e-ISSN : 2829-4149
p-ISSN : 2829-4157

COELITE

JOURNAL OF COMPUTER ENGINEERING, ELECTRONICS AND INFORMATION TECHNOLOGY

OJS :

Publisher :

The Education University

Vol. 05| Issue. 01 | Apr 2026

Volume 5, Issue 1, April 2026

e-ISSN : 2829-4149
p-ISSN : 2829-4157

Journal of
Computer Engineering, Electronics and Information Technology
Universitas Pendidikan Indonesia

Jl. Pendidikan No.15, Cibiru Wetan, Kec. Cileunyi, Kab. Bandung, Jawa Barat 40625
Phone: (022) 7801840, Fax: (022) 7830426
E-Mail: coelite@upi.edu, Website: https://ejournal.upi.edu/index.php/COELITE

Editor in Chief
Dr. Eng. Munawir, S.Kom., M.T., (Scopus ID 57208758022)
Expertise: Data Mining, Text Mining.
Universitas Pendidikan Indonesia, Indonesia.

Executive Editor
Anugrah Adiwilaga, S.ST., M.T., (Scopus ID 57194714935)
Expertise: Robotic, Control and Intelligent Systems, AI, IoT.
Universitas Pendidikan Indonesia, Indonesia.

Editorial Board
Dr. Eng. Dinda Pramanta, M.Eng.
Expertise: Artificial Intelligence, FPGA, Robot and Neural Network.
Associate Professor at Kyushu Institute of Information sciences Japan.

Wirmanto Suteddy, S.T., M.T.,
Expertise: Embedded System, IoT, AI, VLSI, Computer Vision.
Universitas Pendidikan Indonesia, Indonesia.

Dr. Eng. Mia Rizkinia, S.T.,M.T.
Expertise: Optoelectrotechnique and Remote Sensing, Vision, Image and Graphics Processing.

Editorial Assistant
Deden Pradeka, S.T., M.Kom., (Scopus ID 57200731472)
Expertise: Software Engineering Development, Mobile Development, Cryptography, Steganography.
Universitas Pendidikan Indonesia, Indonesia.

Reviewers

Dr. Eng. Md. Zaman Molla
Expertise: Solar Cell.
Associate Professor of Physics at Ahsanullah University of Science and Technology Bangladesh.

Dr. Eng. Alfian Akbar Gozali, S.T., M.T
Expertise: Data Analytics, Machine Learning, Optimization Algorithm, Mobile Development.
Telkom University, Indonesia.

Dr. Eng. Bobby Kurniawan, S.T., M.T
Expertise: Production system and control, mathematical modeling, scheduling metaheuristics, AI.
Universitas Sultan Ageng Tirtayasa, Indonesia.

Syamsul Rizal S.T, M.T., PhD.
Expertise: Image Processing, IoT,Faulty Node, Fault Detection.
Telkom University, Indonesia.

Devi Aprianti Rimadhani Agustini, S.Si, M.Si
Expertise: Modeling and Simulation.
Universitas Pendidikan Indonesia, Indonesia.

Muhammad Taufik Dwi Putra, S.Tr.Kom., M.T.I
Expertise: Computer vision, deep learning, Machine Learning.
Universitas Pendidikan Indonesia, Indonesia.

Ana Rahma Yuniarti, S.T., M.Eng., (SCOPUS ID 57192948351)
Expertise: Biomedical Engineering, Machine Learning, Artificial Inteligence.
Universitas Pendidikan Indonesia, Indonesia.

Volume 5, Issue 1, April 2026

e-ISSN : 2829-4149
p-ISSN : 2829-4157

published by **Universitas Pendidikan Indonesia.**

Journal of

Computer Engineering, Electronics and Information Technology

COELITE

Volume 5, Issue. 1, April 2026

Table of Contents

Optimization and Monitoring of Biogas Power Plants for Energy Management Using Genetic Algorithms and the Internet of Things (Case Study: Kendari City Slaughterhouse) 01-16
Al Ubai, Asminar Asminar, Abdul Djohar, Muhammad Nadzirin Anshari Nur, Bunyamin Bunyamin, Agustinus Lolok

Implementation of the Jaro-Winkler Algorithm for Automatic Grading of Short Answer Questions in the Kifayatul Azhar Elementary School Exam 17-28
Rizki Halomoan, Benny Yoga, Nauval Ardana, Trisna Gelar, Santi Sundari

Double Proxy Architecture for Data Center DDoS Mitigation 29-42
Rahmawan Bagus Trianto, Kukuh Muhammad, Dwi Novia Prasetyanti

A Comparative Evaluation of State-of-the-Art Deep Face Recognition Models on the Indonesian Muslim Student Face Dataset 43-62
Mokhamad Arfan Wicaksono, Ibrahim Faizal Burhan, Hadi Permana, Muhung Anggarawan, Bakhriyah Firdausi, Muhammad Rafid Amrullah, Ratnasari Nur Rohmah

Graph Theory-Based Simulation for Optimal Access Point Selection in Smart Buildings 63-70
Devi Aprianti Rimadhani Agustini, Deva Naufal, Muhammad Zirly Atthaulah, Halimah Kiysa Fiddien, Alaqso Fattar Djuson

Indexed by.

Google Scholar GARUDA GARBA RUJUKAN DIGITAL sinta S4 Science and Technology Index

Volume 5, Issue 1, April 2026

e-ISSN : 2829-4149
p-ISSN : 2829-4157

Published by:
Universitas Pendidikan Indonesia
E-Mail: coelite@upi.edu
Website: https://ejournal.upi.edu/index.php/COELITE

Journal of Computer Engineering, Electronics and Information Technology (COELITE)

Journal homepage: https://ejournal.upi.edu/index.php/COELITE

Optimization and Monitoring of Biogas Power Plants for Energy Management Using Genetic Algorithms and the Internet of Things (Case Study: Kendari City Slaughterhouse)

Al Ubai[1], Asminar [2], Abdul Djohar[3], Muhammad Nadzirin Anshari Nur [4], Bunyamin [5], Agustinus Lolok [6]*

[1,2,3,4,5,6] Department of Electrical Engineering, Universitas Halu Oleo, Indonesia

Correspondence E-mail: alubaiwork@gmail.com

ABSTRACT

This project seeks to build and optimize a Biogas Power Plant system connected with Internet of Things (IoT) technology and Genetic Algorithms (GA) for energy management purposes. The system is executed as a simulation at the Kendari City Slaughterhouse (RPH), employing organic waste as biogas feedstock. IoT-based monitoring is utilized to track essential data, such as methane concentration, temperature, pressure, voltage, and current, in real time. Genetic Algorithms are utilized to optimize load distribution, guaranteeing efficient energy consumption. The findings demonstrate that the monitoring system operates dependably, exhibiting great sensor precision, with voltage readings attaining an average accuracy of 98.91% and pressure sensors achieving 99.1% accuracy. Additionally, GA optimization effectively identifies an appropriate load supply combination that aligns closely with the generator capacity, avoiding overload problems. The integration of IoT and GA augments the performance, efficiency, and reliability of the PLTBG system, exhibiting its capacity for sustainable energy management in biogas-powered generation.

ARTICLE INFO

Article History:
Submitted/Received 02 Mar 2026
First Revised 16 Mar 2026
Accepted 30 Mar 2026
First Available online 01 Apr 2026
Publication Date 01 Apr 2026

Keyword:
Biogas power plant,
Genetic Algorithm,
IoT, Monitoring System,
Renewable Energy

© 2026 Universitas Pendidikan Indonesia

1. INTRODUCTION

The demand for sustainable energy sources is increasing in the contemporary era, propelled by swift globalization and a rising worldwide population. Conversely, fossil fuel reserves are diminishing and contributing to numerous adverse environmental effects, including heightened greenhouse gas emissions. As fossil fuel supplies deplete and global energy consumption escalates, the prospect of an energy crisis has emerged as a significant concern in numerous nations. The International Energy Agency (IEA) reports that more than 80% of global energy requirements continue to depend on non-renewable resources, including natural gas, coal, and crude oil. This scenario demonstrates that excessive reliance on fossil fuels poses a threat of an impending energy crisis and intensifies environmental challenges like climate change, necessitating a shift to more sustainable renewable energy sources [1].

The disparity between demand and the supply of conventional energy has catalyzed the advancement of many alternative energy sources that are more ecologically sustainable. A notable renewable energy source with considerable potential is biogas, which can be utilized via biogas power plants. PLTBG employs organic waste as its principal feedstock; hence, this technology not only produces power but also enhances sustainable waste management. Indonesia, possessing a substantial agricultural and animal sector, has an ample source of biogas feedstock; nevertheless, its utilization is poor. Consequently, the advancement of biogas power plant technology represents a viable approach to fulfill energy requirements while mitigating the environmental repercussions of organic waste [2].

Nonetheless, the deployment of biogas power plants continues to encounter numerous obstacles, especially with energy conversion efficiency and system reliability. Innovative and integrated methodologies are essential to enhance the optimal functioning of biogas systems, particularly within the slaughterhouse sector, which produces substantial quantities of organic waste. Inadequately handled slaughterhouse waste can adversely affect the environment, rendering its usage as biogas feedstock a viable solution [3]. Research by Asminar et al. (2024) on the modeling and simulation of off-grid PV–biogas systems at the Kendari City slaughterhouse reveals that the system utilizes basic rule-based control, such as connecting or disconnecting batteries at predetermined levels, without accommodating fluctuations in biogas production or alterations in load. In that investigation, biogas generated approximately 1.175 kW of the total power during operation, demonstrating that the current system has limits in optimizing energy output and load control [4][5].

An optimization strategy is required to adaptively modify operational parameters to enhance the performance of the PLTBG system. A commonly employed technique in power system optimization is the genetic algorithm (GA). This method was selected because of its capability to address both confined and unconstrained optimization problems. Genetic algorithms operate by iteratively altering a population of solutions via selection, crossover, and mutation processes until the optimal solution is achieved. By emulating the principle of natural selection, genetic algorithms can identify the optimal solution within an extensive search space [6].

The development of integrated monitoring systems is also a key factor in improving the performance of biogas power plants. Research conducted by Junus et al. (2025) shows that the integration of sensors and Internet of Things (IoT) technology into biogas production monitoring and control systems can enhance the efficiency and reliability of biogas power plants in Indonesia [7]. IoT sensors such as methane (CH4) sensors, pressure sensors, and temperature sensors connected to microcontrollers and communication modules are capable

DOI: https://doi.org/10.17509/coelite. v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

of collecting operational data in real-time and displaying this information on the monitoring system. The integration of this technology not only improves data visibility and responsiveness to changes in system conditions but also supports decision-making processes in energy management[8].

2. METHODS

This study was conducted at the Electrical Power Systems Laboratory, Faculty of Engineering, Halu Oleo University, and at the Kendari City Slaughterhouse over a period of six months, from September to March 2026. The research phases included a literature review, data collection, instrument design and development, and instrument testing. Each phase was carried out systematically to ensure the validity and accuracy of the data obtained.

2.1. Biogas power plant system configuration

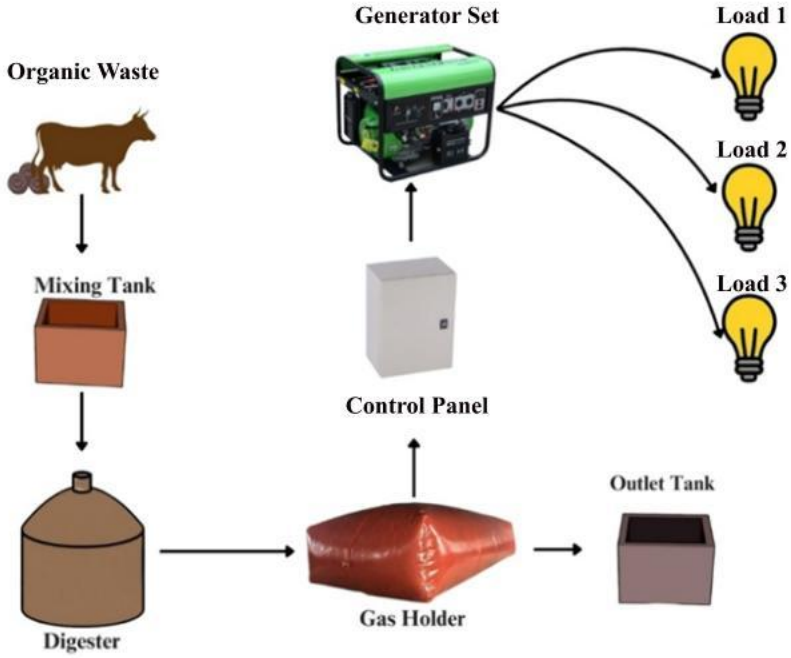


Figure 1. Biogas Power Plant System Configuration

Figure 1 above shows the design of a biogas power plant configuration, illustrating the structured process of converting organic waste into electrical energy through anaerobic digestion for 3 to 14 days [9][10][11]. The process begins by feeding organic waste, such as cow or livestock manure, into a mixing tank to be blended with water[12]. Next, the mixture is fed into the digester, an anaerobic reactor where methanogenic bacteria carry out fermentation without oxygen. These bacteria produce biomethane gas as the primary product of the fermentation process. To maintain stable pressure, the resulting gas is then directed into a gas storage tank. Afterward, the gas is routed through a pressure transmitter and controlled by a control panel to regulate the gas supply to the generator set. The generator produces electrical energy from the chemical energy of the biogas [13]. This electrical energy is then transmitted to loads such as lights or other electrical equipment. Meanwhile, the residual slurry from the digester is discharged into an outlet tank and can be used as liquid organic fertilizer [14].

DOI: https://doi.org/10.17509/coelite.v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

2.2. Design and theoretical planning

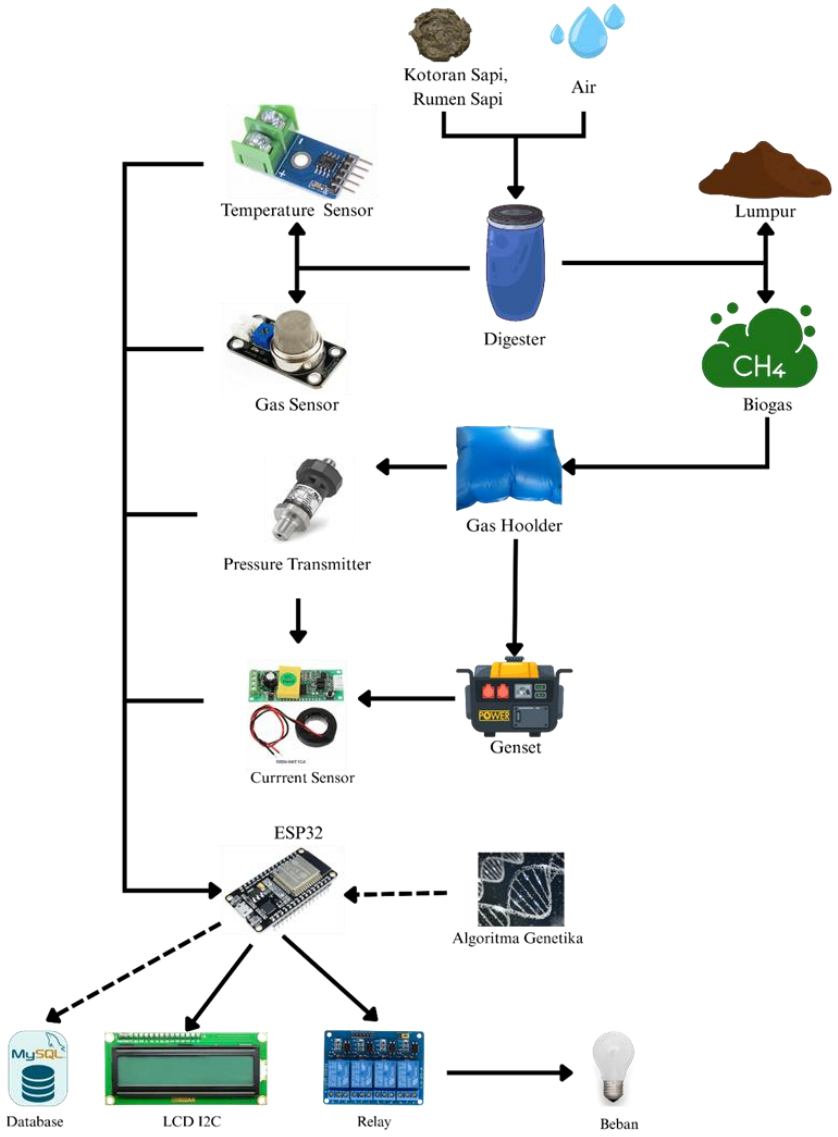


Figure 2. Architecture of The Proposed Biogas Monitoring System

The **Figure 2** system begins by introducing cow manure and rumen contents as input, mixed with water in a 1:1 ratio to remove coarse solids [15]. The next step evaluates the biogas potential of the input materials. The material is then fed into a digester for anaerobic fermentation. Inside the digester, an MQ4 sensor is used to detect the methane gas produced, and a temperature sensor is used to measure the temperature to optimize fermentation results [16]. The digester produces biogas as well as residual slurry that can be used as organic fertilizer. The biogas output is channeled through a pipe to a gas storage tank for collection, and the pressure inside is measured using a pressure transmitter[17]. The biogas is then routed to a generator, and a current sensor measures the electrical current output from the generator. The system is optimized using a genetic algorithm to determine the optimal load configuration. The data obtained is displayed on an LCD screen, stored in a MySQL database, and displayed in real-time via a monitoring website.

2.3. Genetic algorithm flowchart

DOI: https://doi.org/10.17509/coelite. v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

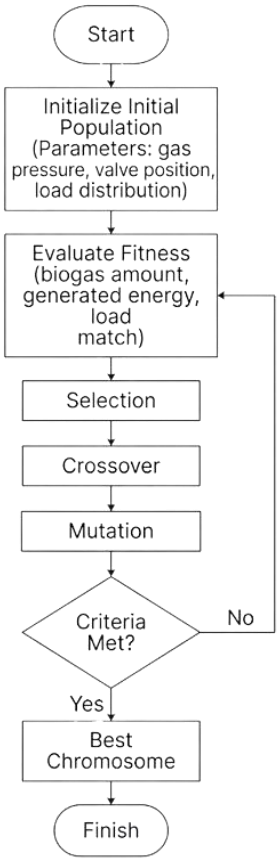


Figure 3. Flowchart of The Genetic Algorithm Process

In **Figure 3**, a Genetic Algorithm is used to optimize the configuration of a biogas power generation system. The process begins with population initialization, where a set of initial solutions (chromosomes) is randomly generated with predetermined parameters such as gas pressure, valve positions, and power distribution across the three load combinations [18]. Each individual in the population is then evaluated using a fitness function, which indicates the solution's success based on parameters such as the amount of energy generated, methane utilization, and the ability to meet load requirements. Selection is then performed to identify the best solutions, which are subsequently replicated through the crossover mechanism—the combination of two chromosomes to produce a new one. This is followed by a mutation process, where random changes are made to a small portion of the genetic values to prevent getting stuck in local optima. The results of the new crossover and mutation are then re-evaluated. If neither of the two termination criteria is met—whether the maximum number of generations or the conversion of an unstable solution—the process proceeds to the next cycle. Once the optimal configuration is found, the real-world system is configured, and the GA process concludes [19].

2.4. Sensor Performance Test Planning

System testing was conducted to evaluate the accuracy and reliability of the sensors in measuring environmental parameters in grain storage. This testing involved comparing sensor readings with those from reference instruments to determine the error rates and accuracy of each sensor. Measurement accuracy was calculated using the following formula [20]:

DOI: https://doi.org/10.17509/coelite.v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

$$Accuracy = \left(\frac{ReferenceValue - SensorValue}{ReferenceValue}\right) \times 100\% \quad (1)$$

This formula indicates the degree of closeness between the sensor's measurement results and the reference value, expressed as a percentage. The higher the accuracy value obtained, the better the sensor's performance in providing precise data. This testing includes a thermocouple temperature sensor compared to a digital thermometer, a Wisner WPT83G pressure sensor compared to a pressure gauge, and a PZEM sensor compared to a multimeter.

2.5. Data Collection Methods

Data collection in this study was conducted using several approaches aimed at obtaining accurate and relevant information regarding the optimization of biogas power plants and Internet of Things-based monitoring systems. One of the methods used was a literature review, which involved examining and analyzing various scientific references related to the research topic. These sources include scientific journals, books, conference proceedings, and research articles from reputable institutions. Through this literature review, the researcher can understand the basic concepts of biogas power plants, optimization methods that have been used previously, and various challenges that may arise in the implementation of technology-based energy monitoring and optimization systems.

In addition to a literature review, data collection was also conducted through consultations and field observations. Consultations were held with the supervising lecturer as well as staff and field supervisors at the Kendari City Slaughterhouse to obtain the information needed for the study, such as data on cattle weight and the number of cattle slaughtered each day. Meanwhile, observations were conducted directly at the research site to understand the actual conditions of the environment to be monitored. These observation activities included monitoring the average weight of incoming cattle, machine operating times at the slaughterhouse, total electricity consumption in the surrounding area, and various challenges that might be encountered in implementing a Biogas Power Plant system. By combining these methods, the data collected is expected to support a more comprehensive analysis and system design process.

3. RESULTS AND DISCUSSION

3.1. Design of an Internet of Things-Based Monitoring System for a Biogas Power Plant

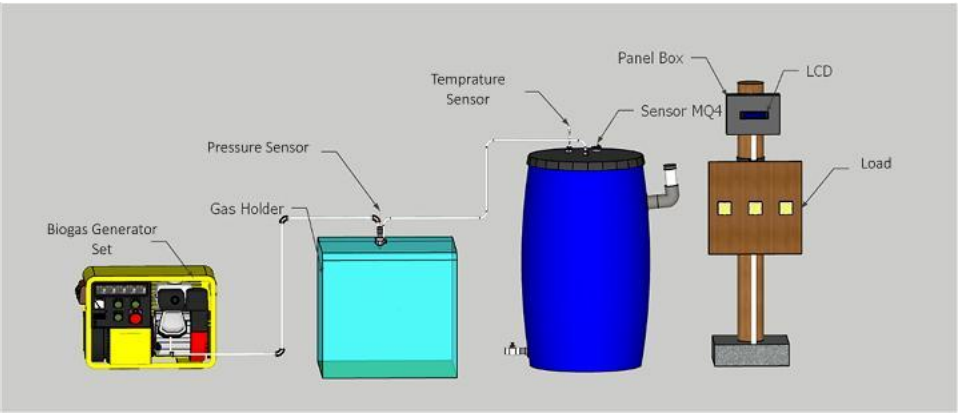


Figure 4. Visual Representation of The Proposed Biogas Power Generation System

Figure **4** shows a visual representation of the prototype design for the proposed biogas power generation system. This system integrates the process of converting organic waste into

DOI: https://doi.org/10.17509/coelite. v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

electrical energy through several interconnected main components. These include a digester, a gas holder, a biogas generator set, a control panel, a load, and several sensors installed within the system.

