information Systems and Industrial Management*. Springer International Publishing, Cham, pp. 131-140, 2017.
- [25] D. C. K. Divyakant and T. Meva. "Performance Measurement of Face and Fingerprint Recognition System." In: *RK University First International Conference On Research and Entrepreneurship*, 2016.
- [26] M. Abd Al Rahim, W. El-Shafai, E. S. M. El-Rabaie, O. Zahran and F. E. Abd El-Samie. "Comb filter approach for cancelable face and fingerprints recognition." *Menoufia Journal of Electronic Engineering Research*, vol. 28, no. 1, pp. 89-94, 2019.
- [27] G. S. G. Anjaneyulu and V. Jalaja. "Novel Authentication Process of the Smart Cards Using Face and Fingerprint Recognition." In: *Advances in Automation, Signal Processing, Instrumentation, and Control*. Springer, Singapore, pp. 2547-2556, 2021.
- [28] B. O. Mohammed and S. M. Shamsuddin. "Twins multimodal biometric identification system with aspect united moment invariant." *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 4, pp. 788-803, 2017.
- [29] B. O. Mohammed and S. M. Shamsuddin. "A multimodal biometric system using global features for identical twins identification." *Journal of Computational Science*, vol. 14, no. 1, pp. 92-107, 2018.

Classification of Acute Lymphoblastic Leukemia through the Fusion of Local Descriptors



Shakhawan Hares Wady^{1,2}

¹Department of Applied Computer, College of Medicals and Applied Sciences, Charmo University, Chamchamal, Sulaimani, KRG, Iraq, ²Department of Information Technology, University College of Goizha, Sulaimani, KRG, Iraq

ABSTRACT

Leukemia is characterized by an abnormal proliferation of leukocytes in the bone marrow and blood, which is usually detected by pathologists using a microscope to examine a blood smear. Leukemia identification and diagnosis in advance are a trending topic in medical applications for decreasing the death toll of individuals with Acute Lymphoblastic Leukemia (ALL). It is critical to analyze the white blood cells for the identification of ALL for which the blood smear images are utilized. This paper discusses and presents a micro-pattern descriptor, called Local Directional Number Pattern along with Multi-scale Weber Local Descriptor for feature extraction task to determine cancerous and noncancerous blood cells. A balanced dataset with 260 blood smear images from the ALL-IDB2 dataset was used as training data. Consequently, a proposed model was constructed by applying different individual and combined feature extraction methods, and fed into the machine learning classifiers (Decision Tree, Ensemble, K-Nearest Neighbors, Naïve Bayes, and Random Forest) to determine cancerous and noncancerous blood cells. Experimental results indicate that the developed feature fusion technique assured a reasonable performance compared to other existing works with a testing average accuracy of $97.69 \pm 1.83\%$ using Ensemble classifier.

Index Terms: Leukemia Diagnosis, Blood Smear, Feature Extraction, Machine Learning

1. INTRODUCTION

Hematology is the branch of medicine concerned with the study of blood and blood-forming organs, including the diagnosis, treatment, and prevention of illnesses of the blood, bone marrow cells, an immune-mediated disease, hemostatic, and vascular systems. Hematologists conduct and analyze a wide range of laboratory tests to assist clinicians in the diagnosis and treatment of disease. They deal with the blood and bone marrow to provide patients with immediate clinical care [1]. The amount and type of cells produced at any given time is controlled by your body's requirements. In

some cases, estimating how bone marrow cells contribute to a clinical disease may be more essential than determining the patient's hematologic status. Hematologists typically receive blood smear samples and study them for abnormality; if they identify the presence of diseases, they perform a bone marrow biopsy and provide a diagnosis in a short period of time.

The bone marrow's primary function is to produce red blood cells (RBCs or erythrocytes), platelets (or thrombocytes), and white blood cells (WBCs or leukocytes) [2], as shown in Fig. 1. Complete Blood Count (CBC) is a laboratory hematology analyzer medical test that provides information that can be used to diagnose a disease. The CBC measures the production of all the cellular components, identifies the patient's oxygen-carrying capacity by evaluating RBC counts, and allows for the estimation of the immune system by evaluating WBC counts with differential. This test supports in the diagnosis of anemia, certain cancers, infections, and a variety of other

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp21-33 E-ISSN: 2521-4217
P-ISSN: 2521-4209

Copyright © 2022 Wady. This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: shakhawan.hares@charmouniversity.org

Received: 22-01-2022

Accepted: 17-02-2022

Published: 26-02-2022

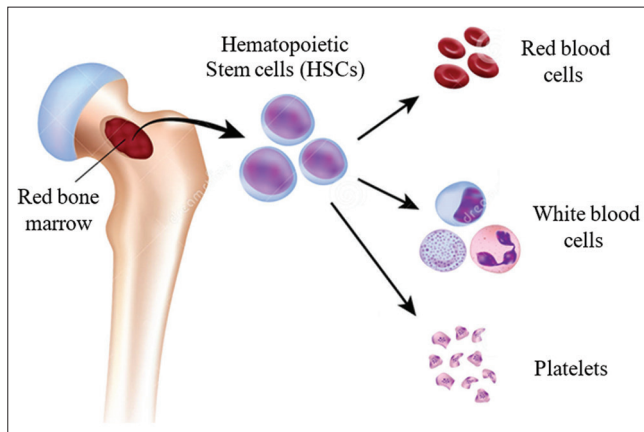


Fig. 1. Blood smear components (©Alila Medical Media, used with permission).

conditions, as well as monitoring the side effects of certain medications [3].

The absolute or relative count of WBCs, as well as their appearance on a blood smear, can be influenced by a variety of disorders. This is the case with parasite infections that cause an increase in WBCs, while the most severe cases are certainly leukemias [1]. Leukemia also known as Acute Lymphoblastic Leukemia (ALL), is a form of cancer of the blood or bone marrow that causes the body to produce cancerous WBCs called lymphocytes. These uncontrolled blood cells divisions harm the blood, lymphatic system, and bone marrow, throwing the immune system at risk [4]. They can also inhibit bone marrow's ability to produce RBCs and platelets. In addition, these cancerous WBCs can enter the bloodstream and cause harm to other regions of the human body, including as the liver, kidney, spleen, brain, and other organs, leading to the development of other deadly cancers.

In terms of how quickly it develops or gets worse, leukemia is characterized as either acute (which appears suddenly and grows quickly) or chronic (which develops more slowly). Acute leukemia can also be classified as lymphoblastic (ALL) or myelogenous depending on the cell type that is affected. Chronic lymphocytic leukemia (CLL) and chronic myeloid leukemia are two types of chronic leukemia [5-7]. ALL is the primary target of this research since it is predicted to have a superior survival rate than other categories. Only a suspicious and meticulous microscopic investigation of stained blood smears can accurately identify leukemia. Manual examination, due to the complicated nature of WBCs, results in inconsistencies in slide processing, resulting in non-standardized, unreliable, and subjective observations. Consequently, a cost-effective and robust automated system

is crucial to satisfy the demand for accurate diagnosis and classification without being impacted by experts. For this reason, several Computer-Aided Diagnosis (CAD) systems have been developed for identifying blast cells from microscopic blood images.

This paper proposes a new technique for detecting ALL-leukemia cells from microscopic blood images using a publicly available and widely used dataset (ALL-IDB). The rest of the paper is organized as follows. Section 2 puts forward a literature survey. Section 3 presents a complete CAD system for the detection of ALL, including sections such as an overview of system architecture, ALL-IDB dataset description, data preprocessing, feature extraction, feature fusion, classification, and performance metrics. Section 4 discusses the experimental results achieved after applying different feature extractors and comparing them with the existing approaches. Finally, Section 5 provides the conclusion of the work.

2. LITERATURE SURVEY

In biomedical image analysis and processing, machine learning and image processing methods have produced outstanding consequences, particularly in the field of ALL [8]. These methods are frequently performed in the classification on microscopic blood smear for ALL detection. For automatic classification of these disease-causing infections, a number of approaches have been effectively applied. Various procedures include convolutional neural networks (CNNs), ensemble learning, feature extraction, and feature selection, among others [9]. Recently, a number of works were executed with the help of numerous machine learning based procedures to detect and classify of leukemia microscopic blood smear images. Several transfer learning approaches, different system designs, and ensemble solutions were optimized to increase system performance in classifying acute leukemias, normal, and other disorders of the bone marrow and blood cells. A short review of some substantial contributions from the existing literature is provided.

Based on peripheral blood smear images, Al-Tahhan [10] intended to develop an improved classification model that is capable to categorize the ALL subtypes. In that article, the cytoplasmic vacuoles and the regularity of the nucleus membrane of ALL cells are the only geometric features that this approach relied on. Support Vector Machine (SVM) with multiple kernels, K-Nearest Neighbors (KNN) with various metric functions, and Artificial Neural Network (ANN)

were chosen and fine-tuned to identify automatically the subtypes of ALL utilizing ALL-IDB2 dataset. The authors in Shafique and Tehsin [4] developed a deep CNN (DCNN) model, trained on ALL-IDB dataset augmented with 50 private images, for the classification of ALL and its subtypes using pre-trained AlexNet and fine-tuning. Sharif *et al.* [11] recommended a YOLOv2-Nucleus-cytoplasm based scheme for WBCs localization using blood smear images. In that article, the Bag-of-Features were extracted from blood smear images of WBCs and optimized by using Particle Swarm Optimization for the classification task. On two challenging datasets, Leukocyte-Images for-Segmentation-Classification (LISC) and ALL-IDB, the classification results were computed. The experimental results reveal that the Optimized Naïve Bayes (O-NB) classifier outperformed the Optimized Discriminant Analysis (O-DA) classifier on ALL-IDB1, and ALL-IDB2 datasets. On the LISC dataset, however, the O-DA classifier outperformed the O-NB classifier.

Jha and Dutta [12] introduced the Sine Cosine Algorithm-based actor-critic neural network technique for leukemia classification. In that article, the developed entropy-based hybrid approach was used to segment the blood smear images, and the image-level features and statistical features were extracted from the segments. The relevant features were then fed into the designed classifier, which diagnoses leukemia using the ALL-IDB2 database. In another work [13], the authors proposed a new approach for detecting leukemia in blood images using transfer learning in CNNs and SVM for classification task. The procedure used pre-trained CNN methods (AlexNet, CaffeNet, and VGG-F) to extract images features directly without any prior preprocessing using three heterogeneous datasets. The authors of Di Ruberto *et al.* [1] developed a novel approach for leukocyte counting and classification of WBCs as healthy or affected by leukemia from microscopic blood images. The System, Man and Cybernetics-Image Data Base, Iran University of Medical Science-Image Data Base, and ALL-IDB public datasets for leukemia identification were utilized to evaluate the proposed approach. In Muthumayil *et al.* [14], the authors addressed a computer-based application technique based on Enhanced Virtual Neural Network classification for identifying and classifying CLL utilizing microscopic images of WBCs. The proposed technique attained the optimum accuracy in terms of detecting and classifying leukemia using WBCs images. In terms of sensitivity, specificity, accuracy, and misclassification error, the presented approach reached 97.8%, 89.9%, 76.6%, and 2.2%, respectively.

The authors in Singhal and Singh [15] proposed a new method for distinguishing ALL lymphoblast cells from healthy lymphocytes. Leukocytes were isolated from other blood cells first, and then lymphocytes were extracted. In this context, a novel CAD based on Gray Level Co-occurrence Matrices (GLCM) and shape-based features was developed for the diagnosis of hematological disorders such as leukemia (blood cancer). The auto SVM binary classifier was used to classify the retrieved features to determine the existence of leukemic cells. A new technique for segmenting and classifying ALL using the input blood smear images was proposed in Praveena and Singh [16]. In that article, the DCNN was trained using an optimization technique called Grey wolf-based Jaya Optimization Algorithm, which was built using the Grey Wolf Optimizer and Jaya Optimization Algorithm. Using the ALL-IDB2 dataset, the investigation was conducted by means of performance measures. The maximum accuracy, sensitivity, and specificity of the recommended approach were 0.9350, 0.9528, and 0.9389, respectively. The authors of Singhal and Singh [15] addressed an automatic leukemia detection system that used two texture descriptors, namely, Local Binary pattern (LBP) and GLCM generated from the nucleus image to detect ALL. The ALL-IDB2 dataset with 260 blood smear images (130 normal and 130 blast) was used as training data for a two-class classification. The experimental results from this paper demonstrated that LBP performed better than GLCM texture features with the classification accuracies of 93.84% and 87.30%, respectively.

The study Umamaheswari and Geetha [17] developed a scheme for optimized identification and detection of ALL using a novel customized-KNN classification model. During the preprocessing stage of this work, the medical image was prepared for segmentation through changing its size, brightness, and contrast. The nucleus portion was segmented using mathematical operators and Otsu's thresholding during the segmentation phase. Afterward, in the post-processing phase, mathematical morphological operators were used to make the nucleus portion appropriate for feature extraction task. Finally, using the developed customized KNN classification technique, the segmented portions were categorized into ALL affected and normal cells. This work was tested on over 80 images from the ALL-IDB2 dataset and achieved an average accuracy of 96.25%, sensitivity of 95%, and specificity of 97%. The authors in Tuba and Tuba [18] introduced Generative Adversarial Optimization (GAO)-based scheme for detecting acute lymphocytic leukemia as normal cells or blasts in microscopic digital blood images. In this study, the ALL-IDB2 image dataset was employed, and each blood image was characterized by 11 features

for classification tasks, along with 5 shape and 6 texture features. As a classifier, the SVM was used, and its criteria were tuned using an innovative optimization technique: The GAO technique.

Hariprasath *et al.* proposed a methodology for detection of acute lymphocytic leukemia from largest publicly accessible standard ALL-IDB dataset using statistical features. To differentiate between benign and leukemic cells, morphological and statistical features of blast cells were examined. SVM-R, SVM-L, and KNN classifiers were utilized and compared for classification process [19]. A unique strategy to automatically identify and classify ALL from peripheral blood smear images based on traditional image processing and machine learning approaches was proposed in Bodzas *et al.* [20]. In this work, to achieve the best segmentation results, substantial pre-processing and a three-phase filtration procedure were performed. Furthermore, sixteen robust features were extracted from microscopic images to discriminate between cancerous and noncancerous blood cells. Two popular machine learning classifiers, the ANN and the SVM, were computed to perform the classification task. A dataset of 31 peripheral blood smear images from a local dataset was utilized to test this approach, and they achieved a sensitivity of 100% and an average accuracy of 97.52%. The research paper [21] addresses a strategy for automatically detecting WBCs in complicated blood smear images based on Watershed Transform and circle fitting method. For the separation of overlapped WBCs, the presented technique employed segmentation and edge map extraction in preprocessing steps, in addition to parametric circle approximation, which identified both separated and overlapped WBCs. A dataset of 384 WBC images, including some overlapping cases, from ALL-IDB and ASH image bank was utilized to test this method.

This paper discusses and presents a micro-pattern descriptor, called Local Directional Number Pattern (LDNP) along with MDWT for feature extraction task. A balanced dataset with 260 blood smear images from the ALL-IDB2 dataset was used as training data. Consequently, a proposed model was constructed by applying different individual and combined feature extraction methods to overcome the most challenging parts of the detection of ALL-leukemia cells from microscopic blood images.

3. PROPOSED METHODOLOGY

3.1. System Architecture

To classify ALL-leukemia cells, the proposed approach used microscopic blood images as input data. To begin,

this method transformed RGB color images to gray scale and defined the areas of interest of healthy and blast cells by removing unwanted areas. Moreover, the system studied two feature descriptors: LDNP and MDWT. First, a feature vector was extracted from microscopic blood images of the ALL IDB2 database using the LDNP descriptor. Then the MDWT descriptor was performed to extract another feature vector from the same microscopic blood images. Various individual and combined feature extraction methods were constructed and fed into the classification model as input data. The number of features extracted by one method was insufficient to accurately classify ALL-leukemia cells. Nevertheless, using two distinct strategies to extract features could lead to a large number of features for classification task. In this circumstance, fusion was considered as a combination of the distinct feature vectors. Finally, the fused features were then classified microscopic blood images to determine cancerous and noncancerous blood cells using five well-known classifiers (Decision Tree [DT], Ensemble, KNN, Naïve Bayes [NB], and Random Forest [RF]). The significant stages of the proposed system design are shown in Fig. 2.

3.2. Dataset Description

The proposed model was trained as well as tested on a the ALL-IDB dataset containing images of leukemic blood smears along with images of non-leukemic blood smears. ALL-IDB [22] is a public image dataset which divided into two distinct versions: ALL-IDB1 and ALL-IDB2, and its images are in JPG format with a color depth of 24 bits. ALL-IDB was rated more accurate since professional oncologists classified/positioned ALL lymphoblasts for each image in the dataset. ALL-IDB1 is composed of 108 original RGB images taken with an optical laboratory microscope and either an Olympus Optical C2500L camera or a Canon PowerShot G5 camera. The first 33 images of the ALL-IDB1 are 1368×1712 pixels in size, while the rest of the images are 1944×2592 pixels in size; additionally, the images of these two image sets were collected under various magnifications, brightness, and hue staining. ALL-IDB1 depicts entire images comprising both cells or agglomerates; thus, it can be used to analyze the segmentation performance of techniques, image pre-processing methodologies, or classification approaches. ALL-IDB2 is a collection of cropped areas of interest derived from the ALL-IDB1 dataset of healthy and blast cells; furthermore, the gray level characteristics of ALL-IDB2 images are similar to those of ALL-IDB1 images. This dataset contains 260 microscopic images, 50% of these represent non-cancerous lymphocyte cells, and 50% cancerous (lymphoblast) cells. In this study, as a benchmark for evaluating the intelligence system's performance, the second version was considered.

In Fig. 3, representative example images from ALL-IDB2 dataset: Images of healthy lymphocyte and probable blast cells are given, respectively.

3.3. Preprocessing

Image preprocessing is a substantial step to improve the visualization of leukocytes in a blood image and accurate classification by removing noisy or unwanted pixels from each image. To eliminate superfluous text and machine annotation around images, the area of interest was localized (Fig. 4), cropped and resized images to a standard size 128×128 pixels after the raw images were converted from RGB to gray scale (Fig. 5). The area of interest on the microscopic blood images was determined by an area covering mostly WBC nuclei region to attain valuable information. As the raw images were captured in real life with a large variance

in exposure and contrast, image enhancement was required for superior classification performance. Consequently, the contrast enhancement of gray scale images was employed using Contrast Limited Adaptive Histogram Equalization technique and the median filter for a proper brightness and enhancement. Before importing the input images into a feature extraction stage, image adjustment was applied to improve the quality of microscopic blood smear images as shown in Fig. 6.

3.4. Feature Extraction

Feature extraction is the process of converting the raw pixel values from an image into a set of features; normally, this step comprises obtaining important features extracted from input patterns that can be used in the classification tasks [23]. In this study, two groups of features, including

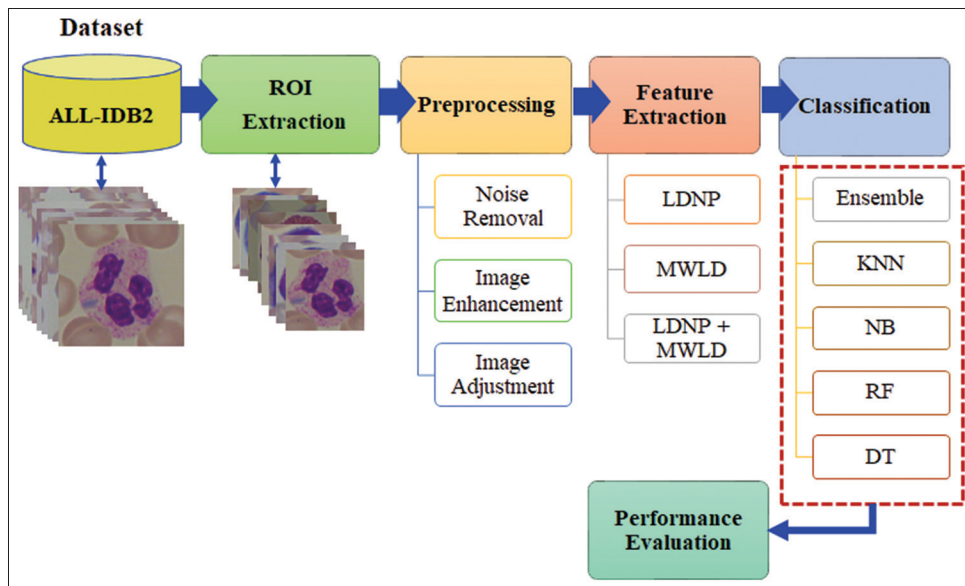


Fig. 2. Workflow of proposed system.

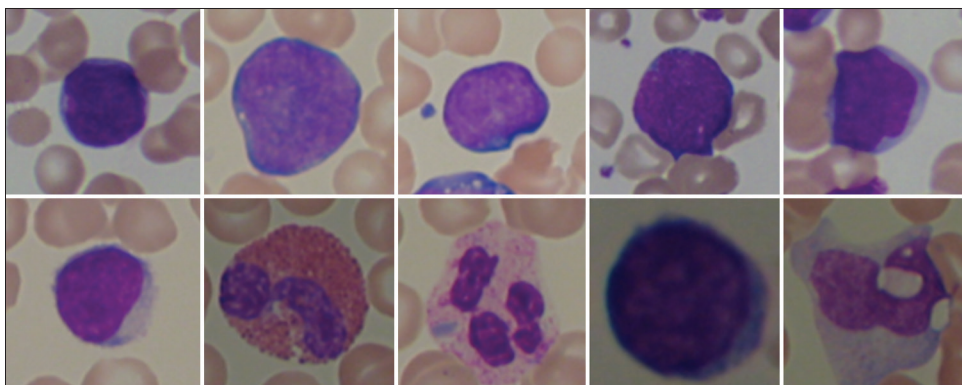


Fig. 3. Example images contained in the ALL-IDB2: healthy lymphocyte (first row) and lymphoblast cells (second row).

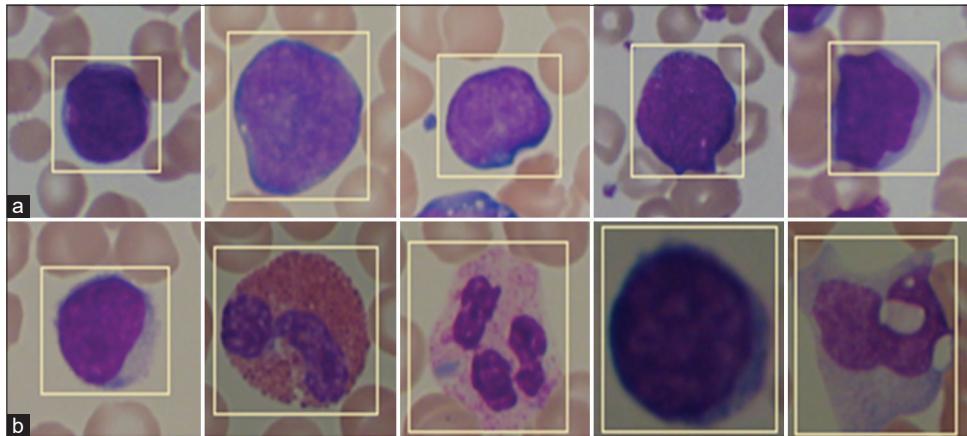


Fig. 4. Samples of blood smear images dataset used for proposed scheme; (a) localized ROI area of non-leukemic blood smears (first row), and (b) localized ROI area of leukemic blood smear cases (second row).

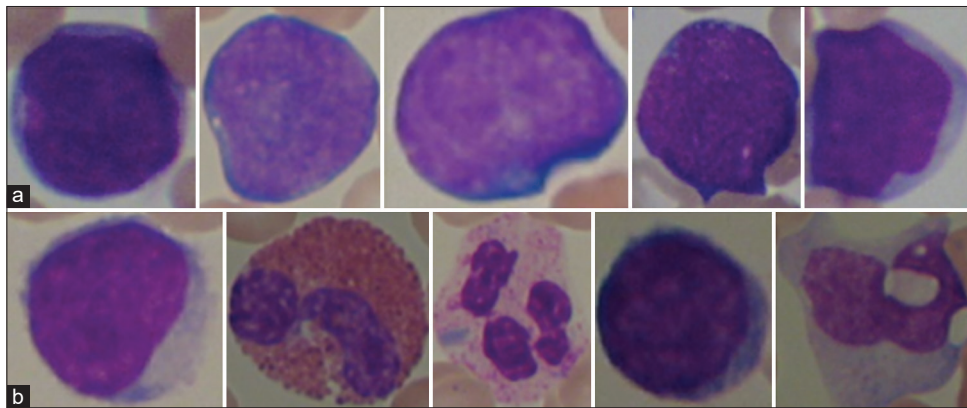


Fig. 5. Samples of blood smear images dataset used for proposed scheme; (a) cropped ROI area of healthy individuals (first row), and (b) cropped ROI area of leukemic blood smear cases (second row).

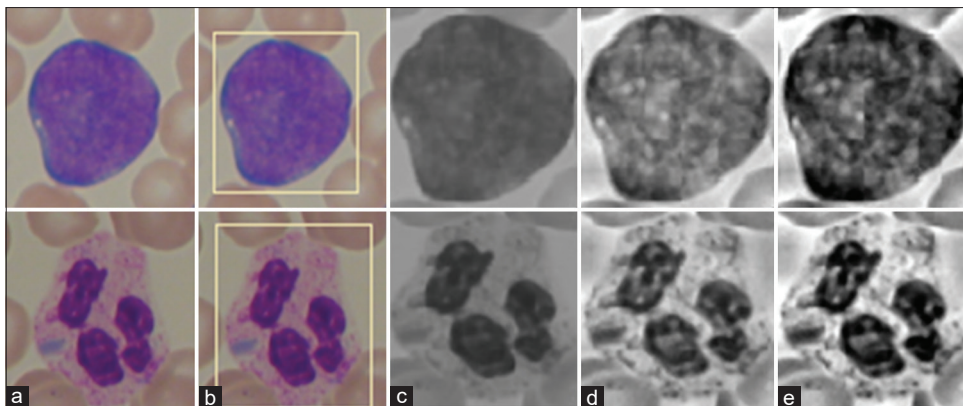


Fig. 6. Samples of blood smear images dataset used for proposed scheme; (a) original blood smear images, (b) localized ROI area, (c) denoised blood smear images, (d) contrast enhancement of blood smear images, and (e) adjusted blood smear images.

LDNP and MDWT are designed and proposed to distinguish noncancerous and cancerous blood cells from microscopic blood smear images.

3.4.1. LDNP

LDNP descriptor, introduced by Rivera *et al.* [24], is a six-bit binary code assigned to each pixel of an input image to

encode the structure information and intensity variations of a local texture pattern. The patterns are constructed using a compass mask to obtain the edge response value of the neighborhoods, as well as the direction of the top most positive and top most negative responses [25,26]. The positive and negative responses are critical in identifying the gradient direction of the neighborhood's prominent bright and dark areas. LDNP decreases the bit size per pixel from eight to a 6-bit binary code, where the first three bits represent the top positive directional number and the next three bits represent the top negative directional number. To generate the LDNP code, Kirsch compass mask (M_0, M_1, \dots, M_7) was experimented to calculate the edge response values and produce eight directional edge images or eight directional numbers for each pixel in microscopic blood smear images. Fig. 7 explains the working process of calculating bit for each pixel, and the Kirsch compass masks for eight directions [27] are revealed in Fig. 8.

To determine the meaningful descriptor for each region, LDNP is encoded using the maximum positive and maximum negative position values. The sign information is being used to determine the prominent values: The most significant bit is allocated to positive direction and the least significant bit is allocated to negative direction. Therefore, for the 3×3 neighborhood pixel window, the decimal LDNP coding is defined as:

$$LDN(x, y) = 8i_{(x,y)} + j_{(x,y)} \quad (1)$$

Where (x, y) is the central pixel value of the region, $i_{(x,y)}$ is the direction index number of the maximum edge response, and $j_{(x,y)}$ is the top most negative direction. These directional numbers can be defined by equations 2 and 3, respectively.

$$i_{(x,y)} = \text{arg,max} \{ i(x,y) | 0 \leq i \leq 7 \} \quad (2)$$

$$j_{(x,y)} = \text{arg,max} \{ j(x,y) | 0 \leq j \leq 7 \} \quad (3)$$

Where i is the convolution of the original image, I and the i^{th} mask, M^i , characterized by:

$$i = I * M^i \quad (4)$$

In this work, LDNP descriptor was performed to extract a set of 56 features for each image was extracted from the dataset of microscopic blood images.

3.4.2. Multi-scale weber local descriptors

WLD is a powerful local descriptor that is based on Weber's Law and consists of two differential excitation and

orientation component [29,30]. WLD has a number of unique features, including edge detection and resistance to changes in illumination and noise. The differential excitation component is being used to extract the most important information from an image. For each pixel, the gradient orientation component of WLD is calculated. The differential excitation $D(p_c)$ of a current pixel p_c is computed as:

$$D(p_c) = \arctan \left[\sum_{i=0}^{N-1} \left(\frac{p_i - p_c}{p_c} \right) \right] \quad (5)$$

Where p_c represents the center pixel value, p_i denotes the value of i^{th} pixel in the neighborhood, N is the total number of pixels in the neighborhood. If $D(p_c)$ is positive, then center pixel is darker respect to the neighbor pixels and if $D(p_c)$ is negative, then current pixel is lighter respect to the neighbor pixels. The orientation component of WLD is the gradient orientation that determines the directional property of the pixels. For pixel p_c , it is calculated as follows: in Equation 6.

$$\varnothing(p_c) = \arctan \left(\frac{K_s^{11}}{K_s^{10}} \right) \quad (6)$$

Where, K_s^{11} and K_s^{10} are the outputs of the filters f_{11} and f_{10} (Fig. 9), respectively.

After calculating differential excitation and gradient orientation, a concatenated two-dimensional WLD histogram using D and \varnothing is created to produce the multi-WLD that forms the representation of the image. In this work, multi-WLD was computed to extract a set of 32 features for each image was extracted from the dataset of microscopic blood images.

3.5. Feature Fusion and Classification

Data fusion has been applied in several applications for machine learning and computer vision. Feature fusion strategy, in particular, can concatenate multiple feature vectors. The multi-feature fusion can improve the robustness of the model predictions [32]. This work proposed a fusion of feature vectors attained by a combination of LDNP (1×56) and MWLD (1×32) approaches. Equations 7 and 8 represent features extracted by LDNP and MWLD, respectively. The extracted feature vectors were combined by concatenation and represented by Equation 9.

$$F_{LDNP_{1 \times n}} = \{LDNP_{1 \times 1} + LDNP_{1 \times 2} + LDNP_{1 \times 3} \dots \dots \dots LDNP_{1 \times n}\} \quad (7)$$

$$F_{MWLD_{1 \times n}} = \{MWLD_{1 \times 1} + MWLD_{1 \times 2} + MWLD_{1 \times 3} \dots \dots \dots MWLD_{1 \times n}\} \quad (8)$$

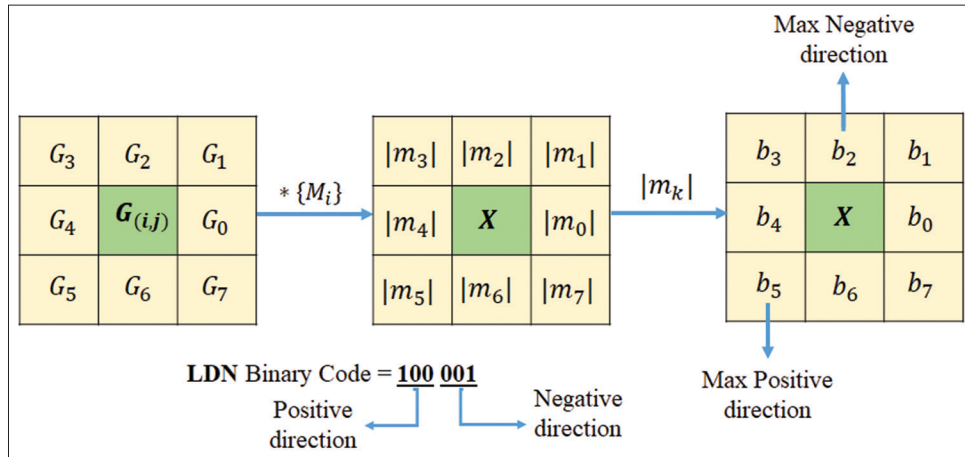


Fig. 7. The procedure of calculating LDNP code.

