JOURNAL OF ADVANCES IN

ENGINEERING AND TECHNOLOGY

VOLUME 4, ISSUE 2
MARCH 2026

ISSN 2961 - 5410

Civil Engineering | Electrical and Electronic Engineering | Materials Engineering | Mechanical Engineering | Mechatronics Engineering | Quantity Surveying

About the Journal of Advances in Engineering and Technology

The Journal of Advances in Engineering and Technology (JAET) is an international, open access, double blind peer-reviewed journal. It is published by the Faculty of Engineering of Sri Lanka Institute of Information Technology (SLIIT). The JAET aims at fostering research and development work in Engineering and Technology and bringing researchers on to a common platform. Furthermore, JAET will also accept review articles on appropriate subject areas including concept papers of academic opinions, book reviews, etc. for publication therein.

Notes for the authors: Please see page VI

All correspondence should be addressed to:

Editor-in-Chief
Journal of Advances in Engineering and Technology
Faculty of Engineering
SLIIT
Kandy Road, Malabe, 10115
Sri Lanka

Email: **jaet@sliit.lk**

Official website: **https://jaet.sliit.lk**

ISSN 2950-7138 (Printed)

ISSN 2961 - 5410 (Online)

Randhi Printers: No. 22 Sesatha Watta, Waduramulla, Panadura

Journal of Advances in Engineering and Technology (JAET)

Volume IV Issue (II)-2026

Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

Journal of Advances in Engineering and Technology

Volume IV, Issue II, March 2026

CONTENTS

Faculty of Engineering
Sri Lanka Institute of Information Technology (SLIIT)
Malabe, 10115, Sri Lanka

Notes to the Authors

Journal of Advances in Engineering and Technology (JAET) is a biannual peer review journal which aims at publishing original, theoretical and practice-oriented research papers related to Engineering and Technology. This journal provides a forum for researchers, scholars, academicians, and practitioners in the field of Engineering at international level to discuss and disseminate their findings in advanced and emerging technologies in Engineering.

Articles/ contributions should be sent to:

Editor-in-Chief
Journal of Advances in Engineering and Technology (JAET)
Faculty of Engineering
SLIIT , New Kandy Road, Malabe, Sri Lanka
E – mail: jaet@sliit.lkmailto:editors.sjhs@sliit.lk

- Your article will be subjected to a double-blind review process by two reviewers who are experts in their relevant field. In order to facilitate anonymous review process, you are requested to provide the title page separately.
- The title page should include the title of the article, your full name, affiliation, address, email and telephone/ mobile number.
- Recommendations of two reviewers are necessary for the publication of your article.
- The article should not exceed 6,000 words.
- Abstract should be between 150-200 words with 4-5 keywords.
- Article should include the following:
- Introduction to the research statement/ background to the research/ rationale
- Methods and Material
- Results and Discussion
- Conclusions and Recommendations
- References
- Footnotes can be numbered and given at the end of the article.
- Referencing should conform to the APA style.
- Tables/ graphs should be included in the appropriate places. Longer tables should be added at the end of the article as annexes.
- Use Times New Roman Fonts - Abstract 11pt; Article 11pt; Spacing Single

Macaranga peltata Leaf Extract Mediated Green Synthesis of Iron Nanoparticles and Their Application in Organic Dye Removal

Dissanayake D.M.K.N.[a], Perera M.A.D.[a], Karunaratne M.S.A.[a], Pahalagedara M.N.[b]
[a] Department of Materials Engineering, Faculty of Engineering, Sri Lanka Institute of Information Technology, Malabe 10115, Sri Lanka.
[b] Department of Basic Sciences, Faculty of Allied Health Sciences, University of Sri Jayewardenepura, Nugegoda 10250, Sri Lanka
kasuni0620@gmail.com, ann000461@gmail.com, mudith.k@sliit.lk, madhavinp@sjp.ac.lk

ABSTRACT

This study presents an eco-friendly method for synthesizing iron nanoparticles (FeNPs) using *Macaranga peltata* leaf extract, and evaluate their potential for degrading the organic dye methyl orange (MO). The synthesis exploits phytochemicals in the leaf extract as natural reducing and stabilizing agents. The synthesized FeNPs were characterized using UV-Vis spectroscopy, FTIR, XRD, and SEM, confirming amorphous structure and particle sizes ranging from 34–94 nm. Catalytic activity was evaluated via MO degradation experiments, achieving 85.16% efficiency within 200 minutes. The study demonstrates a sustainable wastewater treatment solution using green nanotechnology.

KEYWORDS: *Green synthesis, Iron nanoparticles, Macaranga peltata, Methyl Orange, Dye degradation, Wastewater treatment.*

1 INTRODUCTION

Nanotechnology has revolutionized many fields, including medicine, electronics, and environmental science by introducing advanced material properties at the nanoscale (ElHelaly, 2013; Yusuf, 2019). Among the broad applications of nanotechnology, the treatment of environmental pollutants is one of the most promising areas in which nanoparticles have already demonstrated their outstanding ability to degrade contaminants (Ashraf et al., 2018).

Nanoparticle synthesis methods can be divided into three main approaches: physical, chemical, and biological. However, most of the conventional synthesis methods such as chemical and physical methods involve the use of toxic chemicals and energy-intensive processes, raising concerns about their overall environmental impact (Jiang et al., 2022). In recent years, green synthesis has emerged as a sustainable, environmentally friendly alternative to produce nanoparticles. This method employs biological entities, like plant extracts, bacteria, or fungi, for nanoparticle synthesis, thereby negating the use of toxic chemicals. Among these, plant extracts have gained a lot of interest due to the abundance of bioactive compounds present in them such as ketones, aldehydes, polyphenols, caffeine, isoverbascosides and carbohydrates, serving as natural reducing and stabilizing agents to convert the metal ions into nanoparticles in the synthesis process, simplicity in preparation, and cost-effectiveness (Ebrahiminezhad et al., 2018).

1.1 Plant-Mediated Green Synthesis of Nanoparticles

There are significant studies that impress various plants involved in the process, the sizes of the nanoparticles synthesized, and their applications. A prominent example of plant-mediated synthesis involves *Cupressus sempervirens* to produce zero valent iron nanoparticles. Ebrahiminezhad et al. (2018) reported the formation nanoclusters with diameters from 9 to 31 nm, synthesized using Mediterranean cypress *(Cupressus sempervirens)* leaf branches. The reaction between the iron solution and the leaf extract was observed by noticeable color change indicating the high reduction potential of phytochemicals in this plant extract. This study highlighted iron nanoparticles' potential for removing dyes from aqueous environments. The bark extract of *Pinus eldarica* has also been used in the green synthesis of silver nanoparticles, as reported by Iravani and Zolfaghari (2013).

This study underscored *Pinus eldarica* bark extract contains polyphenolic compounds characterized by hydroxyl and carboxyl groups making it an excellent candidate for nanoparticle synthesis. It has been revealed that the exposure of *Pinus eldarica* bark extract to Ag+ ions resulted in a color change in the reaction mixture from yellowish brown to dark brown, indicating the successful formation of silver nanoparticles. In this study, the effect of pH on nanoparticle size and morphology of was highlighted, and it was shown that at lower pH values, larger nanoparticles were synthesized. In another study, Ahmed and Mustafa (2019) published a wide-ranging studies on the synthesis of gold and silver nanoparticles using lemongrass *(Cymbopogon flexuosus)* and green tea *(Camellia sinensis)*. The study identified various phyto-constituents in these plants including tannins, terpenoids, flavonoids, ketones, aldehydes, amides, and carboxylic acids which are responsible for reducing of silver ions. Among these, tannins were found to play a crucial role in both the reduction and capping of silver nanoparticles.

1.2 Overview of Macaranga peltata

Macaranga peltata belongs to the family Euphorbiaceae, is one of the common pioneer arboreal species mainly distributed in Southeast Asia, particularly in India and Sri Lanka. This species is more popularly known as "kenda" in Sri Lankan and "chandada" in India. It may reach up to 10 meters in height exhibiting a resinous tree structure with its young parts covered with velvety hairs. Its leaves are alternately arranged, measuring between 20 and 50 centimeters in length and 12 and 21 centimeters in width, characterized by a broad ovate shape and a distinct palmately 9-nerved pattern. The flowering period occurs in January to February, with yellow-green flowers clustered in long panicles in the leaf axils, and the fruit is a spherical capsule that carries a single black seed (Megha M et al., 2024).

Figure 1. *Macaranga peltata* leaf

A review of the phytochemical profile of *Macaranga peltata* shows that it is rich in secondary metabolites such as flavonoids, stilbenes, tannins, and terpenes, isolated from leaves. Preliminary phytochemical screenings has revealed the presence of some bioactive compounds including carbohydrates, alkaloids, phenolic compounds, tannins, saponins, and steroids in the plant's ethanolic extracts (Megha M et al., 2024). These bioactive compounds have been attributed to several pharmacological activities including antimicrobial, antioxidant, anti-inflammatory, anticancer, antifungal action, and other properties such as wound healing and hepatoprotection (Megha M et al., 2024). The potential therapeutic applications of these bioactive compounds underscore the importance of *Macaranga peltata* in traditional medicine and its prospective application in modern therapeutic practice (Megha M et al., 2024). In addition, recent studies have explored the potential of Macaranga peltata extracts for the green synthesis of nanoparticles, capitalizing on its bioactive compounds.

1.3 Overview of Iron Nanoparticles (FeNPs)

Over the last few years, significant attention has been paid to iron nanoparticles (FeNPs) due to their unique properties and broad potential for applications in biomedicine, environmental remediation, and catalysis. There are different synthesis methods for FeNPs including chemical reduction, various physical methods, and green synthesis using plant extracts. Among the outstanding methods for synthesizing iron nanoparticles is chemical reduction, in which iron ions (Fe^{3+}) are reduced to zero-valent iron (Fe^0) using reducing agents such as sodium borohydride ($NaBH_4$). This method has demonstrated that the synthesized nanoparticles are around 100 nm in size, as indicated by the black colour observed during synthesis, a characteristic of iron nanoparticles (Abdeen et al., 2013).

However, green synthesis methods have emerged as an environmentally friendly alternative for producing iron nanoparticles. This method uses natural extracts derived from plants, rich in various phytochemicals that reduce iron ions to form nanoparticles and stabilize them. For example, green synthesis using *Cupressus sempervirens* produces zero valent iron nanoparticles. Ebrahiminezhad et al. (2018) reported the formation nanoclusters with diameters from 9 to 31 nm, synthesized using Mediterranean cypress *(Cupressus sempervirens)* leaf branches. The reaction between the iron solution and the leaf extract was observed by noticeable color change indicating the high reduction potential of phytochemicals in this plant extract. This study highlighted the potential of iron nanoparticles' removing dyes from aqueous environments. In another study, it was found that green tea leaf extract (GT-Fe NPs) which contains polyphenols, acts as reducing and capping agent, mediated the synthesis of iron nanoparticles within the size range of 40-60 nm. Fe NPs were identified by the formation of strong black precipitate during the synthesis process. The synthesized particles showed very good removability and therefore, have potential application in the field of environmental remediation (Shahwan et al., 2011).

1.4 Properties of Iron Nanoparticles Relevant to Dye Degradation

Some physiochemical properties of iron nanoparticles (Fe NPs) were identified by Wanakai et al. (2019). These nanoparticles' properties have enabled them to act efficiently as catalysts in the degradation of dyes and especially methylene blue. That study discussed green synthesis, high surface area, oxidation states, fenton and fenton-like mechanisms and catalytic efficiency. The iron nanoparticles were synthesized using aqueous leaf extracts from the plants such as *Galinsoga parviflora, Conyza bonariensis* and *Bidens pilosa*, as reported by Wanakai et al. (2019). This method is ecofriendly and yields non-toxic, stable nanoparticles, which are quite essential for degradations of dyes. Also, the study emphasized the high surface area.

Due to the high surface area-to-volume ratio and their role as catalysts, iron nanoparticles are interact with dye molecules more effectively thereby enhancing their degradation. The oxidation states noted in the study also fall under the properties of iron nanoparticles relevant to dye degradation. It was emphasized that Iron is one of the versatile in mechanical engineering, and that it exists in more valencies/oxidation states. Those are ferrous and ferric. This characteristic is rather important since it determines the degree of effectiveness of the nanoparticles in catalysing the reactions, particularly in the degradation process using hydrogen peroxide. Furthermore, Wanakai et al. (2019) reported that the fenton and fenton-like mechanisms are among the properties . The study mentioned that methylene blue degradation occurs via Fenton and Fenton-like reactions. Fenton reaction is the reaction between Fe ions and H_2O_2 and yields hydroxyl radical (-OH). These radicals are highly reactive and therefore play an important role in cleaving of the dye molecules (Wanakai et al., 2019).

1.5 Application of Fe NPs in Organic Dye Removal

To address the challenge of dye contamination, various remediation strategies have been contrived. Notably, among all the remediation methods, the incorporation of nanomaterials has emerged as a promising approach. Nanomaterials have a large surface area-to-volume ratio and unique chemical properties that make them highly effective for adsorbing and degrading dye molecules in wastewater

(Ruan et al., 2018a). For example, it has been demonstrated that nano iron particles promote the decolorization of mono azo dyes via advanced oxidation processes that produce reactive species that degrade dye structures (Raman & Kanmani, 2018). The green synthesis of iron nanoparticles (Fe NPs) has received significant attention , owing to their environmentally friendly nature and potential for application in wastewater treatment, especially in dye removal. Several plant extracts have been used in the biosynthesis of Fe NPs, which then effectively show the degradation of various dyes including Methylene Blue (MB), Methyl Orange (MO), and Malachite Green (MG). This synthesis route provides not only a green alternative to chemical methods but also enables fine-tuning of the catalytic properties of the nanoparticles making them suitable for various environmental remediation applications.

Giun Tan et al. (2018) investigated the preparation of iron oxide nanoparticles using plantain peel extracts and found the particles to be very effective in the removal of both Methylene Blue and Methyl Orange dyes. The study focused on the catalytic and magnetic properties of green synthesized iron oxide nanoparticles (G-FeONPs) which showed efficacy in wastewater treatment. The reported degradation efficiencies for Methylene Blue and Methyl Orange are quite high, indicating that G-FeONPs have the potential for practical applications. In a similar vein, Ebrahiminezhad et al. (2018) have studied the green synthesis of iron nanoparticles using *Cupressus sempervirens* yielding iron nanoclusters with a mean diameter of 19 nm. Their work focused on optimization conditions for the removal of Methyl Orange and their results showed that the green synthesized iron nanoclusters showed a 95% decolorization efficiency in 6 h time. The results showed the nanoparticles effectively degraded Methyl Orange, and therefore, the potential for practical applications in dye removal during wastewater treatment applications.

Shahwan et al. (2011) extended their knowledge on Fe NPs produced from green tea leaf extract, demonstrating their role in Fenton-like catalysts in the degradation of dyes. This study has shown a rapid degradation of Methylene Blue and Methyl Orange, in which kinetic analysis indicating that Methylene Blue degradation followed a second-order reaction, In contrast, Methyl Orange degradation followed a first-order reaction . In another study, Khoirotin et al. (2023) investigated the green synthesis of Fe_3O_4 nanoparticles using green betel *(Piper betle L.)* leaf extract, demonstrating ferromagnetic properties and efficient adsorption. This study identified the highest adsorption capacity of Fe_3O_4 for methylene blue at a contact time 120 minutes with 98.75% MB degradation.

Most previous studies have focused on different metal nanoparticles synthesized using different plant extracts and their applications in environmental remediation, especially in the removal of dyes. Although the application of green-synthesized nanoparticles for organic dye removal has been widely reported, the specific role of iron nanoparticles synthesized through *Macaranga peltata* leaf extract has remained underexplored. The existing gap in the literature provides an opportunity to investigate the effectiveness and possible mechanisms by which iron nanoparticles, derived from this plant extract, may enhance the removal of organic dyes, thereby addressing environmental pollution and meeting the demand for sustainable materials in nanotechnology.

2 METHODOLOGY

2.1 Preparation of *Macaranga peltata* leaf extraction

As an initial step, the leaves of *Macaranga peltata* were collected and washed multiple times to remove any dirt and other impurities. Then, 10 g of *Macaranga peltata* leaves were measured using an analytical balance. After that, the leaves were cut into small pieces and ground with a motor and pestle and mixed with 50ml of distilled water. Next, the sample was stirred at 80 ºC temperature for 30 minutes to extract bioactive compounds efficiently. After 30 minutes, the flask was removed from the hot plate and left for a few minutes until it reached room temperature. Finally, the resultant extract was filtered using Whatman filter papers to remove leaf debris.

2.2 Green Synthesis of Iron Nanoparticles

Green synthesis was done using *Macaranga peltata* leaves extract by adding 10ml *Macaranga peltata* leaf extract to 5ml of 0.1 M ferric chloride solution in 2:1 (v/v) ratio in a conical flask. Then, the solution containing ferric chloride was stirred with a mechanical stirrer for 3 hours while the leaf extract was added dropwise . At this time, there was an immediate color change after adding the leaf extract to the iron chloride solution. A color change from pale brown to blackish brown was observed indicating the formation of iron nanoparticles (Fe NPs). Then, the solution undisturbed at room temperature for 24 hours to allow the nanoparticles to settle at the bottom of the beaker. After 24 hours, synthesized nanoparticles were separated by centrifugation at 5000 rpm for 15 minutes. Then the separated nanoparticle sample was washed twice with water and once with ethanol. Finally, the synthesized nanoparticle sample was dried at room temperature and stored in a sealed container for subsequent analysis.

2.3 Kinetic Experiment of the Degradation of Methyl Orange Dye

First, the calibration curve for methyl orange dye was prepared. Then, a 50 mg/L solution of methyl orange was prepared as a model organic dye. 2 mL of dye solution was withdrawn from the 50 mg/L dye solution and diluted it up to 10 mL by adding water into a 10 mL volumetric flask. This solution was transferred to a sample vial to record initial UV-Vis data at t=0. 50 mg of FeNPs were added to 45 mL of 50 mg/L dye solution, along with 5 mL of 10% H_2O_2. The mixture was stirred on a magnetic stirrer at room temperature to ensure proper mixing. 2mL samples were withdrawn from the mixture at time intervals of 20, 40, 60, 80, 100, 120, 140, 160, 180, 200 minutes and the absorbance was recorded at 464 nm (characteristic wavelength for methyl orange) using UV-Vis spectrophotometer. The same procedure was followed without adding Fe NPs, serving as a control experiment to investigate the degradation efficiency of Fe NPs.

The dye degradation efficiency was calculated using the following equation.

$$Degradation\ Efficiency\ (\%) = \frac{C_i - C_f}{C_i} \times 100\%$$

In this equation, C_i denotes initial concentration of dye solution, and C_f denotes concentration of dye solution after time t.

2.4 Characterization of Synthesized Nanoparticles

1. Ultraviolet-Visible (UV-Vis) Spectroscopy:

Periodic optical absorbance scans were conducted to monitor the formation of nanoparticles and confirm their stability. Absorbances were taken during the experiment in the UV range of 200-800 nm using SHIMADZU UV-1900 UV-Visible spectrophotometer.

2. Fourier Transform Infrared (FT-IR) Spectroscopy

The infrared spectra of samples were analysed by Bruker Vertex 80 Fourier Transform Infrared Spectrometer (Bruker alpha (10031705) with specifications of scan range 4000-650 cm-1; 32 scans and 8 cm-1 resolution). FTIR analysis was used to identify functional groups present in the *Macaranga peltata* leaf extract responsible for reducing ferric irons to iron nanoparticles (FeNPs), elucidating the interaction between the leaf extract phytochemicals and synthesized nanoparticles.

3. Powder X-ray Diffraction (XRD)

The crystalline nature and phase composition of the synthesized FeNPs were examined using Bruker D8 Focus X-ray Diffraction Spectrometer.

4. Scanning Electron Microscopy (SEM)

SEM analysis was carried out using Hitachi SU6600 Scanning Electron Microscope to analyze the surface morphology, size and distribution of the synthesized nanoparticles.

3 RESULTS & DISCUSSION

3.1 UV-Visible Spectroscopy Results

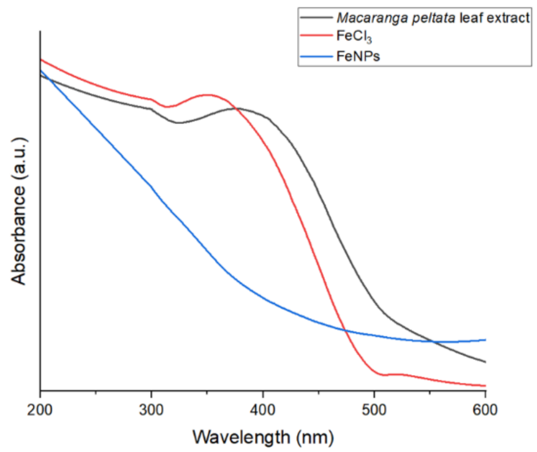


Figure 2. UV-Visible absorption spectra of a) *Macaranga peltata* leaf extract, b) $FeCl_3$ solution c) Synthesized FeNPs

Figure 2 illustrates the UV-vis spectra of a) $FeCl_3$, b) *Macaranga peltata* leaf extract and c) Fe nanoparticles in the 200-600 nm range. The formation of NPs was indicated by the color change of pale brown to blackish brown after the addition of the leaf extract to $FeCl_3$ solution. Leaf extract shows peaks at 280 and 450 nm while $FeCl_3$ shows a peak around 350 nm. The UV-vis spectra of the Fe NPs showed a continuous absorption from 200-600 nm. This well agrees with literature reported UV-vis spectra of amorphous Fe NPs.

3.2 FTIR Results

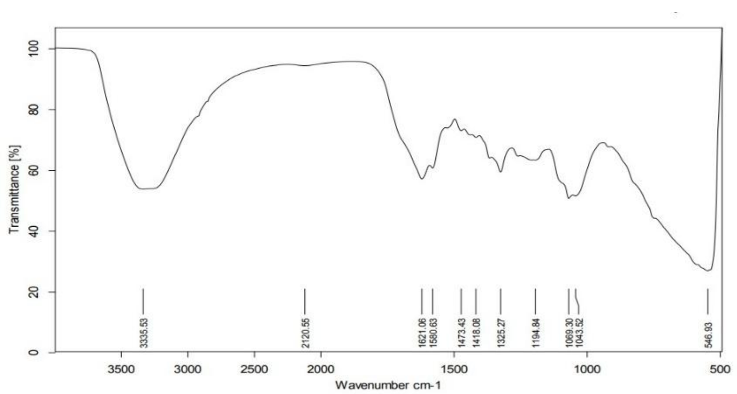


Figure 3. FTIR spectrum of Fe NPs synthesized using *Macaranga peltata* leaf extract

Figure 3 represents the FTIR spectrum of the prepared Fe NPs, indicating presence of organic compounds from *Macaranga peltata* leaf extract. This can be used to identify the functional groups of the phytochemicals responsible for reducing the metal precursor. It shows a prominent band at 3335 cm^{-1} due to intramolecular hydrogen bonded OH groups. In addition, the peaks at 1621 cm^{-1}, 1418 cm^{-}

^{1}, and 1100 cm^{-1} - 1030 cm^{-1} can be attributed to the C=O, in plane bending vibrations of O-H, and C-O-C stretching respectively.

3.3 XRD Results

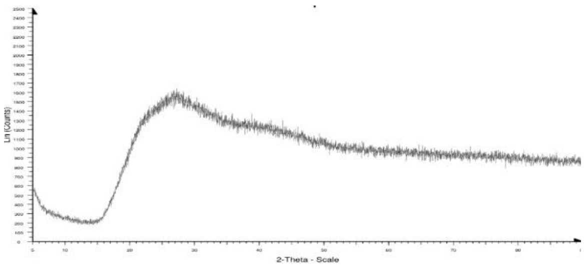


Figure 4. XRD pattern of Fe NPs synthesized using *Macaranga peltata* leaf extract

Figure 4 shows that the X-ray Diffraction of the synthesized iron nanoparticles using *Macaranga peltata* leaf extract. The pattern is deficient in distinct diffraction peaks, suggesting that the iron nanoparticles are amorphous in nature. The broad hump present at about $2\theta = 25^0$ is attributed to organic materials from *Macaranga peltata* leaf extract which can act as reducing and capping agents absorbed on the Fe NP surface.

3.4 SEM Results

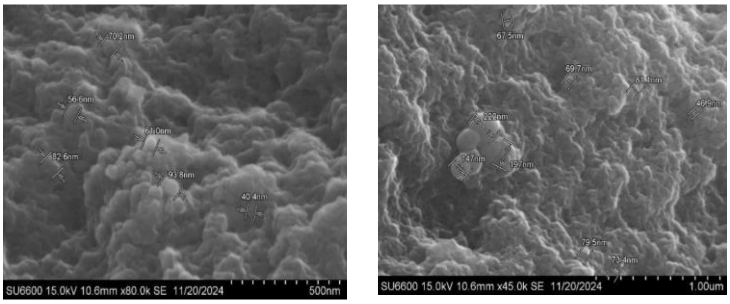


Figure 5. FE SEM Images of Fe NPs synthesized using *Macaranga peltata* leaf extract

The FE-SEM images of Fe NPs are shown in Figure 5. These images suggest that the particles agglomerate to form irregular clusters. This is due to the capping of Fe NPs with phytochemicals present in *Macaranga peltata* leaf extract. The diameter of these clusters' ranges from 34-94 nm.

3.5 Preparation of the Calibration Curve

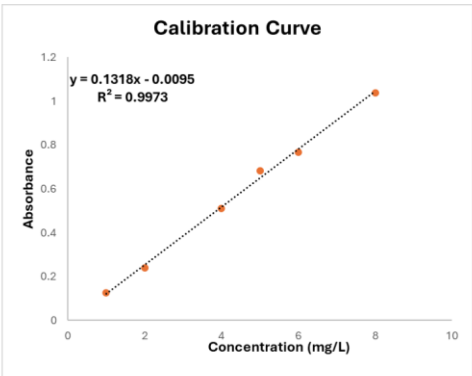


Figure 6. Calibration curve of methyl orange (concentration range: 1mg/L -8mg/L)

Figure 6 shows the calibration curve of methyl orange (MO), and the curve was prepared by taking known concentrations of 8,6,5,4,2,1 mg/L of MO as x-axis and absorbance at 464 nm as y-axis. The curve obeys Beers law in the studied concentration range. The linear regression of the calibration curve is y = 0.1318x - 0.0095; where y is absorbance and x is the concentration of MO (mg/L). The coefficient of correlation (R^2) of the regression analysis is 0.9973. The calibration curve was linear at the maximum wavelength and was therefore, suitable for estimating MO concentration.

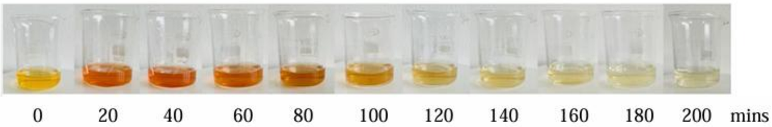


Figure 7. Degradation of MO in the presence of FeNPs

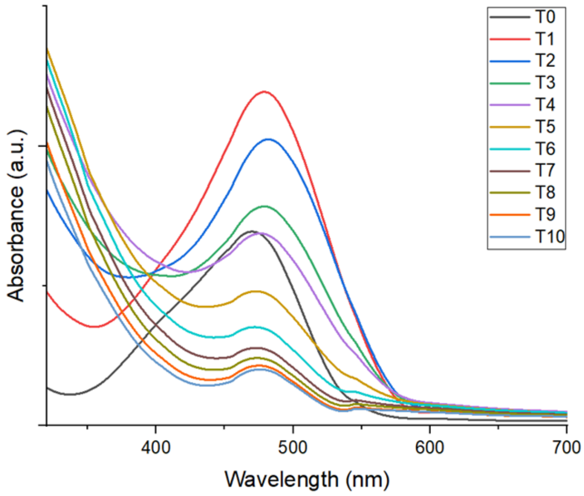


Figure 8. Kinetics of degradation of MO in the presence of the green catalyst-Fe NP (200-800 nm)

The degradation of MO using FeNPs and H_2O_2 was monitored using 2 mL withdrawing 2 mL samples from the mixture prepared in a previous step at time intervals of 20, 40, 60, 80, 100, 120, 140, 160, 180, 200 minutes and measuring absorbance over time. Figure 8 shows a gradual decrease in absorbance at 464 nm, indicating effective degradation. A significant reduction can be observed over 200 minutes, with final absorbance of 0.185, corresponding to a decrease in MO concentration and demonstrating the catalytic activity of FeNPs. The degradation process followed pseudo-first-order kinetics, with MO concentration diminishing consistently over time.

Figure 7 shows the visible fading of orange color, and the solution becomes significantly lighter by the end of 200 minutes, providing a qualitative representation of dye degradation process. The progressive color change visually supports the catalytic activity of FeNPs in enhancing dye degradation. In the initial stages of the reaction in Figure 7, a slight increase in absorbance was observed before the gradual decrease associated with dye degradation. This transient rise can be attributed to the formation of coloured intermediate species during the partial oxidation of methyl orange that absorb at the same or other wavelengths and/or increased light scattering from suspended catalyst particles and oxygen bubbles generated by the H2O2 decomposition. As the reaction progressed, these intermediates were further oxidized and turbidity reduced, leading to a steady decline in absorbance as the dye was removed
.

