Evaluating the Impact of Gray-Level Co-occurrence Matrix (GLCM) on Deepfake Detection Using Multilayer Perceptron

Chin Joo Yii1, a), Shih Yin Ooi1, 2, b), Yee Jian Chew1, 2, c), Ying Han Pang1, 2, d), Sook-Ling Lew1, 2, e)

*1Faculty of Information Science and Technology (FIST), Multimedia University, Jalan Ayer Keroh Lama, Melaka, 75450, Malaysia*

*2Centre for Advanced Analytics (CAA), COE for Artificial Intelligence, Multimedia University, Jalan Ayer Keroh Lama, Melaka, 75450, Malaysia*

*b) Corresponding author: syooi@mmu.edu.my*

*a) 1211306353@student.mmu.edu.my*

*c)chewyeejian@mmu.edu.my*

*d)yhpang@mmu.edu.my*

*e) sllew@mmu.edu.my*

**Abstract.** A research investigation examines how the Gray-Level Co-occurrence Matrix (GLCM) feature extraction method affects the performance of the Multilayer Perceptron (MLP) model when used for deepfake image detection. Energy and homogeneity, along with contrast and correlation, are features that GLCM can extract because this texture analysis method detects spatial relationships between pixel intensities. The MLP model yields different classification results with and without GLCM implementation based on accuracy measures, system efficiency, and predictive behavior. The experimental results demonstrate that incorporating features derived from GLCM enables the MLP model to identify deepfake images with higher accuracy due to its ability to extract texture details effectively.