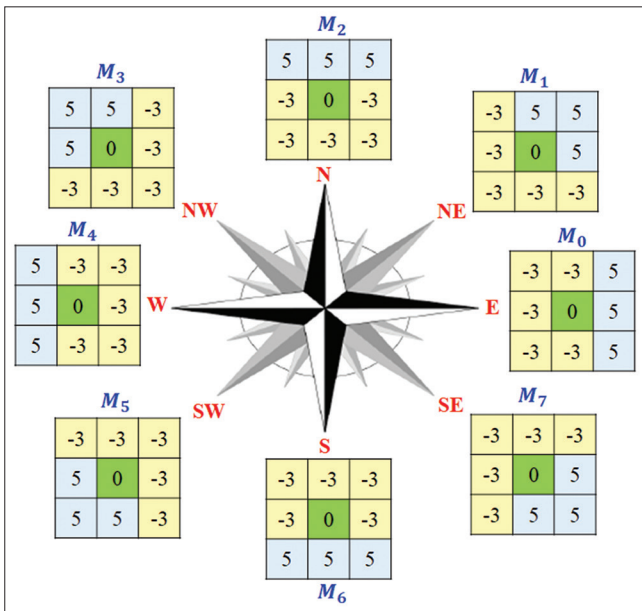


Fig. 8. Kirsch compass masks for eight directions [28].

f_{00}			f_{10}			f_{11}		
+1	+1	+1		-1				
+1	-8	+1				+1		-1
+1	+1	+1		+1				

Fig. 9. Filters used in simple WLD calculation [31].

$$Fused (Features vector)_{1 \times p}^{cat} = \{F_{LDNP_{1 \times n}}, F_{MWLD_{1 \times m}}\} \quad (9)$$

Then the features extracted by LDNP and MWLD were fused with 88 features. This fusion vector, which considered as the

final input for the training and testing dataset, was fed to the classifiers to validate the proposed approach and identify blast cells from microscopic blood images. In the proposed workflow, machine learning models were used to identify patients affected by leukemia. To achieve the objective of identifying leukemia patients among normal healthy individuals, five pattern recognition classifiers, namely, DT, Ensemble, KNN, NB, and RF classifiers were separately executed.

3.6. Model Evaluation

Model evaluation is used to estimate the parameter space and feature extraction results from various models. For the classification of ALL images, five widely used performance metrics were applied to assess the proposed model's performance: accuracy, recall, precision, F-measure, and MCC metrics. To calculate the metrics specified by Equations 10-14, four distinct performance parameters were used: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (13)$$

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (14)$$

4. RESULTS AND DISCUSSION

The performance of the proposed system was evaluated based on extracted features derived from LDNP and

MWLD coefficients to identify automatically the class of the ALL images. The evaluation was performed using a dataset of 260 blood smear images, comprising 130 blasts and 130 non-blast cells. All experiments were conducted in MATLAB (R2021b) environment using microscopic blood images described in Section 3.2. Different scenarios using different features individually and combination of (LDNP + MWLD) features have been suggested. For each scenario, these extracted features were classified using five pattern recognition classifiers (i.e., DT, Ensemble, KNN, NB, and RF) to see which scenario could perform better classification performance. Furthermore, the entire dataset was divided into two groups: About 80% for training the model and 20% for evaluation of the classification performance using holdout cross-validation method. Performances of the suggested scenarios were analyzed through a number of different measures including accuracy, recall, precision, F-measure, MCC, and misclassification error rate computed from confusion matrix.

From the experiments performed on ALL-IDB2 dataset, the detailed category/class wise analysis of each scenario was estimated in terms of accuracy and average accuracy as (mean ± SD) with DT (Table 1), Ensemble (Table 2), KNN (Table 3), NB (Table 4), and RF (Table 5) classifiers respectively. According to the results in Tables 1-5, it is concluded that combining the LDNP and MWLD features

(scenario 3) together attain the highest average accuracy of $97.69 \pm 1.83\%$ followed by LDNP features (scenario 1) with $97.11 \pm 1.44\%$ along with Ensemble classifier, while the classification average accuracy of features derived from MWDT method had the lowest scoring ($79.42 \pm 4.8\%$) with NB classifier.

Based on the experimental outcomes for all three scenarios depicted in Fig. 10, it can be concluded that the fusion of features extracted from LDNP and MWLD methods has a positive impact on the performance and outperformed the other two scenarios with all classifiers. The consequences revealed that the fusion LDNP and MWLD techniques attained the highest average accuracy of $97.69 \pm 1.83\%$, $96.92 \pm 1.15\%$, $95.19 \pm 1.63\%$, $93.84 \pm 3.11\%$, and $89.03 \pm$

TABLE 1: Performance analysis on average accuracy with Decision Tree classifier

Descriptors	Features	Per Class Accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	82.3±8.54	83.46±7.03	82.88±5.62
MWDT	32	89.61±6.02	85.38±8.06	87.50±5.96
LDNP+ MWDT	88	91.92±5.57	86.15±6.58	89.03±3.63

The highlighted accuracy in bold indicates the best classification result. LDNP: Local directional number pattern

TABLE 2: Performance analysis on average accuracy with Ensemble classifier

Descriptors	Features	Per Class Accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	97.3±1.07	96.92±1.03	97.11±1.44
MWDT	32	94.99±2.07	95.76±2.13	95.38±2.59
LDNP+ MWDT	88	99.23±0.02	96.15±1.73	97.69±1.83

The highlighted accuracy in bold indicates the best classification result. LDNP: Local directional number pattern

TABLE 3: Performance analysis on average accuracy with KNN classifier

Descriptors	Features	Per class accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	99.23±0.12	92.69±3.27	95.96±1.78
MWDT	32	96.15±1.14	84.99±4.39	90.57±3.89
LDNP+ MWDT	88	96.15±1.24	97.69±1.34	96.92±1.15

The highlighted accuracy in bold indicates the best classification result. KNN: K-Nearest Neighbors, LDNP: Local directional number pattern

TABLE 4: Performance analysis on average accuracy with NB classifier

Descriptors	Features	Per class accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	91.15±3.45	93.07±4.22	92.11±4.29
MWDT	32	75.38±7.51	83.46±5.74	79.42±4.80
LDNP+ MWDT	88	94.23±2.88	93.46±4.02	93.84±3.11

The highlighted accuracy in bold indicates the best classification result. NB: Naïve Bayes, LDNP: Local directional number pattern

TABLE 5: Performance analysis on average accuracy with RF classifier

Descriptors	Features	Per Class Accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	96.92±1.05	90.38±6.34	93.65±3.39
MWDT	32	93.84±3.19	92.69±2.23	93.26±2.60
LDNP+ MWDT	88	97.69±1.68	92.69±2.36	95.19±1.63

The highlighted accuracy in bold indicates the best classification result. RF: Random forest, LDNP: Local directional number pattern

3.63% using Ensemble, KNN, RF, NB, and DT classifiers, respectively. On the other hand, the fusion of fusion LDNP and MWLD techniques with Ensemble classifier was sufficient to record maximal average accuracy performance of $97.69 \pm 1.83\%$ among the remaining classifiers for all scenarios.

The same fact has been concluded by examining other performance measures (precision, recall, F-measure, and MCC) to estimate the proposed framework. For all classifiers, the best precision rate was achieved with a set of features fusion LDNP and MWLD methods (scenario 3) and outperformed the other scenarios. The results of scenario 3 provided that all five classifiers reached the highest precision of $98.05 \pm 1.74\%$, $97.77 \pm 1.09\%$, $93.85 \pm 5.21\%$, $93.15 \pm 2.90\%$, and $87.25 \pm 5.33\%$ using Ensemble, KNN, NB, RF, and DT classifiers, respectively; however, the lowest precision rate of $82.14 \pm 5.39\%$ was recorded using MWDT method with NB classifier. Furthermore, the experiment verified maximum precision performance of $98.05 \pm 1.74\%$ with Ensemble classifier among the remaining classifiers for all scenarios. Comparison results of precision rates of all system scenarios with all classifiers are displayed in Table 6.

Furthermore, the fusion of features from LDNP and MWDT methods also performed the best in terms of recall rates as $97.69 \pm 1.68\%$, $97.30 \pm 1.14\%$, $96.15 \pm 1.14\%$, and $94.23 \pm 1.88\%$ was attained using RF, Ensemble, KNN, and NB classifiers, respectively (Table 7); furthermore,

MWDT performed the best recall rate of $93.46 \pm 2.81\%$ with DT classifier; conversely, the lowest recall rate of $75.38 \pm 7.51\%$ has been recorded when MWDT method was utilized with NB classifier. Based on the experimental results for all five classifiers depicted in Table 7, it can be verified that the value of recall rate with RF classifier was overall superior and outperformed the other classifiers. With regard to the F-measure rates, the results presented in Table 8 demonstrate the superiority of the fusion of LDNP and MWDT scenario and it was undoubtedly yielded excellent results which defiantly go beyond the other scenarios. The best performance with F-measure rate of $97.63 \pm 1.62\%$ was achieved using fusion of LDNP and MWDT features with Ensemble classifier, which surpassed other classifiers with F-measure rates of $96.89 \pm 1.86\%$, $95.31 \pm 1.56\%$, and $93.88 \pm 3.11\%$ for KNN, RF, and NB classifiers, respectively; however, the F-measure rate of MWDT method using NB classifier had the lowermost recording of $78.45 \pm 5.43\%$.

With respects to the MCC rates, the outcomes depicted in Table 9 prove the superiority of the fusion of LDNP and MWDT scenario and it was unquestionably returned outstanding outcomes which defiantly go beyond the other scenarios. The finest performance with F-measure rate of $95.46 \pm 2.85\%$ was achieved using fusion of LDNP and MWDT features with Ensemble classifier, followed by F-measure rate of $93.97 \pm 3.67\%$ along with KNN classifier, while the classification F-measure rate of features derived from MWDT method had the lowest scoring ($59.26 \pm 9.59\%$) with NB classifier.

The experimentations from Fig. 11 besides obviously confirmed that the features derived by fusion of LDNP and MWDT scenario outperformed other scenarios and recorded the highest precision, recall, F-measure, and MCC rates with Ensemble classifier. Considering the obtained outcomes, the highest precision, recall, F-measure, and MCC scores of the features extracted using fused LDNP and MWDT methods were 98.05% , 97.30% , 97.63% , and 95.46% respectively, and were achieved using 88 effective features. While, the lowest precision, recall, F-measure, and MCC rates were achieved

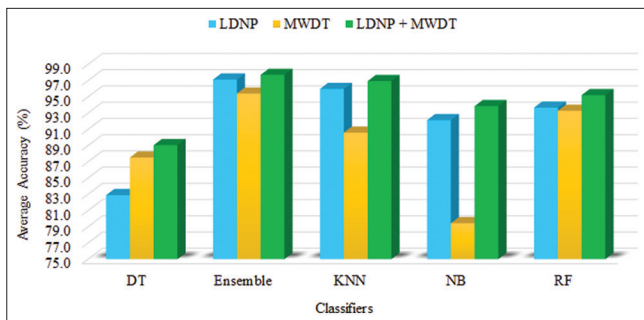


Fig. 10. Comparison of average accuracies for modeling the various scenarios using different classification methods.

Methods	DT	Ensemble	KNN	NB	RF
LDNP	83.51±5.90	97.01±1.89	93.33±4.41	93.25±5.96	91.24±5.02
MWDT	87.03±6.30	96.24±1.99	86.30±5.93	82.14±5.39	92.96±3.71
LDNP+MWDT	87.25±5.33	98.05±1.74	97.77±1.09	93.85±5.21	93.15±2.90

The best result per row is highlighted in bold. DT: Decision tree, KNN: K-Nearest Neighbors, NB: Naïve Bayes, RF: Random forest, LDNP: Local directional number pattern

TABLE 7: The average classification recall in % over 10 runs for each scenario scenarios using different classification algorithms

Methods	DT	Ensemble	KNN	NB	RF
LDNP	82.30±8.54	97.20±1.07	96.13±1.02	91.15±5.45	96.92±2.03
MWDT	93.46±2.81	96.53±2.36	96.15±1.44	75.38±7.51	93.84±3.19
LDNP+MWDT	91.92±5.57	97.30±1.14	96.15±1.14	94.23±1.88	97.69±1.68

The best result per row is highlighted in bold. DT: Decision tree, KNN: K-Nearest Neighbors, NB: Naive Bayes, RF: Random forest, LDNP: Local directional number pattern

TABLE 8: The average classification MCC in % over 10 runs for each scenario scenarios using different classification algorithms

Methods	DT	Ensemble	KNN	NB	RF
LDNP	66.11±11.17	94.34±4.78	92.22±5.21	84.49±8.61	87.67±6.37
MWDT	79.62±10.19	92.78±4.57	81.27±7.68	59.26±9.59	86.78±5.17
LDNP+MWDT	78.55±7.27	95.46±2.85	93.97±3.67	88.00±6.04	90.60±3.18

The best result per row is highlighted in bold. DT: Decision tree, KNN: K-Nearest Neighbors, NB: Naive Bayes, RF: Random forest, LDNP: Local directional number pattern

TABLE 9: The average classification MCC in % over 10 runs for each scenario scenarios using different classification algorithms

Methods	DT	Ensemble	KNN	NB	RF
LDNP	66.11±11.17	94.34±4.78	92.22±5.21	84.49±8.61	87.67±6.37
MWDT	79.62±10.19	92.78±4.57	81.27±7.68	59.26±9.59	86.78±5.17
LDNP+MWDT	78.55±7.27	95.46±2.85	93.97±3.67	88.00±6.04	90.60±3.18

The best result per row is highlighted in bold. DT: Decision tree, KNN: K-Nearest Neighbors, NB: Naive Bayes, RF: Random forest, LDNP: Local directional number pattern

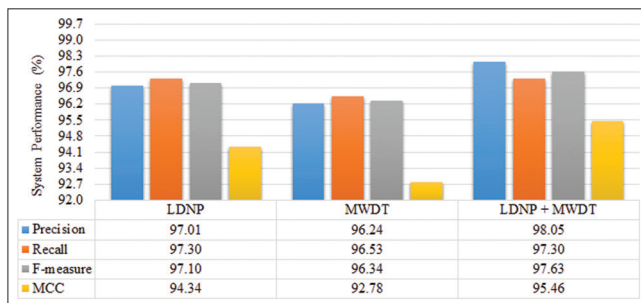


Fig. 11. Comparison of system performance for a different scenario using Ensemble classifier.

using MWDT method with a score of 96.24 %, 96.53%, 96.34%, and 92.78%, respectively, and were achieved with 32 extracted features.

In this work, the performance of the proposed scenarios was also evaluated through misclassification error rate metric used the same dataset and computing environment. As confirmed by Fig. 12, the misclassification error rates for the proposed scenarios were measured. The results prove that the fusion of LDNP and MWDT scenario with Ensemble

classifier outcomes in a lower misclassification error of 2.31% rate which confirms that the proposed scenario performs considerably much better than other suggesting scenarios. Thus, this scenario was chosen as a proposed technique for the classification of ALL images.

Finally, the performance of the proposed fusion scheme was also compared with some existing state-of-the-art methods as revealed in Table 10. The proposed scheme provides a promising outcome particularly in terms of average classification accuracy when comparing with the existing approaches. This is due to the combination carried out between LDNP and MWDT approaches which led to gaining their advantages. Nevertheless, the other researchers used some huge number of features, whereas in the proposed scheme, 88 features were utilized with the best performance results attained.

From the above experimental outcomes, it is obviously noticeable that the proposed system can effectively applied discriminates the normal cell cases from blasts in microscopic blood images more precisely, which might support the clinicians to made up a clear diagnosis based on their clinical specialists as well as the proposed tool as a second opinion.

TABLE 10: Comparison of proposed classification accuracy with recent techniques

Author	ALL-IDB2 datasets		Accuracy (%)
	Method	Classifier	
Umamaheswari and Geetha [17]	GLCM+Statistical+Geometrical Features	C-KNN	96.25
Tuba and Tuba [18]	Shape+Texture features	GAO-based methods	93.84
Singhal and Singh [15]	LBP	-----	93.84
	GLCM		87.30
Praveena and Singh [16]	GreyJOA	Deep CNN	93.50
Proposed work	LDNP+MWDT	Ensemble	97.69

LBP: Local binary pattern, GLCM: Gray level co-occurrence matrices, C-KNN: C-K-Nearest Neighbors, GAO: Generative adversarial optimization, LDNP: Local directional number pattern, CNN: Convolutional neural network, GreyJOA: Grey wolf-based Jaya optimization algorithm. The best result per row is highlighted in bold.

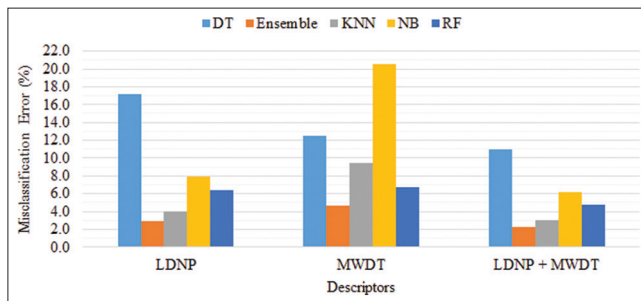


Fig. 12. Comparison of misclassification error rate for a different scenario.

5. CONCLUSION

Early diagnosis of leukemia especially acute leukemia in patients, gives the chance to cure cancer with the right treatment. Image processing-based technologies can be used to automatically and effectively analyze microscopic smear images in order to detect the incidence of leukemia. The main objective of this paper is to use feature fusion and a machine learning model to perform leukemia classification using input blood smear images. Each trained model was evaluated using benchmark performance metrics, for example, accuracy, precision, recall, F-measure, MCC, and misclassification error rate under three different scenarios concerned with balanced learning and classification method. The proposed method was tested on images of microscopic thin blood smears obtained from the publicly available leukemia benchmark dataset, that is, ALL-IDB2. The proposed feature fusion pipeline presented a higher average classification accuracy ($97.69 \pm 1.83\%$) compared to the accuracies attained using features achieved by individual feature extraction approaches, such as LDNP and MWDT. Furthermore, experimental outcomes revealed that the proposed model is more effective than previous works carried out for the classification of ALL.

REFERENCES

[1] C. Di Ruberto, A. Loddo and G. Puglisi G. "Blob detection and deep learning for leukemic blood image analysis". *Applied Sciences*,

vol. 10, no. 3, p. 1176, 2020.

- [2] G. Dralus, D. Mazur and A. Czml. "Automatic detection and counting of blood cells in smear images using retinanet". *Entropy*, vol. 23, no. 11, p. 1522, 2021.
- [3] B. George-Gay and K. Parker. "Understanding the complete blood count with differential". *Journal of PeriAnesthesia Nursing*, vol. 18, pp. 96-117, 2003.
- [4] S. Shafique and S. Tehsin. "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks". *Technology in Cancer Research and Treatment*, vol. 17, pp. 1-7, 2018.
- [5] D. A. Arber, A. Orazi, R. Hasserjian, J. Thiele, M. J. Borowitz, M. Le Beau, C. D. Bloomfield, M. Cazzola and J. W. Vardiman. "The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia". *Blood*, vol. 127, pp. 2391-2405, 2016.
- [6] F. Huang, P. Guang, F. Li, X. Liu, W. Zhang and W. Huang. "AML, ALL, and CML classification and diagnosis based on bone marrow cell morphology combined with convolutional neural network: A STARD compliant diagnosis research". *Medicine (Baltimore)*, vol. 99, p. e23154, 2020.
- [7] Y. Dong, O. Shi, Q. Zeng, X. Lu, W. Wang, Y. Li and Q. Wang. "Leukemia incidence trends at the global, regional, and national level between 1990 and 2017". *Experimental Hematology and Oncology*, vol. 9, pp. 1-11, 2020.
- [8] M. Ghaderzadeh, F. Asadi, A. Hosseini, D. Bashash, H. Abolghasemi and A. Roshanpour. "Machine learning in detection and classification of leukemia using smear blood images: A systematic review". *Scientific Programming*, vol. 2021, p. 9933481, 2021.
- [9] M. Kim, K. Chae, S. Lee, H. J. Jang and S. Kim. "Automated classification of online sources for infectious disease occurrences using machine-learning-based natural language processing approaches". *International Journal of Environmental Research and Public Health*, vol. 17, pp. 1-13, 2020.
- [10] F. E. Al-Tahhan, M. E. Fares, A. A. Sakr and D. A. Aladle. "Accurate automatic detection of acute lymphatic leukemia using a refined simple classification". *Microscopy Research and Technique*, vol. 83, pp. 1178-1189, 2020.
- [11] M. Sharif, J. Amin, A. Siddiq, H. U. Khan, M. S. A. Malik, M. A. Anjum and S. Kadry. "Recognition of different types of leukocytes using YOLOv2 and optimized bag-of-features". *IEEE Access*, vol. 8, pp. 167448-167459, 2020.
- [12] K. K. Jha and H. S. Dutta. "Nucleus and cytoplasm-based segmentation and actor-critic neural network for acute lymphocytic leukaemia detection in single cell blood smear images". *Medical and Biological Engineering and Computing*, vol. 58, pp. 171-186, 2020.

- [13] L. H. S. Vogado, R. M. S. Veras, F. H. D. Araujo, R. R. V. Silva and K. R. T. Aires. "Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification". *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 415-422, 2018.
- [14] K. Muthumayil, S. Manikandan, S. Srinivasan, J. Escorcia-Gutierrez, M. Gamarra and R. F. Mansour. "Diagnosis of leukemia disease based on enhanced virtual neural network". *Computers, Materials and Continua*, vol. 69, pp. 2031-2044, 2021.
- [15] V. Singhal and P. Singh. "Texture Features for the Detection of Acute Lymphoblastic Leukemia." 2016.
- [16] S. Praveena and S. P. Singh. "Sparse-FCM and deep convolutional neural network for the segmentation and classification of acute lymphoblastic leukaemia". *Biomed Technologies*, vol. 65, pp. 759-773, 2020.
- [17] D. Umamaheswari and S. Geetha. "A framework for efficient recognition and classification of acute lymphoblastic leukemia with a novel customized-KNN classifier". *The Journal of Computer Engineering and Information Technology*, vol. 26, pp. 131-140, 2018.
- [18] M. Tuba and E. Tuba. "Generative adversarial optimization (GOA) for acute lymphocytic leukemia detection". *Studies in Informatics and Control*, vol. 28, pp. 245-254, 2019.
- [19] S. Hariprasath, T. Dharani, S. Mohammad and N. Bilal. "Automated Detection of Acute Lymphocytic Leukemia Using Blast Cell Morphological Features". 2nd International Conference on Advances in Science and Technology (ICAST) 2019 on 8th, 9th April 2019 by K J Somaiya Institute of Engineering and Information Technology, Mumbai, India, 2019.
- [20] A. Bodzas, P. Kodytek and J. Zidek. "Automated detection of acute lymphoblastic leukemia from microscopic images based on human visual perception". *Frontiers in Bioengineering and Biotechnology*, vol. 8, pp. 1-13, 2020.
- [21] K. N. Sukhia, M. M. Riaz, A. Ghafoor and N. Iltaf. "Overlapping white blood cells detection based on watershed transform and circle fitting". *Radioengineering*, vol. 26, pp. 1177-1181, 2017.
- [22] Conference II, Processing I. *ALL-IDB : The Acute Lymphoblastic Leukemia Image Database for Image Processing Ruggero Donida Labati*. Vincenzo Piuri, Fabio Scotti Università degli Studi di Milano, Department of Information Technology, IEEE International Conference on Image Processing, pp. 2089-2092, 2011.
- [23] F. H. Ahmad and S. H. Wady. "COVID-19 infection detection from chest X-ray images using feature fusion and machine learning". *Scientific Journal*, Vol. 5, pp. 10-30, 2021.
- [24] A. R. Rivera, J. R. Castillo and O. Chae. "Local directional texture pattern image descriptor". *Pattern Recognition Letters*, vol. 51, pp. 94-100, 2015.
- [25] R. R. Rose, A. Suruliandi and K. Meena. "Local texture description framework-based modified local directional number pattern: A new descriptor for face recognition". *International Journal of Biometeorology*, vol. 7, pp. 147-169, 2015.
- [26] S. P. Ramalingam. "Paturu Venkata Subbu Sita Rama CM. Dimensionality reduced local directional number pattern for face recognition". *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, pp. 95-103, 2018.
- [27] R. A. Kirsch. "Computer determination of the constituent structure of biological images". *Computers and Biomedical Research*, vol. 4, pp. 315-328, 1971.
- [28] A. El Idrissi, Y. El Merabet and Y. Ruichek. "Palmprint recognition using state-of-the-art local texture descriptors: A comparative study". *IET Biometrics*, vol. 9, pp. 143-153, 2020.
- [29] Sachinkumar and S. Raga. "Breast cancer detection technique based on multi-subspace randomization and collaboration". *International Journal of Advanced Research in Engineering and Technology*, vol. 11, pp. 687-701, 2020.
- [30] A. Banerjee, N. Das and K. C. Santosh. "Weber local descriptor for image analysis and recognition: A survey". *The Visual Computer*, vol. 38, pp. 321-243, 2020.
- [31] M. Hussain, S. Qasem, G. Bebis, G. Muhammad, H. Aboalsamh and H. Mathkour. "Evaluation of image forgery detection using multi-scale weber local descriptors". *International Journal on Artificial Intelligence Tools*, vol. 24, pp. 416-424, 2015.
- [32] S. Ahmed. "Ethnicity identification based on fusion strategy of local and global features extraction". *International Journal of Multidisciplinary and Current Research*, vol. 4, pp. 200-205, 2016.

An Intelligent and Precise Method Used for Detecting Gestational Diabetes in the Early Stages



Safa Abdul Wahab Hameed, Alaa Badeea Ali

Department of Computer Science, Faculty of Engineering and Science, Bayan University, Erbil, Iraq

ABSTRACT

This paper suggests a Naive Bayes classifier technique for identifying and categorizing gestational diabetes mellitus (GDM), GDM is a kind of diabetes mellitus that affects a small proportion of pregnant women but recovers to normal once the baby is born. The Pima Indians Diabetes Dataset was chosen for a comprehensive analysis of this critical and pervasive health disease because it contains 768 patient characteristics acquired from a machine learning source at the University of California, Irvine. The goal of the study is to apply smart technology to categorize diseases with high accuracy and precision, practically free of conceivable and potential faults, to provide satisfying findings. The approach is based on eight major characteristics that are present in the operations that are required to establish a precise and reliable categorization system. This approach involves training and testing on real data, as well as for deciding whether or not to construct a categorization model. The work was compared to earlier work and had a 96% accuracy rating.

Index Terms: Classifier, Feature Selection, Gestational Diabetes, Machine Learning, Naïve Bayes

1. INTRODUCTION

In today's world, diabetes is one of the most frequent diseases [1], whereas diabetes is a non-communicable disease that has a significant impact on people's health today [2]. It is a chronic condition or collection of metabolic diseases in which a person's blood glucose levels remain elevated for an extended period due to insufficient insulin synthesis or improper insulin response by the body's cells [3]. Diabetes mellitus (DM) is a condition that affects more than 60% of the population and has a high mortality rate [4], this disease has increased at an exponential rate in recent years, according to a statistical study published on the World Health Organization's website.

The number of diabetic patients worldwide has increased significantly, from 108 million in 1980 to 422 million in 2014 [5]. A type of diabetes is gestational diabetes. Gestational diabetes is a kind of diabetes that develop during pregnancy [6]. Changes in dietary habits, increased spending power, and climate change, among other factors, are all contributing to an increase in the number of women with gestational diabetes aggravated by pregnancy [7]. Gestational DM (GDM) is a type of glucose intolerance that develops during pregnancy and can cause difficulties for both the mother and the fetus [8]. Women are also more likely to have diabetes-related comorbidities such as renal disease, depression, and poor vision [9]. It can be detected early, which reduces the patient's health risk [10]. However, if the illness is not treated promptly, it can have serious consequences for the kidneys, brain system, retina of the eyes, and heart problems [11], to diagnose diabetes, medical experts require a technique of prediction [12].

Using the PIMA dataset and a variety of machine learning (ML) algorithms, it is feasible to give an advanced technique

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp34-42

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Copyright © 2022 Hameed and Ali. This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Safa Abdul Wahab Hameed, Alaa Badeea Ali, Department of Computer Science, Faculty of Engineering and Science, Bayan University, Erbil, Iraq. E-mail: safa.hamid@bnu.edu.iq; alaa.baban@bnu.edu.iq

Received: 07-12-2021

Accepted: 18-02-2022

Published: 20-03-2022

for diabetes prediction. Because deep learning algorithms may be used in a variety of ways in this industry, Models based on Artificial Neural Networks (ANN) and the Quasi-Newton technique, for example [13]. This area lends itself well to ML methods. Models are trained using ML techniques. There are three types of ML algorithms: Supervised learning (in which datasets are labeled and Regression and Classification techniques are used), unsupervised learning (in which datasets are not labeled and techniques such as dimensionality reduction and clustering are used), and reinforcement learning (in which the model learns from its every action). ML is a rapidly developing new technology with several applications [14]. ML techniques have advanced at a breakneck pace and are now widely used in a variety of medical applications [15], one of the most commonly explored challenges by DM and ML researchers is classification [16]. One of the most crucial parts of supervised learning is classification. Picking the proper classification model is a trade-off between performance, the execution time of models, and scalability. Parameter adjustment should also be considered in order to improve model performance. ML training data are an important input to an algorithm that comprehends and memorizes information from such data to predict the future. Understanding the significance of the training set in ML can assist you in obtaining the appropriate quality and amount of training data for your model training. Once you understand why it's essential and how it influences model prediction, you'll be able to select the best method based on the availability and compatibility of your training data set [17]. It is vital to develop predictive algorithms that are both accurate and simple to use when evaluating large amounts of data and converting it into useful information [18] ML methods are commonly utilized for detection and classification [19]. This diagnosis allows for proper treatment to begin as soon as feasible, avoiding deaths [20], [21]. People will be able to seek treatment for this condition if it can be detected and predicted at an early stage [22]. This type of disease is the gestational diabetes is the focus of this research, in this research an effective method was used, which is a Naïve Bayes classifier, it used to detect and identify gestational diabetes and give high performance and accurate results.

2. LITERATURE REVIEW

Similar works on diabetes analysis, prediction, and diagnosis are reviewed in this section. It uses a variety of classification and ML algorithms to handle diabetes management prediction problems.

In Pradhana *et al.* [5], the suggested methodology analyzed publically accessible data collected from diabetic patients to

identify the causes of diabetes, the most afflicted age groups, job styles, and eating patterns. ANN are used in the model to detect diabetes and determine its kind. The authors utilized the “Pima Indian Diabetes” dataset, which has the maximum accuracy of 85.09%. 768 patients’ medical histories are included in the dataset. Where as in the Filho *et al.* [7], the suggested method improved the accuracy of the classification techniques by focusing on identifying the features that fail in early diagnosis of Diabetes Miletus utilizing Predictive analysis using Support vector machine (SVM) and Naïve Base (NB) algorithms. The accuracy of the improved SVM is 77%, while the accuracy of the NB is 82.30%. In Prasanth *et al.* [2] explained, the captured data were fed into supervised ML techniques. The Pima Indians Diabetes Dataset was utilized in this study, and a model was created using SVM, CatBoost, and Relative frequency to predict DM, with an accuracy of 86.15%. And in Rawat and Suryakant [18], a comparison was done between suggested approaches and previously published studies. In this work, five ML algorithms, AdaBoost, LogicBoost, RobustBoost, Naive Bayes, and Bagging, were proposed for the analysis and prediction of DM patients. The suggested methodologies were tested on a data set of Pima Indians with diabetes, and the proposed algorithm, Bagging, was applied on the same database with an accuracy of 81.77%.

