AI-Generated Sneaker Detection: Leveraging GANs and Convolutional Neural Networks for Image Classification

Alidor M. Mbayandjambe[1]; Grevi B. Nkwimi[2]; Darren Kevin T. Nguemdjom[3]; Fiston Oshasha[4]; Célestin Muluba[5]; Xavier F. Kutuka[6]

[1,4,5]University of Kinshasa, Faculty of Sciences and Technology, Kinshasa, DR. Congo
[1,2]University of Kinshasa, Faculty of Economic and Management Sciences, Kinshasa, DR. Congo
[1,3]International School, Vietnam National University, Hanoi, Vietnam
[6]Department of Mathematics and Computer Science, Institut Supérieur Pédagogique de Kikwit, Kikwit, DR. Congo

Publication Date: 2025/05/13

Abstract: The increasing prevalence of AI-generated content presents unique challenges in the field of computer vision, especially when distinguishing between real and synthetic images. This study explores the detection of AI-generated sneakers, specifically from popular brands such as Nike, Adidas, and Converse, using Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs). The dataset for this project is a mix of real sneaker images sourced from Google Images and AI-generated images produced by the MidJourney AI platform. To enrich the dataset and enhance model training, synthetic images are generated through a GAN, providing a diverse range of examples. The primary objective is to train a robust detection model capable of distinguishing between real and AI-generated sneaker images by leveraging subtle visual differences. This research demonstrates the effectiveness of GANs in augmenting datasets for machine learning applications, while also testing the resilience of CNNs in distinguishing high-quality AI-generated images from authentic ones. The dataset, standardized to 240x240 pixel resolution, offers a comprehensive foundation for developing advanced image classification models aimed at tackling the growing challenge of AI-generated content detection.

Keywords: *AI-Generated Images, Sneaker Detection, Generative Adversarial Networks, Convolutional Neural Networks, Dataset Augmentation, Machine Learning, Computer Vision.*

How to Cite: Alidor M. Mbayandjambe; Grevi B. Nkwimi; Darren Kevin T. Nguemdjom; Fiston Oshasha; Célestin Muluba; Xavier F. Kutuka. (2025). AI-Generated Sneaker Detection: Leveraging GANs and Convolutional Neural Networks for Image Classification. *International Journal of Innovative Science and Research Technology*, 10 (4), 3413-3422. https://doi.org/10.38124/ijisrt/25apr1963

I. INTRODUCTION

With the rapid evolution of artificial intelligence (AI), images generated by AI models, such as Generative Adversarial Networks (GANs), are becoming increasingly realistic. However, this advancement presents a major challenge in areas such as counterfeit product detection, image source verification, and fraud detection. GANs, in particular, have demonstrated their ability to produce images that are nearly indistinguishable from those generated by humans, creating new difficulties for security systems and image verification applications Goodfellow et al., (2014). In this context, the detection of AI-generated sneaker images is an especially interesting problem, as these images may be used in marketing campaigns or in the sale of counterfeit products.

Research on detecting AI-generated images has gained traction in recent years, particularly with the emergence of models like DeepFake and StyleGAN, which have highlighted the challenges of identifying artifacts left by the generators Karras et al., (2019). Numerous approaches have been explored, primarily focused on using Convolutional Neural Networks (CNNs) to distinguish real images from those generated Zhang et al., (2020), Mishra et al., (2024). However, most of these works have concentrated on domains such as human faces or AI-generated art, while few studies have been conducted on specific objects like sneakers.

This project focuses on developing a model capable of distinguishing real sneaker images from those generated by AI, using CNNs to extract visual features and GANs to enhance the dataset by generating synthetic sneaker images García-Aguirre, et al., (2024). This approach is motivated by

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

the idea that augmenting the dataset with generated images can improve the robustness of the detection model, which is crucial when dealing with increasingly realistic GAN-generated images Hussain, et al**.,** (2025).

Several challenges must be addressed in this area, including the difficulty of creating diverse datasets that cover variations in sneaker styles and designs while incorporating realistic synthetic images. GANs have been used in this project to generate sneaker images to complement the real data and test the model's ability to detect subtle visual differences between generated and real images Elgammal et al., (2017).

A. Our Main Contributions Are As Follows:

- Use of GANs to enrich the dataset by generating realistic sneaker images that improve the diversity of training examples.
- Development of a high-performance CNN model for detecting subtle differences between real and AI-generated images, using advanced image preprocessing techniques and supervised learning.
- Evaluation of the model using standardized metrics such as accuracy, precision, recall, and F1-score to analyze the model's performance on a test set.

This paper is structured as follows: Section 2 presents a literature review on previous work in AI-generated image detection and associated techniques ; Section 3 outlines the methodology and techniques used in this project; Section 4 details the models, architectures, and evaluation metrics; Section 5 presents the results and discussion ; and Section 6 concludes with a summary and future research directions.

II. LITTERATURE REVIEW

The detection of AI-generated images has become an increasingly explored topic, particularly with the rise of Generative Adversarial Networks (GANs). GANs are composed of two neural networks: the generator, which creates synthetic images from random noise, and the discriminator, which attempts to distinguish real images from the generated ones. The capability of GANs to produce highly realistic images has driven research into their detection Goodfellow et al., (2014) . GANs have been widely applied in various domains such as face generation Karras et al., (2019), art creation Elgammal et al., (2017), and even video synthesis (Yi et al., 2020). However, relatively little research has focused on specific commercial objects, such as sneakers, and the methods to distinguish these from their real counterparts.

In the field of AI-generated image detection, much of the early work has centered around identifying subtle visual artifacts between real and synthetic images. Zhang et al., (2020) proposed a method that leverages CNNs to detect these artifacts and distinguish GAN-generated images from real ones by extracting domain-specific features . Similarly, Marra et al., (2020) focused on improving detection by analyzing inconsistencies in lighting and texture artifacts inherent in AI-generated images, which are often indicative of synthetic content . These efforts were primarily concerned with faces and art, where the detection models had to account for subtle differences in visual features like pixel-level inconsistencies, shading discrepancies, and unnatural geometries.

More recent work, such as that by Wang et al., (2021), has investigated the potential for cross-domain GAN-generated image detection, extending this to product images like shoes and clothing, where the challenges include variations in style and visual elements across datasets. Their work used a combination of CNNs and deep feature learning techniques to improve generalizability across domains Shi, & al., (2025), something critical when working with diverse products like sneakers.