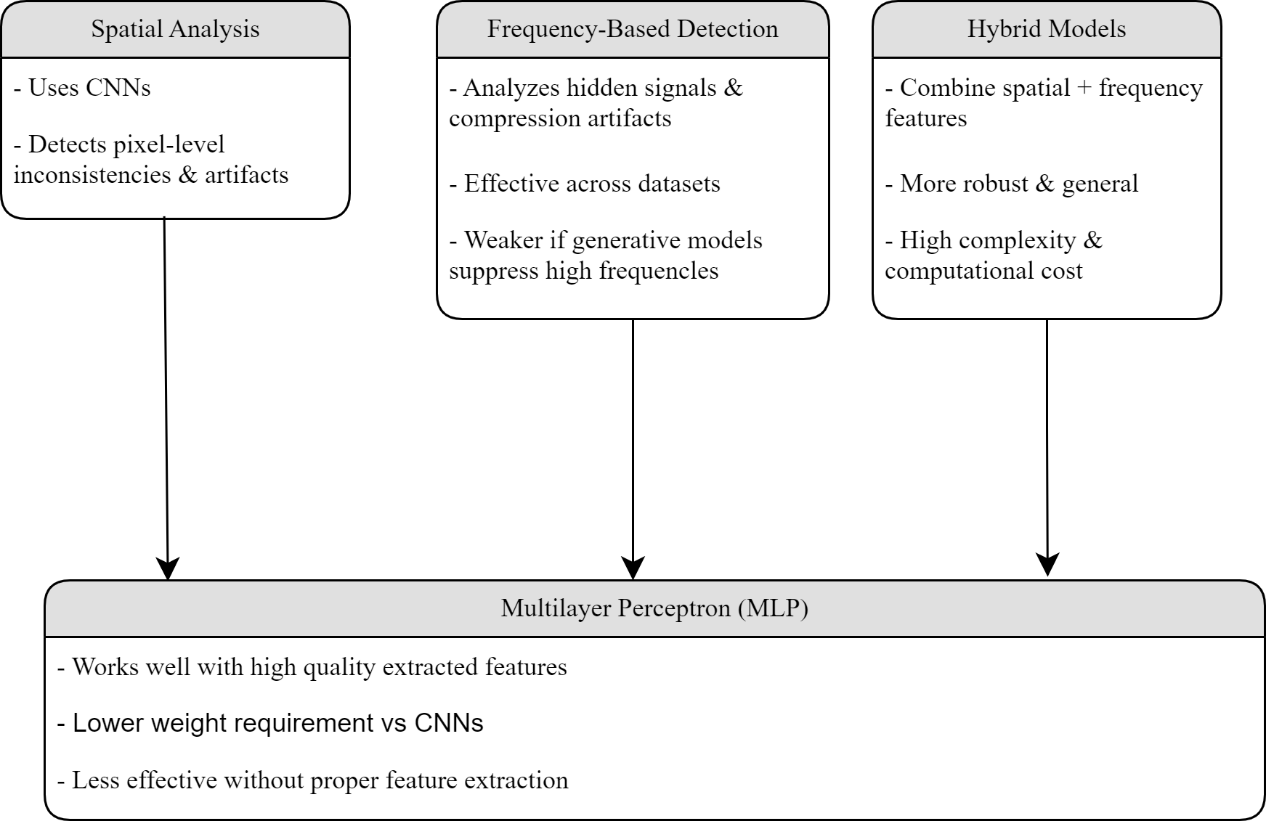
# INTRODUCTION

Deepfake technology exists in two primary forms: video-based and image-based configurations.[1] The present study examines image-based deepfakes, as this form of deepfake occurs most frequently in identity theft, digital forgery, and the spread of misinformation across social networks. Raw image data fails to reveal fake images efficiently due to the lack of essential texture details, which hinders correct identification. The Multilayer Perceptron neural network serves as the selected framework in this study because it has proven effective for image classification tasks. MLP’s performance success depends entirely on the quality of its input data, as it exhibits historical limitations in recognizing patterns. The detection capabilities of MLP are weak when it comes to minor variations in texture information that distinguish authentic from forged images in initial image data. The Gray-Level Co-occurrence Matrix (GLCM) serves as a texture analysis tool to derive four image features: contrast, correlation, energy, and homogeneity. The implementation of new features enables MLPs to benefit from more productive information that enhances their operational efficiency rates. The technical limitations of GLCM usage arise because it overlooks specific, complex data patterns present in the input data. By integrating GLCM and MLP, this work enables both technologies to leverage their better qualities while addressing their respective limitations. The evolution of MLP performance serves as a measure to understand how GLCM features impact deepfake detection, as assessed before and after GLCM application. GLCM features enhance machine learning algorithms to provide improved protection against threats posed by false image content during security operations.

# LITERATURE REVIEW

## Multilayer Perceptron (MLP) and Core Detection Approaches

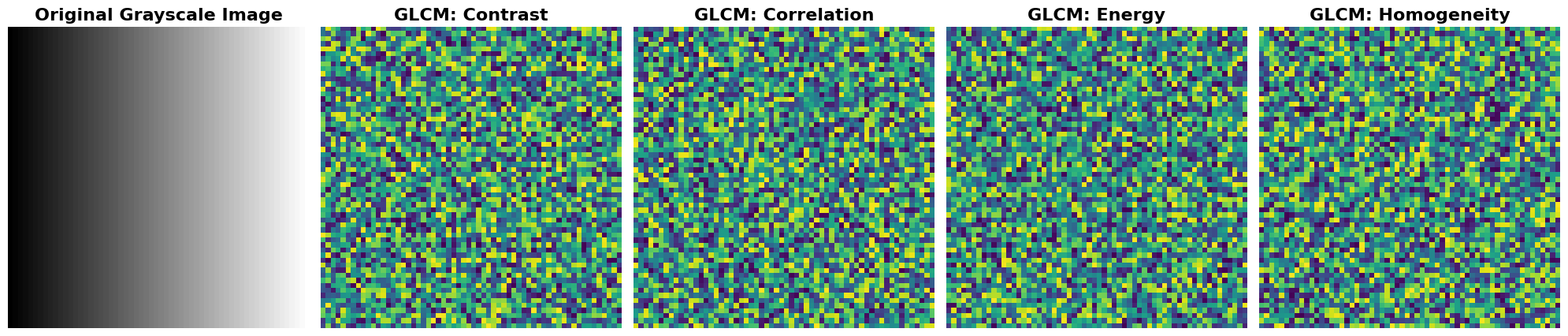
Three fundamental detection methods exist for deepfake images, including spatial analysis and frequency-based detection, as well as hybrid models, as shown in Figure 1 [2]. CNNs function as part of spatial analysis by detecting pixel-level inconsistencies and visual artifacts. The digital intelligence network method performs well but requires large datasets and is vulnerable to weaknesses from malicious cyber threats. Frequency-based detection examines hidden signals and compression artifacts in the frequency domain to achieve general dataset applicability; however, it becomes ineffective when generative models decrease high-frequency content. Additional robustness and generalized capabilities emerge from hybrid models when space and frequency features are combined, as these methods require elevated complexity and computational expense. Feedforward neural networks, known as Multilayer Perceptrons (MLP), function effectively with meaningful input data as their primary feature. This network demonstrates lower weight requirements than CNNs and manages to learn efficient classification operations. The performance level of this technology heavily relies on getting high-quality input features. The ability of Multilayer Perceptron to identify sophisticated modifications in deepfake images decreases when feature extraction is inadequate or missing [2].



