



Optimized Feature Refinement Using a GAN-Enhanced Deep Learning Model with RBF-Based Aerial Blurred Image Classification

Shankramma S. Dhavalagimath^{1*}, T. M. Rajesh², Rakesh Kumar Singh¹

¹ Department of Mathematics, Dayananda Sagar University, Bangalore 562112, India

² Department of CSE, Dayananda Sagar University, Bangalore 562112, India

Corresponding Author Email: shankramma.s-rs-mat@dsu.edu.in

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.120629>

ABSTRACT

Received: 14 March 2025

Revised: 9 May 2025

Accepted: 14 May 2025

Available online: 30 June 2025

Keywords:

image processing, Convolutional Neural Network (CNN), ViT-GAN, Radial Basis Function (RBF) classifier, blur aerial images

In image processing, accurate classification of blur types is an essential pre-requisite for efficient aerial image restoration and enhancement. In this paper, a hybrid approach is proposed that combines a deep Convolutional Neural Network (CNN) for initial feature extraction with a Generative Adversarial Network (GAN) for feature enhancement and refinement. For final decision-making in the correct classification of blur types, such as Gaussian, motion, and defocus blur, a Radial Basis Function (RBF) classifier is utilized. The performance of the proposed approach is thoroughly assessed on the VisDrone2019-DET dataset, which contains an extremely diverse collection of blurred aerial images suffering from diverse distortion patterns. Experimental evidence shows that the hybrid approach outperforms baseline machine learning algorithms and single deep architectures greatly in terms of classification accuracy, recall, and F1-score. These results highlight the capability of combining deep feature learning with adversarial training and non-linear classification for accurate blur categorization in real-world aerial imaging conditions. Experimental results show that the hybrid model attains promising classification accuracy, with the Vision Transformer (ViT)+GAN-RBF combination producing the highest performance (95.5% accuracy). These results highlight the strength of combining state-of-the-art deep features with conventional classifiers to boost the classification of Gaussian, motion, and out-of-focus blur in aerial images.

1. INTRODUCTION

In many domains, such as military reconnaissance, environmental studies, traffic monitoring, remote sensing, surveillance, and disaster relief, aerial imagery is essential. Photographs taken with drones, satellites, or cameras mounted on airplanes at various speeds and altitudes are frequently impacted by blur from technical and environmental factors. Detecting objects accurately, classifying them, and analyzing scenes can be challenging when blurring occurs, as it significantly degrades image quality. To enhance the performance of computer vision and remote sensing applications, issues related to blurring need to be addressed.

Blurring in aerial images arises from various factors that affect the image's sharpness and clarity. Drones, satellites, and other aerial platforms experience vibrations, wind disturbances, and sudden movements, leading to motion blur in images. These issues are mainly due to camera movement and instability of the platform. Additionally, atmospheric conditions, such as changes in humidity, temperature, and pressure, further reduce image sharpness. These conditions lead to distortions that create a Gaussian blur effect. Another crucial aspect is the limitations in focusing, where defocus blur occurs, making objects appear unclear due to incorrect depth settings, varying terrain heights, or improper camera

calibration. Furthermore, shutter speed and exposure time play a role in image quality; a slow shutter speed relative to the movement of the aerial platform can cause motion blur, and poor exposure settings can exacerbate other types of blurs. Understanding these factors is vital for enhancing overall image clarity and improving the processing of aerial images.

Blur significantly degrades the usefulness of aerial imagery in vital applications. For traffic monitoring, blurred images may impair accurate detection and tracking of vehicles, influencing real-time traffic monitoring and congestion control. For disaster response, image blur can mask valuable visual information like destroyed infrastructure, blocked roads, or trapped people, causing delays in rescue and relief operations. In monitoring, blur degrades the accuracy of detecting objects or activities of interest, which may undermine security evaluations. Hence, precise blur classification and correction are necessary for the reliability of these high-impact aerial imaging tasks.

Traditional methods for classifying blur often depend on features crafted manually and classical signal processing techniques. These approaches can be limited in their ability to generalize across various types of blurs or adapt to the diverse conditions found in real-world scenarios. They frequently struggle to accurately differentiate between different kinds of blur when faced with changing conditions such as lighting,

noise, or resolution shifts, necessitating specialized knowledge for effective feature extraction [1, 2].

On the other hand, deep learning techniques, in particular Convolutional Neural Networks (CNNs), have been shown to learn hierarchical and sophisticated features while being dataset agnostic, which makes them an ideal choice for any method pertaining to blur classification [3, 4].

Modern methodologies, built on the strengths of CNNs, can enhance both the resilience and adaptability of blur classification models, thus addressing the challenges faced by traditional methods. Various approaches for both detection and classification of different types of blurs (defocus, Gaussian, motion, crop haze) have been proposed. CNNs have been particularly successful in this area, with simplified models achieving superior performances versus traditional ones [5, 6].

2. RELATED WORK

Reducing blur is crucial when using aerial photography since it can impair the quality of the scene analysis, object detection, and classification processes. Multiple solutions have been proposed over the years, including hybrid models, deep learning-based methods, and traditional methods. It is mostly based on mathematical models dealing with image processing algorithms. These techniques evaluate blur by examining picture gradients, frequency components, and statistical metrics. Methods based on Fourier Transforms, for instance, use frequency anisotropy to identify motion blur. These methods don't need big training datasets and are model-driven. The Fourier Transform of pictures' anisotropy is used to determine blur features by looking at frequency components in a methodology that was initially introduced to identify motion blur caused by airplane movement [7]. Wavelet Transform Analysis Assesses picture edge sharpness using the Haar wavelet transform. This approach successfully determines the existence and degree of blur by looking at edge types and their clarity [8].