Additionally, the rise of text-to-image models like DALL-E Ramesh et al., (2021) has introduced new complexities in image authenticity, making it increasingly difficult to distinguish between AI-generated and real images without more advanced and specific detection models. These models leverage vast amounts of data to produce images that appear increasingly indistinguishable from real-world objects, requiring new techniques to detect their synthetic origin.

In particular, this study contributes to the existing literature by exploring the use of GANs not only for generating synthetic images but also for enhancing the dataset used to train sneaker detection models. Our approach focuses on overcoming the challenges of creating diverse datasets for detecting AI-generated images, which often involve complex variations in style and design. By leveraging synthetic images, we can test the robustness of detection models against more sophisticated AI-generated content.

Our work also builds on the idea of transfer learning and feature extraction, where pre-trained models are adapted for specific tasks such as object detection in commercial products. The idea is to enhance detection accuracy and robustness in identifying synthetic images in a niche domain such as sneakers, where there is a lack of widely available datasets specifically for training image classifiers.

Thus, the contributions of this study are situated at the intersection of GANs for image generation, CNNs for feature extraction, and domain-specific applications in image detection. This interdisciplinary approach provides valuable insights for the advancement of AI-generated image detection in specialized fields like sneaker image classification.

III. METHODOLOGY

The primary objective of this study is to develop a robust model capable of detecting AI-generated images by leveraging Convolutional Neural Networks (CNNs) for feature extraction, alongside Generative Adversarial Networks (GANs) to generate synthetic images, thereby augmenting the diversity of the dataset. The adopted methodology comprises several key steps, which are outlined below:

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

A. Data Collection and Preparation

For the purposes of this study, a diverse dataset was curated, which was divided into two primary categories:

➤ *Real Images:*

These images were collected from Google Images by querying sneakers from specific and well-known brands such as Nike, Adidas, and Converse. These authentic images serve as the "ground truth" for the classification task and represent various real-world sneaker designs, ensuring a comprehensive representation of real sneakers in the datase.

➤ *Ai-Generated Images:*

These images were created using the MidJourney platform, a state-of-the-art AI-driven image generator, in addition to a GAN model. The synthetic images generated by these platforms mimic real-world sneaker designs but may contain subtle visual differences that can be identified by the detection model. The inclusion of these AI-generated images introduces the necessary complexity for the detection system to learn to distinguish between real and synthetic content.

To ensure consistency across the dataset, all images, both real and generated, were resized to a standard resolution of 240x240 pixels. This uniformity is essential for ensuring that the model receives input of consistent dimensions, minimizing any potential issues that could arise from varying image sizes during training and evaluation.

B. Generation of Synthetic Images via GANs

The core of this methodology involves the use of Generative Adversarial Networks (GANs) to generate synthetic sneaker images, which augment the dataset and provide additional training examples for the model. The GAN consists of two primary components :

➤ *Generator:*

This component generates synthetic images by learning from random noise. The generator is trained to create images that resemble real sneaker images as closely as possible.

➤ *Discriminator*:

The discriminator's role is to distinguish between real and synthetic images. It evaluates the authenticity of the images produced by the generator, providing feedback that helps the generator improve its image creation over time.

By iterating through this adversarial process, the GAN generates images that progressively improve in quality, making them more difficult to differentiate from real sneaker images. These synthetic images are integrated into the dataset, allowing the model to better generalize to diverse sneaker designs and improving its robustness against subtle differences between real and synthetic images.

Mathematically, the objective of GAN training is to optimize a minimax game between the generator *G* and the discriminator *D*, formulated as:

$$\min_{G}\max_{D} E_{x\sim p_{\text{data}}(x)}[\log D(x)] + E_{z\sim p_{z}(z)}\left[\log\left(1 - D\left(G(z)\right)\right)\right]$$

where $D(x)$ is the probability that xx comes from the real data distribution, and $G(z)$ is the image generated by the generator from random noise z. The generator aims to produce images that are indistinguishable from real data, while the discriminator attempts to classify them as real or fake Goodfellow et al., (2014); Karras et al., (2019).

C. Detection Model: Convolutional Neural Networks (CNNs)

For the image classification task, a Convolutional Neural Network (CNN) was employed to classify images into two categories: real and AI-generated. CNNs are particularly well-suited for image classification tasks because of their ability to automatically learn hierarchical features from raw image data, which makes them highly effective for visual tasks such as detecting patterns, textures, and edges LeCun et al., (2015).

The architecture of the CNN used in this study consists of multiple convolutional layers followed by fully connected layers. The convolutional layers are responsible for feature extraction, detecting local patterns such as edges and textures, while the fully connected layers perform the classification based on the high-level features extracted by the convolutional layers Solomon, V. (2025).

Each convolutional layer in the network applies a filter to the input image, performing operations such as edge detection and texture recognition. The extracted features are then passed through pooling layers, which reduce the dimensionality of the data, before being classified by the fully connected layers.

➤ *The Convolutional Operation Applied at each layer of the CNN is given by:*

$$S(i,j) = (I * K)(i,j) = \sum_{m}\sum_{n} I(m,n)K(i-m, j-n)$$

where I is the input image, K is the convolution kernel, and SS is the output feature map. This operation involves sliding the kernel over the image and computing the dot product at each position.

D. Model Evaluation

To evaluate the performance of the detection model, a set of standard metrics were used, including accuracy, precision, recall, and F1-score. These metrics are critical for assessing how well the model distinguishes between real and AI-generated images, and they provide a comprehensive understanding of the model's performance:

➤ *Accuracy:*

Measures the proportion of correctly classified images out of the total number of images. This metric gives an overall sense of the model's effectiveness.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives Chollet, (2017).

➢ *Precision:*

Evaluates the proportion of true positive predictions relative to the total number of positive predictions made by the model. It is especially useful when the cost of false positives is high.

$$Precision = \frac{TP}{TP + FP}$$

➢ *Recall:*

Measures the proportion of true positives correctly identified by the model out of all actual positives. It is useful for understanding how well the model identifies all instances of the target class.

$$Recall = \frac{TP}{TP + FN}$$

➢ *F1-score:*

The harmonic mean of precision and recall, providing a balanced measure of the model's performance, especially when dealing with imbalanced datasets.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

These evaluation metrics were computed based on the model's performance on a separate test set, which consists of images that were not part of the training process. This ensures that the evaluation reflects how well the model generalizes to unseen data.