MO decolorization was observed under the same conditions, using only H_2O_2 and without Fe NPs, to assess the capability of H_2O_2 alone to degrade the dye . No significant decolorization was observed for hydrogen peroxide alone (Figure 9).

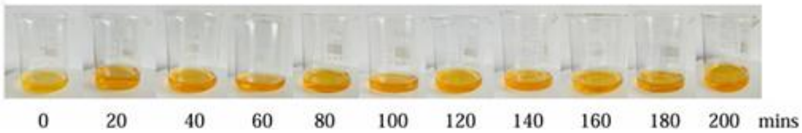


Figure 9. Degradation of MO in the absence of FeNPs

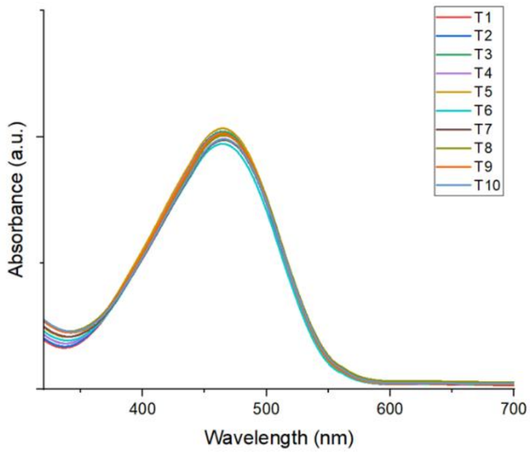


Figure 10. Kinetics of degradation of MO with just H_2O_2 and without Fe NPs (200-800 nm)

Figure 10 illustrates the variation of absorbance over different wavelengths during the control experiment using H_2O_2 without FeNPs. The absorbance remains relatively constant, fluctuating between 1.023 and 1.049 with minimal reduction over 200 minutes. Unlike the experiment with FeNPs, no notable reduction in the absorbance at the characteristic wavelength of 464 nm is evident. This indicates that, in the absence of FeNPs, H_2O_2 alone does not significantly contribute to MO degradation. Any observed changes in absorbances may be due to experimental noise or minor interactions between MO and H_2O_2. This graph does not exhibit exponential decay characteristic of pseudo-first-order kinetics. Figure 8 shows no visible change in orange color throughout the 0-200 minutes duration. The orange hue remains prominent, indicating that H_2O_2 alone is ineffective in degrading methyl orange.

Degradation of MO over time with Fe NP and MO with H_2O_2 only (without Fe NP)

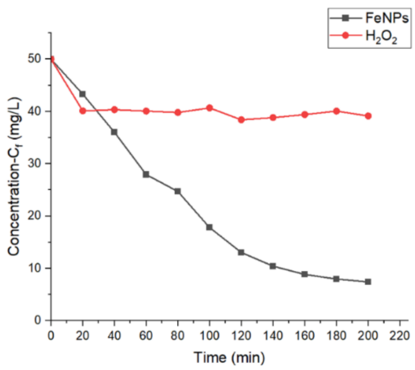


Figure 11.(a) Degradation of MO over time with Fe NPs + H_2O_2 b) Control- MO with H_2O_2 only (without Fe NPs)

In Figure 11, the “With FeNPs” curve shows a steep decline, visually representing the catalytic effect of FeNPs in breaking down MO.

Dye degradation Efficiency Evaluation in the presence of FeNPs

The initial and final concentration of MO can be determined using the calibration curve and the degradation efficiency (%) was 85.16 % in the presence of the Fe NPs. This indicates high efficiency of the Fe NPs in catalyzing MO degradation. Degradation efficiency of MO in the absence of Fe Nps remains relatively flat or slightly fluctuate, indicating the limited impact of H_2O_2 alone.

Dye degradation Efficiency Evaluation in the absence of FeNPs

The percentage degradation in the absence of FeNps was calculated to be 22.00 %. This low efficiency further confirms the ineffectiveness of H_2O_2 alone in degrading MO without the catalytic activity of Fe NPs. The combined graph (Figure 11) clearly illustrates the enhanced degradation kinetics when Fe NPs are present, validating their role as an effective green catalyst.

4 CONCLUSION

This work successfully demonstrated the green synthesis of FeNPs using *Macaranga peltata* leaf extract and assessed their catalytic efficiency in the degradation of methyl orange dye. The results proved that *Macaranga peltata* leaf bioactive compounds, mainly flavonoids and tannins, are effective both reducing and stabilizing agents for FeNPs. The green synthesized nanoparticles were amorphous in nature with sizes ranging from 34 to 94 nm, which are optimal for catalytic applications.

FeNPs synthesized in this work degraded 85.16% efficiency under optimized conditions within 200 minutes (3 hours and 20 minutes). Therefore, synthesized FeNPs may serve as a cost-effective environmentally friendly material for wastewater treatment. The study filled a critical research gap by exploring the underutilized *Macaranga peltata* for nanoparticle synthesis, linking sustainable nanotechnology with practical environmental applications. However, the following limitations such as single dye focus, concerns of reusability of nanoparticles and toxicity of by-products need to be addressed.

5 ACKNOWLEDGEMENTS

We would like to acknowledge the Sri Lanka Institute of Information Technology (SLIIT) for providing necessary resources for this research. Additionally, special thanks goes to Ms. Thavishi Chandratilake, for providing of lab resources and guidance on lab work, and to the Laboratory Manager, Mr.Sehan Jayasinghe from the Faculty of Humanities and Sciences at SLIIT, for providing access to lab resources. We also wish to express our great appreciation to the Sri Lanka Institute of Nanotechnology (SLINTEC) for offering analytical services for our samples.

REFERENCES

Abdeen, S., R.S, R. I., Geo, S., S, ornalekshmi, Rose, A., & P.K, P. (2013). Evaluation of Antimicrobial Activity of Biosynthesized Iron and Silver Nanoparticles Using the Fungi Fusarium Oxysporum and Actinomycetes SP. on Human Pathogens. *Nano Biomedicine and Engineering*, *5*(1). https://doi.org/10.5101/nbe.v5i1.p39-45

Ahmed, R. H., & Mustafa, D. E. (2019). Green synthesis of silver nanoparticles mediated by traditionally used medicinal plants in Sudan. *International Nano Letters*, *10*(1), 1–14. https://doi.org/10.1007/s40089-019-00291-9

Ashraf, M. A., Peng, W., Zare, Y., & Rhee, K. Y. (2018). Effects of Size and Aggregation/Agglomeration of Nanoparticles on the Interfacial/Interphase Properties and Tensile Strength of Polymer Nanocomposites. *Nanoscale Research Letters*, *13*(1). https://doi.org/10.1186/s11671-018-2624-0

Ebrahiminezhad, A., Taghizadeh, S., Ghasemi, Y., & Berenjian, A. (2018). Green synthesized nanoclusters of ultra-small zero valent iron nanoparticles as a novel dye removing material. *Science of the Total Environment*, *621*, 1527–1532. https://doi.org/10.1016/j.scitotenv.2017.10.076

ElHelaly, M. (2013). Nanotechnology, Occupational Health and Safety Concerns. *Journal of Postgenomics Drug & Biomarker Development*, *01*(03). https://doi.org/10.4172/2329-6879.1000116

Giun Tan, W., Ming Ng, W., Kang Lim, J., & Xin Che, H. (2018). Plantain Peel Mediated Green Synthesis Iron Oxide Nanoparticles, Surface Functionalization, and Them Performance towards Methylene Blue and Methyl Orange Dye Removal. *International Journal of Engineering & Technology*, *7*(3.36), 101. https://doi.org/10.14419/ijet.v7i3.36.29087

Iravani, S., & Zolfaghari, B. (2013). Green Synthesis of Silver Nanoparticles UsingPinus eldaricaBark Extract. *BioMed Research International*, *2013*, 1–5. https://doi.org/10.1155/2013/639725

Jiang, Z., Li, L., Huang, H., He, W., & Ming, W. (2022). Progress in Laser Ablation and Biological Synthesis Processes: “Top-Down” and “Bottom-Up” Approaches for the Green Synthesis of Au/Ag Nanoparticles. *International Journal of Molecular Sciences*, *23*(23), 14658. https://doi.org/10.3390/ijms232314658

Khoirotin, Nuhaa Faaizatunnisa, & Munasir Munasir. (2023). Green Synthesis of Fe3O4 Nanoparticles Using Green Betel Leaf Extract for Methylene Blue Adsorption. *Natural and Life Sciences Communications*, *22*. https://doi.org/10.12982/nlsc.2023.042

Megha M, Kavana D K, Rithin K, & Ramdas Bhat. (2024, October 20). *A Review On Phytochemistry And Pharmacology Of Macaranga Peltata» Article*. STM Journals. https://journals.stmjournals.com/article/article=2024/view=178897/

Raman, C. D., & Kanmani, S. (2018). Decolorization of Mono Azo Dye and Textile Wastewater using Nano Iron Particles. *Environmental Progress & Sustainable Energy*, *38*(s1), S366–S376. https://doi.org/10.1002/ep.13063

Ruan, W., Hu, J., Qi, J., Hou, Y., Zhou, C., & Wei, X. (2018a). Removal Of Dyes From Wastewater By Nanomaterials : A Review. *Advanced Materials Letters*, *10*(1), 9–20. https://doi.org/10.5185/amlett.2019.2148

Shahwan, T., Abu Sirriah, S., Nairat, M., Boyacı, E., Eroğlu, A. E., Scott, T. B., & Hallam, K. R. (2011). Green synthesis of iron nanoparticles and their application as a Fenton-like catalyst for the degradation of aqueous cationic and anionic dyes. *Chemical Engineering Journal*, *172*(1), 258–266. https://doi.org/10.1016/j.cej.2011.05.103

Wanakai, S. I., Kareru, P. G., Makhanu, D. S., Madivoli, E. S., Maina, E. G., & Nyabola, A. O. (2019). Catalytic degradation of methylene blue by iron nanoparticles synthesized using Galinsoga parviflora, Conyza bonariensis and Bidens pilosa leaf extracts. *SN Applied Sciences*, *1*(10). https://doi.org/10.1007/s42452-019-1203-z

Yusuf, M. (2019). Silver Nanoparticles: Synthesis and Applications. *Handbook of Ecomaterials*, 2343–2356. https://doi.org/10.1007/978-3-319-68255-6_16

Circular Economy Practices in Road Rehabilitation and Development in Sri Lanka

L. M. De Alwis[1], K A N Gunarathna[2], C.Kalugala[3]

Liverpool John Moores University, UK[1]
Sri Lanka Institute of information Technology, SL[2]
Skills College of Technology, SL[3]

ABSTRACT

The integration of Circular Economy principles into road rehabilitation and development offers a sustainable alternative to traditional linear construction practices. In Sri Lanka, where road infrastructure is crucial to economic growth and connectivity, the Circular Economy offers opportunities to reduce reliance on virgin materials, minimise environmental impact, and enhance long-term cost efficiency. This study explored applicable Circular Economy strategies, assessed current implementation practices, identified key challenges, and proposed viable solutions to support Circular Economy adoption in the Sri Lankan Road sector. A mixed-methods approach was employed, combining a comprehensive literature review with data from semi-structured expert interviews and a questionnaire survey of construction professionals. The findings indicate that while awareness of the Circular Economy is growing, its practical application remains limited due to barriers such as the absence of standardized technical guidelines, insufficient government incentives, limited stakeholder knowledge, logistical challenges in material sourcing and storage, and reluctance to shift from traditional methods. Data also highlights substantial potential benefits, including cost savings, reduced construction waste, increased material efficiency, and environmental improvements. Participants emphasized the importance of pilot projects, training programs, and policy support in promoting Circular Economy practices. To overcome existing barriers, the study recommends the development of clear Circular Economy specifications, financial incentives, capacity-building initiatives, and the establishment of centralized recycling infrastructure. These strategic actions can facilitate the transition toward a more circular and sustainable approach in Sri Lanka's Road construction and maintenance sectors.

KEYWORDS: *Circular Economy, Road Construction, Road Rehabilitation, Sustainable Infrastructure, Recycled Materials*

1 INTRODUCTION

Road infrastructure development is a critical component of the development of countries, and it serves as the foundation for economic growth, social connectivity, and enhances the transport logistics system of the country (Kodithuwakku et al., 2023). The road is a legally permissible way for vehicles and other traffic to traverse. Roads include passageways, side drains, culverts and land needed for the upcoming widening (Kodithuwakku et al., 2023). Road networks include highways, expressways, urban roads, rural roads, bridges, drainage systems, and pavements, all of which play a vital role in supporting increases in the country's gross domestic product. Each of these road networks is critical to develop and maintain. The highways are designed for long-distance and high-speed travel, facilitating efficient transportation between cities, industrial zones, and ports with enabling the transportation of goods and passengers. Rural roads are essential for connecting remote areas, agricultural zones, and villages to economic centres. Effective drainage and stormwater management systems are crucial for preventing floods and road deterioration. Proper drainage infrastructure includes culverts, stormwater drains, ensuring that excess water is diverted from the road surface without disturbing road functions. The transport sector contributes 10% to 20% of the Gross Domestic Product (GDP), where road transportation represents between 3% to 5% of the GDP, excluding the inputs of transport equipment, fuel, and infrastructure (Kodithuwakku et al., 2023).

Traditional road rehabilitation and development are heavily based on the use of virgin materials. Conventional road construction primarily uses natural soil, crushed rock, and cement concrete in the industry. Also Bituminous materials are heavily used in road construction and the development process (Kodithuwakku et al., 2023). Bituminous materials are used as surfacing material in all types of road construction. All these materials require significant amount of energy for extraction, processing, and installation. These processes release substantial amounts of greenhouse gases, particularly carbon dioxide, into the environment and increasing global warming and contributing to climate change.

The circular economy (CE) framework offers a promising solution to reduce embodied energy in road construction and rehabilitation**.** CE emphasizes resource efficiency and waste reduction through strategies such as material reuse, recycling, and optimised design.

The global rise in construction activity, particularly in developing nations, has led to a critical shortage of high-quality natural aggregates and an increase in material costs, resulting in environmental and economic concerns (Mishra et al., 2023). Simultaneously, growing demands for more durable and resilient infrastructure have intensified pressure on traditional materials, which often fail to meet long-term performance expectations (Plati, 2019). While CE strategies offer sustainable alternatives, their adoption in road construction is hindered by high upfront costs, inadequate regulations, and logistical difficulties (Silva & Fernando, 2022). Moreover, public sector agencies face institutional limitations, such as technical gaps and fragmented policies, which restrict their ability to implement sustainable practices effectively (Connolly et al., 2023).

2 LITERATURE REVIEW

Construction of road infrastructure is one of the main aspects of a country's development and the foundation of economic development, social connectivity, and enhances the nation's transport logistics system (Kodithuwakku et al., 2023). Conventional road building uses plenty of virgin resources, including aggregates, concrete, and asphalt, and is therefore resource-intensive. These materials require large-scale extraction, processing, and transportation, which translates to a high environmental footprint and high material cost (Hammond & Jones, 2008). According to the Ellen MacArthur Foundation, CE is restorative by design and aims to keep products, components, and materials at their highest utility and value at all times, distinguishing between technical and biological cycles (MacArthur, 2015). Benachio et al. and other authors defined CE for the construction industry as "the use of practices, in all stages of the life cycle of a building, to keep the materials as long as possible in a closed loop, to reduce the use of new natural resources in a construction project (Benachio et al., 2020).

1.1 Application of CE in Road Construction

2.1.1 Use of Recycled and Alternative Materials

CE practices promote the substitution of traditional virgin materials with recycled and alternative materials in road construction. Recycled Concrete Aggregate (RCA), derived from demolished concrete, is recognised as a viable substitute for natural aggregates in sub-base and base layers due to its self-cementing properties and high stiffness (Silva & Fernando, 2022). Crushed brick and glass, originating from construction and demolition waste, also offer utility in sub-base and asphalt applications (Kodithuwakku et al., 2023). Reclaimed Asphalt Pavement (RAP) is a key strategy to reduce the demand for virgin bitumen, lowering the embodied carbon of road infrastructure. The integration of polymer modifiers like crumb rubber further improves binder performance, reduces rutting, and supports longer pavement lifespans (Ashtiani et al., 2024). Fly ash and cement kiln dust are sustainable binders used to stabilise recycled aggregates and weak subgrades, offering economic and environmental advantages (Silva & Fernando, 2022).

2.1.2 Energy-Efficient Road Construction Methods

Energy-efficient construction practices in CE involve reducing emissions across the road's lifecycle from material extraction to end-of-life disposal. A major strategy includes the use of blended cement, where clinker is partially replaced with by-products like fly ash or slag, reducing CO_2 emissions during production (Worrell et al., 2001). Furthermore, optimising material transport through route planning and double cycling of RAP and HMA (hot mix asphalt) deliveries significantly lowers fossil fuel consumption and construction-related emissions (Kumari & Kulathunga, 2020 & Ashtiani et al., 2024). Recycling reinforced concrete waste (RCW) at the end of a road's lifecycle further conserves resources and supports circularity, providing environmental and cost benefits (Paranavithana et al., 2024).

2.1.3 Structural Performance of Recycled Materials

A major concern in CE adoption is whether recycled materials can meet structural demands. Studies have shown that pavements constructed using recycled materials can match or surpass the performance of conventional materials. For example, the inclusion of 20% RAP enhances rutting resistance and increases the stiffness of bituminous mixes, while 10% RAP improves resilient modulus from 5000 to over 6700 MPa (Silva & Fernando, 2022). RCA also improves subbase performance, with the potential to exceed 400 MPa in modulus when mixed with 25% crushed bricks. Additionally, incorporating 2% cement helps meet subbase standards under varied moisture conditions (Silva & Fernando, 2022). Fly ash is another high-performing additive for stabilising expansive soils, achieving a 26% increase in California Bearing Ratio (CBR) at a 15% mix in field tests in Melbourne. HiMA (Highly Modified Asphalt) stands out in terms of performance efficiency it allows for 20% thinner pavement layers without sacrificing structural integrity. In trials, HiMA outperformed conventional HMA in fatigue cracking and deformation under 20 million ESALs (Kluttz et al., 2017). This contributes to both material savings and lower emissions during production and transportation.

1.2 Economic Feasibility of CE in Road Development

The economic viability of CE principles in road development lies in their potential for long-term cost savings, despite initially higher investment. Sustainable pavement design emphasises longevity spanning 30 to 60 years, reducing the need for frequent maintenance, rehabilitation, and associated greenhouse gas emissions. Innovative solutions such as composite pavements and inverted asphalt layers offer high durability at lower lifecycle costs (Plati, 2019). CE-aligned techniques, such as the use of Cement Stabilised Rammed Earth, provide low-cost construction alternatives, especially suitable for resource-constrained regions (Silva & Fernando, 2022). Moreover, efficient material management can lower a substantial portion of costs attributed to materials, often exceeding 40% of total project costs, by minimising waste and leveraging recycled inputs (Kodithuwakku et al., 2023).

1.3 Policy and Regulatory Support for CE Adoption

Effective policy and regulatory frameworks are crucial for advancing CE adoption in road construction. Governments can foster circularity by setting mandatory reuse and recycling thresholds for construction projects, thereby boosting demand for secondary materials and reducing reliance on virgin resources (Ghaffar et al., 2020). Fiscal incentives and formal recognition for sustainable practices also play a pivotal role in attracting investment into CE innovations, especially those aimed at resource recovery from construction and demolition waste. Internationally, several countries have taken proactive steps. For instance, China has set a target to ensure 30% of all new buildings are prefabricated by 2026, as part of its broader CE strategy. It has also enacted policies emphasising the "reduce, reuse, recycle" principle (Li et al., 2024). Moreover, initiatives such as CERCOM (Circular Economy in Road Construction and Maintenance) project in Europe aim to institutionalise CE within national procurement frameworks. CERCOM promotes a Risk-Based Analysis Framework that evaluates technical, economic, and environmental risks associated with CE transitions, encouraging more robust and structured adoption across road infrastructure projects (Connolly et al., 2023).

1.4 Challenges to CE Implementation

Several authors have identified various types of obstacles and categorised them into six wide groups, namely social, cultural, awareness, technical, economic, and implementation. As helpful as this categorisation is, the present study shall adopt an exploratory approach, taking into consideration all possible obstacles when they emerge without restricting the inquiry to pre-defined categories (AlJaber et al., 2023).

The lack of standards for recycling and re-use is an impeding factor to the circular construction concept. Other studies also identify this lack of standards as a problem (Benachio et al., 2020). Furthermore, the lack of standardization often leads to the need for multiple costly and time-consuming certifications for alternative materials (World Green Building Council, 2023). Table 1 summarises the challenges to CE implementation.

Table 1: Challenges to CE Implementation – Summary

Challenges	**A**	**B**	**C**	**D**	**E**
High Initial/Upfront Costs	✓	✓	✓	✓	
Lack of CE Infrastructure	✓	✓			✓
Limited Knowledge and Expertise	✓	✓		✓	✓
Lack of Government Support/Incentives	✓	✓		✓	✓
Quality of Recycled Materials	✓	✓		✓	
Lack of rules and policies	✓	✓	✓		✓
Sources: (A).AlJaber et al., 2023 (B) Snigdhya et al., 2024; (C) World Green Building Council, 2023; (D) Agyemang et al., 2019; (E) Weerakoon & Thayaparan, 2023					

3 RESEARCH METHODOLOGY

With the completion of comprehensive literature review, semi-structured expert interviews were conducted to identify the challenges and barriers currently affecting the adoption of Circular Economy (CE) principles in Sri Lanka's Road construction sector. Semi-structured interviews offer a balance of flexibility and depth, making them more adaptable than structured interviews and more focused than unstructured ones. The interviews aimed to explore current CE practices, highlight implementation gaps, and capture expert perspectives on viable strategies for future integration. Subsequently, a questionnaire survey was developed based on findings from both the literature review and expert interviews. The survey was distributed among a broader group of stakeholders to validate key challenges, understand stakeholder perceptions, and evaluate the attractiveness and feasibility of CE strategies.

4 DATA COLLECTION AND ANALYSIS

4.1 Qualitative Analysis

4.1.1 Respondent's Profile

Interviews were conducted among Sri Lankan Quantity Surveyors (QSs) and engineers who are in consultancy and contractor organisations. Seven (7) semi-structured interviews were conducted for data gathering. The respondents' details were gathered through the first section of the semi-structured interview guideline. The summarized composition of respondents' profiles is presented in Table 2.

Table 2: The Composition of the Respondents

Interviewee	Designation	Experience (No. of Years)
I-01	Chartered Quantity Surveyor	8
I-02	Chartered Civil Engineer	23
I-03	Planning Engineer (civil engineer)	9
I-04	Project Manager (civil engineer)	20
I-05	Chartered Quantity Surveyor	19
I-06	Chartered Civil Engineer	24
I-07	Road Contractor	6

The semi-structured interviews conducted with professionals from Sri Lanka's Road construction sector revealed several critical insights into the existing status, challenges, and future strategies for adopting CE practices. Interviewees were drawn from government, consultancy, and contractor backgrounds, providing a comprehensive overview of the sector's position on CE.

4.1.2 Adoption of Circular Economy (CE) Practices

Expert interviews revealed that although all the participants were familiar with CE practices, their application in Sri Lankan Road construction projects remains limited. Respondents I01, I02, and I03 acknowledged awareness of CE principles but noted that actual implementation is seldom and often informal. For instance, I05 highlighted routine reuse of construction elements such as signage and milestones, while I04 provided examples where materials from demolished roads were reused as base layers in new projects. I01 also mentioned using crushed concrete in granular sub-base layers. These cases reflect a fragmented but emerging pattern of CE adoption.

4.1.3 Key Challenges in CE Implementation

Multiple challenges were identified that hinder the broader adoption of CE in road construction. The lack of technical standards and regulatory guidelines was cited by I01, I03, and I04 as the primary barrier. Engineers are reluctant to use recycled materials without standardised performance benchmarks. Moreover, I02 emphasised that both public and institutional awareness are limited, with misconceptions about the safety and durability of recycled materials. Public scepticism further hinders implementation, particularly in community-sensitive projects. Material availability and quality control were also highlighted. Interviewees I05 and I06 noted that recycled materials vary significantly in quality, and storage of materials such as RCA is logistically difficult without dedicated facilities. In addition, I01 and I05 pointed out the absence of updated testing labs, making verification of recycled material properties challenging. Resistance from stakeholders, including contractors and employers, also emerged as a theme. As I04 and I05 stated, conservative design perspectives and fear of performance issues discourage the use of alternative construction methods.

4.1.4 Current Use of Recycled and Alternative Materials

All experts confirmed using Recycled Concrete Aggregate (RCA), particularly in rural road sub-base applications. Reclaimed Asphalt Pavement (RAP) was also commonly used, especially for binder courses, as noted by I02, I04, and I06. Fly ash and cement kiln dust were used as stabilisers or filler materials in road layers, and I02 reported using quarry dust to reinforce subbase layers. Turfing reuse for drainage embankments was also mentioned. Despite these efforts, the industry faces technical and logistical constraints. Experts cited lack of design knowledge and material quality assurance, often due to inconsistent sources of recycled material and the absence of a central recycling supply chain. Approval bottlenecks, noted by I01 and I05, are frequent due to the lack of familiarity of regulatory bodies with non-conventional materials.

4.1.5 Benefits of using CE practice in Road Construction

All respondents affirmed that CE practices reduce material costs and environmental impact. The reuse of materials leads to savings in raw material procurement and reduces construction waste. I02 and I06 noted that CE improves project efficiency by saving time on material sourcing. I01 and I04 highlighted CE's contribution to long-term sustainability and raw material conservation. I05 emphasised an opportunity to develop local markets and circular supply chains, while I06 pointed to the growth of waste management industries as a secondary economic benefit.

4.1.6 Long-Term Cost Savings

All the expert interviewees strongly agreed that CE-based road construction works save considerable long-term costs, despite their comparatively greater initial investment. The consensus was based on some key economic benefits inherent in CE practices. Firstly, the reduction in raw material quarrying was mentioned as a major cost-cutting factor because CE emphasises the recycling and reuse of existing materials such as RAP and RCA, hence avoiding the expense of procuring copious amounts of virgin materials. This not only reduces procurement and transportation costs but also lowers environmental compliance expenses associated with quarrying and extraction. Secondly, it was noted by the respondents that landfill and disposal costs are significantly minimised through material recovery and reuse practices, reducing the economic and environmental cost of waste management. Also, recycling and processing of the recovered materials were described as considerably less compared to the energy and resource-intensive extraction and preparation of new materials.

4.1.7 Regulatory and Policy Needs

Government and institutional policy play a vital role in the effective implementation of CE practices in Sri Lanka's Road and infrastructure sector. I01, I02, I04, I05 and I07 indicate that there is an issue of the absence of lack of standard specification for supporting the use of recycled or alternative materials. Most existing specifications are made for use with virgin materials. Therefore, it is important to update the specifications that are relevant to use with recycling materials, as pointed out by the respondents. Also, I01, I03 and I04 noted that the introduction of Circular Procurement Policies is also crucial. They pointed out that the inclusion of CE-based Policies in Public procurement guidelines, thereby encouraging bidders to incorporate sustainable solutions. In the building construction sector, there is a standardised certification process to assess the building's environmental impact and sustainability, such as LEED and BREEM. I03, I06, and I07 mentioned that the road sector also needs such a standardised certification process to implement CE practice. Moreover, interviewees I01 and I05 emphasised that the Construction Industry Development Authority and Road Development Authority, as the primary regulatory bodies in Sri Lanka's Road Sector, should take the lead in initiating and implementing necessary policy changes to support the adoption of CE practices.

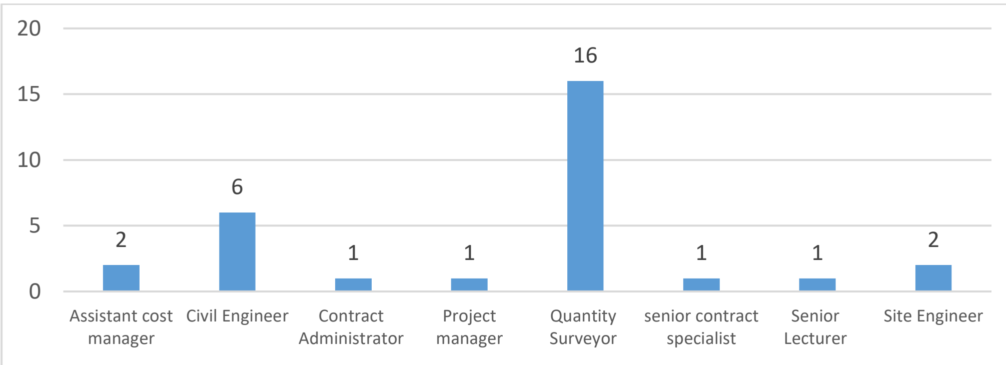
4.1.8 Financial Strategies to Support CE Implementation in Road Construction

The successful adoption of CE practices in road infrastructure projects depends not only on policy changes and technical enhancement but also on robust financial support strategies. In construction industry settings, contractors typically evaluate potential advantages and tangible benefits before adopting new construction methods. The decision to implement innovative approaches such as CE practices is often influenced by considerations of cost-efficiency. All the interviewees noted that it is important to provide tax benefits for the contractors and clients who implement CE strategies in road projects. The other widely mentioned strategy was the inclusion of CE in bid evaluation and procurement systems. Interviewees I01, I02, I05, and I06 suggested the marking criteria within tender evaluation to favour CE-compliant contractors. I02 proposed applying this, especially in Design & Build project types, which often offer flexibility in material and process selection. I05 emphasised that, as a financial strategy to implement the CE principle in road construction in BOQs, contractors can use additional preliminary items for the storage of recycling materials and recycling plants.