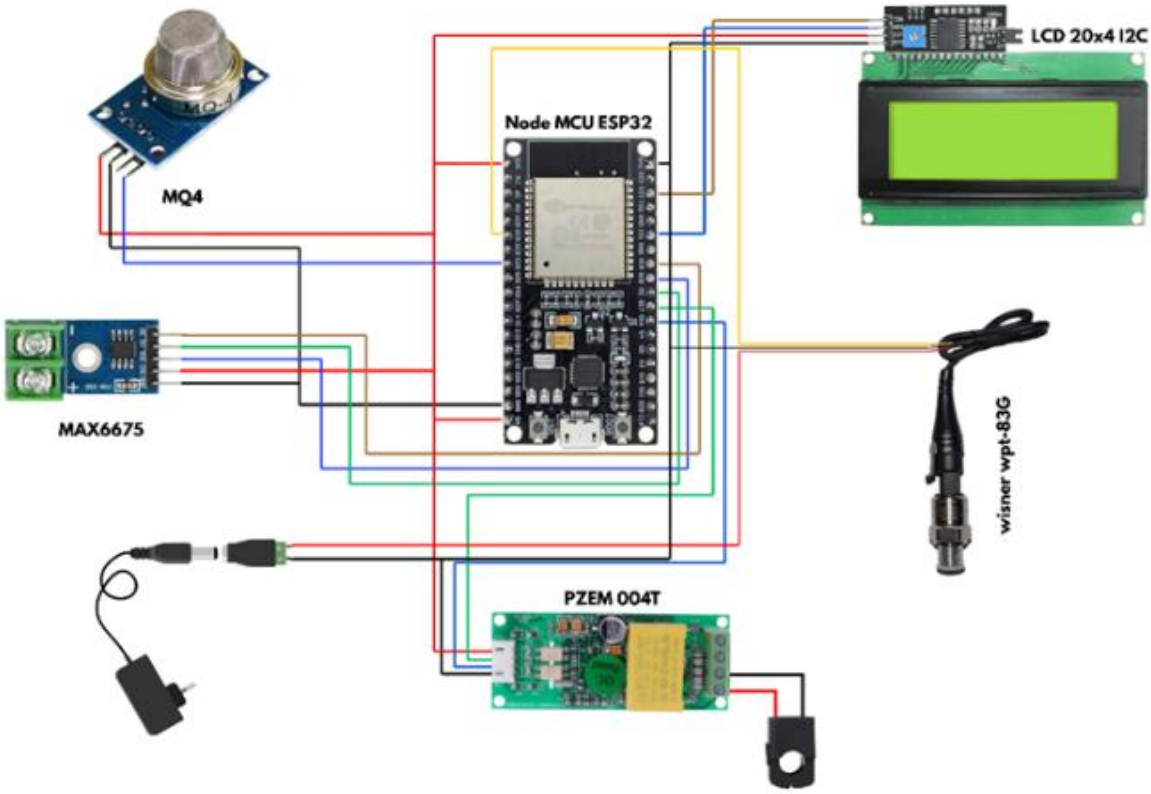


Figure 5. Wiring Diagram of The Monitoring System

In **Figure 5**, based on the wiring diagram shown, this monitoring system draws power from the mains electrical grid, which is converted to DC voltage via an adapter to serve as the primary power source for the ESP32 microcontroller, sensors, and other supporting devices. All components share the same ground to ensure circuit stability. This system utilizes several sensors to monitor key parameters in the biogas plant, such as the MQ-4 sensor to detect methane (CH4) concentration from the digester fermentation process, a Type-K thermocouple connected to the MAX6675 module to measure digester temperature via SPI communication, and the WPT-83G pressure sensor to measure gas pressure in the biogas bag. Additionally, the PZEM-004T module is used to monitor electrical parameters such as voltage, current, and power at the generator's output to the load via TX and RX serial communication with the ESP32 and a CT current sensor. The readings from all sensors are displayed locally via a 20×4 I2C-based LCD connected to the SDA and SCL pins, while the acquired data is also transmitted by the ESP32 via a WiFi network to a MySQL database so it can be monitored in real-time via a website.

DOI: https://doi.org/10.17509/coelite.v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

3.2. Evaluation

Table 1. Testing voltage with a multimeter

No	Voltage Measurement Results (V)		Error Value (%)	Accuracy (%)
	PZEM Sensor	Multimeter		
1	11.80	11.92	1.01	98.99
2	11.83	11.96	1.09	98.91
3	11.91	12.04	1.08	98.92
4	12.17	12.32	1.22	98.78
5	12.24	12.37	1.05	98.95
Average Error (%)			1.09	
Average Accuracy (%)			98.91	

Table 1 shows the results of the voltage comparison between the multimeter and the PZEM-004T sensor was conducted five times, yielding an average error of 1.09% and an average voltage reading accuracy of 98.91%, indicating that the sensor performs very well in measuring voltage and comes very close to the actual value.

Table 2. Testing current values with a multimeter

No	Current Measurement Results (A)		Error Value (%)	Accuracy (%)
	PZEM Sensor	Multimeter		
1	1.18	1.12	5.09	94.91
2	0.8	0.76	5.26	94.74
3	0.89	0.78	14.10	85.90
4	0.93	0.88	5.68	94.32
5	0.91	0.8	13.75	86.25
Average Error (%)			8.78	
Average Accuracy (%)			91.22	

Table 2 shows the result of the current comparison between the multimeter and the INA219 sensor was performed five times, yielding an average error of 8.78% and an average current reading accuracy of 91.2%, indicating that the sensor performs very well in measuring current and comes very close to the actual value.

DOI: https://doi.org/10.17509/coelite. v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

Table 3. MQ4 sensor response test results

Gas Exposure Duration	Gas Concentration Reading MQ-4 Sensor (ppm)
1 Minute	4273.20
	4262.20
	4284.30
	4328.60
	4339.80
Average	4404,73

Table 3 shows the results of a 1-minute test of the MQ-4 methane sensor in the digester, methane gas concentration readings ranged from 4262.20 to 4557.20 ppm, with an average of 4404.73 ppm. This average value was used as the baseline methane concentration under normal digester conditions.

Table 4. Testing the WPT-83G sensor with a pressure gauge

No	Pressure Measurement Results (BAR)		Error Value (%)	Accuracy (%)
	SWPT-83G Sensor	Pressure Gauge		
1	1,01	1	1,0	99,0
2	2,03	2	1,5	99,5
3	3,02	3	0,7	99,3
4	5,03	5	0,6	99,4
5	7,05	7	0,7	99,3
Average Error (%)			0,90	
Average Accuracy (%)			99,1	

Table 4 shows a comparison of pressure readings between the WPT-83G sensor and a pressure gauge was conducted on five samples, yielding an average error of 0.90% and an average pressure reading accuracy of 99.1%. These results indicate that the sensor effectively measures pressure within a chamber and closely approximates the actual value.

DOI: https://doi.org/10.17509/coelite.v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

3.3. Analysis of Energy Management Optimization Using Algorithms

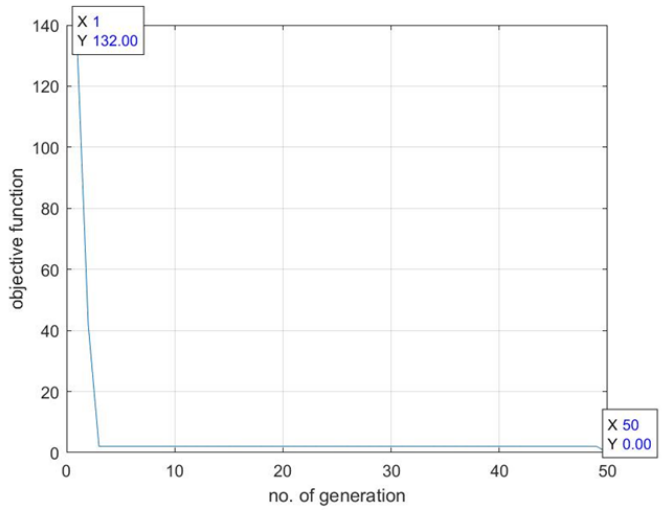


Figure 6. Optimization Results for Run 2

Based on the results of the GA optimization simulation from **Figure 6**, the best chromosome percentages were 61, 100, and 89 in the 50th generation, and validation was performed over 0 iterations. In the second simulation run, the best chromosome percentages were 61, 100, and 89.

400 A + 500 B + 400 C = 1100

400 (61%) + 500 (100%) + 400 (89%) = 1100

244 + 500 + 356 = 1100 W

The resulting combination is: Load A = 244 W; Load B = 500 W; Load C = 356 W.

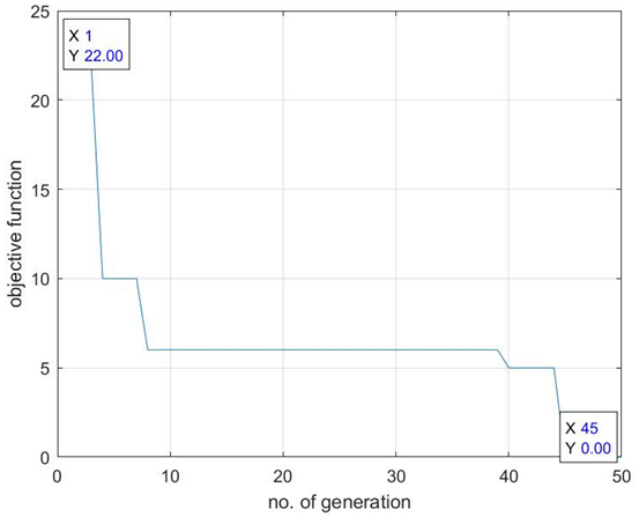


Figure 7. Optimization Results for Run 5

Based on the results of the GA optimization simulation from **Figure 7**, the best chromosome percentages were 78, 80, and 97 in the 39th generation, and validation was performed over 5 iterations. In the fifth simulation run, the best chromosome percentages were 78, 80, and 97

400 A + 500 B + 400 C = 1100

DOI: https://doi.org/10.17509/coelite. v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

400 (78%) + 500 (80%) + 400 (97%) = 1100

312 + 400 + 388 = 1100 W

The resulting combination is load A = 312 W, load B = 400 W, load C = 388 W.

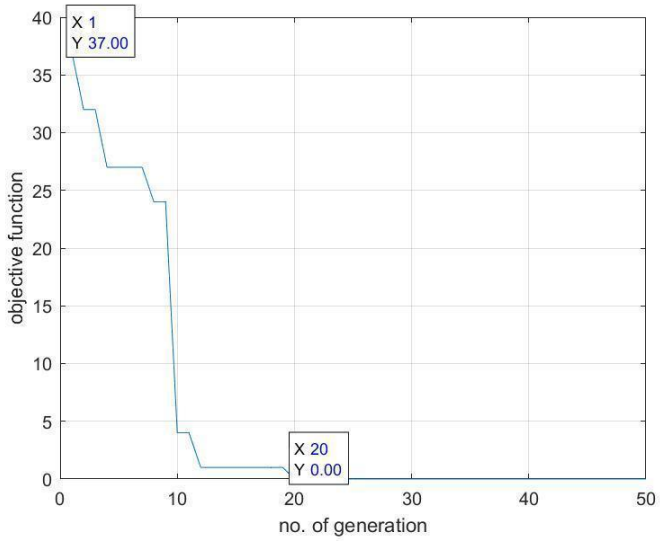


Figure 8. Top Optimization Results 3

Based on the results of the GA optimization simulation from **Figure 8**, the best chromosome percentages were 84, 76, and 96 in the 20th generation, and validation was performed over 30 iterations. In the first simulation run, the best chromosome percentages were 84, 76, and 96

400 A + 500 B + 400 C = 1100

400 (84%) + 500 (76%) + 400 (96%) = 1100

336 + 380 + 384 = 1100 W

The resulting combination is load A = 336 W, load B = 380 W, load C = 384 W.

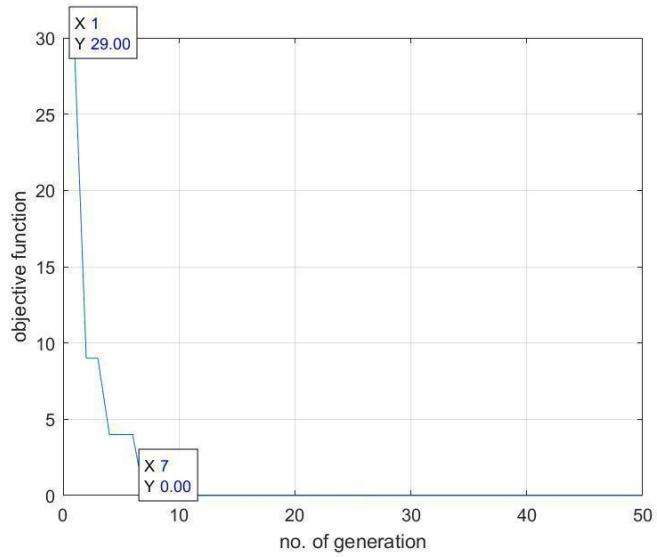


Figure 9. Top optimization results 4

Based on the results of the GA optimization simulation from **Figure 9**, the best chromosome percentages were 96, 64, and 99 in the 7th generation, and validation was performed over 93 iterations. In the third program run, the best chromosome percentages were 96, 64, and 99

DOI: https://doi.org/10.17509/coelite.v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

400 A + 500 B + 400 C = 1100
400 (96%) + 500 (64%) + 400 (99%) = 1100
384 + 320 + 396 = 1100 W
The resulting combination is load A = 384 W, load B = 320 W, load C = 396 W.

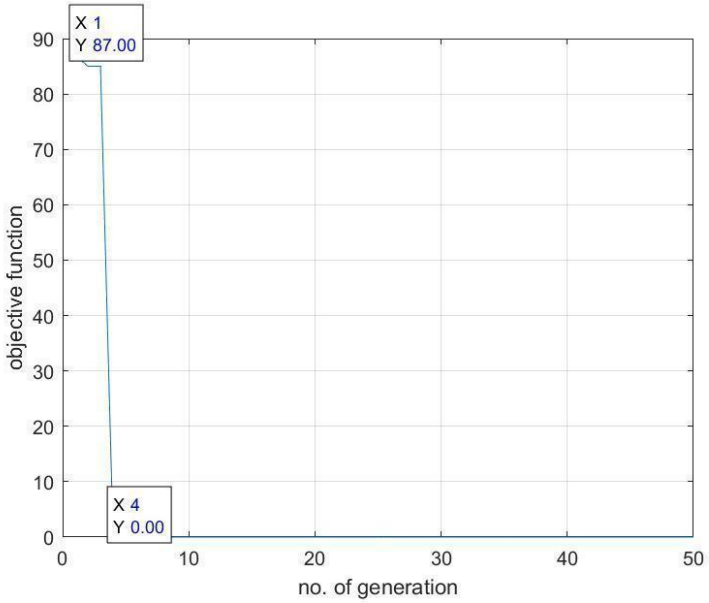


Figure 10. Top optimization results 5

Based on the results of the GA optimization simulation from **Figure 10**, the best chromosome percentages were 98, 66, and 97 in the 4th generation, and validation was performed over 96 iterations. In the first program run, the best chromosome percentages were 98, 66, and 97
400 A + 500 B + 400 C = 1100
400 (98%) + 500 (66%) + 400 (97%) = 1100
392 + 330 + 388 = 1100 W
The resulting combination is load A = 392 W; load B = 330 W; load C = 388 W.

3.4. Prototype Load Supply Combination

Table 5. Results of the genetic algorithm prototype simulation

Load	Supply Percentage (%)	Maximum Power (W)
Load A	89 %	20 W
Load B	60 %	5 W
Load C	100 %	10 W
Total		35 W

Table 5 shows that the optimization results show that the Genetic Algorithm successfully identified the optimal load-supply combination. In this combination, Load A is supplied at 89%, Load B at 60%, and Load C at 100%. This combination yields a total power of 26W. This condition still meets the system's operational limits since the value remains below the power plant's maximum capacity of 35W. The percentage results of the optimization are then transmitted from MATLAB to the ESP32 microcontroller via serial communication, processed

DOI: https://doi.org/10.17509/coelite. v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

into load control logic displayed on the LCD, and sent to the monitoring website. Thus, the system can monitor the results of the Genetic Algorithm optimization in real-time, demonstrating that this method is capable of providing an efficient power supply control solution for a PLTBG system with multiple load groups.

3.5. Data Collection

Data collection was conducted using a 1:2:2 ratio (feces:water:rumen) in a 120-liter digester. Samples were collected every hour for a total of 8 hours over a 15-day period, specifically from December 1 to 15, 2025. From this series of tests, a total of 120 data points were obtained, each reflecting the temperature, gas presence, gas pressure, voltage, and current during the testing period.

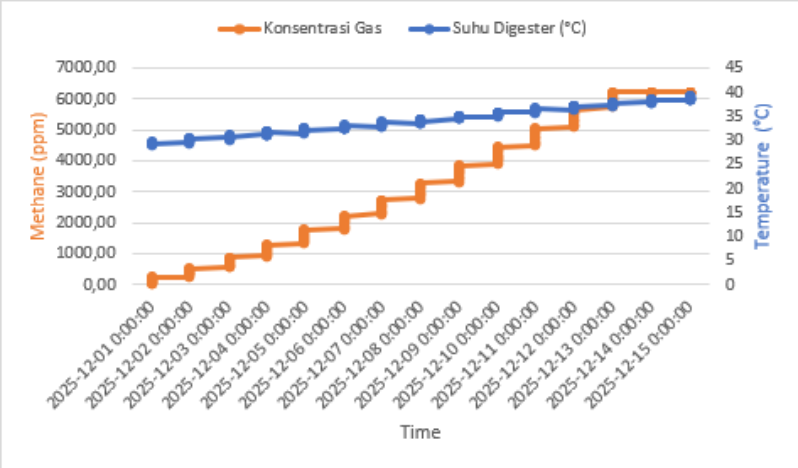


Figure 11. Correlation Between Methane Concentration and Digester Temperature

Figure 11 graph shows a positive correlation between methane production and temperature. As methane gas production increases, the temperature within the system also tends to rise. The anaerobic processes occurring within the digester across all treatments took place within a temperature range of ±29 °C to 39 °C, known as the mesophilic range. This temperature range is considered optimal for the activity of biogas-producing microorganisms, although various literature sources offer differing views regarding the ideal temperature for maximum biogas production.

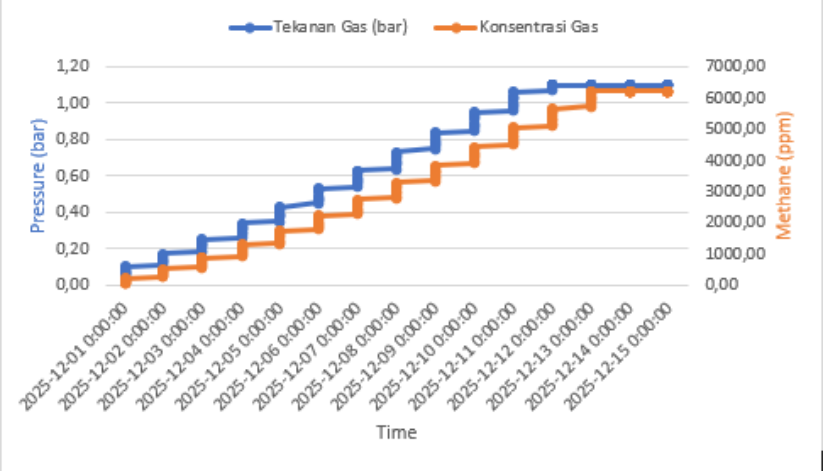


Figure 12. Correlation Between Gas Pressure and Methane Concentration

DOI: https://doi.org/10.17509/coelite.v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

The **Figure 12** graph shows that pressure has a positive correlation with methane production. As methane production in the biodigester increases, pressure also rises slightly, although this does not significantly affect the course of the anaerobic process. This increase in pressure reflects the accumulation of methane produced by the ongoing anaerobic fermentation process. Therefore, pressure is considered one of the key parameters in assessing the success of the biogas production process.

4. CONCLUSION

Based on the results of testing and research conducted on the optimization and monitoring of biogas power plants using Genetic Algorithms and the Internet of Things, it can be concluded that an Internet of Things (IoT)-based biogas power plant monitoring system can be effectively designed and utilized to monitor parameters such as gas pressure, temperature, methane content, and electrical parameters in real time. Additionally, this system can display measurement data remotely, which facilitates the monitoring and supervision of system performance. On the other hand, the energy management system in the biogas power plant is capable of effectively regulating the power supply combination to loads A, B, and C in accordance with the plant's capacity. The use of Genetic Algorithms to calculate the supply percentage for each load yields good results, where the total power generated can approach the plant's capacity of 1.1 kW without causing an overload condition. Consequently, the resulting load supply combination is in an optimal state, and the designed energy management system is capable of maximizing power utilization from the plant. In addition, the results of testing the Internet of Things-based monitoring system indicate that all major components function properly, although there are some limitations with certain sensors that may affect measurement accuracy; however, overall, the designed system has met the research objectives and can be used to monitor and manage an IoT-based biogas power plant. Furthermore, the optimization process conducted using the Genetic Algorithm showed that the best fitness values improved in quality with each generation until optimal conditions were reached when the total power supplied approached the maximum generator capacity.

5. ACKNOWLEDGMENT

The author would like to express sincere gratitude to all my friends, especially the 021 group members who accompanied me throughout my college journey. Special thanks are also extended to the academic advisors and stakeholders who provided guidance and support throughout this research process.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

7. REFERENCES

[1] International Energy Agency, "World Energy Outlook 2023 | Enhanced Reader," hal. 23–28, 2023, [Daring]. Tersedia pada: https://www.iea.org/news/the-energy-world-is-set-to-change-significantly-by-2030-based-on-today-s-policy-settings-alone

[2] N. F. T. Danun, G. P. Pasak, B. M. Simak, dan J. Andi, "Potensi Pengembangan Biomassa

Sebagai Alternatif Energi Terbarukan," *Bakar*, no. 1, 2022, [Daring]. Tersedia pada: https://fisika.ukitoraja.ac.id/wp-content/uploads/2022/04/17.pdf

[3] Asminar, A. Suyuti, S. M. Said, dan Syafaruddin, "Biogas Control As a Backup System on a PV/Wind Hybrid Microgrid With a Battery System," *ICIC Express Lett. Part B Appl.*, vol. 15, no. 1, hal. 1–9, 2024, doi: 10.24507/icicelb.15.01.1.

[4] Asminar, A. Suyuti, S. M. Said, and Syafaruddin, "Modelling and simulation of microgrid systems for powering a slaughterhouse," *AIP Conf. Proc.*, vol. 2952, no. 1, 2024, doi: 10.1063/5.0211996.

[5] Asminar, Abdul Djohar, Mansur, dan Bunyamin, "Pemodelan dan Simulasi PLTS , PLTBG Off Grid pada Rumah Potong Hewan Kota Kendari," hal. 4–7, 2024.

[6] Asminar, "DISERTASI Desain Koordinasi Dan Optimisasi Pembangkit Listrik Berbasis Energi Terbarukan Dengan Sistem Baterai Dalam Konteks Microgrid Design Coordination and Optimization of Renewable Energy-Based Power Plants with Battery Systems in The Context of Micro," 2023.

[7] M. Junus, A. D. Rahman, dan R. F. Shodiq, "Micro-Controller Based Biogas Production Monitoring System," vol. 8, no. 2, hal. 69–75, 2025.

[8] R. A. Nugraha, Y. H. Arita, M. Rifqy, A. Marendra, dan A. Aria, "Desain sistem pemantauan tekanan digester biogas berbasis internet of things studi kasus: digester biogas 10 m3 di Desa Rejosari, Kabupaten Demak," *J. Teknol.*, vol. 12, no. 1, hal. 68–78, 2024.

[9] E. A. Priantoro, E. S. Butar Butar, dan T. Sembiring, "PLTBg investment analysis from agricultural organic waste to produce electrical energy: preliminary potential analysis study and investment in the Caringin Main Market, Bandung," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1388, no. 1, 2024, doi: 10.1088/1755-1315/1388/1/012035.

[10] R. N. Utami, "Estimasi Potensi Pembangkit Listrik Tenaga Biogas (PLTBG) Dari Palm Oil Mill Effluent (POME) Pabrik Kelapa Sawit Pt Bina Mitra Makmur, Bungo," hal. 1–60, 2023.

[11] M. Mahmoodi-Eshkaftaki dan G. Mockaitis, "Structural optimization of biohydrogen production: Impact of pretreatments on volatile fatty acids and biogas parameters," *Int. J. Hydrogen Energy*, vol. 47, no. 11, hal. 7072–7081, 2022, doi: 10.1016/j.ijhydene.2021.12.088.

[12] L. G. S. A.-A. Michelle C. Almendrala, Kyle Adrienne T. Valenzuela, Steffany Marie Niña B. Santos, "Enhanced Biogas Production via Anaerobic Co-Digestion of Slaughterhouse and Food Waste Using Ferric Oxide as a Sustainable Conductive Material Michelle C. Almendrala *," 2025, doi: https://doi.org/10.48550/arXiv.2505.04635.

[13] R. Alieffiaul, Jamaluddin, dan Azhar, "Rancang bangun sistem kendali proses produksi biogas," vol. 8, no. 1, hal. 1–7, 2024, doi: http://dx.doi.org/10.30811/tektro.v8i1.5662.

[14] J. Full *et al.*, "Perspectives of Biogas Plants as BECCS Facilities: A Comparative Analysis of Biomethane vs. Biohydrogen Production with Carbon Capture and Storage or Use (CCS/CCU)," *Energies*, vol. 16, no. 13, 2023, doi: 10.3390/en16135066.