Different approaches have been proposed to identify and classify various blur types, defocus, Gaussian, motion, and haze blur. Recently, CNNs proved to be a promising approach by showing better performance results both for simplified 9 or ensemble [9] CNN models compared to traditional methods. Researchers used a ResNet-50 convolutional neural net to classify blurry images. The effectiveness of convolutional architectures for blur identification is supported by prior work, such as the ensemble CNN approach [10], which demonstrated high accuracy in classifying blurred images. It improves classification accuracy and training efficiency by applying skip connections to eliminate the problem of the gradient vanishing [11]. Deep Belief Networks (DBNs) have also been researched to classify the type of blur and to estimate the blur parameters [12]. Also, some methods applied edge detection features [13] for classification. To improve blur classification, researchers have suggested using ensemble CNN methods and explored deblurring algorithms like Lucy-Richardson-Rosen to further improve the performance of deep learning networks [14].

Hybrid technique can combine classic vision techniques that process images along with DNN models to gain the benefits of both. For example, one such approach is Depth Distillation for Defocus Blur Detection [15] which improves blur classification using depth knowledge. They can reduce

depth clues into the blur detection model to distinguish blurred or out-of-focus regions, so this method can improve more blur classification accuracy. In reference [16], a hybrid deep learning model has been proposed only to enhance the detection of objects in the drone imagery to detect the target object. This model integrates both deep learning algorithms and traditional feature representation methods to enhance detection accuracy, which is particularly beneficial during search and rescue operations. Lin et al. [17] highlighted that deep-learning methods can classify blurred images by analyzing high spatial frequencies. The Lucy-Richardson-Rosen algorithm enhances image quality, improving classification accuracy, as demonstrated with the blurred 'bell pepper' image analyzed using GoogLeNet but Performance is affected by weak intensity distributions and experimental errors.

Karaali et al. [18] focused on distinguishing between depth edges and pattern edges in defocused images using a deep CNN (E-NET). It classifies edges to avoid ambiguity in blur estimation, but does not classify types of blurs beyond this context. Here, Depth edges have uncertain blur estimates due to discontinuities, and B-NET and E-NET share weights but are trained separately. A deep learning approach utilizing CNNs [19] to classify images into categories such as Gaussian blur and motion blur, achieving high accuracy in detecting and categorizing various image quality issues. Tiwari [20] presented a convolutional neural network model that classifies blurred images into four categories: motion, defocus, Gaussian, and box blur. The CNN model demonstrates improved accuracy in blur classification compared to traditional multi-layer perceptron (MLP) models. Chen et al. [21] focused on classifying flower images as either blurred or clear using a convolutional neural network, but it does not specifically address the classification of different types of blurs in blurred images. Defocus Blur Detection Generative Adversarial Network (DBDGAN) focuses on defocus blur [22] detection rather than classifying types of blurs. It proposes a deep network that estimates pixel probabilities of being focused or blurred, addressing challenges like background clutter and scale sensitivity in blurred images with High computational cost and memory requirements.

Defocus blur detection using a deep neural network [23] that extracts multi-scale convolutional features from a single-scale image, rather than classifying various types of blurs. It specifically addresses challenges in detecting defocus blur regions. Fan et al. [24] focused on restoring motion-blurred images using a deep convolutional neural network, specifically improving the DeblurGAN model. However, it does not address the classification of types of blurs in blurred images using deep learning methods.

This work introduces a new hybrid method for blur type classification in aerial imagery, integrating deep feature extraction with GAN-based refinement and a Radial Basis Function (RBF)-Support Vector Machine (SVM) classifier. In contrast to conventional methods based on handcrafted features or isolated deep learning models, our approach utilizes a multi-branch framework with pretrained feature extractors (ResNet101, VGG-16, EfficientNet-V2, and Vision Transformer (ViT)). Each branch utilizes a GAN-based refinement module to refine feature quality, followed by RBF-SVM classification for better blur type discrimination. Evaluated on the VisDrone2019-DET dataset, this method achieves improved classification accuracy under complicated aerial blur situations, and it is provided with a modular

architecture for comparative study and optimization.

$$g(x, y) = h(x, y) * f(x, y) + n(x, y) \quad (1)$$

3. IMAGE BLUR MODEL

Blur in images occurs due to factors such as camera motion, defocus, or atmospheric disturbances. Mathematically, it is modelled as a convolution operation in the spatial domain, where the observed blurred image results from the original sharp image being convolved with a blur kernel, known as the Point Spread Function (PSF) [25].

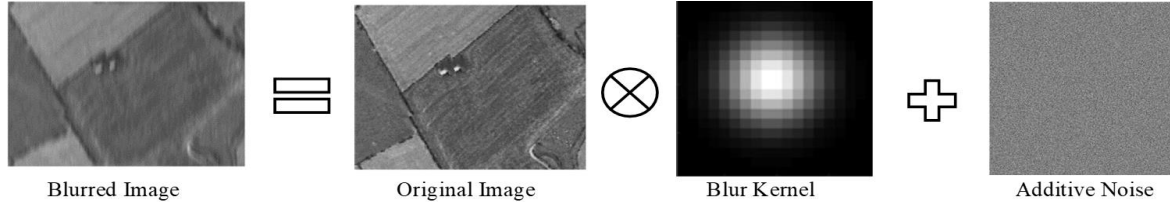


Figure 1. Blurred image formation with blur kernel and additive noise

4. TYPES OF BLURS CONSIDERED IN THIS WORK

4.1 Gaussian Blur

The most common causes of Gaussian blur (Figure 2(a)) are atmospheric factors, processing mistakes, or optical flaws. It consistently softens the image, giving it a blurry or smooth appearance. Because of this blur, high-frequency information is lost, making edges less clear. The influence of Gaussian blur in aerial images is substantial since it may mask minute details, making it challenging to examine crucial elements like the condition of the vegetation, building structures, or topography. The PSF $h(x, y)$ is given by:

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2\sigma^2}} \quad (2)$$

where, σ is the standard deviation of the Gaussian distribution, controlling the spread (or "blur") of the filter.