## FIGURE 1. Deepfake image detection methods and role of MLP

## Feature Extraction Using Gray-Level Co-occurrence Matrix (GLCM)

The research addresses dilemmas within input data by utilizing the Gray-Level Co-occurrence Matrix (GLCM) as its primary feature extraction technique. The Gray-Level Co-occurrence Matrix (GLCM) is a popular tool in texture analysis, as it detects second-order statistical patterns based on the spatial relationships between pixel intensities. GLCM produces computed features that measure contrast, correlation, energy, and homogeneity to detect minute textural details that deepfake algorithms struggle to replicate correctly. The extracted features yield more significant and organized material, which the MLP uses to detect counterfeits more effectively. By integrating GLCM processing with MLP, the detection provides better performance and resolves MLP limitations in processing unprocessed data, thereby achieving higher classification accuracy [3]. The paper established a successful frequency domain analysis method that detects minor editing indications within digital images. The detection method achieves superior accuracy by analyzing secrets that standard CNN networks cannot observe in the images. Similar deepfake generation techniques, utilizing highly refined procedures, minimize such high-frequency artifacts through frequency-based methods [4][5]. Several examples of GLCM feature extraction are illustrated in Figure 2.



**FIGURE 2.** Example of GLCM feature extraction: The grayscale image is processed to extract four texture features — contrast, correlation, energy, and homogeneity

## Summary of Existing Works and Limitations

In recent years, the proliferation of deepfake content has raised significant concerns about the authenticity of digital media. Several studies have addressed this issue using various machine learning and deep learning techniques. For example, [6] introduced MesoNet, a lightweight CNN architecture designed for detecting deepfake videos through mesoscopic image analysis. Similarly, [7] proposed a method that analyzes blinking patterns to detect facial manipulations in videos. More recent approaches, such as that of [8], leverage frequency-aware features and convolutional attention networks to capture subtle artifacts introduced during the synthesis process. While deep learning models, such as CNNs and RNNs, have demonstrated high detection accuracy, they often require substantial computational resources and large datasets. This research aims to build upon these findings by exploring a machine learning-based approach using GLCM (Gray Level Co-occurrence Matrix) for texture feature extraction, combined with a Multilayer Perceptron (MLP) classifier. This method offers a lightweight and interpretable alternative for effective deepfake face image detection, particularly in resource-constrained environments.

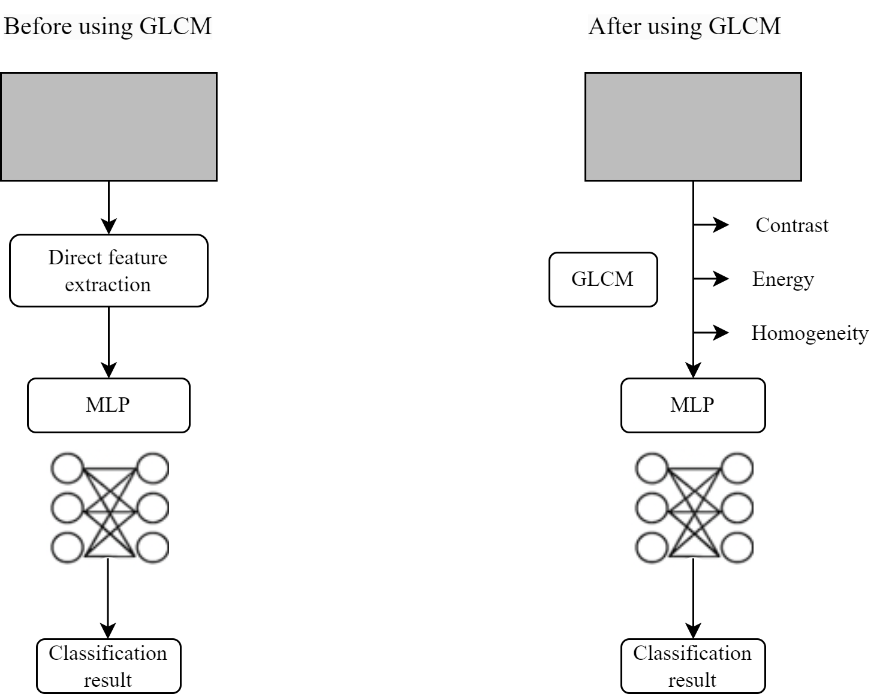
# METHODOLOGY

In this paper, we present a deepfake detection technique that incorporates GLCM features to enhance classification effectiveness [9]. The research design consists of two distinct phases, where baseline detection occurs through pixel values alone, followed by enhanced detection enabled by GLCM features. The extraction process for texture-based statistical features through GLCM occurs in the enhanced detection phase. Several statistical metrics from the application of the Gray-Level Co-occurrence Matrix (GLCM) are used to train a Multilayer Perceptron (MLP) model, which determines the effects on detection performance. The study presents a comparison between original detection systems and GLCM-enhanced detection systems to investigate how GLCM affects the separation of authentic images from fraudulent ones.