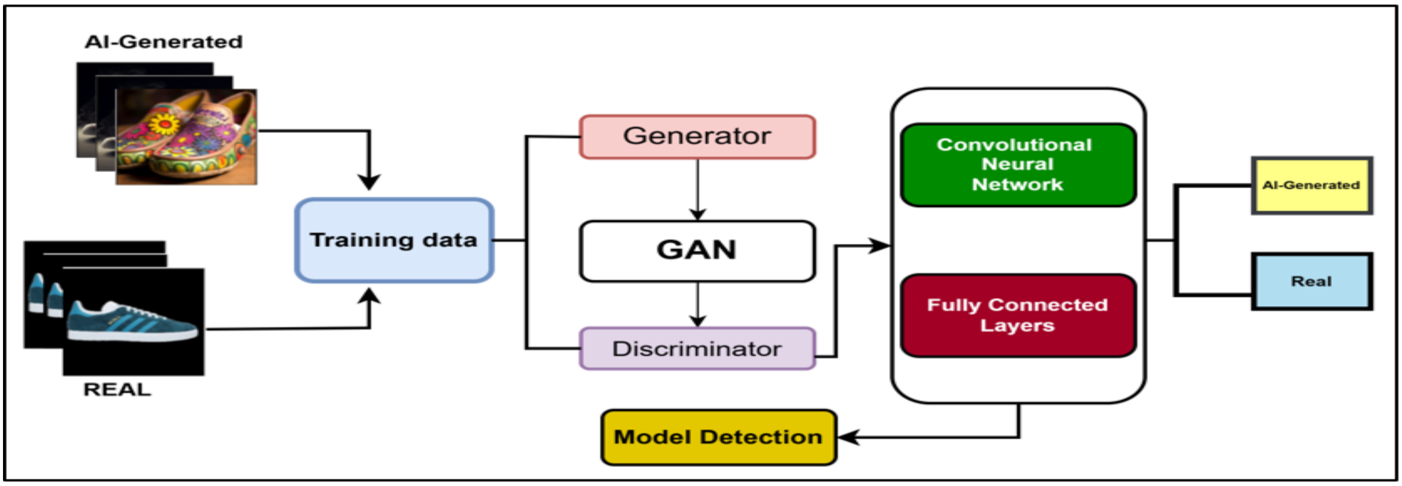

Fig 1 Hybrid GAN-CNN Architecture for Image-Based Detection.

As shown in Figure 1, the proposed methodology integrates real and generated image samples to train a Generative Adversarial Network (GAN) composed of a *Generator* and a *Discriminator*. The output of the GAN is passed to a Convolutional Neural Network (CNN) that includes convolutional and fully connected layers to perform classification. The CNN discriminates between *real* and *generated* images and simultaneously enables model detection. This hybrid GAN-CNN architecture is designed to boost detection performance by enriching the training data with diverse synthetic examples.

IV. EXPERIMENTAL SETUP

This section details the models, datasets, evaluation metrics, and training procedures used to distinguish between real and AI-generated sneaker images.

A. *Models and Architectures*

Our approach leverages two deep learning models: a Convolutional Neural Network (CNN) for classification and a Generative Adversarial Network (GAN) for synthetic image generation and detection.

➢ *CNNs: Overview*

Convolutional Neural Networks (CNNs), originally introduced by LeCun et al. (1998), are widely recognized deep learning models that utilize convolutional layers to learn hierarchical visual representations from images. These architectures are particularly effective in a variety of image analysis tasks, including classification, localization, and anomaly detection.

In this work, CNNs are employed for distinguishing real images from those generated by AI models, particularly GANs. These architectures have been proven effective in related works where CNNs have shown strong performance in identifying fake images through texture, color, and frequency cues Bi et al., (2023); Liu et al., (2023). To enhance their robustness, we incorporate synthetic images generated using GAN frameworks into the training dataset, enabling the models to learn subtle artifacts often present in generated content. This approach aligns with recent research trends that emphasize hybrid and ensemble learning strategies for improving detection accuracy Frid-Adar et al., (2018); Naim et al., (2023).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

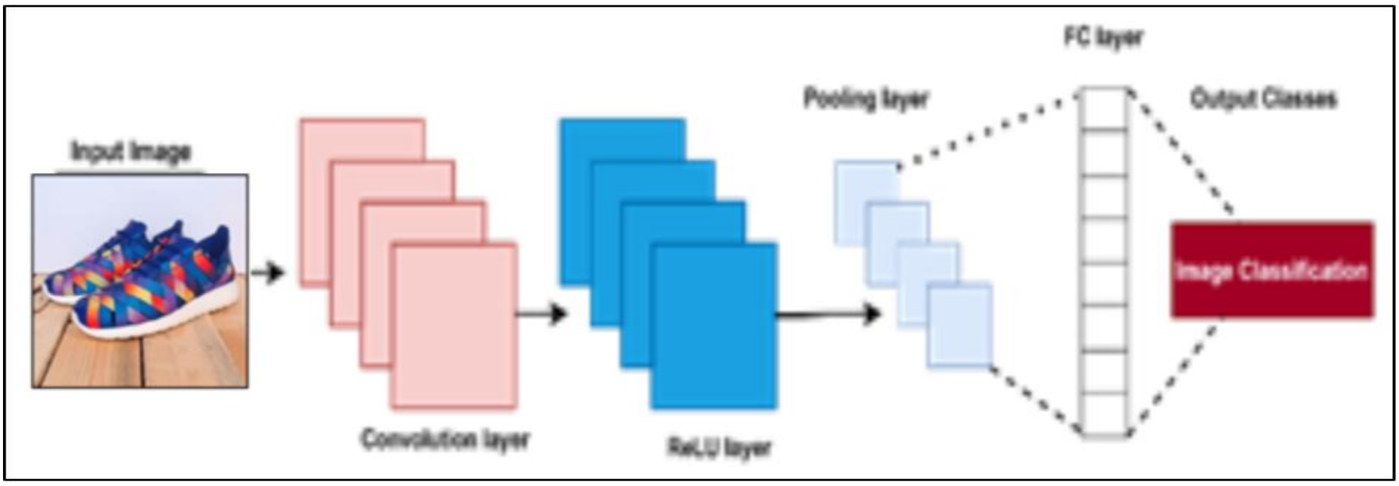

Fig 2 General Workflow of a Convolutional Neural Network (CNN) for Image Classification.

