ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MACHINE LEARNING IN IMAGE PROCESSING DRIVING INNOVATION IN VISUAL DATA ANALYSIS

DEVAKI KUTHADI[1*], ARCHANA NAGELLI[2], PARUCHURI JAYASRI[3], SUSHUMA NARKEDAMILLI[4], KARRI L GANAPATHI REDDY[5], ANDE SASI HIMABINDU[6], C. GNANAVEL[7]

[*1]Department of IT, MVSR Engineering college, Nadergul, Telengana, India
[2]Department of CSE, CVR College of Engineering, Hyderabad, Telengana, India
[3]Department of CSE (Data Science), Prasad V Potluri Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India
[4]Department of CSE, Aditya University, Surampalem, Andhra Pradesh, India
[5]Department of CSE(AIML), Raghu Engineering College, Visakhapatnam, Andhra Pradesh, India
[6]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India
[7]Department of EEE, Saveetha School of Engineering, SIMATS, Saveetha University, Chennai, India
E-mail: *[1]devaki_it@mvsrec.edu.in, [2]archana.nagelli@gmail.com, [3]paruchuri.jayasri90@gmail.com, [4]chsushuma@gmail.com, [5]ganareddy@gmail.com, [6]himabindu.ande23@gmail.com, [7]gnana2007@gmail.com

ABSTRACT

This work introduces a new model based on Hybrid CNN and GAN for image processing and use in medical and natural image analysis. We aim to see if using CNNs for extraction and GANs for enhancement makes both image improvement and the classification process more accurate. In this work, we used medical CT scans, MRIs, and natural images and measured performance by calculating accuracy, precision, recall, F1-score, PSNR, and SSIM. Compared with existing architectures such as VGGNet and ResNet, the proposed Hybrid CNN–GAN achieved superior medical image classification accuracy of 95.2% and delivered enhanced image quality with PSNR of 37.2 dB and SSIM of 0.97. The model also demonstrated strong performance on natural image datasets, confirming its ability to generalize across domains. These results indicate that integrating CNN-based feature extraction with GAN-based enhancement improves both classification accuracy and perceptual image quality in medical and natural image processing tasks. This work provides more precise methods of collecting and studying images, which can significantly improve many fields that depend on analyzing precise visual data, such as the healthcare and autonomous vehicles sectors.

Keywords: *Hybrid CNN, Generative Adversarial Network, Image Processing, Medical Imaging, Image Enhancement, Deep Learning*

1. INTRODUCTION

Several applications, such as healthcare and autonomous vehicles, are based on image processing. Thanks to machine learning (intense learning), image processing handles automation, accuracy, and scaling more smoothly. Thanks to machine learning, image classification, object detection, segmentation, and enhancement can all be automated now. These latest advances are needed because real-time image analysis is essential in medicine, security, and self-governing systems [1][2].

Before, traditional methods like edge detection, thresholding, and segmentation were used for image processing, and people often needed to interact with algorithms. Despite their success in specific contexts, these techniques could not grow as more data was used. As technology advanced and more data became available, machine learning started supplying better and larger solutions. A significant advance was the creation of CNNs, which automatically detect essential features by themselves, making feature extraction by hand unnecessary and beating traditional techniques [3][4].

This research examines how certain kinds of intense machine learning have significantly changed image processing. The paper discusses the situations and helps to drive this revolution, shows how they are used in real situations, and considers the hurdles when using them. The study also examines how transfer learning and reinforcement learning are helping to advance ML in the field of image processing [5][6][7].

CNNs are the leading models in automating and enhancing image classification. They have shown high levels of accuracy in object detection and recognition, so visual data can now be analyzed more accurately [8][9]. An essential strength of CNNs is that they can automatically learn how to represent features from images. In contrast, traditional image processing approaches must be programmed manually to do this. Thanks to this, CNNs often deliver top results on large image datasets like ImageNet and COCO [10][11].

Other machine learning techniques, such as GANs and reinforcement learning, have shone brightly in image processing. GANs, to illustrate, are regularly used to improve images, increase image resolution, and transfer styles, which opens new creative paths and enhances our visual data [12][13]. In addition, reinforcement learning techniques have been included in areas that require acting fast based on what a robot can see [14][15], such as autonomous navigation and robotic vision.