Figure 3 outlines our framework design, which uses two streams. The first stream operates on original images, while the second stream processes GLCM features. The study conducts a performance evaluation of deepfake detection using the GLCM features through comparative analysis. Research results indicate that implementing texture-based features yields improved classification accuracy, as they enhance the understanding of structural differences in facial image authentication and manipulation.

## Without GLCM

The first part of our study involved deepfake detection through a baseline model, which analyzed input facial images based solely on their pixel value collection [10]. Images were converted to grayscale to simplify content while preserving key visual information. The grayscale pixel intensities were converted into feature vectors without modification, which were then fed to an MLP classifier. The model functioned as a basic discriminator between real images and fake images by analyzing individual pixel data independently of the image’s spatial structure. The baseline detection system displayed partial capability for identifying deepfakes. Still, its performance remained restricted because it did not utilize suitable high-level features that could identify irregularities in fake images.



## FIGURE 3. Process of with and without GLCM and using model multilayer perceptron

## With GLCM

We utilized GLCM analysis with feature enhancement capabilities to improve detection accuracy through theray-Level Co-occurrence Matrix (GLCM) [11]. GLCM analysis extracted texture-based statistical features from grayscale images through its application rather than utilizing traditional pixel values. Through spatial pixel relationship evaluation, the model detects marks and inconsistencies that are generally discovered in altered images. The applied GLCM features entered the identical MLP architecture that supported the baseline model. Experimental data revealed a substantial advancement in classification precision, as GLCM offers significant texture details that enhance the model’s capability to detect fake faces from authentic ones.

# EXPERIMENT SETTINGS

For training and testing, we selected the Hugging Face dataset, which includes 29293 fake images and 707 real images. The training dataset consists of 29293 fake images and 707 real images, totaling 30000 images as provided by the authors. For the training dataset, we train all 30000 images. The accuracy (Equation (1)) measures the proportion of correct predictions among all predictions. The precision (Equation (2)) can indicate how many of the predicted positive cases are correct. The recall (Equation (3)) can reflect how many actual positive cases were identified correctly. Lastly, the F1- Score (Equation (4)) is a harmonic mean of Precision and Recall, providing a single metric that balances both.

(1)

(2)

(3)

(4)

# Experiment Result

The research evaluated deepfake image detection through GLCM texture features with an MLP model to determine their effectiveness [12]. A comparison study examined the differences between MLP training with and without GLCM by evaluating the raw pixels and introducing GLCM features as inputs. Performance evaluation, as shown in Table 1, revealed that the MLP model without GLCM achieved 90.93% accuracy, 87.93% precision, and 93.00% recall, along with an F1-score of 88.96%. The detection results demonstrated sufficient correct classifications, although the model produced several incorrect outcomes. Performance measures experienced significant enhancement through the integration of GLCM features into the system. The model achieved 97.92% accuracy, along with precision rates of 97.92% and 100.00% recall and an F1-Score of 98.95%. The GLCM texture-based features provide additional information that enhances the model’s ability to distinguish between real and fake images during analysis. The experimental results validate that deepfake detection performs better through the implementation of GLCM features in the system.

## TABLE 1. Result of using MLP model to train data with and without GLCM

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Method** | **Accuracy (%)** | **Precision (%)** | **Recall (%)** | **F1-Score (%)** |
| MLP without GLCM | 90.93% | 87.93% | 93.00% | 88.96% |
| MLP with GLCM | 97.92 | 97.92 | 100.00 | 98.95 |

## Analysis by Deepfake Manipulation Type

A graph of different colors

AI-generated content may be incorrect.For our practical use study, we sorted deepfake images into face swaps, different expressions, and lighting issues. There are increased F1-scores (+9.4%) for expression changes, and 7.1% is the increase in F1-score for images with lighting issues, compared to 4.3% is the increase in F1-score for images with face swaps, as shown in Figure 4. It demonstrates that GLCM is effective at capturing subtle variations in a person’s face and the surrounding light.