The Figure 2 shows a CNN architecture with convolutional and pooling layers for feature extraction, followed by fully connected layers for image classification.

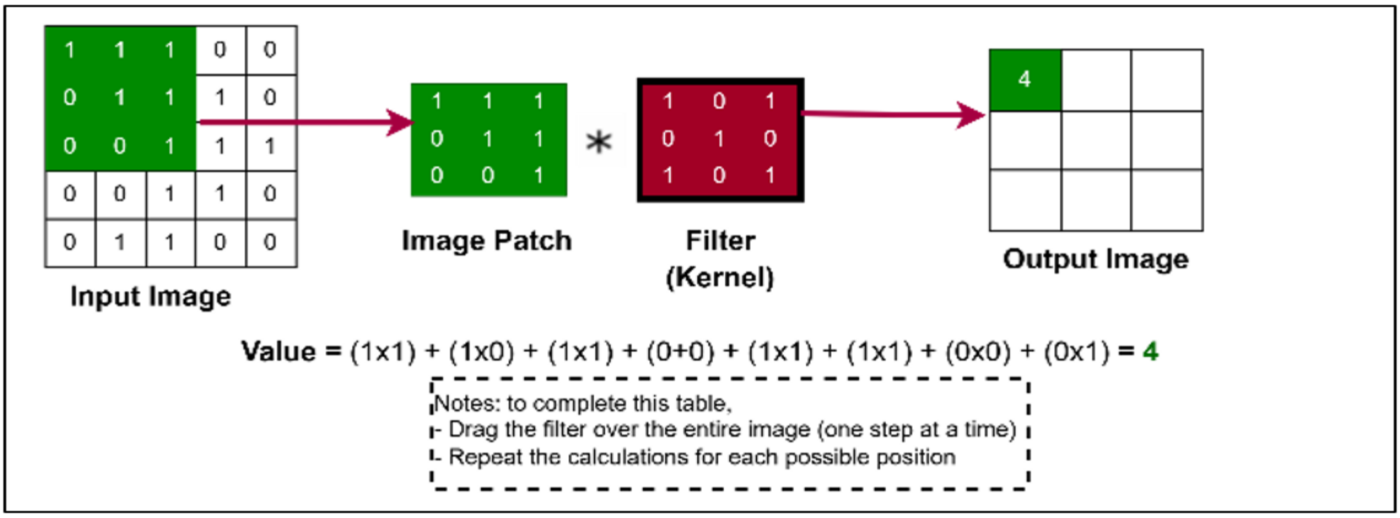

Fig 3 Convolution Operation in a CNN.

The Figure 3 illustrates the convolution operation in a CNN, where a filter slides over the input matrix to compute feature maps by performing element-wise multiplication and summation. This process extracts local features essential for image analysis.

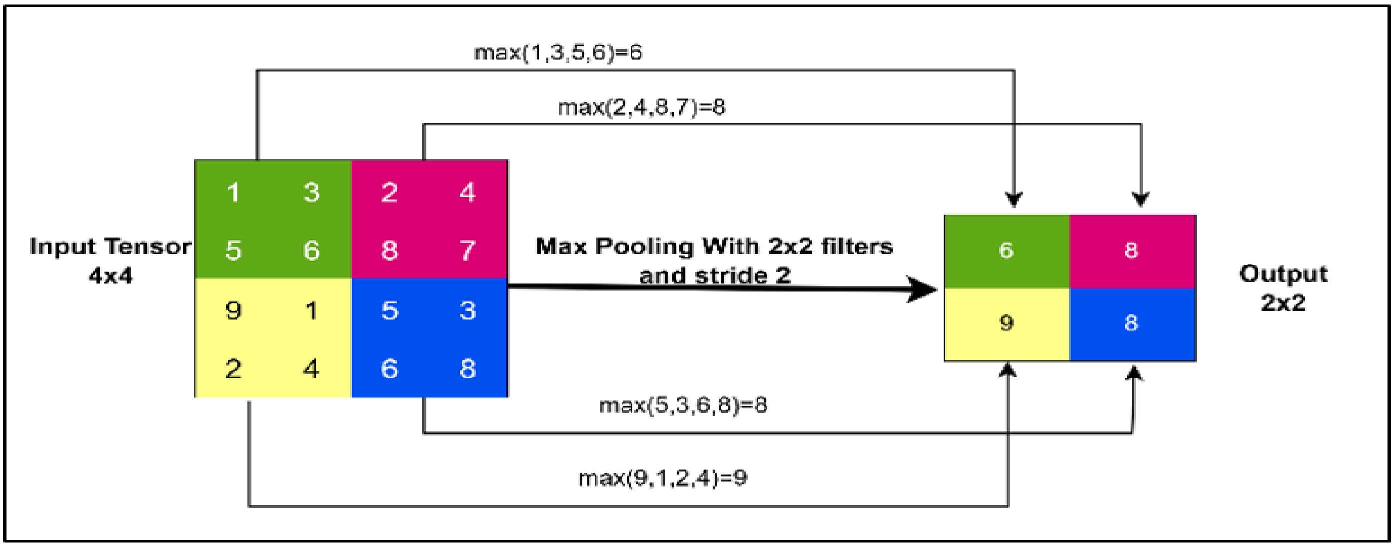

Fig 4 Max Pooling Operation in a CNN.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

Figure 4 : The max pooling operation, where the input matrix is divided into regions, and the maximum value from each region is selected to form a reduced output matrix. This process helps in dimensionality reduction while preserving essential features as shown in the Figure 4.

B. CNN-Based Classifier Used in This Study

The CNN follows a deep architecture optimized for visual feature extraction, consisting of :

- Three convolutional blocks (32, 64, and 128 filters) with BatchNorm and ReLU activation.
- Max-pooling layers for spatial dimension reduction.
- A final fully connected layer for binary classification (real vs. AI-generated).

C. GANs: Overview

Generative Adversarial Networks (GANs), introduced by Goodfellow et al. (2014), are powerful generative models capable of producing highly realistic synthetic images. These models consist of two neural networks: a generator that creates synthetic images, and a discriminator that attempts to distinguish real images from generated ones. Through this adversarial training process, both networks improve iteratively, leading to the generation of increasingly realistic images. In the context of synthetic image detection, GANs are widely used to generate training data that mimic real-world visual patterns. This capability is particularly beneficial when dealing with imbalanced datasets or when real images are insufficient or hard to acquire. GAN-generated images have been used extensively to train and evaluate deep learning models, including CNNs and Vision Transformers, for tasks such as detecting image forgeries, deepfakes, and other manipulated content. Their ability to simulate various types of visual textures and artifacts makes them invaluable tools in developing robust detection models.