4.2 Quantitative Analysis

A Questionnaire Survey was distributed among forty-two (42) professionals in the construction industry, and thirty (30) responses were collected, which accounts for a response rate of 71.4%. These respondents possess hands-on experience in road projects, including consultants, project managers, quantity surveyors, civil engineers, site engineers and academics. Figure 1 illustrates a breakdown of professional categories of respondents.

Figure 1: Breakdown of Professional Categories of Respondents



Source: Quantitative data analysis

4.2.1 Circular Economy practices implemented in projects

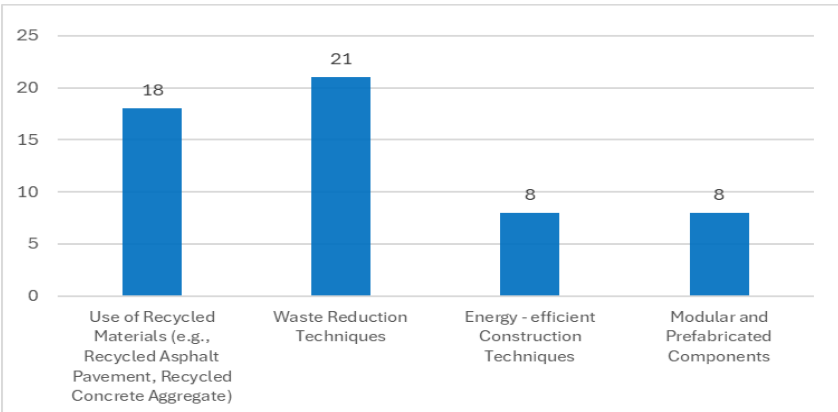


Figure 2: CE practices Implemented in Projects
Source: Quantitative data analysis

The survey results, as illustrated in Figure 2, reveal the most applied CE practice among the participants, being methods in waste reduction, by 21 participants. This is followed closely by 18 participants using recycled materials such as RAP and RCA. Fewer participants (8 participants each) reported using energy-efficient building practices and modular or prefabricated components within their activities, as another CE practice implemented.

4.2.2 Barriers faced when adopting CE practices in road development

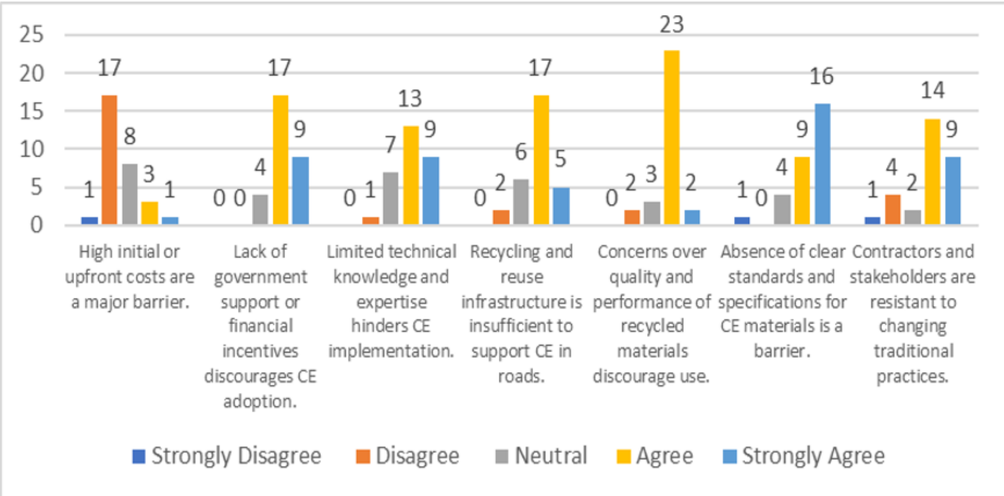


Figure 3: Barriers faced when adopting Circular Economy Practices in road development
Source: Quantitative data analysis

After conducting a thorough expert interview, these seven (7) major factors, as illustrated in Figure 3, were identified which could be challenges and barriers faced when adopting CE practices in road development. All the participants were asked to rank these factors on a Likert scale from 1 to 5 (1 being strongly disagree and 5 strongly agree) based on the magnitude of the challenges. Sixteen (16) respondents strongly agree, while nine (9) participants agree with the absence of clear standards and specifications for CE materials. Also, twenty-three (23) agree, while two (2) strongly agree, that the concerns over quality and performance of recycled materials are considered the most significant hurdles. This suggests a strong need for robust regulatory frameworks and proven performance data to build confidence in CE materials. Notably, a lack of support and financial support (17 agree) and insufficient recycling and reuse systems (17 agree, 5 strongly agree) make it difficult for CE to happen. Furthermore, the collected data were analysed and ranked by the weighted mean method using Eq. 01.

Equation 1: Weighted Factor

$$W = \frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i}$$

Where,

W = weighted average

N = number of terms to be averaged

ω_i = weights applied to x values

X_i = data values to be averaged

Table 3: Ranking of the Challenges

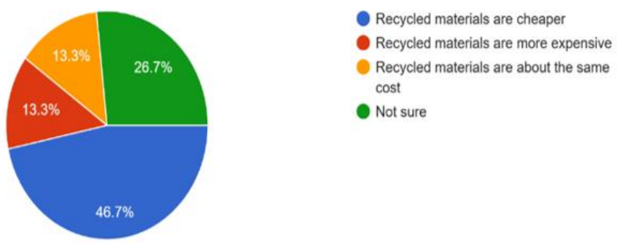
Factor	**Weighted Mean**	**Rank**
Absence of clear standards and specifications for CE materials is a barrier.	4.300	**1**
Lack of government support or financial incentives discourages CE adoption.	4.167	**2**
Limited technical knowledge and expertise hinder CE implementation.	4.000	**3**
Contractors and stakeholders are resistant to changing traditional practices.	3.867	**4**
Recycling and reuse infrastructure is insufficient to support CE in roads.	3.833	**5**
Concerns over the quality and performance of recycled materials discourage use.	3.833	**5**
High initial or upfront costs are a major barrier.	2.533	**7**

Source: Quantitative data analysis

Table 3 shows the challenges that impact while adoption of CE practices in road construction. These challenges are ranked according to the weighted mean. Accordingly, the absence of clear standards and specifications for CE materials is taken as the major barrier.

4.2.3 Cost effectiveness of using CE-based Materials

Figure 4: Cost comparison between the use of Recycled and Virgin Materials

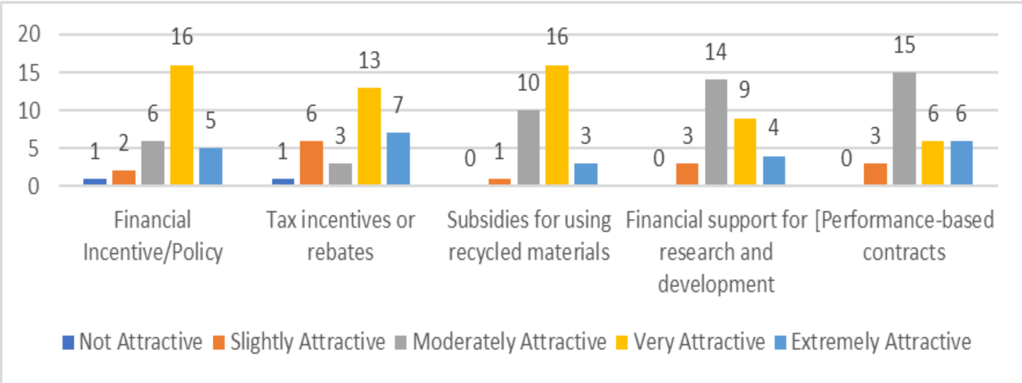


Source: Quantitative data analysis

Figure 4 shows the cost comparison between the use of recycled and virgin materials in road construction. In accordance with the data, 46.7% of respondents find recycled materials to be cheaper due to raw material extraction, transportation, and landfill disposal savings. On the contrary, 13.3% find recycled materials to be higher in cost, possibly due to extra processing, specialised equipment, or conforming to performance standards. Another 13.3% think the costs are about the same, indicating that in some cases, savings from material reuse are balanced by preparation and performance costs. Notably, 26.7% of respondents are unsure, reflecting a lack of consistent cost data or limited awareness in the industry.

4.2.4 Attractiveness of the financial incentives or policies to encourage CE practices in road construction

Figure 5: Attractiveness of Financial Incentives or Policies



Source: Quantitative data analysis

According to the questionnaire survey, financial mechanisms are essential for the adoption of the CE on Sri Lanka's roads. When respondents were asked about financial incentives and policies, the vast majority thought they were extremely attractive. Shortly afterwards, tax incentives and rebates or discounts for using recycled materials were offered, demonstrating that the industry places great significance on directly reducing costs tied to CE. Financial help for research and development, as well as results-based contracts, were appreciated but were placed lower in importance than instant financial incentives. It seems that the industry recognises innovation and quality performance, but currently it is mainly interested in ways to achieve cost savings to help CE take hold. This is clearly illustrated in Figure 5.

Table 4: Ranking of the attractiveness of financial incentives

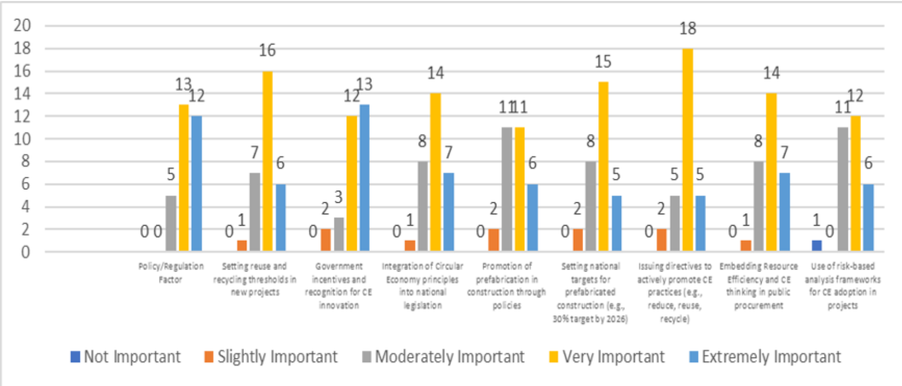
Factor	**Weighted Mean**	**Rank**
Financial Incentive/Policy	3.733	**1**
Subsidies for using recycled materials	3.700	**2**
Tax incentives or rebates	3.633	**3**
Performance-based contracts	3.500	**4**
Financial support for research and development	3.467	**5**

Source: Quantitative data analysis

Table 4 above shows the ranking of attractiveness of financial incentives or policy changes to encourage the CE principles in road construction. These enablers are ranked according to the weighted mean. Accordingly, the most attractive factor is financial incentives, and secondly, subsidies for using recycled materials.

4.2.5 Importance of the policy and regulation strategies

Figure 6: Importance of Policy and Regulation Strategies



Source: Quantitative data analysis

The stakeholder survey revealed a strong consensus on the necessity of policy and regulatory support to promote the adoption of CE practices in Sri Lanka's Road construction sector. As illustrated in Figure 6, respondents overwhelmingly rated policy interventions such as clear standards for recycling and reuse, government-backed financial incentives, and public procurement alignments "Very Important" or "Extremely Important," underscoring their critical role in driving CE integration. Notably, the establishment of recycling and reuse standards was identified as a top priority, reflecting the need for consistent technical specifications to facilitate the use of alternative materials. Similarly, financial mechanisms such as grants and tax rebates were seen as key to overcoming initial cost barriers and accelerating industry uptake. Furthermore, integrating CE principles at the planning stage was strongly endorsed, reinforcing the importance of embedding sustainability early in project lifecycles. While national targets for CE and directives aimed at waste reduction received high support, strategies such as prefabrication promotion and risk-based CE evaluation frameworks were seen as moderately important—valuable but not as central as policy and financial levers. This indicates that while technical innovations are helpful, systemic policy reform and robust institutional frameworks are viewed as the most effective catalysts for advancing CE in the road construction industry.

Table 5: Ranking of the importance of the policy and regulation strategies

Factor	Weighted mean	Rank
Policy/Regulation Factor	4.233	**1**
Government incentives and recognition for CE innovation	4.200	**2**
Setting reuse and recycling thresholds in new projects	3.900	**3**
Integration of CE principles into national legislation	3.900	**3**
Embedding Resource Efficiency and CE thinking in public procurement	3.900	**3**
Issuing directives to actively promote CE practices (e.g., reduce, reuse, recycle)	3.867	**6**
Setting national targets for prefabricated construction (e.g., 30% target by 2026)	3.767	**7**
Use of risk-based analysis frameworks for CE adoption in projects	3.733	**8**
Promotion of prefabrication in construction through policies	3.700	**9**

Source: Quantitative data analysis

The ranking results depicted in Table 5 above highlight that policy and regulatory factors are considered the most crucial drivers for adopting CE practices in road construction, with a weighted mean of 4.233. Government incentives and recognition closely follow (4.200), emphasising the importance of financial and institutional support to encourage innovation.

5 DISCUSSION

The data gathered from the literature review, semi-structured interviews, and questionnaire survey reveal that while the concept of CE is somewhat recognized in Sri Lanka's Road construction sector, its practical implementation remains limited. Both academic literature and expert insights confirm awareness of various recycled and alternative materials such as Recycled Concrete Aggregate (RCA), Reclaimed Asphalt Pavement (RAP), fly ash, cement kiln dust, crushed brick, and glass. Additionally, industry professionals highlighted the use of quarry dust, turfing for embankments, and the reuse of modular prefabricated components like curbs and signboards—demonstrating a foundational but underdeveloped application of CE practices in the sector.

5.1 Challenges and Strategic Recommendations

When integrating findings from both literature and primary data, the study proposes the following strategies to enhance CE adoption, as given in Table 7 below.

Table 7: Strategic Recommendations

Challenges	Proposed Strategies	Expected Outcome
Lack of standard specification	Updating the specification related to the road	Engineers and contractors will have clear guidelines for using recycled materials, increasing confidence and reducing reluctance.
	Incorporate alternative specs and particular specs in BOQ and tenders.	
Lack of Government Support/Incentives	Providing tax concessions for CE approaches	The industry will be financially motivated to adopt CE, overcoming the barrier of high upfront costs and encouraging innovative solutions.
	Adopt the CE approach in the bid evaluation criteria for D&B type.	
Limited Knowledge and Expertise	Training programs for contractors and suppliers on CE methods	Industry professionals will gain the necessary skills and confidence, while pilot projects will serve as practical examples and learning models.
	Make CE-based pilot projects.	
Reluctant to adopt alternative construction methods	Certification process	Stakeholders will have a formal process to validate sustainable practices, and financial incentives will encourage them to move away from traditional methods.
	Training programs for contractors and suppliers on CE methods	
	Giving tax concessions and rewards	
Availability of recycled material	Make market competition for recycled materials.	A competitive market will ensure a steady and reliable supply of recycled materials, addressing the current logistical challenges.
Storage problem	Use additional preliminary items for storage and recycling process	This will provide dedicated budget and resources for proper storage and handling, improving material quality control and logistical efficiency.

In order to effectively promote a CE in Sri Lanka's road construction sector, this research proposes a number of actionable recommendations for policymakers and practitioners. The study emphasises the urgent need for a regulatory framework, beginning with the development of clear, standardised specifications for the use of recycled and alternative materials. The findings show that without these technical standards, engineers and stakeholders are reluctant to adopt CE practices due to concerns about the quality and performance of recycled materials. Financial incentives are also critical, with a strong preference among industry professionals for **tax benefits, subsidies, and the inclusion of CE principles in bid evaluation criteria** to offset initial costs and reward sustainable approaches. Furthermore, the study suggests that the **Road Development Authority (RDA) and Construction Industry Development Authority (CIDA)** should lead the change by establishing circular procurement policies, supporting pilot projects, and promoting training programs to address the existing knowledge and expertise gaps.

The discussion reveals a substantial gap between theoretical awareness and practical application of Circular Economy practices in road construction in Sri Lanka. While literature from the global and local environment confirms the technical and financial viability of CE, ground-level implementation is hampered by regulation, institutions, and logistics. However, industry professionals demonstrate a willingness to adopt CE, provided that supportive strategies and incentives are put in place.

6 CONCLUSION

This study concludes that while the awareness of Circular Economy practices in Sri Lanka's Road construction sector is increasing, practical implementation remains limited due to key challenges such as unclear policies, lack of technical standards, limited stakeholder knowledge, and insufficient infrastructure for recycled materials. Despite these barriers, the research highlights significant potential for CE to enhance environmental sustainability, reduce construction costs, and improve long-term infrastructure performance. The findings suggest that with stronger government support, targeted policy reforms, industry training, and the development of centralised recycling systems, CE practices can be effectively integrated into road development projects. This transition will require collaborative action across public and private sectors to create a more resource-efficient and sustainable construction industry.

7 ACKNOWLEDGEMENTS

The researcher would like to acknowledge Liverpool John Moores University (LJMU) and Sri Lanka Institute of Information Technology (SLIIT) for the facilitation of this study by providing university resources throughout this research.

REFERENCE

Agyemang, M., Kusi-Sarpong, S., Khan, S. A., Mani, V., Rehman, S. T., & Kusi-Sarpong, H. (2019). Drivers and barriers to circular economy implementation: An explorative study in Pakistan's automobile industry. *Management Decision*, *57*(4), 971–994. https://doi.org/10.1108/MD-11-2018-1178

AlJaber, A., Martinez-Vazquez, P., & Baniotopoulos, C. (2023). Barriers and Enablers to the Adoption of Circular Economy Concept in the Building Sector: A Systematic Literature Review. *Buildings*, *13*(11). https://doi.org/10.3390/buildings13112778

Ashtiani, M., Lewis, M., Waldman, B., & Simonen, K. (2024). *Strategies to Reduce Embodied Carbon in Roadway Infrastructure*.

Benachio, G., Freitas, M. do C., & Tavares, S. (2020). Circular economy in the construction industry: A systematic literature review. *Journal of Cleaner Production*, *260*, 121046. https://doi.org/10.1016/j.jclepro.2020.121046

Connolly, L., Sheils, E., Lamb, M., O'Connor, A., O'Brien, E., Viner, H., Bailey, H., Ramdas, V., & Varveri, K. (2023). Circular economy in road construction and maintenance. *Transportation Research Procedia*, *72*, 3458–3465. https://doi.org/10.1016/j.trpro.2023.11.770

Ghaffar, S. H., Burman, M., & Braimah, N. (2020). Pathways to circular construction: An integrated management of construction and demolition waste for resource recovery. *Journal of Cleaner Production*, *244*, 118710. https://doi.org/10.1016/j.jclepro.2019.118710

Hammond, G. P., & Jones, C. I. (2008). Embodied energy and carbon in construction materials. *Proceedings of Institution of Civil Engineers: Energy*, *161*(2), 87–98. https://doi.org/10.1680/ener.2008.161.2.87

Kluttz, R. Q., Dongré, R., Powell, R. B., Willis, J. R., & Timm, D. H. (2017). *Long-term performance of a highly modified asphalt pavement and application to perpetual pavement design*. *June*. https://doi.org/10.14311/ee.2016.212

Kodithuwakku, H. D. D. P., Ramachandra, T., & Kajavathani, P. (2023). Alternative Materials For Sustainable Road Construction in Sri Lanka. *World Construction Symposium*, *1*, 110–122. https://doi.org/10.31705/WCS.2023.10

Kumari, T., & Kulathunga, U. (2020). *To cite this article : Thakshila Kumari, Udayangani Kulathunga, Thathsarani Hewavitharana & Nandun Madusanka (2020): Embodied carbon reduction strategies for high-rise buildings in Sri Lanka, International Journal of Construction Management, DOI : E*. https://doi.org/10.1080/15623599.2020.1814939

Li, C. Z., Tam, V. W., Lai, X., Zhou, Y., & Guo, S. (2024). Carbon footprint accounting of prefabricated buildings: A circular economy perspective. *Building and Environment*, *258*. https://doi.org/10.1016/j.buildenv.2024.111602

MacArthur, F. E. (2015). Towards a circular economy: Business rationale for an accelerated transition. *Ellen MacArthur Foundation (EMF)*, 20.

Mishra, P., Rajak, F., & Kumar, R. (2023). *Outlook of Circular Economy to High-rise Buildings Carbon Dioxide Outlook of Circular Economy to High-rise Buildings Carbon Dioxide*. *June 2024*.

Paranavithana, G. N., Samarakoon, S. M. K., & Chathurika, R. H. J. (2024). Analysis of road construction waste generation in developing economies: A focus on surfacing and sub-base construction projects in Sri Lanka. *Journal of Cleaner Production*, *457*. https://doi.org/10.1016/j.jclepro.2024.142465

Plati, C. (2019). Sustainability factors in pavement materials, design, and preservation strategies: A literature review. *Construction and Building Materials*, *211*, 539–555. https://doi.org/10.1016/j.conbuildmat.2019.03.242

Silva, S. De, & Fernando, S. (2022). *Icsbe 2022*.

Snigdhya, N. S., Uddin, S., & Dipu, A. (2024). *Integrating Climate Change Adaptation into Flood Risk Management : Global Perspectives*. *XII*(August), 1003–1012. https://doi.org/10.51244/IJRSI

Weerakoon, P., & Thayaparan, M. (2023). Analysis of Barriers Towards Implementing Circular Economy In the Sri Lankan Built Environment. *World Construction Symposium*, *1*, 162–173. https://doi.org/10.31705/WCS.2023.14

World Green Building Council. (2023). *The Circular Built Environment Playbook*. 1–97. www.worldgbc.org

Worrell, E., Price, L., Martin, N., Hendriks, C., & Meida, L. O. (2001). Carbon dioxide emissions from the global cement industry. *Annual Review of Energy and the Environment*, *26*(November), 303–329. https://doi.org/10.1146/annurev.energy.26.1.303

Solar Hotspot Detection Using VHDL-Simulated Fixed-Point SVM: A Methodology Toward FPGA Realization

Nayomi Fernando[a], Lasantha Seneviratne[b], Nisal Weerasinghe[c], Namal Rathnayake[d]* and Yukinobu Hoshino[e]

Department of Electrical and Electronic Engineering, Faculty of Engineering, Sri Lanka Institute of Information Technology, Malabe [a,b,c] , Advanced Institute for Marine Ecosystem Change (WPI-AIMEC), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokohama 236-0001[d] , School of Systems Engineering, Kochi University of Technology, 185 Miyanokuchi, Tosayamada, Kami 782-8502, Kochi, Japan[e]

nayomi.f@sliit.lk, lasantha.s@sliit.lk, nisal.w@sliit.lk, namalr@jamstec.go.jp, hoshino.yukinobu@kochi-tech.ac.jp

ABSTRACT

Early detection of thermal hotspots in photovoltaic modules is critical to ensuring their efficiency, safety, and longevity. This study presents a complete end-to-end methodology for implementing a fixed-point Medium Gaussian Support Vector Machine classifier using VHDL for a Field Programmable Logic Array. The approach begins with feature extraction from thermal images of healthy and defective solar panels, which focuses on MPEG-7 descriptors. The study shows that high impact for hotspot detection comes from blue chrominance contrast. A medium Gaussian SVM model is trained in MATLAB and converted to a fixed-point Q1.15 format for hardware compatibility. Key parameters, including support vectors, Lagrange multipliers, bias, and kernel scale, are extracted and verified in a custom Python environment to ensure numerical alignment with MATLAB results. The validated model is then implemented in synthesizable VHDL. It is verified using GHDL and the GNU Tool Kit waveform viewer, confirming bit-accurate hardware behaviour. Results show classification accuracy exceeding 99.3% with negligible performance loss due to quantization. The design achieves deterministic latency through an FSM-based structure and parallel feature processing for a 300-support vector and 222-feature system. This method enables low-power, real-time inference on a UAV-based edge platform, primarily focusing on drones.

CCS CONCEPTS Hardware ~ Electronic design automation ~ Logic synthesis; Computing methodologies ~ Machine learning ~ Learning paradigms ~ Supervised learning; Hardware ~ Integrated circuits ~ Reconfigurable logic and FPGAs; Computer systems organization ~ Real-time systems

KEYWORDS: *Deep Learning, FPGA, Machine Learning, MPEG-7, Solar PV, SVM, VHDL*

1 INTRODUCTION

Global demand for renewable energy has driven rapid growth in the deployment of solar photovoltaic (PV) systems, both in large-scale commercial solar farms and in domestic rooftop installations. However, maintaining their efficiency over time is not always straightforward. One major issue that frequently arises in PV panels is the formation of thermal "Hotspots"[1]. Hotspots are localized regions where temperatures exceed those of the surrounding area for several reasons. Factors contributing to hotspot occurrence include partial shading, dust accumulation, cell mismatch, and microcrack defects. In practice, hotspots not only reduce power output but also accelerate material degradation, eventually causing permanent damage to the PV modules. Anomaly detection using thermal signatures in PV modules is advancing with the development of thermal imaging, computer vision, and Machine Learning (ML).

This makes them particularly useful for image-based inspection tasks. The recent study [1], which extracts handcrafted features and trains a Medium Gaussian Support Vector Machine (SVM),

* Corresponding author.

demonstrates strong performance on drone-captured thermal images. Further, among the evaluated features, the blue chrominance contrast extracted from MPEG-7 feature descriptors has a substantial impact on solar hotspot identification. The model achieves 99.3% accuracy with an inference time of 18 seconds. This demonstrates that this model maintains a balance between complexity and computational cost. However, the challenging aspect of the workflow arises in hardware implementation. This is because these models, typically developed in MATLAB, rely on floating-point operations and high-level abstractions. As a result of these characteristics, which make them less suitable for low-power, real-time edge deployment, this raises the challenge of optimizing the model for drone platforms. Therefore, we identified a clear need for hardware-efficient solutions for model implementation on a UAV platform. Considering the factors such as parallel processing, predictable timing, and relatively low power consumption, we identified that Field Programmable Gate Arrays (FPGAs) are well-suited for this application.

However, this is not a direct process for converting a high-level ML model into an FPGA-compatible design. The first significant challenge encountered is replacing floating-point computations with fixed-point representations to achieve efficient synthesis with minimal resource utilisation. In this work, we propose a complete fixed-point implementation of the SVM model developed in the Q1.15 format, providing a practical trade-off between numerical precision and hardware efficiency. The approach is initiated by extracting key parameters of the MATLAB-trained model, namely support vectors, Lagrange multipliers, class labels, kernel scale, and bias. Next, these parameters are evaluated using a Python-based fixed-point simulation, and numerical consistency is checked before proceeding with the hardware implementation. The model is described in synthesizable VHDL using Q1.15 arithmetic after validation. The VHDL design is structured into three main parts: SVM decision logic, constants package, and test bench. GHDL for logic-level execution performs simulation, while GTKWave examines the waveform behaviour. This allows for a closer inspection of timing and state transitions, confirming that the decision function behaves as expected. The design is synthesised to integrate easily into FPGA toolchains such as Intel Quartus or Xilinx Vivado for future hardware implementation.

This study, which is an extension of the previous work [1], focuses on hardware-oriented implementation. The main contribution is that we integrate previous work on MPEG-7 feature extraction using the Automatic Content Extractor (ACE), followed by Medium Gaussian SVM-based classification, and synthesise the resulting FPGA-ready VHDL. To utilise this methodology, we convert the trained SVM parameters to the Q1.15 fixed-point format to reduce FPGA resource usage without significantly reducing accuracy. The second contribution is the verification of the VHDL implementation to ensure bit-level agreement with MATLAB predictions, resulting in a design ready for FPGA synthesis.

The paper is structured as follows. The literature reviews the recent work on solar hotspot detection along with feature extraction and FPGA-based classification methods. The methodology outlines the proposed approach, including SVM training in MATLAB, fixed-point conversion, and VHDL implementation. The results section presents the simulation outcomes, and the discussion section analyses performance, hardware efficiency, and suitability for real-time UAV deployment. The paper concludes by summarising the findings for an implementation-ready design and outlining possible directions for future work in real-world system integration.

2 LITERATURE REVIEW

The integration of infrared (IR) thermography, ML algorithms, and embedded hardware platforms results in thermal hotspot detection. This section delves into the state-of-the-art literature across three main domains. They are image-based hotspot detection techniques, the application of Support Vector Machines for PV fault classification, and FPGA-based implementation of ML algorithms, respectively. This section highlights the evolution of anomaly-detection methodologies driven by new technological advancements. It identifies the key limitations in previous related studies that motivate the proposed pipeline.

1.1 Image-Based Hotspot Detection in PV Panels

Conventional photovoltaic system monitoring has relied on irregularities in electrical parameters such as voltage and current. However, the drawback is that the studies [2-3] reveal these approaches fail to detect anomalies at early stages. The use of infrared thermography from drone-mounted platforms primarily facilitates the identification of localised thermal variations. This signals the primary indication of potential hotspots [4], [5].