[15] J. Z. Yan, P. Kumar, and W. Rauch, "Effect of turbulent diffusion in modeling anaerobic digestion," *Bioresour. Technol.*, vol. 393, no. November 2023, hal. 130068, 2024, doi: 10.1016/j.biortech.2023.130068.

[16] J. Zulfadli dan H. Habibullah, "Perancangan Sistem Kontrol Kelistrikan Otomatis Kompor Listrik Halogen Berbasis Mikrokontroler," *JTEIN J. Tek. Elektro Indones.*, vol. 2, no. 1, hal. 119–127, 2021, doi: 10.24036/jtein.v2i1.140.

[17] B. P. Ginting, D. Setiawan, dan S. Yakub, "Implementasi IOT Pada Ruang Tangki Gas Menggunakan Pressure Transmitter Berbasis NodeMCU," vol. 3, hal. 109–117, 2024.

DOI: https://doi.org/10.17509/coelite.v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

[18] D. Savitri, N. W. Hidajati, dan H. S. Panigoro, "Implementasi algoritma genetika dalam mengestimasi kepadatan populasi jackrabbit dan coyote," *Jambura J. Biomath.*, vol. 3, no. 1, hal. 23–28, 2022, doi: 10.34312/jjbm.v3i1.11935.

[19] N. Anasta, I. Kesuma, dan A. Wardana, "Optimalisasi Sistem Distribusi Tenaga Listrik Menggunakan Algoritma Genetika," *J. Elektron. List. Dan Teknol. Inf. Terap.*, vol. 6, no. 2, hal. 73–77, 2024, doi: 10.37338/elti.v6i2.389.

[20] R. Suhartina, S. I. P, A. Firdaus, N. Ferdana, dan A. Hidayat, "Evaluasi Kinerja Sensor BH1750 pada Sistem Monitoring Iluminansi berbasis Mikrokontroler," vol. 07, no. 04, hal. 885–892, 2025.

DOI: https://doi.org/10.17509/coelite. v5i1.98498
p- ISSN 2829-4157 e- ISSN 2829-4149

Journal of Computer Engineering, Electronics and Information Technology (COELITE)

Journal homepage: https://ejournal.upi.edu/index.php/COELITE

Implementation of the Jaro-Winkler Algorithm for Automatic Grading of Short Answer Questions in the Kifayatul Azhar Elementary School Exam

Rizki Halomoan[1], Benny Yoga[2], Nauval Ardana[3], Trisna Gelar[4], Santi Sundari[5]*

[1,2,3,4,5]Department of Computer and Informatics Engineering, Politeknik Negeri Bandung, Indonesia
Correspondence E-mail: trisna.gelar@polban.ac.id

ABSTRACT

Kifayatul Azhar Primary School faces significant obstacles in examination administration. High logistical costs and inefficient social media coordination hamper physical exam distribution. Existing digital tools, such as Google Forms, cannot adapt to minor typographical errors in short-answer evaluations. This research develops a web-based automated grading system using the Jaro-Winkler algorithm and the Rational Unified Process (RUP) framework. The RUP methodology manages iterative development through four stages: Inception, Elaboration, Construction, and Transition. The system engine performs text preprocessing tasks such as case folding and tokenization prior to calculating Jaro-Winkler similarity scores. Experimental results indicate that $d_w \geq 0.8$ serves as the optimal threshold for differentiating acceptable typographical variants from substantive errors. The system validation process included 26 integration test cases and 76 User Acceptance Test (UAT) cases, encompassing four distinct user roles: Admin, Teacher, Mentor, and Parent. All tests demonstrated a complete functional success rate. This system provides a centralized, automated environment that reduces administrative workload while maintaining grading objectivity.

© 2026 Universitas Pendidikan Indonesia

ARTICLE INFO

Article History:
Submitted/Received 05 Mar 2026
First Revised 17 Mar 2026
Accepted 29 Mar 2026
First Available online 01 Apr 2026
Publication Date 01 Apr 2026

Keyword:
Automatic Grading,
Elementary Education,
Jaro-Winkler Algorithm,
Rational Unified Process (RUP),
Short Answer Questions.

1. INTRODUCTION

Kifayatul Azhar Primary School implements a flexible learning model through online-based Home Schooling.[1] Talent Class Program. The school organizes this program to serve 157 students scattered throughout various regions in Indonesia. For instance, students outside Java Island participate in all academic activities independently via digital platforms under parental supervision. The existence of this specialized program demands a support system capable of crossing geographical boundaries efficiently.[2], [3].

Digitalization of the evaluation system becomes the primary solution to reduce school operational costs.[4], [5]. The school management has been sending physical exam papers (hardcopy) to each student's home address using expedited services. This procedure causes a surge in logistics costs and delays in question distribution during Midterm (UTS) and Final Exams (UAS). The implementation of web-based exams can cut the physical distribution chain so that the evaluation process runs faster and more affordably.[6].

Online exam execution faces technical obstacles in the automation of short-answer scoring. [7]. The Google Form technology currently used only accommodates accurate assessment for multiple-choice question types. Teachers frequently find cases where the system absolutely penalizes student answers simply due to differences in capitalization or typographical errors (typos)[8]. For example, the system will award a score of zero to the answer "matahari" if the set answer key is "Matahari", even though the substance is identical.

The limitations of said platform increase the administrative workload for educators.[4]. Teachers must manually correct thousands of short-answer entries in subjects such as Islamic Education (PAI), Science (IPA), and Social Studies (IPS) to maintain assessment objectivity.[9]. Previous researchers have explored various string-matching algorithms to handle text writing variations automatically.[10]. One study indicates that the use of the N-Gram method can improve word error detection accuracy in Indonesian text by up to 83.64%[11].

This research proposes the Jaro-Winkler algorithm as an automatic correction engine for short answers.[12]. The Jaro-Winkler algorithm calculates the similarity level between two strings by giving more weight to character matches at the beginning of the word (prefix matching). For example, the system will still provide a high similarity score for the word "Indonesia" even if the student types "Indoensia" due to the similarity of its first four characters. The researcher utilizes the Rational Unified Process (RUP) method in building this application to ensure the system's functionality meets specific field requirements.[13].

The developed application aims to automate assessments in eight core subjects, such as PAI, PPKN, and IPA. This software focuses the evaluation process on string similarity comparison and does not involve semantic or meaning analysis.[14]. The application of this algorithm is expected to significantly cut teacher correction time without ignoring tolerance for student typing errors. This step serves as a crucial strategy for SD Kifayatul Azhar in optimizing the efficiency of educational evaluation and improving corrective feedback in the digital era.[15].

2. METHODS

This research utilizes the Rational Unified Process (RUP) framework to manage the software development life cycle iteratively. The selection of RUP allows the researcher to align the specific needs of Kifayatul Azhar Primary School with the development of the Jaro-Winkler algorithm through four main phases. Figure 1 illustrates the iterative development lifecycle of the system following the RUP framework. The process begins with the Inception phase, focusing on stakeholder requirement analysis at SD Kifayatul Azhar. It proceeds to the

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149

The development process begins with the Elaboration phase, where the system architecture and Jaro-Winkler algorithmic logic are meticulously designed. This architectural blueprint, which serves as the foundation for the entire development lifecycle, is illustrated in **Figure 1**. Following the design, the Construction phase is executed through five distinct iterations, ensuring each module is validated through rigorous unit and integration testing. Finally, the Transition phase involves field implementation and threshold calibration ($d_w \geq 0.8$) through User Acceptance Testing (UAT) to ensure the system meets the required accuracy and performance standards in a real-world environment.

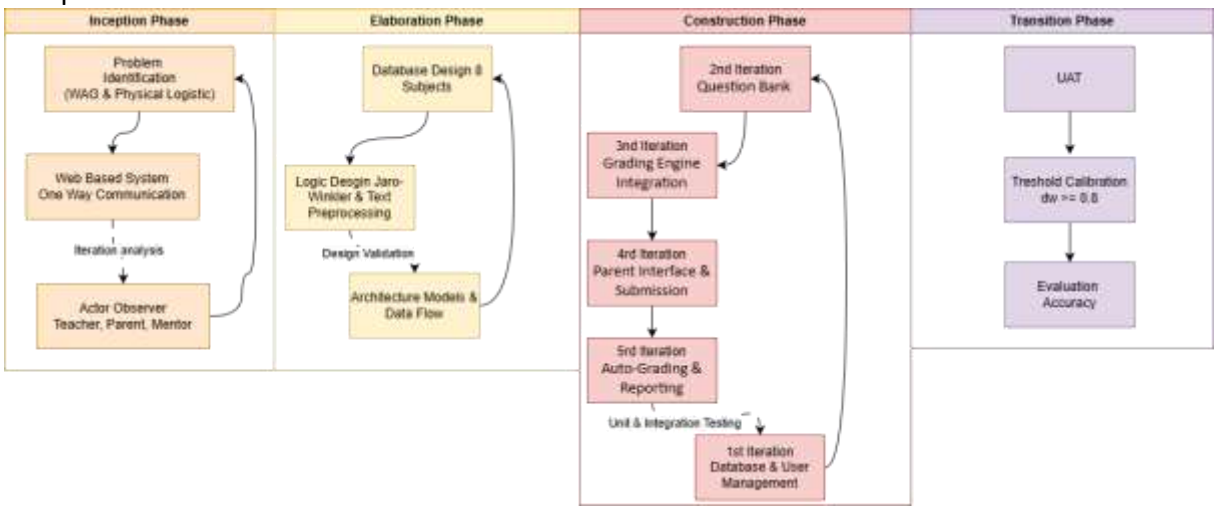


Figure 1. The four phases of the Rational Unified Process (RUP) framework were applied in the development of the automatic grading system.

2.1. Inception Phase

This phase establishes the project boundaries by observing the interactions among three main actors: Teachers, Mentors, and Parents. The learning system in the Talent Class Program at SD Kifayatul Azhar relies on mentors as an information bridge to distribute teaching modules and exam questions from teachers to parents. Current coordination activities rely entirely on WhatsApp Groups (WAG), leading to disorganized message stacks. Technical issues arise regarding the documentation of exam links and modules, which are frequently lost in conversations, making it difficult for teachers to conduct centralized grade recording at the end of the semester.

An in-depth analysis was conducted on the Midterm (UTS) and Final Exam (UAS) procedures, which require the distribution of physical exam papers. Mentors must prepare printed packages and use expedition services to reach 157 students scattered throughout Indonesia. This procedure not only burdens the school with logistics costs but also demands extra workload for mentors to ensure the delivery status of each package. Furthermore, the assignment submission method via photo uploads in WAG is highly ineffective, as mentors must manually verify the validity of files one by one. This condition prolongs the verification process for task submissions in every class.

Based on these findings, the system requirements are focused on a web-based application capable of providing one-way communication from mentors to parents regarding exam schedules and thematic activities. The system must accommodate answer inputs by parents according to the duration determined by the school. Key identified features include digital

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149

exam submission management, a question bank with various types (multiple choice, short answer, and essay), and an automatic grade recording function. This phase concludes that the digitalization of exams is not only a solution to reduce paper distribution costs but also a strategy to improve academic documentation management that was previously fragmented across social media.

2.2. Elaboration Phase

This phase matures the system design through the integration of string-matching logic and database architecture. Activities began with modeling the data flow that connects answer inputs from the parent interface to the automatic grading engine. The architectural structure is designed to manage question variations from eight core subjects, including Islamic Religious Education (PAI), Science (IPA), and Civics (PPKN). The output of this phase consists of database schemas and sequence diagrams that ensure every answer input undergoes a validation process before being calculated by the algorithm.

The short-answer evaluation process applies systematic Text Preprocessing stages. Student answers are processed through two main steps: case folding and tokenization.[10]. Case folding converts all characters to lowercase to eliminate case sensitivity, which was a weakness in the previous platform. Subsequently, tokenization separates words based on spaces and punctuation to facilitate the identification of text units for comparison. Stop-word removal and stemming are not implemented in this system to maintain the integrity of the original words.[16], which often serve as primary keywords in elementary school students' short answers.

The Jaro-Winkler algorithm serves as the core component in determining the similarity score between student answers and the answer key. The calculation begins by determining the Jaro Distance (dj), which considers the number of matching characters (m) and the number of transpositions (t) between two strings (S_1 and S_2):

$$j \quad \frac{1}{3} x \left(\frac{m}{S1} + \frac{m}{S2} + \frac{m-t}{m} \right) \tag{1}$$

The system enhances grading accuracy by applying weighting to the initial characters (prefix matching). The Jaro-Winkler formula d_w provides an additional score based on the length of the common prefix (l) using a constant scaling factor (p=0.1).

$$d_w = d_j + (p (1 - d_j) \tag{2}$$

For instance, if a student's answer contains a typographical error at the end of the word but is correct at the beginning, the algorithm still yields a high score. Algorithmic simulations were conducted on actual answer samples from Talent Class students to determine the assessment threshold. The simulation results indicate that a value of d_w≥0.8 is the optimal threshold for the 'Almost Correct' category tolerable by the system. This stage concluded with the creation of a functional mockup to verify the grading workflow before proceeding to full-scale coding.

2.3. Construction Phase

The system development was carried out gradually through five main iterations. Each iteration produced functional modules that integrated academic management requirements

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149

with the automatic grading algorithm. The development focus during this phase was divided into several activities, which are summarized in **Table 1**.

Table 1. Development Iterations in the Construction Phase

Iteration	Development Focus	Main Activity
1	Database & User Management	Designing the schema for eight subjects and implementing a multi-actor authentication system (Teacher, Mentor, Parent).
2	Question Bank Module	Developing features for inputting and managing various question types: multiple-choice, short answer, and essays.
3	Grading Engine Integration	Integrating the Jaro-Winkler algorithm with the text preprocessing workflow (case folding and tokenization).
4	Parent Interface & Submission	Developing the digital exam submission portal, including real-time countdown timers and answer input handling.
5	Auto-Grading & Reporting	Implementing the automated scoring logic for short answers and generating centralized academic reports.

All modules developed within these five iterations underwent unit testing and integration testing. This process ensured that data synchronization between student answer inputs and the teacher's grade dashboard operated without logical conflicts during mass access. The use of these iterations guarantees that every feature, from the question bank to the automatic grading engine, was validated before proceeding to the field implementation stage.

2.4. Transition Phase

The Transition Phase focuses its activities on final verification and system handover to the users. The primary objective is to ensure that the application can stably and accurately replace the physical exam distribution procedures. Validation was conducted through a User Acceptance Test (UAT) mechanism involving Teachers, Mentors, and Parents as end-users. For instance, users tested the entire workflow, from the distribution of exam schedules by mentors to the submission of answers by parents, to ensure the system interface is easily understood.

Testing the accuracy of the automated grading is a crucial parameter in this validation stage. Teachers performed a comparison between the scores generated by the Jaro-Winkler algorithm and the results of manual assessments across eight core subjects. This activity aimed to calibrate the similarity threshold to ensure fair grading for students despite typographical errors. For example, the system is declared valid if it can recognize variations in short answers that contain typos but maintain the correct substance according to the answer key.

3. RESULTS AND DISCUSSION

3.1. Analysis of Requirements (Inception & Elaboration Outcomes)

As show in **Table 2,** The initial analysis identified that the primary challenge at SD Kifayatul Azhar was the high logistical cost of physical exams and the inefficiency of manual verification via social media. The system was designed to handle eight core subjects.

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149

Table 2. Sample of Short Answer Keys for Assessment Calibration

Subject	Questions	Answer
PPKN	What is the first principle of the Pancasila?	Ketuhanan yang maha esa
PAI	What is the second source of Islamic law after the Al-Quran?	Hadist
Science (IPA)	What is the primary source of light and energy for the Earth?	Matahari
Science (IPA)	What type of mirror is used for vehicle rearview mirrors?	Cembung
Art (SBdP)	What is the term for the speed of a musical beat?	Tempo
Art (SBdP)	Art that has length, width, and height (volume) is called...	3D
Social Studies	Who is the authority responsible for regulating road traffic?	Polisi lalu lintas
PAI	What is the meaning of the word 'Zikir' in a literal sense?	Ingat
PJOK	What is the starting position for running?	Berdiri
Indonesian	What is the antonym of the word "soft" (lembut)?	Keras

The architectural design of the system was matured during the Elaboration phase to ensure seamless integration between the database and the grading engine. **Figure 2** presents the class diagram that defines the relationships between the Teacher, Mentor, and Parent actors.

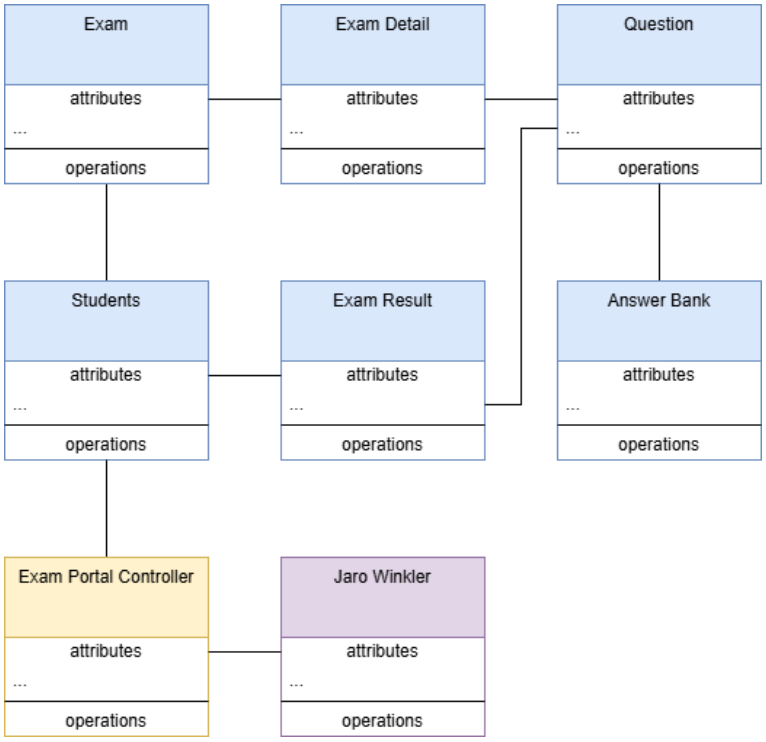


Figure 2. Class Diagram of the Automated Grading System showing the structural relationship between users and the assessment modules.

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149

3.1.1 Algorithmic Comparison and Accuracy Analysis

Based on the data provided by the partner regarding student answers during the final semester evaluation, an analysis was conducted to determine the appropriate method for correcting correct or incorrect answers. Several student answers contained typographical errors, leading to a series of experiments. The system's effectiveness was tested by comparing student inputs that contained typographical errors against the gold standard keys. Table 3 illustrates how the Jaro-Winkler algorithm maintains a high similarity score (d_w) even when student input is slightly inaccurate.

Table 3. Comparison of Jaro-Winkler Scores (dw) for Student Answer Variations

No	Answer Key (S_1)	Student Input (S2)	d_w Score	Assessment Status
1	Matahari	matahari	1.00	Correct (Case folding applied)
2	Matahari	matahair	0.958	Correct (Minor typo)
3	Hadis	Hadis	1.00	Correct
4	Hadis	Habis	0.866	Correct (Minor typo)
5	Cembung	Cembun	0.94	Correct (Missing character)
6	Tempo	Temppo	0.93	Correct (Double character)
7	Ketuhanan yang maha esa	Ketuhanan maha esa	0.88	Incorrect (Missing word)
8	Ketuhanan yang maha esa	Kemanusiaan yang adil dan beradab	0.655	Incorrect (Substantive error)
9	keras	pelan	0.600	Incorrect (Substantive error)
10	ingat	lupa	0.483	Incorrect (Substantive error)
11	Polisi lalu lintas	Polis lalu lintas	0.904	Correct (Minor typo)
12	3D	3 dimensi	0.750	Incorrect (System)

As seen in **Table 3**, the Jaro-Winkler algorithm performs exceptionally well in handling student typos. For the answer "Hadis" instead of "Hadist," the system yields a score of 0.96, which is well above the established threshold of dw≥0.8. This allows the system to grant a score of 100 to the student automatically. Conversely, for substantive errors like pelan (where the key is "keras"), the score drops significantly to 0.60, allowing the system to correctly mark the answer as wrong. A unique case was observed in the comparison between '3D' and '3 Dimensi'. Although both terms are semantically identical in the context of elementary school art subjects, the Jaro-Winkler algorithm yielded a low similarity score (dw≈0.75). This occurs because the algorithm calculates lexical similarity based on character sequences rather than semantic meaning. Since only the character '3' matches, the score falls below the 0.8 threshold. **Figure 3** showcases the standalone prototype developed to verify the Jaro-Winkler calculation logic.

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149



Figure 3. Standalone Prototype of the Jaro-Winkler Distance Calculator used for algorithmic verification during the Elaboration phase.

3.1.2 Analysis of Results

The experimental results indicate that the Jaro-Winkler algorithm effectively handles most student typographical errors. By setting a threshold of dw≥0.8, the system successfully identifies answers that are lexically like the gold standard, such as "Hadis" and "Cembun," which would otherwise be marked incorrect by standard exact-match systems. This flexibility significantly reduces the administrative burden on mentors at SD Kifayatul Azhar, allowing for a more objective and faster grading process.

Despite its effectiveness in handling typos, this study identifies a significant limitation regarding semantic variation. As shown in Table 3 (No. 12), the pair "3D" and "3 Dimensi" yielded a low similarity score of 0.53 despite being semantically identical. This occurs because the Jaro-Winkler algorithm operates solely on character sequences (lexical matching) and does not possess natural language understanding (semantic matching).

Consequently, for subjects requiring technical abbreviations or synonyms, the current system still requires a "multiple-key" database approach where teachers input various acceptable variations (e.g., both "3D" and "3 Dimensi") as separate gold standards to ensure fair grading.

$$True, \quad similarity = 1 \; Expert \; Judgment, \quad 0.8 < similarity < 1 \; False, \quad similarity \leq 0.8$$

3.2. System Implementation and Iterative Development (Construction Outcome)

The Construction phase focused on the phased implementation of the system requirements through five distinct iterations. This approach ensured that the core grading engine was stable before the user interface was finalized.

3.2.1 Functional Iterations

The development was divided into functional modules as detailed previously in **Table 1**. The early iterations established a robust database capable of storing diverse question types (Multiple Choice, Short Answer, and Essay) for eight core subjects. By Iteration 3, the Jaro-Winkler algorithm was integrated into the grading logic, preceded by a mandatory text preprocessing pipeline, as illustrated in **Figure 4** below. This integrated workflow ensures that the input data is properly refined before the algorithm calculates the final similarity Scores

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149



Figure 4. Integration of the Jaro-Winkler Algorithm in the Admin Question Bank Module

3.2.2 Automated Grading Workflow

The core functionality of the system involves comparing student submissions against the stored "Gold Standard" keys. **Figure 5** illustrates the system's ability to calculate similarity scores in real-time. During this phase, unit testing confirmed that the system correctly applied case-folding (e.g., converting "cembung" and "kembung" to a common format) to avoid unnecessary grading penalties.



Figure 5. Real-time Jaro-Winkler Similarity Calculation during the Grading Process

3.2.3 Data Synchronization and Integration Testing

During the final iterations of the Construction phase, the system underwent rigorous Unit Testing and Integration Testing to ensure data synchronization across all modules. A total of 26 test cases were executed using the Black Box method, covering authentication, data management, and the core system functionalities, as summarized in **Table 4** below. This structured evaluation confirmed that each feature performs according to the specified requirements before moving into the deployment stage.

Table 4. Summary of Functional and Integration Testing Results

Test Code	Module / Function	Testing Objective	Status
UT-01 to UT-05	Authentication	Validating login security for Admin, Teacher, and Parents.	PASS
IT-06 to IT-15	User & Data Mgmt	Ensuring seamless data flow between student profiles and class assignments.	PASS

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149

IT-16 to IT-21	Question Bank	Verifying the storage and retrieval of Short Answer keys for 8 subjects.	PASS
IT-22 to IT-26	Grading Engine	Testing the synchronization between student input and automatic d_w calculation.	PASS

The integration tests confirmed that the data entered by Parents (student answers) was accurately captured and processed by the Jaro-Winkler engine. As shown in the test results (e.g., IT-23), modifications to the question bank were immediately reflected in the grading dashboard without logical conflicts. This synchronization is critical for the Transition phase, as it ensures that the teacher's grade recap is always based on the most recent and valid student submissions.