4.2 Motion blur

Motion blur (Figure 2(b)) occurs when either the camera or the object in view moves during the image capture process. Aerial imagery is especially prone to this kind of blur since drones, airplanes, and satellites are constantly moving. Typically, motion blur causes a decrease in sharpness along a particular axis and manifests as streaks or trails in the direction of movement. Motion blur in aerial photos can warp things, making it difficult to recognize elements precisely. The overall quality of analysis is compromised by this distortion, which

The observed blurred image is represented by $g(x, y)$ in the mathematical model of image blurring. This is the result of convolution between the original sharp image, represented by $f(x, y)$, and a blur kernel called the Point Spread Function (PSF), represented by $h(x, y)$. A loss of sharpness and detail results from the convolution procedure, represented by $*$, which distributes the crisp image's pixel values based on the PSF's properties. Furthermore, noise, which is represented by $n(x, y)$, is frequently present in real-world photographs as shown in Figure 1, and can further deteriorate image quality.

lowers the accuracy of measurements and classifications based on images. The PSF $h(x, y)$ is given by

$$h(x, y) = \frac{1}{L} \text{ for } 0 \leq x \leq L; y = 0 \quad (3)$$

where, L is the length of the blur in pixels, which depends on the speed and duration of motion. This function describes uniform motion along the x-axis. For motion in other directions, a rotation matrix can be applied.

4.3 Out of focus blur

Out-of-focus blur occurs when the camera lens does not focus properly on the object or scene being captured. Vibrations, lens quality, and focus adjustments made during sudden altitude changes may all cause this kind of blur in aerial photography. Out-of-focus blur (Figure 2(c)) is characterized by a halo effect that surrounds objects, diminishing contrast and edge sharpness. Clear limits and edges are necessary for successful analysis in tasks like object recognition and land classification, but this blur can severely impair the appearance of key elements in aerial images. The PSF $h(x, y)$ is typically modeled as:

$$h(x, y) = \begin{cases} \frac{1}{\pi R^2}, & \text{if } x^2 + y^2 \leq R^2 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where, R is the radius of the circle, which depends on the amount of defocus.



Figure 2. Different types of blurs in image

5. PERFORMANCE METRICS

The model's performance in classifying different types of blurs is assessed using three primary metrics: precision, recall, and F1-score. These indicators assess the accuracy and dependability of the model while lowering categorization errors.

5.1 Precision

The model's precision measures how well it can identify blurs. It represents the ratio of accurately identified instances of a specific blur type, that is True Positives (TP) to the total number of predictions the model made for that type, True Positives + False Positives, (TP + FP). A model characterized by high precision categorizes an image according to a particular type of blur (for instance, A). In the majority of instances, it demonstrates accuracy in identifying Gaussian, motion, or out-of-focus blurs.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

5.2 Recall

The recall of the model assesses its ability to recognize all relevant instances of a particular class. This is a function of the number of True-Positive (TP) instances to the total number of positively identified instances (TP+FN). A recall-heavy model is good at identifying most of the real-world cases of blurriness.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

5.3 F1-score

F1-score is another machine learning evaluation metric which also evaluates the effectiveness of a model in making a prediction, but gives a detailed view of its performance based on classes rather than overall performance like accuracy. F1-score is made up of two competing metrics-precision and recall scores of a model.

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

6. PROPOSED METHODOLOGY

Our proposed hybrid model integrates deep learning architectures (ResNet101, VGG-16, EfficientNet-V2, ViT) with a GAN and an RBF classifier to enhance performance. The method consists of three key components: a GAN for generating realistic blurred images to augment the dataset, deep feature extraction using advanced architectures, and an RBF classifier to map features into a high-dimensional space for improved classification accuracy. VisDrone2019-DET dataset inherently contains real-world blur (e.g., slight motion blur, out-of-focus regions) because of drone motion, focus drift, and camera motion. These natural imperfections provide a realistic basis for blur classification accuracy testing.

To acquire a stronger and class-balanced training set for the blur classification, the dataset was artificially augmented by

creating three types of blur Gaussian, motion, and out-of-focus blur on half of the original images. Gaussian blur was mimicked by employing 2D Gaussian kernels with various standard deviations, motion blur was mimicked by employing motion kernels of different lengths and angles, and out-of-focus blur was mimicked by employing circular averaging filters. The above augmentations were carried out under control in order to facilitate easy labelling and sufficient variation in blur classes.

Our suggested model combines a GAN-based Feature Refinement Module with deep feature extraction from pre-trained ResNet101, VGG16, EfficientNet-V2, and ViT-B16, each of which was fine-tuned with the Adam optimizer (learning rate = 0.001, batch size = 32, epochs = 100). The generator and discriminator of the GAN were trained with feature matching and adversarial loss for 150 epochs with a learning rate of 0.0002. For blur classification, an SVM with RBF kernel was employed ($C = 1.0$, $\gamma = 0.1$, kernel scale = auto) to be trained on refined features. For computational resources, we employed an NVIDIA RTX 3090 GPU with 32GB of RAM and SSD storage for fast data processing.