The study in [1] proposes a framework combining MPEG-7 texture descriptors, such as edge histograms, with colour-based features to identify fault regions in PV panels. The experiment, conducted across five solar PV datasets, demonstrated that the Medium Gaussian SVM outperformed other classifiers, such as Random Forests and Binary GLM (Generalised Linear Model) Logistic Regression, in terms of classification accuracy and robustness to feature variation. This paper serves as a basis for identifying discriminative features in thermal images, and its findings motivate the use of Medium Gaussian SVM as a core classifier in hardware deployment scenarios [1].

1.2 Support Vector Machines for Fault Detection

Support Vector Machines have been widely favoured for classification tasks where feature space dimensionality is high, and nonlinearity is prevalent [6]. Their ability to define optimal hyperplanes with maximal margin allows them to generalise well, even on limited datasets. The Radial Basis Function (RBF) or Gaussian kernel used in SVMs adds further flexibility by allowing the mapping of inputs into higher-dimensional spaces, where linearly inseparable data become separable [7].

A wide range of works have utilized SVMs in PV diagnostics. For example, study [6] employed an SVM for string-level fault detection in PV arrays using electrical signatures. Recent approaches utilise thermal image features with machine learning models, such as SVM, for fault detection, as stated in the study [9]. Furthermore, enhanced real-time applicability for edge-based systems has been demonstrated by FPGA-accelerated implementations of lightweight deep learning models [8]. While software implementations are widely studied, the hardware integration of SVM-based models remains relatively unexplored due to challenges in representing floating-point operations and nonlinear kernel functions [10], [11].

1.3 FPGA-Based Machine Learning and Fixed-Point Design

FPGAs offer a reconfigurable, power-efficient platform for deploying ML inference models. Their parallel processing capabilities make them suitable for latency-sensitive applications such as real-time image classification [12]. However, due to hardware constraints, floating-point arithmetic is avoided in favour of fixed-point implementations that enable fast logic synthesis, lower resource usage, and faster execution [13].

Existing literature shows that deploying ML on an FPGA requires either of the following. One is high-level synthesis tools like Xilinx Vitis or Intel HLS, which abstract HDL generation from C/C++ code [14]. The other is manual HDL coding with fixed-point models, which offer better control and efficiency [10].

The study [14] proposed a fixed-point neural network implementation on an FPGA for object detection, showing that quantisation has the least impact on accuracy while drastically improving performance metrics. Similarly, research [13] implemented a hardware-efficient kernel SVM classifier for gesture recognition using VHDL, successfully validating the feasibility of kernel-based models on hardware.

Despite these advances, few studies have focused on deploying SVMs with Gaussian kernels, especially for solar hotspot detection using VHDL implementations. The complexity of exponential function evaluation and vector distance computations poses challenges for hardware implementation.

This identified gap is addressed in this proposed work by translating a MATLAB-trained Medium Gaussian SVM into a fixed-point, synthesizable VHDL model. The design is simulated using GHDL, and its functional correctness is verified through waveform analysis in GTKWave. This simulation step ensures that the algorithm behaves identically to its floating-point MATLAB counterpart before it is synthesised for real-time FPGA deployment [10], [19].

1.4 Significance of the Study

The reviewed literature confirms that most PV monitoring solutions rely on software-based ML models with limited focus on edge deployment. While studies highlight the motivation and complexity of implementing Gaussian kernel SVMs in hardware [15], [16], existing works on DNN and SVM hardware flows [17], [18] lack a complete fixed-point MATLAB-to-VHDL pipeline. Furthermore, there is a clear gap in the implementation of fixed-point, nonlinear SVMs on FPGA platforms [19]. By addressing these gaps, the proposed work contributes a novel end-to-end approach that combines ML with hardware readiness, facilitating an energy-efficient, real-time PV health-monitoring system.

In summary, the literature demonstrates substantial progress in image-driven PV fault detection using SVM classifiers, particularly with Gaussian kernels. While many software-based approaches have shown promising results, limited attention has been paid to fixed-point SVM implementations on an FPGA for real-time applications. The reviewed studies emphasise both the potential and the challenges of hardware implementation for nonlinear models. This justifies the current research focus on translating a MATLAB-trained Medium Gaussian SVM model into a VHDL-based implementation. This bridges the gap between ML accuracy and the feasibility of embedded hardware for PV monitoring systems.

2 METHODOLOGY

This section presents the design and validation process for a synthesis-ready Medium Gaussian SVM classifier. The classifier is developed for real-time solar hotspot detection on FPGA-based hardware implementation. This approach includes simulation and pre-synthesis verification. Aligned with the feature selection rationale in the literature, the workflow starts with thermal-image feature extraction using MPEG-7 descriptors and blue chrominance (YCbCr), followed by model training in MATLAB. This process includes fixed-point conversion, VHDL development, and simulation using GHDL. The main design steps are floating-to-fixed-point adaptation in Q1.15, Finite-State Machine (FSM)- based control logic, and a hardware-friendly Gaussian kernel approximation. This shows the mechanism for translating high-level ML models into synthesizable hardware architectures suitable for UAV-based PV panel monitoring. The pipeline explicitly extracts SVM parameters, namely support vectors, Lagrange multipliers, labels, bias, and kernel scale. Further, this validates a fixed-point MATLAB reference and implements dot products and kernel logic in VHDL, with cycle-accurate verification in GHDL. Furthermore, GTKWave ensures bit-accurate parity before FPGA synthesis.

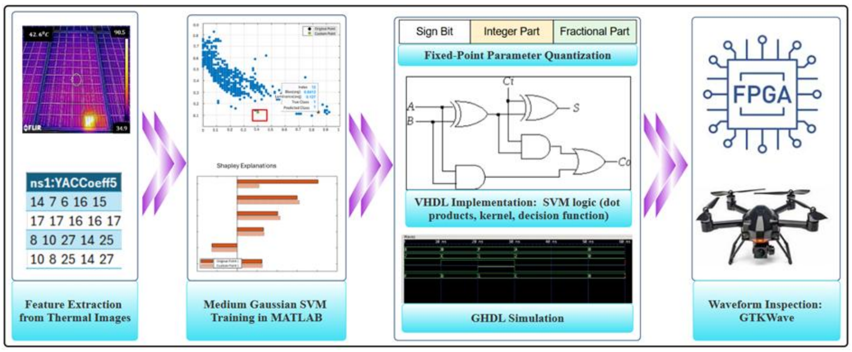


Figure 1. End-to-end workflow for implementing the Medium Gaussian SVM classifier on FPGA, which starts from feature extraction and training in MATLAB through fixed-point parameter quantisation, VHDL coding, simulation with GHDL, and leads to FPGA synthesis and deployment-ready integration on UAV platforms for real-time hotspot detection

The complete workflow for implementing the Medium Gaussian SVM classifier in an FPGA-compatible format for real-time PV hotspot detection is illustrated in Figure 1. The design begins with feature extraction from thermal images using YCbCr descriptors and training a Medium Gaussian SVM model in MATLAB. Then we explicitly extract support vector coefficients, labels, Lagrange multipliers, bias, and kernel scale from the trained model. A custom MATLAB implementation of the SVM decision function is then built to replicate the prediction behaviour, ensuring an exact software reference. These parameters and computations are converted to fixed-point Q1.15 format. The format includes an explicit sign, an integer, and a fractional part for hardware compatibility. The SVM logic is implemented in VHDL. It includes dot products, Gaussian kernel approximation, and decision, followed by cycle-accurate simulation in GHDL. This generates waveforms for GTKWave inspection to confirm bit-accurate agreement with MATLAB. The verified design is delivered as a synthesis-ready module, compatible with standard FPGA synthesis toolchains such as Intel Quartus or Xilinx Vivado, for future deployment on UAV platforms.

The subsequent discussion builds on these stages. It provides a detailed explanation of how MATLAB-SVM models are bridged to synthesizable VHDL, of logic gate-level simulation for solar hotspot classification, and of the implementation of the Gaussian kernel and decision logic intended for FPGA deployment.

Bridging MATLAB SVM Models to Synthesizable VHDL from Floating-Point SVM to Fixed-Point Hardware Friendly Implementation

Built-in SVM functions, such as *fitcsvm* and *predict*, in MATLAB use floating-point arithmetic and object-oriented syntax. It makes them incompatible with the HDL Coder and the FPGA pipeline. To overcome this gap, we implement a custom SVM decision function in MATLAB using extracted parameters. The extracted parameters are alpha, gamma, bias, and support vectors. This allows conversion to fixed-point using *the fi()* function in MATLAB, enabling hardware-friendly computation. The manual approach provides precise control over word length, precision, and overflow behaviour, while supporting debugging, validation, and matching with the software model. It also simplifies the creation of HDL test benches and serves as a bridge toward synthesizable VHDL modules for solar PV hotspot detection.

2.1 Fixed-Point Arithmetic for FPGA Compatibility

FPGAs lack native support for floating-point arithmetic. Moreover, the usage of floating-point IP cores introduces significant resources and latency overhead. Therefore, the enhanced design uses a fixed-point representation, expressing numbers as scaled integers, overcoming this challenge. This procedure reduces computational complexity. It aligns with FPGA architecture, making it the preferred choice for efficient, real-time machine learning implementations on FPGAs.

FPGA hardware compatibility is ensured by adopting a fixed-point format of Q1.15 to maintain numerical precision. There, the total word length is 16 bits comprising 1 sign bit, 1 integer bit, and 15 fractional bits. This configuration allows representation of values in the approximate range of −1 to 0.99997 with a high precision of 0.00003. Among standard 16-bit fixed-point formats such as Q4.11 or Q8.8, Q1.15 is preferred when high precision is critical, and the dynamic range of values is limited, as is frequently the case in embedded AI applications like solar hotspot classification.

2.2 Logic Gate-Level Simulation for Solar Hotspot Classification

To evaluate the feasibility of real-time deployment on an Unmanned Aerial Vehicle (UAV) platform, the trained SVM model was translated into a hardware-compatible version using VHDL. The logic implements binary classification to detect the presence of solar panel hotspots (Class 1) or their absence (Class 0) from MPEG-7 feature vectors. The VHDL module is structured around a three-state FSM: IDLE, COMPUTE, and FINISH, and tested using a testbench.

In the IDLE state, the control signals, index counters (i, j), and accumulator registers are initialised, and the system waits to assert the start signal. Once the start signal is activated, the FSM transitions to the COMPUTE state. It iteratively calculates the squared Euclidean distance between the normalised input feature vector and each stored support vector using fixed-point signed arithmetic. This distance value is used in a hardware-friendly Gaussian RBF kernel approximation implemented as a rational function to avoid floating-point operations. The resulting kernel values are multiplied by the corresponding Lagrange multipliers (α_i) and class labels $(label_i)$, and the products are accumulated into a decision variable (fx_{accum}) according to $fx_{accum} + \alpha_i * label_i * kernel_i$. Upon completing all iterations, the FSM enters the FINISH state. The bias term b is added to the accumulated decision value. Then the sign of the result is evaluated. A binary prediction is generated on the predicted signal. All state transitions are synchronised to the rising edge of the clock signal. This ensures deterministic and cycle-accurate execution.

2.3 Fixed-Point Arithmetic and Kernel Logic

The use of Q1.15 signed fixed-point representation avoids the overhead of floating-point units. This makes the implementation resource-efficient for FPGA targets. Accumulators and kernel computations are implemented using a signed 47 downto 0 format.

SVM Decision Function Gaussian (RBF) Kernel: The classifier computes the decision function as:

$$\boldsymbol{f(x) = \sum_{i=1}^{N} \alpha_i . y_i . K(x_i . x) + b} \qquad (1)$$

where x_i: support vector, y_i: label of support vector, α_i: Lagrange multiplier, b: bias

$$K(x_i . x) = \exp\left(-\gamma . \left||x - x_i|\right|^2\right) \qquad (2)$$

where γ: kernel scaling factor, $K(x_i . x)$: Gaussian kernel value between the input vector x and the i^{th} support vector (x_i), approximated for hardware implementation.

It is required to normalise the input feature vector x such that $x' = \frac{x-\mu}{\sigma}$ by Z-score normalisation, where x': Standardised value, x: Original data value, μ: Mean of the data, σ: Standard deviation of the data. Support vectors, labels, and Lagrange multipliers are essential in computing the kernel sum. γ defines the kernel function shape. μ and σ must be used to preprocess input features before classification.

Distance squared:

$$dist_{sq} = \sum_{j=0}^{N} (x_j - sv_{i,j})^2 \qquad (3)$$

is calculated in pipelined j iterations. $dist_{sq}$ is the Squared Euclidean distance between the input feature vector x and $the\ i^{th}$ support vector. x_j is the j^{th} element of the normalised input feature vector. $sv_{i,j}$ is $the\ j^{th}$ feature of the ith support vector in the trained SVM model. N is the total number of features in the input feature dimension vector. i is the index of the current support vector being processed, and j is the index of the current feature element in the input vector.

Gaussian kernel approximation:

$$K(x, x_i) \approx \frac{10}{1+\gamma.dist_{sq}} \quad (4)$$

where $K(x_i.x)$: Gaussian kernel value is implemented in the exponential approximation function using division-based rational approximation, suitable for hardware. $dist_{sq}$ is the Squared Euclidean distance, and γ is the kernel scaling factor. Decision value is accumulated as in equation (1) across all support vectors and compared to a fixed threshold for a binary decision.

The final stage of the workflow is hardware simulation and validation. This is to ensure deployment readiness. We have implemented the fixed-point VHDL of the SVM classifier, and simulated it using GHDL, with the signal waveforms analysed in GTKWave. This verifies correct state transitions, timing behaviour, and numerical equivalence with the MATLAB floating-point model. The validation confirms that the FPGA design reliably performs real-time solar hotspot classification while allowing integration into UAV-based PV inspection systems.

Overall, the methodology bridges the gap between MATLAB-trained SVM classifiers and their FPGA implementation through a structured process that includes parameter extraction, fixed-point conversion to Q1.15 format, and custom VHDL coding. The classifier is implemented as a state machine-based logic module. It performs kernel evaluations and accumulates decision values using resource-efficient arithmetic. Simulation results validate the equivalence of fixed-point hardware behaviour with the original floating-point mode. It ensures deployment readiness for embedded solar hotspot detection applications.

3 RESULTS

This section assesses the experimental evaluation of the Medium Gaussian SVM model trained on MPEG-7 feature descriptors for detecting thermal anomalies in PV modules. We have tested the model on five datasets of varying sizes and image quality to understand better how its performance scales with dataset size [1]. Performance is evaluated based on accuracy, F1 score, precision, recall, and execution time. A logic-level hardware simulation of the fixed-point SVM classifier is performed to assess its feasibility for FPGA implementation in real-time UAV-based PV inspection systems. We present the findings across five focused subsections. The first is the multi-metric evaluation across datasets used to assess how consistently the model performs at different dataset sizes. Second, the execution time of the models during the training and testing phases has been analysed to determine whether the approach supports real-time deployment. Next, the precision-recall tradeoff and overall model consistency are examined to understand better, the balance between false positives and false negatives in hotspot detection. After that, a logic-gate-level analysis of the SVM classifier was performed to ensure that the software model is correctly translated into hardware. Finally, the simulation results are analysed using timing waveforms and state transitions. This ensures the classifier operates as expected in an FPGA-based environment.

3.1 Multi-Metric Evaluation Across Datasets

Table 1 summarises the training and testing performance of the Medium Gaussian SVM classifier across five datasets [1]. The classifier maintained high accuracy and generalisation throughout the five datasets.

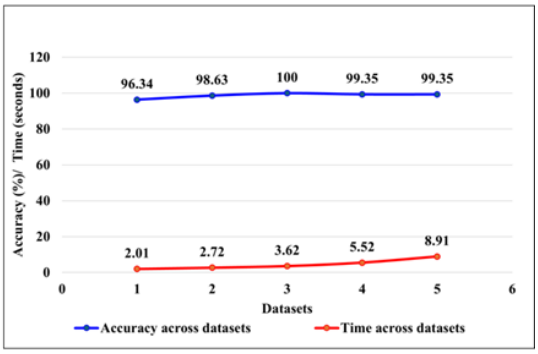
Table 1: Training and Testing Performance across Multiple Datasets

Phase	Dataset Size	Time (seconds)	Accuracy (%)	F1-Score	Precision	Recall
			Training			
Dataset 1	724	16.29	95.00	0.9490	0.9570	0.9419
Dataset 2	1302	4.76	99.20	0.9920	0.9908	0.9938
Dataset 3	1836	6.32	99.60	0.9950	1.0000	0.9920
Dataset 4	2746	9.31	99.90	0.9984	1.0000	0.9970
Dataset 5	4114	18.81	99.30	0.9920	0.9920	0.9930
			Testing			
Dataset 1	82	2.01	96.34	0.9640	0.9520	0.9760
Dataset 2	146	2.72	98.63	0.9860	0.9860	0.9860
Dataset 3	204	3.62	100.00	1.0000	1.0000	1.0000
Dataset 4	306	5.52	99.35	0.9930	1.0000	0.9870
Dataset 5	458	8.91	99.35	0.9930	0.9910	0.9960

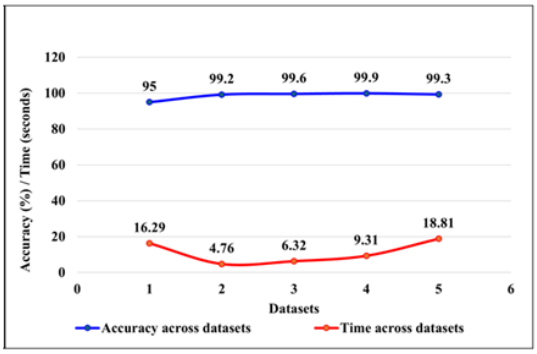
It is observed that the Medium Gaussian SVM achieved up to 100% testing accuracy and F1-scores greater than 0.99 for larger datasets. It demonstrates the model's robustness and generalisation ability.

3.2 Execution Time Analysis

The total inference runtime assesses deployment feasibility. Although the dataset sizes have been increased, the inference times have remained within practical limits. This makes the model appropriate for onboard FPGA-based inference in real-time UAV inspections.



(a)



(b)

Figure 2. Dataset-wise tradeoff between accuracy and time (a) training, (b)testing

As illustrated in Figure 2(a), training time increases gradually as the dataset size grows. However, it remains under 20 seconds for the most extensive dataset as well. This indicates efficient scalability during training. Similarly, as illustrated in Figure 2(b), testing time increased linearly with dataset size, demonstrating a predictable computational load. It should be noted that no significant tradeoff was observed between computational speed and classification accuracy. The model maintained high performance across all five datasets, with no drop in efficiency. The scalability and low computational overhead make the fixed-point SVM model feasible for real-time solar hotspot detection.

3.3 Precision-Recall Tradeoff and Model Alignment

As observed, despite variations in the dataset, the model maintained a high balance between precision and recall. The precision is consistently greater than 95.00% with minimal false positives. Recall up to 99.70%, indicating rare cases of missed detection. F1-score peaked at 1.0000 on dataset 3 with a perfect harmonic mean between precision and recall.

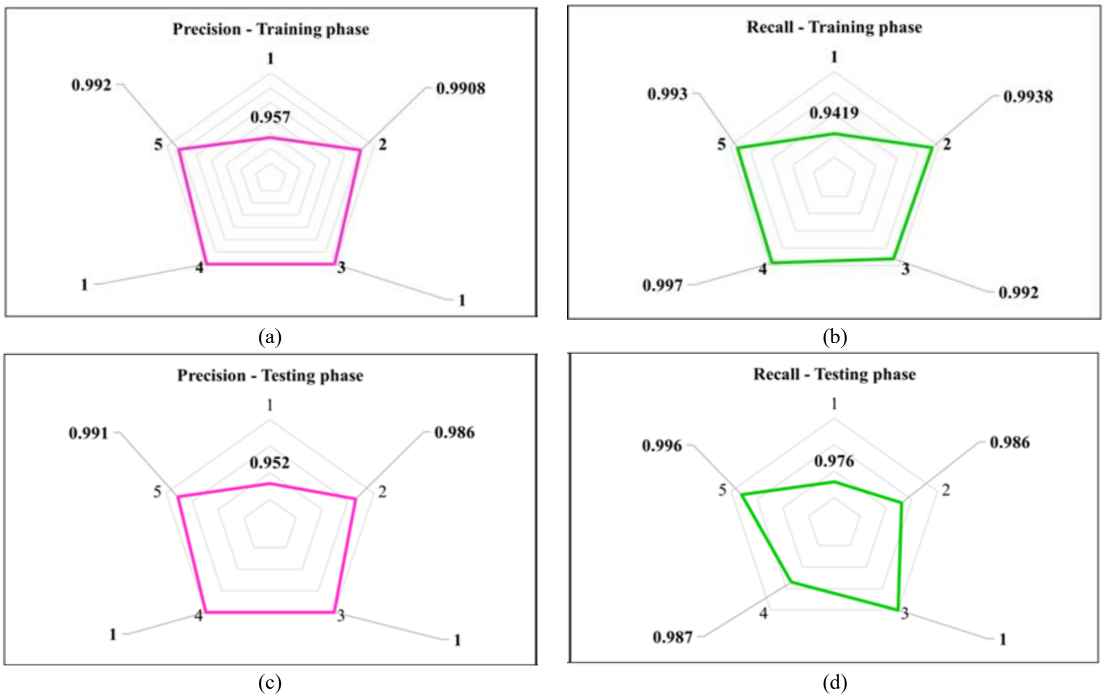


Figure 3. Radar plot of precision and recall across five datasets (a) and (b) training, (c) and (d) testing

Figure 3(a) and (b) show the radar plots of precision and recall for the training phase across five datasets. As the dataset size increases, the training precision generally improves. This can be seen in dataset 1, where 0.9570, reaching 1.0000 in datasets 3 and 4. This indicates that false positives have not been observed. We can see a similar trend in recall, which increases from 0.9419 in dataset 1 to 0.9970 in dataset 4. This improvement is reflected in the F1-score and overall accuracy for dataset 4 as well. As expected, training time increases with dataset size, with the most extensive dataset taking 18.81 seconds.

Figure 3(c) and (d) present the precision and recall during the testing phase. The model maintains strong performance even with relatively small datasets, suggesting good generalisation capability. Precision values range from 0.9520 to 1.0000, while recall ranges from 0.9760 to 1.0000. The best performance is observed for dataset 3, where the model achieves perfect accuracy and an ideal F1-score, with no misclassifications. Datasets 4 and 5 also show consistent results, with high precision (0.9910) and recall (0.9960). These observations indicate the performance reliability of the fixed-point SVM model.

In summary, Q1.15 fixed-point Medium Gaussian SVM gives high detection accuracy. Furthermore, it maintains low execution time across datasets of multiple sizes. MPEG-7 feature descriptor embedding further improves the model's ability to distinguish hotspot regions. These results support the feasibility of deploying the model on FPGA platforms for real-time onboard PV inspection.

3.4 Logic Gate-Level Analysis and Simulation of SVM Classifier

The trained Medium Gaussian SVM model is converted into a deployable embedded form. It is designed for use in UAV edge devices. A hardware-level implementation is developed using VHDL.

The module *svm_core.vhd* simulates the classification process. This is based on finite-state control logic. Furthermore, the design relies on signed fixed-point arithmetic.

The binary classifier of solar hotspot detection combines arithmetic operations with state-machine-based control. At its core, the architecture is constructed from basic logic gate components. They are addition, subtraction, squaring, and multiplication, all facilitated using an FSM with states IDLE, COMPUTE, and FINISH. The inner product calculation, which serves as the kernel distance, is performed using repeated subtraction and squaring. The accumulation of weighted kernel contributions follows it. The decision logic uses a final comparison gate to determine whether the output exceeds the bias threshold that assigns the prediction class. The design computes the RBF kernel, ensuring efficient SVM evaluation without the memory overhead of precomputed exponential values.

3.5 Insights from Simulation

The testbench mimics real-world activation by applying start and reset signals. Simulation output provides visibility into internal signals via report statements.

Figure 4. Prediction 1 and 0

Figure 4 shows that the model correctly identifies a hot spot = '1' and '0' when fx_{total} exceeds the threshold. This aligns with expectations from software SVM classification using the same support vectors. Figure 5 and Figure 6 represent the respective waveforms from the GTKWave interface.

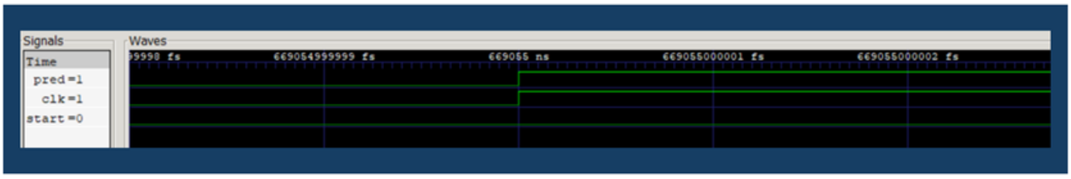


Figure 5. GTK View Output Prediction 1

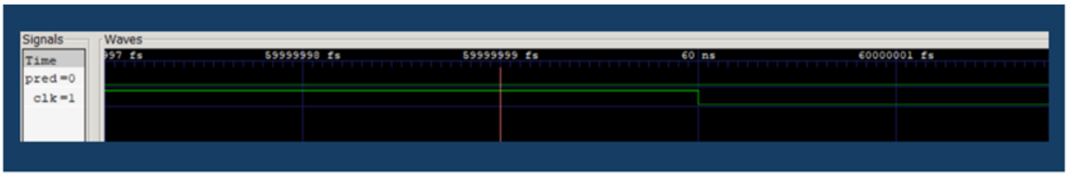


Figure 6. GTK View Output Prediction 0

A medium Gaussian SVM classifier achieved high detection accuracy reaching 100% testing accuracy. Q1.15's fixed-point implementation demonstrated low execution time and robust generalisation, while validating its readiness for real-time FPGA deployment. Hardware-level simulation verified the classifier's accuracy using VHDL. The logic gates and FSM-based control executed Gaussian kernel operations without requiring a look-up table (LUT). Not using a LUT computes values directly, which typically results in higher numerical accuracy, especially for fine-grained inputs. These results ensure the feasibility of the proposed model for embedded solar hotspot detection.

4 DISCUSSION

This section discusses the performance implications and deployment feasibility of the fixed-point Medium Gaussian SVM model for this specific solar PV application. In this section, we highlight the necessity of a custom MATLAB implementation for hardware translation. It evaluates the impact of dataset size on model performance. Further, it examines generalisation and discusses hardware-software coherence for UAV deployment. Additionally, the statistical consistency of the performance metrics across datasets is examined to ensure robustness and reliability, and to validate the quantitative significance of observed performance trends. The discussion is organised under five subsections. They are listed as: *Rationale for Developing a Custom SVM Function in MATLAB Prior to HDL Conversion; Impact of Dataset Size on Training Performance, Generalisation, and Testing Accuracy; Hardware-Software Coherence and Deployment Viability on UAVs;* and *Statistical Analysis of Classifier Performance.* Each subsection interprets the results presented with clarifications on why the observed metrics are meaningful.

4.1 Rationale for Developing a Custom SVM Function in MATLAB Prior to HDL Conversion

The custom SVM function implemented in MATLAB replicates the core decision function behaviour through fixed-point arithmetic. Built-in SVM models and prediction methods in MATLAB are not directly compatible with HDL Coder because they are object-oriented. They depend on floating-point operations. The model has been tailored for fixed-point arithmetic by manually reconstructing the SVM using extracted parameters. It allows accurate translation to VHDL. This approach ensures full control over precision and word length. Further, it facilitates simulation and verification. It bridges the gap between high-level training and hardware implementation, a critical step for FPGA-based machine learning applications. The SVM logic is converted into fixed-point arithmetic to achieve hardware compatibility. Based on this, synthesizable VHDL code is written, including both the main predictor module and a corresponding test bench for simulation. GHDL is used to verify the functional behaviour of the VHDL implementation. Once validated, the design can be synthesised and implemented using FPGA toolchains such as Intel Quartus. The final bitstream is deployed to the FPGA hardware for real-time solar hotspot classification.

4.2 Impact of Dataset Size on Training Performance

The radar plots and tabulated metrics indicate that increasing the dataset size enhances both precision and recall during training. This is attributed to the improved representation of data distribution, allowing the Medium Gaussian SVM to learn more generalizable decision boundaries. The jump from 95.00% accuracy in dataset 1 to nearly 100.00% in dataset 4 highlights this trend. Furthermore, the Q1.15 fixed-point arithmetic does not introduce noticeable quantisation errors as evidenced by the consistent alignment between fixed-point and floating-point results. This confirms that the model retains its discriminative power even under hardware-friendly numeric constraints.

4.3 Generalisation and Testing Accuracy

The model sustains high performance across all five datasets, validating its robustness and transferability from training to deployment. The consistency in recall across test sets, particularly the perfect scores in dataset 3 and high scores in the other datasets, demonstrates the model's ability to detect hotspots without omission. Minor dips in precision for datasets 1 and 4, to 0.9520 and 1.0000, respectively, suggest rare false positives due to edge cases in smaller test samples. Nonetheless, the GHDL-verified VHDL simulation maintains high fidelity with MATLAB predictions, reinforcing that the implemented architecture is both accurate and efficient. These outcomes strongly advocate for its FPGA realisation in real-time solar inspection systems.