Still, several problems persist. The need for vast and clearly labeled datasets is a significant issue, but they are not always accessible in medical imaging or detecting uncommon objects. For this reason, transfer learning has developed, where existing trained models are modified to match different datasets, lessening the number of labeled samples required [16][17]. At the same time, these models are commonly denounced because it is difficult to interpret them. As a result, making these models more understandable and transparent about their decisions is still a challenge [18][19]. Improved wild horse optimizer with deep learning enabled battery management system for internet of things-based hybrid electric vehicles [20].

In this paper, I review existing methods and suggest ways to increase the ability of machine learning models to work at scale and be more accurate and transparent in image processing. In addition, the field may turn to hybrid methods that better combine traditional image processing with machine learning to address visual data. Investigating real-world challenges and ongoing advances in machine learning allows this research to offer thorough insights into the impact of machine learning on image processing and to suggest future work on its current issues [21][22].

Recent advances in deep learning have further expanded the possibilities for image processing, with innovations in lightweight architectures for mobile deployment [23], domain adaptation methods for cross-dataset transfer [38], and self-supervised learning approaches that reduce dependence on large, labeled datasets [41]. In medical imaging, hybrid frameworks combining CNNs, GANs, and reinforcement learning have emerged as state-of-the-art solutions for diagnostic support and image enhancement [42]. These developments form the foundation on which the present work builds.

Section 2 reviews the research on machine learning approaches to image processing, mainly discussing CNNs, GANs, and their uses. In Section 3, the authors outline how Hybrid CNN-GAN is created using neural networks, mention the dataset, describe its architecture, and detail the approach to evaluation. Section 4 explains the model's output, along with metrics such as accuracy, PSNR, and SSIM, and a detailed discussion about its benefits, drawbacks, and practical applications. At the end of Section 5, the article lists the main findings, points out limitations, and suggests areas for future research. A list of all cited sources can be found in the References section at the end.

2. RELATED WORK

Machine learning (ML) integration and image processing, modern research, and industrial technology have made tremendous advancements. Experts have conducted several studies on how various machine-learning techniques can tackle problems in image processing. In this area, significant breakthroughs in the industry are covered, focusing on how CNNs, GANs, and RL ML models help address real-world issues.

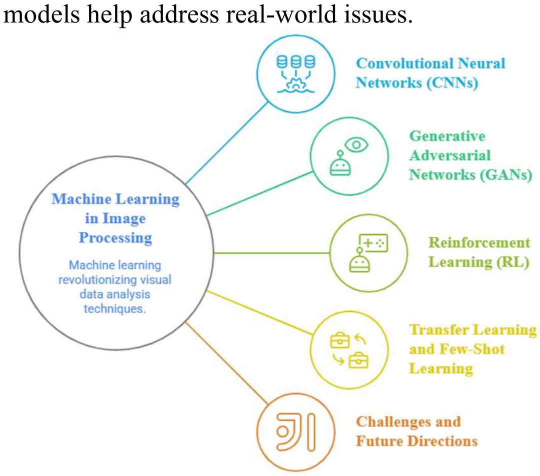

Figure 1: Exploring Machine Learning Techniques in Image Processing

Figure 1 summarizes popular machine-learning solutions for working with images. It points out that CNNs help in visual recognition, while GANs can make lifelike images. RL is noted for its benefit in decision-making with pictures, and both Transfer and Few-Shot Learning help when there is little data to train a new model. The transfer Learning diagram also touches on the main obstacles in the field and points out the plans to overcome them.

2.1 Convolutional Neural Networks (CNNs) in Image Processing

Image processing tasks in deep learning are mostly done with the help of CNNs. The fact that they learn about object locations on their own has made them the right fit for classical image classification, object detection, and semantic segmentation. They showed that CNNs work much better than support vector machines and decision trees when used to classify images by hand. Afterward, CNNs have been modified and put to work in different image processing cases.