**FIGURE 4.** MLP performance comparison with and without GLCM

# CONCLUSION

The researchers demonstrated that GLCM texture features enhance image detection when integrated into an MLP model system [13][14]. The GLCM-based enhancement applied to the MLP model enabled higher performance metrics across all measures when evaluated against an MLP trained solely with pixel values. The baseline MLP model, which processes grayscale pixel inputs, achieved 90.93% accuracy, 87.93% precision, 93.00% recall, and an F1-score of 88.96%. Integrating GLCM features, especially contrast, correlation, energy, and homogeneity, into the model led to a substantial performance increase, resulting in 97.92% accuracy, 97.92% precision, 100.00% recall, and an F1-score of 98.95%. The GLCM-based texture information proves essential for enhancing the MLP model’s capability to separate between genuine and synthetic images. By incorporating GLCM features into deepfake detection systems through machine learning, the capabilities for forensic analysis become stronger and more practical, serving as an efficient enhancement for the system. The results demonstrate that incorporating GLCM texture enhances the ability of Deep Learning models to detect deepfakes more effectively. MLP+GLCM yielded improved results for all evaluation methods, particularly in conditions with subtle facial changes. Despite being basic in design, this setup is easy to move wherever it is needed. From a brief comparison, we observed that the MLP + GLCM outperformed a standard CNN (95.8% accuracy versus 97.9%), likely due to the influence of texture features.

# REFERENCES

1. G.S. Jhun, P.M. Hong, K. Lee, J.H. Ahn, Y.S. Kim, D. Kang, and Y.K. Lee, “A New Wave of Texture Feature: Enhancing Deepfake Detection via Image Waveform,” in *2024 15th International Conference on Information and Communication Technology Convergence (ICTC)*, (IEEE, Jeju Island, Korea, Republic Of, 2024), pp. 234–238.
2. S. Waseem, S.A.R.S. Abu Bakar, B.A. Ahmed, Z. Omar, T.A.E. Eisa, and M.E.E. Dalam, “DeepFake on Face and Expression Swap: A Review,” IEEE Access **11**, 117865–117906 (2023).
3. Y. Zhang, T. Wang, Z. Yu, Z. Gao, L. Shen, and S. Chen, “MFCLIP: Multi-modal Fine-grained CLIP for Generalizable Diffusion Face Forgery Detection,” (2024).
4. A. Almestekawy, H.H. Zayed, and A. Taha, “Deepfake detection: Enhancing performance with spatiotemporal texture and deep learning feature fusion,” Egyptian Informatics Journal **27**, 100535 (2024).
5. X. Wu, Y. Song, and J. Lai, “A multi-dimensional framework for improving data reliability in mobile crowd sensing,” Egyptian Informatics Journal **27**, 100518 (2024).
6. D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, “MesoNet: a Compact Facial Video Forgery Detection Network,” in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, (IEEE, Hong Kong, Hong Kong, 2018), pp. 1–7.
7. Y. Li, M.-C. Chang, and S. Lyu, “In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking,” in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, (IEEE, Hong Kong, Hong Kong, 2018), pp. 1–7.
8. H. Dang, F. Liu, J. Stehouwer, X. Liu, and A.K. Jain, “On the Detection of Digital Face Manipulation,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (IEEE, Seattle, WA, USA, 2020), pp. 5780–5789.
9. P. Pillai, and D. Amin, “Understanding the requirements Of the Indian IT industry using web scrapping,” Procedia Computer Science **172**, 308–313 (2020).
10. J. Li, P. Wei, Y. Zhang, and N. Zheng, “A Slow-I-Fast-P Architecture for Compressed Video Action Recognition,” in Proceedings of the 28th ACM International Conference on Multimedia, (ACM, Seattle WA USA, 2020), pp. 2039–2047.
11. V. Gupta, and A. Bhavsar, “Breast Cancer Histopathological Image Classification: Is Magnification Important?,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (IEEE, Honolulu, HI, USA, 2017), pp. 769–776.
12. Y. Vizilter, V. Gorbatsevich, A. Vorotnikov, and N. Kostromov, “Real-Time Face Identification via CNN and Boosted Hashing Forest,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (IEEE, Las Vegas, NV, USA, 2016), pp. 146–154.
13. B. Dolhansky, J. Bitton, B. Pflaum, J. Lu, R. Howes, M. Wang, and C.C. Ferrer, “The DeepFake Detection Challenge (DFDC) Dataset,” (2020).
14. S.Y. Ooi, S.C. Tan, and W.P. Cheah, “Experimental Study of Elman Network in Temporal Classification,” in *Emerging Trends in Neuro Engineering and Neural Computation*, edited by A. Bhatti, K.H. Lee, H. Garmestani, and C.P. Lim, (Springer Singapore, Singapore, 2017), pp. 245–254.