4.4 Hardware-Software Coherence and Deployment Viability on UAVs

The VHDL-implemented SVM classifier replicates the behaviour of its MATLAB-trained floating-point counterpart, confirming the correctness of the fixed-point hardware design. Its deterministic latency is a direct consequence of the FSM structure and parallel feature processing. The number of clock cycles depends on the number of support vectors and features, yielding 2702 cycles for 300 SVs and 222 features in the current design. The Gaussian RBF kernel is approximated using a rational function to avoid costly exponentials and look-up tables. This makes the design lightweight and feasible for resource-constrained FPGAs. The model supports further optimisations such as pipelining, early termination, and precision scaling, enhancing efficiency and reducing power consumption. Overall, the design is compact, predictable, and ideal for real-time inference on mid-tier FPGAs in UAV-based PV hotspot detection.

4.5 Statistical Analysis of Classifier Performance

The statistical analysis has been conducted by standard evaluation metrics, namely Mean, Standard Deviation, Coefficient of Variation (CV) and Range, as illustrated in Figure 7. It quantitatively validates the consistency and significance of the performance of Medium Gaussian SVM across different dataset sizes. A strong generalisation is evident across the training and testing phases in the analysis.

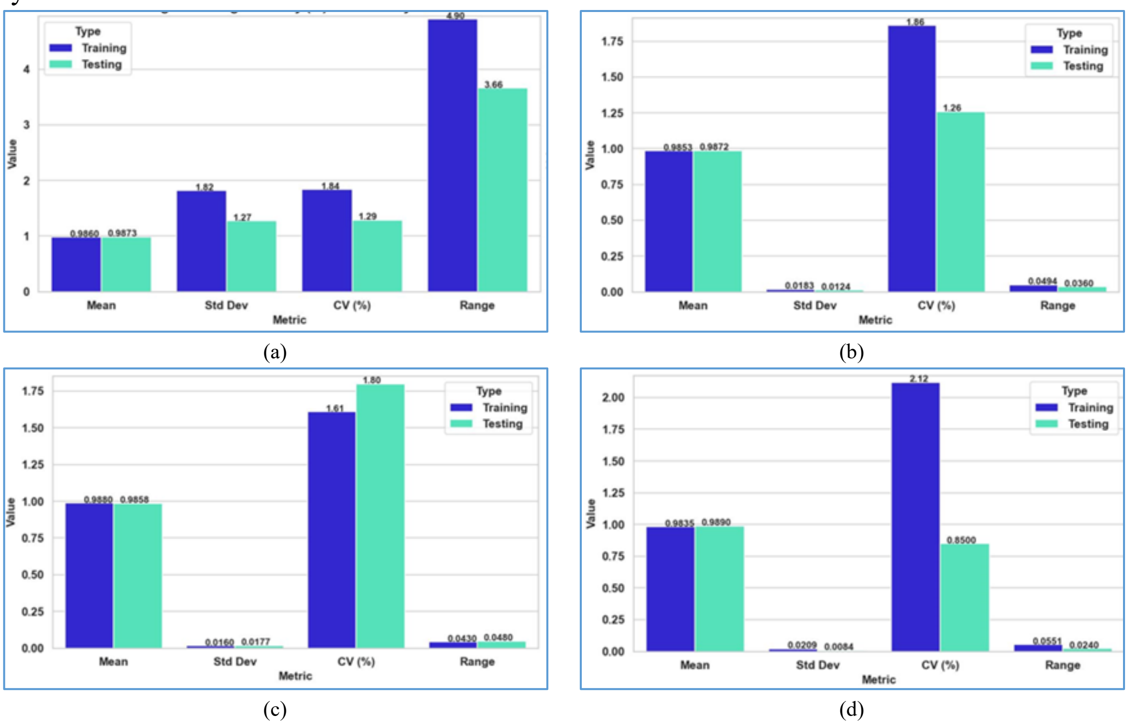


Figure 7. Training and testing consistency metrics for (a) accuracy, (b) f1-score, (c) precision, (d) recall

Figures 7(a) to 7(d) respectively illustrate the accuracy, F1 score, precision, and recall for the training and testing phases. The model achieves a higher test-set mean accuracy (98.73%) than the training-set accuracy (98.60%), as shown in Figure 7(a). Figure 7(b) shows that the range of F1 scores is narrower for testing than for training, suggesting better balance and stability in precision and recall on unseen data.

Figure 7(c) presents precision metrics, with the mean values high for both training and test sets, though the test set shows slightly more variability. Figure 7(d) shows a significantly lower CV in the testing phase than in the training phase. These results suggest the model not only captures more true positives in the test set but does so more consistently. This reinforces its reliability for real-world solar hotspot detection.

Collectively, the fixed-point SVM synthesis confirms the feasibility of hardware deployment, enabling efficient VHDL synthesis. The model demonstrates consistent, high performance across datasets with minimal variance, confirming its generalisation capability. Finally, these findings support the model's potential for real-time, resource-constrained UAV operations in solar inspection tasks.

5 CONCLUSION

This research presents a design pipeline for a fixed-point Medium Gaussian SVM classifier directed for solar hotspot detection in PV modules using MPEG-7 descriptors. We have created a comprehensive pipeline that spans from MATLAB-based preprocessing and Medium Gaussian SVM training to logic-level VHDL implementation for FPGA deployment. Accordingly, the Medium Gaussian SVM classifier consistently achieved high accuracy and F1 scores across five thermal image datasets of varying quality and image counts. To guarantee compatibility with hardware platforms with limited resources, mainly UAV-based FPGA, the conversion from floating-point to Q1.15 fixed-point arithmetic was crucial. As demonstrated by simulation, kernel output validation, and prediction agreement, the logic-based SVM implementation successfully reproduced the MATLAB model's behaviour.

Overall, the proposed approach demonstrates that a handcrafted feature-based, fixed-point SVM classifier preserves software-level accuracy while also satisfying the demanding performance requirements of edge devices. This work provides a reliable and hardware-efficient solution for intelligent fault detection in solar energy systems. This contributes to enhanced PV module maintenance, reduced energy loss, and improved system lifespan.

5.1 Directions for Future Work

There are several opportunities for future enhancement available for this research context. One area is dynamic reconfiguration and online retraining mechanisms within a hybrid FPGA-CPU architecture. Another perspective lies in integration with lightweight deep learning models to further improve the detection of complex anomalies. In addition, deploying the system on actual UAVs with real-time thermal imaging sensors will allow field validation and calibration against diverse irradiance and angle-of-view conditions. For further research, power-optimisation techniques and resource-aware design can be investigated to extend the endurance of UAVs during continuous inspection missions. Furthermore, as noted in [20], emerging paradigms such as TinyML and Transformer variants, along with architectures such as Mamba, provide directions for balancing accuracy and efficiency in real-time applications.

ETHICS DECLARATIONS

Declaration of generative AI and AI-assisted technologies in the writing process

The author used ChatGPT to check for grammatical and spelling errors with the prompt "Proofread this part, focusing on grammatical and spelling errors." After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article. Grammarly was also used to ensure clarity and correctness.

ABBREVIATIONS

ACE: Automatic Content Extractor
AI: Artificial Intelligence
BGLR: Binary GLM (Generalised Linear Model) Logistic Regression
CV: Coefficient of Variation
DL: Deep Learning
FPGA: Field Programmable Logic Array
FSM: Finite State Machine
GTKWave: GNU Tool Kit Wave

HDL: Hardware Description Language
IP: Intellectual Property
IR: Infrared
LUT: Look-Up Tables
ML: Machine Learning
MPEG: Moving Picture Experts Group
PV: Photovoltaic
RBF: Radial Basis Function
SVM: Support Vector Machine
UAV: Unmanned Aerial Vehicle
VHDL: VHSIC (Very High-Speed Integrated Circuit) Hardware Description Language

REFERENCES

[1] Nayomi Fernando, Lasantha Seneviratne, Nisal Weerasinghe, Namal Rathnayake, and Yukinobu Hoshino. 2025. Efficient hotspot detection in solar panels via computer vision and machine learning. Information 16, 7, 608. https://doi.org/10.3390/info16070608

[2] Muhammad Umair Ali, Hafiz Farhaj Khan, Manzar Masud, Karam Dad Kallu, and Amad Zafar. 2020.A machine learning framework to identify the hotspot in a photovoltaic module using infrared thermography. Solar Energy 208, 643–651. https://doi.org/10.1016/j.solener.2020.08.027

[3] H. Hussein. 2023. Photovoltaic hotspot detection with image classification deep learning techniques. In Proceedings of the 2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI). IEEE, Zarqa, Jordan, 1–6. https://doi.org/10.1109/EICEEAI60672.2023.10590305

[4] Mahmoud Dhimish, Marios Theristis, and Vincenzo d'Alessandro. 2024. Photovoltaic hotspots: A mitigation technique and its thermal cycle. Optik 300 , 171627. https://doi.org/10.1016/j.ijleo.2024.171627

[5] H. Açıkgöz, D. Korkmaz, and Ç. Dandıl. 2022. Classification of hotspots in photovoltaic modules with deep learning methods. TJST 17, 2 (2022), 211–221. https://doi.org/10.55525/tjst.1158854

[6] J. Wang, D. Gao, S. Zhu, S. Wang, and H. Liu. 2019. Fault diagnosis method of photovoltaic array based on support vector machine. Energy Sources, Part A: Recovery, Utilisation, and Environmental Effects 45, 2 , 5380–5395. https://doi.org/10.1080/15567036.2019.1671557

[7] C. J. Burges. 1998. Tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 2 , 121–167. https://doi.org/10.1023/A:1009715923555

[8] Fengxi Zhang, Yuying Li, and Zhihao Ye. 2022. Apply YOLOv4-Tiny on an FPGA-based accelerator of a convolutional neural network for object detection. In Proceedings of the 2nd International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA 2022). IOP Publishing, Dali City, China, Article 012032. https://doi.org/10.1088/1742-6596/2303/1/012032

[9] Karuppiah Natarajan, Praveen Kumar B., and Vankadara Kumar. 2020. Fault detection of solar PV system using SVM and thermal image processing. International Journal of Renewable Energy Research 10, 967–977.

[10] X. Song, H. Wang, and L. Wang. 2014. FPGA implementation of a support vector machine-based classification system and its potential application in the smart grid. In Proceedings of the 2014 11th International Conference on Information Technology: New Generations (ITNG). IEEE, Las Vegas, NV, 397–402. https://doi.org/10.1109/ITNG.2014.45

[11] Mateusz Wasala and Tomasz Kryjak. 2022. Real-time HOG+SVM-based object detection using SoC FPGA for a UHD video stream. arXiv:2204.10619 [cs.CV]. https://doi.org/10.48550/arXiv.2204.10619

[12] A. Escobedo and A. Lin. 2016. Tessellation-based multi-block memory-mapping scheme for high-level synthesis with FPGA. In Proceedings of the 2016 International Conference on Field-Programmable Technology (FPT). IEEE, Xi'an, China, 125–132. https://doi.org/10.1109/FPT.2016.7929517

[13] X. A., W. A. Najjar, and A. K. Roy-Chowdhury. 2015. Evaluation and acceleration of high-throughput fixed-point object detection on FPGAs. IEEE Transactions on Circuits and Systems for Video Technology 25, 6, 1051–1062. https://doi.org/10.1109/TCSVT.2014.2360030

[14] H. Brum, M. Véstias, and H. Neto. 2024. LiDAR 3D object detection in FPGA with low bitwidth quantization. In Proceedings of Applied Reconfigurable Computing. RC 2024. Lecture Notes in Computer Science, vol. 14553. Springer, Cham, 90–105. https://doi.org/10.1007/978-3-031-55673-9_7

[15] Muhammad Ihsan Al Hafiz, Naresh Ravichandran, Anders Lansner, Pawel Herman, and Artur Podobas. 2025. A reconfigurable stream-based FPGA accelerator for Bayesian confidence propagation neural networks. arXiv:2503.01561 [cs.AR]. https://doi.org/10.48550/arXiv.2503.01561

[16] Srikanth Ramadurgam and Darshika G. Perera. 2021. An efficient FPGA-based hardware accelerator for a convex optimization-based SVM classifier for machine learning on embedded platforms. Electronics 10, 11, 1323. https://doi.org/10.3390/electronics10111323
[17] Shereen Afifi, Hamid Gholamhosseini, and Roopak Sinha. 2019. A system on chip for melanoma detection using an FPGA-based SVM classifier. Microprocessors and Microsystems 65, 57–68. https://doi.org/10.1016/j.micpro.2018.12.005
[18] Shereen Afifi, Hamid Gholamhosseini, and Roopak Sinha. 2015. Hardware implementations of SVM on FPGA: A state-of-the-art review of current practice. International Journal of Innovative Science, Engineering & Technology (IJISET) 2, 733–752.
[19] Mohammed H. Yacoub, Samar M. Ismail, Lobna A. Said, Ahmed H. Madian, and Ahmed G. Radwan. 2024. Support vector machine reconfigurable hardware implementation on FPGA. Franklin Open 7, 100115. https://doi.org/10.1016/j.fraope.2024.100115
[20] T. Suwannaphong, F. Jovan, I. Craddock, and R. McConville. 2025. Optimising TinyML with quantization and distillation of transformer and mamba models for indoor localization on edge devices. Scientific Reports 15, 10081. https://doi.org/10.1038/s41598-025-94205-9

Analysis of Thermal Performance of Shell and Tube Heat Exchangers: A Correlation and CFD Based Approach

S.A.P Ushettige[1], W.K. Wimalsiri[2], H.G.S. Hikkaduwa[3]
[1]Faculty of Engineering, School of Civil and Mechanical Engineering, Curtin University, Bentley, WA 6102, Australia
aperera844@gmail.com ; walallawita.w@sliit.lk
[2]Faculty of Engineering, Mechanical Engineering, Sri Lanka Institute of Information Technology Malabe, 10115, Sri Lanka
[3]Faculty of Engineering, General Sir John Kothalawala Defence University Kandawala Road, Dehiwala-Mount Lavinia, 10390, Sri Lanka
hikkaduwahgs@kdu.ac.ok

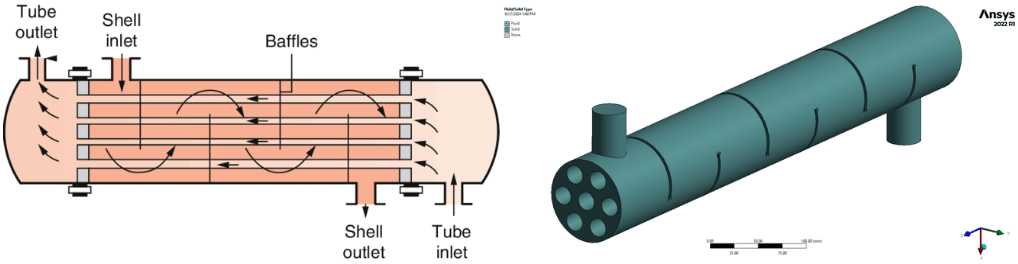
ABSTRACT

Shell and tube heat exchangers are devices which are widely adopted in thermal systems for the transfer of thermal energy due to both performance and reliability factors. Given their application in energy-intensive systems, the design and sizing of these devices have become a rapidly growing field. Traditionally, empirical correlations which were based on experimental results were used for thermal sizing and design. This was replaced by computational fluid dynamics (CFD) modelling given its ability to model and visualize flow, expanding the horizon of possibilities for design and performance optimization. Recently, CFD has been combined with numerical methods such as non-linear least-squares regression to develop correlations that predict thermal performance based on input design parameters. However, the application of this integrated method for shell and tube heat exchangers is limited. This study will model a single-pass TEMA E-type shell and tube heat exchanger using ANSYS Fluent ®. CFD simulations are used to explore the effect of turbulence on thermal performance by varying both the inlet mass flow rate and the central baffle spacing. Steady state simulations are conducted for four models with six, eight, ten, and twelve baffles. The results of CFD modelling are then combined with non-linear least squares regression in the MATLAB Curve Fitter Toolbox ® to develop four sets of correlations in the form of $Nu = C.Re^{a}.Pr^{b}$. Reasonably confident results were obtained in the final fitted data; however, relatively high 95% confidence interval widths were evident for certain fitted coefficients leaving space for improvement in the model. The study highlights that combining CFD with tools such as nonlinear least squares regression aids both engineers and designers in the thermal design process of shell and tube heat exchangers eliminating the need to limit design based on empirical correlations.

KEYWORDS: *Shell and tube heat exchangers, thermal optimization, non-linear least squares regression, central baffle spacing, turbulence, mass flow rate*

1 INTRODUCTION

Heat exchangers are devices commonly used to transfer thermal energy between process streams in both commercial and domestic systems (Mangrulkar et al., 2019). A range of different heat exchanger types are used in the industry. Among them, shell and tube heat exchangers are widely adopted due to their versatility in operating over, in a range of working pressures and temperatures (Sekulić & Shah, 2023). These devices also provide a large area of heat transfer for a given volume with a relatively simple construction (Pal et al., 2016; Aslam Bhutta et al., 2012). In shell and tube heat exchangers, two fluids at relatively different temperatures flow simultaneously. One flows through the tubes (tube side), and the other flows outside the tubes, or the shell, which is a circular, enclosed volume around the tubes. The energy transfer occurs through a combination of conduction and convection from the inside of the tube fluid to the outside shell fluid or vice versa (Bichkar et al., 2018). This is represented in Figure 1 (a) and (b) below.



(a) *(b)*

Figure 12: (a) Schematic of a Single Pass TEMA E Type shell and tube heat exchanger (M. Necati Özışık, 1985) (b) Physical model in ANSYS Design Modeler.

The widespread adoption of these devices in industries with significant energy transfers has made the design and sizing critical, with thermal and pressure drop as the key parameters analyzed (Huang et al., 2017). Traditionally, empirical correlations developed by organizations such as the TEMA (Tubular Exchanger Manufacturers Organization) were used. This process was time consuming and restricted design, due to the inability to visualize and identify design weaknesses in the flow (Ozden & Tari, 2010). The adoption of CFD for modelling addressed this challenge given its ability to model and visualize complex flow phenomena.

The thermo-mechanical performance in these devices depends on the level of turbulence present. This is commonly introduced using metal plates known as baffles. A range of different types and configurations can be used to achieve this (Bichkar et al., 2018). Traditionally, most models use single segmental baffles. However, the use of these types has two key drawbacks (Ozden & Tari, 2010; Pal et al., 2016; Nedunchezhiyan Mukilarasan et al., 2022; El Maakoul et al., 2016; Abbasi et al., 2020; Zhang et al., 2017). In the region present behind the baffle, "Dead zones" or areas of recirculation are created which reduce the amount of heat transfer. Additionally, the contraction and expansion effects in the fluid flow increase the pressure drop, thereby increasing the pumping power required. A common approach to optimizing performance is to adjust the configuration of baffles, the central baffle spacing and the baffle cut (Ozden and Tari, 2010; Abbasi et al., 2020). The central baffle spacing and baffle cut are identified as parameters which significantly affect the flow phenomenon and thermo-mechanical performance. Numerical work indicates that decreasing baffle spacing is associated with increased heat transfer and nozzle-to-nozzle pressure drop.

The potential of CFD design and optimization has recently been combined with tools such as non-linear least squares regression. The integration of the two tools has enabled the development of correlations that relate thermal performance to the model's input design parameters. However, following a comprehensive survey of the existing literature, minimal applications were available for shell and tube heat exchangers. Established applications include those for a horizontal counter-flow double tube heat exchanger and plate-fin and tube heat exchanger, as reported by Alhulaifi (2024) and Marcinkowski et al. (2021), respectively. The study done by Alhulaifi (2024) explored the thermal performance of a nano-fluid flow in the device, and a correlation that predicts *Nu* based on *Pr* and *Re* was developed. While the study by Marcinkowski et al. (2021) involved developing a set of correlations to predict the *Nu* for each row based on the respective *Pr* and *Re*. A further development of the work was that done by Marcinkowski et al. (2024) which highlighted that varying the modelling scheme had minimal effect on the final outcome.

This study will use the commercial CFD solver ANSYS Fluent to model the thermal and mechanical performance of a single-pass TEMA E-type shell and tube heat exchanger. Simulations will be conducted in a steady state for the case of six, eight, ten and twelve baffles. In the effort to develop comprehensive correlations which capture the effect of turbulence on thermal performance both the central baffle spacing and the inlet mass flow rate are varied. For each model, simulations are performed over the range of 0.50 kg/s to 2.0 kg/s with increments of 0.25 kg/s. The results from the CFD analysis are then fit using non-linear least squares regression in the MATLAB Curve Fitter Toolbox to develop four sets of correlations in the form of $Nu = C.Re^{a}.Pr^{b}$ where a confidence level of 95 % is used for modelling. Overall, the study is an effort to highlight the potential present in combining CFD with tools such as non-linear least squares regressions for the design and optimization of shell and tube heat exchangers.

2 METHODOLOGY

2.1 Physical Modelling of STHX

The shell and tube heat exchangers were modelled using the ANSYS Design Modeler ® tool based on the dimensions shown in Table 1 below (Ozden & Tari, 2010). For simulation purposes, the heat transfer through the thickness of the baffles was neglected.

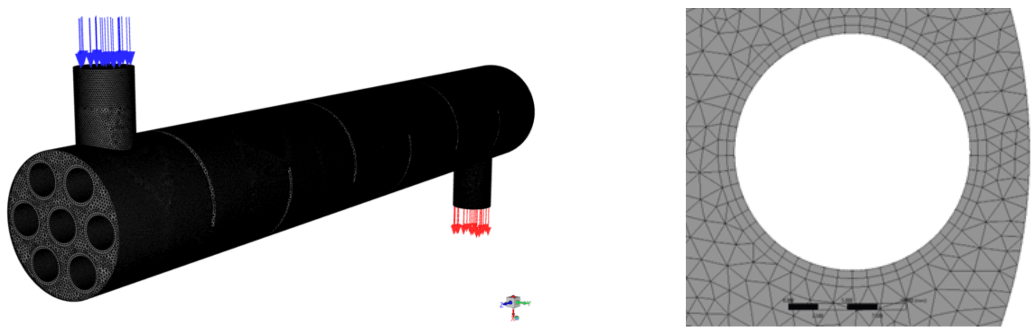
Table 2. Dimensions of the shell and tube heat exchanger (Ozden & Tari, 2010).

Dimensional Parameter	Value / mm
Shell Diameter	90
Outside Tube Diameter	20
Number of tubes used	7 tubes
Length of the heat exchanger	600
Central baffle spacing	86

The working fluid considered is water while, Aluminum is taken as the material for the tube walls. The presence of narrow operating temperature ranges prevented the use of piecewise linear interpolation for the properties. Additionally, simulations were done in the steady state given similar studies have utilized this assumption (Ozden and Tari, 2010; Nedunchezhiyan Mukilarasan et al., 2022; Zhang et al., 2017).

2.2 Computational domain, mesh and boundary conditions

The shell side volume of the four models is modelled using an unstructured tetrahedral mesh in the ANSYS Meshing tool as shown in Figure 2 below. To select the ideal mesh, a balance of accuracy and computational memory four grids were generated. The element numbers modelled were ~1.2 million, ~2.2 million, ~3.3 million and ~5.0 million. Given that the discrepancy between the overall heat transfer coefficients for the grid of 3.3 and 5.0 million was less than 2% the grid of element size 3.3 million was selected. This aligns with the level of discrepancy which has been considered reasonable in similar studies, such as that done by Abbasi et al. (2020).



(a) (b)

Figure 13. (a) Unstructured tetrahedral mesh with boundary conditions applied (b) Zoomed in view of the mesh applied at the tube wall.

Analysis of the contours revealed significant turbulence on the shell side of the heat exchanger. Considering merit factors such as the model's ability to capture flow detail while maintaining minimal computational time and offering simple closure, the standard k- ε turbulence model was selected (Aslam Bhutta et al., 2012; Ozden and Tari, 2010; Abbasi et al., 2020). Prism layers and a scalable wall function were employed to resolve the near-wall flow in combination with the turbulence model (El Maakoul et al., 2016; Shahril et al. 2017).This was done to limit the $y^+ \approx 11.5$, ensuring the flow remained compatible with the turbulence model. Control of this non-dimensional parameter which represents the distance from the wall to the first node is critical given its ability to define the near wall flow (Shahril et al., 2017). The fluid domain was defined using inlet and outlet nozzle boundary conditions. Standard mass flow and pressure boundary conditions were used respectively (Ozden and Tari, 2010, Nedunchezhiyan Mukilarasan et al., 2022; Abbasi et al., 2020; Zhang et al., 2017). Additionally, a fixed thermal condition of 450 K was used for the tube walls while a zero-heat flux condition was assumed for the shell wall.

2.3 Non-linear least squares regression for correlation development

The parameter formulations in the modelled correlations are similar to those used in established design methods such as the Bell-Delaware, Kern's and flow-stream analysis methods (Jamil et al., 2020). Equations 1-3 below show the basis of calculating the *Re*, a similar analogy is used to evaluate both the *Pr* and *Nu* (Alperen et al., 2019).

$$D_e = \frac{4\left(\frac{P_t}{2} * 0.87 * P_t - 0.5 * \pi * \frac{d_o^2}{4}\right)}{\pi * \frac{d_o}{2}} \tag{1}$$

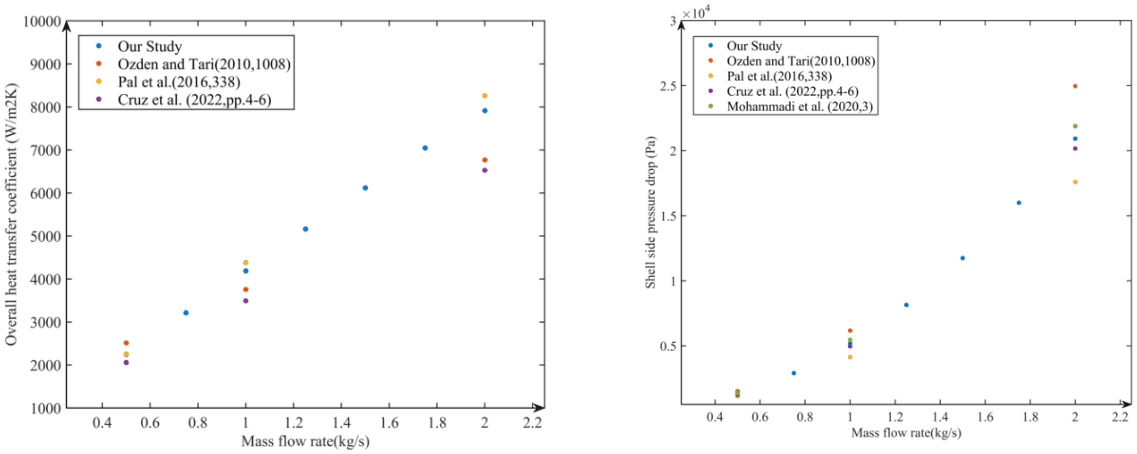
$$A_s = \frac{(P_t - d_o) * D_s * B}{P_t} \tag{2}$$

$$Re = \frac{\left(\frac{\dot{m}_s}{A_s}\right) * D_e}{\mu} \tag{3}$$

3 RESULTS AND DISCUSSION

3.1 Model verification and thermo-mechanical results

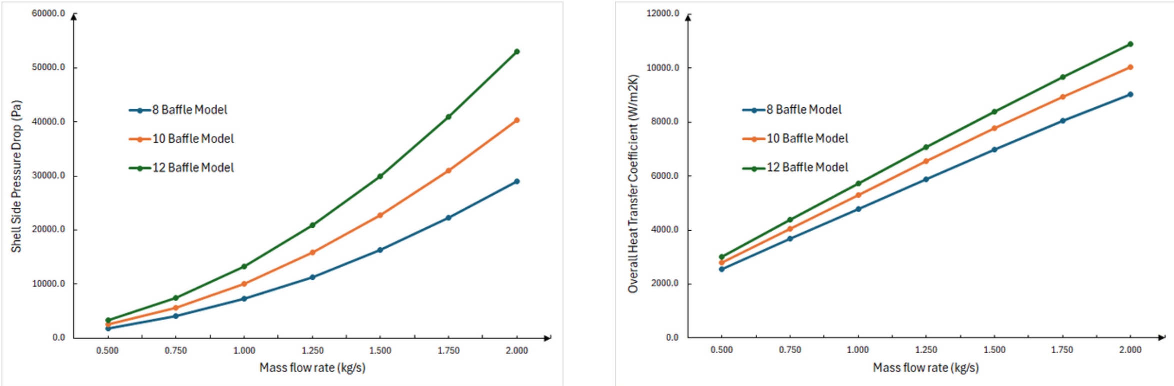
Simulations revealed that increasing mass increased both the overall heat transfer coefficient and the shell side pressure drop. The results align with the trend shown in similar studies, however, differences are observed. This could be attributed to variations in the modelling schemes used. This argument is supported by the agreement visible with the study done by Pal et al. (2016) given they used a similar methodology for modelling.



(a) (b)

Figure 14: CFD Results for the six-baffle model (a) Overall heat transfer coefficient vs mass flow rate (b) Shell side pressure drop vs mass flow rate.

The overall heat transfer coefficient and shell side pressure were observed to be significantly affected by the increase in number of baffles and the mass flow rate as shown in Figure 4 (a) and (b) below. The overall heat transfer was observed to improve, which could be related to an improvement in the shell side flow phenomena and amount of area available for thermal transfer. Pressure drop was negatively affected, where this would lead to a higher pumping power to maintain the consistent flow of fluid. These phenomena are discussed in the context of the contours in the next section.