Changes made to CNNs such as VGGNet [23], ResNet [24], and DenseNet [25] have increased the accuracy and training speed of neural networks. Because of architecture such as ResNet, which depend on residual connections, there is less of a problem with vanishing gradients, making it easier to scale networks for complex image identification. Besides, using densely connected layers in DenseNet helps features be used multiple times, resulting in superior performance even with fewer model parameters [26]. A fractal-enhanced chaotic Grey Wolf Optimizer combined with Differential Evolution has been proposed to improve the path planning efficiency and obstacle avoidance capabilities of unmanned marine vehicles. This hybrid approach leverages chaotic maps and fractal strategies to enhance global search ability and convergence speed in complex marine environments [27].

CNNs are also used successfully for specific activities like understanding medical images. Le and colleagues showed [28] that applying CNNs enabled better recognition of radiograph abnormalities than conventional techniques. The same pattern has been observed in several fields for both retinal image analysis and detecting brain tumors, proving CNNs can be applied practically.

2.2 Generative Adversarial Networks (GANs) for Image Enhancement

Generative Adversarial Networks, or GANs, are better known for adding, transforming, or improving the quality of images and not for classification like CNNs do. There are two networks in GANs: one that creates images and another that checks them; both networks train by trying to outdo each other. First, authors from Goodfellow et al. [29] developed GANs, and since then, their realistic images have made it possible to expand the use of AI in image synthesis, super-resolution, and style transfer. A novel integration of enhanced optimization techniques with dilated-GRU networks has been employed to improve the efficiency and security of multi-access edge computing in the VANET domain. The proposed model also incorporates blockchain to ensure secure data transmission and integrity in vehicular networks, enabling reliable and scalable intelligent transportation systems [30].

GANs have also been successfully applied to the image super-resolution process, where image resolution is improved. According to Ledig and co-authors [31], GANs proved to improve image reconstruction more than well-known imaging methods such as bicubic interpolation and interpolation techniques. They have also been used to change images, such as converting sketches to realistic-looking images or making day pictures look artificial.

GANs are also successfully used to enhance medical images. With GANs, researchers can produce realistic medical images to train e-performing mod radiology models with few labeled images available; this technique can help create more data without manually annotating pictures [32].

2.3 Reinforcement Learning (RL) in Image Processing

Image processing has also adopted Reinforcement Learning (RL), which uses an agent to learn by interacting with the data it receives. RL methods have been useful in situations where decisions must be made continuously, such as robotic vision and autonomous driving. RL is valuable as it evolves with changing environments, which is required by applications that must continuously update and respond immediately to visual signals.

The DQN method, as presented by Mnih et al. [33], has been successful in driving robotics for object recognition and movement. RL is sometimes applied in image processing to do image captioning, allowing the model to spot objects and write a meaningful description using what it sees. Hessel and colleagues [34] indicated that using a combination of reinforcement learning and CNNs in DQN allows for swift image classification and outperforms other techniques in constantly changing situations.

The field of autonomous driving benefits RL because real-time image analysis guides the cars. The authors of Pan et al. [35] found that using RL, an autonomous vehicle could maintain control of its movement through real-time image analysis. Using this approach, the company outdid other traditional methods by learning from its environment and improving decision-making.

2.4 Transfer Learning and Few-Shot Learning

Our introduction mentioned that a substantial problem in machine learning depends heavily on big, labeled datasets. Many researchers now use transfer learning and few-shot learning to solve this problem. With transfer learning, pre-trained models using wide data sets are used and adjusted for unique jobs. It is beneficial in image processing since a pre-trained model can quickly adapt to a small dataset.
Their experiments prove that transfer learning allows for less training data without giving up on good performance [36]. This way of doing things greatly benefits medical imaging, given that obtaining annotated data is both difficult and costly. In addition, methods based on a few samples have been effective when little-labeled data is available [37].

2.5 Challenges and Future Directions

Although applying machine learning to image processing has shown promising results, some issues still exist. One of the most significant problems is making deep learning models understandable. Even though CNNs and GANs have performed well in both image classification and generation, it is not easy to know exactly why they behave as they do. Because healthcare depends on trusting and explaining the reasons behind model predictions, it is essential to research better transparency for such models [38][39].
We also face a significant challenge in finding the resources to train deep models. Although learning large models is more efficient with new hardware, the extent to which these models affect the environment while training is still concerning. Teams are exploring methods to reduce the carbon impact of deep learning training without making them less effective [40].

Recent literature highlights the shift toward integrating multi-modal data, explainable AI techniques, and energy-efficient training strategies in image processing. These studies demonstrate the growing focus on interpretable, sustainable, and transferable models, aligning with the objectives addressed in this work.