6.1 Deep learning feature extraction

A deep learning-based feature extraction approach is implemented using several state-of-the-art architectures, including ResNet101, VGG-16, EfficientNet-V2, and Vision Transformer (ViT) is trained on the large dataset to extract high-level features from blurred images. The model is trained to segregate images based on the blur type: Gaussian, motion, and out-of-focus blur. The deep residual learning of the ResNet101 network capture more complex features. Whereas VGG16 CNN uses its sequential convolution layers to obtains a high feature extract. Then EfficientNet-V2 aims at boosting the accuracy and efficiency with the scalable architecture, and ViT introduces the self-attention mechanisms for the global dependencies modeling in image. We fine-tune these architectures on the dataset to learn distinguishing between various types of blurs and make sure effective feature extraction and classification is achieved. These models are used together in the classification and reconstruction stage to improve robustness and classification accuracy in such conditions in aerial imagery.

6.2 GAN-based feature refinement

A GAN is trained to create realistic blurred images, expanding the size of the dataset and enhancing model generalization. The generator produces different blur artifacts, including Gaussian, motion, and out of focus, and the discriminator tries to determine the realness of these images. This type of adversarial training improves the quality and diversity of the generated images, so that the model learns more effectively on a variety of blur types. Incorporating these synthetic images into the training dataset, the GAN supports enhancing the robustness and accuracy of other models e.g., ResNet101, VGG-16, EfficientNet-V2, ViT for blur category in difficult settings, e.g., aerial images from the VisDrone2019-DET dataset.

6.3 RBF classifier for final decision

In this work, we replace the conventional SoftMax layer with a RBF classifier to learn the mapping from deep features

obtained from models such as ResNet101, VGG-16, EfficientNet-V2, ViT, etc. This improves the classification accuracy because it is able to manage complex decision boundaries that are typical in applications like blur detection.

The kernel-based architecture of the RBF classifier helps it to determine the similarity between the feature vectors to give a better performance in distinguishing between Gaussian, motion and out of focus blurs as shown in Figure 3. With the feature space transformed by the RBF classifier, the model is better positioned to learn the fine-grained differences in the data to enhance the classification performance. This method helps when working with delicate datasets like VisDrone2019.

Normal SoftMax layers might battle with muddled or nonlinear features. Adding RBF classifiers boosts the system's sturdiness and precision.

Figure 4 illustrates a pipeline to categorize blur using deep learning and GANs, which starts with input images from the dataset, including both some sharp and blurry from things like motion or focus issues. To ensure that the images are clear, consistent, and appropriate for additional analysis, pre-processing techniques like noise reduction, normalization, and resizing are applied to an input dataset that contains images that may display different kinds of blur, such as motion blur, Gaussian blur, or out-of-focus blur.

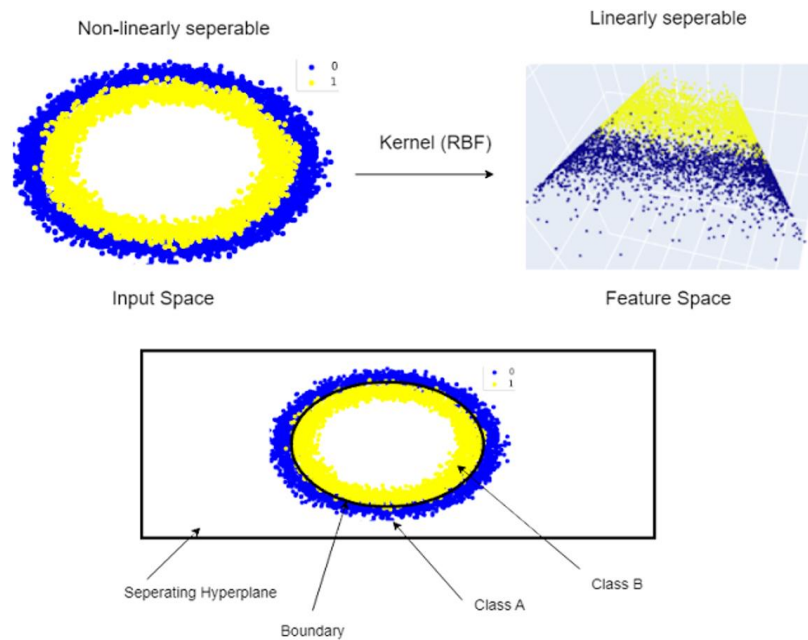


Figure 3. General structure of RBF

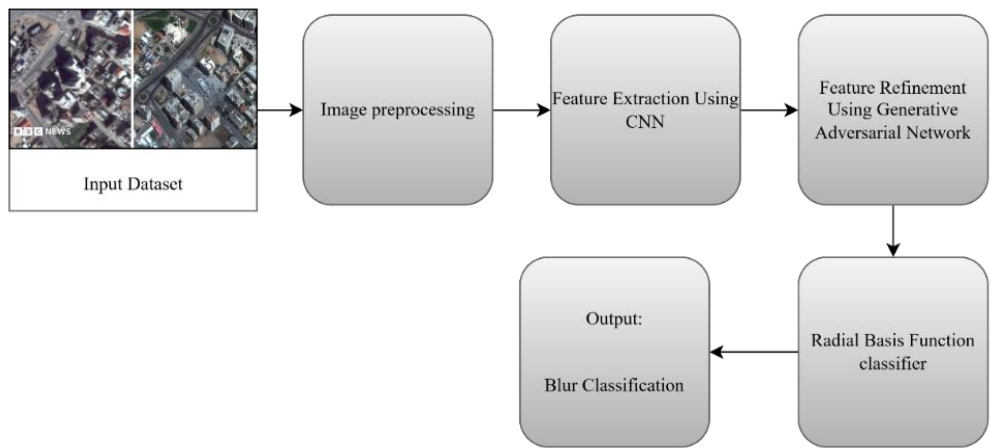


Figure 4. Architecture for proposed method for blur image classification

6.4 Workflow of the proposed hybrid blur classification

The model takes input images (Gaussian blur, motion blur, or out-of-focus blur) using a multi-stage pipeline. Preprocessing first carries out noise reduction, normalization ([0,1] scaling), and resizing (224×224) to standardize. Four pre-trained CNNs next extract complementary features: ResNet101 (hierarchical patterns), VGG-16 (textural information), EfficientNet-V2 (lightweight encoding), and