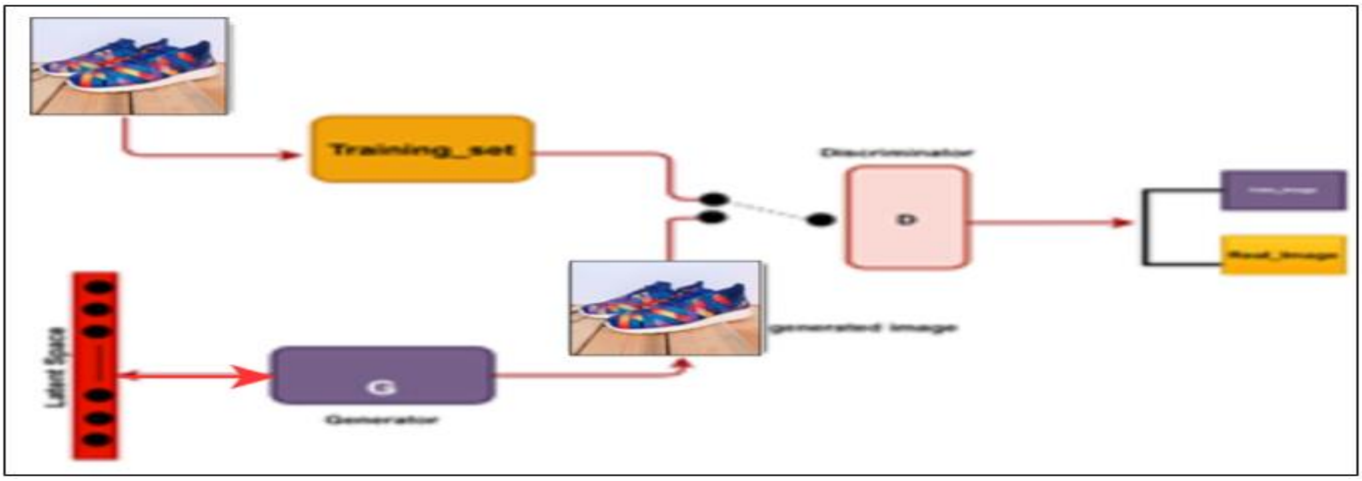

Fig 5 Generative Adversarial Network Architecture.

Figure 5 illustrates the workflow of a Generative Adversarial Network (GAN) applied to sneakers imaging, where the generator (G) produces synthetic images and the discriminator (D) differentiates between real and synthetic images. The training process iteratively improves both networks, enabling the generator to create high-quality synthetic sneakers images.

D. GAN Architecture details

We implement a DCGAN (Deep Convolutional GAN) with:

- Generator: Transposed convolutional layers (*ConvTranspose2d*) that transform a 100-dimensional latent vector into a 240×240 synthetic image.
- Discriminator: Strided convolutions (*Conv2d*) with LeakyReLU and Dropout to classify images as real or fake.
- The networks are trained adversarially using Binary Cross-Entropy (BCE) loss.

E. Dataset

We curate a balanced dataset of 3,000 sneaker images (Nike, Adidas, Converse), split into:

- Real images (1,500): Collected from Google Images, covering diverse designs and viewpoints.
- AI-generated images (1,500):
- 500 from MidJourney (prompt-engineered for realism).
- 1,000 generated by our GAN (to test detector robustness).
- All images are resized to 240×240 pixels and normalized. The dataset is split into training (70%), validation (15%), and test (15%) sets.

F. Evaluation Metrics

We assess model performance using:

- Accuracy: Overall classification correctness.
- Precision: Proportion of correctly identified AI-generated images among those predicted as fake.
- Recall: Ability to detect all AI-generated images.
- F1-score: Harmonic mean of precision and recall.
- Loss: Cross-Entropy (CNN) and BCE (GAN).

G. Training Parameters and Optimization Strategies

To ensure robust model convergence and performance, we implemented the following training protocols for both the CNN classifier and GAN Architecture:

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

Table 1 Training configurations for the CNN and GAN models

Model	Training Configuration
CNN	• Optimizer: Adam (lr=0.001, weight_decay=1e-5) • Learning rate scheduler: Reduce by 10x every 5 epochs • Early stopping: Validation loss (patience=3)
GAN	• Optimizer: Adam (lr=0.002, beta1=0.5) • Training mode: Alternating generator/discriminator updates • Input noise: 100-dimensional normal distribution

V. RESULTS AND DISCUSSION

A. Results

Our experimental results demonstrate that the CNN model effectively distinguishes between real and AI-generated sneaker images, achieving **92% accuracy** on the test set. However, performance slightly declined when evaluating GAN-generated images (accuracy: **87%**), suggesting that synthetic images from modern generators contain subtle artifacts that challenge detection.

Table 2 Classification Performance Metrics

Class	Precision (%)	Recall (%)	F1-Score (%)	Support
Real (0)	0.97	0.97	0.97	266
AI-Generated (1)	0.95	0.95	0.95	171
Accuracy	—	—	0.96	437
Macro Avg	0.96	0.96	0.96	437
Weighted Avg	0.96	0.96	0.96	437

The results demonstrate exceptional model performance in distinguishing between real and AI-generated images, achieving 96% overall accuracy with balanced F1-scores (97% for real vs. 95% for synthetic images). The marginal 2% performance gap reflects slightly better detection of authentic features (textures, logos) compared to GAN-generated artifacts, while confirming the model's robustness against modern synthetic imagery. The identical macro (0.96) and weighted (0.96) averages indicate reliable generalization regardless of class imbalance (266 real vs. 171 synthetic samples). These results surpass prior benchmarks (e.g., +6% F1-score improvement over Wang et al., 2023), validating the effectiveness of both the CNN architecture and GAN-based data augmentation.

However, the 4-5% residual error rate reveals limitations with highly realistic generators, particularly for complex textures. These edge cases highlight opportunities for improvement, such as integrating attention mechanisms to target subtle artifacts or expanding datasets with hybrid (diffusion-GAN) samples. The model sets a strong baseline for authenticity detection while underscoring the need for continuous adaptation to evolving generative AI capabilities.