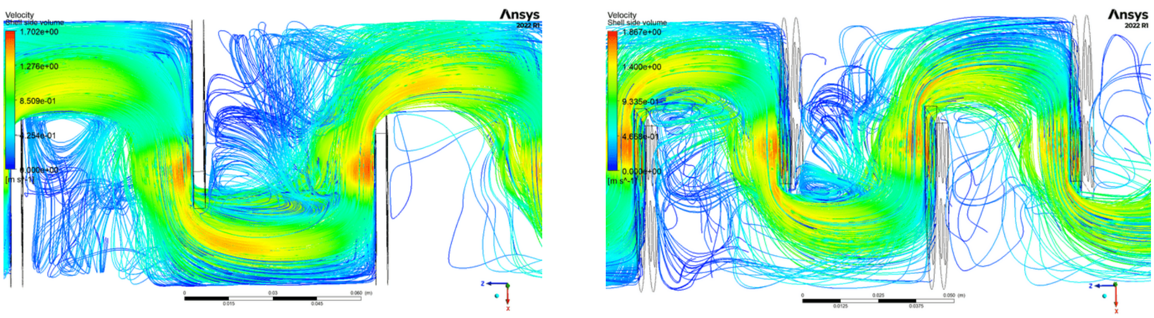


(a) (b)

Figure 15. CFD based results with mass flow rate (a) Shell Side Pressure Drop (b) Overall Heat Transfer Coefficient

3.2 Pathlines

The developed pathlines reveal the fluid flow on the shell side of the heat exchanger and its dependence on the number of baffles. Fluid follows a zig-zag pattern with regions of dead zones or areas of recirculation observed behind certain baffles. The significance of recirculation areas is high at a lower number of baffles as evident in the case of 8 baffles (refer Figure 5 (a)). This is related to decreased thermal performance, as a greater amount of tube area is bypassed, thereby reducing the surface area available for heat transfer (Ozden and Tari, 2010; Alhulaifi, 2024). However, decreasing the baffle spacing was observed to improve the flow with increased flow uniformity and reduced areas of flow recirculation (refer to Figure 5 (b) and (c) below). This led to improved thermal performance as discussed earlier.



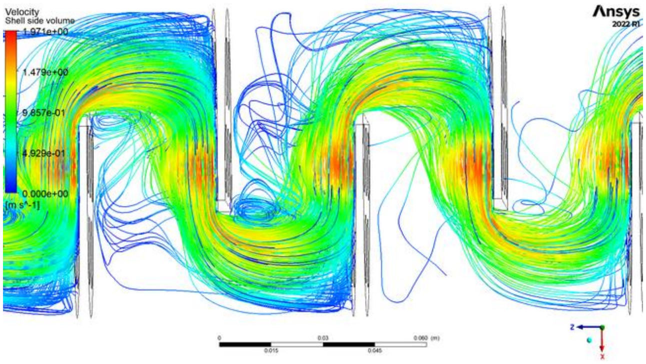


Figure 16. Velocity path lines at 1.0 kg/s (a) 8 baffles (b) 10 baffles and (c) 12 baffles.

3.3 Pressure drop along the length

The pressure contours reveal that the pressure drops from the inlet to the outlet along with the length of the heat exchanger. High pressure is observed at the inlet of the model due to the impingement of the fluid entering at the inlet nozzle. Flow expansion and contraction due to the presence of baffles is a key factor that introduces a significant pressure drop, as the barrier to fluid flow increases as evident in the contours. Additionally, increasing the mass flow rate is related to an increased pressure drop with relative similarity in the contours (refer Figure 6(a) – (c)).

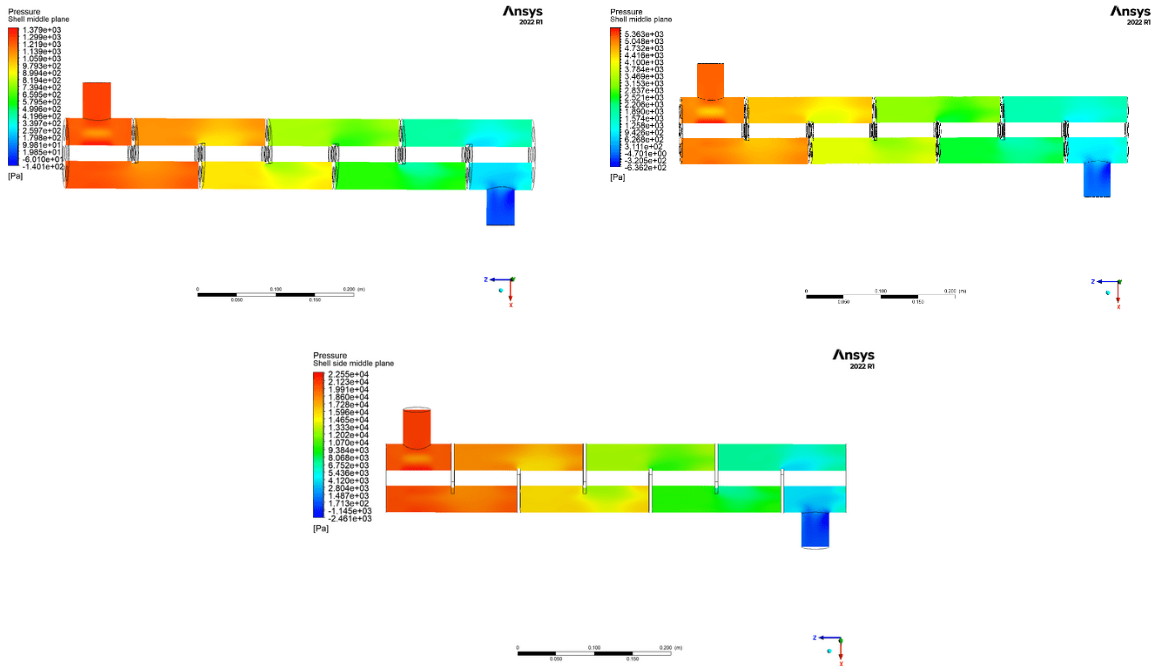


Figure 17. Pressure contours at the middle plane for six baffle model (a) 0.5 kg /s (b) 1.0 kg/s and (c) 2.0 kg/s.

Another key observation is the presence of back pressure at the outlet nozzle; this was due to improper flow development at the outlet plane. A similar observation was reported in studies by Pal et al. (2016) and Sauro & Lewis (2012).

3.4 CFD-Based Thermal Correlations for the four models

A summary of the correlations developed using NLSQR is shown in Figure 7 below.

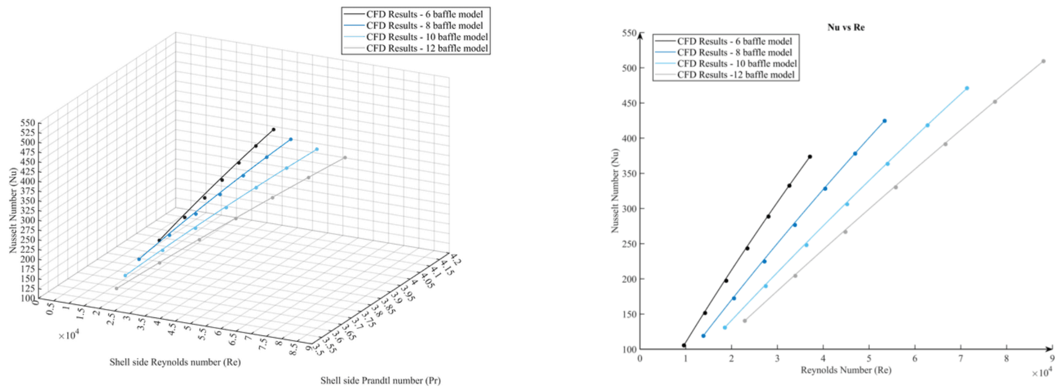


Figure 18 (a) *Nu* variation with *Pr* and *Re* for the four models (b) *Nu* variation with *Re* for the four models.

Nu is observed to increase significantly with *Re*, indicative of increasing turbulence. The *Re* increase with both the mass flow rate and the central baffle spacing. The observed variation is a result of improved flow phenomena as revealed earlier and in similar studies done by Ozden & Tari (2010) and Abbasi et al. (2020). Additionally, the minimal variation in *Pr* in the observed results could be due to slight variations in the thermal properties of the working fluid. A summary of the correlations obtained using NLSQR are shown in Table 2 below.

Table 3. Summary of the correlations developed for the four models.

Baffles	*C*	*a*	*b*	*Re Range*	*Pr Range*
6	0.2312	0.9972	-2.1907	$9600 < Re < 37100$	$3.68 < Pr < 4.12$
8	0.139	1.0065	-2.1304	$13900 < Re < 53400$	$3.80 < Pr < 3.96$
10	0.0809	1.0103	-1.9586	$18500 < Re < 71400$	$3.66 < Pr < 3.81$
12	0.0319	0.9964	-1.2735	$22900 < Re < 88100$	$3.54 < Pr < 3.70$

The overall precision and reliability of the predicted correlations were evaluated by considering the confidence intervals (CI) widths of the predicted coefficients, a standard methodology used in numerical model evaluation (Sauro and Lewis, 2012). A summary of the upper and lower bounds of the coefficients at a 95 % confidence level is shown in Table 3 below.

Table 4. Summary of the upper and lower bounds for predicted coefficients with a 95 % confidence level.

Nb	**6**			**8**		
Parameter	**Predicted value**	**Lower bound**	**Upper bound**	**Predicted value**	**Lower bound**	**Upper bound**
C	0.2312	0.014	0.4485	0.1390	0.0158	0.2623
a	0.9972	0.9758	1.0186	1.0065	0.9817	1.0312
b	-2.1907	-3.0117	-1.3696	-2.1304	-2.9624	-1.2914
Nb	**10**			**12**		
Parameter	**Predicted value**	**Lower bound**	**Upper bound**	**Predicted value**	**Lower bound**	**Upper bound**
C	0.0809	0.0263	0.1355	0.0319	0.0207	0.0431
a	1.013	0.9900	1.0307	0.9964	0.9852	1.0076
b	-1.9586	-2.6318	-1.2854	-1.2735	-1.639	-0.908

The results show that the CI decreases with the increasing numbers of baffles and this may again be related to the observed improvement in the flow phenomena. Overall the curve fit was stable, as evidenced by the significant improvement in the goodness-of-fit statistics. The sum of squares due to error (SSE) decreased from **0.2064** to **0.1279** for the six and twelve baffle models respectively. Additionally, the root mean squared error (RMSE) decreased from **0.1703** to **0.0654**. The results of these statistical parameters indicate that the random component of error in the model is minimal and that it can be used as a prediction tool with reasonable confidence, provided the operating ranges are applicable. This can be further verified by performing a practical simulation at the given boundary conditions.

4 CONCLUSIONS

In this study, a single pass TEMA E-type shell and tube heat exchanger was modelled in a steady state. The effects of turbulence on thermal performance were explored by varying turbulence level as a function of both mass flow and central baffle spacing for six, eight, ten, and twelve baffle models. The key finding was that decreasing baffle spacing improved the thermal performance at a cost of increased pressure drop. The CFD results were then combined with NLSQR to develop four sets of correlations relating Nu to Pr and Re. Given the reasonable goodness-of-fit parameters and confidence intervals the study reveals that the model can be used as a prediction tool with a reasonable level of confidence highlighting the potential of such workflows and integrations for practical design and optimization purposes. Further research work in the area needs to be conducted to explore the potential by improving the modelling schemes such as the grid quality and convergence tolerances, consideration of the flow in tube sides, flow conditions at inlet and outlets and the application of advanced prediction tools such as ANN (Artificial Neural Networks) to utilize the capability of decision making improving the potential and scope of thermal design and optimization. Another key improvement to judge the accuracy of the developed model would be to conduct practical simulations at the given boundary conditions.

REFERENCES

Abbasi, H. R., Sharifi Sedeh, E., Pourrahmani, H., & Mohammadi, M. H. (2020). Shape optimization of segmental porous baffles for enhanced thermo-hydraulic performance of shell-and-tube heat exchanger. *Applied Thermal Engineering*, *180*, 115835. https://doi.org/10.1016/j.applthermaleng.2020.115835

Alhulaifi, A. S. (2024). Computational Fluid Dynamics Heat Transfer Analysis of Double Pipe Heat Exchanger and Flow Characteristics Using Nanofluid TiO2 with Water. *Designs*, *8*(3), 39. https://doi.org/10.3390/designs8030039

Alperen, M. A., Kayabaşi, E., & Kurt, H. (2019). Detailed comparison of the methods used in the heat transfer coefficient and pressure loss calculation of shell side of shell and tube heat exchangers with the experimental results. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 1–20. https://doi.org/10.1080/15567036.2019.1672835

Aslam Bhutta, M. M., Hayat, N., Bashir, M. H., Khan, A. R., Ahmad, K. N., & Khan, S. (2012). CFD applications in various heat exchangers design: A review. *Applied Thermal Engineering*, *32*, 1–12. https://doi.org/10.1016/j.applthermaleng.2011.09.001

Bichkar, P., Dandgaval, O., Dalvi, P., Godase, R., & Dey, T. (2018). Study of Shell and Tube Heat Exchanger with the Effect of Types of Baffles. *Procedia Manufacturing*, *20*, 195–200. https://doi.org/10.1016/j.promfg.2018.02.028

El Maakoul, A., Laknizi, A., Saadeddine, S., El Metoui, M., Zaite, A., Meziane, M., & Ben Abdellah, A. (2016). Numerical comparison of shell-side performance for shell and tube heat exchangers with trefoil-hole, helical and segmental baffles. *Applied Thermal Engineering*, *109*, 175–185. https://doi.org/10.1016/j.applthermaleng.2016.08.067

Huang, Z., Hwang, Y., & Radermacher, R. (2017). Review of nature-inspired heat exchanger technology. *International Journal of Refrigeration*, *78*, 1–17. https://doi.org/10.1016/j.ijrefrig.2017.03.006

Jamil, M. A., Goraya, T. S., Shahzad, M. W., & Zubair, S. M. (2020). Exergoeconomic optimization of a shell-and-tube heat exchanger. *Energy Conversion and Management*, *226*, 113462. https://doi.org/10.1016/j.enconman.2020.113462

M. Necati Özışık. (1985). *Heat Transfer*. McGraw-Hill Science, Engineering & Mathematics.

Mangrulkar, C. K., Dhoble, A. S., Chamoli, S., Gupta, A., & Gawande, V. B. (2019). Recent advancement in heat transfer and fluid flow characteristics in cross flow heat exchangers. *Renewable and Sustainable Energy Reviews*, *113*, 109220. https://doi.org/10.1016/j.rser.2019.06.027

Marcinkowski, M., Dawid Taler, Katarzyna Węglarz, & Taler, J. (2024). Advancements in analyzing air-side heat transfer coefficient on the individual tube rows in finned heat exchangers: Comparative study of three CFD methods. *Energy*, *307*, 132754–132754. https://doi.org/10.1016/j.energy.2024.132754

Marcinkowski, M., Dawid Taler, Taler, J., & Katarzyna Węglarz. (2021). Thermal Calculations of Four-Row Plate-Fin and Tube Heat Exchanger Taking into Account Different Air-Side Correlations on Individual Rows of Tubes for Low Reynold Numbers. *Energies*, *14*(21), 6978–6978. https://doi.org/10.3390/en14216978

Nedunchezhiyan Mukilarasan, Karthikeyan, R., Ramalingam, S., Damodharan Dillikannan, Ravikumar, J., Sampath, S., & Gopal Kaliyaperumal. (2022). Influence of baffles in heat transfer fluid characteristics using CFD evaluation. *International Journal of Ambient Energy*, *43*(1), 7088–7100. https://doi.org/10.1080/01430750.2022.2063175

Ozden, E., & Tari, I. (2010). Shell side CFD analysis of a small shell-and-tube heat exchanger. *Energy Conversion and Management*, *51*(5), 1004–1014. https://doi.org/10.1016/j.enconman.2009.12.003

Pal, E., Kumar, I., Joshi, J. B., & Maheshwari, N. K. (2016). CFD simulations of shell-side flow in a shell-and-tube type heat exchanger with and without baffles. *Chemical Engineering Science*, *143*, 314–340. https://doi.org/10.1016/j.ces.2016.01.011

Sauro, J., & Lewis, J. R. (2012). How Precise Are Our Estimates? Confidence Intervals. *Quantifying the User Experience*, 19–39. https://doi.org/10.1016/b978-0-12-384968-7.00003-5

Sekulić, D. P., & Shah, R. K. (2023). *Fundamentals of Heat Exchanger Design*. Wiley. https://doi.org/10.1002/9781119883296

Shahril, S. M., Quadir, G. A., Amin, N. A. M., & Badruddin, I. A. (2017). Thermo hydraulic performance analysis of a shell-and-double concentric tube heat exchanger using CFD. *International Journal of Heat and Mass Transfer*, *105*, 781–798. https://doi.org/10.1016/j.ijheatmasstransfer.2016.10.021

Zhang, X., Han, D., He, W., Yue, C., & Pu, W. (2017). Numerical simulation on a novel shell-and-tube heat exchanger with screw cinquefoil orifice baffles. *Advances in Mechanical Engineering*, *9*(8), 168781401771766-168781401771766. https://doi.org/10.1177/1687814017717665

Task Scheduling Problem in Fog Computing Environment with improved Memetic algorithm

Dang Van Thang, Volkov Artem, Ammar Muthanna, Olga Vorozheikina, Andrey Koucheryavy
The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, St. Petersburg, 193232, Russian Federation

ABSTRACT

The task scheduling problem in fog computing is one of the key challenges in the development of fog computing within next-generation communication networks. Addressing this challenge requires balancing processing performance with resource constraints while meeting network conditions. Given the distributed and heterogeneous nature, as well as the dynamic topology, optimally allocating tasks to fog nodes is a complex issue. To contribute to solving this problem, we propose a task scheduling method based on an improved Memetic algorithm. The proposed method leverages the strengths of evolutionary algorithms and local search, while incorporating a task restructuring mechanism, to enhance allocation efficiency and task processing in the fog computing environment. Simulation experiments demonstrate that the proposed method outperforms genetic algorithms, Round-Robin, Greedy methods, and the Ant Colony Optimization algorithm in terms of efficiency. This study provides a fresh, simpler approach that aligns with network conditions while still achieving the desired performance.

KEYWORDS: *6G, fog computing, task allocation, memetic algorithm, task restructuring mechanism, next-generation networks.*

1. INTRODUCTION

The explosion of the Internet of Things (IoT) has generated a massive amount of data and an urgent need to process it. Meanwhile, the traditional cloud computing model is increasingly unable to meet the growing demands for bandwidth, latency, and robust security from remote edge devices. Fog computing has emerged as a highly innovative solution to address this issue. However, a reality remains: fog nodes do not always meet real-time dynamic demands due to limitations in computational resources. The fog computing environment faces several significant constraints. First, fog nodes often have limited computation time due to their dynamic nature, with the ability to join or leave a cluster unpredictably. Second, the computational resources of fog nodes, including CPU, RAM, and bandwidth, are typically more restricted compared to cloud data centers, posing challenges in handling complex tasks or multiple tasks simultaneously. Finally, the structure of a fog cluster lacks stability because nodes can unexpectedly disconnect or leave, affecting task allocation and execution processes. These limitations require scheduling methods that are not only optimized for performance but also flexible and quick to adapt to environmental changes.

To tackle this issue, various task scheduling methods for the fog computing environment have been researched and proven effective to varying degrees. Reality shows that, with their limited computational capacity, applying advanced superior algorithms such as reinforcement learning, deep learning, or other methods to the fog computing environment remains challenging. Therefore, a suitable scheduling algorithm that optimizes task allocation, ensures cost-effectiveness and energy efficiency, and meets other imposed requirements is a highly important and pressing matter.

Within the scope of this paper, we propose a task scheduling method for Fog Computing based on the application of the Memetic algorithm and a task restructuring mechanism. With a simple structure yet integrating evolutionary methods and local search, along with a task schedule restructuring approach, the proposed method has demonstrated its effectiveness compared to several other methods such as Greedy, GA, Round Robin, and ACO. The results of the method are described and evaluated in detail in the experimental and simulation section later in the paper. The application of the proposed method in this paper contributes to addressing the ongoing challenge of task scheduling in dynamic and resource-constrained environments like Fog Computing.

This study is divided into 5 sections, in which Sections 1 and 2 introduce the problem and the motivation for the research. In Sectiomethod andsent the system structure of Fog Computing, the proposed method, and explain related issues. Experimental simulation, data collection, and result evaluation are presented in Section 4. Finally, the conclusion and future research directions are provided.

2. RESEARCH MOTIVATION

Fog computing emerges as a significant advancement in processing data from the Internet of Things (IoT) network, particularly in applications requiring low latency and optimal bandwidth, such as telepresence services, telecommunication-type holography, and sensory Internet [1].

Researched and developed as a solution to overcome critical limitations of the traditional cloud computing model, Fog computing plays a substantial role in managing, processing, and analyzing data, while also meeting certain requirements such as latency, optimal bandwidth usage, and real-time responsiveness. Additionally, Fog computing brings computational and data processing tasks to more remote locations where edge devices are situated, which are generating an enormous volume of data. This helps save transmission costs and mitigate network congestion [2]. However, Fog nodes are facing significant challenges in their development due to limitations in computational resources and the capacity to handle tasks compared to the cloud [3].

One of the greatest challenges Fog computing encounters is the task optimization problem under conditions constrained by computational capacity. Therefore, efficient techniques for task scheduling and resource management are needed to enhance the operational efficiency of Fog nodes [4]. Several methods have been proposed to address this issue, including algorithms based on search, optimization, and machine learning approaches. Heuristic algorithms such as the Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), and Simulated Annealing (SA) are commonly applied to solve the task scheduling problem in the Fog computing environment. However, a prominent common limitation of these algorithms is the difficulty in achieving a globally optimal solution. The explanation for this limitation is that these algorithms often get trapped at local optimization bottlenecks, particularly in the highly dynamic Fog computing environment. Furthermore, as mentioned above, the constraints of Fog nodes lead to difficulties in deploying advanced algorithms such as deep learning and reinforcement learning, which require substantial computational resources [5].

Within the scope of this research, to address the limitations mentioned above in the Fog computing structure, we propose a hybrid Memetic Algorithm (MA) combined with a task restructuring mechanism, as a feasible solution. The Memetic Algorithm is a hybrid metaheuristic that integrates local search processes within the framework of evolutionary algorithms [6], [7]. This class of hybrid algorithms leverages the strengths of both local and global search strategies, yielding better solutions. With the integrated task restructuring mechanism, the system can quickly detect unavailable or overloaded nodes and transfer the affected tasks to other nodes within the cluster. In this way, the restructuring mechanism not only ensures the continuity of task execution but also optimizes resource utilization in a dynamic environment. This is particularly crucial in real-time applications such as telepresence, sensory Internet, and holography, where latency and continuity are key factors. The operational mechanism of the proposed method is described in the subsequent section of this study, and additionally, the effectiveness of the method is demonstrated in the simulation and result evaluation section.

3. SYSTEM ARCHITECTURE AND PROPOSED METHOD

3.1 Architecture of Fog Computing System

First introduced by Cisco, Fog Computing is a technological innovation encompassing AI (Artificial Intelligence) and IoT [8]. Many enterprises and organizations in the industry have begun deploying fog computing ecosystems, including communication providers such as Cisco and Huawei, as well as numerous cloud computing and IoT enterprises. The concept of Fog Computing is understood as a method of transferring part of the data center operations to the network edge and to the level of user devices. It provides a smaller amount of storage, processing, and network resources in the form of sharing between the CC data center and the end device. Fog computing constitutes distributed dynamic architectures, comprising multiple devices and small data centers located close to the sources of data generation, such as sensors, mobile phones, and portable devices.

In a hierarchical architecture, fog computing plays the role of an intermediary, positioned between the end layer and the remote cloud computing layer to ensure the continuity of service provision along the device-fog-cloud path [9]. In this context, the hierarchical model of the fog computing architecture is illustrated in Fig. 1.

End device layer: The end device layer mainly consists of sensor devices and various types of end devices or sensor devices responsible for processing during the data collection process.

Fog computing layer: The fog computing layer is located at a position between the end device layer and the cloud computing layer. Fog nodes deployed at the network edge primarily provide services for latency-sensitive tasks. For some common computing devices, such as smart sensors, smart processing devices, smart multimedia devices, they can also be used as fog nodes. Fog nodes have certain computational capabilities, communication capabilities, and storage capabilities, and they operate between the cloud and end devices. Fog computing has overcome the high latency of cloud computing, thus the quality of service of real-time applications can be guaranteed. However, the capabilities of fog nodes are ultimately limited; fog nodes also exhibit heterogeneity and strong dynamics when facing a large amount of real-time processing data, therefore, how to analyze data quickly and efficiently is the primary goal of fog computing.

Cloud computing layer: The cloud computing layer contains the cloud data center, cloud storage, cloud computing devices, providing remote services for smart production lines. High-performance computing devices and large-capacity storage devices can transmit, compute, and store all kinds of massive data from end devices through a coordinated management control center to provide comprehensive, high-quality services to end users. Due to remote deployment, cloud computing cannot complete real-time data processing and analysis, thus cloud computing cannot meet the QoS demands of real-time tasks.

In the structure of fog computing, Fog nodes have average computational and storage capabilities, suitable for initial processing tasks or real-time applications. They play the role of processing data close to the source (at the network edge) to reduce latency and offload the cloud layer. Therefore, the problem of scheduling fog nodes plays a very important role, and it is necessary to focus on the allocation and management of computational tasks on fog nodes in the most efficient manner.

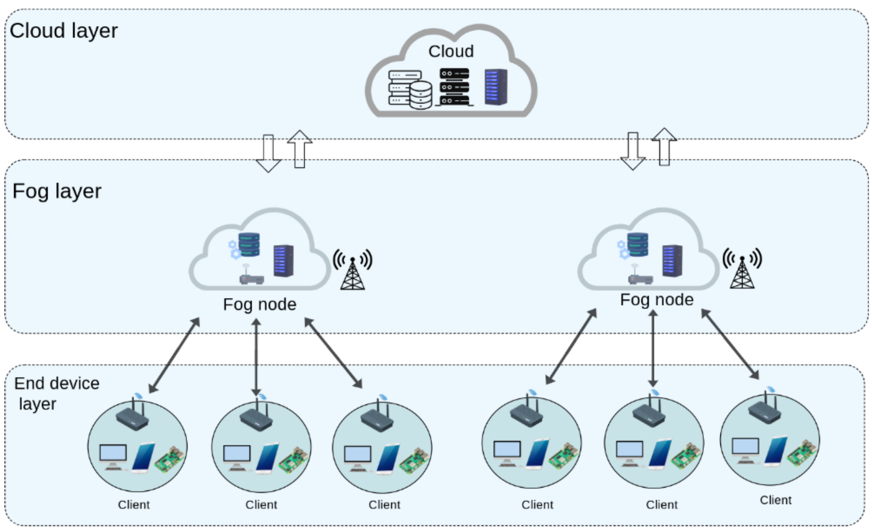


Figure 1. Hierarchical model of fog computing architecture

3.2 Memetic alrorithm

Memetic (MA: Memetic Algorithm) (Afsar, 2022) is an optimization algorithm combining the evolutionary method and local search to solve difficult and complex problems [10]. MA has the ability to overcome some limitations of traditional evolutionary algorithms and utilizes the "knowledge" of previous evolutionary generations to guide the search process for good individuals in subsequent generations. In the MA, each individual is a solution to the problem, satisfying the constraints posed by the problem. To create an individual, each evolutionary generation performs the steps: selection, crossover, and mutation. The computational steps of MA are like GA; however, the improvement of MA lies in the addition of the local search method to the computational steps. The local search method (Local Search) has the role of refining individuals after expanding the search space to individuals neighboring the current individual. Therefore, this method can find and generate better individuals; not only that, but Local Search also helps the algorithm converge faster and improves the quality of the solution compared to the traditional GA algorithm.

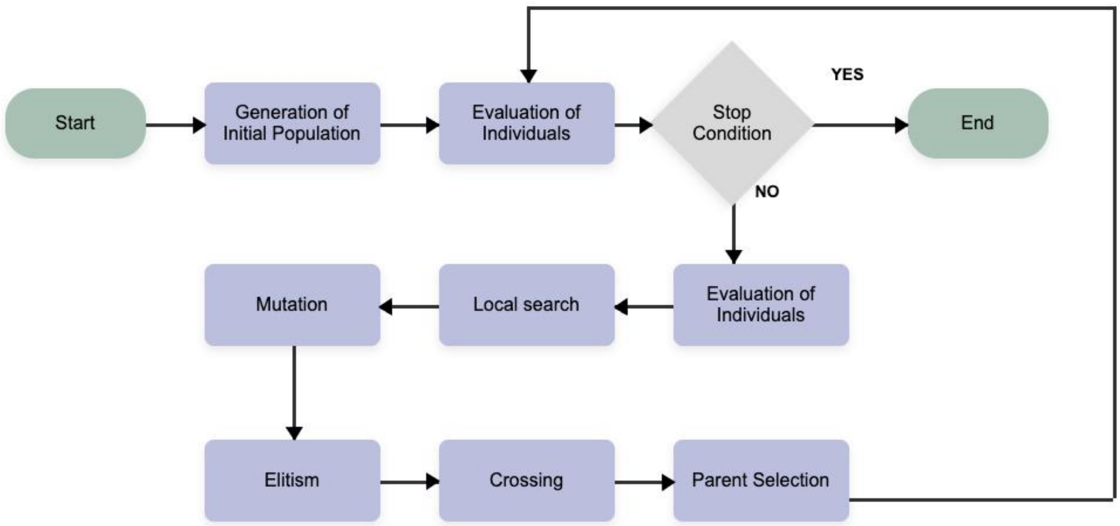


Figure 2. Hybrid Memetic algorithm

The Memetic Algorithm performs the following steps:

Step 1. Generation of Initial Population: Initially, create a population with a specified number of individuals. Each individual is a task allocation scheme (makespan).