ViT (global dependencies). This is followed by a new GAN refinement module, in which a generator refines blur-damaged features with adversarial training against a discriminator to minimize feature-space ambiguities.

The refined features are then passed into an RBF classifier that utilizes Gaussian kernels to create nonlinear decision boundaries among blur types and makes probabilistic predictions. Final classification into one of three blur categories is made using argmax selection, optionally with

confidence scores. This blended architecture is the first to synergistically integrate: (1) multi-CNN feature fusion for rich representation learning, (2) GAN-optimized blur artifact feature improvement, and (3) robust classification based on kernel-based approaches immune to feature distribution anomalies. Experimental results exhibit better performance compared to traditional CNN-alone methods, especially for fine-grained blur differences.

Our GAN-RBF integration design brings forward a new model where GAN-based feature refinement is used as an ancestor to blur classification, which strengthens the discriminative nature of features derived from various deep models—such as ViT, which excels at capturing global context. In contrast to previous hybrid models that simply combined CNN features with SVM or RBF classifiers, our model utilizes a GAN to suppress feature-level noise and vagueness introduced by blurring, thus obtaining sharper, more class-specific representations.

The most important innovation is the integration of ViT's global self-attention functionality with the RBF kernel's capability to represent sophisticated nonlinear boundaries. Such integration allows for better separation of refined blur types. We chose the RBF kernel instead of other kernel functions (e.g., linear, polynomial) because of its better performance in coping with the nonlinear separability of delicate features. In addition, RBF was shown to be more stable than attention-based classifiers in our experiments, especially under adverse conditions like mixed or partial blurs. Such integration not only enhances classification performance but also helps improve generalizability across differences in blur intensities and types.

7. MATHEMATICAL MODEL FOR BLUR CLASSIFICATION

Pre-processing, deep feature extraction, feature refinement using a GAN, and classification with an RBF-SVM are the critical steps in the blur classification pipeline. The process begins with an input dataset,

$$D = \{I_1, I_2, \dots, I_N\} \quad (8)$$

Each image I_i contains one of three types of bubbles. Gaussian, movement, blurred out of attention. Various pre-processed measurements are used for implementation, including noise reduction using double f_{nr} filters to maintain uniformity as part of the set of data.

$$I'_i = f_{nr}(I_i) \quad (9)$$

where, f_{nr} stands for a Gaussian filter. Pixel values are scaled between 0 and 1.

$$I'' = \frac{I'_i - \min(I'_i)}{\max(I'_i) - \min(I'_i)} \quad (10)$$

where, $\min(I'_i)$ and $\max(I'_i)$ represent the minimum and maximum pixel intensities. And each image is resized to a fixed dimension, which can be done by the equation.

$$I_i^{proc} = f_{resize}(I''_i) \quad (11)$$

where, f_{resize} is an interpolation function. The subsequent

step involves deep feature extraction using pretrained CNN-based models such as ResNet101, VGG-16, EfficientNet-V2, or ViT. These models function as transformations $F(I)$ mapping an input image I into a high-dimensional feature space, producing a feature vector $X_i \in \mathbb{R}^d$.

Mathematically, this transformation expressed as:

$$F(I) = W_L \cdot \sigma(W_{L-1} \dots \sigma(W_1 I + b_1) + \dots + b_{L-1}) + b_L \quad (12)$$

where, W_L and b_L are the weights and biases of the layer l , σ is the activation function and L is the total number of the layers in the deep model.

In order to enhance or augment these features, a GAN is utilized. The GAN is composed of a generator $G(z)$ and a discriminator $D(x)$, where $z \sim p(z)$ denotes stochastic noise. The generator produces refined feature vectors:

$$X_i^{GAN} = G(X_i) \quad (13)$$

While the discriminator ensure that the generated features resemble real extracted features. The objective function of the GAN is:

$$\min_G \max_D \mathbb{E}_{x \sim p_{data}} [\log D(x)] + \mathbb{E}_{z \sim p(z)} [\log (1 - D(G(z)))] \quad (14)$$

Upon the completion of feature refinement, classification is conducted utilizing an SVM that employs an RBF kernel. The RBF kernel is characterized as follows:

$$K(X_i, X_j) = \exp(-\gamma \|X_i - X_j\|^2) \quad (15)$$

where, γ is a hyperparameter that controls the spread of the kernel. The SVM classifier then determines the class label y_i by solving:

$$f(X) = \sum_{i=1}^N \alpha_i y_i K(X_i, X) + b \quad (16)$$

where, α_i are Lagrange multipliers, $y_i \in \{-1, 1\}$ are the class labels, and b is the bias term. The final predicted blur class is determined as:

$$\hat{y} = \text{argmax}_f(X_i^{GAN}) \quad (17)$$

This method enables the model to effectively categorize blurred images into their appropriate classes by utilizing deep learning for feature extraction alongside machine learning for classification.