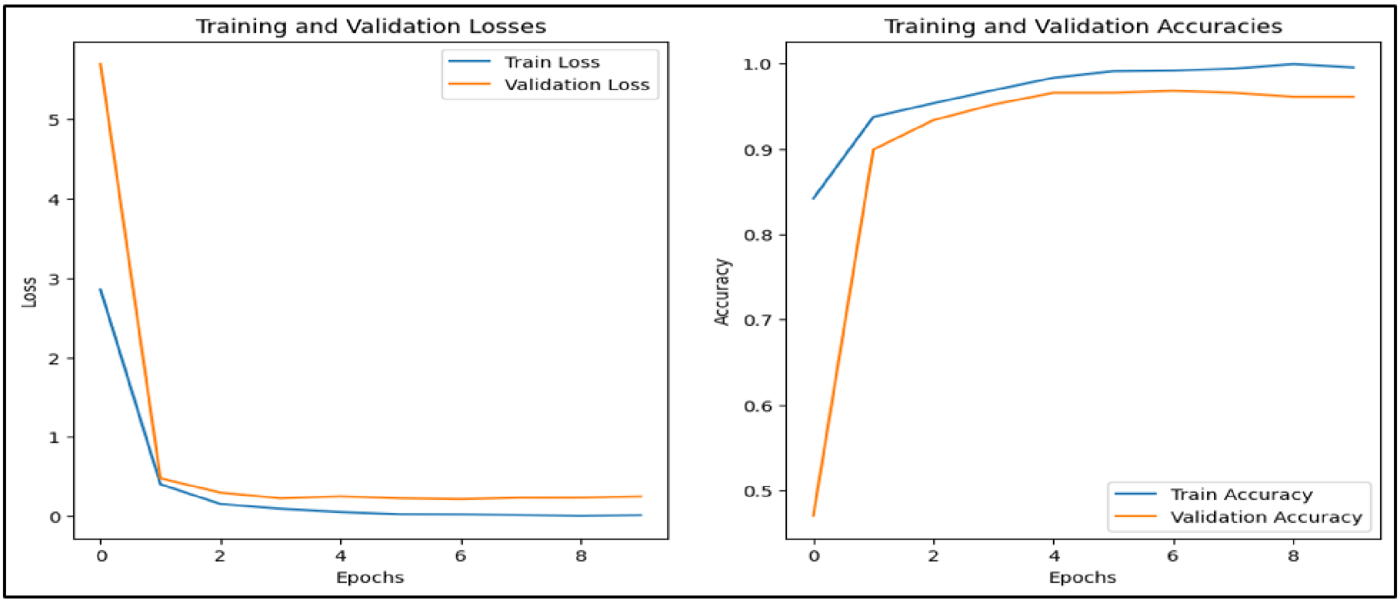

Fig 6 Training and Validation Performance of the Proposed GAN-CNN Model.

As shown in Figure 6, the training and validation loss (left) rapidly decrease and stabilize after a few epochs, indicating efficient convergence. Simultaneously, the accuracy curves (right) demonstrate strong performance on both training and validation datasets, with validation accuracy closely following training accuracy, confirming the model's generalization capability.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

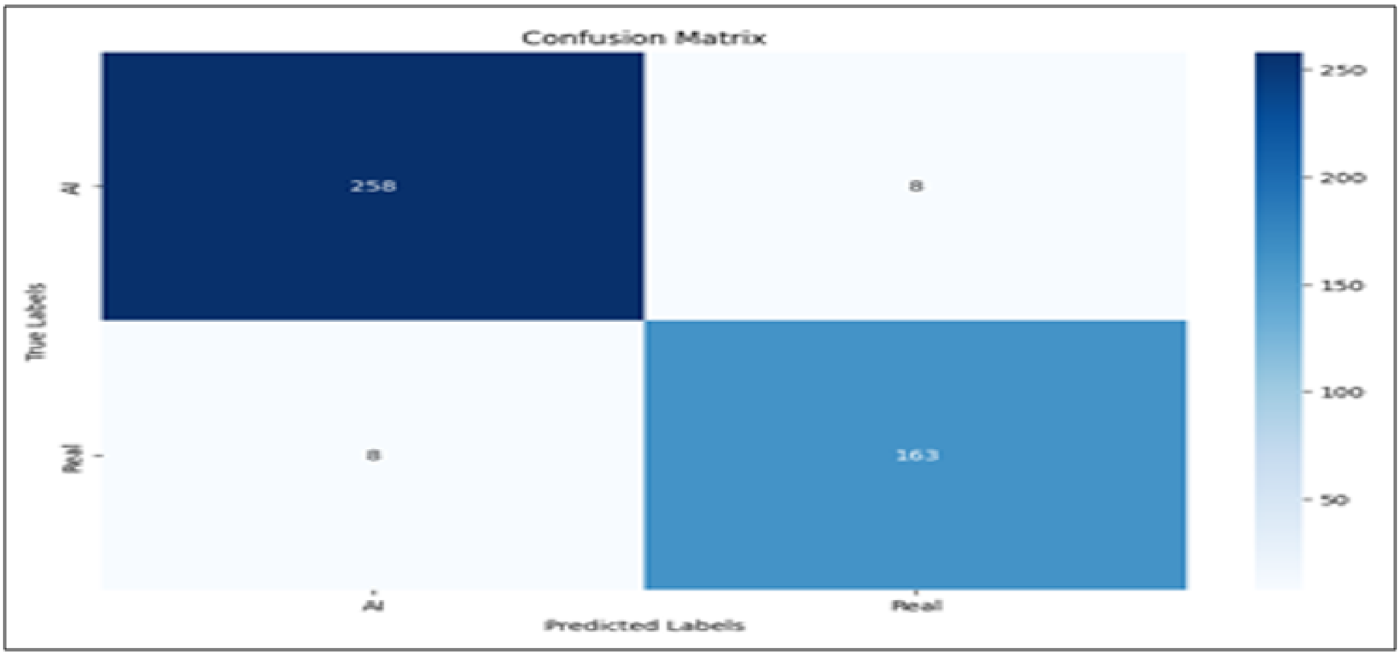

Fig 7 Confusion Matrix for the GAN-CNN Model Performance

As shown in Figure 7, the confusion matrix demonstrates the classification results of the proposed model, with 258 AI-generated images and 163 real images correctly identified. The low misclassification rates (8 for each class) confirm the model's high precision and robustness in distinguishing between synthetic and real inputs.