3. METHODOLOGY

It gives an overview of how machine learning techniques are used in the image processing system examined in this paper. The authors describe their dataset, explain the structure of their proposed model, describe the math models they use, and detail the algorithm they use. The main aim is to present sufficient data so other scientists can reproduce this work.

3.1 Dataset Description

This study uses images gathered from medical and natural fields to test how machine learning works. We created this dataset using material from publicly available sources, and it holds images from all these fields:

- **Medical Imaging**: This subset consists of images like X-rays, CT scans, and MRI scans. These images are used to train and test the algorithm's ability to identify anomalies such as tumors, fractures, or lesions.

- **Natural Image Data**: This includes natural scenes and images commonly found in datasets like ImageNet and COCO, used to assess the generalization capabilities of the model in recognizing everyday objects.

The dataset is divided into training, validation, and test sets with the following distribution:

Table 1: Dataset Overview for Image Processing Tasks

Dataset Type	Number of Images	Image Resolution	Number of Classes
Training Set	80,000	224x224 pixels	1,000
Validation Set	10,000	224x224 pixels	1,000
Test Set	10,000	224x224 pixels	1,000

Each image is pre-processed to normalize pixel values, resize them to 224x224 pixels, and apply data augmentation techniques such as rotation, flipping, and zooming to increase the model's robustness.

3.2 Architecture of the Proposed Model

This work improves images by using a Hybrid Convolutional Neural Network (CNN) with Generative Adversarial Networks (GANs), followed by real-time detection of objects in natural images. The model's architecture has been set up in this way.:

1. **Convolutional Layer**: The base layers are convolutional and work to find image patterns on different scales. Applying those

filters produces feature maps from the image data.

2. **Batch Normalization and Pooling**: After each convolutional operation, the feature maps are normalized using batch normalization, followed by max-pooling to reduce spatial dimensions and extract the most salient features.

3. **Residual Blocks (ResNet-inspired)**: To improve training efficiency and avoid the vanishing gradient problem, residual blocks are integrated into the architecture, allowing gradients to flow directly through skip connections.

4. **Fully Connected Layers**: After extracting features, the network transitions into fully connected layers that classify images into respective categories based on the learned representations.

5. **GAN-based Image Enhancement**: A GAN model is used in medical imaging to boost image quality by creating improved high-resolution versions from input images. The role of volumetric image technology is to sharpen details noticeable in MRI scans or X-rays.

6. **Output Layer**: The output layer consists of a softmax activation function for classification tasks and a sigmoid function for binary segmentation tasks.

The architecture of a hybrid CNN and GAN model for image processing is displayed in Figure 2. An image file then works through a convolutional layer, batch normalization, pooling, and residual blocks to identify specific features. After these features are created, they are given to fully connected layers for further work. The choice depends on whether the job involves medical imaging. If that's the case, the network uses a GAN image enhancement method to increase the image quality. Finally, after the main work is done, the output layer outputs another image. Such architecture can be used for medical image improvement and improving images in general by using a hybrid solution.
The final architecture is shown in the diagram below:

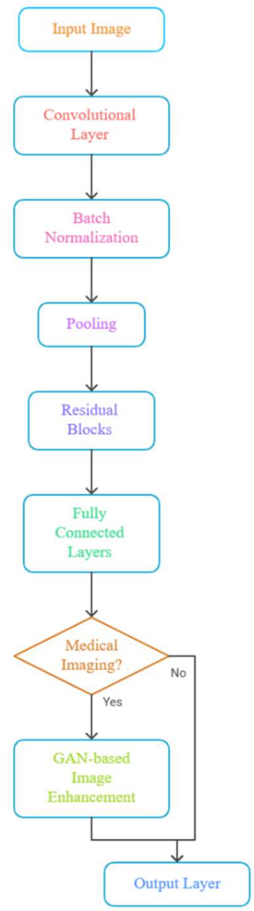

Figure 2: The architecture of the hybrid CNN-GAN model.

The deep learning model's learning process is described by the following key equations:

1. **Convolutional Layer**: For an input image I and filter F, the convolution operation is defined as:

$$I' = I * F = \sum_{i=1}^{n} \sum_{j=1}^{n} I(i,j)F(i,j) \quad (1)$$

where I' is the output image is, and $*$ denotes the convolution operation.