The two algorithms Naive Bayes and SVM used in Gupta *et al.* [1] as classification models, and feature selection to improve the model's accuracy. The accuracy, precision, and recall values were used to evaluate the results. The model's improved performance was calculated using the k-fold cross-validation technique. In Rajivkannan and Aparna [22], the aim of this study was to build an objective method to evaluate DM risk from past GDM data recorded 15 years ago and find a shortlist of the most informative indicators. The research steps involve pre-processing data to evaluate missing values MVs, finding the most informative attributes, and testing standard classification algorithms to combine into the most effective voting meta-algorithm. Meta-algorithm-based classification of limited anamnestic GDM related data for DM prediction is proving. Relative frequency of occurrence (RFO) analysis of attributes combined with voting meta-algorithm helped find the optimal amount of attributes giving the best possible classification result. The algorithm applied to two-class data set with 12 selected attributes produced an accuracy of 75.85. In [17], the major goal of this research is to look at different forms of machine-learning classification algorithms and compared them. In this study, use machine-learning classification algorithms to detect the start of diabetes in diabetic patients. The top performing algorithm, Logistic Regression, has an accuracy of 80%. In Moon [21], the research provided a biological ML

method that is both efficient and effective which is applied on Pima Indian diabetic database (PIDD). The proposed ensemble of SVM and back-propagation neural network (BP-NN) tested on diabetes diagnosis; one of the most frequently investigated topics in bioinformatics. The findings reveal an accuracy of 88.04% on this problem. In Sanakal and Jayakumari [16], the application of Fuzzy C-means clustering (FCM), FCM, and SVM on a collection of medical data linked to diabetes diagnostic difficulties was the subject of this work. The medical data comprises nine input variables relating to diabetes clinical diagnosis and one output attribute that indicates whether or not the patient has been diagnosed with diabetes. FCM had the best outcome, with an accuracy of 94.3%. In Lavanya and Rani [23], this work provided a quick overview of the old and new data mining approaches utilized in diabetes. This study used Fcm and Svm procedures and got an accuracy of 94.3%. In Sarwar *et al.* [4], the results showed that the ensemble approach had a 98.60% accuracy, which combines the predictive performance of numerous AI-based algorithms and is superior to all other individual competitors. ANN, Naive Bayes, SVM, and K-Nearest Neighbor are the techniques that were more precise than the others (K-NN). About 400 persons were included in the database, which came from all across the world. And in Resti *et al.* [9] a model validation built which based on 5-fold cross-validation, which divided the data into training and test data. The Gaussian Nave Bayes was the best strategy for predicting diabetes diagnosis, according to the model validation results. The contribution of this research was that the Multinomial Naïve Bayes method's performance metrics all exceed 93%. With the same explanatory factors, these findings are useful in predicting diabetes status. In this paper, the proposed method is implemented on Pima Indians Diabetes Dataset. Here in this paper, the work was presented using a Naive Base algorithm. The work was carried out in stages that relied on training and testing on real data based on certain characteristics as explained later, and a classification result was obtained with high efficiency and accuracy.

3. METHOD

This work is done to diagnose gestational diabetes by classification using a Naive Base algorithm. This work includes several stages, including training and testing on real data, to adopt and use the system. Following the stages to implement the method:

3.1. System Diabetic Detection

The Diabetes Classification System includes two phases; the first is the training stage, which has particular functions such as

reading the diabetes dataset, feature selection, discretization, and the classifier model used to create the decision rules. The second stage is the testing stage, which includes the following particular functions: read data set, discretization, decision rules, and output [1], [24], as illustrated in Fig. 1.

3.2. Data Set

The data set has been taken from 768 women, (500 negative cases and 268 positive cases) from 21 years old and above, and eight recorded features as follows:

- Number of previous pregnancies
- Plasma glucose concentration at 2-h in an oral glucose tolerance test
- Diastolic blood pressure (mm Hg)
- Triceps skinfold thickness (mm)
- 2-H serum insulin (μ U/ml),
- Body mass index (weight in kg/[height in m]²)
- Diabetes pedigree function
- Age (years)

There are a variety of causes for incomplete data, including patient death, device problems, and respondents' reluctance to answer particular questions.

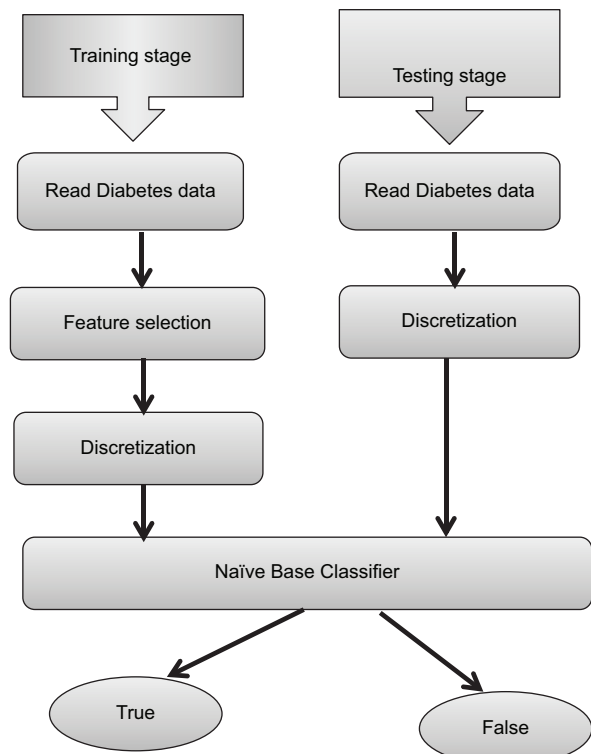


Fig. 1. The structure of the training and testing stages of the diabetes classification system.

Every patient in the database is a Pima Indian lady of at least 21 years of age who lives in or around Arizona. There are eight properties in each dataset sample (attributes), as shown in Table 1.

A sample of the data used is shown in Table 2: The row header in the table related to the column header (features) in Table 1.

There are 768 samples in all, divided into two groups. The following is the distribution of classes:

1. Normal group: (500) samples
2. Abnormal group: (268) samples.

Fold creation: The complete labeled dataset is separated into mutually exclusive folds to perform the cross validation procedure. The 768 cases in this study were chosen from the PIDD database, which means:

1. The training-set contains 499 occurrences, 325 of which are normal and 174 of which are aberrant (i.e., 65%)
2. There are 269 examples used in the testing, 175 of which are normal and 94 of which are abnormal (i.e., 35%) as shown in the Table 3.

3.3. Feature Selection Stage

The PIDD data set has eight attributes in each sample. According to the entropy of the property with class, Let S (training data) be a data-set of C outputs. For the classification issue with C classes, let P (I) indicate the proportion of S that belongs to Class I, where I is distinct from one to C [1], [23], [24].

$$p \text{ is the simple diversity index (I)} \tag{1}$$

The information-theoretical approach of assessing the quality of the split is entropy. It calculates the quantity of information in a given attribute.

$$\text{Entropy (S)} = \sum_{I=1}^C (-P(I) \log_2 P(I)) \tag{2}$$

The information gain of the example set S on the attribute A is defined as Gain(S,A).

$$\text{Gain (S, A)} = \text{Entropy (S)} - \left(\sum_{V \in \text{values of A}} \left(\frac{|SV|}{|S|} \right) \times \text{Entropy (SV)} \right) \tag{3}$$

Where SV = subset of S in which feature A has value V, |SV| = amount of data in SV, and |S| = amount of elements in S,

TABLE 1: Number of Feature of PIDD

Feature	Description	Range
1	Number of times pregnant	1-4
2	Plasma glucose concentration a 2 h in an oral glucose tolerance test.	120-140
3	Diastolic blood pressure (mm Hg).	80-90
4	Triceps skin fold thickness (mm).	12 mm (male)-23(female)
5	2-h serum insulin (IU/ml).	16-166 mlu/L
6	Body mass index (weight in kg/(height in m)^2)	18-24.9 kg/m2
7	Diabetes pedigree function Feature 8: 2-h serum insulin (IU/ml).	1-3
8	Age (years).	21 and above

TABLE 2: A sample of PIMA Indian diabetes dataset

1	2	3	4	5	6	7	8	Class
1	116	78	29	180	36.1	0.496	25	Normal
2	130	96	0	0	22.6	0.268	21	Normal
0	129	110	46	130	67.1	0.319	26	Abnormal
0	135	68	42	250	42.3	0.365	24	Abnormal
4	112	78	40	0	39.4	0.236	38	Normal

TABLE 3: Number of cases tested and trained

Training	No. Of cases	Normal	Abnormal
	499	325	174
Testing	269	175	94

and is over every value V of every conceivable value of the attribute A, a sub-set property is chosen from eight options. It takes a long time to choose more than five properties. Selecting <5 attributes, on the other hand, make the Diabetes Classification System less time-consuming but less accurate. According to the greater entropy, the ideal five attributes (1, 3, 4, 5, and 7) for describing diabetes have been determined and will be used as an input to the classification step.

3.4. Discretization

Before the categorization procedure in the Diabetes Classification System, a crucial phase must be completed. It is necessary to transform the numerical values of eight attributes to category values. This is accomplished by splitting the range of values for eight characteristics into k equal-sized bins, or equal width intervals, where k is a user-selected quantity based on the length of data. Where it can be done in a mechanism of the Equal Width Interval Discretization (EWID) algorithm, by the method used in this algorithm,


```

Algorithm (1): Equal Width Interval Discretization (EWID)

Input: Eight attributes have numerical values.
Output: Eight attributes have categorical values.
Begin
minimum1 = Op1(0) : maximum1 = Op 1(0) : minimum 2 = Op 2(0) : maximum
2 = Op 2(0) : minimum 3 = Op 3(0) : maximum 3 = Op 3(0)

minimum4 = Op 4(0) : max4 = Op 4(0) : minimum 5 = Op 5(0) : maximum 5
= Op 5(0) : minimum 6 = Op 6(0) : maximum 6 = Op 6 (0)

minimum 7 = Op 7(0) : maximum 7 = Op 7(0) : minimum 8 = Op 8(0) :
maximum 8 = Op 8(0)

For x = 1 To Len of data set {
If Op1(x) > maximum1 Then maximum1 = Op1(x) : If Op1(x) < minimum 1
Then minimum 1 = Op1(x)
If Op2(x) > maximum2 Then maximum2 = Op2(x) : If Op2(x) < minimum 2
Then minimum 2 = Op2(x)
If Op3(x) > maximum3 Then maximum3 = Op3(x) : If Op3(x) < minimum 3
Then minimum 3 = Op 3(x)
If Op4(x) > maximum4 Then maximum4 = Op4(x) : If Op4(x) < minimum 4
Then minimum 4 = Op4(x)
If Op5(x) > maximum5 Then maximum5 = Op5(x) : If Op5(x) < minimum 5
Then minimum 5 = Op5(x)
If Op6(x) > maximum6 Then maximum6 = Op6(x) : If Op6(x) < minimum 6
Then minimum 6 = Op6(x)
If Op7(x) > maximum7 Then maximum7 = Op7(x) : If Op7(x) < minimum7
Then minimum 7 = Op7(x)
If Op8(x) > maximum8 Then maximum8 = Op8(x) : If Op8(x) < minimum8
Then minimum8 = Op8(x) }

K=3 // possibilities number
G1 = Round(((maximum1 - minimum1) / k), 3) : G2 = Round(((maximum2 -
minimum2) / k), 3)
G3 = Round(((maximum3 - minimum3) / k), 3) : G4 = Round(((maximum4 -
minimum4) / k), 3)

G5 = Round(((maximum5 - minimum5) / k), 3) : G6 = Round(((maximum6 -
minimum6) / k), 3)

G7 = Round(((maximum7 - minimum7) / k), 3) : G8 = Round(((maximum8 -
minimum8) / k), 3)

// Identify k ranges for each of the eight attributes.

LowOp1 = (minimum1 + G1) : medOp1 = (minimum 1 + (2 * G1)) : highOp1
= (minimum1 + (3 * G1))
LowOp2 = (minimum2 + G2) : medOp2 = (minimum2 + (2 * G2)) : highOp2
= (minimum2 + (3 * G2))

```

```

LowOp3 = (minimum3 + G3) : medOp3 = (minimum3 + (2 * G3)) : highOp3
= (minimum3 + (3 * G3))
LowOp4 = (minimum4 + G4) : medOp4 = (minimum4 + (2 * G4)) : highOp4
= (minimum4 + (3 * G4))
LowOp5 = (minimum5 + G5) : medOp5 = (minimum5 + (2 * G5)) : highOp5
= (minimum5 + (3 * G5))
LowOp6 = (minimum6 + G6) : medOp6 = (minimum6 + (2 * G6)) : highOp6
= (minimum6 + (3 * G6))
LowOp7 = (minimum7 + G7) : medOp7 = (minimum7 + (2 * G7)) : highOp7
= (minimum7 + (3 * G7))
LowOp8 = (minimum8 + G8) : medOp8 = (minimum8 + (2 * G8)) : highOp8
= (minimum8 + (3 * G8))

End

```

Fig. 2. The structure for the EWID algorithm code.

we have obtained the traits under a specified range. It is an important step before classifying to convert it into categories, to be used for training. [25], [26] [27], which consists of steps as shown in Fig. 2.

3.5. Classifier Model

The most well-known task classification is the constructing classifier model. This structure is used to determine the diabetes class, which can be either normal or abnormal. The Diabetes (Training-Set) database is made up of attribute-value representations for a large number of patients, with five categorical characteristics (1, 3, 4, 5, and 7) and class attributes. Those characteristics are fed into the learning classifier model. The classifier model is used to predict the new case. The decision is used to build the classifier model in this work, which is based on training diabetes [1], [25]. This mechanism is implemented using a Naive base algorithm, where the algorithm is fed by the characteristics from the training set, and this helps in building a classifier model for prediction based on the decision, the steps of this algorithm and the mechanism used in it are explained as shown in Fig. 3.

4. RESULTS

As previously stated, the diabetes categorization system's initial data included Pima Indian diabetes illness measures. Before classification, the retrieved numerical values from attributes must be transformed into categorical values, which will be used to train the classifier using the EWID method.

Table 4 shows the categorical values of the five attributes according to the Diabetes Classification System. Attributes, attribute-value, and range of values are the three fields that

Algorithm (2): Naïve Base Algorithm

Output: decision

Begin

For i = 0 To 1 // two group

Q = 0

For j = 0 To LenTr

If Opr(i) = Op(j) Then Q = Q + 1

End for

Opr (i) = Math.Round((Q / CT), 3)

end for

For G = 1 To 8

If G = 1 Then a = b1 : If G = 2 Then a = b2 : If G = 3

Then a = b3

If G = 4 Then a = b4 : If G = 5 Then a = b5 : If G = 6

Then a = b6

If G = 7 Then a = b7 : If G = 8 Then a = b8

For i = 0 To a

For k = 0 To 1

Q = 0 : Y = 0

For j = 0 To LenTr

If G = 1 Then If P1(i) = F1(j) And Opr (k) = Op(j) Then Q = Q + 1

If G = 2 Then If P2(i) = F2(j) And Opr(k) = Op(j) Then Q = Q + 1

If G = 3 Then If P3(i) = F3(j) And Opr(k) = Op(j) Then Q = Q + 1

If G = 4 Then If P4(i) = F4(j) And Opr(k) = Op(j) Then Q = Q + 1

If G = 5 Then If P5(i) = F5(j) And Opr(k) = Op(j) Then Q = Q + 1

If G = 6 Then If P6(i) = F6(j) And Opr(k) = Op(j) Then Q = Q + 1

If G = 7 Then If P7(i) = F7(j) And Opr(k) = Op(j) Then Q = Q + 1

If G = 8 Then If P8(i) = F8(j) And Opr(k) = Op(j) Then Q = Q + 1

If G = 9 Then If P9(i) = F9(j) And Opr(k) = Op(j) Then Q = Q + 1

If G = 10 Then If P 10(i) = F 10(j) And Opr(k) = fc(j) Then

Q = Q + 1

Next

// likelihood

If G = 1 Then pp1(i, k) = Math.Round((Q / (QP(k) *CT)), 3)

If G = 2 Then pp2(i, k) = Math.Round((Q / (QP(k) *CT)), 3)

If G = 3 Then pp3(i, k) = Math.Round((Q / (QP(k) *CT)), 3)

If G = 4 Then pp4(i, k) = Math.Round((Q / (QP(k) *CT)), 3)

If G = 5 Then pp5(i, k) = Math.Round((Q / (QP(k) *CT)), 3)

If G = 6 Then pp6(i, k) = Math.Round((Q / (QP(k) *CT)), 3)

If G = 7 Then pp7(i, k) = Math.Round((Q / (QP(k) *CT)), 3)

If G = 8 Then pp8(i, k) = Math.Round((Q / (QP(k) *CT)), 3)

End for

End for

End for

End

Fig. 3. Naïve Base Algorithm.

make up this table. The five attributes are shown in the first field (How many times pregnant, Diastolic blood pressure, Triceps skin fold thickness, serum Insulin and diabetes pedigree function Feature). The categorical values of five attributes are shown in the second field. The range of values obtained by the EWID method is represented in the third field.

Table 5 shows the categorical values samples that were obtained by changing numerical property values of the diabetes case with the class attribute according to the range of value field in Table 4.

Table 6 shows the results of the top five attributes chosen for use in the Classifier model. The following properties were used to train and evaluate the classifier model: Preg signifies the number of pregnancies a woman has had, Pres the Diastolic blood pressure, Skin the thickness of the Triceps skin folds, Insu the serum insulin, and Pedi the Diabetes pedigree function feature. The entropy for each feature is calculated using the entropy equation (2).

For the diabetes instances indicated in the tables, the following naïve Bayés classifier was trained: The likelihood of the number of times pregnant features for the three ranges of two classes (normal and abnormal) is represented in Table 7.

The normal class is represented by the number 0.322, while the abnormal class of the low range is represented by the value 0.551. The abnormal class of medium-range is represented by the value 0.301. The number 0.241 represents the typical normal class, whereas the value 0.148 represents the aberrant high-range class. The likelihood of each class being equal to one is shown by the end row.

The likelihood of the number of skin characteristics for the three ranges of two classes (normal and abnormal) is represented in Table 8. The number 0.35 represents the normal class, whereas the value 0.431 represents the low-range abnormal class. The values 0.276 and 0.179 represent the normal and abnormal classes in the medium range, respectively, while the values 0.403 and 0.39 represent the normal and abnormal classes in the high range, respectively. The likelihood of each class being equal to one is shown by the end row.

The likelihood of the number of ins features for the three ranges of two classes is represented in Table 9. (Normal, abnormal) The values 0.977 and 0.865 reflect

TABLE 4: Categorical features

Attributes	Attribute-value	Range of Values
Number of times pregnant	Low	(0–2)
	Medium	(3–5)
	High	(6–17)
Diastolic blood pressure	Low	(0–80)
	Medium	(80–100)
	High	(100–122)
Triceps Skin fold thickness	Low	(0–20)
	Medium	(20–60)
	High	(60–99)
Serum insulin	Normal	(0–280)
	Abnormal	(280–860)
Diabetes pedigree	Low	(0.084–1.251)
	High	(1.251–2.42)

TABLE 5: Samples of categorical features values

Id	Preg	Press	Skin	Insu	Pedi	Class
1	Low	Low	Low	Normal	Low	Normal
2	Low	Medium	Medium	Normal	High	Normal
3	High	Low	Medium	Normal	Low	Normal
4	High	Medium	Medium	Abnormal	High	Abnormal
5	Low	High	Medium	Normal	High	Abnormal
6	Medium	High	High	Abnormal	High	Abnormal

TABLE 6: Entropy of categorical features values

Entropy	Preg	Plas	Skin	Skin	Ins	Mass	Pedi	Age
	0.71	0.13	0.49	0.88	0.55	0.02	0.64	0.08

the normal and abnormal classes, respectively, of the normal –range. The normal class is represented by the number 0.023, whereas the abnormal class is represented by the value 0.006.

The likelihood of the number of pedi characteristics for the three ranges of two classes (abnormal and normal) is shown by Table 10, where the value 0.874 represents the normal class and the value 0.935 represents the abnormal class of the low –range. The number 0.126 represents the normal class, whereas the value 0.065 represents the high-range abnormal class. The probability of each class that is equal to one is summed in the last row.

Table 11 provides the confusion matrix of classifier implementation retrieved from the testing stage using NB.

The accuracy and error rate for diagnosed cases are calculated using Table 11. The values used in calculating the accuracy of the NB classifier using the accuracy equation (4),

Where the:

TP: true positive

TN: true negative

TP and TN are added together, then divided by the sum of all with FP (false positive) and FN (false negative), and the error is computed using the equation (5) [28].

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

$$Accuracy = (173 + 87) / (173 + 4 + 5 + 87)$$

$$= 0.96 \tag{4}$$

$$Error = (FP + FN) / (TP + TN + FP + FN)$$

$$= 0.334 \tag{5}$$

TABLE 7: Preg Feature probability

Preg	Normal	Abnormal
Low	0.322	0.551
Medium	0.437	0.391
High	0.241	0.148
SUM	1	1

TABLE 8: Skin feature probability

Preg	Normal	Abnormal
Low	0.35	0.431
Medium	0.276	0.179
High	0.403	0.39
SUM	1	1

TABLE 9: Ins Feature probability

Preg	Normal	Abnormal
Normal	0.977	0.865
Abnormal	0.023	0.006
SUM	1	1

TABLE 10: Pedi feature probability

Preg	Normal	Abnormal
Low	0.874	0.935
High	0.126	0.065
SUM	1	1

TABLE 11: The confusion matrix using Naïve Bayes classifier

Predicate Class	Actual Class	
	Normal	Abnormal
Normal	(TP) 173	(FP) 4
Abnormal	(FN) 5	(TN) 87

TABLE 12: Comparison of accuracy with previous works

Authors	The year	dataset	The method used	Accuracy
Sneha N. and Gangil T	2019	Pima Indians Diabetes Dataset	1) Support vector machine SVM 2) Naïve Base NB	77% 82.30%
Islam M. A. and Jahan N.	2017	Pima Indians Diabetes Dataset	Logistic Regression algorithm	80%
Rawat V and Suryakant S.	2019	Pima Indians Diabetes Dataset	Bagging method	81.77%
Pradhana N , Rania G, Singh V, Dhaka V. S and Pooniab R. C.	2020	Pima Indian Diabetes dataset	Artificial neural networks	85.09%
Prasanth S, Banujan K and Btgs K	2021	Pima Indians Diabetes Dataset	The adaptation model of SVM, CatBoost, and Random Forest (RF)	86.15%.
Zolfaghar	2012	Pima Indians Diabetes Dataset	the support vector machine (SVM) and back-propagation neural network (BP-NN)	88%
Resti Y., Kresnawati E. S., Dewi N. R., Zayanti D. A. and Eliyati N.	2021	Pima Indians Diabetes Dataset	Naive Bayes, Discriminant Analysis, and Logistic Regression	93%
Sanakal R. and Jayakumari S. T.,	2014	Pima Indians Diabetes Dataset	Fuzzy C Means Clustering	94.3&
Jayanthi N. , Babu .V. B. and Rao S.	2016	Pima Indians Diabetes Dataset	FCM and SVM	94.3%
Hameed S. A. and Baban A. B. (The proposed method in this paper)	2021	Pima Indians Diabetes Dataset	Naive Base classifier	96%

These values are compensated according to the work performed and the results obtained.

This proposed work was used to enhance accuracy using the Naive Base classifier method, where the mechanism was implemented with a performance that gives higher accuracy than the accuracy obtained from previous studies. This work has been compared with the previous works as shown in Table 12, as each work used the appropriate mechanism for diagnosing and classifying this disease and obtained an appropriate accuracy rate for the work. After studying and analyzing this problem, a high percentage of accuracy was obtained.

5. CONCLUSION

In this study, the NB classifier was used, we attempted to provide an approach for identifying the classification method for detecting and classifying diabetes at an early stage. There are eight properties in each dataset sample (attributes), divided into two classes normal and abnormal, and used in two stages (training and testing), during the training stage, specific functions were performed, such as reading the diabetes dataset, feature selection, discretization, and the classifier model used to create the decision rules, and during the testing stage, specific functions were performed, such as reading the diabetes dataset, discretization, decision rules, and output. The findings of the experiments were run on the dataset and compared with the previous works, and a system that can

reliably diagnose and categorize gestational diabetes was shown to be 96% accurate.

6. REFERENCES

- [1] S. Gupta, H. K. Verma and D. Bhardwaj. "Classification of diabetes using naïve Bayes and support vector machine as a technique". *Operations Management and Systems Engineering*, pp. 365-376, 2020.
- [2] S. Prasanth, K. Banujan and K. Btgs. "Hyper Parameter Tuned Ensemble Approach for Gestational Diabetes Prediction". International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), IEEE. pp. 18-23, 2021.
- [3] N. Sneha and T. Gangil. "Analysis of diabetes mellitus for early prediction using optimal features selection". *Journal of Big Data*, vol. 6, p. 13, 2019.
- [4] A. Sarwar, M. Ali, J. Manhas and V. Sharma. "Diagnosis of diabetes Type-II using hybrid machine learning based ensemble model". *International Journal of Information Technology*, vol. 12, pp. 419-428, 2020.
- [5] N. Pradhana, G. Rania, V. Singh, V. S. Dhaka and R. C. Pooniab. "Diabetes prediction using artificial neural network". In: *Deep Learning Techniques for Biomedical and Health Informatics*, ScienceDirect, pp. 327-339, 2020.
- [6] M. D. Okpor. "Prognostic diagnosis of gestational diabetes utilizing fuzzy classifier". *International Journal of Computer Science and Network Security*, vol. 15, no. 6, pp. 44-48, 2015.
- [7] E. G. Filho, P. R. Pinheiro, M. C. D. Pinheiro, L. C. Nunes and L. B. G. Gom. "Heterogeneous methodology to support the early diagnosis of gestational diabetes". *IEEE Access*, vol. 99, p. 1, 2019.
- [8] M. Marozas, S. Sosunkevič, M. Francaitė-Daugėlienė, D. Veličkienė and A. Lukoševičius. "Algorithm for diabetes risk evaluation from past gestational diabetes data". *Technology and Health Care*, vol. 26, no. 4, pp. 637-648, 2018.
- [9] Y. Resti, E. S. Kresnawati, N. R. Dewi, D. A. Zayanti and N. Eliyati. "Diagnosis of diabetes mellitus in women of reproductive age using the prediction methods of naive bayes, discriminant analysis, and logistic regression". *Science and Technology Indonesia*, vol. 6, no. 2, pp. 96-104, 2021.
- [10] M.A. Islam and N. Jahan. "Prediction of Onset Diabetes using Machine Learning Techniques". *International Journal of Computer Applications*, vol. 180, no. 5, pp. 7-11, 2017.
- [11] R. Saxena, S. K. Sharma and M. Gupta. "Analysis of machine learning algorithms in diabetes mellitus prediction". *Journal of Physics: Conference Series*, vol. 1921, p. 012073, 2021.
- [12] N. Jayanthi, V. B. Babu and S. Rao. "Data mining techniques for CPD of diabetes". *International Journal of Engineering Computational Research and Technology*, 2014.
- [13] K. Lakhwani, S. Bhargava, K. K. Hiran, M. M. Bundeale and D. Somwanshi. "Prediction of the Onset of Diabetes Using Artificial Neural Network and Pima Indians Diabetes Dataset". 5th IEEE International Conference on Recent Advances and Innovations in Engineering, pp. 1-6, 2020.
- [14] R. Zolfaghar. "Diagnosis of diabetes in female population of pima indian heritage with ensemble of BP neural network and SVM". *International Journal of Computational Engineering and Management*, vol. 15, no. 4, pp. 115-121, 2012.
- [15] A. Kaushik, A. Sehgal, S. Vora, V. Palan and S. Patil. "Presaging The Signs Of Diabetes Using Machine Learning Algorithms". 12th International Conference on Computing Communication and Networking Technologies, 2021.
- [16] R. Sanakal and S. T. Jayakumari. "Prognosis of diabetes using data mining approach-Fuzzy C Means clustering and support vector machine". *International Journal of Computer Trends and Technology*, vol. 11, no. 2, pp. 94-98, 2014.
- [17] H. Naz and S. Ahuja. "Deep learning approach for diabetes prediction using PIMA Indian dataset". *Journal of Diabetes and Metabolic Disorders*, vol. 19, no. 1, pp. 391-403, 2020.
- [18] V. Rawat and S. Suryakant. "A classification system for diabetic patients with machine learning techniques". *International Journal of Mathematical, Engineering and Management Sciences*, vol. 4, no. 3, pp. 729-744, 2019.
- [19] P. Kaur and R. Kaur. "Comparative analysis of classification techniques for diagnosis of diabetes". In: Jain, L., Virvou, M., Piuri, V. and Balas, V. (eds.), *Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals Advances in Intelligent Systems and Computing*. Vol. 1064. Springer, Singapore, 2020.
- [20] L. Jonk. "Chronic Disease Prevention a Vital Investment". World Health Organization, Geneva, Switzerland, 2005.
- [21] L. Moon. "Prevention of Cardiovascular Disease, Diabetes and Chronic Kidney Disease: Targeting Risk Factors". Vol. 118. AIHW, 2009. Available from: <http://www.aihw.gov.au/publications/index.cfm>. [Last accessed on 2022 Mar 09].
- [22] A. Rajivkannan and K. S. Aparna. "A survey on diabetes prediction using machine learning techniques". *International Journal of Research in Engineering, Science and Management*, vol. 4, no. 11, pp. 51-54, 2021.
- [23] D. Lavanya and K. U. Rani. "Performance evaluation of decision tree classifiers on medical datasets". *International Journal of Computer Applications*, vol. 26, no. 4, pp. 1-4, 2011.
- [24] R. Raja, I. Mukherjee and B. K. Sarkar. "A machine learning-based prediction model for preterm birth in Rural India". *Journal of Healthcare Engineering*, vol. 2021, p. 6665573, 2021.
- [25] A. Saleha and F. Nasari. "Implementation of equal-width interval discretization in naive bayes method for increasing accuracy of students' majors prediction". *Lontar Komputer Jurnal Ilmiah Teknologi Informasi*. Vol. 9, no. 2, pp. 104-113, 2018.
- [26] R. Dash, R. L. Paramguru and R. Dash. "Comparative analysis of supervised and unsupervised discretization techniques". *International Journal of Advances in Science and Technology*, vol. 2, no. 3, pp. 29-37, 2011.
- [27] J. Dougherty, R. Kohavi and M. Sahami. "Supervised and Unsupervised Discretization of Continuous Features". In: Proceedings of the Twelfth International Conference on International Conference on Machine Learning (ICML'95). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1995, pp. 194-202.
- [28] J. Han and M. Kambar. "Data Mining: Concepts and Techniques". 2nd ed. Morgan Kaufmann Publisher, Burlington, Massachusetts, 2006.

Eye Tracking Technique for Controlling Computer Game Objects

Tara Qadir Kaka Muhammad¹, Hawar Othman Sharif¹, Mazen Ismaeel Ghareb²

¹Department of Computer, College of Science, University of Sulaimani, Sulaimani, Iraq, ²Department of Computer Science, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq



ABSTRACT

The study explored the employment of associate in accessible eye tracer with keyboard and mouse input devices for video games. An interactive game has been developed using unity with multiple balls objects and by hitting they could collect more point for each player. It has been used different techniques to hit the balls using mouse, keyboard, and mixed. Eye tracker input has been help to increase the performance of collected the player points. The research explains how the eye tacking techniques can be used in widely in video game and it is very interactive. Finally, we examine the use of visual observation in relevancy the keyboard and mouse input control and show the difference. Our results indicate that the employment of a watch huntsman will increase the immersion of a computer game and considerably improve the video game technology.