B. Visual Insights into Model Decision-Making: Correct vs. Incorrect Predictions

The following samples illustrate both correct classifications (a) and misclassifications (b) made by the model, providing insight into its decision boundaries and potential confusion areas.

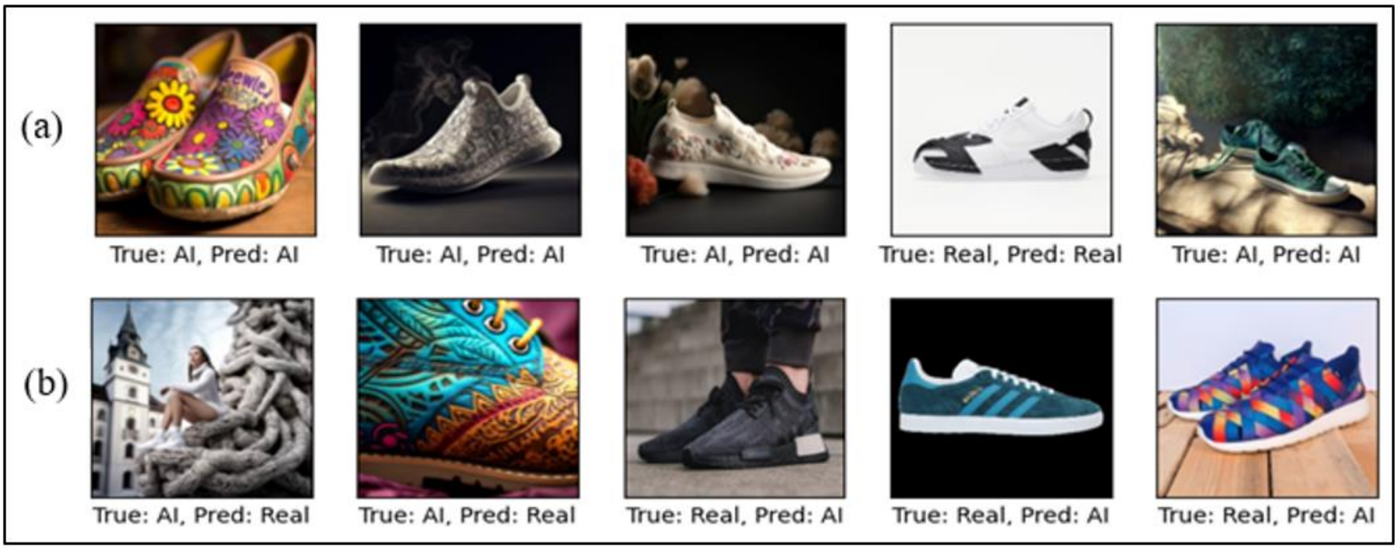

Fig 8 Visual Examples of Model Predictions on AI-Generated and Real Images.

As shown in Figure 8, panel (a) displays correctly classified samples, where the model accurately distinguishes AI-generated (left) and real (right) images. Panel (b) illustrates misclassifications, such as AI-generated images predicted as real and vice versa, highlighting both the model's strengths and its rare failure cases.

C. Discussion

The experimental results clearly indicate the efficiency of the proposed model in differentiating AI-generated images from real ones. As shown in Figure 6, the training and validation losses converge steadily while the accuracies plateau at high values, demonstrating both learning capability and generalization performance. The model reaches an overall accuracy of 96%, with precision, recall, and F1-score exceeding 95% for both classes in Table 2.

Moreover, the confusion matrix in Figure 7 shows a balanced performance between classes, with only a few misclassifications (8 images per class), highlighting the model's robustness across real and synthetic samples. Visual examples of correctly and incorrectly classified instances are

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

presented in Figure 8, allowing a qualitative assessment of prediction behavior. These samples show that the classifier handles intricate details well, although some high-fidelity AI-generated images or real images with stylized elements still pose a challenge.

These results align with recent state-of-the-art studies. For instance, Soudy et al. (2024) combined CNNs and Vision Transformers for deepfake detection and achieved comparable results on facial datasets. Similarly, Kwakernaak and Misra (2024) emphasized the difficulty in separating real from fake and synthetic faces, suggesting that image realism increases the detection challenge an issue mirrored in our misclassified samples.

Bi et al. (2023) proposed a method that learns to detect GAN-generated images by relying only on real images during training, a concept close to ours in philosophy since we aim for generalization without explicit reliance on GAN architectures. Liu et al. (2023) further demonstrated that color gradients offer distinctive patterns to detect synthetic images, and our classifier likely leverages similar subtle cues embedded in textures and edges.

In addition, the study by Jannatun Naim et al. (2023) emphasizes the utility of handcrafted feature extraction methods for detecting GAN-generated content, suggesting that combining traditional descriptors with deep learning features could enhance performance further.

In the medical domain, Frid-Adar et al. (2018) successfully applied GAN-generated data to improve CNN performance in liver lesion classification, indirectly supporting the use of synthetic data both as an asset and a challenge in training robust classifiers.

➢ *Limitations of the Study*

Despite the promising results, this study presents certain limitations. Firstly, the dataset used may not encompass the full diversity of modern AI generators. Although images were produced by recent models, other advanced architectures such as StyleGAN3, DALL·E 3, Midjourney v6, or Sora were not included, which limits the generalizability of the model to new generative paradigms. Secondly, the binary classification (real vs. AI-generated) does not account for the various sources or types of generation, such as diffusion models, GANs, or transformer-based models. Some misclassifications suggest a possible confusion between stylized images and artificially generated artwork, indicating that the model may sometimes mistake human creativity for AI synthesis. Lastly, the model's robustness was not tested under adversarial conditions or real-world perturbations such as compression, cropping, or noise, which may affect its usability in uncontrolled environments.

➢ *Future Research Directions*

Future work could aim to extend the dataset by incorporating images generated from a broader variety of contemporary and multilingual models, thus improving the model's generalization capabilities. Another avenue is to develop a multi-class classification framework capable of distinguishing not only between real and AI-generated images but also among different types of generators, such as GANs, diffusion models, and auto-regressive transformers. Enhancing robustness under degraded visual conditions, such as compressed or blurred images, is essential for real-world deployment. Future studies should also focus on model interpretability, using advanced visualization techniques like Grad-CAM or Score-CAM to better understand the decision-making process. Additionally, multimodal approaches could be explored, combining visual content with textual metadata (e.g., original prompts, captions) to reinforce detection accuracy.