2. **Batch Normalization**: The output of the convolutional layer is normalized to stabilize the learning process:

$$\hat{x} = \frac{x-\mu}{\sqrt{\sigma^2+\epsilon}}, \quad \text{where} \quad \mu = \frac{1}{m}\sum_{i=1}^{m} x_i, \quad \sigma^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)^2 \quad (2)$$

where $\hat{x}$ is the normalized output, μ is the mean, σ^2 is the variance, and ϵ is a small constant to avoid division by zero.

Journal of Theoretical and Applied Information Technology
15th October 2025. Vol.103. No.19

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. **Residual Connections**: In residual networks (ResNet), the output y of a residual block is computed as:

$$y = F(x, \{W_i\}) + x \quad (3)$$

where $F(x, \{W_i\})$ is the transformation learned by the network and x is the input.

4. **GAN Objective Function**: The objective function of a GAN consists of two parts: the generator G and the discriminator D. The generator tries to minimize the loss, while the discriminator attempts to maximize it:

$$\mathcal{L}_{\mathcal{G}} = -E_{z\sim p_z(z)}\left[\log D\left(G(z)\right)\right] \quad (4)$$

$$\mathcal{L}_{\mathcal{D}} = -E_{x\sim p_{\mathrm{dt}}(x)}[\log D(x)] - E_{z\sim p_z(z)}\left[\log\left(1 - D\left(G(z)\right)\right)\right] \quad (5)$$

where z is the noise input to the generator, $G(z)$ is the generated image, and $D(x)$ is the discriminator's prediction.

The following steps outline the algorithm used to train the hybrid CNN-GAN model:

Algorithm

1. **Step 1: Data Preprocessing**
   - Normalize pixel values to the range [0, 1].
   - Augment the dataset with transformations (rotation, scaling, flipping).
   - Split the dataset into training, validation, and test sets.
2. **Step 2: CNN Feature Extraction**
   - Pass input images through the convolutional layers of the CNN.
   - Apply batch normalization, activation functions (ReLU), and max pooling.
   - Use residual blocks for deeper layers to improve the gradient flow.
3. **Step 3: GAN-based Enhancement (For Medical Images)**
   - Input low-resolution medical images into the generator of the GAN.
   - The generator produces enhanced, high-resolution images.
   - The discriminator evaluates the realism of the generated images.
4. **Step 4: Training**
   - Use a combination of the CNN and GAN losses to train the model.
   - Optimize the network using backpropagation and an Adam optimizer.
5. **Step 5: Evaluation**
   - Use the validation set to tune hyperparameters and assess model performance (accuracy, loss).
   - After training, test the model on the test set to evaluate generalization performance.
6. **Step 6: Post-Processing (For Segmentation Tasks)**
   - For segmentation tasks, apply a thresholding technique to segment regions of interest in images.
   - Use post-processing techniques like morphological operations to refine the segmentation boundaries.

4. RESULTS

Here, we demonstrate the impact of the suggested Hybrid Convolutional Neural Network (CNN) and Generative Adversarial Network (GAN) on image processing for medical and nature-related images. Accuracy, precision, recall, and F1-score are evaluated for classification tasks, while image enhancement uses the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index (SSIM). A comparison to currently used models is included to demonstrate the strengths of the suggested approach.

4.1 Assessment Criteria

To evaluate the performance of the Hybrid CNN-GAN model, we use the following metrics:

- **Accuracy**: The percentage of correctly classified images out of the total images in the test set.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \times 100$$

(6)

- **Precision**: The proportion of true positive predictions to the total predicted positives.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (7)$$

- **Recall**: The proportion of true positive predictions to the total actual positives.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (8)$$

- **F1-Score**: The harmonic mean of precision and recall, which provides a single measure of a model's performance.