3.3. Final Validation and User Acceptance (Transition Outcome)

The Transition phase concluded with a comprehensive User Acceptance Test (UAT). The objective was to ensure that the system met the operational needs of all four primary actors at SD Kifayatul Azhar. A total of 76 test cases were executed across different user roles, covering end-to-end workflows from exam scheduling to automated grading. The results of these evaluations, as detailed in **Table 5**, demonstrate a 100% success rate in functional execution, with 90.7% of cases passing flawlessly and 9.3% passing with minor UI/UX notes (such as minor layout adjustments).

Table 5. Summary of User Acceptance Test (UAT) Results

No	User Role	Total Test Cases	Status: Pass
1	Teacher	24	21
2	Parent	13	13
3	Admin	12	9
4	Mentor	27	26

Specifically, the Teacher role confirmed that the automated grading correctly applied the Jaro-Winkler logic, and the Parent role successfully submitted answers without technical hurdles. The notes provided by the Admin and Mentor roles were primarily focused on decorative interface elements, which did not affect the core functionality of the exam distribution or the accuracy of the similarity scores. This phase confirms that the system is ready for full-scale implementation, effectively replacing the previous manual WhatsApp-based coordination with a structured, automated academic environment.

4. CONCLUSION

This study employed a web-based automated grading system at Kifayatul Azhar Primary School, utilizing the Rational Unified Process (RUP) and the Jaro-Winkler algorithm. The system was designed to mitigate the significant logistical expenses and disorganized coordination that had previously been managed through WhatsApp Groups. The resulting platform effectively centralized examination management and enhanced the precision of documentation. Experimental findings validated that a similarity threshold of $dw \geq 0.8$ offered the most effective equilibrium for identifying student typographical errors. Furthermore, validation through 76 User Acceptance Test (UAT) cases revealed a 100% functional success rate across all user roles. Although the Jaro-Winkler algorithm facilitates objective lexical assessment, it is constrained by its inability to recognize semantic synonyms. Consequently,

DOI: https://doi.org/10.17509/coelite.v5i1.98351
p- ISSN 2829-4157 e- ISSN 2829-4149

future iterations should incorporate natural language processing to improve the system's comprehension of contextual variations within student responses.

5. ACKNOWLEDGMENT

The authors would like to express their gratitude to Politeknik Negeri Bandung (POLBAN) for the institutional support provided during this research. We also extend our sincere appreciation to the faculty and staff of SD Kifayatul Azhar for their collaboration as research partners and for providing the necessary data and environment for the system's implementation and testing.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

7. REFERENCES

[1] Nor Annisa, Noor Padilah, Reni Rulita, Renny Yuniar, and Nita Priyanti, "Model Pembelajaran Homeschooling Pendidikan Anak Usia Dini," *EDUKASIA J. Pendidik. dan Pembelajaran*, vol. 4, no. 1, pp. 89–100, Aug. 2022, doi: 10.62775/edukasia.v4i1.106. Available: https://jurnaledukasia.org/index.php/edukasia/article/view/106

[2] M. Munawarah, I. Ikhsan, L. N. W. Fajzrina, R. Aulia, and C. N. Fadillah, "Strategy for Accepting New Students Post-COVID-19: A Case Study," *Al-Athfal J. Pendidik. Anak*, vol. 8, no. 2, pp. 105–116, Mar. 2023, doi: 10.14421/al-athfal.2022.82-03. Available: https://ejournal.uin-suka.ac.id/tarbiyah/alathfal/article/view/6330

[3] M. Lutviatiani, N. C. Shofiyah, C. F. Wijaksono, and A. A. Zawawi, "The Implementation of an Online Distance Learning Program with an Independent Curriculum-Based on Improving Child-Parent Relations," *Educ. J. Ilm. Pendidik.*, vol. 1, no. 2, pp. 51–62, Jul. 2023, doi: 10.37985/educative.v1i2.56. Available: https://publikasi.abidan.org/index.php/educative/article/view/56

[4] H. W. Dhany and S. Sapriadi, "Pengembangan Perangkat Lunak Penilaian Otomatis Ujian Pilihan Ganda Menggunakan Algoritma String Matching," *J. Minfo Polgan*, vol. 14, no. 2, pp. 2768–2774, Nov. 2025, doi: 10.33395/jmp.v14i2.15452. Available: https://jurnal.polgan.ac.id/index.php/jmp/article/view/15452

[5] M. Sabri and A. Wais, "Artificial Intelligence dalam Pengukuran dan Penilaian Pendidikan," *J. Eval. Pendidik*, vol. 16, no. 2, pp. 95–107, Oct. 2025, doi: 10.21009/jep.v16i2.60738. Available: https://journal.unj.ac.id/unj/index.php/jep/article/view/60738

[6] D. P. Sidik, A. Rozak, F. Fatqurhohman, and H. Fatkurochman, "Literature Review of Artificial Intelligence in Learning: Trends and Opportunities," *RESET Rev. Educ. Sci. Technol.*, vol. 1, no. 1, pp. 43–54, Aug. 2025, doi: 10.66031/reset.v1i1.19. Available: https://jurnal.ihsancahayapustaka.id/index.php/reset/article/view/19

[7] K. Harefa and A. J. Lubis, "Aplikasi Sistem Automated Essay Scoring untuk Jawaban Soal Ujian dengan Menerapkan Algoritma Jaro-Winkler," *J. Ilmu Komput. Dan Sist. Komput. Terap.*, vol. 4, no. 2, pp. 109–119, 2022, doi: https://doi.org/10.35447/jikstra.v4i2.611

[8] A. R. Amri, C. Charlina, and M. Sinaga, "Kesalahan Pemakaian Ejaan Bahasa Indonesia

dalam Soal Try Out Ujian Sekolah Berstandar Nasional Sekolah Dasar Tahun Pelajaran 2017/2018," *J. TUAH Pendidik. dan Pengajaran Bhs.*, vol. 2, no. 1, pp. 1–9, 2020, doi: https://dx.doi.org/10.31258/jtuah.2.1.p.1-9

[9] N. S. Lagutina and K. V. Lagutina, "A Survey of Models for Automatic Assessment of Similarity of Students' Answer to the Reference Answer," *Autom. Control Comput. Sci.*, vol. 59, no. 7, pp. 1152–1169, Dec. 2025, doi: 10.3103/S0146411625700427. Available: https://link.springer.com/10.3103/S0146411625700427

[10] Y. Rochmawati and R. Kusumaningrum, "Studi Perbandingan Algoritma Pencarian String dalam Metode Approximate String Matching untuk Identifikasi Kesalahan Pengetikan Teks," *J. Buana Inform.*, vol. 7, no. 2, Jan. 2016, doi: 10.24002/jbi.v7i2.491. Available: https://ojs.uajy.ac.id/index.php/jbi/article/view/491

[11] H. Jayadianti, B. Santosa, J. Cahyaning, S. Saifullah, and R. Drezewski, "Essay Auto-Scoring Using N-Gram And Jaro Winkler Based Indonesian Typos," *MATRIK J. Manajemen, Tek. Inform. Dan Rekayasa Komput.*, vol. 22, no. 2, pp. 325–338, Mar. 2023, doi: 10.30812/matrik.v22i2.2473. Available: https://journal.universitasbumigora.ac.id/index.php/matrik/article/view/2473

[12] W. E. Winkler, "Record Linkage," 2009, pp. 351–380. doi: 10.1016/S0169-7161(08)00014-X. Available: https://linkinghub.elsevier.com/retrieve/pii/S016971610800014X

[13] T. K. Tia, I. Nuryasin, and M. Maskur, "Model Simulasi Rational Unified Process (RUP) Pada Pegembagan Perangkat Lunak," *J. Repos.*, vol. 2, no. 4, Jan. 2024, doi: 10.22219/repositor.v2i4.30511. Available: https://ejournal.umm.ac.id/index.php/repositor/article/view/30511

[14] A. Ayaan and K.-W. Ng, "Automated Grading Using Natural Language Processing And Semantic Analysis," *MethodsX*, vol. 14, p. 103395, Jun. 2025, doi: 10.1016/j.mex.2025.103395. Available: https://linkinghub.elsevier.com/retrieve/pii/S2215016125002419

[15] I. M. A. Purwadi, "Pengaruh Pemberian Umpan Balik Dari Teman Sejawat Terhadap Kemampuan Berpikir Kritis Dan Komunikasi Matematis Siswa Smp Laboratorium Undiksha Singaraja," *J. Pendidik. Mat. Undiksha*, vol. 11, no. 1, pp. 63–72, 2020, doi: https://doi.org/10.23887/jjpm.v11i1.24382

[16] T. Mardiana, T. Bharata Adji, and I. Hidayah, "Stemming Influence on Similarity Detection of Abstract Written in Indonesia," *TELKOMNIKA (Telecommunication Comput. Electron. Control*, vol. 14, no. 1, p. 219, Mar. 2016, doi: 10.12928/telkomnika.v14i1.1926. Available: http://telkomnika.uad.ac.id/index.php/TELKOMNIKA/article/view/1926

Journal of Computer Engineering, Electronics and Information Technology (COELITE)

Journal homepage: https://ejournal.upi.edu/index.php/COELITE

Double Proxy Architecture for Data Center DDoS Mitigation

Rahmawan Bagus Trianto[1], Kukuh Muhammad[2], Dwi Novia Prasetyanti[3]*

[1,2,3] Department of Informatics Engineering, Politeknik Negeri Cilacap, Indonesia
Correspondence E-mail: rahmawanbagustrianto@pnc.ac.id

ABSTRACT

The increasing frequency and complexity of Distributed Denial-of-Service (DDoS) attacks present serious challenges to data center availability. Conventional single-layer defense mechanisms are often insufficient to handle high-volume attacks, resulting in excessive resource consumption on critical servers. This study proposes a data center security architecture based on a double proxy system to enhance resilience. The architecture consists of two sequential reverse proxy layers, namely a public-facing proxy and an internal proxy, which collaboratively perform traffic filtering, request validation, and load isolation before forwarding requests to the application servers. Experimental evaluations were conducted under simulated TCP- and UDP-based DDoS attack scenarios to assess system performance. The results indicate that the proposed double proxy architecture significantly reduces CPU utilization by up to 76% and decreases network bandwidth consumption by up to 352 Mbps compared to a traditional single-proxy configuration. Memory and disk usage remain stable during attack conditions, demonstrating minimal performance overhead. Furthermore, the architecture effectively absorbs malicious traffic, prevents direct exposure of backend servers, and preserves data integrity. These findings suggest that the double proxy approach provides a practical, efficient, and scalable solution for mitigating modern DDoS attacks in data center environments and can be adapted to evolving cyber threat landscapes.

ARTICLE INFO

Article History:
Submitted/Received 07 Feb 2026
First Revised 20 Feb 2026
Accepted 17 Mar 2026
First Available online 01 Apr 2026
Publication Date 01 Apr 2026

Keyword:
Data Center Security,
DDoS Mitigation,
Double Proxy,
Network Resilience,
Reverse Proxy.

© 2026 Universitas Pendidikan Indonesia

1. INTRODUCTION

A Distributed Denial of Service (DDoS) attack is a cyberattack designed to make an online service inaccessible by inundating it with a deluge of internet traffic. It's like flooding a highway, causing gridlock and preventing legitimate traffic from passing through [1]–[3]. The increasing sophistication of cyber threats, particularly Distributed Denial of Service (DDoS) attacks, has made data centers vulnerable to unprecedented levels of disruption. DDoS attacks exploit network bandwidth, processing power, or application resources to overwhelm targeted systems and render them unresponsive. The consequences of such attacks can be severe, including financial losses, damage to brand reputation, and loss of customer trust [4]–[6]. As a result, the need for effective DDoS mitigation strategies has become more critical than ever.

Existing single-proxy solutions offer basic DDoS protection but are increasingly ineffective against sophisticated volumetric and application-layer attacks. The need for a more robust, multi-layered defense is critical. This study aims to address this gap by proposing a double proxy architecture that enhances traffic filtering and reduces the burden on backend servers.

The application of proxy servers as a defense mechanism against Distributed Denial of Service (DDoS) attacks. Proxy servers act as intermediaries, filtering and inspecting inbound traffic to identify and block malicious requests, while forwarding legitimate traffic to the backend server. Furthermore, proxy servers can be configured to implement rate-limiting policies, thus preventing excessive traffic from reaching the server [7]–[10].

The research compares the performance of Network Load Balancing (NLB) and High Availability Proxy (HAProxy) in mitigating Distributed Denial of Service (DDoS) attacks, specifically SYN floods. NLB is utilized on Windows Server 2016 with IIS 10.0, while HAProxy is employed on Ubuntu 16.04 Linux with Apache 2. Experiments were conducted in a live network environment to evaluate the efficiency of each load balancer. Average response time and CPU utilization were used as performance metrics. The findings reveal that NLB on Windows platforms exhibits superior performance in mitigating SYN flood DDoS attacks compared to HAProxy on Linux platforms [4][11][12].

The other research at Torabelo Hospital, the current server system faces security and optimization challenges. This research analyzes the impact and recommends solutions to enhance server security and optimization. Research findings indicate that the server system is vulnerable to various types of attacks and performance degradation. This can negatively impact hospital operations and jeopardize patient care. The recommended solution is to implement Squid as a reverse proxy, WAF (Web Application Firewall), and Snort as an IDS (Intrusion Detection System). System testing demonstrates that this solution effectively detects and prevents various common attacks [13]–[15].

Next paper investigates the role of proxy servers in educational settings and proposes innovative enhancements to optimize their performance and functionality. The study explores various cache optimization strategies to accommodate diverse organizational requirements and introduces methods for automatically reducing student bandwidth consumption, thereby enabling more efficient allocation of bandwidth resources for administrative tasks. The paper focuses on improving cache management, bandwidth control, and network traffic management through user and thread management. A practical implementation of a basic proxy server, Peroxy, is presented, demonstrating the potential of integrating the proposed features into an academic proxy server. Peroxy showcases the

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

feasibility of efficient and effective Internet bandwidth management for institutions, preventing excessive bandwidth consumption by students and employees [8][16].

The study examines DDoS attacks that overload network resources and backend servers. A proxy-based architecture was implemented to manage traffic by dynamically adding servers during traffic spikes. However, attackers can redirect DDoS attacks to the proxy servers once their IP addresses are exposed. To address this issue, the study proposes an authentication mechanism to ensure balanced request distribution and block unauthorized requests. Simulation results show that the proposed approach effectively detects and mitigates DDoS attacks [17].

The other research employing the robust ModSecurity web application firewall (WAF) provides comprehensive protection against a wide range of web application attacks and facilitates HTTP traffic monitoring. ModSecurity has demonstrated its efficacy in mitigating DDoS attacks, achieving a 100% success rate when evaluated against three attack tools: Bash scripts, Golden Eye, and HULK. Furthermore, the implementation of a reverse proxy enhances server security by creating an additional layer of defense. While reverse proxies cannot prevent attacks, they can increase the time it takes for an attacker to reach the target, as evidenced by the average 53ms increase in attack duration using the three aforementioned tools. The combination of ModSecurity and a reverse proxy offers a dual-layered defense mechanism against DDoS attacks [18][19].

Data centers, as the backbone of modern digital infrastructure, are increasingly targeted by DDoS attacks that can disrupt critical services and cause significant operational impacts. While various mitigation techniques have been developed, most existing studies rely on single-layer proxy solutions. To address this limitation, this study proposes a double proxy architecture that combines a public reverse proxy and a local reverse proxy with a custom firewall [20]. This layered approach aims to improve system resilience by filtering malicious traffic before it reaches core systems and providing stronger protection against volumetric and application-layer DDoS attacks.

2. METHODS

This research aims to enhance the security of data centers and networks by proposing a novel architecture. The research process involved the following stages: conducting a literature review, designing the data center and network security architecture, implementing the designed architecture, testing the design, and evaluating the results. Through a comprehensive evaluation, the effectiveness of the proposed architecture in mitigating various cyber threats was assessed. **Figure 1** provides a visual representation of the research process. The steps involved in this research are as follows:



Figure 1. Research workflow.

Figure 1 illustrates the research methodology, which encompasses a literature review, data center and network design, implementation, testing, and evaluation of the data center security architecture's performance against DDoS attacks. A detailed explanation of each stage will be provided in the subsequent sections.

2.1. Literature Review

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

This research commences with a comprehensive literature review. The review serves to provide a thorough understanding of previous studies, identify knowledge gaps, and lay a solid foundation for the proposed research. The review focuses on the threats posed by DDoS attacks and the critical need for mitigation strategies in data centers to ensure the continued security of systems.

2.2. Data Center and Network Design

This study proceeded to design a secure data center and network architecture, innovating with a double proxy approach. This architecture utilizes both a public reverse proxy and a local reverse proxy, with an additional layer of security provided by a network firewall. **Figures 2 and 3** represent the initial state where a traditional single proxy configuration on the local network was used, prior to the introduction of the double proxy system.

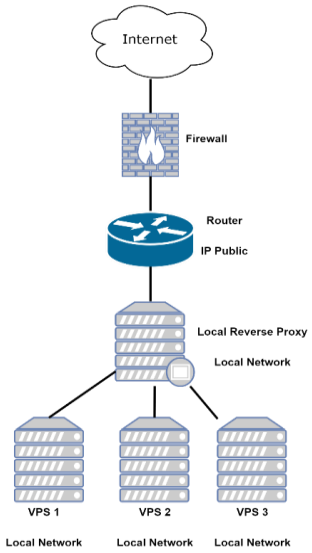


Figure 2. First state design without a public proxy and using the public IP at the router's public interface.

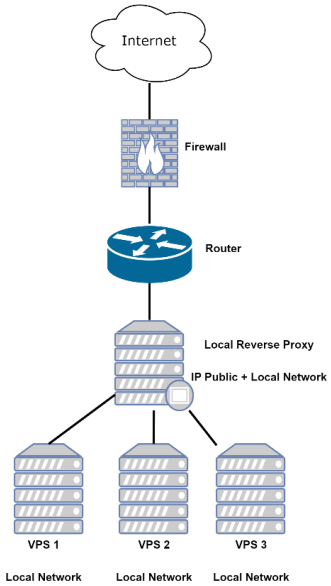


Figure 3. Second state design without public proxy and using IP Public at local proxy.

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

Figure 2 illustrates the initial design without a public proxy. The public IP was assigned to the router's boundary interface, and a firewall was deployed on the router to provide an additional layer of security against DDoS attacks, in combination with the local proxy. **Figure 3** depicts the second design where the public proxy was omitted. The local proxy was assigned a public IP, allowing it to function on both the local and public networks. The local network serves as a reverse proxy to the VPS, while the public network enables global internet access. A firewall on the router provided an additional layer of security against DDoS attacks, working in tandem with the local proxy. The key difference between the two figures is the location of the public IP, which determines how incoming traffic is initially handled. In **Figure 2**, the router acts as the initial point of contact for public traffic, while in **Figure 3**, the local proxy assumes this role.

Figures 4 and 5 below illustrate the innovation proposed in this research, with a primary focus on the double proxy approach.

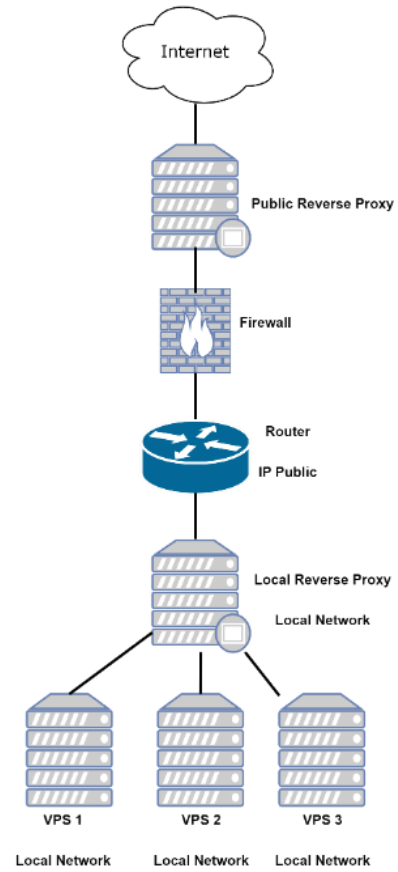


Figure 4. First proposed design with public proxy and using IP Public at router public interface.

Figure 4 presents the first network security architecture design to secure the data center using a double proxy. Both the public and local proxies are reverse proxies. The public proxy is placed above the firewall in the network architecture, aiming to protect against DDoS attacks originating from the internet towards applications on the local VPS. The local proxy is used for load balancing, hardening against local network DDoS attacks, and as a reverse proxy for requests to the local VPS. The router is configured with a public IP on the interface facing the internet.

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

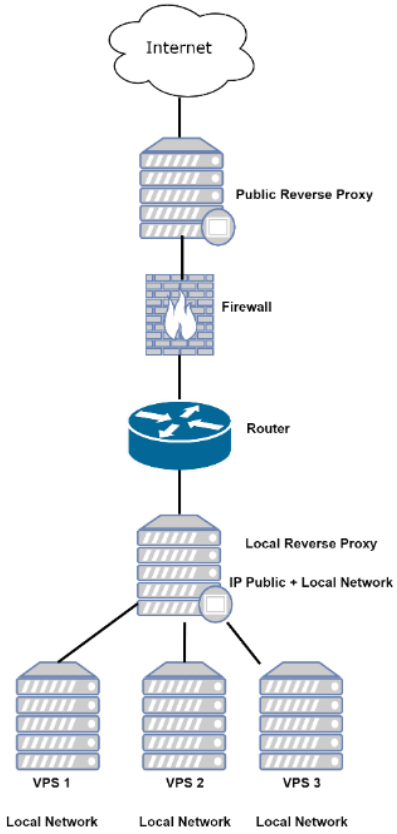


Figure 5. Second proposed design with public proxy and using IP Public at local proxy.

Figure 5 presents the second network security architecture design to secure the data center using a double proxy. Both the public and local proxies are reverse proxies. The public proxy is placed above the firewall in the network architecture, aiming to protect against DDoS attacks originating from the internet towards applications on the local VPS. The local proxy is used for load balancing, hardening against local network DDoS attacks, and as a reverse proxy for requests to the local VPS. The local proxy is configured with a public IP directly connected to the internet. The difference between the two figures lies in the placement of the public IP address. In **Figure 4**, the public IP is located on the router, while in **Figure 4**, it is placed on the local proxy.

2.3. Design Implementation

The previously designed data center and network were implemented on physical servers and virtualized environments. By using server virtualization techniques, it can improve security, reliability, flexible resource usage and better scalability [21][22]. The data center server had the following specifications: 16 x Intel(R) Xeon(R) Silver 4110 CPU @ 2.10GHz, 32 GB Memory, and 2TB Hard disk. To test the design model, a local reverse proxy with 2 Cores, 1 GB of memory, and 3 GB of hard disk was used. The local reverse proxy utilizes the Nginx web server, which offers better performance compared to Apache [5][23][24]. The local reverse proxy was configured for security by applying hardening techniques. Additionally, a virtual server was used as a local server with 1 Core, 512 MB Memory, and 3 GB Hard disk. All servers, including the data center, local reverse proxy, and virtual server, were connected to the local network. This was also a security measure for the data center, as access to the virtual server from the public network or the internet required going through the local reverse proxy. For the public reverse proxy, the highly reliable third-party platform Cloudflare was employed

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

[2]. For the network model used in this study, the design of the network topology uses a tree topology, which is widely used and has many advantages, is more efficient, and easier to maintain [25].

2.4. Testing the Design

The fourth phase of this research consists of testing the designed and deployed system. Building upon the second step, four data center and network security architectures will undergo comparative analysis under two attack scenarios, employing both TCP and UDP protocols. The testing will involve subjecting the system to DDoS attacks utilizing both TCP and UDP protocols [4][7][18][17], with each attack lasting for 30 seconds. The study will evaluate the impact of these attacks on the data center, local reverse proxy, local virtual server, and network bandwidth. The performance of the third-party public reverse proxy will not be within the scope of this research. The primary focus will be on assessing the impact on the internal data center. The DDoS attacks will be conducted using the MHDDoS tool, a tool that provides 56 attack methods, including TCP and UDP-based attacks, which will be evaluated in this study. **Figure 6** provides a visual representation of the MHDDoS tool interface and its configuration for the attack testing phase.