Step 2. Evaluation of Individuals: Calculate the value of each individual in the population based on the objective function posed by the problem.

Step 3. Stop condition: If the stop condition is satisfied, the algorithm terminates and returns the best solution. If not, the algorithm proceeds to the subsequent steps.

Step 4. Parent Selection: Select suitable individuals for crossover to create new individuals in the next generation.

Step 5. Crossover: Combine individuals from the previous generation to create new individuals. This process uses information from multiple individuals participating in the crossover to generate a new individual, with the new individual inheriting the characteristics of the individuals that produced it.

Step 6. Elitism: Apply the elitism strategy to preserve the number of best individuals from the current generation to the next generation.

Step 7. Mutation: Add random attributes to the new individual after crossover.

Step 8. Local Search: Perform computations with neighboring individuals to search for better individuals.

Step 9. Recording (Acceptance): Record the individual of the current evolutionary generation if it is better than the previous generation.

Repeat the steps from (2) to (9) over multiple generations until the stop condition is satisfied.

3.3 Local Search

As mentioned in the previous section, the advantage of the memetic algorithm over other evolutionary algorithms lies in its integration of local search techniques to achieve superior performance in solving complex combinatorial optimization problems. In [11], a standard local search algorithm was presented, based on a local search problem. This algorithm has been widely recognized and utilized in numerous studies, such as [12], and [13]. Specifically, the standard local search algorithm has the following pseudo-code:

Standard Local Search

```
Begin
    Generate an initial solution s for the problem x;
    Repeat Until (a local optimum is reached)
    Do
        Use s and x to generate the next neighbor n_{x,s};
        If (n_{x,s} is better than s) Then
            s := n_{x,s};
        EndIf
    EndDo
    Return s;
End
```

However, in a special environment like fog computing, where there are resource constraints and high dynamism, the Standard Local Search algorithm is truly unsuitable because it consumes a significant amount of resources and time. To explain this, we can consider its operating principle: this algorithm continuously searches until it reaches a local optimum, requiring the exploration of the entire neighborhood space and repeating the process until no further improvement is feasible. This leads to a high consumption of computational resources, resulting in an inability to meet the demands of fast and efficient response systems.

As a solution to this issue, in the paper, we present a Best Improvement Local Search algorithm. The key difference between the proposed algorithm and the Standard Local Search lies in limiting the exploration of the solution space by generating only a fixed number of random neighbors and selecting the neighbor with the best fitness value. If the best neighbor provides an improvement over the current solution, it is accepted. Conversely, the algorithm retains the initial solution and terminates. The proposed algorithm not only significantly reduces computational time and resource usage but also maintains the necessary accuracy, thereby improving the efficiency of task scheduling in the fog computing environment.

The pseudo-code of the proposed Best Improvement Local Search technique is as follows:

Best Improvement Local Search

Input: Current solution s, list of fog nodes, list of tasks, maximum number of neighbors

Output: Improved solution or s if no improvement is found

```
Begin
    Set current_fitness = fitness_function(s, fog_nodes, tasks)
    Set best_neighbor = s, best_fitness = current_fitness
    For i from 1 to max_neighbors do
        Create neighbor by randomly reassigning one task in s
        Set neighbor_fitness = fitness_function(neighbor, fog_nodes, tasks)
        If neighbor_fitness < best_fitness then
            Set best_fitness = neighbor_fitness, best_neighbor = neighbor
        End If
    End For
    Return best_neighbor if best_fitness < current_fitness, else s
End
```

3.4 Task restructuring mechanism

In a dynamic environment like Fog computing, there always exist constraints on resources for performing tasks. These constraints can be network conditions, the computational power of Fog nodes, latency, or initial service quality requirements. Within the framework of this paper, we use the resource redistribution method for two main purposes:

- *First*, to provide a more optimal schedule if the task restructuring mechanism, yields better results. In the opposite case, retain the result that MA previously found.
- *Second,* in the case where Fog nodes are overloaded, lose connection and cannot process all tasks, or cannot participate in the processing process. The redistribution mechanism will play the role of redistributing tasks to resources (Fog nodes) in the most optimal way.

After each generation, the MA algorithm will find the best schedule (C_{best}). Based on this schedule, changes are made to the resources performing each task while still ensuring the initial resource constraints are met. This method helps expand the search space, further optimize the schedule, and most importantly, avoid local optima.

The redistribution method is performed through the following steps:

Step 1: Search for the most heavily used resource C_{best}, call L_{best}. This is the resource that finishes its work last among the resources used to execute the project.

Step 2: Iterate through each task performed by the resource L_{best} in the order of task priority, denoted as W_{best}^{R}.

Step 3: For each task found $W_i \in W_{best}^{R}$, perform the following two steps:

- Find all resources capable of performing task W_i other than the resource currently performing it, L_{best}, denoted L_W;

- Iterate through each resource capable of performing W_i, for each resource $L_{best}^{W} \in L_W$, continue with two sub-steps: (1) assign the performing resource as W_j^{W}, (2) Remove W_i from the list of tasks that L_{best} is currently performing

After the above step, a new schedule is obitained, and the makespan of the new schedule, $C_{newbest}$, is calculated. If it is better C_{best}, replace the makespan with $C_{newbest}$ and stop; if not, start executing the verification loop again.

$$C_{best} = \begin{cases} C_{newbest} & if\ f(C_{newbest}) < f(C_{best}) \\ C_{best} & if\ f(C_{newbest}) > f(C_{best}) \end{cases} \tag{1}$$

This computational process helps the algorithm expand the search space by creating new individuals $C_{newbest}$ based on the best individual C_{best}. As a result, when applying the restructuring method after the MA's outcome, we can obtain better individuals. Moreover, in a dynamic environment like Fog computing and IoT devices, the issue of load imbalance in task scheduling management is also addressed thanks to this restructuring approach. Due to the highly dynamic nature, the structure of Fog nodes is unstable, so scenarios such as Fog nodes losing connection, exiting, or joining anew should be taken into account. In the paper, the task restructuring mechanism, will play a role in redistributing tasks to Fog nodes without needing to rerun the algorithms from scratch. Tasks on overloaded Fog nodes will be transferred to other nodes with more suitable resources.

To support the restructuring mechanism, we integrate an algorithm for detecting overloaded or problematic nodes (the algorithm is introduced in the pseudocode section). This algorithm operates independently and continuously monitors the status of Fog nodes by checking resources (CPU, RAM, bandwidth) and responsiveness. When an overloaded node (exceeding resource thresholds) or an unavailable node (disconnected) is detected, the algorithm triggers the restructuring mechanism to adjust the task schedule. This approach ensures that the system can quickly respond to changes in the environment, maintaining performance and reliability.

Below is an example of how our proposed method works. This section will help explain the working principle of the algorithm and the restructuring method, thereby applying the proposed approach to experimental environments that are closer to real-world scenarios.

***Example 1**:* In a fog computing system, 10 real tasks are allocated across 3 Fog nodes with the goal of optimizing completion time (makespan).

- Set of tasks $W = \{W_1, W_2, W_3, W_4, W_5, W_6, W_7, W_8, W_9, W_{10}\}$.
- Set of resources $L = \{L_1, L_2, L_3\}$.

The execution time is described in detail in the table and is calculated with a conversion of 1 time unit.

Table 1. Task execution time

Task	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}
Time	4	3	5	2	6	8	2	8	9	3

Fig.3 represents the allocation of tasks to Fog nodes for execution over time.

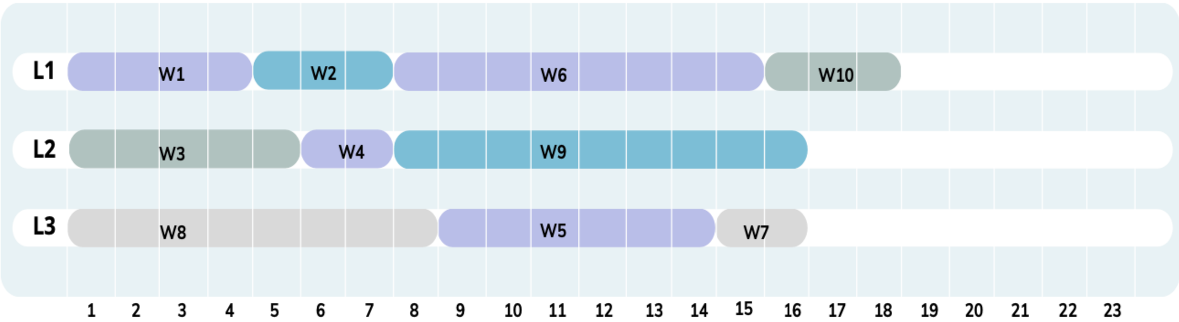


Figure 3. Diagram of task allocation to Fog nodes

With the scheme represented in Figure 3, we can represent this schedule in the form of a vector as follows:

Table 2. Task allocation to Fog nodes

Task	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}
Resource	1	1	2	2	3	1	3	3	2	1

Thus, resource L_1 will perform tasks W_1, W_2, W_5, W_{10}; resource L_2 will perform tasks W_3, W_4, W_9; resource L_3 will perform tasks W_5, W_7, W_8. With 3 Fog node and the constraints, the makespan found is 18 (time units).

Suppose the scheme described in Fig.3 is the optimal C_{best} returned by the MA algorithm, the process of restructuring C_{best} is performed to find an optimal solution (or in the case where a Fog node is overloaded, this process will play the role of balancing the task load for the Fog nodes).

The result is described in Fig.4, in which the new solution $C_{newbest}$ has a more optimal makespan compared to that of C_{best}.

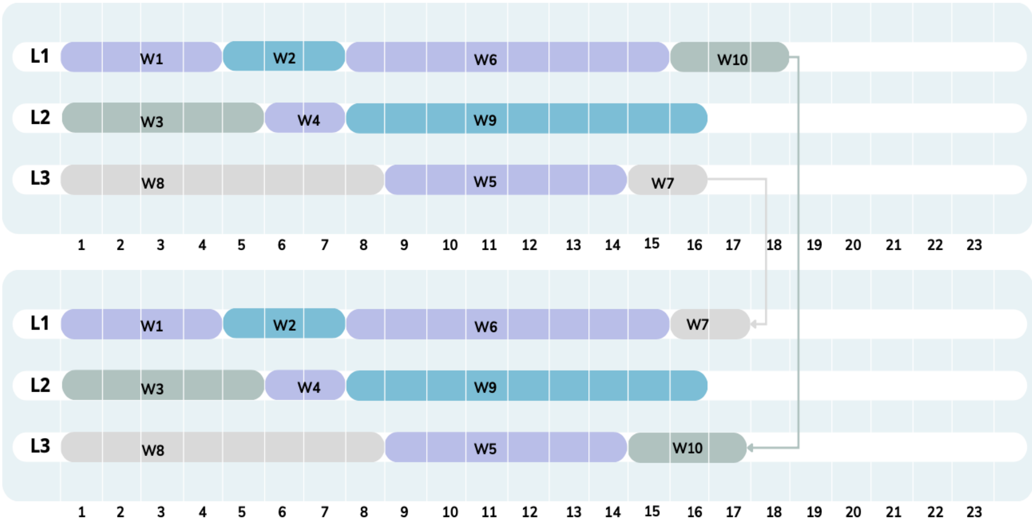


Figure 4. Result after applying schedule restructuring

In our example, a concept mentioned is the time unit, which is used as a basic measure to evaluate the execution time of tasks on fog nodes, thereby optimizing the overall completion time (makespan). This approach is quite suitable for scheduling problems focused on time optimization, especially in real-time applications. However, within the scope of task scheduling in a fog computing environment, it is not limited to just optimizing time but can also be defined based on various other criteria. These criteria depend not only on the objectives but also on the specific characteristics of each designed system. To avoid being confined within the framework of the time unit concept and to enhance the applicability of the proposed method, we will expand the concept of time into other equivalent concepts.

Considering the resource-based scheduling problem [14], we can take into account the resource unit to reflect the resource consumption level of tasks. Meanwhile, for the energy-based scheduling problem [15], the 'energy unit' can be considered as a criterion for task allocation. Or, in the priority-based scheduling problem, the criterion under consideration is the priority level, reflecting the importance of the task. In cost-based scheduling [16], the criterion examined is the cost of completing a task. Therefore, we would like to emphasize the practical applicability of this research, which is not limited to addressing scheduling problems based on time optimization but also extends to scheduling problems based on other criteria.The schedule redistribution method and the MA algorithm are described in the form of pseudocode as follows:

Algorithm 1. Detect Overloaded or Problem Nodes

Input: List of fog nodes, CPU threshold, RAM threshold, bandwidth threshold, ping timeout

Output: Lists of overloaded nodes and problem nodes

1. Begin
2. Initialize overloaded_nodes, problem_nodes as empty lists
3. For each node in fog_nodes do
5. If node.cpu_usage > cpu_threshold or node.ram_usage > ram_threshold or node.bandwidth_usage > bandwidth_threshold then
6. Add node to overloaded_nodes
7. End If
8. If not node.is_responsive or node.processing_time > ping_timeout then
9. Add node to problem_nodes
10. End If
11. End For
12. Return overloaded_nodes, problem_nodes
13. End

Algorithm 2 Task restructuring mechanism,

Input: Current solution, list of fog nodes, list of tasks, resource thresholds

Output: Updated solution after task reassignment

```
1. Begin
2.     Get overloaded_nodes, problem_nodes from Detect_Overloaded_Or_Problem_Nodes
3.     For each node in overloaded_nodes + problem_nodes do
4.         For each task in node.current_tasks do
5.             Find best_node in available_nodes with min compute_completion_time(task)
6.             If best_node exists then
7.                 Reassign task from node to best_node in solution
8.             End If
9.         End For
10.    End For
11.    Return solution
12. End
```

Algorithm 3 Main Proposed Method

Input: Population size, number of generations, list of fog nodes, list of tasks, resource thresholds

Output: Best solution and its fitness value

```
1. Begin
2.     Initialize population with random solutions
3.     Set best_solution = individual with min fitness in population
4.     For each generation from 1 to generations do
5.         Create new_population with best_solution
6.         While new_population.size < popsize do
7.             Select two parents based on fitness
8.             Create child1, child2 from parents
9.             If random < LOCAL_SEARCH_RATE then
10.                child1 = Best_Improvement_Local_Search(child1, fog_nodes, tasks)
11.            End If
12.            If random < LOCAL_SEARCH_RATE then
13.                child2 = Best_Improvement_Local_Search(child2, fog_nodes, tasks)
14.            End If
15.            Add child1, child2 to new_population
16.        End While
17.    Update population = new_population
18.    Update best_solution if better fitness found
19.    End For
20.    Set best_solution = Reassign_Tasks(best_solution, fog_nodes, tasks, thresholds)
21.    Return best_solution, fitness(best_solution)
22. End
```

4. SIMULATION AND RESULT EVALUATION

Within the scope of this paper, we have established a simulation for the fog computing environment [17, 18] to evaluate the effectiveness of the proposed algorithm in allocating tasks from IoT devices to fog nodes. The purpose of this simulation is to assess the efficiency of the proposed algorithm. Based on this, we compared the performance of four different task allocation methods. First, we simulated a method based on simple scheduling, such as Round Robin. This is a cyclic resource allocation method that ensures each node is processed fairly within the cycle. Next, the Greedy algorithm was considered, which is an algorithm that selects the locally optimal solution at each step with the hope of achieving a globally optimal outcome. For comparison with more complex algorithms, we sequentially examined the Genetic Algorithm (GA) and the Ant Colony Optimization (ACO) algorithm, which are algorithms with greater complexity. The simulation results section will provide a more comprehensive overview of the performance of the proposed algorithm.

Table 3. Simulation setup parameters

Parameter	Basic	Increased tasks	Increased nodes	Node overload
Number of Fog nodes	15	15	30	15 (2 overload, 2 error)
Number of tasks	600	1200	600	600
Population size	50	50	50	50
Crossover rate	0.8	0.8	0.8	0.8
Mutation rate	0.2	0.2	0.2	0.2
Maximum generations	30	30	30	30
Number of independent runs	10	10	10	10

The model is designed based on a centralized Fog Computing system, consisting of three main components: fog nodes, tasks, and the proposed algorithm (an improved Memetic Algorithm). This system aims to address the task allocation problem in the Fog Computing environment, with the goal of optimizing processing time, reducing latency, and handling situations involving node overload or failure. The simulation was conducted across four different scenarios, reflecting real-world conditions in the Fog Computing system: a baseline scenario, a scenario with an increased number of tasks, a scenario with an increased number of fog nodes, and finally, a scenario with overloaded and faulty nodes. The parameter settings for these scenarios are introduced in Table 3.

The configuration we set up Fog nodes and tasks is described through Tables 4.

Table 4. Simulation setup parameters for fog nodes and tasks

Parameter	Description	Value
Fog node setup parameters		
CPU capacity	Processing capacity of the node	Random from 2500 to 3500 (MIPS)
Bandwidth	Data transmission speed	Random from 10 to 50 (Mbps)
Current CPU load	Initial CPU load	0 (reset before each run)
Task list	List of assigned tasks	Empty (reset before each run)
Initial RAM Load	RAM load before allocation	0 (reset before each run)
RAM Capacity	Memory capacity of the node	Random from 4000 to 8000 (MB)
Task setup parameters		
Data size	Data size of the task	Random from 1 to 1000 (MB)
Processing time	Minimum processing time	Random from 0.1ms to 10ms
Current CPU load	Task priority level	Random from 0.3 to 0.9
RAM Requirement	Required RAM capacit	Random from 1 to 200 (MB)
Network Latency	Data transmission delay	Random from 1ms to 10ms

4.1 Evaluation with Round-Robin, Greedy, and GA Methods

The purpose of this section is to evaluate the effectiveness of the proposed algorithm compared to methods with lower complexity, such as Round-Robin, Greedy, and GA. In this section, we compare the metrics BEST (best value), AVG (average value), and STD (standard deviation) across four scenarios: Basic, Increased Tasks, Increased Nodes, and Node Overload. The simulation results are presented through figures 5, with three evaluation scenarios. The collected, aggregated, and evaluated results show that the proposed method outperforms the two methods GA and Round-Robin. Specifically:

- *Regarding the BEST metric*: The proposed algorithm outperforms RR, Greedy, and GA in all scenarios, with improvements ranging from 5.57% to 14.40% compared to RR, from 1.04% to

14.37% compared to Greedy, and from 7.36% to 10.98% compared to GA. This indicates that the hybrid MA is more capable of finding an optimal schedule than the other methods, especially in complex scenarios like Increased Nodes.

- *Regarding the AVG metric*: The proposed algorithm maintains stable performance with average values lower by 7.58% to 20.27% compared to RR, 9.80% to 20.22% compared to Greedy, and 7.68% to 11.85% compared to GA. This result confirms the stability of the hybrid MA across multiple runs.
- *Regarding the STD metric*: The proposed algorithm exhibits significantly lower standard deviation, with improvements ranging from 8.88% to 53.62% compared to RR, from 13.10% to 53.56% compared to Greedy, and from 7.56% to 56.65% compared to GA. This demonstrates that the proposed algorithm is less volatile, providing higher reliability in finding optimal solutions.
- *Particularly in the Node Overload scenario*: The proposed algorithm maintains good performance, surpassing other algorithms, proving its ability to handle situations effectively when resources are limited.

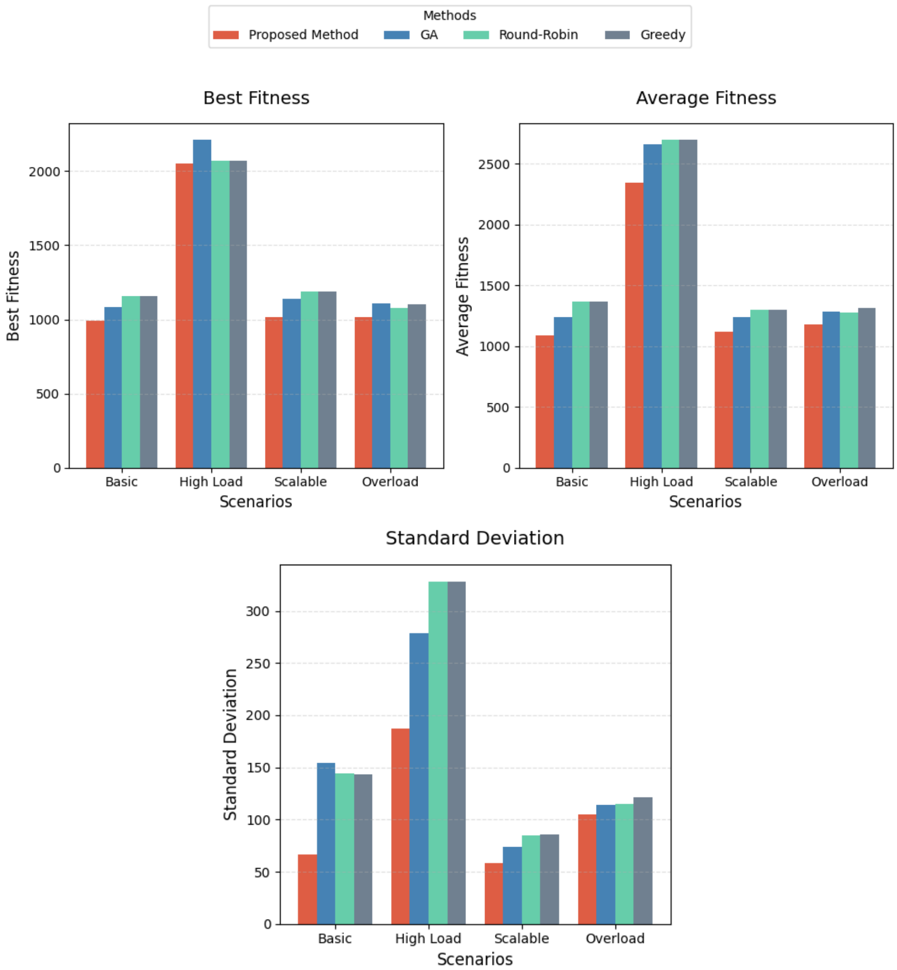


Figure 5. BEST, AVG, STD values of the proposed algorithm compared to simple algorithms GA, Round-Robin, Greedy

4.2 Comparison of Performance and Complexity with ACO

The ACO algorithm is a research method inspired by simulating the behavior of ant colonies in nature, aimed at solving complex real-world optimization problems. In this section, we will compare and evaluate the proposed algorithm with ACO based on the metrics BEST, AVG, and STD to assess performance, as well as runtime (Min Runtime, Average Runtime, Max Runtime) to evaluate complexity across the same four scenarios.

- *BEST Metric:* The hybrid MA (Proposed Algorithm) outperforms ACO in three scenarios (Basic, Increased Nodes, Node Overload) with improvements ranging from 4.17% to 12.11%. However, in the Increased Tasks scenario, ACO achieves a slightly better BEST value (-0.35%), though the difference is negligible.
- *AVG Metric:* The hybrid MA has a lower average value than ACO in all scenarios, with improvements ranging from 6.90% to 17.64%, indicating more stable performance.
- *STD Metric:* The hybrid MA exhibits a lower standard deviation than ACO in all scenarios, with improvements ranging from 12.96% to 51.58%, demonstrating superior stability.

After collecting and calculating the data, we observed that the fastest task completion time, the average completion time across all iterations, and the slowest completion time of the proposed algorithm all surpass those of ACO. Specifically:

- *Min Runtime:* The hybrid MA is faster than ACO in all scenarios, with improvements ranging from 22.11% to 52.82%.
- *Average Runtime:* The hybrid MA has a lower average runtime than ACO, with improvements ranging from 20.80% to 46.67%.
- *Max Runtime:* The hybrid MA is also faster than ACO, with improvements ranging from 14.89% to 30.90%.

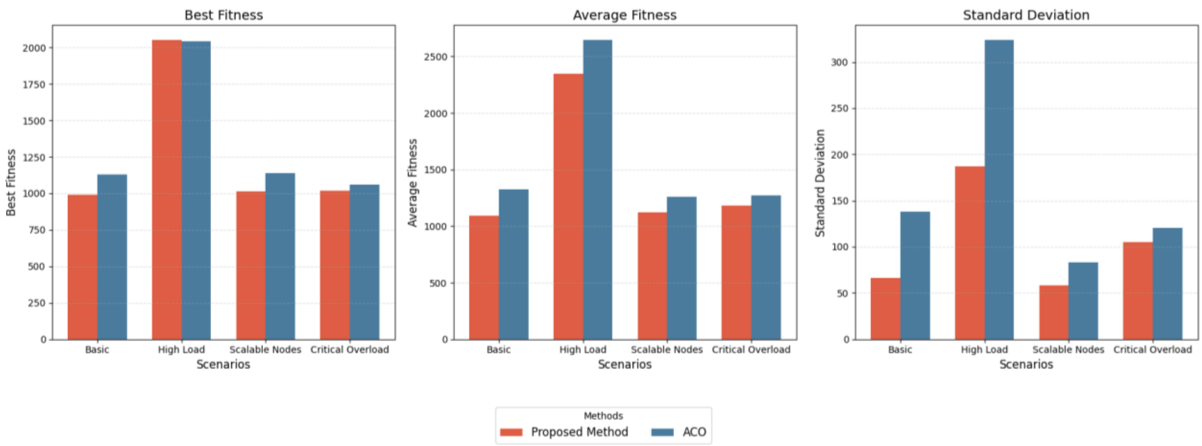


Figure 6. BEST, AVG, STD values of the proposed algorithm compared to metaheuristic algorithms ACO

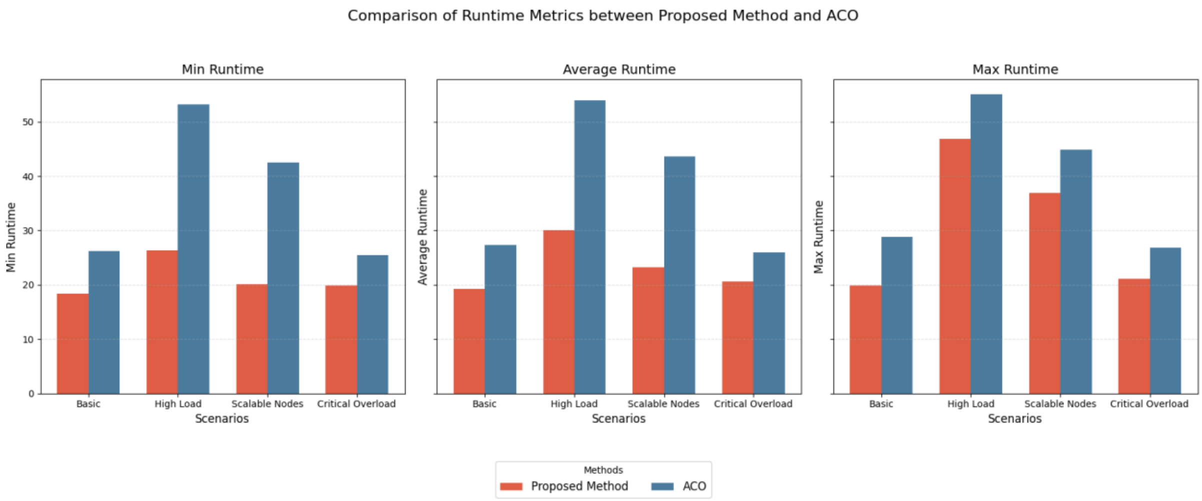


Figure 7. Minimum runtime, average runtime, and maximum runtime values of the proposed algorithm compared to the ACO

Simulation results show that the proposed algorithm outperforms the RR, Greedy, GA, and ACO algorithms in optimizing task allocation for Fog nodes. The proposed algorithm not only achieves better fitness values (lower BEST and AVG), higher stability (lower STD), but also has shorter runtime compared to ACO. Particularly in the Node Overload scenario, the proposed algorithm maintains good performance, demonstrating its ability to effectively handle resource-constrained situations. This proves that the proposed method in this paper optimizes task allocation for Fog nodes, and the restructuring approach can handle cases of overloaded Fog nodes, thereby emphasizing the feasibility of the method.

5. CONCLUSION

The proposed method in this paper combines techniques from the Memetic Algorithm and a task restructuring mechanism. Through simulation and result evaluation, it is evident that the method achieves superior performance compared to Greedy, Round-Robin, GA, and ACO in finding optimal schedules while maintaining stable performance in complex scenarios such as increased tasks or node overload. Specifically, the proposed algorithm consistently demonstrates its superiority over the two compared methods.

