ISSN: 2583-7354

International Journal of Emerging Knowledge Studies

Publisher's Home Page: https://www.ijeks.com/

Fully Open Access

Research Paper

Skin Lesion Classification Using CAF-IDNet (Cross-Attention Fusion InceptionV3 and DenseNet201)

S. Kalyan[1*] A. Swapna[2] P . Lakshmi Prasanna[3] K. Ganesh Kumar[4] T. Likhitha[5]

[1]Department of Computer Science and Engineering (Data Science), Aditya Institute of Technology and Management, Tekkali, India.

[2]Department of Computer Science and Engineering (Data Science), Aditya Institute of Technology and Management, Tekkali, India.

[3]Department of Computer Science and Engineering (Data Science), Aditya Institute of Technology and Management, Tekkali, India.

[4]Department of Computer Science and Engineering (Data Science), Aditya Institute of Technology and Management, Tekkali, India.

[5]Department of Computer Science and Engineering (Data Science), Aditya Institute of Technology and Management, Tekkali, India.

DOI: https://doi.org/10.70333/ijeks-04-10-043

***Corresponding Author:** kalyan826@gmail.com

Article Info:- **Received** : 07 August 2025 **Accepted** : 25 September 2025 **Published** : 30 September 2025

Abstract

Skin lesion classification is an important task in medical image analysis for the early detection of skin cancer and other dermatological diseases. Accurate classification of skin lesions is challenging due to variations in lesion shape, size, color, and the presence of noise such as hair and low contrast in dermoscopic images. To address these challenges, this study proposes a hybrid deep learning model called CAF-IDNet (Cross-Attention Fusion InceptionV3 and DenseNet201) for skin lesion classification. The proposed model combines the strengths of InceptionV3 and DenseNet201 to extract multi-scale and deep features from skin lesion images. A cross-attention fusion module is introduced to improve feature interaction and enhance important feature representation. The fused features are then passed to fully connected layers for classification. The proposed model is trained and tested on a skin lesion dataset, and its performance is evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. The experimental results show that the proposed CAF-IDNet model achieves higher classification accuracy compared to existing deep learning models. The results demonstrate that the proposed cross-attention fusion approach improves classification performance and reduces misclassification. Therefore, the proposed model can be used as an efficient computer-aided diagnosis system for skin lesion classification.

Keywords: *Skin Lesion Classification, Deep Learning, InceptionV3, DenseNet201, Cross-Attention, Feature Fusion.*

ISSN: 2583-7354

1. Introduction

Skin lesion classification is an essential task in medical image analysis, particularly for the early detection of skin cancer and other dermatological diseases. Early diagnosis of skin lesions such as melanoma, basal cell carcinoma, and benign keratosis can significantly improve patient survival rates. However, manual diagnosis by dermatologists is time-consuming, subjective, and dependent on clinical expertise. Therefore, automated skin lesion classification using deep learning techniques has become an important research area in medical image analysis.

In recent years, deep convolutional neural networks (CNNs) have demonstrated significant performance improvements in image classification tasks, including medical image analysis. CNN-based models such as DenseNet, Inception, ResNet, and EfficientNet have been widely used for disease classification due to their ability to automatically extract deep features from images **(Bansal et al., 2021; Srinidhi et al., 2021; Vishnoi et al., 2023).** Transfer learning approaches using pre-trained CNN models have further improved classification accuracy, especially when working with limited medical datasets **(Fan et al., 2022; Khan et al., 2022).**

Several researchers have proposed improved CNN architectures to enhance classification performance. For example, lightweight CNN models were introduced to reduce computational complexity while maintaining high **accuracy (Fu et al., 2022).** Multi-scale CNN and deformable convolution networks have also been proposed to improve feature extraction from complex background images **(Liu et al., 2024; Gao et al., 2023).** Additionally, image segmentation techniques such as DeepLabV3+ have been used to improve lesion detection and classification performance **(Ding et al., 2025).**

Feature fusion and hybrid deep learning models have recently gained attention because single CNN models may fail to capture both low-level and high-level features effectively. Feature fusion techniques combine features from multiple CNN models to improve classification performance **(Fan et al., 2022).** Hybrid models combining multiple architectures such as DenseNet, Inception, and Vision Transformers have shown improved performance in medical image classification tasks **(Jayaraman et al., 2025).** Furthermore, attention mechanisms and cross-attention fusion models have been introduced to improve feature representation by focusing on important regions of the image. Cross-attention models allow interaction between feature maps extracted from different networks, improving classification accuracy **(Wei et al., 2025; Xie & Zhao, 2024).** These methods demonstrate that combining multiple deep learning architectures with attention mechanisms can significantly enhance classification performance.