Index Terms: Computer Games, Eye Tracking, Eye Gaze Interaction, Facial Expressions as Game Input, Evaluating Peripheral Interaction

1. INTRODUCTION

It has been described two experiments that examine our technique of selecting eye objects using traditional mouse selection. We have already looked at how people behave when interacting with their eyes in the demos. The next step is to show that our method can withstand tougher use and that people like to select objects using their gaze for an extended period.

We rated the overall performance of gaze interaction with that of a widely used and widely used device: The mouse. The interaction of the eye requires different hardware and software. It's a question of whether it's worth it. If it works properly, we could also get some secondary benefits that

are difficult to quantify through an additional, passive, or mild input channel. For example, we've found that when the visual interaction works well, the device senses almost as if it were waiting for user controls. Just like you are studying the mind of the user. You want no more guidance input and you have your palms free from various tasks. It slows down the interaction and can "cover costs" in a simple experimental comparison with the mouse, regardless of the immaturity of current eye tracking technology.

The eye interaction approach is faster, but we see it as an advantage; however, it is not now the primary motivation for using eye monitoring in most environments.

Our experiments have measured the time required to perform simple and representative direct manipulative arithmetic tasks. One asked to select a highlighted circle from a circle grid. The second asked the test person to select the named letter on a loudspeaker from a letter grid. Our results show a wonderful and measurable pacemaker benefit for searching the mouse in the same experimental setting, persevering in every

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp43-51

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Copyright © 2022 Muhammad, *et al.* This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Tara Qadir Kaka Muhammad, Department of Computer, College of Science, University of Sulaimani, Sulaimani, Iraq. E-mail: tara.qadir@univsul.edu.iq

Received: 01-09-2021

Accepted: 30-03-2022

Published: 20-04-2022

experiment. The key points of the test allow us to understand how our method of visible interaction works and why it is effective. As expected, the method is a little faster than the mouse. Our search suggests that the eye can also go faster than the hand. Our method verification is how our entire interaction method and algorithm preserve this eye speed advantage on a proper object. We study the physiology of the eye and use these records to extract useful facts about the user's overall intentions from noisy and fearful eye movement data. Even if it does, this algorithm is primarily based on understanding eye movement. It used to be no longer clear that our eye interaction approach would keep pace with the eye, as the eye monitoring hardware entails additional latencies [1]. The overall performance of any interaction science results from its software and hardware program. The previous experiments show that keyboard, mouse and joystick considered as traditional game inputs and there are many new techniques need to be considers [2], [3], [4].

Manipulate the voice [5] and monitor the head [6]. For the past few years, while using various entry level techniques, the goal of researchers has been to find out which technique is most accurate, most immersive, and most convenient for users. Here, we explain the benefits of using eye movements as a game controller. The eye tracking techniques have been compared to mouse input and have showed the results. Eye tracking technology has been shown to increase immersion and make games more fun for the player. As in Ivanchenko *et al.* [7], Almansouri [8], it shows that eye tracking as game input is very precise. The comparison of mouse, keyboard, and appearance in Jiménez-Rodríguez *et al.* [9] is also used as an entry level solo controller.

In contrast to these studies, which deal specifically with the comparison between mouse and eye control in terms of precision and effectiveness, many studies have focused on investigating the game experiment. In article [10], you focused on this immersive evaluation and the user experience. However, their research showed that when comparing the game with mouse data, the players were more immersed in the game, but in paper [11], they achieved a reliable questionnaire evaluation and a high score in terms of stamina. Feelings of fluidity and immersion in gaze-controlled play compared to the study by Modi and Singh [10], the result of which had to be further investigated. The way people treat the computer as human-computer interaction (HCI) has user actions on three different levels: Physical, cognitive, and emotional; however, the emotional level is a new topic that not only tries to make the interaction experience pleasant but also affects the further use of the machine by the user [12].

However, to better understand the emotional level at HCI, that is, the user's involvement in the machine's use, an evaluation of the user is required for the use of a peripheral device with test emotions. Emotional use of a peripheral device at the same time is difficult. Research [13] examined this primary function that requires continuous interaction and secondary work that takes place on the periphery. Some research on HCI has used emotions as a starting point. In research [14], Bernhaupt *et al.* designed an emotion flower set and used facial emotions as input. They used positive emotions (joy and surprise) to grow flowers and negative emotions (disgust, anger, sadness, and fear) to slow growth. Their game was intended for the workplace and they understood that their game improved the player's emotional state while playing, while the game did not affect people's general mood, but their work has become a fundamental work for Lankes *et al.* group [15]. His research works alongside the redesign of the emotion flower game; they tried wearing it in a mall and examined the player's emotional feedback to add more contrast to their basic work. Our approach is to combine existing input techniques such as (mouse, keyboard, and eye tracking). Add more facial emotions (joy, anger, and surprise) than inputs to a main game used in the proposed game [16]. The user can choose one. The game is using all three types of inputs such as (mouse, keyboard and eye tracker). Later, the damaged balloons are collected, and the score is increased. However, the emotion of the face has a peripheral role that helps the user control the speed of the balloons and gets more points in a time [17].

In addition, evaluation and effectiveness are important to us; measure effectiveness by comparing recorded results from different users. However, the assessment is made by comparing the input parameters in two different categories (emotional and unemotional).

2. RELATED WORK

2.1. Review Stage

Since there is extensive research on evaluating emotions and emotions with the help of users, research on emotions is not limited to facial expression, as the ability to become aware of people's emotions has an impact on social interaction and human behavior [18], moreover, some researchers worked on the body as in Ghareb [19] argue that recognition of emotions through non-verbal communication can be achieved by sensing body expression. However, our research on facial emotions touches a small portion of this area.

Ekman in Chittaro and Sioni [20] defined that facial expression is an example of things one can do through the face, his dialogue focused on the set of facial expressions – happiness, surprise, anger, sadness, worry, and disgust – that are culturally international and on which cultures depend to decide, to show rules. As in Emotion Recognition and its Application in Software Engineering [21], it examines a number of scenarios to assess the possibility of applying emotional cognition strategies in four areas: Software programming, website personalization, school, and games. Video games are contingencies that can dynamically respond to the emotions of the diagnosed contemporary gamer. Massive investigations worked on a specific reenactment for capturing emotions as discussed in Ekman *et al.* [22] an open-source EVG (Emotion Evoking Game) and a first and formative comparison strange end result ordinary variations were determined by comparison and facial expressions of surprise, joy, and disappointment been. There is a lot of research on recovery that has been used in focusing emotions in deferent approaches [18], [23], [24]. In addition to the emotion, the evaluation of peripheral units is a trend phase that has led to extensive discussions, countless findings focused on a unique type of peripheral evaluation. Some of the systems in the literature were evaluated on the basis of studies of test subjects [25], [26]. In addition, a learnable subject [27] using unique modalities (tangible, tactile, and hands-free); however, for the peripheral interaction, all peripheral works are ultimately a simple interaction, unless the focus is on controlling the audio participant. Collecting the points using user friendly method and easy way for evaluation different technology in the game. Besides working on the contrast of consumer emotions at a certain point in the game, our device focuses on evaluating the input devices (keyboard, mouse, and gaze); some research has observed the usefulness of the gaze as an access system [28], [29], [30], [31], [32], [33]. The rating of the normal mouse with momentary was examined in Almansouri [8], among the recently introduced (adult, middle aged, and the elderly) the rating of the eye over size for middle aged and the elderly; while Jacob [34], the evaluation was equated and the operability, in contrast to the menu resolution selection technique of a developed web browser, was experienced again as an operator together with the component, because it results in the system's operability. In Roose and Veinott [3], Ivanchenko *et al.* [7], Sibert and Jacob [35], Murata *et al.* [36], as a tutorial on how the view can be combined with different input methods. However [9], he worked on using the gaze as a solo input and studied the general performance variations for gaze, mouse, or keyboard for a similar project in the game. Today, the most common

structure of the eye tracker is the “corneal reflection” unit of the laptop. These structures place the surroundings of the user's gaze as a display screen coordinate on a monitor. To decide where the consumer is looking. These structures sing to one or each of the eyes they look at with a digital camera equipped with an infrared (IR) filter. The previous game have configured the input of eye using eye cornea and most be configured near the eye using new camera. Because the user's corneal floor is roughly spherical, the corneal reflex area remains constant as the user's eyes move relative to the head, the role of the student in relation to this reflex creates the position of the wearer's eye. A calibration sequence is used to map eye movements to display screen coordinates. There are also portable structures that are beneficial for ubiquitous computing scenarios. These structures use the same method as computer systems, but the archive view as a coordinate in a digital camera built into the user's head [37]. There are several games has been research of have performance issues of using keyboard and mouse as an input. By testing eye input versus mouse input in three extraordinary PC games, Smith and Graham [40] concluded that using eye monitoring can guarantee the player an additional immersive journey using the eye tracker in first-person shooters. Each test participant was asked to play the same sport using three specific input techniques: (1) Mouse, keyboard, and eye tracker; (2) mouse and keyboard only; or (3) a console gamepad. The results are not exactly encouraging now, suggesting that the overall performance with the eye tracker was once well below the two different ones. However, Isokoski and Martin attributed these results to the players' greater experience and expertise. The concern about the players who using typical input methods and when offering alternating input for playing the game it need further training. Other studies came to comparable results. The authors in research [39] created a simple look in which the participant was once asked to remove 25 balloons that were moving across the screen at extraordinary speeds. The participant would move the mouse or the eye tracker over the pointer and remove the balloons with the help of a mouse click. Two prerequisites were tested: With and besides the time limit for completing the task. The results confirmed that besides the time limit, the accuracy and time to complete the task were earlier worse when using the eye tracker than when using a mouse [40]. Performance was once based solely on the percentage of balls that the player wiped out. Michael The paper [43] ended with clearly contrary results. Players also mentioned that the eye tracker used to be exceptionally fun to use. These contradicting research consequences suggest that exercising and the approach to enhancing fair recreation are key factors in achieving a continued satisfactory outcome [44], [45].

2.2. Designing a Proposed Game

The comparison of peripheral interplay requires at least two tasks: A foremost task, which must be the focal point of the participant's attention, and a secondary task, which needs to be carried out in the periphery. This secondary project is normally a given: The assignment supported using the peripheral gadget being evaluated.

2.2.1. Designing the primary task

The tasks are to tell the player to play with keyboard first, then focused on playing with keyboard and mouse if possible, then calculating the results and timing from the players. How many balls have been crashed and how much point has been collected with timing for each techniques. Later on, it has been concentrated on eye tracking for each ball movement and how can control the speed of the ball to hit and crash as much balls as possible. Most of the players have been collected full points but with more time is needed it.

3. METHODOLOGY

The methodology of this study is to observe the players interaction with the games before and after using eye tracking as an input methods for the game. The players all have experience in gaming which they have training for using this game and how can used eye tracker device. The eye tracking techniques help the players to be more interaction with the game. The observation has been conducted and extracted the data from the player before and after using the eye tracker and illustrates the difference between the results using several statistical measurements.

3.1. Game Scenario

The game is designed and implemented to destroy difference color balloons to collect points. The balloons have been destroyed by three means of inputs, mouse, keyboard, and eye. The main objective of the game is finding the difference efficiency of several inputs such as mouse, keyboard, and eye focused.

3.1.1. Hardware requirements

The game has developed on PC with these requirements, CPU Intel i7-2600, 3.6 GHZ, Ram 4GB, Hard desk 512 GB, and graphic card NVID QUARD 600. Tobii Eye Tracker has been used in this game. The only devices are capable of tracking both head and eye movements for game interaction, exports training, and streaming.

The Tobii 4C eye tracker is the hardware device which will track eye movement in the game. It has a driver define it and

then define the eye of the player and later will detect the eye movement in the game and program it in game source code.

Fig. 1 shows the Tobii 4C eye tracker.

3.1.2. Game design

The figures below have been explained the game interface and game rules how player can use the game and what is the interactive with players. Final figure explain how the interaction between eye tracker with ball movement to score point for the player.

The game has been developed using Unity version 18.3, C# ultimate 2012 and using Microsoft Windows 10 64 bit. Fig. 1 shows the game interface and how can user hit the balloons and collects points. Fig. 2 explains users using mouse and Fig. 3 shows how user can use eye movement to hit the balloons and score point.



Fig. 1. Tobii 4C eye tracker device.

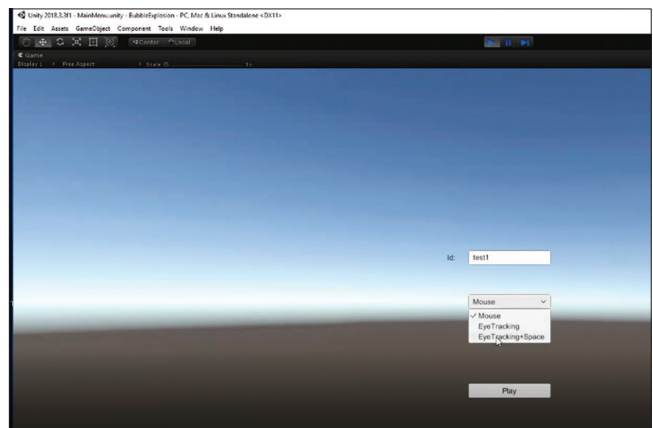


Fig. 2. Users select input option for playing the game.

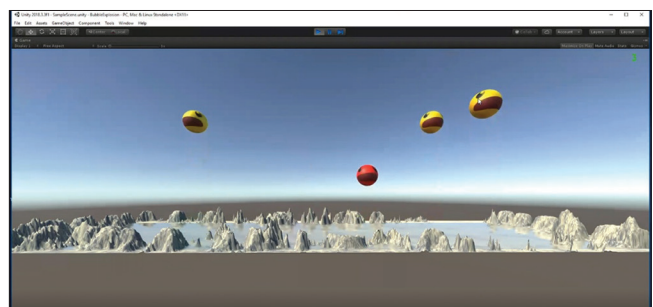


Fig. 3. Mouse input for users.

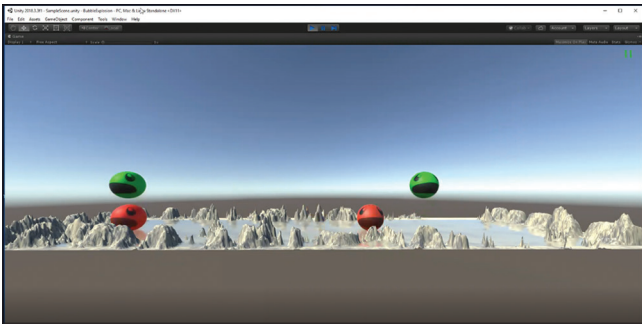


Fig. 4. Eye input for users.

4. DATA COLLECTION

Collecting data about the player have been conducted selecting 48 users. The player was undergraduate students from computer science department and all have been trained to use the game. Each user has been played with mouse, eye, and eye with space for collecting the points. Table 1 shows the time collected using mouse, eye, or combination of the two.

The results have been showed that 45% eye input performance is better than mouse input and the performance of eye input is 66% which is better performance than combination input. This indicates that the eye input has acceptable values as input for the players.

The linear Pearson's correlation has been used for this study. We have been used Pearson correlation coefficient (PCC) to measure a liner correlation between eye control dataset and mouse and other two correlations between eye and eye input control. The measure calculates the covariance of two variables and the product of standard deviations. The results are between -1 and 1 . The covariance only can reflect the relationships or correlations.

Pearson correlation coefficient = $\rho(x,y) = \frac{\sum[(x_i - \bar{x}) * (y_i - \bar{y})]}{(\sigma_x * \sigma_y)}$ [46]

Table 2 has shown that the PCC [46] between mouse and eye input is 0.44 which indicates that the timing for eye input is related to mouse and has acceptable values as input.

Table 3 has shown that the PCC [46] between eye and eye space is 0.54 which indicates that the timing for eye input is related to mouse and has acceptable values as input.

Table 4 shows the regression statistics of difference between eye and mouse tracker and has been generated using SPSS statistical tool. All the results indicate significant results for

user timing compare to mouse input. These statistical results indicate that eye input has slightly better performance than mouse. This means that game industry can use eye interaction techniques beside mouse input. P values show significant values of eye movement regarding effect of the eye input to the user interaction in the game with mouse input also.

Table 5 shows the regression statistics of difference between eye tracker input and keyboard input; these results have been extracted from SPSS statistical tool. All the results indicate that eyes tracker input slightly has less performance for user timing compare to mouse input. These statistical results indicate that eye tracker input has slightly slower performance than keyboard. This means that game industry can use eye interaction techniques beside mouse input.

Descriptive statistics for the three games input for mean, standards error, median, standard deviation, sample variance, and confidence level are explained in Table 6. Fig. 4 has been explain the using of eye tracking as it shows the ball speed has been changes according of eye focus. The results explain better results for eye tracker in some of statically factors. These results indicate that users can use eye tracking and eye tracking with keyboard combination and make it the game more interactive and better performance.

Figs. 5-7 have shown the histogram for speed performance of three different method of game input eye, mouse, and eye keyboard.

5. RESULTS

An evaluation of Pearson correlation [46] and regression records used to be carried out on the overall performance measures for every sport to realize any big variations between the two entry modalities. Users carried out visually well. However, no sizable overall performance variations had been found for each mouse and keyboard. For pointing tasks, for example, the consumer will frequently appear at the goal and then go the cursor solely when he picks a target. However, with the eye pointer, the cursor strikes every time the consumer strikes their eyes. These effects in a widespread expand in the quantity of remarks the person receives from the game, even if the person does now not consciously raise out an express action. Users additionally confirmed a robust choice for the eyepiece tracker throughout playback. We suppose this is because of the decreased quantity of effort it takes to go through the persona when the use of the eyes. To entire the task, customers had to make over one cursor

TABLE 1: Player time difference for different inputs

No. of players	Mouse	Eye	Combination of two
1	46.29168	50.68896	72.26347
2	49.36595	45.9901	47.98271
3	47.02565	54.16142	74.47321
4	35.69894	96.90728	64.29069
5	36.75071	45.54176	39.36242
6	40.79372	68.80875	63.11514
7	47.87828	43.61938	55.02164
8	62.88597	52.35537	52.68043
9	65.5258	70.47275	56.20034
10	42.11875	44.93463	43.45302
11	63.0501	77.57199	90.01323
12	42.8105	24.0855	41.06178
13	39.97028	72.89788	71.18179
14	64.43954	70.54054	52.99361
15	80.71351	91.87479	83.32175
16	36.42005	23.09589	46.32096
17	47.7062	64.47396	60.83725
18	66.36465	61.36431	44.36486
19	69.42387	72.35069	81.43118
20	49.9991	70.50659	78.45703
21	44.26379	51.86754	60.57484
22	57.42333	24.09982	67.70646
23	48.06614	38.37986	46.19806
24	76.02664	64.09002	102.3053
25	40.44497	40.8985	38.73352
26	81.27607	79.20071	38.13029
27	32.56762	24.599	27.99892
28	60.89792	43.21954	81.91699
29	66.81228	53.31702	69.39323
30	68.82234	31.56778	62.89731
31	80.09913	62.05175	74.31324
32	48.12384	81.68127	84.82584
33	30.01091	58.38366	60.69171
34	56.67857	54.97182	62.73943
35	63.83644	77.20864	98.52775
36	71.37659	63.88522	73.57787
37	32.53166	34.77013	40.04244
38	70.75024	47.13963	51.47002
39	84.64526	61.72007	73.09223
40	47.7101	70.00181	76.68742
41	31.31399	39.43198	34.08674
42	35.38079	36.04421	48.4115
43	47.42	53.64031	40.08095
44	38.87392	46.62032	55.43325
45	53.62331	53.33913	58.65588
46	133.2718	85.60275	62.73277
47	48.19358	67.05249	62.56892
48	64.12367	58.64465	68.15075

TABLE 2: Pearson correlation between mouse and eye input

	Mouse input	Eye input
Mouse input	1	0.44
Eye input	0.44	1

TABLE 3: Pearson correlation between eye and eye space input

	Eye	Eye space
Eye	1	0.54
Eye space	0.54	1

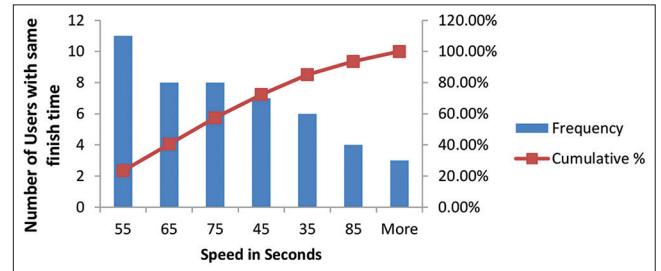


Fig. 5. Players eye game input performance.

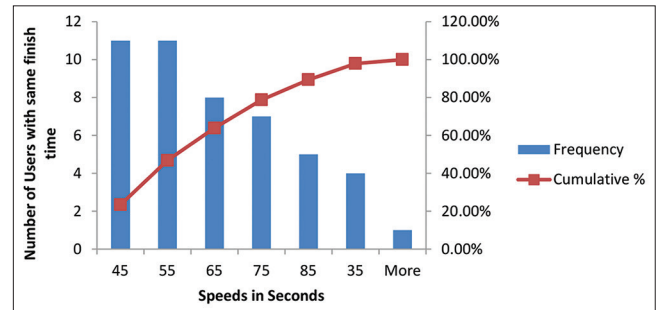


Fig. 6. Mouse game inputs performance.

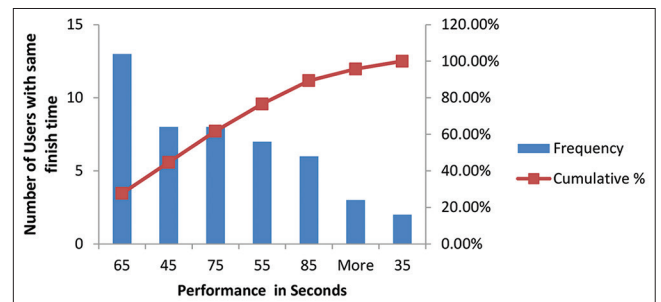


Fig. 7. Eye and keyboard input performance.

moves throughout the whole screen. When the use of the mouse, the consumer appears at the favored goal and then explicitly acts with the mouse to pass the cursor. However, with the eye pointed, truly searching at the favored spot shifted the cursor, casting off the want for any hand motion. A participant commented that, "I ought to discover with the view freely and only clicked on the mouse in case of need." People naturally use eye actions to factor when speaking with

TABLE 4: Regression statistics for eye and mouse inputs

Regression statistics								
Multiple R	0.442484013							
R square	0.195792102							
Adjusted R square	0.178309321							
Standard error	17.0641992							
Observations	48							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	3261.042739	3261.043	11.19914	0.001638			
Residual	46	13394.59714	291.1869					
Total	47	16655.63987						
	Coefficients	Standard error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	29.07932151	8.185906615	3.552364	0.000895	12.60195	45.5567	12.60195	45.55669665
Eye input	0.463467353	0.138492678	3.346512	0.001638	0.184696	0.742239	0.184696	0.742238651

TABLE 5: Regression statistics for keyboard and eye tracker input

Regression statistics								
Multiple R	0.549515							
R square	0.301967							
Adjusted R square	0.286792							
Standard Error	15.17813							
Observations	48							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	4584.343	4584.343	19.89942	5.23E-05			
Residual	46	10597.28	230.3757					
Total	47	15181.62						
	Coefficients	Standard error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	20.99721	8.226233	2.55247	0.014081	4.438661	37.55576	4.438661	37.55576
Eye space	0.577334	0.129422	4.460877	5.23E-05	0.316822	0.837846	0.316822	0.837846

TABLE 6: Descriptive statistics for the three game input mouse, keyboard, and eye tracker

Mouse		Keyboard		Eye tracking	
Mean	55.20413	Mean	56.36817	Mean	61.26604458
Standard error	2.717134	Standard error	2.594117	Standard error	2.469118725
Median	48.77977	Median	54.56662	Median	61.703085
Standard deviation	18.82486	Standard deviation	17.97257	Standard deviation	17.10655632
Sample variance	354.3753	Sample variance	323.0133	Sample variance	292.6342692
Kurtosis	4.833343	Kurtosis	-0.42544	Kurtosis	-0.35056645
Skewness	1.58376	Skewness	0.050186	Skewness	0.27816714
Range	103.2609	Range	73.81139	Range	74.30638
Minimum	30.01091	Minimum	23.09589	Minimum	27.99892
Maximum	133.2718	Maximum	96.90728	Maximum	102.3053
Sum	2649.798	Sum	2705.672	Sum	2940.77014
Count	48	Count	48	Count	48
Largest (1)	133.2718	Largest (1)	96.90728	Largest (1)	102.3053
Smallest (1)	30.01091	Smallest (1)	23.09589	Smallest (1)	27.99892
Confidence level (95.0%)	5.466169	Confidence level (95.0%)	5.21869	Confidence level (95.0%)	4.967226088

different people and eye monitoring permits the identical visible cues to be prolonged in the digital world. We consider that the distinction in overall performance between the eye tracer and the mouse throughout the sport is because of the latency that took place when taking pictures on a target. Remember, it took about a 2d for a shot to attain the place it was fired. Users regarded to have a hard time getting “led” the missiles through searching out into empty area in the front of the cellular target.

6. CONCLUSION

This paper has focused on introducing new input for video games. Eye movement is one of the important inputs in video games it increases the interactive between the players and makes it more interesting and challenging. The case study has been conducted on 48 players (undergraduate BSc students) to test the game with different input mouse, keyboard, mouse keyboard, and eye movement. The outputs have been shown the significant effect on the playing the game regarding scoring the point using eye input techniques and this adds privilege to using more input and makes the game more interactive. However, the game will slow down but this depends on game scenario. Finally, the results show significant correlation between all inputs eye, mouse, and keyboard.

ACKNOWLEDGMENT

We would like to thank all the opportunity and support University of Human Development for usual support, Head of Computer Science Department of University of Sulaimani and all students of Computer Science Department that participate in playing the game.

REFERENCES

- [1] M. R. Mine. “Virtual environment interaction techniques.” *UNC Chapel Hill Computer Science Technical Representative*, vol. 18, pp. 1-18, 1995.
- [2] B. Yuan, E. Folmer and F. C. Harris. “Game accessibility: A survey.” *Univers. Access Information Society*, vol. 10, no. 1, pp. 81-100, 2011.
- [3] K. M. Roose and E. S. Veinott. “Understanding game roles and strategy using a mixed methods approach”. In: *ACM Symposium on Eye Tracking Research and Applications*. Association for Computing Machinery, New York, United States, pp. 1-5, 2021.
- [4] Z. Li, P. Guo and C. Song. “A review of main eye movement tracking methods.” *Journal of Physics: Conference Series*, vol. 1802, No. 4, p. 042066, 2021.
- [5] C. Biele. “Eye movement.” In: *Human Movements in Human-Computer Interaction*. Springer, Cham, pp. 23-37, 2022.
- [6] A. Goettker and K. R. Gegenfurtner. “A change in perspective: The interaction of saccadic and pursuit eye movements in oculomotor control and perception.” *Vision Research*, vol. 188, pp. 283-296, 2021.
- [7] D. Ivanchenko, K. Rifai, Z. M. Hafed and F. Schaeffel. “A low-cost, high-performance video-based binocular eye tracker for psychophysical research.” *Journal of Eye Movement Research*, vol. 14, no. 3, p. 3, 2021.
- [8] A. S. Almansouri. “Tracking eye movement using a composite magnet.” *IEEE Transactions on Magnetics*, vol. 58, no. 4, p. 3152085, 2022.
- [9] C. Jiménez-Rodríguez, L. Yélamos-Capel, P. Salvestrini, C. Pérez-Fernández, F. Sánchez-Santed and F. Nieto-Escámez. Rehabilitation of visual functions in adult amblyopic patients with a virtual reality videogame: A case series. *Virtual Reality*, vol. 2021, pp. 1-12, 2021.
- [10] N. Modi and J. Singh. “A review of various state of art eye gaze estimation techniques.” In: *Advances in Computational Intelligence and Communication Technology*. Springer, Germany, 2021, pp. 501-510.
- [11] L. E. Nacke, S. Stellmach, D. Sasse and C. A. Lindley. “Gameplay experience in a gaze interaction game”. *arXiv*, vol. 2010, pp. 49-54.
- [12] K. Saroha, S. Sharma, G. Bhatia and A. Professor. “Human computer interaction: An intellectual approach.” *International Journal of Computer Science and Management Studies.*, vol. 11, no. 2, p. 2, 2011.
- [13] S. Bakker, E. Van Den Hoven and B. Eggen. “Evaluating peripheral interaction design.” *Human-Computer Interact*, vol. 30, no. 6, pp. 473-506, 2015.
- [14] R. Bernhaupt, A. Boldt, T. Mirlacher, D. Wilfinger and M. Tscheligi. “Using emotion in games: Emotional flowers.” *ACM International Conference Proceedings Series*, vol. 203, pp. 41-48, 2007.
- [15] M. Lankes, S. Riegler, A. Weiss, T. Mirlacher, M. Pirker and M. Tscheligi. “Facial expressions as game input with different emotional feedback conditions.” In: *ACE '08: Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*. Association for Computing Machinery, New York, United States, pp. 253-256, 2014.
- [16] A. Covaci, G. Ghinea, C. H. Lin, S. H. Huang, and J. L. Shih. “Multisensory games-based learning -lessons learnt from olfactory enhancement of a digital board game.” *Multimed. Tools Appl.*, vol. 77, no. 16, pp. 21245-21263, 2018.
- [17] Y. A. Sekhavat and P. Nomani. “A comparison of active and passive virtual reality exposure scenarios to elicit social anxiety.” *International Journal of Serious Games*, vol. 4, no. 2, pp. 3-15, 2017.
- [18] A. M. Darwesh, M. I. Ghareb. and S. Karimi. “Towards a serious game for kurdish language learning.” *Journal of University of Human Development*, vol. 1, no. 3, pp. 376-384, 2015.
- [19] M. I. Ghareb. HTML5, future to solve cross-platform issue in serious game development. *Journal of University of Human Development*, vol. 2, no. 4, pp. 443-450, 2016.
- [20] L. Chittaro and R. Sioni. “Affective computing vs. Affective placebo: Study of a biofeedback-controlled game for relaxation training.” *International Journal of Human-Computer Studies*, vol. 72, no. 8-9, pp. 663-673, 2014.
- [21] A. Ahmed and M. Ghareb. Design a mobile learning framework for students in higher education. *Journal of University of Human Development*, vol. 3, no. 1, p. 288, 2017.