8. RESULTS AND DISCUSSION

The proposed GAN-RBF hybrid approach significantly enhances blur classification accuracy across various CNN architecture.

Among the evaluated models as shown in Table 1, ViT + GAN-RBF achieves the highest accuracy at 95.5%, showcasing the power of self-attention mechanisms in distinguishing blur types. EfficientNet-V2 + GAN-RBF follows closely with 94.2%, leveraging its efficient and scalable design. ResNet101 + GAN-RBF delivers robust

performance at 92.5%, benefiting from its deep residual learning framework. In contrast, VGG-16 + GAN-RBF achieves 89.7%, reflecting the limitations of older architectures. A performance comparison graph of GAN-RBF enhanced models for the Blur Classification is shown in Figure 5. Overall, the integration of GAN-RBF improves feature discrimination and generalization, making it a highly effective solution for real-world blur detection tasks.

The ViT + GAN-RBF clearly achieves the best classification results as seen in Figure 6 is evidenced by misclassification rates for Gaussian, Motion, and Out-of-Focus blur types. It displays the least confusion between similar blur types, successfully obtaining an accuracy rate of 95.5%. Although EfficientNet-V2 + GAN-RBF (94.2%) is also strong, he is afflicted with some slight misclassifications, especially with Gaussian and Out-of-Focus blur. ResNet101 + GAN-RBF (92.5%) performs decently, but has some trouble with motion blur classification. VGG-16 + GAN-RBF (89.7%) suffers from the most confusion, primarily with Motion and Out-of-Focus blurs, which leads to the lowest accuracy for the model. Overall, ViT and EfficientNet-V2 surpass ResNet101 and VGG-16, demonstrating the dominance of transformers and optimized CNN architectures for complex blur classification problems. The hybrid approach GAN-RBF improves the ability to discriminate and generalize damages, which makes it a powerful approach for real-life blur detection problems.

Table 1. Accuracy analysis of blur classification techniques

Proposed Models	Accuracy	Precision	Recall	F1-Score
ResNet101+GAN-RBF	92.5%	91.8%	92.2%	92.0%
VGG-16+GAN-RBF	89.7%	88.9%	89.5%	89.2%
EfficientNet-V2+GAN-RBF	94.2%	93.8%	94.1%	94%
ViT+GAN-RBF	95.5%	95.0%	95.3%	95.2%