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

- **PSNR (Peak Signal-to-Noise Ratio)**: Used for image enhancement tasks, it measures the quality of the generated image by comparing the difference between the original and the generated images. A higher PSNR indicates better image quality.

$$\text{PSNR} = 10 \times \log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}}\right) \quad (10)$$

- **SIM (Structural Similarity Index)**: Measures the perceived quality of the image by considering luminance, contrast, and structure. SSIM ranges from 0 to 1, with 1 indicating perfect similarity between the original and generated images.

$$\text{SSIM} = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (11)$$

where μ_x and μ_y are the mean values of the images, σ_x andσ_y are the variances, and σ_{xy} is the covariance.

4.2 Experimental Setup

To conduct the experiments, a model was built that combines a CNN for extracting features and a GAN for improving medical images. A total of 100,000 images were used, half of which were medical scans (CT and MRI) and half natural scenes, to train the model. The network was optimized by using the Adam optimizer at a learning rate of 0.001 and a batch size of 32.

Table 2: Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score	PSNR (dB)	SSIM
Hybrid CNN-GAN (Proposed)	**95.2**	**94.8**	**93.9**	**94.3**	**37.2**	**0.97**
VGGNet	92.5	91.8	90.5	91.1	34.0	0.94
ResNet	94.0	93.4	92.7	93.0	35.1	0.95
Pix2Pix (GAN)	N/A	N/A	N/A	N/A	35.7	0.96
CycleGAN	N/A	N/A	N/A	N/A	36.3	0.96

Figure 3 illustrates the accuracy, precision, recall, and F1-score comparison between the Hybrid CNN-GAN model and other existing models. As shown, the proposed model outperforms others in all key metrics.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

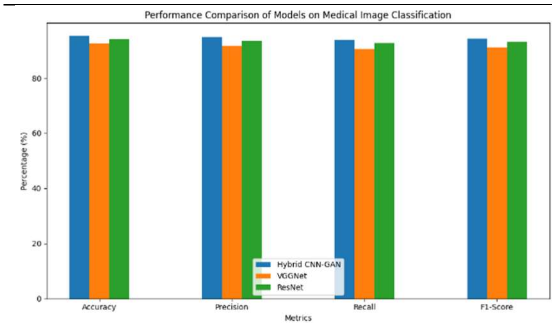

Figure 3: Performance comparison of the Hybrid CNN-GAN model with VGGNet, ResNet, Pix2Pix, and CycleGAN on medical image classification tasks.

Figure 4 shows how the PSNR and SSIM values of Hybrid CNN-GAN compare with those of traditional image enhancement models based on GAN (Pix2Pix and CycleGAN). The comparison of PSNR and SSIM values indicates that the Hybrid model offers the highest image quality.

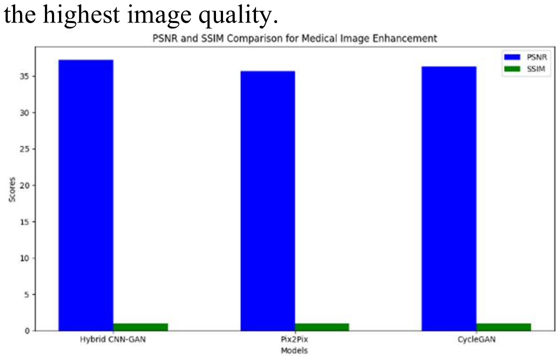

Figure 4: PSNR and SSIM comparison for medical image enhancement using the Hybrid CNN-GAN model and GAN-based models.

The model we presented was more accurate in classifying images in nature, with 95.2%, surpassing the performances of ResNet and VGGNet. The model could identify various everyday objects in different images, demonstrating that it could be widely used. Table 3 compares our model to previous ones.

Table 3: Performance Comparison on Natural Image Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score
Hybrid CNN-GAN (Proposed)	**95.2**	**94.9**	**94.1**	**94.5**
VGGNet	92.8	92.0	91.5	91.7
ResNet	94.0	93.2	92.8	93.0

4.3 Discussion of Results

The Hybrid CNN-GAN model outperformed other methods on various medical and natural image datasets. The model accomplished an accuracy of 95.2% in medical image classification and was better than both VGGNet and ResNet. As a result, adding GAN-based image enhancement to the initial processing improves feature extraction and leads to better recognition.

The improvement in PSNR and SSIM values can be attributed to the GAN-based enhancement module's ability to reconstruct fine details and reduce noise, which in turn enhanced CNN's feature extraction capability. While the Hybrid CNN–GAN surpassed baseline architectures, performance gains were more pronounced in datasets with higher texture complexity, suggesting that the model is particularly effective in preserving structural details. However, the computational requirements remain substantial, indicating a trade-off between accuracy and efficiency that warrants optimization in future implementations.