Figure 6. MHDDoS tool for attack testing.

MHDDoS is a command-line-driven tool that is developed using the Python programming language. It can attack at layer 4 and layer 7 using various methods. The tool's popularity is evident from its over 13,000 stars on GitHub, attesting to its robust capabilities. Attack scenarios were designed utilizing MHDDoS across two primary protocols (TCP and UDP) with a constant attack intensity (30 seconds, 20 threads). The effectiveness of traffic filtering was evaluated not only based on CPU load and bandwidth utilization but also on the system's success in blocking packets before reaching the VPS. Evaluation criteria included: CPU utilization on the router and proxies, bandwidth usage, latency (not directly tested in this research but proposed as a future metric), and service request success rate.

2.5. Evaluation

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

The research concludes with an evaluation phase. This phase involves a comprehensive assessment of the experimental results to compare the performance of each designed data center security architecture. The objective is to identify the most suitable design for institutional implementation. However, it is essential to note that ongoing security enhancements will be required to ensure sustained user satisfaction, data integrity, and overall system security.

3. RESULTS AND DISCUSSION

DDoS attack tests were conducted using both TCP and UDP protocols for each of the previously mentioned architecture designs. Each attack was carried out for 30 seconds on each architectural design. The attacks utilized 20 threads, enabling the transmission of numerous large data packets. The DDoS attacks consumed approximately 300-400 Mbps of bandwidth. In the DDoS attack trials using the TCP protocol, packet sizes varied between 11,000 and 44,000 packets per second. From this packet count, the data size sent was equivalent to the attacker's server bandwidth, ranging between 350 Mbps and 400 Mbps, which is equivalent to approximately 45 Mbps. The DDoS attack trials were conducted sequentially under identical environmental conditions on all four architectural designs depicted in **Figures 2, 3, 4, and 5** above using the MHDDoS tool. The research results (the average CPU load was taken, and if it was below 2%, it was written as NaN, as it was not considered a DDoS attack) are detailed in **Table 1** below.

Table 1. Result of DDoS attack testing.

No	Attack	Design	Router			Data Center		Local Proxy		Local VPS	
			CPU	BW-1	BW-2	CPU	BW	CPU	BW	CPU	BW
1	TCP	1	67%	120M	120M	5.7%	42M	5.2%	42M	NaN	< 1K
2	UDP	1	80%	352M	< 1K	NaN	< 1K	NaN	< 1K	NaN	< 1K
3	TCP	2	67%	120M	120M	3.2%	36M	5.4%	120M	NaN	< 1K
4	UDP	2	80%	352M	351M	NaN	346M	4.6%	< 1K	NaN	< 1K
5	TCP	3	4%	< 1K	< 1K	NaN	< 1K	NaN	< 1K	NaN	< 1K
6	UDP	3	4%	< 1K	< 1K	NaN	< 1K	NaN	< 1K	NaN	< 1K
7	TCP	4	4%	< 1K	< 1K	NaN	< 1K	NaN	< 1K	NaN	< 1K
8	UDP	4	4%	< 1K	< 1K	NaN	< 1K	NaN	< 1K	NaN	< 1K

The TCP-based DDoS attack on both architecture designs showed that multi-threaded attacks significantly increased CPU usage on the local proxy while having no impact on the target VPS, as the attacks were successfully terminated at the preconfigured proxy layer. In architecture design 1, the attack generated approximately 120 Mbps of bandwidth on the internet interface, with 42 Mbps reaching the local proxy network and increasing the proxy CPU load to 5.2% from a pre-attack level of around 0.03%. In architecture design 2, similar behavior was observed, with 120 Mbps traffic forwarded to the local proxy network and the proxy CPU load reaching 5.4%. In both cases, the router experienced an increase in CPU utilization to an average of 67%, depending on the number of attack threads and the attacker's bandwidth capacity.

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

The UDP-based DDoS attack also increased system resource usage but remained contained at the local proxy layer, preventing any impact on the target VPS. In architecture design 1, the attack caused an average bandwidth increase of approximately 352 Mbps on the internet interface, although the traffic was not forwarded to the local proxy network. The maximum CPU load on the local proxy reached 3.7%, slightly higher than the pre-attack level. In architecture design 2, the UDP attack produced similar bandwidth spikes, with traffic forwarded to the local proxy network and proxy CPU usage reaching 4.6%. For both architectures, router CPU utilization increased significantly, averaging around 80%, indicating that the router handled most of the attack traffic load.

DDoS attacks utilizing both TCP and UDP protocols on the proposed data center security architectures (**Figures 4 and 5**) demonstrated significant improvements. The public proxy placed above the firewall exhibited exceptional performance, effectively preventing DDoS attacks from reaching the firewall layer and below. This is evident from the absence of any spikes in CPU load and bandwidth on the router, data center, local proxy, or local VPS. Both proposed data center security architectures demonstrated equally impressive performance. The only difference lies in the placement of the public IP, which is located on the router in the design depicted in **Figure 4** and on the local proxy in the design depicted in **Figure 5**. This provides stakeholders with options that can be implemented in their network security systems.

To further illustrate these improvements, **Figure 7** presents a comparative analysis between design 1 and design 3, highlighting the performance gap under identical attack conditions.

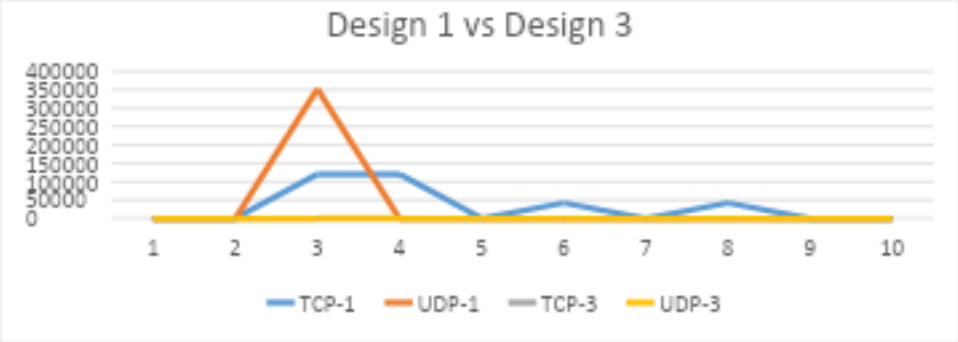


Figure 7. Comparison between design 1 and design 3.

A comparison between design 3 and design 1 (the initial design without a double proxy) in mitigating DDoS attacks using the TCP protocol reveals a significant performance improvement. Specifically, there was a 63% reduction in the router's CPU load, a 120 Mbps bandwidth saving on the internet-connected interface, a 120 Mbps bandwidth saving on the local network, a 5.7% reduction in overall data center CPU load, a 42 Mbps bandwidth saving on the data center, and a 5.2% reduction in local proxy CPU load with a 42 Mbps bandwidth saving. When comparing design 4 and design 2 in mitigating DDoS attacks using the UDP protocol, there was a 76% improvement in the router's CPU load and a 352 Mbps bandwidth saving on the internet-connected interface.

A detailed side-by-side comparison of these performance metrics between design 4 and design 2 is visualized in **Figure 8**.

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

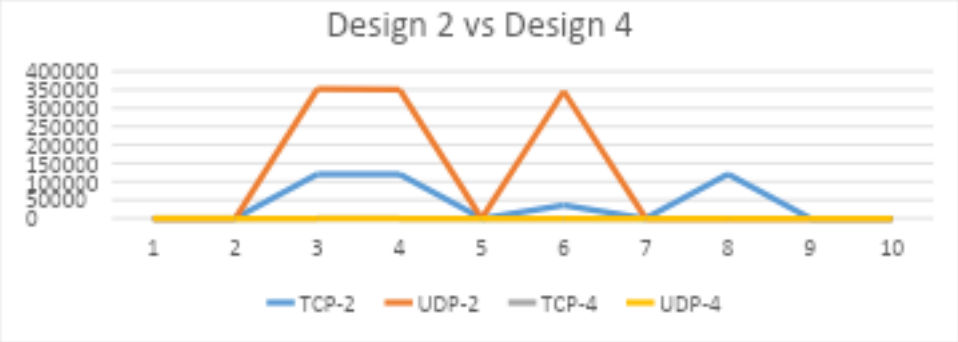


Figure 8. Comparison between design 2 and design 4.

Comparing designs 4 and 2, the new design (5) demonstrated a significant improvement in mitigating DDoS attacks. Using TCP, design 5 reduced router CPU load by 63%, saved 120 Mbps of bandwidth on both the internet and local networks, and decreased data center and local proxy CPU load and bandwidth by 3.2% and 5.4%, respectively, with corresponding bandwidth savings. For UDP attacks, design 5 reduced router CPU load by 76%, saved 352 Mbps of bandwidth on the internet, and significantly reduced bandwidth consumption on the local network and data center. To enhance the clarity and impact of the findings, four bar charts have been added to illustrate the resource usage across the four architectural designs. These graphs allow for direct visual comparison between traditional single-proxy designs and the proposed double-proxy implementations. The comprehensive data for these metrics are visualized in **Figures 9 and 10**, which display the router and data center resource usage, respectively.

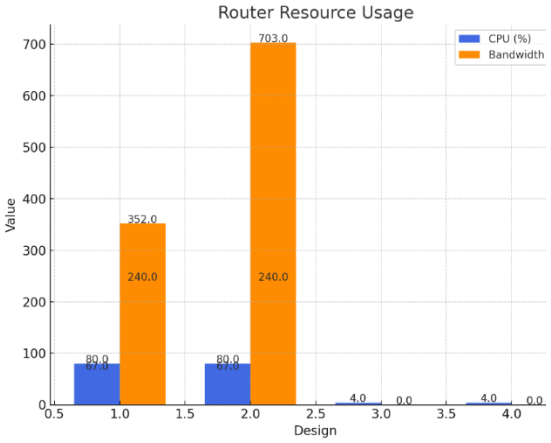


Figure 9. Router Resources Usage.

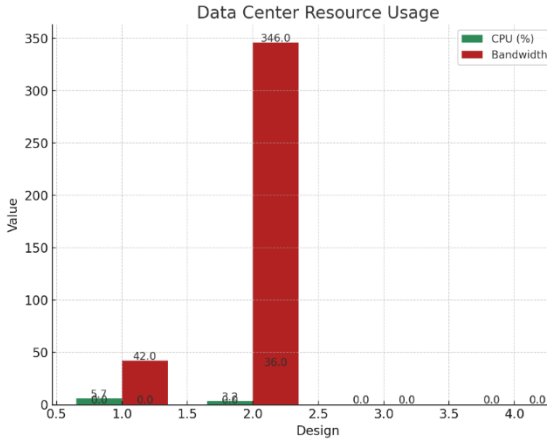


Figure 10. Data Center Resource Usage.

Figure 9 highlights the CPU and bandwidth utilization of the router, designs 1 and 2, CPU usage reaches 67%–80%, and bandwidth spikes to 703 MB. Designs 3 and 4, CPU usage drops to 4%, and bandwidth consumption is 0 MB. In single proxy configurations, the router is the first component to face the full intensity of a DDoS attack. The double proxy setup, especially with a cloud-based public proxy, effectively shields the router, drastically reducing its load.

Figure 10 displays CPU and bandwidth usage at the Data Center server. Designs 1 and 2 show CPU usage of up to 5.7% and bandwidth consumption as high as 346 MB. Designs 3 and 4 show no measurable resource usage, with values at or near zero. Without the double proxy system, malicious traffic is able to reach the data center server, causing performance degradation. The double proxy structure effectively intercepts and blocks harmful traffic before it impacts core infrastructure. Furthermore, the efficiency of this mitigation strategy

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

across other critical components is detailed in **Figures 11 and 12**, which illustrate the local proxy resource usage and local VPS bandwidth usage, respectively.

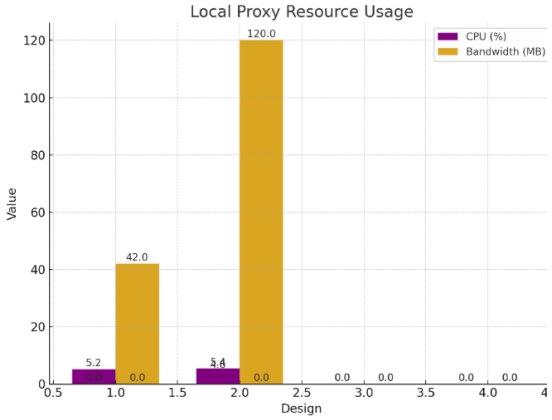


Figure 11. Local Proxy Resource Usage.

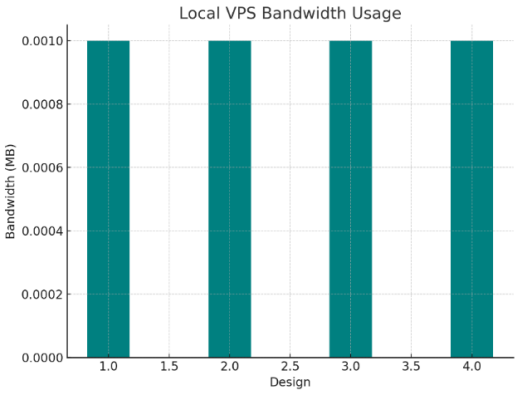


Figure 12. Local VPS Bandwidth Usage.

Figure 11 compares the CPU usage and bandwidth consumption of the Local Proxy. Designs 1 and 2 (without a public proxy) show CPU loads of up to 5.2%–5.4% and bandwidth usage reaching 120 MB. Designs 3 and 4 (with a double proxy setup) show near-zero CPU and bandwidth usage. In single proxy designs, the local proxy directly absorbs the DDoS attack load. The addition of a public reverse proxy in Designs 3 and 4 shifts that burden to the outer layer, allowing the local proxy to focus on legitimate traffic only.

Figure 12 illustrates the bandwidth usage of the Local VPS across the four architectural designs (Designs 1 to 4). The bandwidth recorded for all designs is nearly 0.001 MB, indicating minimal data transmission to the local VPS. All designs successfully prevent DDoS traffic from reaching the Local VPS. However, the double proxy architectures (Designs 3 and 4) enhance traffic isolation and filtering, ensuring that the VPS remains fully protected without experiencing any significant load.

The proposed double proxy design outperforms the single-proxy models due to its layered defense mechanism. The public reverse proxy, hosted via Cloudflare, acts as a buffer zone that absorbs volumetric traffic and blocks known threats even before reaching the institutional network. The local reverse proxy further inspects, filters, and balances the incoming requests within a secure internal boundary. This two-tier separation significantly reduces load on downstream components (router, local proxy, data center), as malicious traffic is halted in earlier stages. In contrast, single-proxy systems expose internal resources more directly to attack loads. Implementing the double proxy architecture provides substantial long-term benefits. It improves operational stability by reducing hardware stress, which helps extend the lifespan of network equipment. In addition, the architecture enables resource savings by lowering CPU and bandwidth demands, thereby reducing energy consumption and infrastructure costs. The modular design also enhances scalability, allowing the system to be adapted to various network sizes and environments. Furthermore, the implementation strengthens the overall security posture by establishing a multi-layer protection mechanism that aligns with zero-trust principles and supports compliance with cybersecurity standards.

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

4. CONCLUSION

This study proposed and evaluated a double proxy architecture to enhance data center security against DDoS attacks. The results show that the proposed approach improves system performance compared to traditional single-proxy designs, achieving CPU usage reductions of up to 76% and bandwidth savings of up to 352 MB. By combining a public reverse proxy and a local reverse proxy, the architecture effectively filters malicious traffic and maintains system stability during volumetric attacks.

However, this study has several limitations. The evaluation focused only on TCP and UDP-based DDoS attacks (Layer 3 and 4), while the effectiveness against more complex Layer 7 attacks was not examined. In addition, scalability in large-scale enterprise or multi-site data center environments has not yet been evaluated. Future research should explore the integration of machine learning-based traffic classification and anomaly detection, as well as real-world deployment to further assess system performance, reliability, and operational feasibility.

5. ACKNOWLEDGMENT

This study is supported by the Department of Computer and Business, Politeknik Negeri Cilacap, in providing the necessary network equipment.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

7. REFERENCES

[1] Wikrama, K. S. M., Firdaus, R., Mendrofa, L. Z. M., Saskara, G. A. J., & Listartha, I. M. E. (2023). DDoS Attack Using GoldenEye, DAVOSET, and PyLoris Tools. Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi, 9(2), 43–52. https://doi.org/10.24014/coreit.v9i2.20020.

[2] Imthiyas, M., Wani, S., Abdulkhaleq, R., Abdulghafor, A., Ibrahim, A. A., & Hafeez, A. (2020). DDoS Mitigation: A review of Content Delivery Network and its DDoS Defence techniques. International Journal of Perceptive and Cognitive Computing, 6(2), 67–76. https://journals.iium.edu.my/kict/index.php/IJPCC/article/view/160.

[3] Wani, S., Imthiyas, M., Almohamedh, H., Alhamed, K. M., Almotairi, S., & Gulzar, Y. (2021). Distributed Denial of Service (DDoS) Mitigation Using Blockchain—A Comprehensive Insight. Symmetry, 13(2), 1–21. https://doi.org/10.3390/sym13020227.

[4] Zebari, R. R., Zeebaree, S. R. M., Sallow, A. B., Shukur, H. M., Ahmad, O. M., & Jacksi, K. (2020). Distributed Denial of Service Attack Mitigation using High Availability Proxy and Network Load Balancing. 3rd International Conference on Advanced Science and Engineering (ICOASE), 174–179. https://doi.org/10.1109/ICOASE51841.2020.9436545.

[5] Saputra, P. S., Pratama, P. A., & Tjahyanti, L. P. A. S. (2023). Perancangan Dan Komparasi Web Server Nginx Dengan Web Server Apache Serta Pemanfaatan Reverse Proxy Server Pada Nginx. Jurnal Komputer dan Teknologi Sains, 2(1), 16–21.

[6] Singh, A., & Gupta, B. B. (2022). Distributed Denial-of-Service (DDoS) Attacks and Defense Mechanisms in Various Web-Enabled Computing Platforms. International Journal on Semantic Web and Information Systems, 18(1), 1–43. https://doi.org/10.4018/IJSWIS.297143.

[7] Bamane, A. (2023). Detecting and Mitigating Network Attacks Using Network Simulator 3. International Journal of Innovative Research in Technology, 9(12), 1285–1292.

[8] Dangi, R., Jha, V. K., Choudhary, M., & Bhavsar, P. D. (2019). An Enhanced Proxy Server for Better Performance & Security of the Network for Academics. International Journal for Research in Applied Science and Engineering Technology, 7(6), 278–281. https://doi.org/10.32622/ijrat.76S201955.

[9] Shah, Z., Ullah, I., Li, H., Levula, A., & Khurshid, K. (2022). Blockchain-Based Solutions to Mitigate Distributed Denial of Service (DDoS) Attacks in the Internet of Things (IoT): A Survey. Sensors, 22(3), 1–26. https://doi.org/10.3390/s22031094.

[10] Wright, M., Venkatesan, S., Albanese, M., & Wellman, M. P. (2016). Moving Target Defense against DDoS Attacks. Proceedings of the 2016 ACM Workshop on Moving Target Defense, (October), 93–104. https://doi.org/10.1145/2995272.2995279.

[11] Karnani, S., & Shakya, H. K. (2023). Mitigation strategies for distributed denial of service (DDoS) in SDN: A survey and taxonomy. Information Security Journal: A Global Perspective, 32(6), 444–468. https://doi.org/10.1080/19393555.2022.2111004.

[12] Wang, J., Wang, L., & Wang, R. (2023). A Method of DDoS Attack Detection and Mitigation for the Comprehensive Coordinated Protection of SDN Controllers. Entropy, 25(8), 1–26. https://doi.org/10.3390/e25081210.

[13] Rifai, M. F., Hendra, S., Ngemba, H. R., Azhar, R., & Laila, R. (2024). Enhancing Web Server Security against Layered Cyber Threats in Healthcare. Advances in Sustainable Science and Engineering Technology, 6(2), 1–9. https://doi.org/10.26877/asset.v6i2.18307 .

[14] Orosz, P., Nagy, B., & Varga, P. (2025). Real-Time Detection and Mitigation Strategies Newly Appearing for DDoS Profiles. Future Internet, 17(9), 1–32. https://doi.org/10.3390/fi17090400.

[15] Khan, F. F., Hossain, N. M., Shanto, M. N. H., Anwar, S. B., & Noor, J. (2022). Mitigating DDoS Attacks Using a Resource Sharing Network. Proceedings of the 9th International Conference on Networking, Systems and Security, (December), 1–11. https://doi.org/10.1145/3569551.3569560.

[16] Ouhssini, M., Afdel, K., Akouhar, M., Agherrabi, E., & Abarda, A. (2024). Advancements in detecting, preventing, and mitigating DDoS attacks in cloud environments: A comprehensive systematic review of state-of-the-art approaches. Egyptian Informatics Journal, 27(100517), 1–37. https://doi.org/10.1016/j.eij.2024.100517.

[17] Dharam, P., & Musarrat, J. (2020). An Authentication Technique to Handle DDoS Attacks in Proxy-Based Architecture. ICNS 2020: The Sixteenth International Conference on Networking and Services, (16), 49–54.

[18] Zain, A. R., Muhamad, I., Matin, M., & Kautsar, D. K. (2023). Analisis Implementasi Modsecurity dan Reverse Proxy Untuk Pencegahan Serangan Keamanan DDoS pada Web Server. SNIV Seminar Nasional Inovasi Vokasi, 2(1), 118–127.

[19] Mahjabin, T., Xiao, Y., Sun, G., & Jiang, W. (2017). A survey of distributed denial-of-service attack, prevention, and mitigation techniques. International Journal of Distributed Sensor Networks, 13(12), 1–33. https://doi.org/10.1177/1550147717741463.

[20] Amalia, E. R., Nurheki, Saputra, R., Ramadhana, C., & Yossy, E. H. (2022). Computer network design and implementation using a load balancing technique with per connection classifier (PCC) method based on a MikroTik router. Procedia Computer Science, 216, 103–111. https://doi.org/10.1016/j.procs.2022.12.116.

[21] Satwika, I. K. S., & Andika, I. G. (2024). Performance Analysis of an E-commerce Website

Using Distributed Servers (Case Study: E-commerce Bumdes Sarining Kukuh Winangun). Journal of Computer Networks, Architecture and High Performance Computing, 6(3), 1390–1398. https://doi.org/10.47709/cnahpc.v6i3.4383.

[22] Manalu, A. S., & Sitanggang, S. S. (2019). Perancangan Dan Implementasi Private Cloud Storage Dengan Owncloud Pada Jaringan Lokal Menggunakan Virtualbox. Journal of Computer Networks, Architecture and High Performance Computing, 1(2), 60–71. https://doi.org/10.47709/cnahpc.v1i2.244.

[23] Apriliansyah, F., Fitri, I., & Iskandar, A. (2020). Implementasi Load Balancing Pada Web Server Menggunakan Nginx. Jurnal Teknologi dan Manajemen Informatika, 6(1), 18–26. https://doi.org/10.3997/2214-4609.201801770.

[24] Wijaya, G., & Franklyn, F. V. (2021). Perancangan Dan Implementasi Desain Arsitektur Cloud Yang Aman Untuk Sistem Stok Pada Halutime Thrift Shop. Conference on Business, Social Sciences and Technology (CoSST), 1(1), 39–46. https://journal.uib.ac.id/index.php/conescintech/article/view/5828.

[25] Fa'iz, N. F., Irfawan, D. R., Putri, A. W., & Hidayatullah, S. (2024). Arsitektur Jaringan Menggunakan Topologi Tree. Jurnal Riset Sistem Informasi dan Teknik Informatika (Merkurius), 2(3), 98–104. https://doi.org/10.61132/merkurius.v2i3.112.