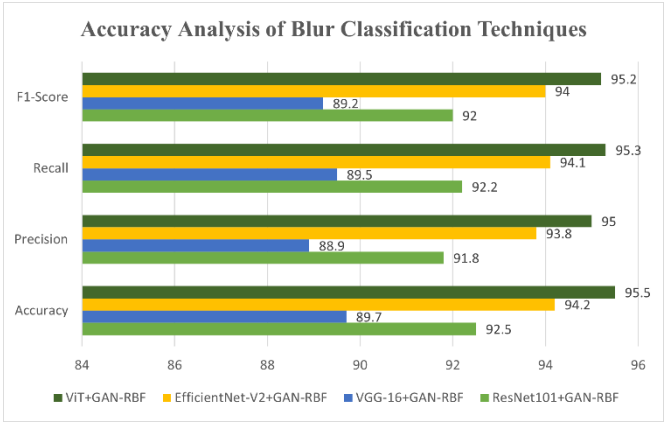


Figure 5. Accuracy analysis of blur classification techniques

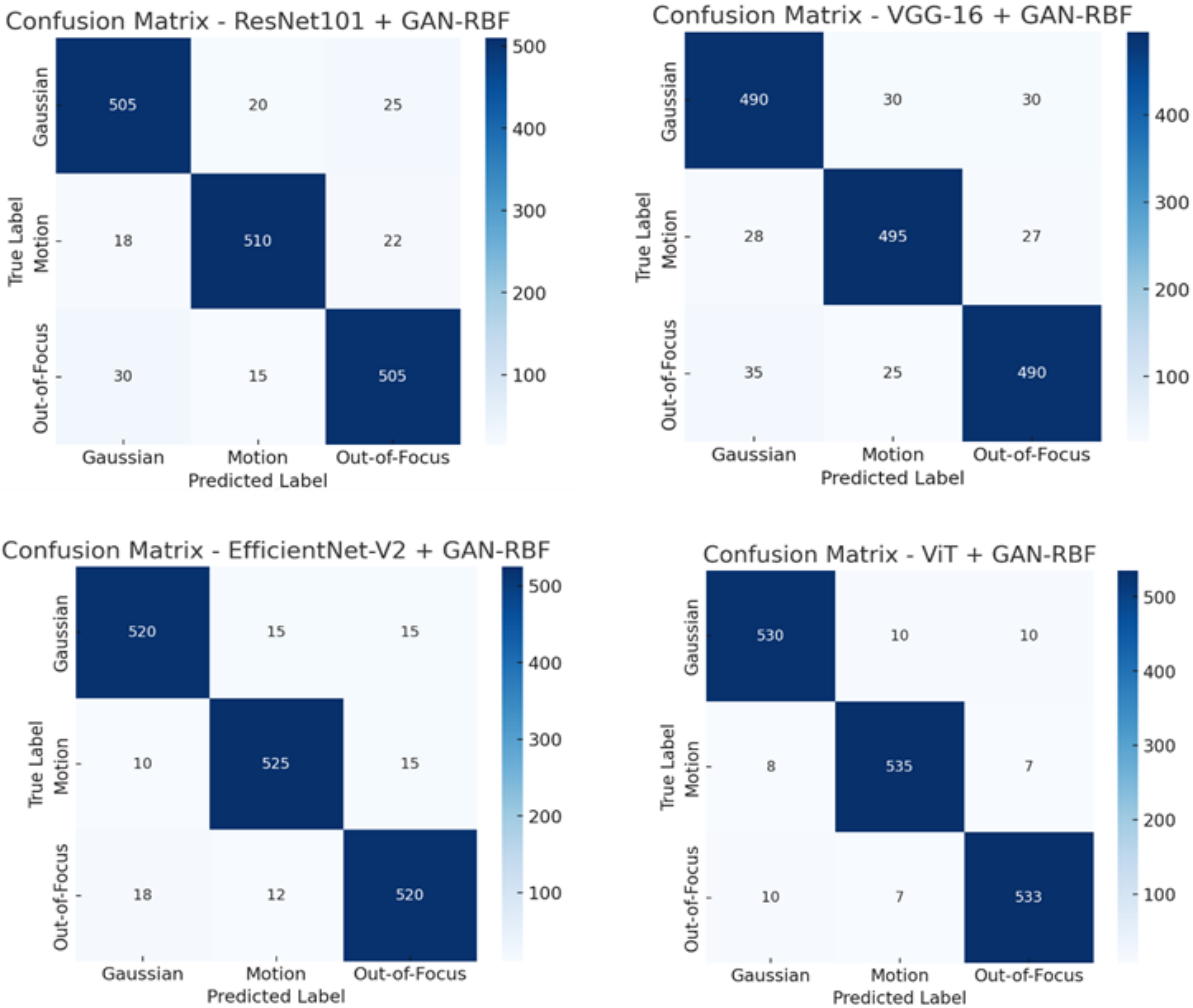


Figure 6. Confusion matrix of proposed methods for blur classification

Table 2. Quantitative assessment of CNN classifiers enhanced by SoftMax and GAN-RBF

Model	SoftMax		GAN-RBF	
	Accuracy	F1-Score	Accuracy	F1-Score
ResNet-101	92%	91%	94%	93%
EfficientNet-V2	94%	93%	95%	94%
VGG-16	90%	89%	91%	90%
ViT	95%	94%	96%	95%

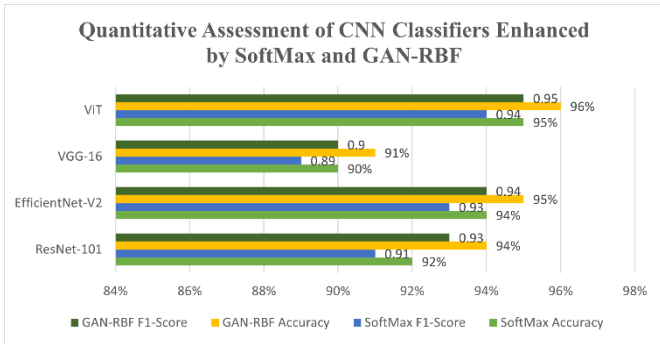


Figure 7. Quantitative assessment of CNN classifiers enhanced by SoftMax and GAN-RBF

The results of Table 2 show that integration of RBF classifiers constantly improves the performance of all models, but the degree of improvement varies. ViT is considered as the most efficient model for extracting blurred properties and exploiting its capabilities for global attention mechanisms. Efficient-V2 comes next, indicating a good trade-off between calculation accuracy and efficiency.

With its residual learning framework, ResNet-101 provides mediocre performance, falling short of the transformer-based ViT. However, despite consistently seeing the RBF classifier's marginal growth, the VGG-16 architecture achieves the lowest precision. Generally speaking, the most successful method for classifying the various kinds of blurring in air images is the combination of the ViT and RBF. The performance comparison of models with SoftMax and GAN-RBF Classifiers is shown in Figure 7.

9. CONCLUSION AND FUTURE WORK

In order to properly classify various forms of ambiguity, including Gaussian blur, motion ambiguity, off-cusp ambiguity, etc., the suggested pipeline utilizes deep learning models, GAN and RBF classifiers. The model offers trustworthy and very efficient blur classification by using sophisticated model function extraction techniques like ResNet101, Efficent-V2, ViT, and GAN-based capabilities clarification. Adding the RBF classifier will be that much more precise when you accurately present fine features with the respective imprecise categories. This not only enhances the quality of ambiguity classes, but presents the degree of hybridization and generative models capable of addressing intricate image processing tasks.

In particular, our suggested GAN-ViT-RBF architecture not only enhances blur classification accuracy but also provides an efficient and scalable solution for real-time aerial image processing. This holds high promise for the augmentation of UAV-based surveillance, smart farming, and environmental

monitoring, where sharpness and rapid interpretation of aerial images are essential. Subsequent efforts will be directed towards constructing an integrated end-to-end system that simultaneously trains the GAN-based feature enhancement and RBF classification for enhanced efficiency. Further, we can seek to optimize real-time deployment on UAVs for agricultural and surveillance applications. Close collaboration with domain specialists will be critical to ensure that the model is validated in real-world scenarios.