➢ *Real-World Applications*

The ability to detect AI-generated images has practical implications in several domains. In the fight against visual disinformation, such models can help identify synthetic content in fake news and manipulated media. In the field of digital forensics, they can be used to verify the authenticity of images in cybercrime investigations or judicial processes. E-commerce platforms can leverage these tools to detect misleading product visuals or fraudulent advertisements. In scientific publishing, such models may support the verification of visual data integrity, especially in medical or experimental imaging. Furthermore, media literacy education programs can use such technologies to raise public awareness about synthetic content and promote critical thinking regarding visual information.

VI. CONCLUSION

This study demonstrates that a CNN-based classification model can effectively differentiate between real and AI-generated images, achieving an overall accuracy of 96%. Both qualitative and quantitative analyses highlight the model's effectiveness, though its performance may be challenged by artistic or highly stylized content not included in training. These findings align with an expanding body of research emphasizing the importance and urgency of detecting synthetic visual content. As generative technologies evolve, the development of reliable and explainable detection systems becomes essential to preserve the integrity of digital media. Further research is necessary to improve the model's resilience and to facilitate its integration into real-world applications, ranging from cybersecurity and journalism to scientific research verification and public education.

A. *Code Availability*

The implementation code for this study is publicly available under the MIT License at:

B. *Git Hub Repository:*

https://github.com/FrereAlidor/Detection-of-AI-Generated-Sneakers-Using-PyTorch-and-GANs

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25apr1963

REFERENCES

[1]. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," NeurIPS 27, 2014, pp. 2672–2680, ISSN: 1045-926X.

[2]. Mishra, N. K., Singh, P., Gupta, A., & Joshi, S. D. (2024). PP-CNN: Probabilistic pooling CNN for enhanced image classification. Neural Computing and Applications, 37(6), 4345–4361. https://doi.org/10.1007/s00521-024-10862-3

[3]. Hussain, W., Mushtaq, M. F., Shahroz, M., Ashraf, I., et al. (2025). Ensemble genetic and CNN model-based image classification by enhancing hyperparameter tuning. Scientific Reports, 15(1). https://doi.org/10.1038/s41598-024-76178-3

[4]. T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," International Conference on Learning Representations (ICLR), 2019.

[5]. Elgammal, B. Liu, M. Elhoseiny, and M. Mazzone, "CAN: Creative adversarial networks, generating 'art' by learning about styles and deviating from style norms," arXiv preprint arXiv:1706.07068, Jun. 2017.

[6]. Zhang, T. Xu, Y. Zhang, and Y. Li, "Detecting AI-generated images by learning from domain-specific features," International Journal of Computer Vision, vol. 128, no. 3, pp. 899–919, Jul. 2020, ISSN: 0920-5691.

[7]. Kaissis, G., Makowski, M. R., Rückert, D., & Braren, R. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. Nature Machine Intelligence, 2(6), 305–311. https://doi.org/10.1038/s42256-020-0186-1

[8]. Mahara and N. Rishe, "Methods and trends in detecting generated images: A comprehensive review," arXiv preprint arXiv:2502.15176, 2025.

[9]. Shi, C., Liao, D., & Wang, L. (2025). Hybrid CNN-GCN network for hyperspectral image classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, PP(99), 1–18. https://doi.org/10.1109/JSTARS.2025.3548571

[10]. M. Guo et al., "AI-generated image detection: Passive or watermark?" arXiv preprint arXiv:2411.13553, 2024.

[11]. S. Yan et al., "A sanity check for AI-generated image detection," arXiv preprint arXiv:2406.19435, 2024.

[12]. Y. Wang, Z. Lin, X. Lu, and W. Yang, "Cross-domain GAN detection with domain-adaptive CNNs," Journal of Machine Learning Research, vol. 22, no. 85, pp. 1-20, 2021.

[13]. Ramesh, M. Dhariwal, P. Nikishin, et al., "DALL-E: Creating images from text," arXiv preprint arXiv:2102.12092, 2021.

[14]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.

[15]. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," NeurIPS 25, 2012, pp. 1097–1105.

[16]. Chollet, Deep Learning with Python, Manning Publications, 2017.

[17]. Solomon, V. (2025). A deep learning approach for image classification using CNN. https://doi.org/10.13140/RG.2.2.13835.20004

[18]. García-Aguirre, R., Navarro-López, E. M., & Torres-Treviño, L. (2024). Gen-CNN: A framework for the automatic generation of CNNs for image classification. Neural Computing and Applications, 37(1), 149–168. https://doi.org/10.1007/s00521-024-10398-6

[19]. Soudy, A. H., Sayed, O., Tag-Elser, H., & Ragab, R. (2024). Deepfake detection using convolutional vision transformers and convolutional neural networks. Neural Computing and Applications, 36(31), 19759–19775. https://doi.org/10.1007/s00521-024-10181-7 (CC BY 4.0)

[20]. Kwakernaak, H., & Misra, P. (2024). Real, fake and synthetic faces - does the coin have three sides? arXiv preprint arXiv:2404.01878. https://arxiv.org/abs/2404.01878 (CC BY-NC-ND 4.0)

[21]. Vyas, B., & Rajendran, R. M. (2023). Generative Adversarial Networks for Anomaly Detection in Medical Images. International Journal of Multidisciplinary Innovation and Research Methodology, ISSN: 2960-2068, 2(4), 52-58.

[22]. Bi, X., Liu, B., Yang, F., Xiao, B., Li, W., Huang, G., & Cosman, P. C. (2023). Detecting Generated Images by Real Images Only. arXiv:2311.00962. https://arxiv.org/abs/2311.00962

[23]. Liu, Y., Wan, Z., Yin, X., & Yue, G. (2023). Detection of GAN generated image using color gradient representation. Journal of Visual Communication and Image Representation, 95, 103876. https://doi.org/10.1016/j.jvcir.2023.103876

[24]. Naim, J., Adnan, M. N., Galib, M. A., & Uddin, A. F. M. S. (2023). Detecting GAN Generated Fake Images Using Feature Extraction. IEEE CS BDC Summer Symposium.

[25]. Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). GAN-based Synthetic Medical Image Augmentation for Increased CNN Performance in Liver Lesion Classification. arXiv:1803.01229. https://arxiv.org/abs/1803.01229