DOI: https://doi.org/10.17509/coelite.v5i1.98157
p- ISSN 2829-4157 e- ISSN 2829-4149

Journal of Computer Engineering, Electronics and Information Technology (COELITE)

Journal homepage: https://ejournal.upi.edu/index.php/COELITE

A Comparative Evaluation of State-of-the-Art Deep Face Recognition Models on the Indonesian Muslim Student Face Dataset

Mokhamad Arfan Wicaksono[1], Ibrahim Faizal Burhan[2], Hadi Permana[3], Muhung Anggarawan[4], Bakhriyah Firdausi[5], Muhammad Rafid Amrullah[6], Ratnasari Nur Rohmah[7]*

[1,2,3,4,5] Department of Software Engineering Technology, Akademi Digital Bandung, Indonesia
[6] Software Engineer, Avey, Qatar
[7] Faculty of Engineering, Universitas Muhammadiyah Surakarta, Indonesia
Correspondence E-mail: marfanw@gmail.com

ABSTRACT

Face recognition systems perform well on Western-centric benchmarks, but their reliability on diverse populations is less studied. This paper evaluates six state-of-the-art models on the Indonesian Muslim Student Face Dataset (IMSFD), containing 26,760 images of 68 individuals under varied conditions, including hijab use. Models tested include FaceNet, AdaFace, MagFace, and ElasticFace-Arc using identification and verification protocols, with metrics such as CMC, ROC/DET, EER, and decidability index, along with computational profiling. Results show AdaFace-R50 achieves the best rank-1 accuracy at 91.70% (95% CI: [90.98%, 92.38%]), significantly outperforming others. Performance varies across subsets, with A2 reaching ≥99.40% accuracy, while B1 is most challenging (≤86.05%). These findings stress the importance of diverse datasets and highlight the effectiveness of modern margin-based loss functions.

ARTICLE INFO

Article History:
Submitted/Received 16 Jan 2026
First Revised 23 Feb 2026
Accepted 09 Mar 2026
First Available online 01 Apr 2026
Publication Date 01 Apr 2026

Keyword:
face recognition, deep learning, comparative evaluation, Indonesian faces, angular margin loss, biometric performance.

© 2026 Universitas Pendidikan Indonesia

1. INTRODUCTION

Face recognition has become one of the most widely deployed biometric technologies, with applications spanning security, surveillance, mobile device authentication, and identity verification [1]. Recent advances in deep learning, particularly the development of discriminative loss functions such as ArcFace [2], CosFace [3], and their variants, have pushed face recognition accuracy to near-perfect levels on established benchmarks like Labeled Faces in the Wild [4] and MegaFace [5].

However, multiple studies have demonstrated that face recognition performance varies significantly across demographic groups defined by race, ethnicity, and gender [6]. Models trained predominantly on Western faces often exhibit degraded performance on underrepresented populations. This disparity is particularly acute for individuals who wear religious head coverings, such as the hijab, which partially occludes the face boundary and alters the overall facial appearance context [7].

Indonesia, the world's most populous Muslim-majority country, presents a unique evaluation context. The Indonesian Muslim Student Face Dataset [8] was specifically curated to address the scarcity of face datasets representing Indonesian individuals, many of whom wear hijab. Despite the availability of IMSFD, no comprehensive comparative evaluation of modern face recognition models has been conducted on this dataset.

This paper addresses this gap by presenting a rigorous, reproducible evaluation of six state-of-the-art face recognition models on IMSFD. Our contributions are as follows:

1. We conduct the first comprehensive comparative evaluation of six modern deep face recognition models—FaceNet [9] (two variants), AdaFace [10] (two architectures), MagFace [11], and ElasticFace [12] —on the IMSFD dataset.
2. We provide a thorough evaluation protocol encompassing both closed-set identification (CMC rank-1 through rank-20 with bootstrap 95% confidence intervals) and face verification (AUC, EER, d', TAR@FAR at multiple operating points, ROC and DET curves).
3. We perform per-subset and per-identity analyses that reveal significant performance heterogeneity across capture conditions, with statistical significance established through McNemar's test.
4. We benchmark inference speed and computational cost—including parameter counts, FLOPs, and GPU memory consumption—on modern GPU hardware (NVIDIA RTX 5060 Ti), providing practical deployment guidance.
5. We release our evaluation code and all experimental configurations for full reproducibility.

Related work

Deep face recognition

The evolution of deep face recognition can be characterized by advances in both network architectures and training loss functions. Early deep learning approaches used contrastive [13] and triplet losses [14] to learn discriminative embeddings. FaceNet [15] demonstrated that a 128-dimensional embedding trained with triplet loss on a large-scale dataset could achieve state-of-the-art verification accuracy. However, triplet mining is computationally expensive, and convergence can be slow.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

The introduction of angular margin losses marked a paradigm shift. SphereFace [16] proposed a multiplicative angular margin, while CosFace [17] and ArcFace [18] introduced additive cosine and additive angular margins, respectively. ArcFace, in particular, has become a de facto standard due to its clear geometric interpretation and strong performance across benchmarks.

Building on this foundation, several variants have been proposed to address specific limitations. AdaFace [19] introduced a quality-adaptive margin that emphasizes hard samples when image quality is high and de-emphasizes them when quality is low, leading to improved robustness on mixed-quality datasets. MagFace [20] proposed a magnitude-aware angular margin that encourages the model to produce larger embedding norms for higher-quality images, providing a built-in quality indicator. ElasticFace [13] replaced fixed margins with random margins sampled from a distribution during training, promoting more flexible decision boundaries.

Backbone architectures

The dominant backbone architecture in face recognition has shifted from VGG-style networks [17] to ResNet variants [21]. The Improved ResNet (IResNet), introduced by the InsightFace framework [16], uses a modified residual block structure with BatchNorm–Conv–BatchNorm–PReLU–Conv–BatchNorm ordering and produces 512-dimensional L2-normalized embeddings from 112×112 input images. Both IResNet-50 and IResNet-100 are widely used, with the latter generally achieving higher accuracy at the cost of increased computation.

FaceNet [10] uses the InceptionResNet-V1 architecture [18], which processes 160×160 inputs and produces 512-dimensional embeddings. While architecturally older, FaceNet models pretrained on VGGFace2 [19] and CASIA-WebFace [20] remain popular due to their accessibility and reasonable performance.

Face recognition for diverse populations

Several studies have highlighted performance disparities across demographic groups. The NIST Face Recognition Vendor Test [6] found that many commercial algorithms exhibit higher false positive rates for African-American and Asian faces compared to Caucasian faces. Buolamwini and Gebru (2018) [7] demonstrated significant gender and skin-type bias in commercial facial analysis systems.

For Muslim populations, the presence of the hijab introduces additional challenges. As noted in the occlusion survey [8], head coverings reduce the available facial region for feature extraction and can confuse models that rely on hair and ear features. Despite these challenges, few studies have systematically evaluated modern face recognition models on datasets specifically curated for Muslim populations.

The IMSFD dataset [9] was created to address this gap, providing images of Indonesian Muslim students across multiple capture conditions. However, existing evaluations on IMSFD have been limited in scope, typically testing only one or two models without comprehensive statistical analysis. Our work provides the first large-scale, statistically rigorous comparison across six modern architectures.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

2. METHODS

Dataset

The Indonesian Muslim Student Face Dataset [9] consists of pre-cropped face images of Indonesian Muslim students, organized into three subsets—A1, A2, and B1—each captured under different conditions. The dataset includes both male and female subjects, with many female participants wearing hijab.

Table 1 IMSFD Dataset Statistics.

Subset	Identities	Training Images	Testing Images
A1	24	8,236	2,419
A2	23	6,354	1,666
B1	21	6,458	1,627
Total	68	21,048	5,712

Table 1 summarizes the dataset composition. The total dataset contains 26,760 face images spanning 68 unique identities. Each identity is assigned to exactly one subset, with no identity overlap between subsets. The images are pre-cropped face regions with resolutions in the range of 410–615 pixels.

The three subsets differ markedly in recognition difficulty, as evidenced by the per-subset results in Section 4.4:

- Subset A1 (24 identities, 10655 images): Presents moderate recognition difficulty across all evaluated models.
- Subset A2 (23 identities, 8020 images): The most tractable subset, with near-perfect Rank-1 accuracy achieved by most models.
- Subset B1 (21 identities, 8085 images): The most challenging subset, with consistently lower Rank-1 accuracy across all models.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Figure 1 illustrates the distribution of images and identities across the training and testing splits.

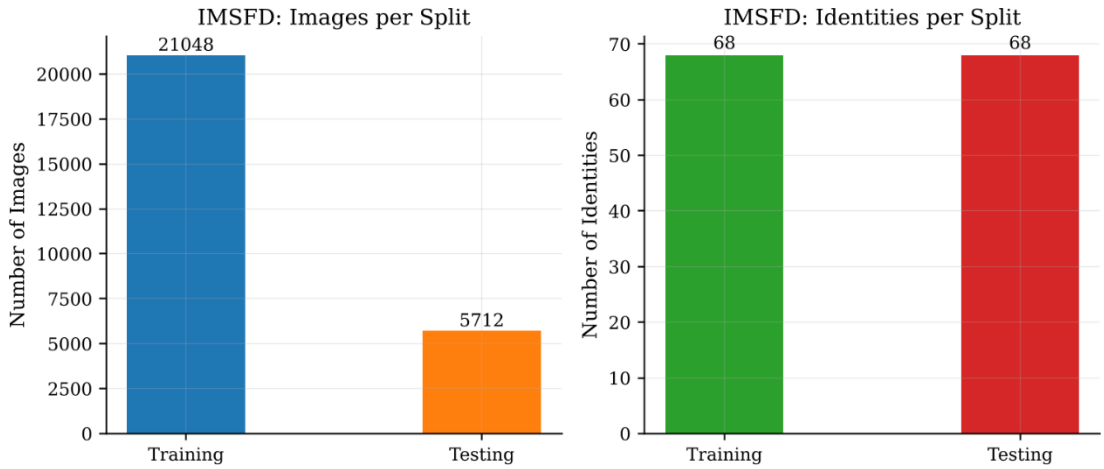


Figure 1 IMSFD dataset statistics showing the distribution of images (left) and identities (right) across training and testing splits.

Evaluated models

We evaluate six pretrained face recognition models spanning four distinct loss function families. All models produce 512-dimensional L2-normalized embedding vectors. **Table 2** summarizes the evaluated configurations.

Table 2. *Summary of evaluated face recognition models.*

Model	Backbone	Inputs	Training Data
FaceNet-VGG2	IncResNet-V1	160^2	VGGFace2
FaceNet-CASIA	IncResNet-V1	160^2	CASIA-WebFace
AdaFace-R100	IResNet-100	112^2	WebFace4M
AdaFace-R50	IResNet-50	112^2	MS1MV2
MagFace-R100	IResNet-100	112^2	MS1MV2
ElasticFace-R100	IResNet-100	112^2	MS1MV2

FaceNet

FaceNet [10] pioneered the use of deep metric learning for face recognition. We evaluate two variants using the InceptionResNet-V1 architecture pretrained on VGGFace2 [19] (3.31M images, 9,131 identities) and CASIA-WebFace [20] (0.49M images, 10,575 identities), respectively. These models accept 160160 inputs normalized to −1,1.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

AdaFace

AdaFace [11] introduces an adaptive margin function that adjusts the emphasis on hard samples based on image quality, estimated via feature norm. We evaluate two configurations: (1) IResNet-100 pretrained on WebFace4M (4.2M images; [22]) and (2) IResNet-50 pretrained on MS1MV2 (5.8M images; [19]). Both accept 112 x 112 inputs.

MagFace

MagFace [12] extends angular margin losses by incorporating magnitude awareness—the magnitude of the feature vector serves as a quality indicator, with the margin adaptively scaled based on the magnitude. We evaluate an IResNet-100 model pretrained on MS1MV2.

ElasticFace

ElasticFace [13] replaces the fixed angular margin in ArcFace with a margin sampled from a Gaussian distribution during training. This elastic margin encourages more flexible and generalizable decision boundaries. We evaluate the ElasticFace-Arc variant using IResNet-100 pretrained on MS1MV2.

Evaluation protocol

We adopt a two-pronged evaluation protocol comprising closed-set identification and face verification.

Closed-set identification

In the closed-set identification protocol, the gallery consists of mean centroid embeddings computed from the training set for each identity. Let $\{x_i\}_{i=1}^{N_k}$ denote the training embeddings for identity k. The gallery centroid is:

$$c_k = \frac{\bar{x}_k}{\| \bar{x}_k \|_2}, \quad \bar{x}_k = \frac{1}{N_k} \sum_{i=1}^{N_k} x_i$$

Each probe image from the testing set is compared against all gallery centroids using cosine similarity. The predicted identity is the one with the highest similarity. Accuracy is reported as the cumulative match characteristic (CMC) at ranks 1 through 20.

Face verification

For face verification, we generate 10000 positive (genuine) pairs and 10000 negative (impostor) pairs from the combined gallery and probe embeddings. For each pair, we compute the cosine similarity score. We report:

- AUC: Area under the ROC curve.
- EER: Equal error rate, the operating point where false positive rate equals false negative rate.
- *d'* (decidability index): A distribution-separation metric defined as:

$$d' = \frac{|\mu_g - \mu_i|}{\sqrt{0.5(\sigma_g^2 + \sigma_i^2)}}$$

  where μ_g, σ_g and μ_i, σ_i are the mean and standard deviation of the genuine and impostor score distributions, respectively.

- TAR@FAR: True accept rate at fixed false accept rate thresholds of 0.1%, 1%, and 10%.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Statistical analysis

Bootstrap confidence intervals

We compute 95% confidence intervals for rank-1 accuracy using the bootstrap method with 2000 resampling iterations. For each iteration, we sample n probe results with replacement (where n is the probe set size) and compute the mean rank-1 accuracy. The 2.5th and 97.5th percentiles of the bootstrap distribution define the 95% CI.

McNemar's test

To assess whether performance differences between model pairs are statistically significant, we apply McNemar's test with continuity correction [23] to the per-probe rank-1 classification outcomes. For each pair of models A and B, we construct the 2×2 contingency table of concordant and discordant probe classifications and compute the McNemar chi-squared statistic. Significance is determined at $\alpha=0.05$.

Per-subset analysis

We repeat the closed-set identification protocol separately for each subset (A1, A2, B1), using only the identities and images belonging to that subset for both gallery construction and probe evaluation.

Speed benchmarking and computational cost

Inference throughput is measured on an NVIDIA GeForce RTX 5060 Ti (16 GB VRAM) using a batch size of 64 with 50 measurement batches after 10 warmup batches, with CUDA synchronization for accurate timing. In addition, we profile the computational cost of each model by reporting: (1) the total number of learnable parameters, (2) floating-point operations (FLOPs) for a single forward pass at the model's native input resolution, and (3) GPU peak memory consumption during inference at batch size 64. These metrics quantify the resource requirements for deployment and enable practitioners to select models that match their hardware constraints.

Implementation details

All experiments are implemented in PyTorch 2.11.0 with CUDA 12.8. Input images are resized to the model-specific input resolution (112×112 for IResNet-based models, 160×160 for FaceNet) and normalized to [-1,1]. Embeddings are L2-normalized before similarity computation. All random seeds are fixed for reproducibility.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

3. RESULTS AND DISCUSSION

Closed-set identification

Table 3 presents the closed-set identification results alongside verification metrics. **Figure 2** shows the CMC curves for all models.

Table 3: closed-set identification and face verification results on IMSFD. Best results in Each column is shown in **bold**.

Model	Rank-1 (%)	95% CI	Rank-5 (%)	Rank-10 (%)	AUC	EER (%)	d'	TAR@FA R=0.1%
FaceNet-VGG2	90.28	[89.51, 91.05]	**94.75**	**96.46**	**0.8861**	**15.81**	1.970	60.41
FaceNet-CASIA	89.16	[88.32, 89.95]	92.70	96.64	0.8465	18.24	1.423	42.05
AdaFace-R100	91.32	[90.58, 92.02]	94.07	95.96	0.8791	16.35	**2.396**	42.76
AdaFace-R50	**91.70**	[**90.98**, **92.38**]	94.17	95.89	0.8710	17.16	2.210	**73.57**
MagFace-R100	51.42	[50.09, 52.70]	69.98	77.19	0.6504	40.66	0.661	19.96
ElasticFace-R100	90.70	[89.93, 91.42]	93.59	95.34	0.8790	16.49	2.354	69.45

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

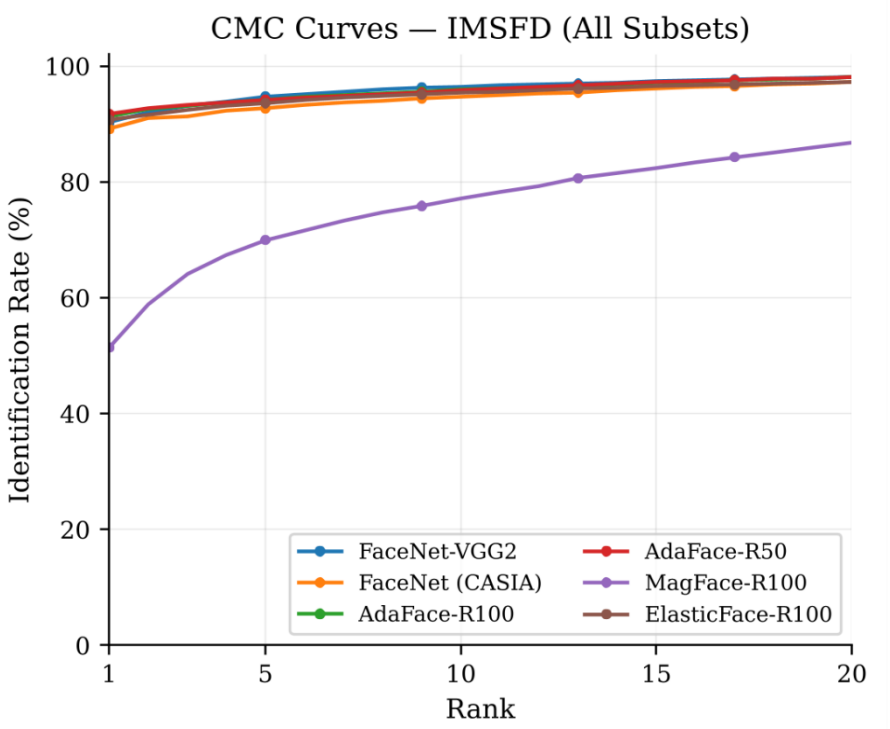


Figure 2. Cumulative Match Characteristic (CMC) curves for all evaluated models on the IMSFD dataset. Higher curves indicate better identification performance across ranks.

AdaFace-R50 achieves the highest rank-1 accuracy at 91.70% (95% CI: [90.98%, 92.38%]), followed closely by AdaFace-R100 at 91.32%. ElasticFace-R100 and FaceNet-VGG2 perform comparably at 90.70% and 90.28%, respectively. FaceNet-CASIA, despite its older architecture, achieves a competitive 89.16%.

The most notable finding is that MagFace-R100 significantly underperforms all other models, achieving only 51.42% rank-1 accuracy. While MagFace has demonstrated strong results on standard benchmarks [12], its magnitude-aware margin mechanism may not generalize well to the specific characteristics of IMSFD, where head coverings and capture conditions may confound the learned quality-magnitude relationship. The poor performance is consistent across all ranks, suggesting a fundamental mismatch rather than a threshold effect.

The non-overlapping 95% confidence intervals between AdaFace-R50 (90.98–92.38%) and FaceNet-CASIA (88.32–89.95%) confirm that the top-performing model significantly outperforms the lower-tier models. The CIs for AdaFace-R50, AdaFace-R100, and ElasticFace-R100 overlap slightly, indicating that the differences among these top performers are small, though McNemar's test (Section 4.5) confirms their statistical significance.

At higher ranks, performance converges: rank-5 accuracies range from 92.70% to 94.75% (excluding MagFace), and rank-10 accuracies range from 94.64% to 96.46%. FaceNet-VGG2 achieves the highest rank-10 accuracy at 96.46%, suggesting that while its top-1 discrimination is slightly weaker than AdaFace, the correct identity is frequently ranked within the top 10.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Face verification

Figure 3 shows the ROC curves in both linear and semi-logarithmic scales. **Figure 4** presents the DET curves.

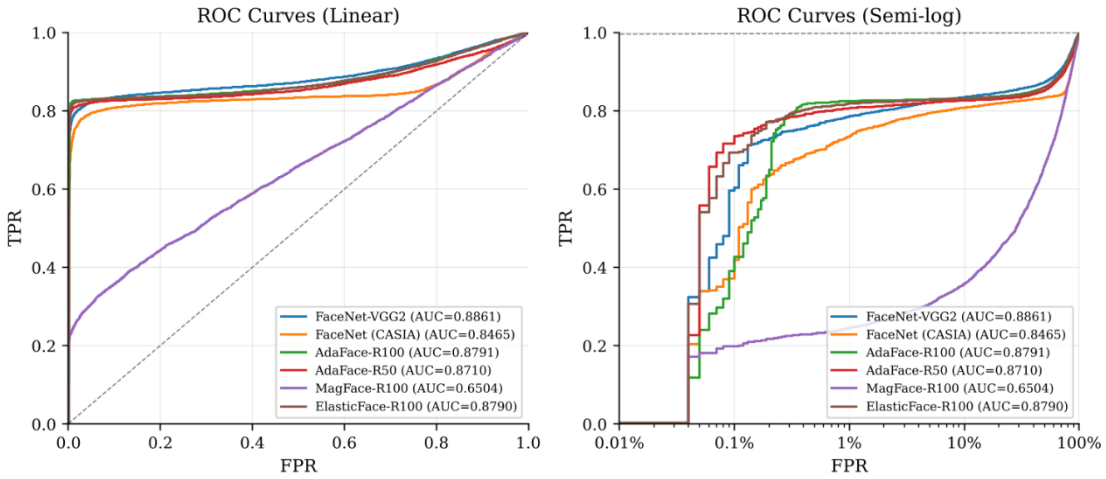


Figure 3. ROC curves in linear (left) and semi-logarithmic (right) scales. The semi-log view reveals differences at low false positive rates.

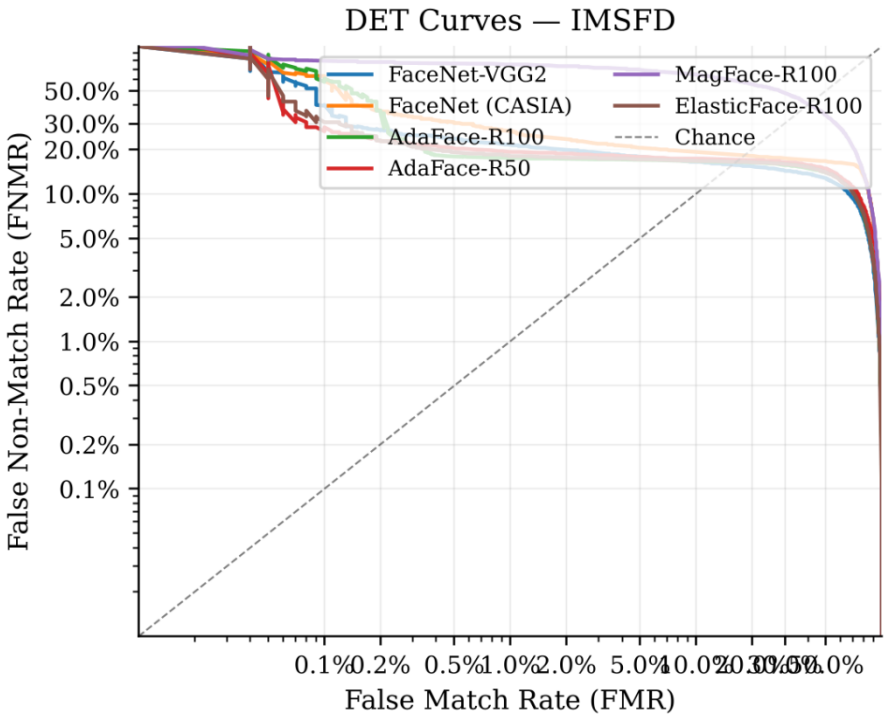


Figure 4. Detection Error Trade-off (DET) curves. Lower and further-left curves indicate better performance.

FaceNet-VGG2 achieves the highest AUC (0.8861) and lowest EER (15.81%), while AdaFace-R100 achieves the highest decidability index (*d'*=2.396), indicating the best separation between genuine and impostor score distributions. The discrepancy between AUC and *d'* rankings arises because AUC captures the entire ROC curve, while *d'* measures distributional separation—AdaFace produces more compact, well-separated clusters despite a slightly worse threshold-independent performance.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

At the critical TAR@FAR=0.1% operating point, AdaFace-R50 achieves the best performance (73.57%), followed by ElasticFace-R100 (69.45%) and FaceNet-VGG2 (60.41%). This metric is particularly important for high-security applications where the false accept rate must be extremely low.