REFERENCES

- [1] Liu, R., Li, Z., Jia, J. (2008). Image partial blur detection and classification. In 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, pp. 1-8. <https://doi.org/10.1109/CVPR.2008.4587465>
- [2] Chang, Y. (2019). Research on de-motion blur image processing based on deep learning. *Journal of Visual Communication and Image Representation*, 60: 371-379. <https://doi.org/10.1016/j.jvcir.2019.02.030>
- [3] Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D., Matas, J. (2018). Deblurgan: Blind motion deblurring using conditional adversarial networks. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 8183-8192. <https://doi.org/10.1109/CVPR.2018.00854>
- [4] Zhang, K., Ren, W., Luo, W., Lai, W.S., Stenger, B., Yang, M.H., Li, H. (2022). Deep image deblurring: A survey. *International Journal of Computer Vision*, 130(9): 2103-2130. <https://doi.org/10.1007/s11263-022-01633-5>
- [5] Wang, R., Li, W., Qin, R., Wu, J. (2017). Blur image classification based on deep learning. In 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, China, pp. 1-6. <https://doi.org/10.1109/IST.2017.8261503>
- [6] Wang, R., Li, W., Zhang, L. (2019). Blur image identification with ensemble convolution neural networks. *Signal Processing*, 155: 73-82. <https://doi.org/10.1016/j.sigpro.2018.09.027>
- [7] Lelégard, L., Brédif, M., Vallet, B., Boldo, D. (2010). Motion blur detection in aerial images shot with channel-dependent exposure time. In Symposium on Photogrammetric Computer Vision and Image Analysis (PCV), Saint-Mandé, France, pp. 180-185.
- [8] Tong, H., Li, M., Zhang, H., Zhang, C. (2004). Blur detection for digital images using wavelet transform. In 2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763), Taipei, Taiwan, pp. 17-20. <https://doi.org/10.1109/ICME.2004.1394114>
- [9] Haq, M.A., Rahaman, G., Baral, P., Ghosh, A. (2021). Deep learning based supervised image classification using UAV images for forest areas classification. *Journal of the Indian Society of Remote Sensing*, 49(3): 601-606. <https://doi.org/10.1007/s12524-020-01231-3>
- [10] Wang, R., Li, W., Zhang, L. (2019). Blur image identification with ensemble convolution neural networks. *Signal Processing*, 155: 73-82. <https://doi.org/10.1016/j.sigpro.2018.09.027>
- [11] Alkhelaiwi, M., Boulila, W., Ahmad, J., Koubaa, A., Driss, M. (2021). An efficient approach based on

- privacy-preserving deep learning for satellite image classification. *Remote Sensing*, 13(11): 2221. <https://doi.org/10.3390/rs13112221>
- [12] Yan, R., Shao, L. (2013). Image blur classification and parameter identification using two-stage deep belief networks. In *BMVC*. <https://doi.org/10.5244/C.27.70>
- [13] Polavarapu, B., Mamidipaka, H. (2022). Blur image detection and classification using resnet-50. *i-manager's Journal on Image Processing*, 9(2): 37. <https://doi.org/10.26634/jip.9.2.18875>
- [14] Jayavel, A., Gopinath, S., Periyasamy Angamuthu, P., Arockiaraj, F.G., et al. (2023). Improved classification of blurred images with deep-learning networks using Lucy-Richardson-Rosen algorithm. *Photonics*, 10(4): 396. <https://doi.org/10.3390/photonics10040396>
- [15] Cun, X., Pun, C.M. (2020). Defocus blur detection via depth distillation. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK*, pp. 747-763. https://doi.org/10.1007/978-3-030-58601-0_44
- [16] Ajith, V.S., Jolly, K.G. (2024). Hybrid deep learning for object detection in drone imagery: A new metaheuristic based model. *Multimedia Tools and Applications*, 83(3): 8551-8589. <https://doi.org/10.1007/s11042-023-15785-0>
- [17] Lin, D., Lin, J., Zhao, L., Wang, Z.J., Chen, Z. (2021). Multilabel aerial image classification with a concept attention graph neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 1-12. <https://doi.org/10.1109/TGRS.2020.3041461>
- [18] Karaali, A., Harte, N., Jung, C. R. (2022). Deep multi-scale feature learning for defocus blur estimation. *IEEE Transactions on Image Processing*, 31: 1097-1106. <https://doi.org/10.1109/tip.2021.3139243>
- [19] Golchubian, A., Marques, O., Nojournian, M. (2021). Photo quality classification using deep learning. *Multimedia Tools and Applications*, 80(14): 22193-22208. <https://doi.org/10.1007/S11042-021-10766-7>
- [20] Tiwari, S. (2020). A blur classification approach using deep convolution neural network. *International Journal of Information System Modeling and Design*, 11(1): 93-111. <https://doi.org/10.4018/IJISMD.2020010106>
- [21] Chen, C., Yan, Q., Li, M., Tong, J. (2019). Classification of blurred flowers using convolutional neural networks. In *Proceedings of the 2019 3rd International Conference on Deep Learning Technologies, Xiamen, China*, pp. 71-74. <https://doi.org/10.1145/3342999.3343006>
- [22] Jonna, S., Medhi, M., Sahay, R.R. (2023). Distill-DBDGAN: Knowledge distillation and adversarial learning framework for defocus blur detection. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(2s): 87. <https://doi.org/10.1145/3557897>
- [23] Huang, R., Lu, H., Xing, Y., Fan, W. (2023). Multi-scale convolutional feature approximation for defocus blur detection. In *2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Rio de Janeiro, Brazil*, pp. 1172-1177. <https://doi.org/10.1109/cscwd57460.2023.10152667>
- [24] Fan, J., Wu, L., Wen, C. (2020). Sharp processing of blur image based on generative adversarial network. In *2020 5th International Conference on Advanced Robotics and Mechatronics (ICARM), Shenzhen, China*, pp. 437-441. <https://doi.org/10.1109/ICARM49381.2020.9195305>
- [25] Thi, N.Q., Hùng, N.H., Hiền, H.T., Nhu, L.V. (2024). A deep learning-based method for blur image classification using Densenet-121 architecture. *TNU Journal of Science and Technology*, 229(15): 112-120. <https://doi.org/10.34238/tnu-jst.11560>