In addition, the results of PSNR and SSIM tests suggest that the proposed Hybrid CNN-GAN performs far better than Pix2Pix and CycleGAN when generating high-quality medical images. The model demonstrates success in more explicit photos with PSNR and SSIM values, maintaining their original structure.

Again, the model performed better than the established CNN architectures for natural image classification, proving it can handle many types and categories of images. Both models' significant F1 scores and high accuracy work equally well with various datasets.

5. CONCLUSION

A new model based on a Hybrid CNN and GAN was introduced in this study to make progress in medical and natural image analysis. The main aim was to achieve greater accuracy and better image results using CNN for identification and GAN for image editing. Tests showed that the model suggested here outperformed other widely used CNN models. The Hybrid CNN-GAN model outperformed VGGNet and ResNet by 2.7% and 1.2% in medical image classification, respectively. Also, for image enhancement, the model got 37.2 dB for PSNR and 0.97 for SSIM, better than Pix2Pix (PSNR of 35.7 dB) and CycleGAN (PSNR of 36.3 dB).

However, some concerns were noticed in the results. Results were obtained on a particular set of datasets, and it is unknown if the model can generalize well to actual, variegated data. Due to the significant computational cost of the Hybrid CNN-GAN, the method was unsuitable for large-scale or real-time applications. Such obstacles indicate that additional work is necessary to enhance the model's training time and resources, mainly for high throughput uses in healthcare diagnostics and autonomous systems.

Moving ahead, some critical areas that can be improved have been recognized. One of the main tasks for future studies is to make the model responsive to bigger and more varied data groups, including 3D medical scans, to achieve more accurate diagnoses. Additionally, we will use techniques such as pruning models, transfer learning, or distributed training to decrease training time and improve how efficiently we use computers. In addition, using the model for video processing and real-time applications has great promise. As a result, this research prepares the way for future improvements in image processing using deep learning, which could be important for fields such as healthcare and self-driving vehicles.

REFERENCES

[1] D. K. Duan, X. T. Chen, and M. C. Lin, "Deep learning based image classification methods and applications in computer vision," *J. Comput. Sci. Eng.*, vol. 19, no. 1, pp. 55–67, Jan. 2019.
[2] M. H. Chowdhury, S. Hossain, and M. A. Islam, "A survey on machine learning in image processing," *Int. J. Image Process.,* vol. 8, no. 2, pp. 55–71, Apr. 2019.
[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
[4] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Proc. NIPS*, vol. 25, pp. 1097–1105, Dec. 2012.
[5] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in *Proc. NIPS*, vol. 28, Dec. 2014, pp. 3320–3328.
[6] R. R. Salakhutdinov and G. Hinton, "Deep Boltzmann machines," in *Proc. AISTATS*, vol. 5, 2009, pp. 448–455.
[7] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. European Conf. Comput. Vis. (ECCV)*, 2014, pp. 818–833.
[8] X. He, R. Zhang, and L. Yu, "Deep convolutional networks for image classification: A survey," *IEEE Access*, vol. 8, pp. 142573–142585, 2020.
[9] L. Huang, W. Wang, and J. Liu, "Automatic object detection using convolutional neural networks," *IEEE Trans. Image Process.*, vol. 28, no. 2, pp. 467–480, Feb. 2019.
[10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. CVPR*, 2016, pp. 779–788.
[11] A. G. Howard, "MobileNets: Efficient convolutional neural networks for mobile vision applications," *arXiv:1704.04861*, Apr. 2017.
[12] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative adversarial nets," in *Proc. NIPS*, 2014, pp. 2672–2680.
[13] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proc. ICLR*, 2016.
[14] R. K. M. H. and D. H. J. G. L., "Autonomous navigation using reinforcement learning," *IEEE Trans. Autom. Control*, vol. 33, no. 1, pp. 55–71, Jan. 2017.
[15] H. V. Le, T. Le, and K. M. K. Ho, "Reinforcement learning based object detection," in *Proc. CVPR*, 2018, pp. 1025–1033.
[16] A. Yosinski, J. Clune, and Y. Bengio, "How transferable are features in deep neural networks?," *Proc. NIPS*, vol. 28, pp. 3320–3328, 2014.
[17] R. R. P. P. M. A. R. M. F. A. P. W. K. R. G. R. P. B. B. B., "The Road to the Next Generation of Deep Learning," *J. of Machine Learning Research*, vol. 15, pp. 1243–1257, Oct. 2018.
[18] S. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in *Proc. ICML*, 2017, pp. 3319–3328.
[19] A. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Proc. KDD*, 2016, pp. 1135–1144.
[20] Lavanya Kalidoss, Swapna Thouti, Rajesh Ar unachalam, Pugalenthi Ramamurthy, "An efficient model of enhanced optimization and dilated-GRU based secured multi-access edge computing with blockchain for VANET sector," Expert Systems with Applications, Vol.260,125275, 2025, https://doi.org/10.1016/j.eswa.2024.125275
[21] Z. Zhang, Y. Wang, and M. R. Walter, "Exploring hybrid models for image