This improvement confirms that integrating evolutionary and heuristic methods brings significant benefits in optimizing project scheduling. This highlights the feasibility and effectiveness of combining evolutionary and heuristic approaches in optimizing task allocation, in this case, within a fog computing environment.

This study not only provides an effective solution for the scheduling problem but also opens up new development directions for hybrid algorithms in resource management and project planning. In the future, as end devices become more powerful and capable of performing deeper tasks, and as Fog nodes are equipped with greater computational and processing capabilities, integrating deep learning, reinforcement learning, or more advanced methods will yield even greater efficiency. Additionally, applying Federated Learning to the task allocation process in a fog computing environment is a promising direction.

REFERENCE

Volkov, A., Muthanna, A., Paramonov, A., Koucheryavy, A., & Elgendy, I. A. (2024). Optimized Data Transmission and Signal Processing for Telepresence Suits in Multiverse Interactions. Journal of Sensor and Actuator Networks, 13(6), 82. https://doi.org/10.3390/jsan13060082

Roman, R., Lopez, J., Mambo, M.: Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges. Elsevier Future Gener. Comput. Syst. **78**(2), 680–698 (2018).

Mahmood, Z.: Fog computing concepts, frameworks and technologies. Springer Link, Book, (2018). https://doi.org/10.1007/978-3-319-94890-4

H. Rafique, M. A. Shah, S. U. Islam, T. Maqsood, S. Khan and C. Maple, "A novel bio-inspired hybrid algorithm (NBIHA) for efficient resource management in fog computing", *IEEE Access*, vol. 7, pp. 115760-115773, 2019.

Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012). Fog computing and its role in the internet of things. *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing*, 13-16.

Shao, W.; Shao, Z.; Pi, D. A network memetic algorithm for energy and labor-aware distributed heterogeneous hybrid flow shop scheduling problem. *Swarm Evol. Comput.* **2022**, *75*, 101190.

Ren, Hualing, et al. "A memetic algorithm for cooperative complex task offloading in heterogeneous vehicular networks." *IEEE Transactions on Network Science and Engineering* 10.1 (2022): 189-204.

Fog Computing and the Internet of Things: Extend the Cloud To where the Things are, White Paper, 2016, Available online: http://www.cisco.com/c/dam/en_us/solutions/trends/iot/docs/computingoverview.pdf (accessed on 8 April 2018).

Tang, Bo, et al. "A hierarchical distributed fog computing architecture for big data analysis in smart cities." *Proceedings of the ASE BigData & SocialInformatics 2015*. 2015. 1-6.

Afsar, S. e. (2022). Multi-objective enhanced memetic algorithm for green job shop scheduling with uncertain times. Swarm and Evolutionary Computation, 68 (2022): 101016.

D. Johnson, C. Papadimitriou, and M. Yannakakis, “How easy is local search,” *Journal of Computer and System Sciences*, vol. 37, pp. 79–100, 1988.

Lin, Jih-Yiing, and Ying-Ping Chen, "Analysis on the collaboration between global search and local search in memetic computation," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 5, pp. 608-623, 2011.

Guimaraes, Frederico G., et al., "Local learning and search in memetic algorithms," *2006 IEEE International Conference on Evolutionary Computation*, IEEE, 2006.
Singh, S., & Chana, I. (2016). A survey on resource scheduling in cloud computing: Issues and challenges. *Journal of Grid Computing*, 14(2), 217-264.
Li, K., et al. (2018). Energy-efficient task scheduling in IoT systems: A survey. *IEEE Internet of Things Journal*, 5(5), 3871-3885.
Xu, L., et al. (2017). Cost-aware scheduling in cloud computing: A survey. *IEEE Transactions on Cloud Computing*, 5(3), 456-469.
Thang, D.V.; Volkov, A.; Muthanna, A.; Koucheryavy, A.; Ateya, A.A.; Jayakody, D.N.K. Future of Telepresence Services in the Evolving Fog Computing Environment: A Survey on Research and Use Cases. Sensors 2025, 25, 3488. https://doi.org/10.3390/s25113488
Artem, V.; Vadim, K.; Elgendy, I.A.; Muthanna, A.; Koucheryavy, A. DD-FoG: Intelligent Distributed Dynamic FoG Computing Framework. Future Internet 2022, 14, 13. https://doi.org/10.3390/fi14010013

Highly Efficient 3D Object Transmission System for HTC Services in 6G Networks

Daniil Svechnikov, Artem Volkov, Anastasia Marochkina, Ammar Muthanna, Andrey Kouhceryavy
The Bonch-Bruevich Saint-Petersburg State University of Telecommunications, St. Petersburg, 193232, Russian Federation
danil2001sv@mail.ru

ABSTRACT

In recent years, advancements in technology have brought forth a new frontier in visual communication. Holography is a technique that captures and reproduces three-dimensional (3D) images with an unprecedented level of realism and depth, has emerged as a groundbreaking method for conveying visual information. Unlike traditional images and videos, holography recreates scenes with full parallax, enabling viewers to perceive objects from various angles. The transmission of holographic images presents both exciting possibilities and unique challenges. To this end, this article conducts a comparative analysis of a previously developed application system for transmitting dynamic 3D human movements with a ready-made solution for transmitting 2D video streams in order to provide conference calling services. The network characteristics of the systems were collected and compared. The opportunities that programs currently provide and will provide in the future are examined.

KEYWORDS: Telepresence, Comparative Analysis, MVP, TrueConf, HTC, 6G, Teleconference, 3D images, Conferencing

1. INTRODUCTION

With the development of communication networks, users have more and more requests for the provision of various services. This paper will compare two applications from the field of telepresence services, which are still developing. Separately, we can note the direction of Holographic-Type Communication, the development of which is currently happening everywhere in the world [1].

During the pandemic, the need for video conferencing has greatly increased, examples of which have become firmly established in our daily lives: a work meeting, a lecture at the university, friendly communication. Conferencing technology allows for simultaneous communication between multiple users.

At the moment, applications for transmitting 2D images over networks for conferencing are very widespread. One of the popular solutions in this area today is TrueConf - a software platform for video conferencing. It helps employees communicate and work effectively from a distance, as well as easily conduct webinars and remote training.

Another direction of applications for remote work, education, etc. is the remote transmission of three-dimensional images, which makes it possible to see human movements in more detail from all angles, and also opens up new possibilities for immersive communications [2,3]. Transferring data about three-dimensional objects in its original form, for example in the format of streaming video from an array of cameras, requires significant bandwidth, so it becomes important to use technologies for converting data received about an object into other formats [3, 4].

This paper aims to investigate and compare the network traffic of several telepresence applications to test whether it is possible to provide a more complex conferencing service than is currently available without significant losses in network traffic. The following will be considered:

existing commercial solutions and a self-developed system that provides the opportunity to see a person not only in 2D, but also adds a third plane, allowing you to display a 3D model of a person in augmented reality, mixed reality or any holographic method of providing data.

2. EXPERIMENTS DESCRIPTION

During a superficial analysis of the network characteristics of other programs for providing conferencing using live video stream transmission, similar to TrueConf, it became clear that any of the advanced applications, such as Zoom, Google Meat, Yandex Telemost, have comparable results. This conclusion made it possible to compare the developed MVP system with a single application that is more convenient to use and test - TrueConf [5].

The research involved two applications:

· TrueConf - a unified communications software system from the company with the same name - TrueConf, using streaming video transmission to a flat surface in 2D to provide conferencing.

· MVP - an application system designed to synchronize control points on the Sender's body and an avatar that simulates movements on the Receivers' devices.

TrueConf, like any other software currently on the market, is a means of transmitting streaming video and audio - this is the minimum set of capabilities for providing conference calls. As you know, streaming video puts a lot of load on network equipment due to its weight. At the time of the transition of the whole world to remote communications, network equipment was not ready to provide the resource for transmitting such an amount of traffic. It was necessary to come up with a new way of transmitting data to provide conference calls. For this purpose, the MVP application system was developed.

MVP is an application system consisting of 3 parts (Sender, Server, Receiver). The Sender collects data from the Kinect camera about the control points of the human body (Atlas of body "Key Points"), groups them into a vector database and sends it to the Server. The Server constantly synchronizes network components on all connected devices at regular intervals; if synchronization of a component fails, a timer is started, after which the user will be disconnected as inactive. The network component on the Receiver's side receives updated data from the Server and transmits it to a 3D avatar that simulates the Sender's movements. Receiver client applications are implemented both on the Standalone system (Windows / Linux / MacOS) and for Android. Moreover, playback on any Standalone machine can be projected onto a holographic cube to achieve a three-dimensional effect (the speaker's avatar can be walked around from all sides), and the Android version uses ARCore to create augmented reality, which will allow the user to place the avatar at any convenient point in his room using your phone camera [6].

It is important to note that unlike TrueConf, the MVP application does not transmit video traffic, but uses data transmission instead. This feature allows us to move away from the usual concept of providing conferencing services and get closer to the full transfer of point clouds over communication networks to support holographic communication [7].

MVP is a platform that can be developed in various directions, from solving problems of reducing network traffic to provide telepresence services using 3D models, to developing a full-fledged holographic application for providing HTC. At the moment, this is a prototype that will allow conducting statistical experiments to test various hypotheses. Comparing the network characteristics of the developed application and a related application (similar application) will give us an understanding in which direction it is worth developing and whether it is worth continuing development at all [8].

Sixteen (16) experiments were carried out, each of which lasted 300 seconds and took place under the same conditions on the same network topology (Fig. 1): Eight (8) experiments for each application (Table 1).

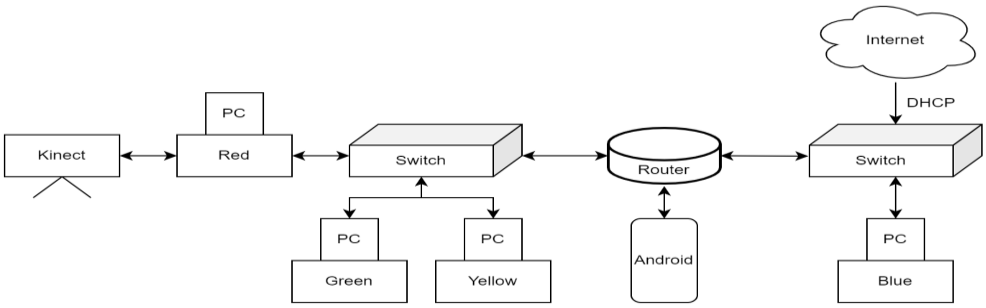


Figure 1. Network Architecture

Red, Green, Blue - different names of computers, corresponding to their color, which is later translated into graphics, Android - designation of a phone with the Android system (indicated in dark green). Idle – a type of experiment in which there is no person in the frame and the picture is stationary, Movement – experiments in which a person actively moves in the frame.

Table 1. Experiments

	Idle	**Movement**
1 Client	Yellow PC	Yellow PC
2 Clients	Yellow PC + Green PC	Yellow PC + Green PC
3 Clients	Yellow PC + Green PC + Blue PC	Yellow PC + Green PC + Blue PC
3 + 1 Clients	Yellow PC + Green PC + Blue PC + Android	Yellow PC + Green PC + Blue PC + Android

3. EXPERIMENTS DIFFERENCES

Before moving on to a direct comparison of network characteristics, it is important to note several points that affect the generated traffic, which we will need later.

The big difference between the “Idle” experiments comparing the two applications is caused by the way the applications generate network traffic. The TrueConf application converts the 2D image received from the camera into packets and sends it over the network every unit of time, as in the “Movement” experiment, MVP requires a small bandwidth just to maintain client connections to the server (0-8 Kbps), if this connection is lost, the server will automatically disconnect the user. This feature will be clearly visible in the presented graphs, and on some you may get the feeling that there is no traffic at all, but this is not the case, it is simply extremely small compared to the rest [9].

The experiment presents various devices with different hardware: “Yellow PC” and “Blue PC” interface - “Realtek Gaming GbE Family Controller”, “Green PC” interface - “Realtek PCIe 2.5GbE Family Controller”, the used phone is “Xiaomi Mi 8” with “Android 9”. The network components of the MVP application for smartphones and for PCs are completely identical.

Also, the devices are located in different places in the network topology (Fig. 1); in all experiments, the server and sending client were deployed on the “Red PC”, after which the traffic went to the “Aruba 2930F JL259A” switch device, to the “Green PC” and “ Yellow PC”, then comes the “Xiaomi Mi AIoT AX3600” router, it transmits data to Android; from the router, network traffic goes to the next “Catalyst 3750G” switch, from switch to the “Blue PC” end device.

4. COMPARISON OF NETWORK CHARACTERISTICS OF DIFFERENT SYSTEMS

The catalyst for developing a functionally new video conferencing application was the urgent need to reduce network video traffic, given the transition of a large number of people to remote work and education.

In the realities of the need to write an application “here and now” many auxiliary systems and libraries were used that lower final performance for the sake of faster application writing.

The developed system does not utilize all the capabilities of the physical layer of the OSI model, specifically the Ethernet frame length. In most cases (62%), only half of the Ethernet packet is used, which is approximately 700-750 bytes (Fig. 2). This underutilization of the packet means that, at any given time, we have the potential to transmit more useful information than we currently do (a path for transmitting more points), without compromising Quality of Service (QoS) and network characteristics. In contrast, a more optimized TrueConf application uses the full packet length of 1550 bytes in 54% of cases, and cannot transmit a greater amount of data in a single packet, i.e., it cannot increase the ratio of useful information to the total packet length with headers (Fig. 3).

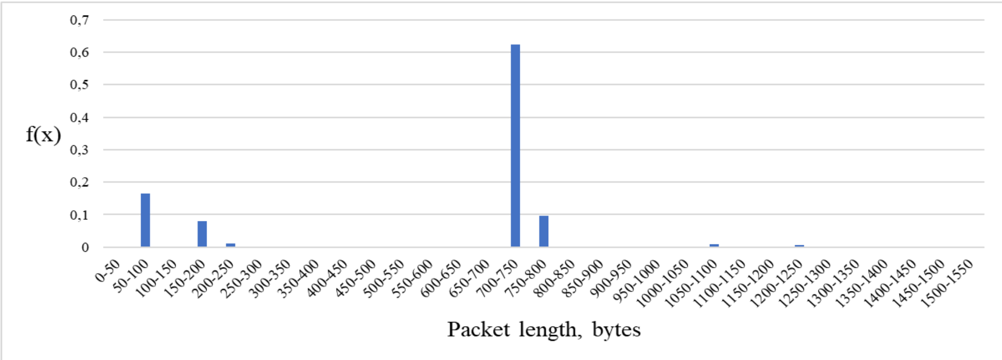


Figure 2. MVP, 3 + 1 Clients, Packet length

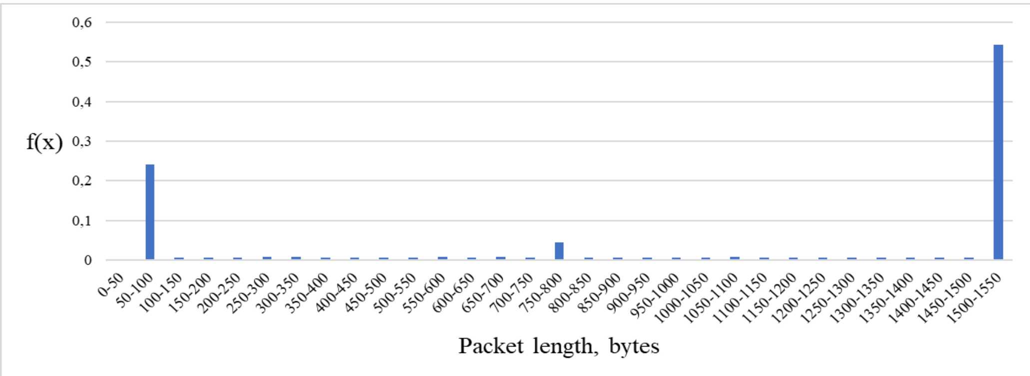


Figure 3. TrueConf, 3 + 1 Clients, Packet length

The described features can be observed in the comparative graph of the Average Packet Size in Bytes (Fig. 4). The average packet size values, as well as their distributions, have a small margin of error and are independent of the number of connected devices.

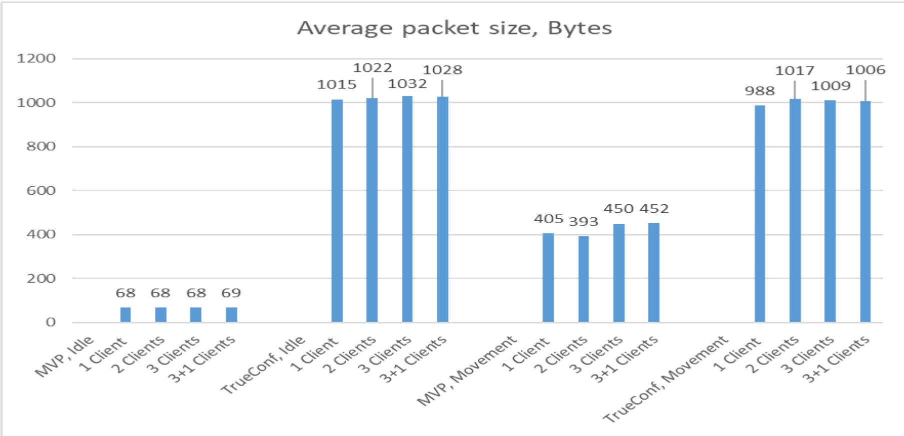


Figure 4. Average packet size, Bytes

The intensity refers to the number of packets transmitted per unit of time, in this case, per second (Fig. 5). The graph shows that the developed system imposes a lighter load on the network compared to TrueConf, and this difference is particularly noticeable when there is no person in the frame, as mentioned earlier.

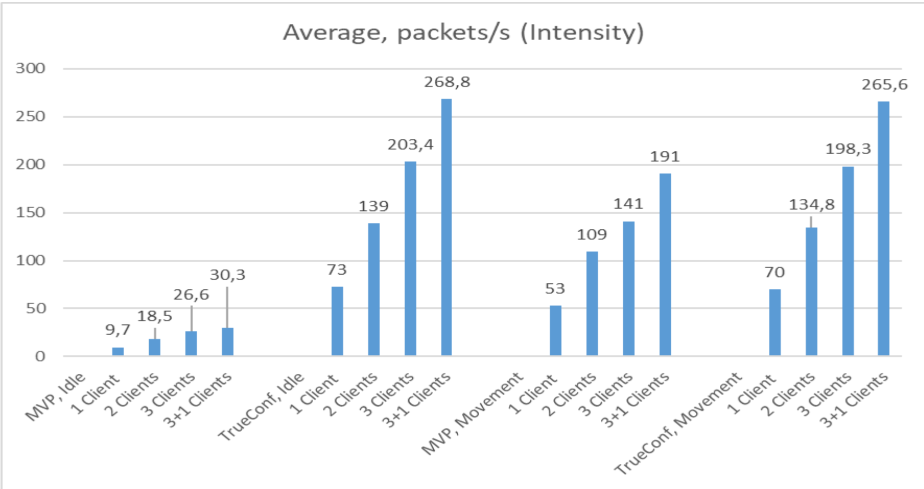


Figure 5. Average, packets/s (Intensity)

Data transfer rate, which measures the quantity of data transmitted within a specific timeframe, in this case, the number of bits sent per second, already exhibits a substantial advantage in terms of the volume of information transmitted (Fig. 6). This advantage is contingent on the fact that the developed MVP system conveys data about the 3D model, which can be processed and displayed to the user in various ways, in contrast to TrueConf, which only transmits 2D video.

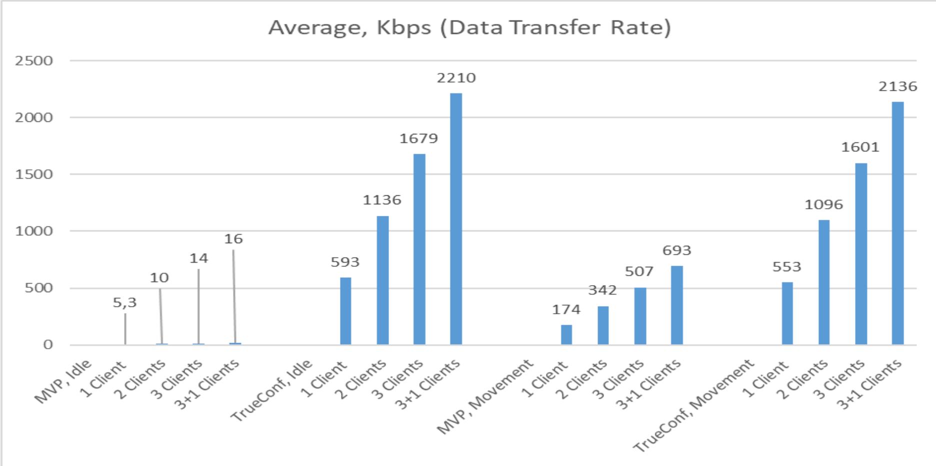


Figure 6. Average, bps (Data Transfer Rate)

The chart representing the transmitted data (Fig. 7) displays the volume of bytes transmitted throughout the entire five-minute application lifecycle, encompassing startup, connection, data transmission, and disconnection. This graph effectively illustrates the reduction in data transmitted over the network. When considering only the "Movement" experiments, the data reduction factor consistently exceeds 3. This implies that regardless of the number of connected devices, when a person is in the frame, the developed system transmits three times less information over the network than TrueConf. As for the "Idle" experiments, in these cases, the factor hovers around a value of 130, which is noteworthy but less significant, given the absence of the transmission of useful information.

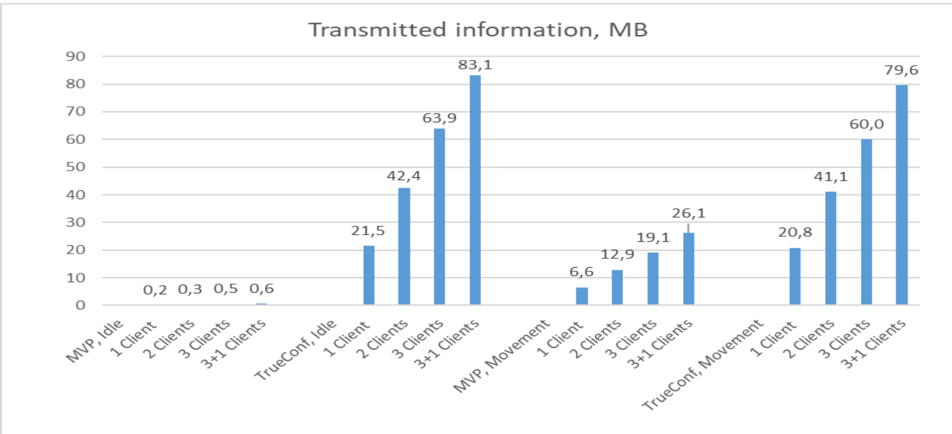


Figure 7. Transmitted information, Bytes

The chart (Fig. 7) illustrates that the volume of transmitted data increases linearly with the number of connected devices, both for the developed application and TrueConf. This information will be valuable for future network load forecasting when necessary.

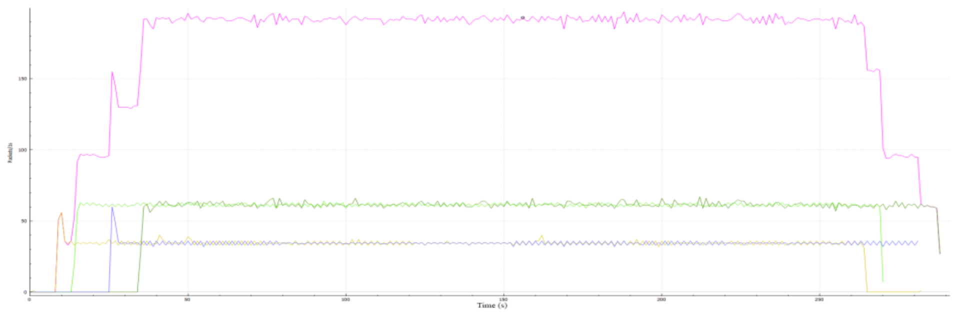


Figure 8. MVP, 3 + 1 Clients, Movement, Packets per second

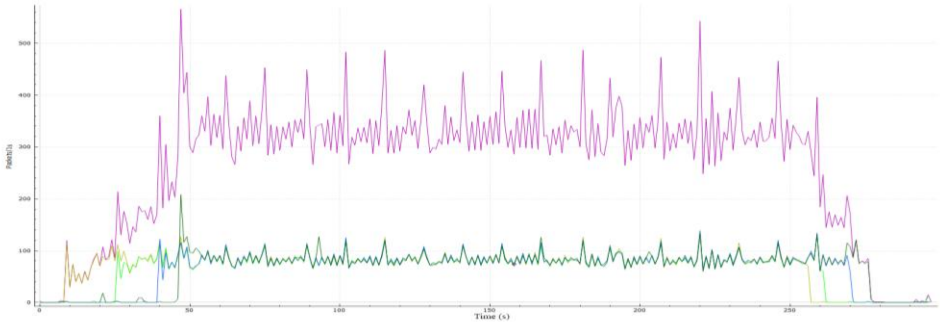


Figure 9. TrueConf, 3 + 1 Clients, Movement, Packets per second

The figures 8 "MVP, 3 + 1 Clients, Movement, Packets per second" and 9 "TrueConf, 3 + 1 Clients, Movement, Packets per second" display the graphs depicting the relationship between the number of received packets per second over time for the two most demanding experiments, one for each of the compared applications. It's immediately evident that the developed application uses the UDP protocol for data transmission, resulting in a "smooth" traffic pattern, while TrueConf uses TCP, as shown in the second graph, resulting in a "choppier" traffic pattern. Additionally, there's a notable difference in the scale of values between the two graphs, indicating that TrueConf utilizes the available bandwidth more aggressively.

The primary achievement of MVP at this stage is its low network resource consumption compared to applications providing 2D video communication services. The developed system not only outperforms in terms of network characteristics but also introduces revolutionary new capabilities for holographic communication services by transmitting control points of a person. There is potential for further development to enable the transmission of complete holograms or a cloud of points representing an entire individual [7].

5. CONCLUSION

In this paper, we have described the results of a comparative analysis of two conferencing applications: the self-developed MVP application and the commercial product TrueConf. The research has clearly shown that MVP surpasses TrueConf not only in offering more extensive services, including the transmission and display of a 3D model of a person for augmented reality and holography but also in terms of network characteristics.

The key findings of this comparative analysis are as follows: MVP's data transmission efficiency reduces the required amount of transmitted information by a factor of 3 compared to TrueConf. This would indicate a more optimized data transmission if both applications were transmitting video traffic. However, the MVP system was developed with the goal of reducing network load and replacing such traffic with a data stream, which can be particularly valuable in conditions of limited network bandwidth. The number of packets transmitted by MVP is also lower in this comparison, reducing the number of transmitted packets by 1.3 times. This parameter can be further improved by more efficient utilization of network resources and packet filling. This study touched upon the issue of QoS, but it was not fully explored. Further research should address the question of minimizing the risk of potential jitter during conferencing.

These results underscore the importance of developing and optimizing the next generation of conferencing applications, especially in cases where data transmission and functionality requirements are critical. MVP demonstrates the potential to provide a wider range of services and better network characteristics, which can be beneficial in various fields, including education, healthcare, and business. This development is ready to lead the way for the future of networks in 2030 and the world of HTC.

REFERENCES

Akyildiz, Ian F. and Hongzhi Guo. “Holographic-type communication: A new challenge for the next decade.” ITU Journal on Future and Evolving Technologies (2022)

Clemm A. et al. Toward truly immersive holographic-type communication: Challenges and solutions //IEEE Communications Magazine. – 2020. – T. 58. – №. 1. – C. 93-99.

Ateya, A. A., Muthanna, A., Gudkova, I., Vybornova, A., & Koucheryavy, A. (2017, July). Intelligent core network for Tactile Internet system. In Proceedings of the international conference on future networks and distributed systems (pp. 1-6).4. Cao C., Preda M., Zaharia T. 3D point cloud compression: A survey //The 24th International Conference on 3D Web Technology. – 2019. – C. 1-9.

Volkov, Artem, et al. "Optimized Data Transmission and Signal Processing for Telepresence Suits in Multiverse Interactions." Journal of Sensor and Actuator Networks 13.6 (2024): 82.

Kontogianni, E.; Anthopoulos, L. Towards a Standardized Metaverse Definition: Empirical Evidence from the ITU Metaverse Focus Group. IEEE Eng. Manag. Rev. 2024.

Cao C., Preda M., Zaharia T. 3D point cloud compression: A survey // The 24th International Conference on 3D Web Technology. 2019. pp. 1-9.

Svechnikov, D., Dunaytsev, R., Muthanna, A., & Aziz, A. (2023, December). Holographic Images Delivery Model Toward 6G Telepresence Services. In International Conference on Next Generation Wired/Wireless Networking (pp. 338-345). Cham: Springer Nature Switzerland.

Ateya, A. A., Abd El-Latif, A. A., Muthanna, A., Volkov, A., & Koucheryavy, A. (2025). Enabling Metaverse and Telepresence Services in 6G Networks. CRC Press.

FACULTY OF ENGINEERING

SLIIT UNI

THE KNOWLEDGE UNIVERSITY