- [22] P. Ekman, W. V. Friesen and P. Ellsworth. *Emotion in the Human Face*. Elsevier, Netherlands, 1972.
- [23] A. Kolakowska, A. Landowska, M. Szwoch, W. Szwoch and M. R. Wróbel. "Emotion recognition and its applications." *Advances in Intelligent Systems and Computing*, vol. 300, pp. 51-62, 2014.
- [24] N. Wang and S. Marsella. "Introducing EVG: An emotion evoking game." *Lecture Notes in Computer Science*, vol. 4133, pp. 282-291, 2006.
- [25] A. Landowska and M. R. Wrobel. "Affective reactions to playing digital games." In: *Proceedings-2015 8th International Conference on Human System Interaction*. IEEE, United States, pp. 264-270, 2015.
- [26] W. Szwoch. "Model of emotions for game players." In: *Proceedings - 2015 8th International Conference on Human System Interaction*. IEEE, United States, pp. 285-290, 2015.
- [27] S. Bakker, E. Van Den Hoven, B. Eggen and K. Overbeeke. "Exploring peripheral interaction design for primary school teachers," *Proceedings 6th International Conference Dedicated to Research in Tangible, Embedded*, vol. 1, no. 212, pp. 245-252, 2012.
- [28] S. Bakker, E. Van Den Hoven and B. Eggen. "FireFlies: Supporting primary school teachers through open-ended interaction design." In: *Proceedings 24th Australian Computer Interaction Conference*. Association for Computing Machinery, New York, United States, pp. 26-29, 2012.
- [29] D. Hausen, H. Richter, A. Hemme, and A. Butz. "Comparing input modalities for peripheral interaction: A case study on peripheral music control." *Lecture Notes in Computer Science*, vol. 8119, no. 3, pp. 162-179, 2013.
- [30] R. J. K. Jacob. "What you look at is what you get: Eye movement-based interaction techniques." *Proceedings ACM*, vol. 90, pp. 11-18, 1990.
- [31] R. J. K. Jacob. "No TitleThe use of eye movements in human-computer interaction techniques: What you look at is what you get." *ACM Transactions on Information Systems*, vol. 9, pp. 152-169, 1991.
- [32] R. J. K. Jacob. "Eye movement-based human-computer interaction techniques: Toward non-command interfaces". In: H. R. Harst and D. Hix. (eds.), *Advances in Human/Computer Interaction*. vol. 4. Hindawi, United Kingdom, pp. 151-190, 1993.
- [33] R. J. K. Jacob. "What you look at is what you get: Using eye movements as computer input." *Proc. Virtual Real. Syst.*, vol. 93, pp. 164-166, 1993.
- [34] R. J. K. Jacob. *Eye Tracking in Advanced Interface Design*. In: *Virtual Environments and Advanced Interface Design*. Vol. 258. Oxford University Press, Inc., Oxford, p. 288, 1995.
- [35] L. E. Sibert and R. J. K. Jacob. "Evaluation of eye gaze interaction." In: *Conference on Human Factors in Computing Systems Proceedings*. Association for Computing Machinery, New York, United States, pp. 281-288, 2000.
- [36] A. Murata, T. Miyake and M. Moriwaka. "Effectiveness of the menu selection method for eye-gaze input system." *Japanese Journal of Ergonomics*, vol. 47, no. 1, pp. 20-30, 2011.
- [37] P. Isokoski and B. Martin. "Eye tracker input in first person shooter games." In: *Proceedings 2nd Conference Communication by Gaze Interact*. Association for Computing Machinery, New York, United States, pp. 78-81, 2006.
- [38] P. Isokoski, A. Hyrskykari, S. Kotkaluoto and B. Martin. "Gamepad and eye tracker input in FPS games: Data for the first 50 min." *Proceedings 3rd Conference Communication by Gaze Interact*. COGAIN, Denmark, pp. 1-5, 2007.
- [39] A. T. Duchowski. *Eye Tracking Methodology: Theory and Practice*. Springer Verlag, London, UK, 2003.
- [40] J. D. Smith and T. C. N. Graham. "Use of eye movements for video game control." In: *Proceedings of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*. ACE, Hollywood, CA, 2006.
- [41] J. Leyba and J. Malcolm. "Eye tracking as an aiming device in a computer game." In: *Course work (CPSC 412/612 Eye Tracking Methodology and Applications by A. Duchowski)*. Clemson University, Clemson, CA, 2004.
- [42] P. Isokoski and B. Martin. "Eye tracker input in first person shooter games." In: *Proceedings of the 2nd Conference on Communication by Gaze Interaction: Communication by Gaze Interaction-COGAIN 2006 Gazing into the Future*. COGAIN, Turin, Italy, pp. 78-81, 2006.
- [43] P. Isokoski, M. Joos, O. Spakov, and B. Martin. "Gaze controlled games." *Universal Access in the Information Society*, vol. 8, pp. 323-337, 2009.
- [44] E. Lacorte, G. Bellomo, S. Nuovo, M. Corbo, N. Vanacore and P. Piscopo. "The use of new mobile and gaming technologies for the assessment and rehabilitation of people with ataxia: A systematic review and meta-analysis." *Cerebellum*, vol. 20, no. 3, pp. 361-373, 2021.
- [45] M. Dorr, L. Pomarjanschi and E. Barth. "Gaze beats mouse: A case study." *Psychology Journal*, vol. 7, pp. 197-211, 2009.
- [46] B. Jacob, J. Chen, Y. Huang and I. Cohen. "Pearson correlation coefficient." In: *Noise Reduction in Speech Processing*. Springer, Berlin, Heidelberg, pp. 1-4, 2009.

Modified Advanced Encryption Standard for Boost Image Encryption

Nigar M. Shafiq Surameery

Department of Information Technology, College of Computer and Information Technology, University of Garmian, Kalar, Sulaimani, Kurdistan Region, Iraq



ABSTRACT

Cryptography is a field of study that deals with converting data from a readable to an unreadable format. It can provide secrecy, data integrity, authenticity, and non-repudiation services. Security has become a concern for the community because of the technology's potential use in numerous sectors of any company, market, agency, or governmental body, information. The cryptosystems ensure that data are transported securely and only authorized individuals have access to it. Deeply encrypted data that cannot be deciphered through cryptanalysis are in high demand right now. There are a variety of encryption algorithms that can guarantee the confidentiality of data. For multimedia data, standard symmetric encryption algorithms (AES) can give superior protection. However, using the symmetric key encryption approach on more complicated multimedia data (mainly photos) may result in a computational issue. To address this issue, the AES has been modified to satisfy the high computing requirements due to the complex mathematical operations in MixColumns transformation, which slow down the encryption process. The modified AES uses bit permutation to replace the MixColumns transformation in AES because it is simple to construct and does not require any complex mathematical computation. This research focuses on using the Modified Advanced Encryption Standard (MAES) algorithm with 128 and 256 bit key sizes to encrypt and decrypt image data. The algorithms were implemented using the Python programming language without complex mathematical computation. By comparing the MAES algorithm with the original AES algorithm, the results showed that the MAES requires less encrypting and decryption time with higher efficiency for all file sizes.

Index Terms: Advanced Encryption Standard, Bit Permutation, Image File Encryption, Symmetric Cipher, Modified Advanced Encryption Standard

1. INTRODUCTION

The world is evolving due to the advancements in technology, attitudes, and investment. Massive amounts of data are being sent, received, and stored by government, commercial, consumer, and manufacturing groups for intelligence, safety, and other purposes such as medical, video monitoring,

and IoT systems. The security requirements for images originate from the need to protect its data from two types of threats: Unintentional loss and corruption, and intentional unauthorized access or manipulation. Cybersecurity is generally recognized as the most basic and effective means of protecting images against both unintentional and intentional fraud [1]. Images data (original data) are converted into an encrypted version (encrypted information) to be saved or sent through an unsecured channel. Encrypted transmission is safe if only the sender and receiver could access the cipher and decipher algorithms and keys. The process of ensuring the integrity, availability, and privacy of data is known as information security.

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp52-59

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Copyright © 2022 Surameery. This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Nigar M. Shafiq Surameery, Department of Information Technology, College of Computer and Information Technology, University of Garmian, Kalar, Sulaimani, Kurdistan Region, Iraq. E-mail: nigar.mahmoud@garmian.edu.krd

Received: 31-12-2021

Accepted: 01-04-2022

Published: 27-04-2022

Nowadays, the majority of businesses and individuals keep their data on computers. Thus, people have more access to information kept in computer databases. Much of the Stored data are very sensitive and should not be shared with the general public. Data encryption is a common technique for protecting secret information by encrypting it with a pre-existing or pre-written algorithm. The key generation is the most potent encryption aspect, which is divided into two parts: Symmetric key generation and asymmetric key generation. With the help of modern high-performance computing devices, hackers may quickly break the key. Therefore, encrypted data that cryptographic analyzers cannot decrypt are currently in demand [2]. The encryption and decryption processes are affected by the encrypted file size and the length of the key [3]. The key used for encryption/decryption in symmetric cryptography is the same. As a result, the required distribution must occur before information transfer. In symmetric key cryptography, as shown in Fig. 1, the key is highly significant because the system's security is directly related to the key's characteristics, such as key length. Advanced Encryption Standard (AES), DES, TRIPLE DES, BLOWFISH, RC4, and RC6 are examples of symmetric cryptography [4].

AES is one of the most used symmetric algorithms and a new generation of data encryption standards. It offers the most outstanding safety and highest operation speed [5]. The AES is the next-generation cryptographic system with many benefits, including strict security, high efficiency, customization, and convenience.

The research establishment has shown keen interest in image encryption to protect valuable images from hackers in recent years. As a result, several encryption algorithms have been used for encrypting images to achieve this goal. This research aims to show the efficiency and speed of the modified AES algorithm compared to the original AES.

2. LITERATURE REVIEW

With the rapid growth of digital data interchange, secure data transfer is becoming increasingly crucial. since images

are increasingly being used in different processes, it is vital to keep confidential image data safe from unauthorized access [6]. In Maolood *et al.* [7], modifications were made to the original AES algorithm to increase security. Using evaluation measures, the effectiveness of the modifications was demonstrated, and the improved algorithm was evaluated. The results showed a small increase in the encryption time of the modified AES algorithm compared to the original AES algorithm because of adding several layers of security. The authors in Ahamad and Abdullah [8], compared the four encryption algorithms (Blowfish, AES, XOR, and RSA), applied to four types of data represented by: Text, image, audio, and video. The simulations demonstrate that AES is more time efficient than all other methods. In another work [9], the AES round numbers have been increased to 16 to be more secure. Theoretical study and practical results revealed that this AES strategy provides great speeds and less data transmission over unprotected networks. Moreover, the research study in Arman *et al.* [10] focuses on symmetric Key cryptosystem represented by the AES algorithm and proposes a significant improvement over AES in terms of privacy and ease of operation. The goal was to achieve a rapid implementation time and low memory consumption while maintaining a high security level. As a result, this solution obtained 88% of efficiency and a faster net processing time. This is a big change in terms of efficiency over the standard AES. In the more advanced researches, to address the poor diffusion rate in the early rounds, the AES cipher round and key schedule have been updated to include more primitive operations such exclusive OR and modulo operations [11]. According to the avalanche effect evaluation results, the results show that the modified AES has enhanced propagation and ambiguity features because there was a 61.98% increase in dispersal in round 1, 14.79% in round 2, and 13.87% in round 3. However, in Shakor and Surameery [12], AES is combined with elliptic curve cryptography (ECC) algorithm which has been used to protect the sensitive COVID-19 files that are stored in the cloud. Although different layers of security have been added, the results indicated that the impact of the hybrid system was more significant than the other secure algorithms without

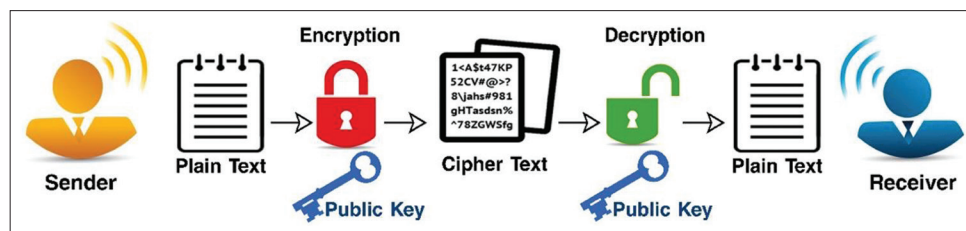


Fig. 1. Symmetric key cryptography.

affecting the consumption time. The author in Mohd and Ashawesh [13] proposed a secured modified AES (MAES) algorithm, which reduces the rounds of AES to 14 rounds to decrease the consumption time of the encryption and decryption process, and improve the digital data security. The results prove that, compared with other studies, the proposed method uses less encryption and decryption time while providing better security, as confirmed by the avalanche effect test. Moreover, in Lin *et al.* [14], an improved AES cryptosystem based on chaotic synchronization with dynamic random keys is proposed. In the traditional AES encryption system, static keys are used, which must be exchanged in advance and confirmed for safe preservation. The chaotic system's synchronization technology, on the other hand, was used to overcome the inadequacies of key storage. In the proposed approach, the static key becomes dynamic and random, and it no longer needs to be kept or transmitted over an open channel. In Hafsa *et al.* [15], the authors use an efficient type of a built-in approach that utilized the AES and elliptic curve encryption (ECC) algorithms for encrypting medical images. It combines the advantages of symmetric AES accelerated data encryption and asymmetric ECC to ensure the exchange of symmetric session keys. The security analysis was successfully performed, and their experiments proved that the proposed technology provides a more straightforward and correct cryptographic basis. Furthermore, the evaluation results demonstrate the proposed algorithm's effectiveness, rapidity, and high security.

Hence, after studying various previous researches on the modification of AES, we came up with a new technique that increases the processing speed and data security.

3. AES

The AES is a symmetric key method that has been developed by the United States National Institute of Standards and Technology as the standard for encrypting digital data. AES is the most powerful algorithm in recent times since it is the only vulnerable to brute force attacks, which makes it difficult for cryptanalysts to crack. It is commonly used in banks and organizations to protect critical information because it promises cybersecurity [16]. The size of the key determines the number of rounds for encryption and decryption in AES, which could be 10, 12, or 14 rounds for 128, 192, and 256 bit keys, respectively [17]. AES is one of the most widely used symmetric key algorithms and a new generation of data encryption standards. It offers the greatest secrecy and highest operation speed [5].

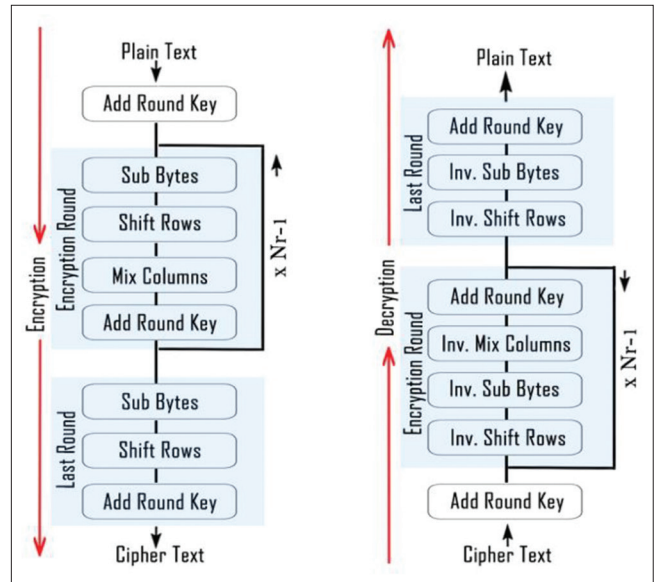


Fig. 2. Structure of AES algorithm.

As shown in Fig. 2, the AES algorithm comprises four invertible transformations: SubBytes, ShiftRows, MixColumns, and AddRoundKey. All of the encryption rounds use these transforms, with the exception of the end round, which removes the MixColumns transformation to make the encryption and decryption methods symmetric.

- SubBytes transform: Which substitutes each byte of the original block of data (D_0, \dots, D_{15}) with the row (initial 4 bits) and column (next 4 bits) of a 16×16 Substitution Box (S-Box). The S-Box includes unique mathematical qualities that guarantee changes in individual state bits spread swiftly throughout the ciphertext, causing confusion. During decryption, a reverse substitution box (InvS-Box) is utilized to reverse the impact of the SubBytes translation.
- ShiftRows Transform: This manipulates the state's rows by shifting the bytes in each row using a specific offset. The first row remains constant during this process, while the second, third, and fourth rows are subjected to 1 byte, 2 byte, and 3 byte circular shift operations. The first row remains untouched throughout the decryption process, while the following rows are moved to the right using the same offset that was used to shift them to the left during the encryption process
- MixColumns transform: Represents the mixing operation that applies the XOR operation on the state's columns, combining the four bytes in each column with the four bytes in a fixed matrix.
- AddRoundKey Transform: This is the last transformation that will be applied for each round. The XOR operation

is employed to execute addition operation between the bytes of the modified state and the round key.

The encryption process consists of a number of different transformations applied consecutively over the data block bits, in a fixed number of iterations, called rounds. The number of rounds depends on the length of the key used for the encryption process. The AES encryption process is as follows

Cipher(byte in $[X*Nb]$, byte out $[X*Nb]$, word $w[Nv*(Nr+1)]$)

```

Begin
Byte state[X, Nb]
State=in
AddRoundKey(state, w[0,Nb-1])
For round=1 step 1 to Nr-1
SubBytes(state)
ShifRows(state)
MixColumns(state)
AddRounKey(state, w[round*Nb, (round+1)*Nb-1])
end for
SubBytes(state)
ShifRows(state)
AddRounKey(state, w[Nr*Nb, (Nr+1)*Nb-1])
Out=state
End
    
```

4. MAES ALGORITHM

To secure the image files and increase their privacy, a modified AES algorithm is employed to encrypt them. The AES is modified to address its higher computational requirement due to the complex mathematical operations in MixColumns transformation making the encryption process slow. Bit permutation includes merely shifting the positions of bits in each stage and does not require any significant mathematical computation. The bit permutation transformation takes the place of the MixColumns transformation in AES, Fig. 3, depicts the bit permutation modification of AES.

For the encryption process, take the state value per column (for example, column 0) from ShiftRows transformation. Each column has four rows, as shown in Fig. 4. Moreover, each state in column 0 is composed of eight bits, represented as $((x,0), b)$ where $x, 0,$ and b indicate the row, column, and bit number for each state, respectively. Subsequently, the states are partitioned into four blocks matrix, called block 0, block 1, block 2, and block 3 and the transpose of each block

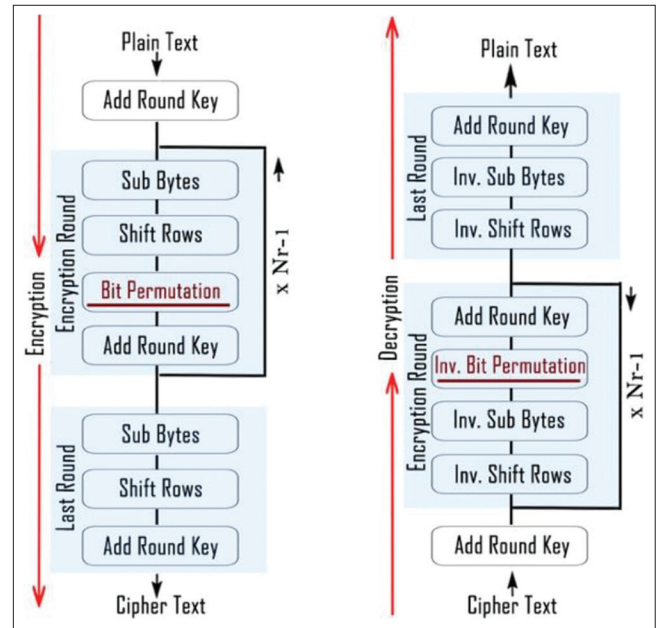


Fig. 3. Structure of MAES algorithm.

matrix is the next step of the process. Finally, to get the new value at state $a'(x,y)$, we need to do a row-wise concatenation of the bit values of the transposed block where x is the column_value and y is the block_value. MAES encryption process is as follows:

Cipher(byte in $[X*Nb]$, byte out $[X*Nb]$, word $w[Nv*(Nr+1)]$)

```

Begin
Byte state[X, Nb]
State=in
AddRoundKey(state, w[0,Nb-1])
For round=1 step 1 to Nr-1
SubBytes(state)
ShifRows(state)
Bit Permutation(state)
AddRounKey(state, w[round*Nb, (round+1)*Nb-1])
end for
SubBytes(state)
ShifRows(state)
AddRounKey(state, w[Nr*Nb, (Nr+1)*Nb-1])
Out=state
End
    
```

For the decryption process, the inverse bit permutation is utilized. There are four columns in each row. In the first steps, starting with row 0, we take the state value from the InvSubBytes transformation. Next, each state is made up of 8 bits, denoted by $((0, y), b)$, where 0 denotes row 0, y

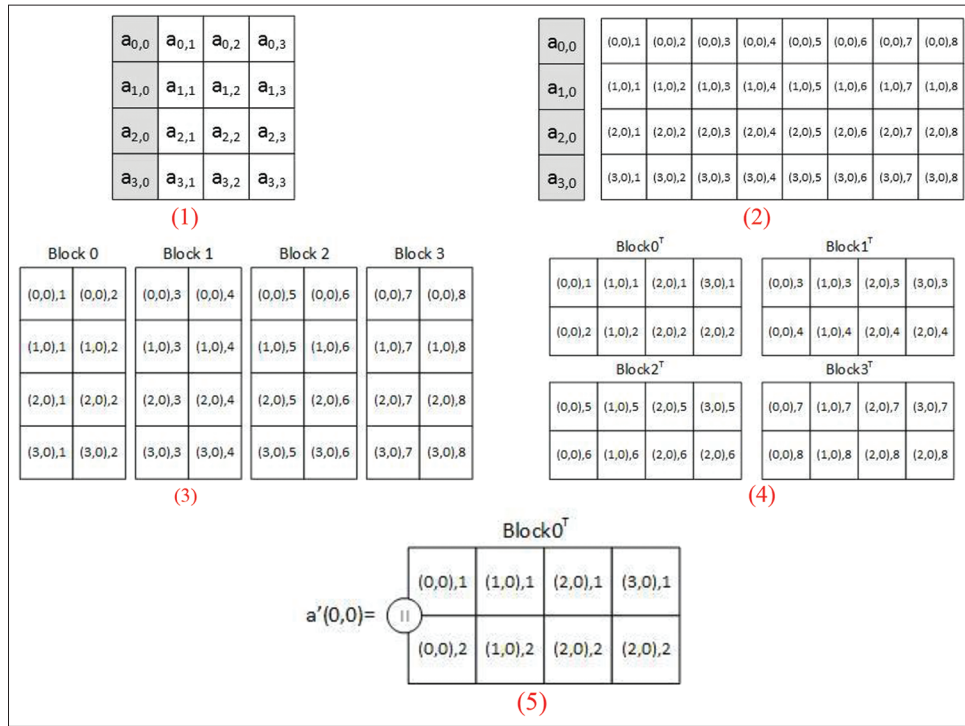


Fig. 4. Bit permutation.

denotes column, and b denotes the bit position in each state. Taking the number of bits per state creates 1×32 matrix. Subsequently, the matrix is partitioned into eight blocks, called block 0 through block 7. After that, the blocks 0 to 7 are transposed, yielding eight 4×1 transposed matrices. Finally, to get the new value at state $a'(x,y)$, we need to do row-wise concatenation of the bit values of the transposed block, where x is the block_value and y is the row_value. Fig. 5.

5. EXPERIMENTAL RESULTS AND DISCUSSION

This section shows the experiment run to derive the effect of image size on the encryption and decryption time and includes a comparison between the modified AES system and the existing standard AES solution. To apply security, AES-128, AES-256, MAES-128, and MAES-256 algorithms were used in the encryption and decryption processes.

These two processes were implemented using the Python programming language. Both standard and modified AES algorithms are written in Spyder, the Scientific Python Development Environment, and simulated on Intel Core i7-2.8 GHz processor with 16 GB RAM and running over Mac OS High Sierra.

The main purpose of this study is to calculate the encryption and decryption speed of each of the algorithm under the study of different image sizes. The AES and MAES algorithms were evaluated for 10 trials to determine the average encryption and decryption time. Their implementation is tried to optimize the maximum performance for the algorithm. Therefore, the throughput for encryption is calculated for each algorithm. Encryption time is used to calculate the throughput of an encryption scheme. The throughput of the encryption scheme is calculated by dividing the total plain image size in kilobytes by total encryption time in second for each algorithm. If the throughput value is increased, the power consumption of this encryption technique is decreased.

$$\text{Throughput} = \text{Plaintext (KB)} / \text{Encryption time (Sec.)}$$

In this research, different sizes of image files were experimented such as 200 KB, 400 KB, 800 KB, and 1600 KB. The performance metrics are analyzed as encryption and decryption time and throughput. The performance of the encryption and decryption time was measured in milliseconds.

For the 128 bit key, it was observed that both the encryption and decryption times increase as the image size increases, as

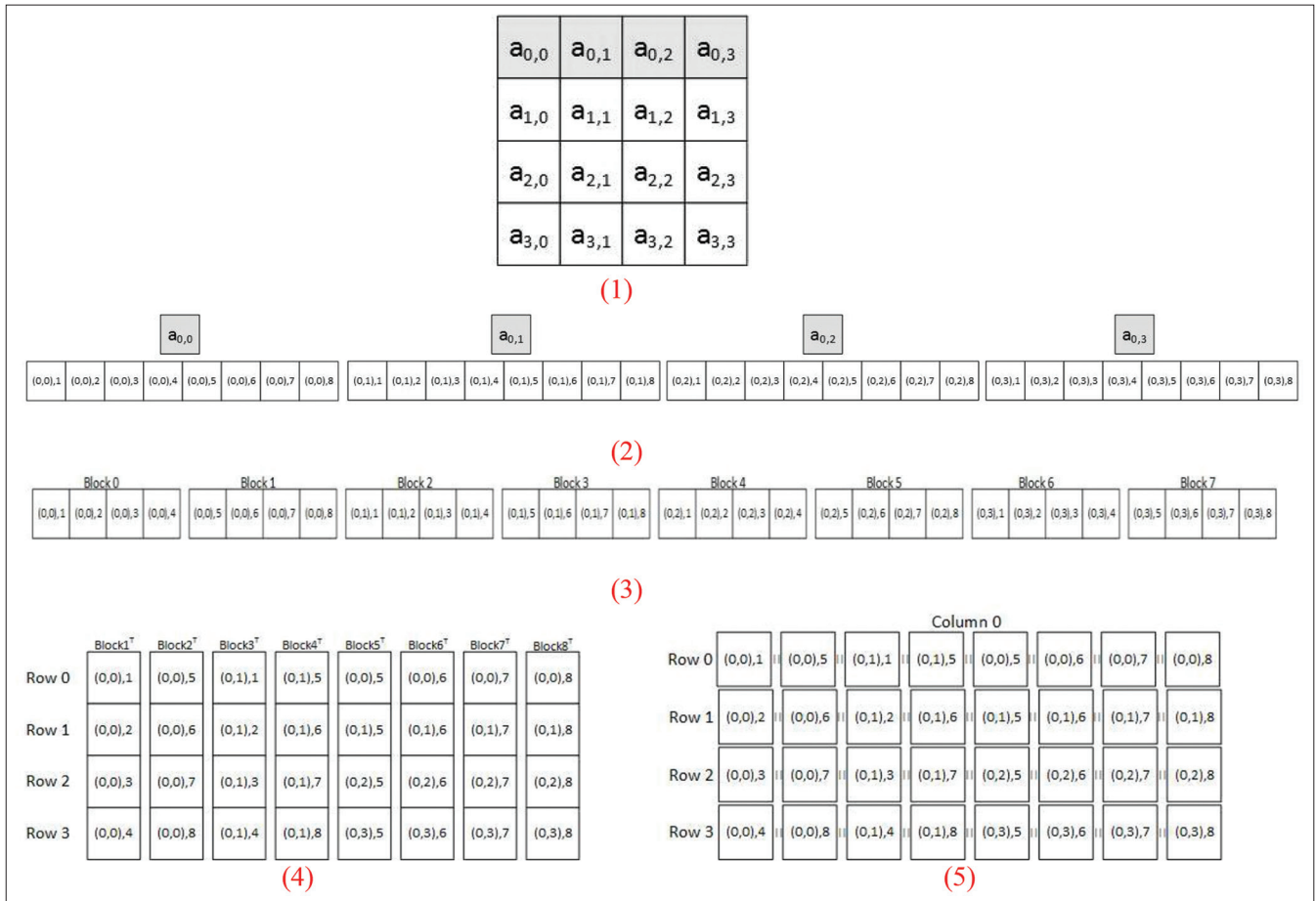


Fig. 5. Inverse bit permutation.

TABLE 1: Comparison of AES and MAES encryption time using 128 bit key size

File size	Encryption time (ms)	
	AES	Modified AES
200 KB	28.24	27.22
400 KB	52.11	51.26
800 KB	113.11	111.45
1600 KB	242.89	240.04
Average time	436.35	429.97
Throughput	6.87	6.97

TABLE 3: Comparison of AES and MAES encryption time using 256 bit key size

File size	Encryption time (ms)	
	AES	Modified AES
200 KB	30.21	26.01
400 KB	58.49	56.11
800 KB	117.24	115.89
1600 KB	249.33	243.34
Average time	455.27	441.35
Throughput	6.589	6.797

TABLE 2: Comparison of AES and MAES decryption time using 128 bit key size

File size	Decryption time (ms)	
	AES	Modified AES
200 KB	29.15	28.70
400 KB	57.02	55.06
800 KB	119.06	117.82
1600 KB	244.88	242.55
Average time	450.11	444.13
Throughput	6.6650	6.754779

TABLE 4: Comparison of AES and MAES decryption time using 256 bit key size

File size	Decryption time (ms)	
	AES	Modified AES
200 KB	30.95	29.95
400 KB	60.40	59.11
800 KB	124.61	122.71
1600 KB	256.24	254.43
Average time	472.2	466.2
Throughput	6.353	6.435

TABLE 5: Comparison of proposed method with another research solutions

Author	Image size	Key size	Encryption time		Decryption time	
			AES (ms)	Modified AES (ms)	AES (ms)	Modified AES (ms)
[18]	200	128 bit	206,016	196,576	176,618	167,532
This study	200	128 bit	28.24	27.22	29.15	28.70

shown in Tables 1 and 2. However, it was noticed that the encryption and decryption processes of the MAES take less time than the standard one for the same image size.

Regarding the enciphering process, for the file sizes of 200, 400, 800, and 1600 KB, the encryption time of the AES was 28.24, 52.11, 113.11, and 242.89 ms, respectively. In comparison, the encryption time of the MAES for the same file sizes were 27.22, 51.26, 111.45, and 240.04 ms, respectively. Moreover, the decryption time of the AES was 28.24, 52.11, 113.11, and 242.89 ms, respectively. In contrast, the decryption time of the MAES for the same file sizes were 27.22, 51.26, 111.45, and 240.04 ms, respectively.

On the other hand, ultra-high security may be achieved using AES-256 due to the number of rounds, which provides privacy to unauthorized users. Therefore, the system has been tested with AES-256 and MAES-256 for different file sizes. For the file sizes of 200, 400, 800, and 1600 KB, the encryption process of MAES-256 was faster than AES-256 by 0.74, 1.91, 7.37, and 6.91 ms, respectively. In the same way, the decryption process of MAES-256 was faster than AES-256 by 3.94, 3.00, 6.82, and 11.0991 ms, respectively, as shown in Tables 3 and 4. Thus, in comparison with AES, the presented results show that slightly less encryption and decryption times in milliseconds were recorded with the bit permutation in MAES for all of the image file sizes, which increases the security, overall performance, and effectiveness of the system.

By considering the results in Tables 1-4, the average execution time for image data types of AES are 436.35, 450.11, 455.27, and 472.2, respectively, and MAES are 429.97, 444.13, 441.35, and 466.2, respectively. Similarly, the throughput are (6.87, 6.6650, 6.589, and 6.353) and (6.97, 6.754779, 6.797, and 6.435) for AES and MAES, respectively. The results show that the standard AES algorithm has higher encryption average time and the MAES has higher throughput. Hence, MAES is more efficient in image data encryption than AES.

Besides, a comparative analysis of the obtained results was carried out with an existing work existing works. This was a little bit of tasking, as there are no standard performance met-

rics that are widely and generally acceptable by all researchers in this regard. While some measured the performance of their modified AES version using text files of different sizes, some used images and video files. However, most authors employed execution time as their performance metrics and just few researches employed the encryption and decryption times, separately.

Therefore, the proposed technique was compared with the recent solutions for the “Implementation of Modified AES as Image Encryption Scheme” that was proposed in 2018 [18]. The research was available to investigate the use of modified AES for image encryption.