Score distribution analysis

Figure 5 shows the genuine and impostor score distributions for each model.

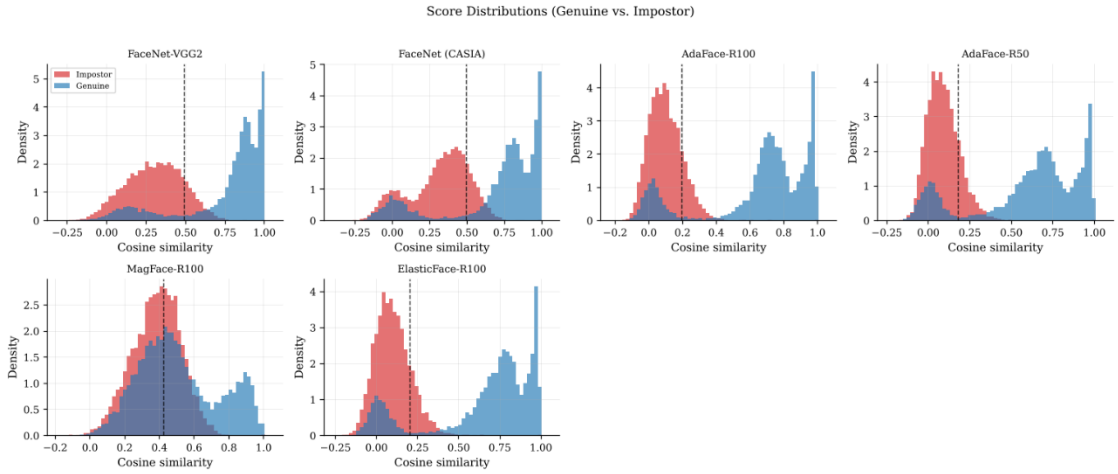


Figure 5. Genuine (blue) and impostor (red) cosine similarity score distributions. The dashed vertical line indicates the EER threshold. Greater separation indicates better discriminability.

The score distributions reveal important characteristics of each model's discriminative behavior. AdaFace-R100 and ElasticFace-R100 show the clearest bimodal separation, consistent with their high *d'* values. FaceNet-VGG2 shows wider genuine distributions, indicating greater intra-class variability but still maintaining good overall separation. MagFace-R100 shows substantial overlap between genuine and impostor distributions, explaining its poor verification performance.

Per-subset analysis

Table 4 and **Figure 6** present the per-subset rank-1 accuracy.

Table 4 *Per-subset rank-1 accuracy (%) on IMSFD.*

Model	A1	A2	B1
FaceNet-VGG2	**91.48**	**99.82**	81.19
FaceNet-CASIA	89.50	99.40	80.02
AdaFace-R100	90.04	99.76	84.94
AdaFace-R50	90.49	**99.82**	**86.05**
MagFace-R100	54.49	62.97	52.80
ElasticFace-R100	90.12	99.76	83.90

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

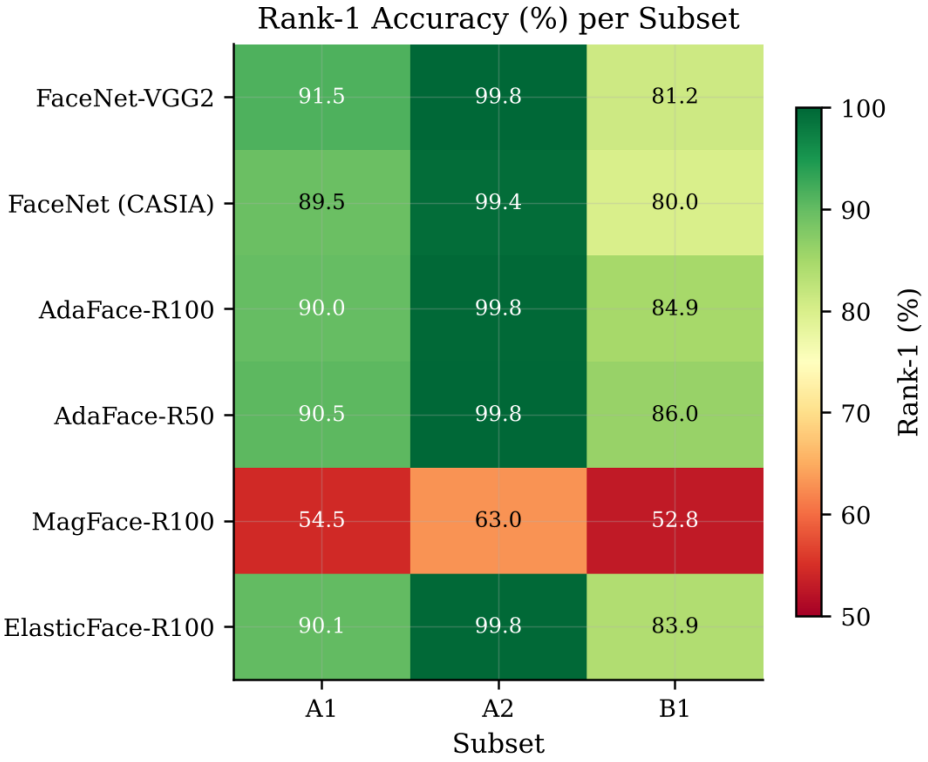


Figure 6. Rank-1 accuracy heatmap across models and subsets. Subset A2 shows near-perfect performance for all competitive models, while B1 is the most challenging.

The per-subset analysis reveals striking performance heterogeneity. Subset A2 is near-trivial for all competitive models, with rank-1 accuracies exceeding 99.40%. This suggests highly controlled capture conditions with minimal intra-class variation. In contrast, subset B1 is consistently the most challenging, with a gap of approximately 10–15 percentage points relative to A2 for the top-performing models.

AdaFace-R50 achieves the best performance on the challenging B1 subset (86.05%), outperforming AdaFace-R100 (84.94%) by over 1 percentage point. This counterintuitive result—where the smaller model outperforms the larger one—may be attributed to the difference in training data: AdaFace-R50 was trained on MS1MV2 (5.8M images), which is larger than the WebFace4M dataset used for AdaFace-R100 (4.2M images), potentially providing better generalization to the challenging conditions in B1.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Statistical significance

Table 5 presents selected pairwise McNemar's test results. The complete set of all 15 pairwise comparisons is available in the accompanying code repository.

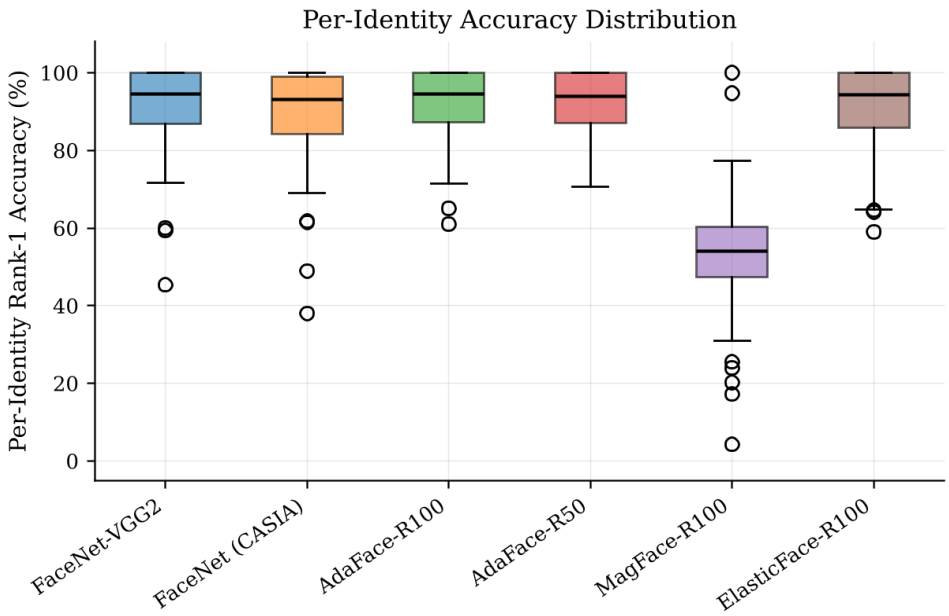
Table 5. *Selected pairwise McNemar's test results (=0.05).*

Model A	Model B	χ^2	p-value
AdaFace-R50	AdaFace-R100	5.96	0.015
AdaFace-R100	ElasticFace-R100	20.28	<0.001
FaceNet-VGG2	ElasticFace-R100	4.07	0.044
FaceNet-VGG2	FaceNet-CASIA	25.12	<0.001
MagFace-R100	ElasticFace-R100	2122.81	<0.001

All 15 pairwise comparisons yield statistically significant differences ($p < 0.05$), confirming that the observed performance differences are not due to chance. The most marginal comparison is FaceNet-VGG2 vs. ElasticFace-R100 ($\chi^2 = 4.07, p = 0.044$), while the largest difference is between MagFace-R100 and any other model ($\chi^2 > 2000, p \ll 0.001$). The comparison between AdaFace-R50 and AdaFace-R100 ($p = 0.015$) confirms that even the modest 0.38 percentage point difference in rank-1 accuracy is statistically significant given the large probe set size.

Per-identity analysis

Figure 7 shows the distribution of per-identity rank-1 accuracy across all 68 identities.



DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Figure 7. Per-identity Rank-1 accuracy distribution. Box plots show the median, interquartile range, and outliers across 68 identities. Models with higher medians and narrower spreads are more consistently accurate.

The per-identity analysis reveals that even the best models exhibit significant variability across identities. The top-performing models (AdaFace-R50, AdaFace-R100, ElasticFace-R100) achieve 100% accuracy on many identities (particularly those in subset A2) but show accuracy as low as 36–72% on the most challenging identities in subset B1. This highlights the importance of reporting distributional statistics rather than aggregate accuracy alone.

Embedding quality

Figure 8 shows the distribution of embedding L2 norms across models for both gallery and probe sets.

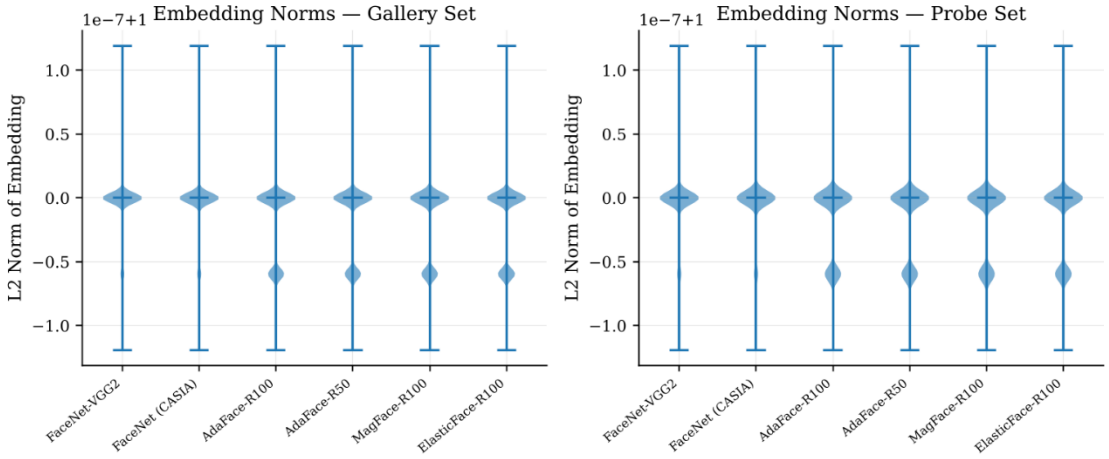


Figure 8. Embedding L2-norm distributions for gallery (left) and probe (right) sets. MagFace's wider norm distribution reflects its magnitude-aware design, while other models produce near-unit-norm embeddings.

Most models produce near-unit-norm embeddings due to L2 normalization in the final layer. MagFace intentionally produces embeddings with variable norms, where the norm encodes an estimate of face quality. However, on IMSFD, this quality estimation does not appear to correlate well with recognition difficulty, contributing to MagFace's poor performance.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Speed–accuracy trade-off

Figure 9 presents the speed–accuracy trade-off measured on the RTX 5060 Ti.

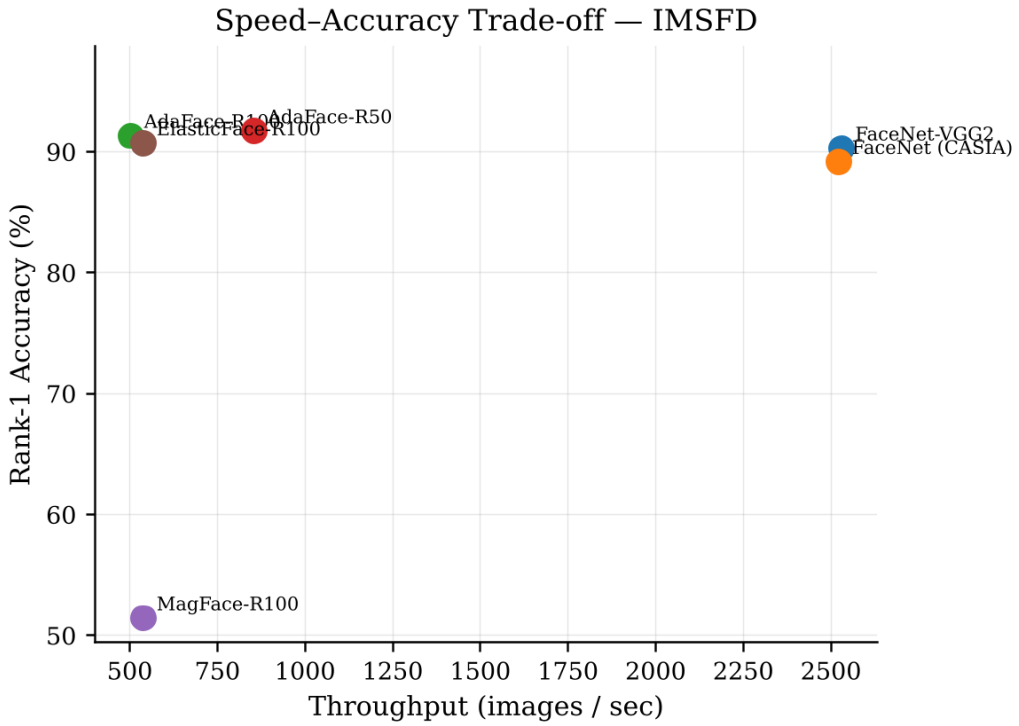


Figure 9. Speed–accuracy trade-off. FaceNet models are approximately 5 times faster than IResNet-100 models, with a modest accuracy penalty.

FaceNet models achieve approximately 2530 images per second, roughly 5 times faster than the IResNet-100-based models (500–540 images/s). AdaFace-R50 offers an attractive middle ground at 854 images/s with the highest rank-1 accuracy, making it the most favorable choice for deployments that require both high accuracy and reasonable throughput.

Computational cost

Table 6 and **Figure 10** present the computational cost of each model in terms of parameter count, FLOPs, and GPU memory consumption.

Table 6. *Computational cost of evaluated models. Params = total learnable parameters; FLOPs = floating-point operations for a single 1122 or 1602 input; Memory = GPU peak memory during inference (batch size 64); Speed = inference throughput.*

Model	Params (M)	FLOPs (G)	Memory (MB)	Speed (img/s)
FaceNet-VGG2	27.91	2.84	414	2531
FaceNet (CASIA)	28.91	2.84	417	2522
AdaFace-R100	65.15	24.15	990	503
AdaFace-R50	43.59	12.59	908	854

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Model	Params (M)	FLOPs (G)	Memory (MB)	Speed (img/s)
MagFace-R100	65.16	24.18	942	538
ElasticFace-R100	65.16	24.18	942	538

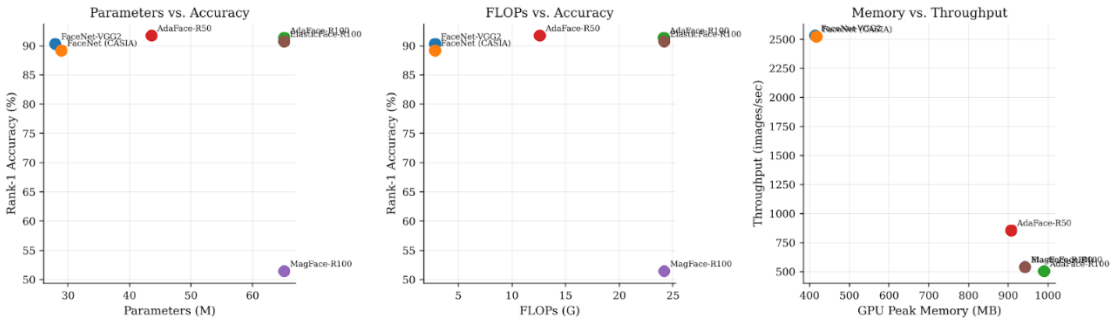


Figure 10. *Computational cost analysis. Left: model size (parameters) vs. rank-1 accuracy. Center: FLOPs vs. accuracy. Right: GPU memory vs. inference throughput. AdaFace-R50 provides the best accuracy-to-cost ratio.*

The IResNet-100-based models (AdaFace-R100, MagFace-R100, ElasticFace-R100) share the same backbone architecture and consequently have similar parameter counts and FLOPs, roughly double those of the IResNet-50-based AdaFace-R50. FaceNet's InceptionResNet-V1 backbone has the fewest parameters among IResNet models but processes larger 160160 inputs, resulting in a different efficiency profile.

Notably, the relationship between model size and accuracy is non-monotonic: AdaFace-R50, with approximately half the parameters of the R100 variants, achieves the highest rank-1 accuracy. This reinforces the finding that training data quality and loss function design can outweigh architectural capacity. From a deployment perspective, AdaFace-R50 offers the most favorable cost–accuracy trade-off, requiring fewer computational resources while delivering superior recognition performance.

GPU memory consumption scales with model size but also depends on input resolution. The peak memory measurements at batch size 64 provide a practical reference for hardware provisioning in deployment scenarios.

Discussion

Key findings

Our evaluation reveals several important findings for deploying face recognition in the context of Indonesian Muslim populations:

Modern margin-based losses outperform triplet loss. AdaFace-R50 and AdaFace-R100 achieve the highest identification rates, confirming that adaptive margin losses provide a meaningful advantage over FaceNet's triplet loss, particularly on challenging subsets.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Training data matters more than architecture depth. AdaFace-R50 (IResNet-50, MS1MV2) outperforms AdaFace-R100 (IResNet-100, WebFace4M), suggesting that the quality and representativeness of the training dataset are more important than backbone capacity for this evaluation context.

Model generalization varies significantly. MagFace's poor performance illustrates that models achieving state-of-the-art results on standard benchmarks may not generalize to datasets with different demographic characteristics and capture conditions. This underscores the importance of dataset-specific evaluation.

Subset difficulty varies dramatically. The 15+ percentage point gap between subsets A2 and B1 demonstrates that recognition performance is highly sensitive to capture conditions. Evaluation on a single condition may give misleading estimates of real-world performance.

Larger models do not guarantee better cost–accuracy trade-offs. Computational cost profiling reveals that AdaFace-R50, with roughly half the parameters and FLOPs of the IResNet-100 variants, achieves the highest accuracy. This demonstrates that practitioners should evaluate cost–accuracy trade-offs rather than defaulting to the largest available model.

Limitations

This study has several limitations. First, we evaluate only pretrained models without fine-tuning on IMSFD, which would likely improve performance for all models. Second, the dataset contains only 68 identities, which limits the generalizability of our findings to larger-scale scenarios. Third, we do not perform face detection or alignment, using only the pre-cropped images provided by IMSFD, which may not reflect realistic deployment conditions. Finally, we do not evaluate the impact of specific demographic attributes (e.g., hijab type, gender) on recognition performance, which would require annotated metadata not available in IMSFD.

Practical recommendations

Based on our findings, we offer the following recommendations:

1. For applications requiring the highest accuracy on Indonesian Muslim populations, **AdaFace-R50** provides the best balance of accuracy and speed.
2. For low-latency applications where speed is critical, **FaceNet-VGG2** offers reasonable accuracy at 5 times the throughput of IResNet-based models.
3. For high-security applications requiring very low false accept rates, **AdaFace-R50** provides the best TAR@FAR=0.1% performance (73.57%).
4. **MagFace should not be deployed** for this population without fine-tuning, as its off-the-shelf performance is inadequate.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

4. CONCLUSION

This paper presented a comprehensive comparative evaluation of six state-of-the-art deep face recognition models on the Indonesian Muslim Student Face Dataset (IMSFD). Our evaluation protocol encompassed closed-set identification with CMC analysis, face verification with multiple metrics, per-subset and per-identity analysis, statistical significance testing, and computational cost profiling.

AdaFace-R50 emerged as the best-performing model with 91.70% rank-1 accuracy (95% CI: [90.98%, 92.38%]), demonstrating the effectiveness of quality-adaptive margin losses for this demographic group. All pairwise model comparisons were statistically significant, and per-subset analysis revealed substantial performance variation across capture conditions.

Our findings highlight the critical importance of evaluating face recognition systems on diverse, non-Western datasets. Models that excel on standard benchmarks—such as MagFace—may perform poorly on underrepresented populations, while less commonly evaluated models may offer unexpected advantages. Future work should explore fine-tuning pretrained models on IMSFD, evaluating the impact of specific demographic factors such as hijab type, and extending the evaluation to open-set identification and cross-dataset generalization scenarios.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The evaluation code and experimental configurations will be made publicly available upon acceptance. The IMSFD dataset is available from the original authors [9].

5. ACKNOWLEDGMENT

This research was supported by Bandung Digital Academy. This work was supported by the Ministry of Education, Culture, Research, and Technology (Kemdiktisaintek) through the *Basis Informasi Penelitian dan Pengabdian kepada Masyarakat* (BIMA) platform.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

7. REFERENCES

[1] Wang, M., Deng, W., 2021. Deep face recognition: A survey. Neurocomputing, 215–244. doi:10.1016/j.neucom.2020.10.081.

[2] Deng, J., Guo, J., Xue, N., Zafeiriou, S., 2019. ArcFace: Additive angular margin loss for deep face recognition, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 4690–4699. doi:10.1109/CVPR.2019.00482.

[3] Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., Liu, W., 2018. CosFace: Large margin cosine loss for deep face recognition, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp.480 5265–5274. doi:10.1109/CVPR.2018.00552.

[4] Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E., 2007. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, 1–11.

[5] Kemelmacher-Shlizerman, I., Seitz, S.M., Miller, D., Brossard, E., 2016. The MegaFace benchmark: 1 million faces for recognition at scale, in: Proc.IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp.4873–4882. doi:10.1109/CVPR.2016.527.

[6] Grother, P., Ngan, M., Hanaoka, K., 2019. Face recognition vendor test 435 (FRVT) part 3: Demographic effects. NIST Interagency Report 8280, 1–82doi:10.6028/NIST.IR.8280.

[7] Buolamwini, J., Gebru, T., 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification, in: Proc. Conf. Fairness, Accountability, and Transparency (FAT*), PMLR. pp. 77–91.

[8] Zeng, D., Veldhuis, R., Spreeuwers, L., 2021. A survey of face recognition techniques under occlusion. IET Biometrics 10, 581–606. doi:10.1049/bme2.12029.

[9] Purnawansyah, P., Wibawa, A.P., Widyaningtyas, T., Haviluddin, H., Darwis, H., Azis, H., Ali, Z., 2023. Indonesian muslim student face dataset (IMSFD). doi:10.17632/f6f3y6ndgw.1.

[10] Schroff, F., Kalenichenko, D., Philbin, J., 2015. FaceNet: A unified embedding for face recognition and clustering, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 815–823. doi:10.1109/CVPR.2015.7298682.

[11] Kim, M., Jain, A.K., Liu, X., 2022. AdaFace: Quality adaptive margin for face recognition, in: Proc. IEEE/CVF Conf. Computer Vision and 450 Pattern Recognition (CVPR), pp. 18750–18759. doi:10.1109/CVPR52688.2022.01819.

[12] Meng, Q., Zhao, S., Huang, Z., Zhou, F., 2021. MagFace: A universal representation for face recognition and quality assessment, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 14225–14234. doi:10.1109/CVPR46437.2021.01400.