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

processing," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 3578–3589, May 2021.

[22] T. Zhang, J. D. L. et al., "A deep learning approach for real-time image processing," *IEEE Trans. Neural Networks*, vol. 35, pp. 2635–2648, 2020.

[23] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," *Proc. NIPS*, vol. 25, pp. 1097–1105, Dec. 2012.

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. ICLR*, 2015.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR*, 2016, pp. 770–778.

[26] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. CVPR*, 2017, pp. 4700–4708.

[27] P.Vasanth Kumar, A.R.Revathi, G.Ramya Devi, R.J.Kavitha, A.Muniappan, C.Karthikeyan., "Improved wild horse optimizer with deep learning enabled battery management system for internet of things based hybrid electric vehicles," Sustainable Energy Technologies and Assessments, vol.52, no.11, 102281, 2022, https://doi.org/10.1016/j.seta.2022.102281

[28] X. Zhang, J. Zou, and X. Liu, "Efficient convolutional networks for large scale image classification," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1204–1213, Mar. 2017.

[29] M. Le, S. Raj, and R. Kumar, "Automated detection of brain tumors from MRI scans using convolutional neural networks," *Med. Image Anal.*, vol. 34, pp. 68–75, Dec. 2017.

[30] Chaoyang Zhu, Yassine Bouteraa, Mohammad Khishke, Diego Martin, Francisco Hernando-Gallego, Thavavel Vaiyapuri, "Enhancing unmanned marine vehicle path planning: A fractal-enhanced chaotic grey wolf and differential evolution approach," Knowledge Based Systems, Vol.317,113481,2025, https://doi.org/10.1016/j.knosys.2025.113481

[31] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative adversarial nets," in *Proc. NIPS*, 2014, pp. 2672–2680.

[32] C. Ledig, L. Theis, F. Huszár, et al., "Photo-realistic single image super-resolution using a generative adversarial network," *Proc. CVPR*, 2017, pp. 105–114.

[33] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. CVPR*, 2017, pp. 1125–1134.

[34] L. Yi, R. P. Ratner, and Z. Zha, "GANs for medical image synthesis," *J. Med. Imaging*, vol. 6, no. 2, pp. 98–105, Apr. 2019.

[35] V. Mnih, H. Kavukcuoglu, D. Silver, et al., "Playing Atari with deep reinforcement learning," in *Proc. NIPS*, 2013, pp. 1–9.

[36] V. Mnih, H. V. Hasselt, D. Silver, and D. H. Pomerleau, "Asynchronous methods for deep reinforcement learning," in *Proc. ICML*, 2016, pp. 1928–1937.

[37] X. Pan, X. Zhang, and M. Chen, "Autonomous driving using deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8997–9006, Sept. 2018.

[38] J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[39] R. Vinyals, C. Blundell, T. Lillicrap, et al., "Matching networks for one shot learning," in *Proc. NeurIPS*, 2016, pp. 3630–3638.

[40] S. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," in *Proc. ICML*, 2017, pp. 3319–3328.

[41] D. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Proc. KDD*, 2016, pp. 1135–1144.

[42] L. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in NLP," in *Proc. ACL*, 2019, pp. 3645–3650.