As presented in Table 5, the proposed technique provides optimistic solutions by recording lower execution time in milliseconds than other research solution. Although this may not be clearly noticeable in real-life application, yet it is significant.

6. CONCLUSION

Cryptography is a widely used approach for ensuring data transfer and image storage security. The proposed algorithm aims to reinforce and improve the standard AES method by replacing the MixColumns transformation with bit permutation. With different file sizes, the new algorithm is tested for encryption/decryption time in milliseconds. The results reveal that the MAES achieves a higher level of security with requiring less time. The key sizes used to compare AES and MAES were 128 bits and 256 bits.

The results showed that the encrypting time in the files whose size was between 200 and 1600 KB in the MAES algorithm was less than AES by 1.595% with the 128 bit key size, while MAES has a lower percentage of encryption time than AES by 3.48% when applying a 256 bit key size. In the same way, MAES decryption time was 1.495% faster than AES with a 128 bit key size and 1.5% faster with a 256 bit key size. In addition, the results show that the standard AES algorithm has higher encryption average time and the MAES has higher throughput. Hence, MAES is more efficient in image data encryption than AES. Furthermore, the proposed MAES cryptosystem provides improved security in terms

of protection against different types of attacks. As a part of future work, we can use additional linear processes experiments to boost AESs confusion property.

REFERENCES

- [1] S. M. Ali Ebrahim. "Hybrid chaotic method for medical images ciphering". *The International Journal of Network Security and Its Applications*, vol. 12, no. 6, pp. 1-14, 2020.
- [2] P. Dixit, A. K. Gupta, M. C. Trivedi and V. K. Yadav. "Traditional and hybrid encryption techniques: A survey". *Lecture Notes on Data Engineering and Communications Technologies*, vol. 4, pp. 239-248, 2018.
- [3] R. Yudistira. "AES (advanced encryption standard) and RSA (rivest-shamir-adleman) encryption on digital signature document: A literature review". *International Journal of Information Technology and Business*, vol. 2, no. 1, pp. 1-3, 2020.
- [4] M. Agrawal and P. Mishra. "A comparative survey on symmetric key encryption techniques". *International Journal of Computational Science and Engineering*, vol. 4, no. 5, pp. 877-882, 2012.
- [5] J. Liu, C. Fan, X. Tian and Q. Ding. "Optimization of AES and RSA algorithm and its mixed encryption system". In: *Smart Innovation, Systems and Technologies*. Vol. 82, pp. 393-403, 2018.
- [6] M. Zeghid, M. Machhout, L. Khriji, A. Baganne and R. Tourki. "A modified AES based algorithm for image encryption". *International Journal of Computational Science and Engineering*, vol. 1, no. 1, p. 70, 2007.
- [7] A. T. Maalood and Y. A. Yasser. "Modifying Advanced Encryption Standard (AES)". Algorithm University of Technology Department of Computer Sciences University of Technology Department of Computer Sciences, pp. 259-285, 2017.
- [8] M. M. Ahamad and M. I. Abdullah. "Comparison of encryption algorithms for multimedia". *Rajshahi University Journal of Science and Engineering*, vol. 44, pp. 131-139, 2016.
- [9] B. N. Rao, D. Tejaswi, K. A. Varshini, K. P. Shankar and B. Prasanth. "Design of modified AES algorithm for data security". *International Journal For Technological Research In Engineering*, vol. 4, no. 8, pp. 1289-1292.
- [10] M. S. Arman, T. Rehnuma and M. M. Rahman. "Design and Implementation of a Modified AES Cryptography with Fast Key Generation Technique". *Proceedings of 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE)*, pp. 191-195, 2020.
- [11] E. M. De Los Reyes, A. M. Sison and R. P. Medina. "Modified AES cipher round and key schedule". *Indonesian Journal of Electrical Engineering and Informatics*, vol. 7, no. 1, pp. 28-35, 2019.
- [12] M. Y. Shakor and N. M. S. Surameery. "Built-in Encrypted Health Cloud Environment for Sharing COVID-19 Data". In: *3rd International Conference on Computer Communication and the Internet, ICCCI 2021*, pp. 96-101, 2021.
- [13] N. A. A. Mohd and A. Y. A. Ashawesh. "Enhanced AES algorithm based on 14 rounds in securing data and minimizing processing time". *Journal of Physics: Conference Series*, vol. 1793, no. 1, p. 012066, 2021.
- [14] C. H. Lin, G. H. Hu, C. Y. Chan and J. J. Yan. "Chaos-based synchronized dynamic keys and their application to image encryption with an improved aes algorithm". *Applied Sciences*, vol. 11, no. 3, pp. 1-16, 2021.
- [15] A. Hafsa, A. Sghaier, J. Malek and M. Machhout. "Image encryption method based on improved ECC and modified AES algorithm". *Multimedia Tools and Applications*, vol. 80, no. 13, pp. 19769-19801, 2021.
- [16] A. Gupta and M. Jaiswal. "*The Safety of Next Generation Internet of Things*". pp. 422-427, 2017.
- [17] P. V. Jaswanth, B. R. Reddy, M. S. P. Kumar and M. J. P. Priyadarsini. "Color image encryption using AES and RSA". *The International Journal of Engineering and Advanced Technology*, vol. 9, no. 5, pp. 547-550, 2020.
- [18] H. V. Gamido, A. M. Sison and R. P. Medina. "Implementation of modified AES as image encryption scheme". *Indonesian Journal of Electrical Engineering and Informatics*, vol. 6, no. 3, pp. 301-308, 2018.

Adaptive Filter based on Absolute Average Error Adaptive Algorithm for Modeling System



Goran Saman Nariman, Hamsa D. Majeed

Department of Information Technology, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq

ABSTRACT

Adaptive identification of the bandpass finite impulse response (FIR) filtering system is proposed through this paper using variable step-size least mean square (VSS-LMS) algorithm called absolute average error-based adjusted step-size LMS as an adapted algorithm. This algorithm used to design an adaptive FIR filter by calculating the absolute averaged value for the recently assessed error with the previous one. Then, the step size has been attuned accordingly with consideration of the slick transition of the step size from bigger to smaller to score an achievement through high convergence rate and low steady-state misadjustment over the other algorithms used for the same purpose. The simulation results through the computer demonstrate remarkable performance compared to the traditional algorithm of LMS and another VSS-LMS algorithm (normalized LMS) which used in this paper for the designed filter. The powerful of the filter has been served in the identification system, bandpass filter has been chosen to be identified in the proposed adaptive system identification. It reports conceivable enhancements in the modeling system concerning the time of convergence, which is well-defined as a fast and steady-state adjustment defined with a low level. The designed filter identified the indefinite system with less than 10 samples; meanwhile, other algorithms were taking more than 20 samples for identification. Alongside the fine behavior of preserving the tradeoff between miss adjustment and the capability of tracking, the fewer calculations and computations regarding the algorithm requirement make the applied real-time striking.

Index Terms: Adaptive Filtering System, Finite Impulse Response design, System Identification, Variable Step-size least mean square algorithm, Absolute Average Error-Adjusted Step Size

1. INTRODUCTION

The main concern of system modeling is a system estimation basis on data observation. This includes the requirement of the model structure, parameters prediction of the unknown plant, and proof of the resulting plant for input-output systems [1]. The main goal of system discovery is to provide a suitable mathematical model or equation for the unknown system that is considered; then, the model is constructed

mathematically as a reference based for the next identification design [2]. For multiparameter systems such as filters, a typical optimization problem occurred, the filter parameters must be chosen to reduce the difference between measured and predicted filters to the minimum state. The difficulty level increased by increasing those parameters, especially with different scales, and the noise and incomplete measurements. Furthermore, the reference filter that has been selected has a major effect on the accuracy and strength of the procedure of identification [3]. The progression of computer algorithms causes an advancement in modeling for a variety of systems. Different intelligent methods are proposed for highly accurate estimation for both non-linear and linear systems [4]. In this work, an adapted filter was designed with different adaptive algorithms and engaged in the modeling system to identify a bandpass filter (BPF) with the minimum difference

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp60-69
E-ISSN: 2521-4217
P-ISSN: 2521-4209

Copyright © 2022 Nariman and Majeed. This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Goran Saman Nariman, Department of Information Technology, College of Science and Technology, University of Human Development, Kurdistan Region, Iraq. E-mail: goran.nariman@uhd.edu.iq

Received: 21-11-2021

Accepted: 03-04-2022

Published: 07-05-2022

between filter coefficients and output signal. An adaptive filter (AF) basically is a digital filter which has a changeable transfer function and adjustable weights (taps), the change is implemented based on the designed filter [5]. The filter weights are adjusted using different techniques being used for implementations in several applications such as system identification, noise cancellation, acoustic echo cancellation, channel estimation, and adaptive line enhancement, system identification is considered as the most crucial application of adapted filters as this work presents. In order of adaptation process, different algorithms are employed for the weights adjustable process useful and effective depending on their concentrated arithmetic structure, leading to efficiency improvement of the performance [6]. This research paper deals with the execution of the LMS, normalized LMS (NLMS), and absolute average error-based adjusted step-size LMS (AAE-ASSLMS) algorithms. The algorithm of LMS is considered an exceptional algorithm of gradient search. The algorithm is often used for many applications on AFs assignable to its simplicity in implementation and the small number of calculation requirements, leading to its usage in countless applications [7]. In most adaptive system applications, fast convergence and low misadjustment are a demand. Although, the algorithm comes down with decelerate convergence time which is conversely followed by the step size. Yet, the selection of large adaptation step size yields into fast convergence but this selection outcomes downslope the steady-state attitude which is directly involved in the misadjustment, this conciliation leads to weakening the LMS functionality. As such, several researchers continually looked for alternative approaches to enhance its performance [8], [9]. From the common approaches, regarding the time-varying step size for updating the weight is the NLMS algorithm, furthermore, the AAE-ASSLMS algorithm is adopted in the filter design, and a comparison is made between the adaptive algorithms to highlight the performance improvement in means of convergence time and the misadjustment in the steady state.

2. LITERATURE REVIEW

The systems with AFs are used for various applications, because of their stability, effectively support many applications, the capableness of implementation, and linear phase response production. Several researchers proposed either new or modified adapted algorithms which are employed in various filter designs then those filters had been served in different systems through many fields as a means of achieving performance enhancement of the designed filter and the

system that filter served in it [10]. Proposed a system to enhance an algorithm for comb adaptive filtering to manage the response equalization for the multipath communication channel, the system provides a high-speed tuning in the receiver side where the multispeed adaptive system is, and the processing of the computational costs in real time is obviously reduced, beside the algorithm analysis. As a result, the developed algorithm provided rate advancement and computational reduction complexity compared with LMS. A boosted AF is proposed by Kari *et al.* [11] using many adapted filtering algorithms those are RLS, LMS, and piecewise linear regression, the proposed algorithm mixture improved the performance opposite to the corresponding adaptive filtering approaches under the same used statistical conventions by establishing gains to mean square error (MSE). Through the years, lots of approaches had been proposed regarding varying the step size [12], presented a combination of adapted filters to enhance its performance, combinations of two algorithms (RLS and LMS) are considered with performance analysis as attested by the time-varying solutions. That convex combination between those algorithms gave the opportunity of achieving performance close to the optimal EMSE and MSD with lower computational complexity. Since LMS has a fixed step size, Si-Min *et al.* [13] proposed a system that enhanced the algorithm by investigating the skewness prediction distribution of LMS parameters as a third-order statistical feature in the purpose of vibration signal filtration system, the outcome shows improvements in error measurement angles and attitude accuracy. Meanwhile, Silva *et al.* [14] adopted (LMS) for adapting a system that assesses the vibration measurement of the drilling tool, the outcome shows measurement accuracy enhancement in the term of steerable drilling tools. The output demonstrated the feasibility of skewness growing limitless despite the existence of adaptive weights which present convergence in average and the mean square attitudes.

Another system built by Yuan and Songtao [15] depending on the non-linear relational band between step size and error, the new variable step-size least mean square (VSS-LMS) proposed and analyzes with different parameters. This brand new algorithm is overcome the step-size adjusting shortage of SVSLMS which is applied on the Sigmoid function, computer simulations and theoretical analysis show advancement with regard to convergence. Another research [16] had the same concern in a try to break the struggle between the time of convergence with the misadjustment of steady state by presenting a control system with a designed filter using an extended varying step-size algorithm, the step-size formula is modeled to improve the algorithm's ability of noise interference

resistance. Else more, a VSS-LMS was proposed [17] using squared error autocorrelation criterion to administrate step-size value variation, a smoothing function employed for such purpose. This approach employs two quantities of step size, a large one for fast convergent in transient state and a small one for low (MSD) level in steady state. The proposal's performances have shown through simulation outcomes illustrated the good performance of the design over other algorithms under the same conditional test. To optimize the issue of both convergence and low misadjustment problems, Rusu *et al.* [18] produced a varying step size-based algorithm to turn this misalignment to its minimum by optimizing both LMS besides NLMS control time-varying parameters. The theoretical outcomes have been confirmed by the output simulation and show well features of that algorithm.

In the frame of the issue, Huang *et al.* [19] produced a novel diffusion (DRVSS-LMS) algorithm which is insensitive to deal with impulsive noise in a network, the proposal is about assigning VSS aiming to improve estimation performance, the algorithm recorded an achievement through high convergence rate and low steady-state misadjustment better than existing diffusion LMS algorithms with impulsive noise. Reviews of seventeen different (VSS-LMS) algorithms introduced by Bismor *et al.* [20], their performances are compared within three various applications systems: System identification, noise removal, and line enhancement, it also presents a modification suggestion for step size that is a central parameter for several LMS algorithms. Surfing through the previous proposing and modifying algorithms to achieve fast convergence with low misadjustment then get the benefit of such improvement of filter performance to boost the efficiency of a specific system, a VSS algorithm is employed in this study to design finite impulse response (FIR) filter then employed in identification system to identify a BPF. A new adjusted VSS algorithm proposed by Jamel and Mohamed [21] called AAE-ASSLMS and employed it into ANC system to enhance the quality of noise cancellation regarding input speech signal, the same algorithm employed in this paper for AF design which is served in modeling system that used to identify a BPF with optimal fast conversion rate, low level of misadjustment, and minimum samples for identification.

3. ADAPTIVE MODELING SYSTEM

The main purpose of the modeling system is to identify unknown systems or plants, the identifier system designed in this paper is to identify a BPF. The system employs an

FIR filter serving the purpose efficiently, giving the best prediction of the unidentified system by providing a linear model. A general block diagram of system modeling is presented in Fig. 1 [1].

Parallel connection takes place gathering both the unidentified system (to recognized) with the designed FIR filter. A general input-output relationship is described by $x(n)$ and $d(n)$, respectively, to be an input to the indefinite system and its output. Besides the mysterious system, $x(n)$ feeds as input to the meant filter that is by its linear model provides the best description to the relationship of input with output to achieve the modeling by producing the nearest matching with the indefinite system output $d(n)$ through its output represented by the signal $y(n)$. On the same end, an error has been computed $e_1(n)$ showing the accuracy of identification by comparing the outputs of that filter in interest and the undefined system [1], [2]. In this work a BPF used as an unknown system to be identified by the required adaptive filter of the system model.

As believed by the major role of the filter adaptation in the identification system, an algorithm adaption should be chosen carefully to give the desired outcome. In this work, the FIR filter is designed with an adaptive plan as said by the structure presented in Fig. 2 [5], [6].

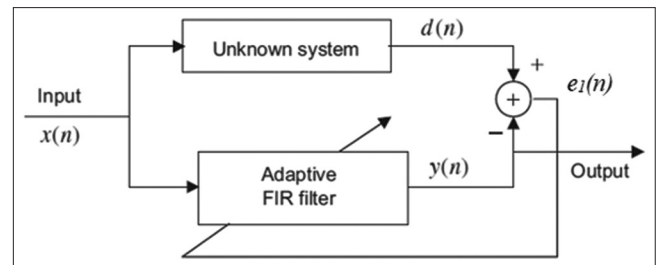


Fig. 1. Adaptive modeling system.

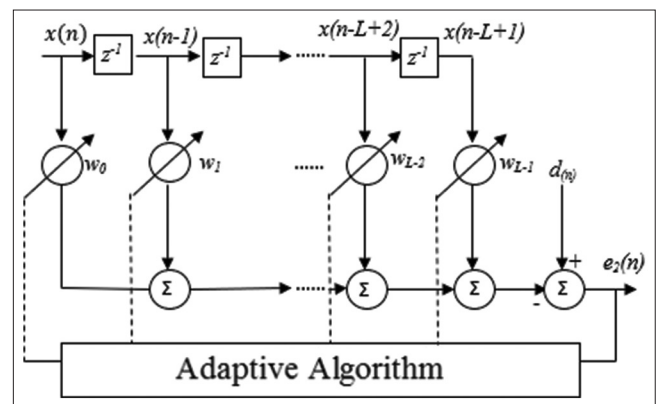


Fig. 2. Structure of adaptive finite impulse response filter.

As part of this research, additional noise (a white noise) has been supplied to the signal $x(n)$, then the input could be displayed as a row vector:

$X(n)=[x(n),x(n-1)\dots x(n-L+1)]$, n is considered as the time index, L is filter length.

Next, the filter weights $w(n)$ can be represented as follows:

$$W(n)=[w_0(n),w_1(n),w_2(n)\dots w_{L-1}(n)]^T$$

The input $X(n)$ is filtered into $e_2(n)$ as output that is produced accordingly by calculating the difference of the adaptive output with the wanted signal, $y(n)$ and $d(n)$, respectively.

The practical sharing piece of the adaptation used an algorithm that is to change the filter weights with length (L) in an iterative method, the weights are changed according to that algorithm every iteration.

This paper employs three different adaptive algorithms to design filter as adaptive with eight as a length to achieve the minimum output error $e_1(n)$ and $e_2(n)$ with the fastest convergence time and lower steady-state misadjustment.

For the mentioned specified modeling system, we are used AAE-ASSLMS adaptive algorithm. To the best of our knowledge, there exist no approaches that aim to get some conclusion or expose a performance result of using this algorithm. For that matter, we have used two other traditional algorithms to make comparisons of the results.

3.1. LMS Adaptive Algorithm

LMS is considered a familiar algorithm that is broadly used because of simple computations and implementation. The filter weights are changed adaptively in harmony with the algorithm at each iteration following the formula [12]:

$$W_{(n+1)} = W_{(n)} + 2\mu e_{(n)} X_{(n)} \quad (1)$$

The step-size parameter μ is the lead player in this formula. To achieve the choicest level of convergence, μ is restricted as a small positive value. The step-size value is still fixed for every iteration despite the weights changing which is considered later as a disability in that algorithm with regard to the misadjustment and time convergence balancing issue. With eight weights to be adjusted and a suitable step-size value, FIR AF was successfully designed and implemented in an identification system to identify a BPF.

3.2. NLMS Adaptive Algorithm

NLMS is considered as the first development or expansion of the algorithm of LMS; practically, both algorithms' implementation is alike. The main intention of this extension is to bypass the convergence issue by making the step-size value depending on the recent input signal values with each iteration, in another word, μ is invulnerable to the changes of that input; therefore, the readjustment is not a demand when such changes occur. As the following equation present [20]:

$$\mu(n) = \frac{\mu}{x^T(n)x(n)} \quad (2)$$

Then, the calculated $\mu(n)$ is applied in Eq. 1 for weight adjustment of filter weights. With identical filter length, the adaptation is applied, the outputs fully will be displayed in the result section.

3.3. AAE-ASSLMS Adaptive Algorithm

Absolute average error-variable step-size LMS (AAE-ASSLMS) is an adjusted method of that algorithm of LMS. A VSS is employed in this algorithm which is adjusted in each iteration along with the filter taps (weights). The adjustment of the VSS is established in means of error relying on the recent and previous and calculating the absolute average value as follows [21]:

$$\mu(n+1) = \mu(n) - \left| \sum_{c=0}^L e(n-c) * \beta \right| \quad (3)$$

Where $0 < \beta < 1$

And

$$\mu(n+1) = \begin{cases} \mu_{min}, & \text{if } \mu(n+1) < \mu_{min} \\ \mu(n+1) & \text{otherwise} \end{cases} \quad (4)$$

This algorithm efficiently contributes to convergence enhancement by starting with a high value of step-size progressively walking through a reduction process as Eq. 3 illustrates. The current step size $\mu(n)$ will be changing accordingly into $\mu(n+1)$ with dependency on the difference of the absolute average value of the recent and previous error until $\mu(n)$ will reach its minimum in the order of low level of misadjustment achievement and the convergence time speed is acceptable and improved through this transition phase smoothly. By this process, $\mu(n+1)$ will maintain a value less than $\mu(n)$ to maintain the optimum system operation as displayed in Eq. 4, also, the equation clarifies that the adapted step size always retains to μ_{min} to prevent $\mu(n+1)$ from going under the bounded level for the filter stability reservation.

The (β) parameter from Eq. 3 is a factor used to keep convergence time in the phase of the steady state under control while the misadjustment level is controlled by preserving the step-size value by μ_{min} . The value β has chosen in between of $0 < \beta < 1$ [21].

The filter weights are now updated based on the following modified formula in preparation for the upcoming iteration.

$$W_{(n+1)} = W_{(n)} + 2\mu_{(n)} e_{(n)}^2 X_{(n)} \quad (5)$$

The convergence factor μ is modified into $\mu(n)$ which can be determined as in Eq. 3 with each iteration will be changed until the algorithm would reach the steady state then the adaptive value reaches its minimum and is reserved at that value to concur with the optimum system operation and maintain its stability with the lowest level of misadjustment.

4. EXPERIMENTAL RESULTS

An adaptive system modeling employing an adaptive FIR filter proposed in this work, designed, and implemented using MATLAB, the AF is designed using three different adaptive algorithms, Fixed step-size LMS, variable step-size NLMS, and adjusted step-size AAE-ASSLMS. The approach parameters are used the same for all the three adaptive algorithms of equal comparison among the output results.

The FIR filter is designed with a length of eight taps, the input for both the indefinite system and the adapted filter is a companion with a white noise signal having zero mean and a unity variance with the main input of SNR (0 dB).

Step-size value selection is set accordingly with the performance and the simulation results examination of the adaptive algorithm AAE-ASSLMS [21]. Of comparison of system performance using the three different adaptive algorithms, the same step value is employed for the other algorithms also.

The filter taps are chosen randomly to show the power of the process of adaptation. The adaptation performance is measured on calculating the normalized difference within the actual taps and the updated ones with a step size equal to 0.05. For a purpose of evaluation, the efficiency of each algorithm and for comparison later, a weight difference (WD) is calculated between the actual and adaptive taps of the filter, as a preparation step for MSE calculation which is aimed at accurate measurement of the resultant output. MSE is calculated for the $e_2(n)$ value in Fig. 2 and according to that

value, a comparison will take place between the algorithms for the optimum accuracy to the system performance achievement.

4.1. Simulation of AF

This section reveals the performance evaluation of all three adaptive algorithms that have been employed in the design filter. Finally, the section is concluded by comparing these performance results to illustrate the quality of the proposed model.

4.1.1. LMS performance evaluation

For LMS, step size is fixed and the updated weights are calculated according to Eq. 1. The results are shown in Fig. 3 which demonstrates the outcome of the applying LMS algorithm for both MSE and WD which are scaled according to the output from -20 to 10 dB and -15 to 5 dB, respectively. The outcomes from the initial point are at their peak, then after around 50 iterations, the value of MSE and WD decreases sharply by around 20 and 15 dB, respectively, until gets to their steady-state phase along with the 1000 iterations with fluctuation values -8 (± 5) dB for MSE and -9 (± 2) for DW, and the recorded minimum values are -13 and -11 dB, respectively. On the other hand, Fig. 4 shows differences between the actual and the evaluated adapted taps, it highlights the performance of the AF in terms of weight adaptation, the figure illustrates how the AF gives weights' values so close to the actual one, while the calculated error is presented with details in Fig. 4 which is the main factor used in the adaptation process.

4.1.2. NLMS performance evaluation

NLMS step size is variable, with regard to this algorithm, the filter taps are updated and weights have gotten from Eq. 2, with the same main parameters, the MSE of output along with WD is illustrated in Fig. 5. Indeed, the figure depicts the minimum recorded values than LMS for both MSE and WD which are -17 and -14 , respectively, and the AF reached its steady-state phase with less number of iterations

NLMS results show obvious improvement over LMS results in both MSE and WD, as a consequence of the adaptive formula for step-size value changing, it was the first step for upcoming formulas of system output enhancement.

4.1.3. AAE-ASS performance evaluation

The step size in the AAE-ASS algorithm is adjusted with each iteration following Eq. 3 to give an output represented with both MSE and WD illustrated in Fig. 6. The optimum value of μ_{max} and μ_{min} for the AAE-ASSLMS was chosen to be 0.05 and 0.0005, respectively. The constant β is chosen to be 0.00018 according to both performance analyses.

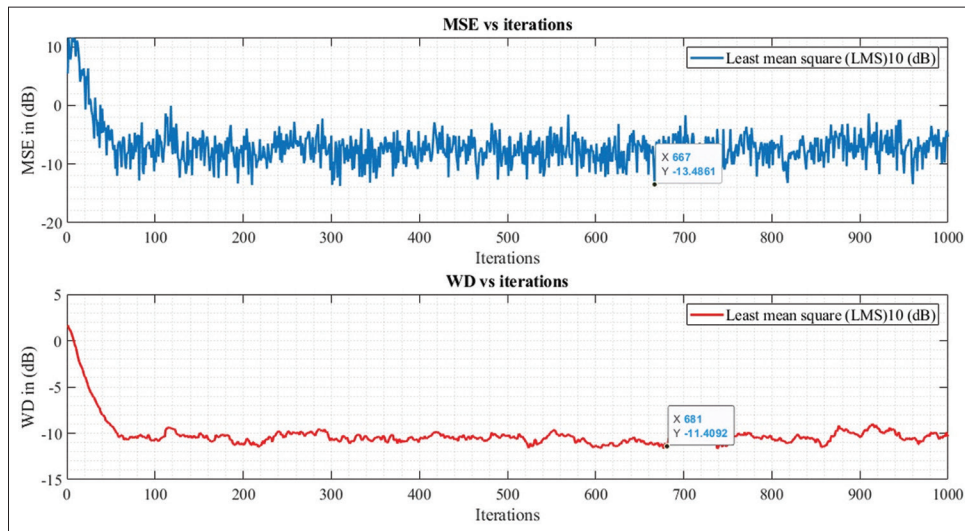


Fig. 3. Least mean square adaptive algorithm results.

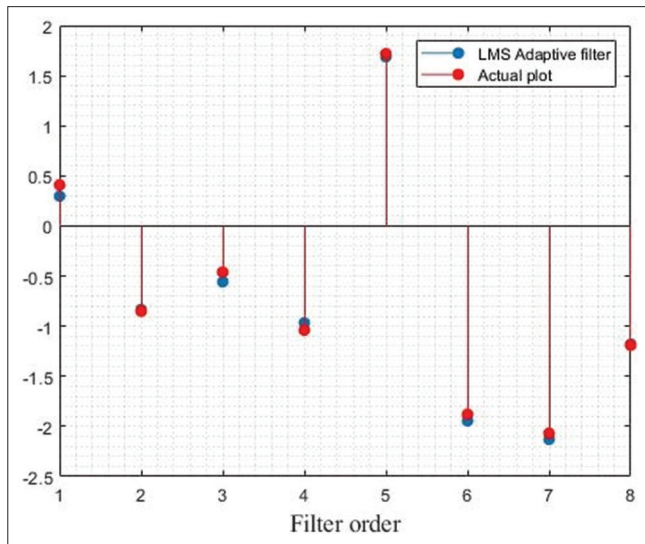


Fig. 4. The actual values of coefficients.

Fig. 6 illustrates the implementation output of the algorithm. Along 1000 iterations, the minimum and maximum recorded MSE are 8 and -46 , respectively. In addition to WD values, the recorded output is -8 dB as the initial point then smartly drops to its steady state with -24 dB as its minimum value with noticeable improvement of reaching steady-state phase with a minimum number of iterations.

4.1.4. Comparison result analysis

From Figs. 3 to Fig. 7, minimum MSE has been recorded for each algorithm beside WD as well, Table 1 represents those records for comparison about the accuracy, speed of convergence, and misadjustment. Fig. 4 represents a chart

to illustrate the differences in performance among the used algorithms that mentioned in Table 1.

NLMS algorithm has the advantage over the LMS in the means of MSE and weight difference, while the AAE-ASSLMS algorithm achieves the greatest results out of them together. A performance summary for all the algorithms is displayed in Table 1 and illustrated in Fig. 7 for more clarity. Besides, the AAE-ASSLMS algorithm attains the best convergence time among the further algorithms as Fig. 6 illustrates. Furthermore, Fig. 8 shows the three algorithms' performances together, the progress of the AAE-ASSLMS algorithm is noticeable over the further algorithms with regard to convergence speed and steady-state misadjustment.

For further elaboration, the MSE results of both LMS and AAE-ASSLMS algorithms were combined into a singular frame as they are shown in Fig. 9. The MSE in dB for the AAE-ASSLMS algorithm is between around 5 and -20 from the beginning and it has decreased to 0—45, while the MSE for LMS is higher accordingly by around 10 dB along with the iterations.

Despite NLMS algorithm progress over LMS, it was not noticeable to be added to Fig. 9 comparing to the achievement of AAE-ASSLMS algorithms over them both.

4.2. Simulation of Adaptive System Identification through a BPF

In this part, the MATLAB program is employed to show the filter adaption and can keep tracking the behavior of the mysterious system which is set to be a BPF whose cutoff

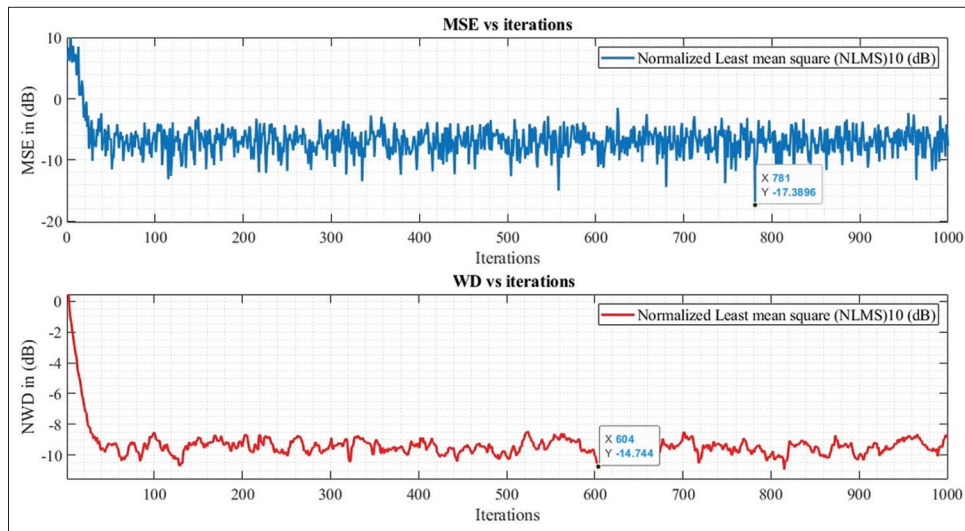


Fig. 5. Normalized least mean square adaptive algorithm results.