[13] Boutros, F., Damer, N., Kirchbuchner, F., Kuijper, A., 2022. Elastic-Face: Elastic margin loss for deep face recognition, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops (CVPRW), pp 1578–1587. doi:10.1109/CVPRW56347.2022.00164.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

[14] Chopra, S., Hadsell, R., LeCun, Y., 2005. Learning a similarity metric discriminatively, with application to face verification, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 539–546. doi:10.1109/CVPR.2005.202.

[15] Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., Song, L., 2017. SphereFace: Deep hypersphere embedding for face recognition, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 6738–6746. doi:10.1109/CVPR.2017.713.

[16] Deng, J., Guo, J., Xue, N., Zafeiriou, S., 2019. ArcFace: Additive angular margin loss for deep face recognition, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 4690–4699. doi:10.1109/CVPR.2019.00482.

[17] Parkhi, O.M., Vedaldi, A., Zisserman, A., 2015. Deep face recognition, in: Proc. British Machine Vision Conference (BMVC), pp. 41.1–41.12. doi:10.5244/C.29.41.

[18] Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2017. Inception-v4, Inception-ResNet, and the impact of residual connections on learning, in: Proc. AAAI Conf. Artificial Intelligence, pp. 4278–4284. doi:10.1609/aaai.v31i1.11231.

[19] Cao, Q., Shen, L., Xie, W., Parkhi, O.M., Zisserman, A., 2018. VGGFace2: A dataset for recognising faces across pose and age, in: Proc. IEEE Int. Conf. Automatic Face & Gesture Recognition (FG), pp. 67–74. doi:10.1109/FG.2018.00020.

[20] Yi, D., Lei, Z., Liao, S., Li, S.Z., 2014. Learning face representation from scratch. arXiv preprint arXiv:1411.7923 doi:10.48550/arXiv.1411.7923.

[21] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 770–778. doi:10.1109/CVPR.2016.90.

[22] Zhu, Z., Huang, G., Deng, J., Ye, Y., Huang, J., Chen, X., Zhu, J., Yang, T., Lu, J., Du, D., Zhou, J., 2021. WebFace260M: A benchmark unveiling the power of million-scale deep face recognition, in: Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 10492–10502. doi:10.1109/CVPR46437.2021.01035.

[23] McNemar, Q., 1947. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12, 153–157. doi:10.1007/BF02295996.

DOI: https://doi.org/10.17509/coelite.v5i1.98428
p- ISSN 2829-4157 e- ISSN 2829-4149

Journal of Computer Engineering, Electronics and Information Technology (COELITE)

Journal homepage: https://ejournal.upi.edu/index.php/COELITE

Graph Theory-Based Simulation for Optimal Access Point Selection in Smart Buildings

Devi Aprianti Rimadhani Agustini[1], Deva Naufal[2], Muhammad Zirly Atthaulah[3], Halimah Kiysa Fiddien[4], Alaqso Fattar Djuson [5]*

[12345]Departement of Computer Engineering, Universitas Pendidikan Indonesia, Indonesia
Correspondence E-mail: deviaprianti@upi.edu

ABSTRACT

This study investigates the optimization of Access Point (AP) selection for microcontroller-based devices within smart building infrastructures using a graph theory approach. The primary challenge in multi-story wireless networks is the topological complexity that triggers signal interference, subsequently degrading the connectivity stability of IoT devices. In this research, the network is modelled as a weighted graph, where vertices represent access points and edges represent transmission paths weighted by distance. The Welch-Powell algorithm is implemented to optimize frequency channel allocation, effectively minimizing co-channel interference between adjacent nodes. Furthermore, Dijkstra's algorithm and Traveling Salesman Problem (TSP) logic are applied to determine the most efficient communication routes. The results demonstrate that transmission distance is the most significant parameter in optimal AP selection; a distance variance of 1.2–1.7 m is proven to correlate with improved signal quality. Meanwhile, frequency channel alignment and interference management serve as crucial secondary factors in optimizing connection stability. This approach provides a mathematical framework for strategic AP placement to ensure reliable data communication for microcontroller-based devices in high-rise environments.

© 2026 Universitas Pendidikan Indonesia

ARTICLE INFO

Article History:
Submitted/Received 10 Mar 2026
First Revised 23 Mar 2026
Accepted 31 Mar 2026
First Available online 01 Apr 2026
Publication Date 01 Apr 2026

Keyword:
Graph Coloring,
Access Point Selection,
Welch-Powell Algorithm,
Smart Building, Simulation.

1. INTRODUCTION

The development of Internet of Things (IoT) technology is driving the transformation of conventional building concepts into smart Buildings where building structures, services, and management are digitally integrated to produce environmentally friendly structures and support user productivity [1]. The integration of various IoT devices as sensor and actuator data processing units allows buildings to operate independently in managing energy efficiency, occupant comfort, and security systems. However, the reliability of this system is highly dependent on the stability of the wireless connection because microcontrollers scattered throughout the building require continuous internet network access to send data to the cloud and receive operational commands in real-time.

The main problem that often arises in the implementation of smart buildings is communication failure due to suboptimal placement of Access Points (AP) [2]. The physical structure of the building such as concrete walls, glass partitions, and the presence of other electronic devices often cause signal attenuation and interference. For IoT devices that generally have limitations in antenna sensitivity and transmission power, even the slightest signal interference can cause high latency or complete disconnection. This not only reduces system efficiency but also poses a fatal risk to critical functions such as fire detection or access security.

To overcome problems related to wireless networks, both regarding access point placement, signal interference management or other wireless network optimizati on, many studies have been conducted. In a study conducted by [3], the Simulated Annealing and tri literation methods were applied to determine the optimal access point position in an office building to reduce blank spots. Another study conducted by [4] used a graph theory approach to select Wi-Fi channels aimed at reducing network interference. A graph theory approach was also used in a study [5] to optimize radio frequency allocation in East Java through simulation. By using the graph colouring method, the number of frequencies produced in the simulation was fewer than the number of frequencies used in reality. The graph colouring method was also used in a study [6] to solve interference problems in an uncoordinated IEEE 802.11 network. Another study [7] applied the Kruskal algorithm in designing an access point network in a university building.

Graph theory is a field of mathematics with a very relevant approach in modeling problems regarding wireless network topology [8]. In this study, simulations are carried out by modeling the network center and access points as graph nodes, while the weights on the graph edges represent the distances between nodes. By utilizing graph coloring, access point channels can be allocated efficiently so that interference on adjacent nodes can be minimized. In certain scenarios, such as when there are mobile IoT devices or when the communication path must pass through several nodes, the Dijkstra algorithm is utilized to find the most efficient communication route. Therefore, this study proposes a graph theory-based simulation to determine the optimal Access Point configuration, in order to ensure connectivity stability in smart building infrastructure.

2. METHODS

In this study, the network is modeled as a weighted graph, which will then implement the Welch-Powell algorithm to group the main access points on each floor and the Dijkstra

DOI: https://doi.org/10.17509/coelite.v5i1.99031
p- ISSN 2829-4157 e- ISSN 2829-4149

algorithm to determine the optimal transmission distance for IoT devices. The stages used in this study can be seen in **Figure 1**.

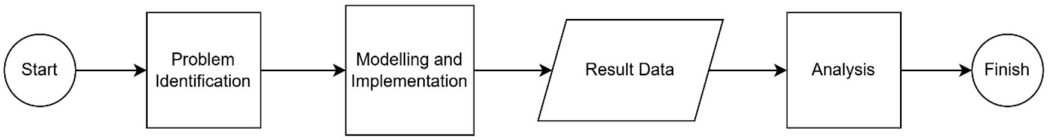


Figure 1 Research Flow

The research phase begins with problem identification, specifically examining the issue of selecting the appropriate Access Point position in buildings, particularly multi-story buildings. Often, installations are carried out without considering the extent to which signals can collide (interference), both on the same floor and between different floors. This results in unstable or slow internet connections in buildings. In this study, a visualization was developed to help map Access Point locations and automatically select the most optimal Access Point (AP).

The second stage is to build a graph model that aims to represent the network structure, so that the relationship between devices can be analyzed systematically and measurably. The graph model to be built is a undirected weighted graph where the nodes represent important elements in the signal network, namely the server as the main controller of the system and the access point, the edge describes the relationship or connection between two nodes in the graph and the weight indicates the transmission distance between the access point to the server or the device to the access point. With this approach, each device involved in the network system is modeled as a node in the graph, so that the interaction and connection between devices can be analyzed comprehensively. Meanwhile, an edge is formed when two nodes are within signal range and allow direct communication.

In this study, edges not only indicate the presence of a connection but also indirectly represent various technical network parameters, such as data transmission path, signal strength, interference level, network traffic load, and distance between devices. Edge weights are determined based on distance, signal quality, or interference level. Therefore, the lower the edge weight, the better the resulting connection quality and the lower the potential for interference between devices.

The next stage of the research is to implement the algorithm used based on the graph model that has been constructed. The Welch-Powell algorithm is used to group nodes based on similar signal characteristics, such as device type, signal strength, and frequency channel used. In theory, in the node coloring process, there is a rule that two neighboring nodes must not have the same color because they have the potential to cause signal interference. Therefore, graph coloring is used to avoid channel interference between access points. This, node coloring helps map network conditions, making it easier to identify areas prone to interference.

Meanwhile, an optimization algorithm is used to determine the most efficient communication path and network configuration. The algorithm used to determine the communication path of the device is the Dijkstra Algorithm [9]. This algorithm is used to find the communication path that minimizes the total distance or path weight between nodes, so that the most efficient network connection path is obtained. In accordance with the research objectives, the implementation of this algorithm aims to minimize the signal transmission

DOI: https://doi.org/10.17509/coelite.v5i1.99031
p- ISSN 2829-4157 e- ISSN 2829-4149

distance so that it can increase network stability for IoT devices. The implementation of the Dijkstra algorithm in determining the communication path follows the flowchart in **Figure 2**.

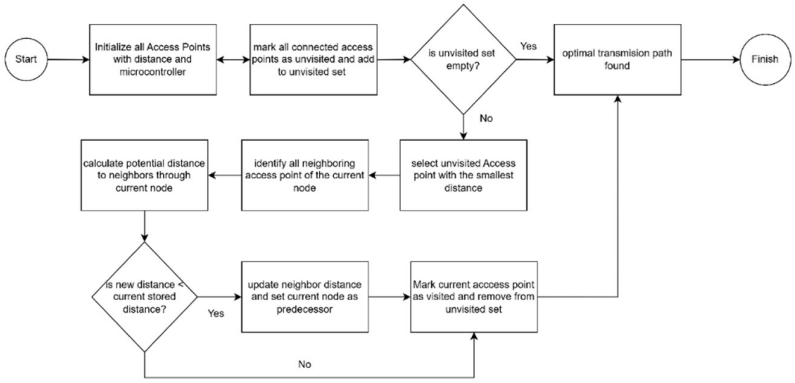


Figure 2. Dijkstra Algorithm for access point selection

The effectiveness of an access point in serving IoT devices is greatly influenced by the distance between the devices. The closer distance between the device and the access point is increase signal quality, smaller the delay, decrease interference level and more efficient the device's power consumption. Conversely, the further device is from the access point, the greater the transmit power required, the lower the signal quality, the greater the potential for interference, and the higher the risk of data loss. By implementing this algorithm, the system will find the optimal path that connects all devices in the network with the minimum total distance, so that the network connection path formed is more efficient, stable, and in accordance with the needs of the smart building system.

The algorithms are implemented in a program using the Python programming language, supported by libraries such as *math* library to perform advanced mathematical operations and the *tkinter* library to build user interface (GUI) applications used to visualize the simulations. Data from the simulations will be analyzed to provide an overview of the application of graph theory approaches to network optimization in smart buildings.

3. RESULTS AND DISCUSSION

The simulation applied several scenarios, including the number of floors in the building and the ideal transmission distance, taking into account the conditions of obstacles in the building. In this study, the simulation was conducted using two main scenarios: a single-story building and a multi-story building. The difference in the number of floors in the simulation was intended to observe the effect of the number of floors on the resulting optimal position. The interference distance, as a safe distance to minimize potential interference between two access point nodes, was a parameter set in this simulation. In multi-story buildings, the effect of placing access points on different floors will also be analyzed.

Figure 3 shows the simulation results for access point placement in a one-story building. Using an interference distance of 100 meters, assuming the building is proportional to the interference distance, with minimal obstructions, and a frequency with long range, the number of access points that can be placed indoors, based on the simulation, is 5. The distance between the access points and the server is calculated using the program, with detailed values shown in **Table 1**.

DOI: https://doi.org/10.17509/coelite.v5i1.99031
p- ISSN 2829-4157 e- ISSN 2829-4149

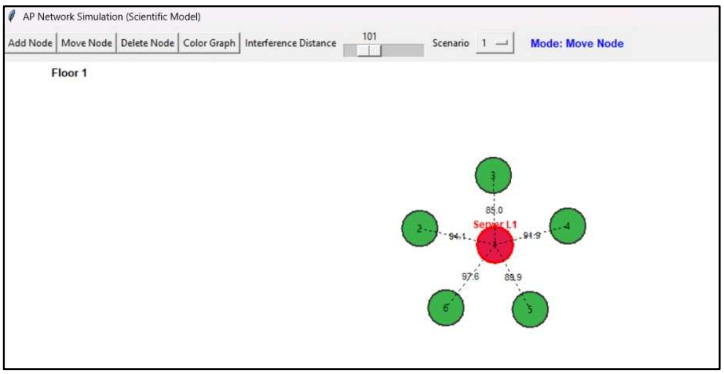


Figure 3. Simulation result of access point placement in a one-story building

For a single-story building, access point placement in this simulation is only affected by the interference distance parameter. The smaller the interference distance parameter, the fewer access points connected to the same channel. The green color shown in Figure 3 indicates that each access point is not adjacent.

Table 1. Access point distance to Server for o ne-story building simulation

Server Number	Access Point Number	Distance (meter)
1	2	94.1
1	3	85
1	4	91.9
1	5	89.9
1	6	97.6

The simulation continues with the next scenario, namely a multi-story building. In this simulation, three floors were selected as an additional simulation parameter, in addition to the interference distance. By using the same interference distance parameter value as the previous scenario, the results of the simulation of access point placement in a three-story building can be seen in **Figure 4.** To visualize the difference in floors in 2D, a horizontal line is used to separate the first floor and the next. The difference in floors is also marked by the difference in connection to the server. The use of 2D takes into account the influence of interference from access points on different floors. Because in reality, access points can transmit signals vertically so the potential for interference is very high [10].

The simulation results show significant differences between single-story and multi-story buildings. One difference is that fewer access points are required per floor in single-story buildings to minimize potential interference.

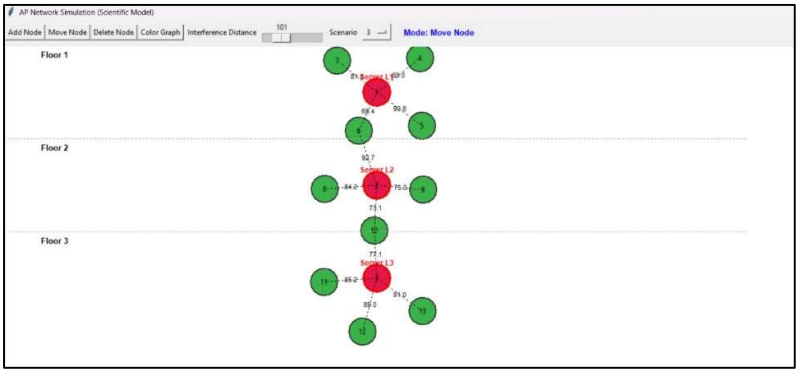


Figure 4. Simulation result of access point placement in a multi-story building

DOI: https://doi.org/10.17509/coelite.v5i1.99031
p- ISSN 2829-4157 e- ISSN 2829-4149

The transmission distance data between the access point and the server on each floor can be seen in **Table 2**. The similarity in color usage shown in the simulation indicates that the generated access points do not cause interference with each other. This is the result of the implementation of graph coloring using the Welch-Powell algorithm. Node coloring helps map network conditions, making it easier to identify interference-prone areas based on the distance parameters implemented in the simulation.

Table 2. Access point distance to Server for multi-story building simulation

Floor	Server Number	Access Point Number	Distance (meter)
1st	1	4	89
	1	5	90.8
	1	6	68.4
	1	7	81
2nd	2	6	92.7
	2	8	84.2
	2	9	75.3
	2	10	73.1
3rd	3	10	77.1
	3	11	85.2
	3	12	89
	3	13	91

In addition to the graph coloring implementation, the Dijkstra algorithm is used to determine the optimal communication path for devices that will connect to the access point to the server. In accordance with the scenario used to determine the placement of the access point, to determine the communication path for the device will be used two scenarios: in a single story building and a multi-story building. The configuration or placement of the access point used in this Dijkstra implementation uses the results obtained in the previous section (the generated access point configuration).

The implementation of searching for a communication path between a device and a server in a single story building with an access point placement configuration as shown in Figure 3 can be seen in **Figure 5**. In this simulation, there is a device whose communication path to the server will be analyzed using the Dijkstra algorithm. Device placement can be done randomly, so the selection of the closest path will depend heavily on the position of the nearest access point.

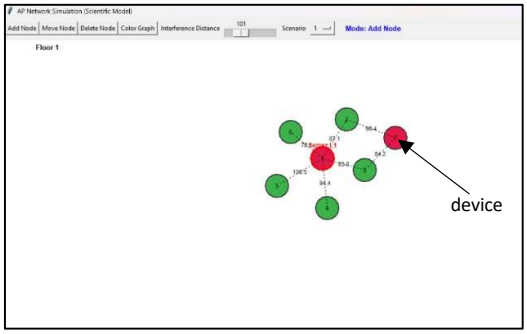


Figure 5. Simulation of communication path in a single-story building

DOI: https://doi.org/10.17509/coelite.v5i1.99031
p- ISSN 2829-4157 e- ISSN 2829-4149

In this simulation, there is one device modeled by node 7. The implementation of Dijkstra's algorithm for this scenario begins with this node (Node 7). Node 7 is connected to the access point modeled by node 3 and node 2. The shortest distance selected is from node 7 to node 3.

The calculation starts from the point with the shortest value in step 1, namely node 3. Node 3 is directly connected to node 1, which is the destination node. The resulting weight of this path is 167.2 m. After that, the calculation will be repeated until all nodes are visited to determine the resulting shortest distance. The resulting path weight will be compared with the existing path. **Table 3** shows the Dijkstra implementation for this scenario.

Table 3. Djikstra Algorithm Implementation for single story building simulation

Step	Node	1	2	3	4	5	6	7	Shortest distance to the next node	
1	7	∞	96.4	84.2	∞	∞	∞	0	96.4	2
2	2	∞	96.4	84.2	∞	∞	∞	0	84.2	3
3	3	167.2	96.4	84.2	∞	∞	∞	0	167.2	1
4	1	167.2	96.4	84.2	261.6	267.7	245.2	0	167.2	6
5	6	167.2	96.4	84.2	261.6	267.7	245.2	0	167.2	4
6	4	167.2	96.4	84.2	261.6	267.7	245.2	0	167.2	5
7	5	167.2	96.4	84.2	261.6	267.7	245.2	0	167.2	

Based on **table 3**, the path that will be selected by the device is through the access point modeled by node 3 with a total distance of 167.2 meters and the selected path is node 7 – node 3 – node 1. The simulation continues with the multi-story building scenario. The simulation results can be seen in **figure 6**.

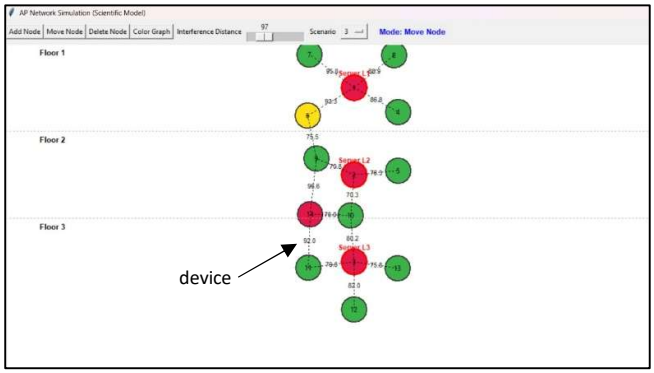


Figure 6. Simulation of communication path in a multi-story building

In this simulation, there is one device modeled by node 14, located on the 2nd floor. The device will connect to the nearest access point to go to the main server on the 1st floor. The shortest path calculation using the Dijkstra algorithm in this scenario produces the shortest path node 14 - node 9 - node 6 - node 1, with a total weight (distance) of 265.4 m.

4. CONCLUSION

The Dijkstra algorithm used to determine the best transmission path from a device connected to an access point to the network hub for each scenario shows that optimal access point selection can be done by considering the transmission distance generated by the program. In this study, the simulation was conducted in 2 dimensions, so the physical representation of wireless signals in 3-dimensional space was only done through a vertical

DOI: https://doi.org/10.17509/coelite.v5i1.99031
p- ISSN 2829-4157 e- ISSN 2829-4149

interference parameter approach. In addition, the coloring algorithm used did not consider other factors, such as external interference or room conditions.

5. ACKNOWLEDGMENT

We would like to thank our colleagues at Computer Engineering Study Program for their helpful discussions and suggestions regarding this research and also to the anonymous reviewers for their constructive comments that helped improve the quality of this manuscript.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

7. REFERENCES

[1] Mangaroska, V., Mangaroski, K. & Mangaroska, L. (2024). *Smart buildings as architectural international technological concept* [Članak]. Zbornik radova sa Nacionalne konferencije sa međunarodnim učešćem – Zelena Gradnja 2024, Niš, Srbija. https://doi.org/10.5937/greenb24023M

[2] Das, S.R., Saha, D., Sinha, B.P. *et al.* (2024). Strategic placement of access points for message communication in a smart building environment. *Innovations Syst Softw Eng* **20**, 425–433. https://doi.org/10.1007/s11334-022-00466-2

[3] Arif, Muhammad. and Utomo, Siswo Mardi. (2022). *Penggunaan Metode Simulated Annealing Dan Trilateration Untuk Optimasi Penempatan Posisi Titik Akses Studi Kasus: Kantor Kesyahbandaran Dan Otoritas Pelabuhan (Ksop) Kelas 1 Tanjung Emas*, Jurnal Komputer Grafis, 15(2):245-253

[4] Orden, David., Guzman, Jose Manuel Gimenez., Mestre, Ivan Marsa-Maestre., and La Hoz, Enrique. (2018). Spectrum Graph Coloring and Applications to Wi-Fi Channel Assignment. *Symmetry*, 10, 65

[5] Putra, Fadhlan Septianto Eka., Darmaji, and Soetrisno. (2017). Implementasi Metode Pewarnaan Graf Menggunakan Algoritma *Welch Powell* untuk Simulasi Penerapan Frekuensi Radio di Jawa Timur, *Jurnal Sains dan Seni ITS*, 6(2), 2337-3520

[6] Gimenez-Guzman, J.M.; Marsa-Maestre, I.; de la Hoz, E.; Orden, D.; Herranz-Oliveros, D. (2023). *Channel Selection in Uncoordinated IEEE 802.11 Networks Using Graph Coloring. Sensors*, 23, 5932. https://doi.org/10.3390/s23135932

[7] Putra, E. Yohanes dan Riti, Yosefina F.(2025). Rancang Jaringan Access Point Menggunakan Algotima Kruskal, *Indexia*, 7(1)

[8] Rusek,K., Suárez-Varela,J., Almasan,P., P. Barlet-Ros and A. Cabellos-Aparicio.(2020) "RouteNet: Leveraging Graph Neural Networks for Network Modeling and Optimization in SDN," in *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 10, pp. 2260-2270, doi: 10.1109/JSAC.2020.3000405

[9] Hadi, H. M., & Ibrahim, I. M. (2025). A Comprehensive Review of Shortest Path Algorithms for Network Routing. *Asian Journal of Research in Computer Science*, *18*(3), 152–175. https://doi.org/10.9734/ajrcos/2025/v18i3584

[10] Hou, B., & Wang, Y. (2023). Positioning by floors based on WiFi fingerprint. *Measurement Science and Technology, 35*. https://doi.org/10.1088/1361-6501/ad179e