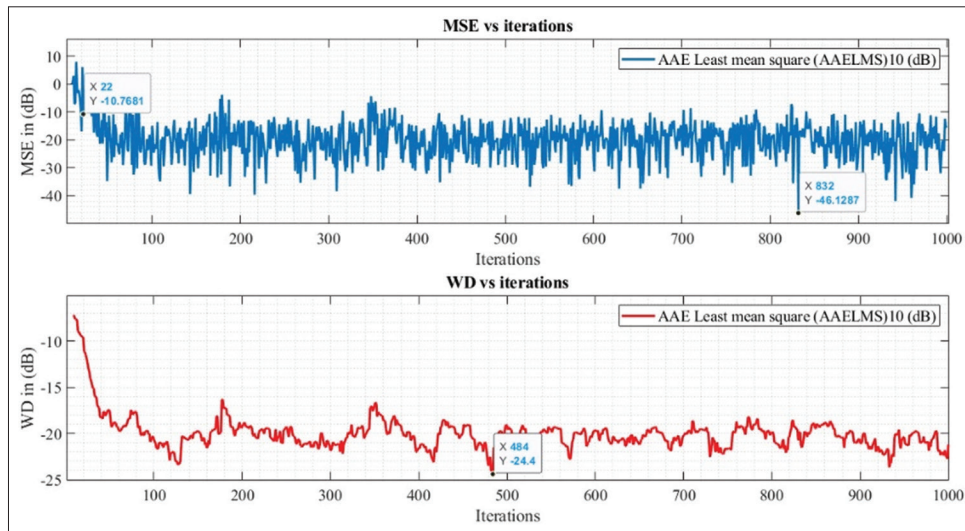


Fig. 6. Absolute average error-based adjusted step-size least mean square adaptive algorithm results.

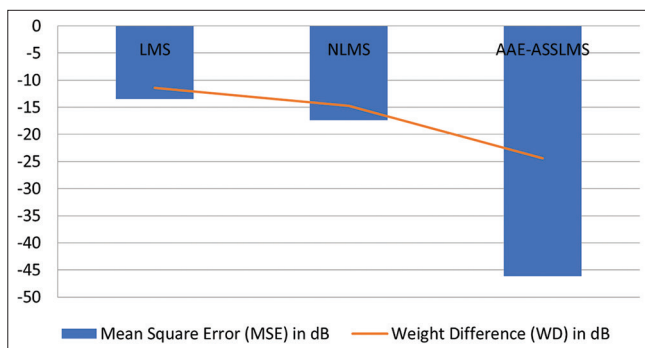


Fig. 7. Summary of adaptive algorithms performance.

frequencies are 1500 Hz and 1700 Hz operating at 8000 Hz. The input consists of tones of 600, 1600, and 2600 Hz. The input with its spectrum is presented in Figs. 10 and 11.

The assumed BPF frequency response is displayed in Fig. 12.

Predictably, the output response of the system should be 1600 Hz tone only since the other tones are rejected by the filter. However, the designed filter of this work along with the three different mentioned algorithms should follow the BPF output. Fig. 13 depicts the frequency domain comparisons; it

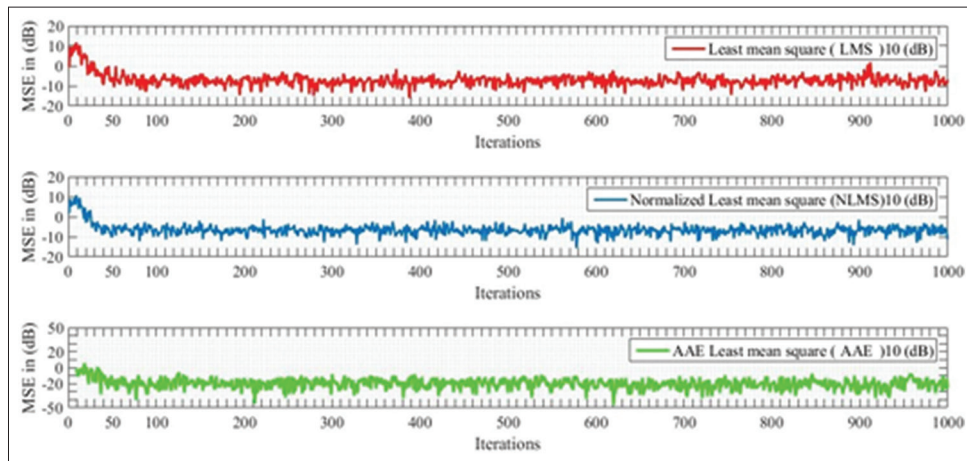


Fig. 8. Algorithms' performance.

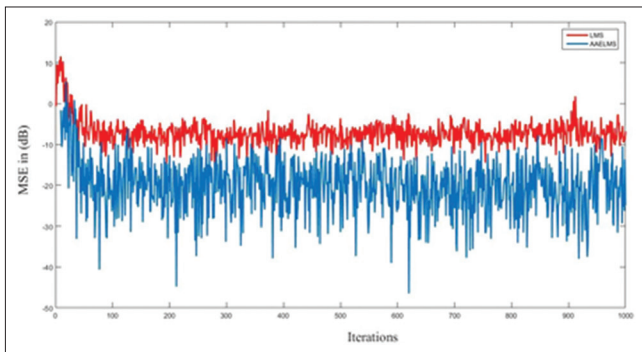


Fig. 9. Mean square error output of least mean square (LMS) and absolute average error-based adjusted step-size LMS algorithms.

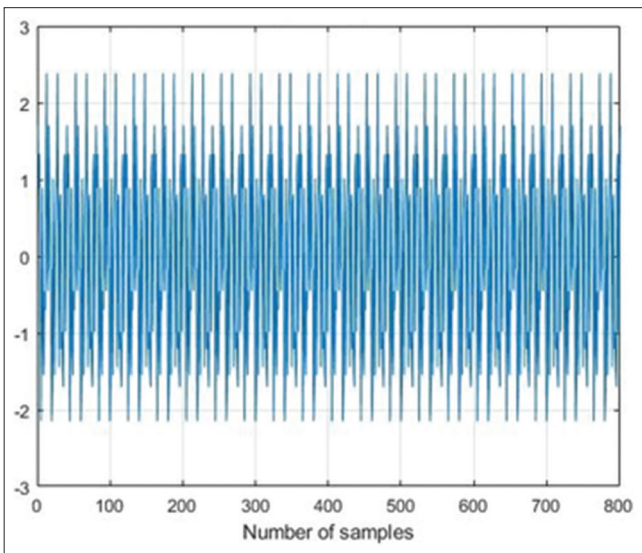


Fig. 10. The input signal to the system.

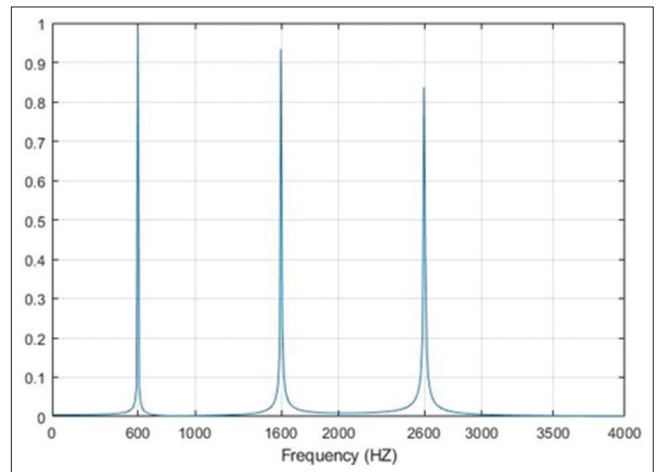


Fig. 11. The spectrum of the input signal.

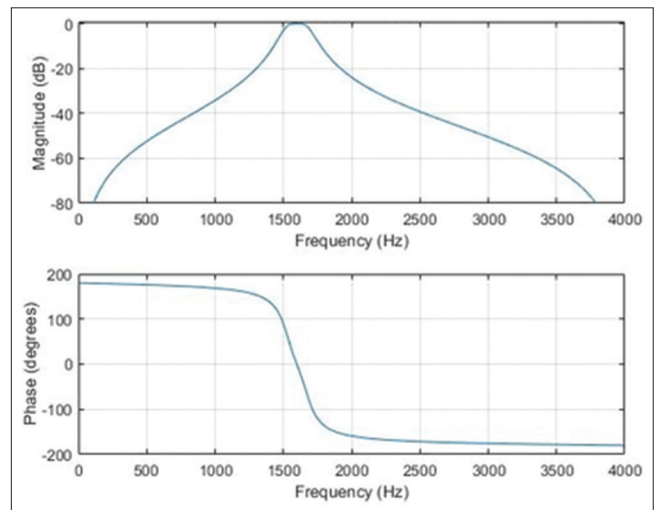


Fig. 12. The frequency response of the unknown system (bandpass filter).

also reveals the matching spectrum output signal for the BPF and the filter adapted using the AAE-ASSLMS algorithm the

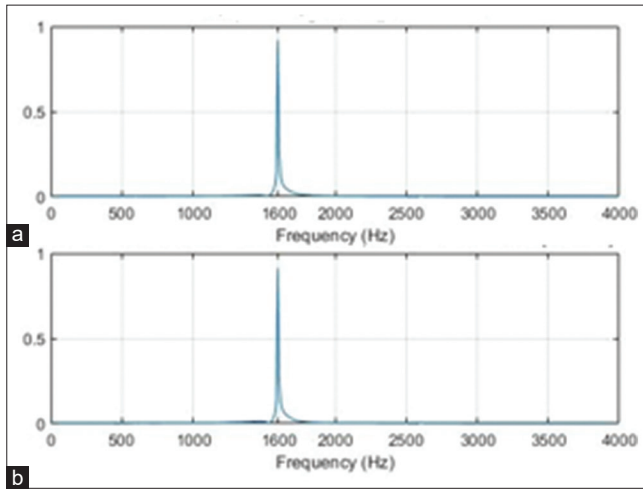


Fig. 13. The spectrum of the output signals for (a) the unknown system (bandpass filter) and (b) the adaptive filter (absolute average error-based adjusted step-size least mean square algorithm).

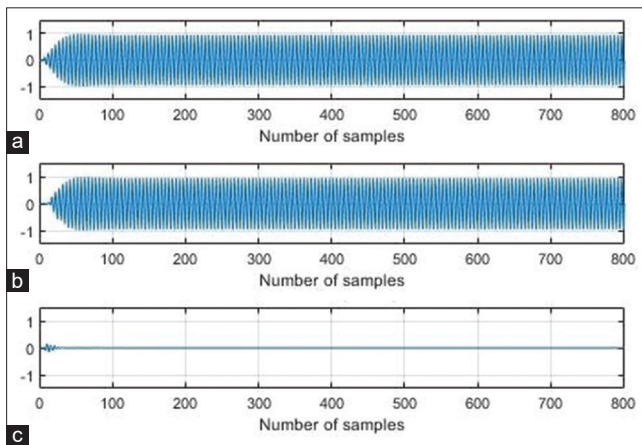


Fig. 14. The output signal of (a) the unknown system (bandpass filter), (b) the adaptive filter (absolute average error-based adjusted step-size least mean square algorithm), and (c) the error signal $e_f(n)$.

output signal spectrum of the three algorithms is identical to 1600 Hz.

Despite the matching spectrums, although an error accrued and calculated ($e_f(n)$ in Fig. 1). Fig. 14 shows the total output signal with the error calculated and presented through the difference between BPF and the design filter output with the (AAE-ASSLMS algorithm) algorithm.

The calculated error for each used algorithm in this study as Fig. 15 illustrated.

Through the simulated output results, it is noticeable that resultant error from the AAE-ASSLMS algorithm is

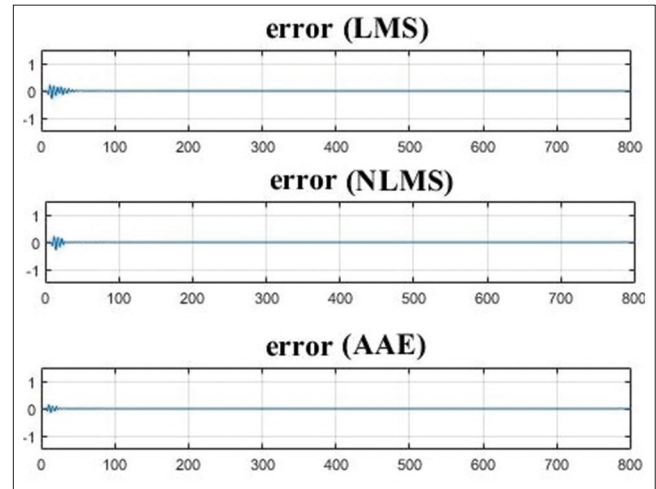


Fig. 15. The error for each algorithm.

TABLE 1: Summary of adaptive algorithms performance

Algorithm	Iteration	Filter order	Mean square error in dB	Weight difference in dB
LMS	1000	8	-13.4861	-11.4092
NLMS	1000	8	-17.3896	-14.744
AAE-ASSLMS	1000	8	-46.1287	-24.4

the minimum in the mean of peak-to-peak value and the system identification time is the shortest among the other algorithms in means of the samples number. In another word, the designed filter identified the indefinite system with <10 samples, meanwhile, other algorithms were taking more than 20 samples for identification.

5. CONCLUSIONS

Two significant parameters that are directly impacted by the step size in VSS algorithms, there are the low-level misadjustment in the steady-state actions and the convergence time speed in the transient actions. The relation between step-size, MSE, and the convergence speed is difficult to find a middle ground to be operated on because of the direct and inverse proportions among them (direct with MSE, inverse with convergence speed). A VSS algorithm is employed within this paper called the absolute average error adjusted step-size LMS algorithm (AAE-ASSLMS) to design an adaptive FIR filter along with LMS and NLMS algorithms. The adaptation step-size value is adjusted based on the absolute and average value for the recently assessed error and the prior one, the adjusting process in the algorithm takes place through

maximum down to minimum step-size selection for optimum achievement in means of fast convergence time and the lowest level of misadjustment. Throughout the simulated results, the algorithms' output shows improvement in those parameters compared with fixed step-size LMS and its modified version (NLMS). Within this work, the designed filter with the AAE-ASSLMS algorithm is employed for the 1st time in a modeling system for higher accuracy of identification. Furthermore, to prove the system efficiency improvement, a BPF is employed in this paper as a system to be identified, successfully, the identification process achieved using the designed adapted FIR filter through the different used algorithms. Again, the outputs of the simulation of the (AAE-ASSLMS) algorithm show improvement in the performance with regard to error and identification time in means of samples numbers over the other algorithms that are used in adaptation within the same parameters.

REFERENCES

- [1] C. Heij, C. M. A. Ran and F. van Schagen. "Introduction to mathematical systems." In: *System Identification*. Birkhäuser, Cham, pp. 137-155, 2021.
- [2] H. Jianwang, R. A. Ramirez-Mendoza, and X. Yan. "Statistical inference for piecewise affine system identification". *Mathematical Problems in Engineering*, vol. 2021, p. 4618030, 2021.
- [3] T. Lee, B. D. Lee, and F. C. Park. "Optimal excitation trajectories for mechanical systems identification". *Automatica*, vol. 131, p. 109773, 2021.
- [4] M. N. Taib and N. Ismail. "System identification makes sense of complex measurements." In: *11th IEEE International Conference on Control System. Computing and Engineering (ICCSCE)*. IEEE, Penang, Malaysia, pp. 240-245, 2021.
- [5] A. Ankur and D. Devi. "Improving Performance of Hardware Adaptive Filter". *Journal of the University of Shanghai for Science and Technology*, vol. 23, pp. 742-745, 2021.
- [6] A. Singh, R. Pandya, A. Sharma, and T. Singh. "Effect of Learning Rates on Adaptive Filter Algorithms and Implementation of Kernel Adaptive Filters". *International Journal of Creative Research Thoughts*, vol. 9, no. 4, pp. 341-350, 2021.
- [7] B. Shen, X. Lv and S. Zhang. "An improved LMS adaptive filtering algorithm and its analysis." In: *2019 International Conference on Intelligent Computing, Automation and Systems (ICICAS)*. IEEE Computer Society, China, pp. 549-551, 2019.
- [8] S. Putluri and M. Zia Ur Rahman. "Novel exon predictors using variable step size adaptive algorithms." In: *Innovative Data Communication Technologies and Application*. Springer Nature, Germany, pp. 750-759, 2020.
- [9] P. Lara, F. Igreja, L. D. T. Tarrataca, D. B. Haddad and M. R. Petraglia. "Exact expectation evaluation and design of variable step-size adaptive algorithms". *IEEE Signal Processing Letters*, vol. 26, no. 1, pp. 74-78, 2019.
- [10] A. Linovich, V. Litvinova and M. Korolev. "Comb adaptive filtering algorithm". *Vestnik of Ryazan State Radio Engineering University*, vol. 77, pp. 3-16, 2021.
- [11] D. Kari, A. H. Mirza, F. Khan, H. Ozkan and S. S. Kozat. "Boosted adaptive filters". *Digital Signal Processing*, vol. 81, pp. 61-78, 2018.
- [12] R. Claser and V. H. Nascimento. "On the tracking performance of adaptive filters and their combinations". *IEEE Transactions on Signal Processing*, vol. 69, pp. 3104-3116, 2021.
- [13] K. Si-Min, W. Yue-Long, S. Tao, G. Yi, and H. Yan. "LMS adaptive filtering of drilling tool vibration signal". *Journal of Physics: Conference Series*, vol. 1237, p. 042018, 2019.
- [14] T. T. P. Silva, F. Igreja, P. Lara, L. Tarrataca, A. Kar and D. B. Haddad. "On the Skewness of the LMS adaptive weights". *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 8, pp. 3022-3026, 2021.
- [15] Z. Yuan and X. Songtao. "New LMS adaptive filtering algorithm with variable step size." In: *2017 International Conference on Vision, Image and Signal Processing (ICVISIP)*. Conal, Osaka, Japan, pp. 1-4, 2017.
- [16] Q. Niu and T. Chen. "A new variable step-size LMS adaptive algorithm." In: *2018 Chinese Control and Decision Conference*. IEEE, United States, pp. 1-4, 2018.
- [17] W. Loedwassana. "A variable step size algorithm of LMS algorithm based on squared autocorrelation criterion." In: *2019 7th International Electrical Engineering Congress (IEECON)*. IEEE, United States, pp. 1-4, 2019.
- [18] A. Rusu, S. Ciocinã and C. Paleologu. "On the step-size optimization of the LMS algorithm." In: *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, United States, pp. 168-173, 2019.
- [19] W. Huang, L. Li, Q. Li and X. Yao. "Diffusion robust variable step-size LMS algorithm over distributed networks". *IEEE Access*, vol. 6, pp. 47511-47520, 2018.
- [20] D. Bismor, K. Czyz and Z. Ogonowski. "Review and comparison of variable step-size LMS algorithms". *International Journal of Acoustics and Vibration*, vol. 21, p. 1392, 2016.
- [21] T. Jamel and H. Mohamed. "Noise canceller using a new modified adaptive step size LMS algorithm". *WSEAS Transactions on Signal Processing*, vol. 10, pp. 637-644, 2014.

Future IoT Software in Healthcare Also Exploring IoT Industry Application



Mustafa N. Rashad¹, Dana L. Hussein¹, Haval D. Abdalkarim², Ribwar R. Azeez²

¹Department of Information Technology, Chamchamal Technical Institute, Sulaimani Polytechnic University, Kurdistan Region, Iraq, ²Department of Database Technology, Computer Science Institute, Sulaimani Polytechnic University, Kurdistan Region, Iraq

ABSTRACT

There has been a great deal of investigation into medical services ability and specialized advancements during the most recent 10 years. To state the obvious, Internet of Things (IoT) has demonstrated insure associating different clinical hardware, sensors, and medical services experts to give top-notch clinical consideration at a distant area. This has upgraded patient security, diminished medical care costs, expanded admittance to medical services benefits, and expanded functional adequacy in the medical care industry. Emerging technologies such as IoT have the potential to transform our lives in many ways. A smart ubiquitous framework can only be built using smart objects in the IoT system, which is its ultimate building pieces. This research surrenders an audit of potential IoT-based innovation applications in medical services conducted to date. This paper records the development of the use of the Healthcare Internet of Things (HIoT) in tending to different medical care worries according to the viewpoints of empowering innovation, medical care administrations, and applications. Besides, potential HIoT framework issues and issues are explored. The current research closes by giving a wellspring of comprehension on the various uses of HIoT with expectations of empowering future scholastics that are quick to chip away at and kick off something new in the field to have a superior handle of the subject. IoT innovation has helped medical care experts in checking and diagnosing an assortment of well-being concerns, estimating an assortment of well-being factors, and giving demonstrative capacities at far-off areas using these standards. The structure and implementation of a specific framework are the subject of this paper. This has moved the medical services industry's concentrate away from clinics and toward patients.

Index Terms: Future IoT Applications, Healthcare Internet of Things, Medical Care, Industrial Applications

1. INTRODUCTION

In current history, medical care industry has encountered rapid development, contributing impressively to income and business creation. A couple of years prior, sicknesses, and inconsistencies in human body must be recognized through

a clinical appraisal [1]. Most of patients ought to visit the emergency clinic for the span of their treatment, which brought about more prominent medical care costs and tension on country and distant medical services organizations [2]. Over the long haul, mechanical progressions have considered the determination of an assortment of problems just as well-being observing by means of minuscule gadgets, for example, smartwatches [3]. Despite of this, innovation has changed the medical care framework from being focused on clinics to being fixated on patients [4], [5]. A few clinical preliminaries, for instance, may be done at home without the assistance of a medical care proficient (e.g., blood glucose level, checking pulse, pO₂, and level) [4], [6], [7].

Access this article online

DOI: 10.21928/uhdjst.v6n1y2022.pp70-75

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Copyright © 2022 Rashad *et al.* This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Mustafa N. Rashad, Department of Information Technology, Chamchamal Technical Institute, Sulaimani Polytechnic University, Kurdistan Region, Iraq. E-mail: Mustafa.rashad@spu.edu.iq

Received: 05-12-2021

Accepted: 07-04-2022

Published: 19-05-2022

The current interchanges innovation may likewise be utilized to send clinical information from remote spots to medical services habitats [5], [8]. The IoT has turned into a vital supporter of worldwide correspondence on account of future conventions and calculations [8]. The IoT has given individuals more autonomy while likewise expanding their capacity to cooperate with the remainder of the world [9]. It associates a wide scope of gadgets to the Internet, including remote sensors, home apparatuses, and electrical gadgets [10]. IoT applications might be found in horticulture, vehicles, the home, and medical services [11]. The IoTs is gathering steam due to its advantages of higher precision, lower costs, and the capacity to more readily anticipate future occasions [12]. Moreover, more prominent information on programming and applications, just as headways in versatile and PC advancements, universal openness of remote innovation, and the development of the computerized economy, have all upheld the fast IoT insurgency [13]. Heating, dampness, electrocardiograph (ECG), electroencephalograph (EEG), and other physiological data from the patient's body are captured using instruments in healthcare applications. Climate, dampness, date, and time may all be recorded as well [5], [14]. These data allow for more precise and relevant conclusions about the patients' health [4], [15]. Because a vast quantity of information is obtained derived from a range of sources, data storage and usability are also crucial in the IoT system. Doctors, caretakers, and other authorized persons have use of obtained information by aforementioned sensing devices. Having the ability to communicate this data relating to healthcare practitioners through cloud-server enables for fast patients' diagnoses and, if necessary, actions in the field of medicine [16]. The IoTs is an idea wherein every one of the gadgets in our day-to-day existence might associate with the web or to one another to communicate the information and execute all positions through the organization [10]. The improvement of new provisions in the manner medical care products is offered has come about due to admittance to versatile clinical clients and portable well-being administrations. Likewise, the development of therapy draws near, just as the spread of arising innovation like robots and man-made brainpower, just as the straightforward exchange, and sharing of clinical information over the web [10], [11]. It is a remarkable idea which spins around utilizing the Internet to better our lives [13]. This fundamental pattern is valuable to patient consideration since it permits specialists to make more exact analyses and, accordingly, accomplishes better treatment results. The utilization of IoT highlights in clinical gear generously expanded the quality and viability of clinical benefits [17].

2. BACKGROUND

IoT is a significant a component of current data innovation. IoT is a framework that spreads over the Internet and is the consequence of ongoing quick development in the field of remote interchanges [10]. To make the "Web of Everything" a reality, it very well might be important to associate different information gathering devices to the Internet [18]. Keen urban communities, shrewd homes, sensors were set, and route frameworks are only a couple of the spaces, where the IoTs is as of now broadly utilized [4]. Perhaps the main application area of all is shrewd well-being [5], [8], [16]. Great many Each year, individuals pass away as a result of different illnesses or medical problems [16]. Individuals' well-being is turning out to be increasingly more of a concern. Subsequently, one of the focal points of study in the field of shrewd well-being is the utilization of IoT advancements to address medical problems [18]. It is the organization with which it communicates the physical, virtual universes of the Internet. The actual world incorporates home devices, autos, modern hardware, structures, clinical gear, and the human body [10]. Individuals' way of life, ongoing sickness the executives, peril ID, and life-saving treatments will all profit from the utilization of IoT innovation in medical services [5]. In medical care, the IoT has a variety of applications: Keep a nearby eye on your well-being. Wearable gadgets would now be able to follow essential human body capacities, examine human conduct, and analyze medical conditions [8]. Wearable innovation devices (smartwatches) can lessen tension and set aside cash for patients [14]. One more delicate well-being observing innovation used in customary emergency clinics isn't care for this. Patient support in well-being related exploration [16]. IoT gadgets might be utilized in medical services settings to remind patients to take as much time as necessary [5], [10]. Electrocardiograms, blood oxygen, and circulatory strain checking hardware can be interconnected to work on patients' and parental figures arranged, observing, and framework expected, prompting further developed treatment results and service enhancements [5]. Autos can be associated with network frameworks utilizing the IoTs. If an auto is occupied with a mishap, the framework can survey the seriousness of the impact and inform the traffic right hand director and the medical care mediation focal point of the mishap area and heading. This will help individuals who have been harmed in looking for brief assistance [10]. A few examination utilizing smartwatches, shrewd wearables, keen wristbands, keen homes, and IoT innovations are presently embraced in the field of brilliant well-being. By the by, no exploration has equitably investigated and pictured all of the writing on the theme and analyze the current state and future

advances of IoT-based wise well-being research, totally, and dispassionately [5].

3. SYSTEM ARCHITECTURE OVERVIEW

3.1. Basics of HIOT Architecture

The IoT worldview for clinical applications supports the coordination of IoT and distributed computing benefits into the field of medication. It likewise gives the conventions to conveying patient information to a clinical association from a scope of sensors and analytic supplies. The geography of HIoT alludes to the association of various parts of an IoT medical care system that is successfully incorporated into medical services setting [19]. The three vital parts of a run of the mill HIoT framework are the maker, representative, and supporter, as shown in Fig. 1. The distributor addresses an organization of connected sensors and other clinical gadgets that may freely or at the same time accumulate the patient's fundamental information. The factors that might be estimated incorporate pulse, temperature, oxygen immersion, ECG, EEG, and EMG. The distributor can send this information to a specialist routinely over an organization [20]. The information that has been gained in the cloud is handled and put away by the representative. At long last, the endorser participates in constant checking of the patient's information, which might be gotten to and seen on a cell phone or PC. The distributor can analyze the

information and proposition criticism in the wake of seeing any physiological irregularities or weakening in the patient's well-being status [1]. Every hub on the IoT organization and server in the medical care network fills a particular need in the HIoT, which unions separate parts into a crossover matrix. Since the geography is relying upon the medical care prerequisite and application, it's hard to offer a uniform establishment for HIoT [12]. For the HIoT framework, a few underlying modifications have been executed before. When fabricating another IoT-based medical care framework for ongoing patient checking, it's basic to make a rundown of all connected activities identified with the planned well-being application [12], [14]. The accomplishment of the IoT framework is controlled by how well it addresses the issues of medical care experts [4]. The geography should follow the clinical standards and stages in the analysis technique since every infection requires a muddled succession of medical care activities [15].

3.2. HIOT Technologies

The innovation important to construct a HIoT framework is critical [1]. This is on the grounds that the arrangement of explicit advancements can upgrade the abilities of an IoT framework [17]. The different state of the art advancements has been joined with an IoT framework to incorporate assorted medical care applications [10]. The three fundamental categorizations where these advances fall are as follows [1]:

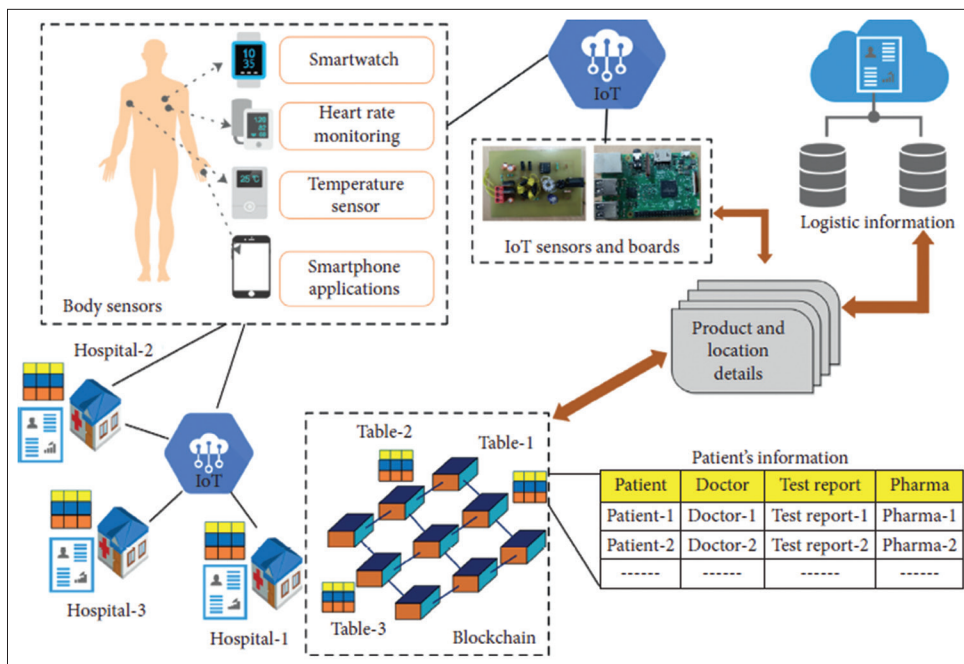


Fig. 1. A typical HIoT framework has three main components [1].

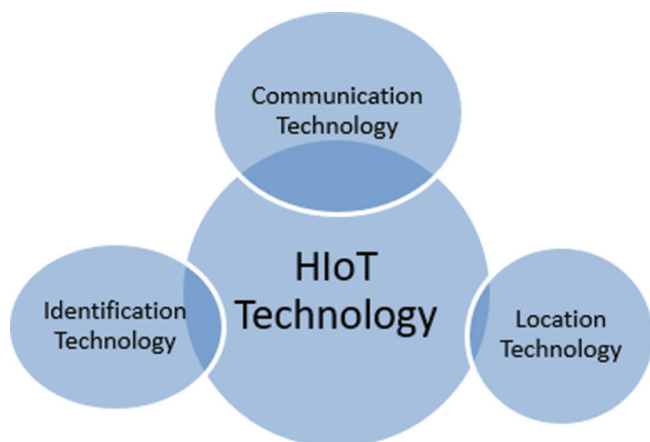


Fig. 2. Categorization of IoT technology.

1. Location technology
2. Identification technology
3. Communication technology, as shown in Fig. 2.

3.2.1. Authentication technology

The availability of the information about the patient derived from endorsed hub sensor device, and possibly situated in remote locations, is a commonsense factor in the plan of a IIoT framework [14]. This can be refined by accurately distinguishing the hubs and sensors that exist in the medical care organization [4], [10]. The act of allocating an interesting personality (ID) to each allowed substance, so it very well may be effectively recognized and steady information transmission can be cultivated, is known as distinguishing proof [1]. A computerized ID is associated with each asset engaged with the medical services framework (emergency clinic, specialist, attendants, cares, clinical gadgets, etc.) [7]. Be that as it may, due to the consistent headway of IoT-based advances, the exceptional character of a part might fluctuate after some time all over the IoT framework's life cycle [11]. To guarantee the uprightness of the medical care gadget/framework, the gadget should offer an arrangement for refreshing patients' information [1], [11]. This is because the arrangement alteration not just influences the most common way of following the organization component(s); however, it might likewise bring about a broken conclusion [11].

3.2.2. Telecommunication technology

Correspondence advancements permit different substances in a IIoT organization to speak with each other [9]. It can be parted into two classes: Short-reach and medium-range correspondence [10]. The conventions used to interface things inside a little scope of body region networks are known as short-range correspondence advancements (BAN) [6], [8]. Medium-

range correspondence frameworks, then again, ordinarily give significant distance correspondence, for example, data trade between a base station and a BAN's focal hub [8], [9], [21]. Zigbee is a typical convention for interconnecting clinical hardware and communicating information [9], [11]. The Zigbee recurrence range is like that of Bluetooth (2.4 GHz) [11], [12]. It does, nonetheless, have a more drawn-out correspondence range than Bluetooth gadgets [6]. The lattice network geography is utilized in this innovation [5]. End hubs, switches, and a handling place make up the framework [22]. Information examination and conglomeration are taken care of by the handling community, regardless of whether a couple of gadgets fall flat, and the cross-section network guarantees that the remainder of the gadgets stays associated. Energy utilization, high transmission rate, and huge organization limit are large benefits of Zigbee [11], [12].

3.2.3. Geolocation technology

The geolocation advance is regularly utilized in medical services organizations to screen and find the whereabouts of an article [6]. It likewise monitors the treatment cycle dependent on how accessible assets are dispersed [10]. The Global Positioning System, otherwise called GPS, is quite possibly the majority typically applied technology [8]. Satellites are utilized for the following purposes [1]. As since a long time ago, there is an unmistakable view between the item and four separate satellites, an article can be identified utilizing GPS [8]. It very well may be utilized in IIoT to distinguish the area of an emergency vehicle, a medical care proficient, cares. Utilization of Global Positioning Systems are, notwithstanding, confined for outside apps since neighboring offices can meddle with correspondence between the item and the satellite. A neighborhood situating framework (LPS) organization can be valuable in these circumstances [1]. LPS can follow an article by recognizing the radio transmission sent by the moving item and sending it to a variety of pre-situated beneficiaries [1]. LPS can likewise be used with an assortment of short-range correspondence innovations such as RFID, Wi-Fi, Zigbee. Super wideband (UWB) radio, then again, is inclined toward due to its predominant transient goal. This permits the collector to work out the appearance time with accuracy [7], [8], [13]. The analysts utilized a UWB-based confinement technique to follow the time distinction of appearance (TDOA). Other estimating rules, such as a family member and differential season of appearance, and full circle beginning speed, have been archived in the writing while developing a UWB-based restriction framework. GPS, just as other high-transmission capacity correspondence innovations, could be utilized to build savvy medical care networks later on [1], [8].

3.2.4. Services and Application of HIoT

Clinical contraptions would now be able to do ongoing investigations that specialists could not direct only a couple of years prior in light of late headways in IoT innovation [8]. It has likewise helped medical services communities connect with more people on the double and give great consideration for a minimal price. The utilization of huge information and distributed computing has fundamentally worked on the dependability and simplicity of correspondence among patients and specialists [11]. Therefore, the patient's commitment to the treatment interaction was expanded, while the patient's monetary weight was diminished [14]. The huge impact of IoT lately has helped the formation of HIoT applications, which incorporate sickness diagnostics, individual consideration for pediatrics and geriatric patients, well-being, and wellness of the executives, and constant illness checking [5], [8], [16]. It has been isolated into two key classes, to be specific administrations and applications, for a superior comprehension of these applications [16]. The previous alludes to the rules that are utilized in the advancement of a HIoT gadget, while the last option alludes to medical care applications that are utilized in either diagnosing a particular medical issue or estimating well-being measurements [8]. By giving answers for various medical care concerns, administrations and ideas have changed the medical services business [14]. With rising medical care requests and innovative headways, more administrations are being presented every day. These are currently turning into a significant piece of the HIoT framework configuration process [11], [12]. In a HIoT setting, each help gives an assortment of medical services arrangements.

4. CHALLENGES

As of late, the medical care business has seen critical mechanical headways and their utilization in the goal of medical services-related challenges [8]. This has significantly further developed medical care administrations, which are presently accessible at the dash of a button. IoT has effectively changed the medical care industry using keen sensors, distributed computing, and correspondence advances [5], [8]. IoT, as different innovations, has its arrangement of obstructions and issues that could be investigated more later on. In the accompanying part, we'll turn out a portion of the worries [13].

- There have been quick specialized enhancements as of late, requiring occasional moves up to HIoT-based gadgets. Countless associated clinical gadgets and sensors are utilized in each IoT-based framework. This involves

costly support, fix, and redesign costs, which might impact the organizations just as end-clients financials. Accordingly, sensors that can be worked with fewer support costs should be incorporated [11]-[13].

- The larger part of IoT gadgets is fueled by means of batteries. It is difficult to modify a sensor's battery whenever it has occurred introduced. Subsequently, a high-limit battery was utilized to control the framework. Subsequently, scientists throughout the planet are endeavoring to fabricate medical services devices that can produce their force. Connecting the IoT framework with environmentally friendly power frameworks is one such likely arrangement. Somewhat, these strategies can help with alleviating the worldwide energy issue [10], [23].
- The thought of ongoing observing has been modified by the coordination of distributed computing. Be that as it may, this has expanded the weakness of medical services organizations to aggressors. This could bring about the bungle of delicate patient information and affect the treatment cycle. A few preparatory insurances should be thought of while fostering a HIoT framework to shield it from this destructive assault [17], [18], [23].
- Real check, secure booting, adaptation to internal failure, approval of the executives, whitelisting, secret key encryption, and secure blending conventions should be in every way assessed and utilized by clinical and sensor gadgets in a HIoT organization to stay away from an assault [8].

5. CONCLUSION AND FUTURE SCOPE

The current research investigated a few features of the HIoT framework. The engineering of a HIoT framework, its parts, and the correspondence among these parts has all been analyzed inside and out here. Furthermore, this article gives information on contemporary medical care benefits that have explored IoT-based innovation. IoT innovation has helped medical care experts in checking and diagnosing an assortment of well-being concerns, estimating an assortment of well-being factors, and giving demonstrative capacities at far-off areas using these standards. This has moved the medical services industry's concentrate away from clinics and toward patients. We've likewise discussed diverse HIoT applications and their latest things. The difficulties and issues identified with the plan, creation and utilization of the HIoT framework have likewise been talked about. Before very long, these troubles will fill in as an establishment for future development and exploration center. Moreover, peruses who

are keen on beginning their exploration as well as making enhancements in the field of HIoT gadgets will get a piece of full modern information on the gadgets.

REFERENCES

- [1] B. Pradhan, S. Bhattacharyya and K. Pal. "IoT-based applications in healthcare devices". *Journal of Healthcare Engineering*, vol. 2021, p. 6632599, 2021.
- [2] M. Javaid and I. H. Khan. "Internet of Things (IoT) enabled healthcare helps to take the challenges of COVID-19 Pandemic". *Journal of Oral Biology and Craniofacial Research*, vol. 11, no. 2, pp. 209-214, 2021.
- [3] Proceedings of the 5th EAI International Conference on Smart Objects and Technologies for Social Good. ACM, New York, USA, 2019.
- [4] I. de Moraes Barroca Filho, G. S. Jr. Aquino and T. B. Vasconcelos. "Extending and instantiating a software reference architecture for IoT-based healthcare applications". In: *Computational Science and Its Applications ICCSA 2019*. Springer International Publishing, Cham, 2019, pp. 203-218.
- [5] S. Ketu and P. K. Mishra. "Internet of healthcare things: A contemporary survey". *The Journal of Network and Computer Applications*, vol. 192, no. 103179, p. 103179, 2021.
- [6] M. M. Alam, H. Malik, M. I. Khan, T. Pardy, A. Kuusik and Y. Le Moullec. "A survey on the roles of communication technologies in IoT-based personalized healthcare applications". *IEEE Access*, vol. 6, pp. 36611-36631, 2018.
- [7] M. A. Akkaş, R. Sokullu and H. E. Çetin. "Healthcare and patient monitoring using IoT". *Internet of Things*, vol. 11, no. 100173, p. 100173, 2020.
- [8] J. Qi, P. Yang, G. Min, O. Amft, F. Dong and L. Xu. "Advanced internet of things for personalised healthcare systems: A survey". *Pervasive and Mobile Computing*, vol. 41, pp. 132-149, 2017.
- [9] Y. A. Qadri, A. Nauman, Y. B. Zikria, A. V. Vasilakos and S. W. Kim. "The future of healthcare internet of things: A survey of emerging technologies". *IEEE Communications Surveys and Tutorials*, vol. 22, no. 2, pp. 1121-1167, 2020.
- [10] R. C. Dharmik, S. Gotarkar, P. Dinesh and H. S. Burde. "An IoT framework for healthcare monitoring system". *Journal of Physics: Conference Series*, vol. 1913, no. 1, p. 012145, 2021.
- [11] N. Gavrilović and A. Mishra. "Software architecture of the internet of things (IoT) for smart city, healthcare and agriculture: Analysis and improvement directions". *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 1315-1336, 2021.
- [12] L. Catarinucci, D. de Donno, L. Mainetti, L. Palano, L. Patrono, M. L. Stefanizzi, L. Tarricone. "An IoT-aware architecture for smart healthcare systems". *IEEE Internet of Things Journal*, vol. 2, no. 6, pp. 515-526, 2015.
- [13] G. Marques, R. Pitarma, N. M. Garcia and N. Pombo. "Internet of things architectures, technologies, applications, challenges, and future directions for enhanced living environments and healthcare systems: A review". *Electronics (Basel)*, vol. 8, no. 10, p. 1081, 2019.
- [14] D. Castro, W. Coral, J. Cabra, J. Colorado, D. Méndez and L. Trujillo. "Survey on IoT solutions applied to Healthcare". *Dyna (Medellin)*, vol. 84, no. 203, pp. 192-200, 2017.
- [15] R. De Michele and M. Furini. "IoT Healthcare: Benefits, Issues and Challenges". In: *Proceedings of the 5th EAI International Conference on Smart Objects and Technologies for Social Good*, 2019.
- [16] S. B. Baker, W. Xiang and I. Atkinson. "Internet of things for smart healthcare: Technologies, challenges, and opportunities". *IEEE Access*, vol. 5, pp. 26521-26544, 2017.
- [17] P. P. Ray, D. Dash and D. De. "Edge computing for internet of things: A survey, e-healthcare case study and future direction". *The Journal of Network and Computer Applications*, vol. 140, pp. 1-22, 2019.
- [18] A. Jain, M. Singh and P. Bhabri. "Performance evaluation of IPv4-IPv6 tunneling procedure using IoT". *Journal of Physics: Conference Series*, vol. 1950, no. 1, p. 012010, 2021.
- [19] L. M. Dang, M. J. Piran, D. Han, K. Min and H. Moon. "A survey on Internet of things and cloud computing for healthcare". *Electronics (Basel)*, vol. 8, no. 7, p. 768, 2019.
- [20] B. Oryema, H. S. Kim, W. Li and J. T. Park. "Design and Implementation of an Interoperable Messaging System for IoT Healthcare Services". In: *2017 14th IEEE Annual Consumer Communications and Networking Conference (CCNC)*, 2017.
- [21] A. Gatouillat, Y. Badr, B. Massot and E. Sejdic. "Internet of medical things: A review of recent contributions dealing with cyber-physical systems in medicine". *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3810-3822, 2018.
- [22] F. Sallabi, F. Naeem, M. Awad and K. Shuaib. "Managing IoT-based smart healthcare systems traffic with software defined networks". In: *2018 International Symposium on Networks, Computers and Communications (ISNCC)*, 2018.
- [23] D. D. Ramlowat and B. K. Pattanayak. "Exploring the internet of things (IoT) in education: A re-view". In: *Advances in Intelligent Systems and Computing*. Springer Singapore, Singapore, 2019, pp. 245-255.

Determination of Potassium Bromate in Bread Brands in Sulaimani City, Kurdistan-Iraq



Sardar M. Weli¹, Sabiha M. Salih², Abdullah A. Hama^{2,3*}, Ary B. Faiq², Fatimah M. Ali⁴

¹Department of Nursing and Research Center, College of Health and Medical Technology, Sulaimani Polytechnic University, Kurdistan Region, Iraq, ²Department of Medical Laboratory and Research Center, College of Health and Medical Technology, Sulaimani Polytechnic University, Kurdistan Region, Iraq, ³Department of Medical Laboratory Science, College of Health Science, University of Human Development, Kurdistan Region, Iraq, ⁴Department of Nursing and Research Center, Sulaimani Technical Institute, Sulaimani Polytechnic University, Kurdistan Region, Iraq

ABSTRACT

Bread is the most consumed and staple food in many countries worldwide. It is made from dough of flour such as wheat and barley, and water. It usually contains flour improver potassium bromate (KBrO_3) which is used by bakers. However, many studies have confirmed the deleterious effects of KBrO_3 on human health. Therefore, this study aimed to determine the rate of KBrO_3 in five main types of bread in Sulaimani city, Kurdistan-Iraq. The duration of the study was from August 2021 to November 2021. Thirty bread samples were collected from five main products that are extremely consumed by Kurdish citizens. The bread-type products were bakery bread (Nani frn), white hamburger bread (Samun), white bread known as Kurdish bread (Nani Hawrami), pizza, and brown barley bread. Single beam UV-visible spectrophotometer APEL-303 was used for the quantification of KBrO_3 in bread samples. The results found that all 30 samples were had KBrO_3 residues in their products with different concentrations. Samples of brown barley bread were having the least content of KBrO_3 while samples from pizza dough were having the highest concentration of KBrO_3 . The present study concludes that all bread samples from five major bread types had potassium bromate above the permitted levels allowed by the United States Food and Drug Agency (FDA).

Index Terms: Potassium Bromate, White Bread, Barley Bread, Flour Improver, Spectrophotometer

1. INTRODUCTION

Bread is the most consumed and staple food in many countries around the world. It is made from dough of flour such as wheat and barley, and water. It usually contains several ingredients such as table salt, sugars, flavors, and flour improver [1]. Potassium bromate (KBrO_3) was commonly

used due to its low cost and acts as a slow oxidizing agent and it makes the dough more strength, and more elastic [2]. Many studies have confirmed the deleterious effects of potassium bromate on human health. For example, according to a study done on mice, potassium bromate administration caused impairment in renal and hepatic tissues. It also increased plasma creatinine levels and decreased antioxidant capacity [3]. Another study found that KBrO_3 exposed mice had increased lipid peroxidation, protein oxidation, and numerous degenerative changes in the cerebellum tissues [4]. In addition, important vitamins in bread such as thiamine (B1) and Niacin (B3) were destroyed by the effects of KBrO_3 . Carcinogenic and mutagenic effects of KBrO_3 were also confirmed in experimental animals [5]. The Center for Science and Environment (CSE) [6] indicated that some

Access this article online

DOI: 10.21928/uhdjst.v1n1y2017.pp76-79

E-ISSN: 2521-4217

P-ISSN: 2521-4209

Copyright © 2022 Weli *et al.* This is an open access article distributed under the Creative Commons Attribution Non-Commercial No Derivatives License 4.0 (CC BY-NC-ND 4.0)

Corresponding author's e-mail: Abdullah A. Hama, Department of Medical Laboratory and Research Center, College of Health and Medical Technology, Sulaimani Polytechnic University, Kurdistan Region, Iraq, and Department of Medical Laboratory Science, College of Health Science, University of Human Development, Kurdistan Region, Iraq. E-mail: abdullah.hama@spu.edu.iq

Received: 04-03-2022

Accepted: 07-04-2022

Published: 19-05-2022

studies found the link of bromate to cancer, so the global scientific expert committees and CSE suggested reducing the allowed limit of use; also, they recommended that KBrO_3 should not be used as a flour treatment agent.

Many studies have proofed that potassium may be cause detrimental health effects in humans [7,8], also in the same area (Metropolis Rivers Stat, Nigeria), the concentration of KBrO_3 in all samples was above the allowed concentration and the authors advised the consumer of read from the study area the bread conduction may be of harmful for our health [9]. Due to the harmful effects of this substance, many countries, including France, the United Kingdom, and Canada, have removed KBrO_3 from the list of acceptable additive substances to flour [10]. However, the maximal permitted dose of KBrO_3 in bread in other countries such as Japan, China, and the USA is 10 mg/kg, 50 mg/kg of flour mass, and 0.02 mg/kg, respectively [11], the most studies indicate that potassium bromate in bread concentration was exceeded the acceptable limit of 0.02 $\mu\text{g/g}$ set by FDA, in Delta State, all 15 bread brand samples was contained a higher concentration of KBrO_3 than permitted range [12] and the authors sate that this can be very dangerous for the bread consumers in the study area, and in the study in Erbil, the level of KBrO_3 was found to be more (6.66 mg/L–67.45 mg/L) than the permissible limit set by FDA [13]. This study aimed to determine the level of KBrO_3 in different types of bread in Sulaimani city, Kurdistan-Iraq.

2. MATERIALS AND METHODS

2.1. Collection of Samples

Bread samples were collected during the day (morning and afternoon) from different bakeries and from different locations in Sulaimani city from August 2021 to November 2021. The locations were the city center, Ibrahim-pasha, Ibrahim-Ahmad, Kani-BA, Sarchnar, Tui-Malik, and Family-mall. Thirty bread samples were collected from five main products that are extremely consumed by Kurdish citizens. The bread-type products were bakery bread (Nani frn), white hamburger bread (Samun), white bread known as Kurdish bread (Nani Hawrami), pizza, and brown barley bread.

2.2. Preparation of Samples

Samples were prepared according to a procedure that has been described and used by Abdulla and Hassan [14]. A small part (about 2 cm) in the center of each bread sample was dried in the oven for 72 h at 55°C. After drying, the sample was ground with an electric grinder to a powder. A 2.5 g of

the powder were dissolved in 25 ml of distilling water. After centrifuging, the liquid fraction was diluted to 50 ml.

2.3. Standard Preparations

A stock solution of 200 ppm potassium bromate (KBrO_3) was prepared by dissolving 0.200 grams of KBrO_3 into 1 L of distilling water. The standard series solutions of potassium bromate were prepared from the stock solution at 0, 4, 12, 20, and 40 ppm.

2.4. Method

A 5 ml of standard or sample solution was mixed with 5 ml of 1% KI, 10 ml of 0.1 N HCl, and then completed to 100 ml. The standards and samples were read after 10 min by a single beam UV–visible spectrophotometer APEL-303 at wavelength 420 nm with a calibration curve used for quantification of the samples.

2.5. Data Analysis

Data were entered into a Statistical Package for the Social Sciences “SPSS” version 26 for the storage and statistical analysis. The one-way ANOVA test was applied to test for association between different groups, with $P = 0.05$ or less considered significant.

3. RESULTS

The results of this study found that all 30 samples from five main types of bread contained different amounts of KBrO_3 residues in their products. Sample number 26 (brown barley bread) had the least content of KBrO_3 while sample number 21 (pizza dough) had the highest concentration of KBrO_3 (Table 1). The calibration curve of this study was shown (Fig. 1).

In addition, this study has found that the concentrations of KBrO_3 were highest in the pizza group and lowest in the brown barley bread group. The means and standard error of all groups with significant differences between each group of bread types are shown in Table 2 and Fig. 2. There were differences in the means of all groups. However, there were no significant differences between Kurdish bread, white bakery bread, and brown barley bread. On the other hand, there was a significant difference between pizza flour, brown barley bread, and Kurdish bread.

4. DISCUSSION

This study was carried out to determine the level of potassium bromate (KBrO_3) in the bread samples and to find the

TABLE 1: Concentrations of potassium bromate (ppm) in all bread samples

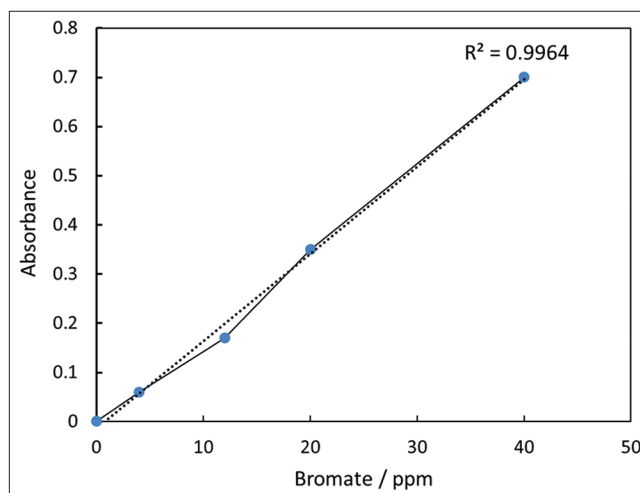
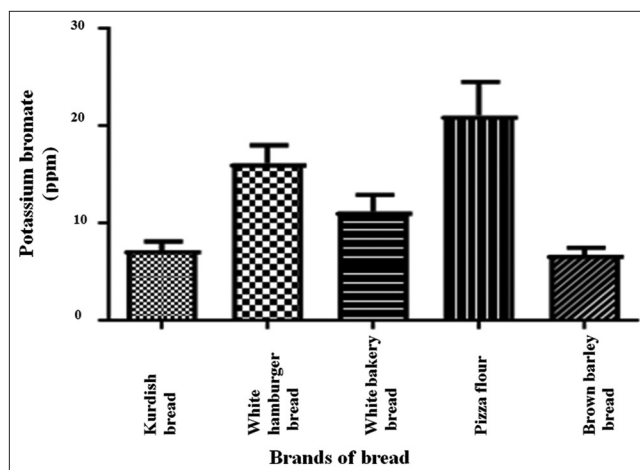
Samples	Type of breads	Quantity of KBrO ₃ (ppm)
1	Kurdish bread (Nani Hawrami)	9.747
2	Kurdish bread (Nani Hawrami)	9.747
3	Kurdish bread (Nani Hawrami)	7.58
4	Kurdish bread (Nani Hawrami)	6.137
5	Kurdish bread (Nani Hawrami)	4.693
6	Kurdish bread (Nani Hawrami)	5.415
7	White hamburger bread (Samun)	14.801
8	White hamburger bread (Samun)	11.913
9	White hamburger bread (Samun)	21.227
10	White hamburger bread (Samun)	10.65
11	White hamburger bread (Samun)	16.968
12	White hamburger bread (Samun)	21.227
13	White bakery bread (Nani frn)	7.581
14	White bakery bread (Nani frn)	13.791
15	White bakery bread (Nani frn)	12.635
16	White bakery bread (Nani frn)	10.65
17	White bakery bread (Nani frn)	16.968
18	White bakery bread (Nani frn)	5.415
19	Pizza flour	11.913
20	Pizza flour	9.747
21	Pizza flour	29.783
22	Pizza flour	26.534
23	Pizza flour	27.076
24	Pizza flour	21.227
25	Brown barley bread	7.581
26	Brown barley bread	4.693
27	Brown barley bread	5.415
28	Brown barley bread	6.859
29	Brown barley bread	9.747
30	Brown barley bread	6.137

TABLE 2: Concentrations of KBrO₃ (ppm) in all five groups of bread samples

Group number	Type of bread	Concentrations of KBrO ₃ Means±SE
1	Kurdish bread (Nani Hawrami)	7.22±2.18 A
2	White hamburger bread (Samun)	16.13±4.52 BE
3	White bakery bread (Nani fern)	11.17±4.22 AB
4	Pizza flour	21.05±8.41 CB
5	Brown barley bread (Nani Jo)	6.74±1.79 AD

Values are presented as means±SE (n=6 sample/group). Different capital letters denote significant differences between groups (P<0.05).

highest and lowest concentrations of KBrO₃ in different types of bread. Thirty samples from five major consumed types of bread were analyzed and KBrO₃ was found in all samples. According to the US Food and Drug Agency (FDA), the amount of KBrO₃ in bread higher than 0.02 µg/g (0.00002 parts per million) is considered not safe for human consumption [15]. All 30 samples of the present study were having concentrations of KBrO₃ higher than the national permitted levels so that none of the bread of all major types in

**Fig. 1.** Calibration curve.**Fig. 2.** Concentrations of KBrO₃ (ppm) in all five groups of bread samples. Values are presented as means ± SE (n=6 sample/group).

Sulaimani city might be unsafe for human consumption. This is in agreement with a study done in Hawler city, Kurdistan Region of Iraq; they found that the residual bromate level in the analyzed bread samples by spectrophotometer was in the range from 6.66 mg/L to 67.45 mg/L [13]. In addition, a study in Iraq (Baghdad city) found that electrical Samun and loaf had 10 and 0.3 µg/g potassium bromate, respectively. These levels were higher than the permissible level by the US Food and Drug Agency (FDA). They also found the exposed bread industry workers were elevated in chromosomal aberrations (CA), represented by chromatic breaks (CB), micronuclei (MN), and ring chromosome (RC) [16]. Another study in Iraq (Basrah city) found the harmful effects of potassium bromate on both hematological and biochemical parameters in rats. Liver enzymes (A.L.T and A.S.T) were increased and blood

parameters (RBC, Hb, WBC, and PCV) were decreased [17]. A recent study which is done in Dhaka city in Bangladesh showed that 67% of collecting samples were had KBrO_3 above the permitted level [18]. The present study also found that there were different concentrations of residues of KBrO_3 in different types of bread. The concentrations of KBrO_3 were highest in the pizza group and lowest in the brown barley bread group. This agrees with a study done in Tunis country. They observed different concentrations of bromate residues in different types of bread. The muffin contained the highest mean concentration of bromate residue as opposed to bread without salt, which had the lowest mean bromate level [6]. Moreover, a study in Nigeria found that 25% of the bread samples were had potassium bromate above the permissible limit allowed by the US Food and Drug Agency (FDA) and explained that these samples are unsafe for human consumption [19].

5. CONCLUSION

The present study concludes that all bread samples from five major bread types had potassium bromate above the permitted levels allowed by the US Food and Drug Agency (FDA). In general, all samples are unsafe for human consumption; however, the riskiest samples that have a greater concentration of potassium bromate were pizza flour and white bakery bread. The Kurdish bread and brown barley bread have a lower concentration of potassium bromate compared to others.

REFERENCES

- [1] M. O. Emeje, S. I. Ofoefule, A. C. Nnaji, A. U. Ofoefule and S. A. Brown. "Assessment of bread safety in Nigeria: Quantitative determination of potassium bromate and lead". *African Journal of Food Science*, vol. 4, no. 6, pp. 394-397, 2010.
- [2] A. Abu-Obaid, S. Abu-Hasan and B. Shraydeh. "Determination and degradation of potassium bromate content in dough and bread samples due to the presence of metals". *American Journal of Analytical Chemistry*, vol. 7, pp. 487-493, 2016.
- [3] N. G. Altoom, J. Ajarem, A. A. Allam, S. N. Maodaa and M. A. Abdel-Maksoud, "Deleterious effects of potassium bromate administration on renal and hepatic tissues of Swiss mice". *Saudi Journal of Biological Sciences*, vol. 25, no. 2, pp. 278-284, 2018.
- [4] H. B. Saad, D. Driss, I. Jaballi, H. Ghazzi, O. Boudawara, M. Droguet, C. Magne, M. Nasri, K. M. Zeghal, A. Hakim and I. B. Amara. "Potassium bromate-induced changes in the adult mouse cerebellum are ameliorated by vanillin". *Biomed Environ Science*, vol. 32, no. 2, pp. 115-125, 2018.
- [5] L. A. Alli, M. M. Nwegbu, B. I. Inyang, K. C. Nwachukwu, J. O. Ogedengbe, O. Onaadepe, M. A. Jamda, G. A. Akintan, S. O. Ibrahim and E. A. Onifade. "Determination of potassium bromate content in selected bread samples in Gwagwalada, Abuja-Nigeria". *International Journal of Health and Nutrition*, vol. 4, no. 1, pp. 15-20, 2013.
- [6] A. Tewari and A. Khurana. "Potassium Bromate/Iodate in Bread and Bakery Products". Centre for Science and Environment, pp. 1-12, 2016. Available from: <https://www.cseindia.org>. [Last accessed on 2022 May 06].
- [7] A. M. Magomya, G. G. Yebpella, U. C. Okpaegbe and P. C. Nwunjuji. "Analysis of potassium bromate in bread and flour samples sold in Jalingo Metropolis, Northern Nigeria". *Journal of Environmental Science*, vol. 14, no. 2, pp. 1-5, 2020.
- [8] N. A. Ugochukwu, O. Elechi and E. A. Ozioma. "Determination of bromate content of selected bread brands consumed within port harcourt and its environs". *Chemistry Research Journal*, vol. 4, no. 3, pp. 86-91, 2019.
- [9] A. U. Naze, O. A. Epete and E. Owhoeke. "Bromate content in thirty different brands of bread baked in Port Harcourt Metropolis Rivers State, Nigeria". *Journal of Applied Sciences and Environmental Management*, vol. 22, no. 8, p. 1321, 2018.
- [10] M. El Ati-Hellal, R. Doggui, Y. Krifa and J. El Ati. "Potassium bromate as a food additive: A case study of Tunisian breads". *Environmental Science and Pollution Research*, vol. 25, pp. 2702-2706, 2018.
- [11] J. El harti, Y. Rahali, M. Ansar, H. Benziane, J. Lamsaouri, M. O. B. Idrissi, M. Draoui, A. Zahidi and J. Taoufik. "A simple and rapid method for spectrophotometric determination of bromate in bread". *Journal of Materials and Environmental Science*, vol. 2, no. 1, pp. 71-76, 2011.
- [12] A. Uwague and O. C. Oghenekohwoyan. "Investigation into the health danger of potassium bromate in bread consumed in Sapele Town, Delta State". *International Journal of Modern Engineering Research*, vol. 7, no. 9, pp. 1-3, 2017.
- [13] S. A. Narmin and A. H. Media. "Spectrophotometric determination of bromate in bread by the oxidation of dyes". *Kirkuk University Journal-Scientific Studies*, vol. 4, no. 1, pp. 31-39, 2009.
- [14] N. S. Abdulla and M. A. Hassan. "Spectrophotometric determination of bromate in bread by the oxidation of dyes". *Journal of Kirkuk University-Scientific Studies*, vol. 4, no. 1, pp. 31-39, 2009.
- [15] A. S. Ekop, I. B. Obot and E. N. Ikpatt. "Anti-nutritional factors and potassium bromate content in bread and flour samples in Uyo Metropolis, Nigeria". *E-Journal of Chemistry*, vol. 5, no. 4, pp. 736-741, 2008.
- [16] A. Haleem. "Cytogenetic effects of potassium bromate KBrO_3 associated with Iraqi baking industry cytogenetic effects of potassium bromate KBrO_3 associated with Iraqi baking industry". *Indian Journal of Applied Research*, vol. 4, no. 6, pp. 10-12, 2015.
- [17] S. A. Zainab and R. F. Ghadhban. "Effect of potassium bromate on some hematological and biochemical parameters and protective role of Vitamin C on laboratory rats (*Rattus rattus*). *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 2, pp. 669-674.
- [18] S. S. Mahmud, M. Moni, A. B. Imran and T. Foyez. "Analysis of the suspected cancer causing potassium bromate additive in bread samples available on the market in and around Dhaka City in Bangladesh". *Food Science and Nutrition*, vol. 9, pp. 3752-3757, 2021.
- [19] H. I. Kelle, V. U. Oguezi and I. P. Udeozo. "Qualitative and spectrophotometric determination of potassium bromate in bread samples sold in Asaba, Delta State, Nigeria". *Pakistan Journal of Chemistry*, vol. 5, no. 2, pp. 93-95, 2015.

p-ISSN 2521-4209
e-ISSN 2521-4217



UHD Journal of Science and Technology

A Scientific periodical issued by University of Human Development

Vol.6 No.(1) June 2022

2022

2722

e.mail:jst@uhd.edu.iq