Despite these advancements, existing models still face challenges such as feature redundancy, poor feature fusion strategies, and insufficient attention mechanisms for effective feature interaction. Therefore, this research proposes a novel CAF-IDNet (Cross-Attention Fusion InceptionV3 and DenseNet201) model to improve skin lesion classification performance by combining the strengths of InceptionV3 and DenseNet201 with a cross-attention fusion mechanism.

Table 1: Summary of Existing Skin Lesion Classification Methods

Author	Year	Method	Model Used	Advantages	Limitations
Liu et al.	2018	CNN	Deep CNN	Good feature extraction	Overfitting
Jiang et al.	2019	CNN	Improved CNN	Real-time detection	Limited feature fusion
Yan et al.	2020	CNN	Improved CNN	Improved accuracy	High computation
Bansal et al.	2021	Deep Learning	CNN	Automatic feature extraction	Requires large dataset

Srinidhi et al.	2021	Transfer Learning	DenseNet, EfficientNet	High accuracy	Training time high
Fan et al.	2022	Feature Fusion	Transfer Learning + Fusion	Better performance	Complex architecture
Fu et al.	2022	Lightweight CNN	CNN	Reduced computation	Slight accuracy reduction
Khan et al.	2022	Deep Learning	CNN	Real-time detection	Less feature interaction
Vishnoi et al.	2023	CNN	CNN	High classification accuracy	No attention mechanism
Gao et al.	2023	Deep Learning	BAM-Net	Works in complex background	Complex model
Liu et al.	2024	Multi-scale CNN	MCDCNet	Multi-scale feature extraction	High computational cost
Ding et al.	2025	Segmentation + CNN	DeepLabV3+	Accurate lesion detection	High training time
Jayaraman et al.	2025	Hybrid Model	CNN + ViT	Improved performance	Complex architecture
Wei et al.	2025	Cross-Attention	Cross-Attention Network	Better feature fusion	Computational complexity
Xie & Zhao	2024	Dual-Branch	Cross-Attention	Improved feature interaction	Complex training

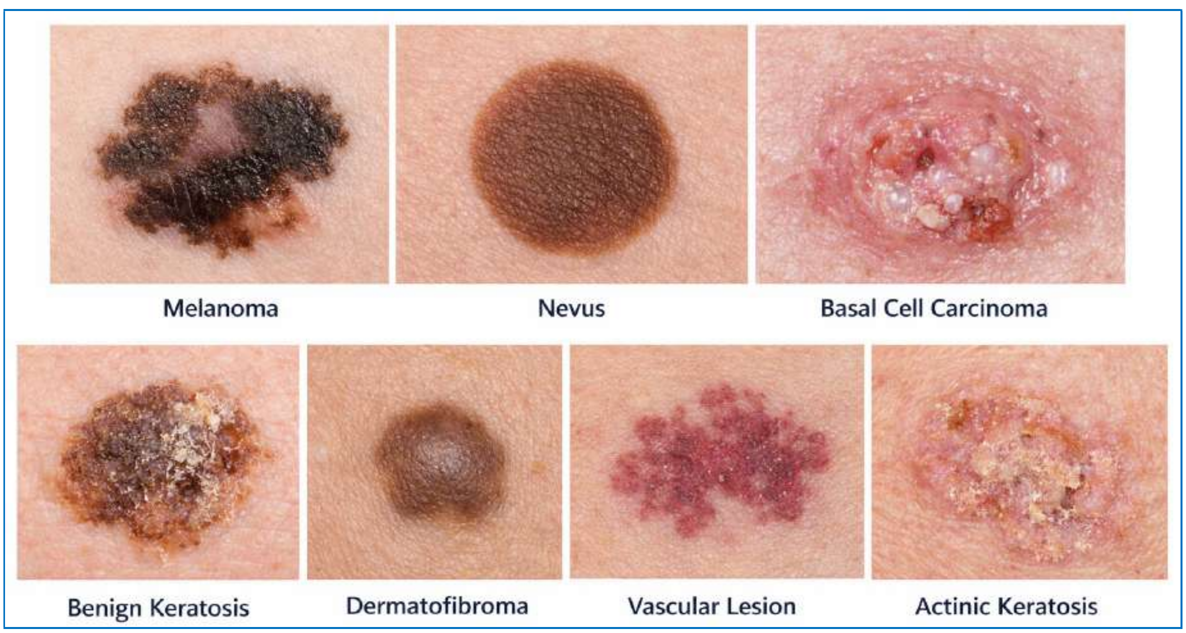

Figure 1: Sample Skin Lesion Images

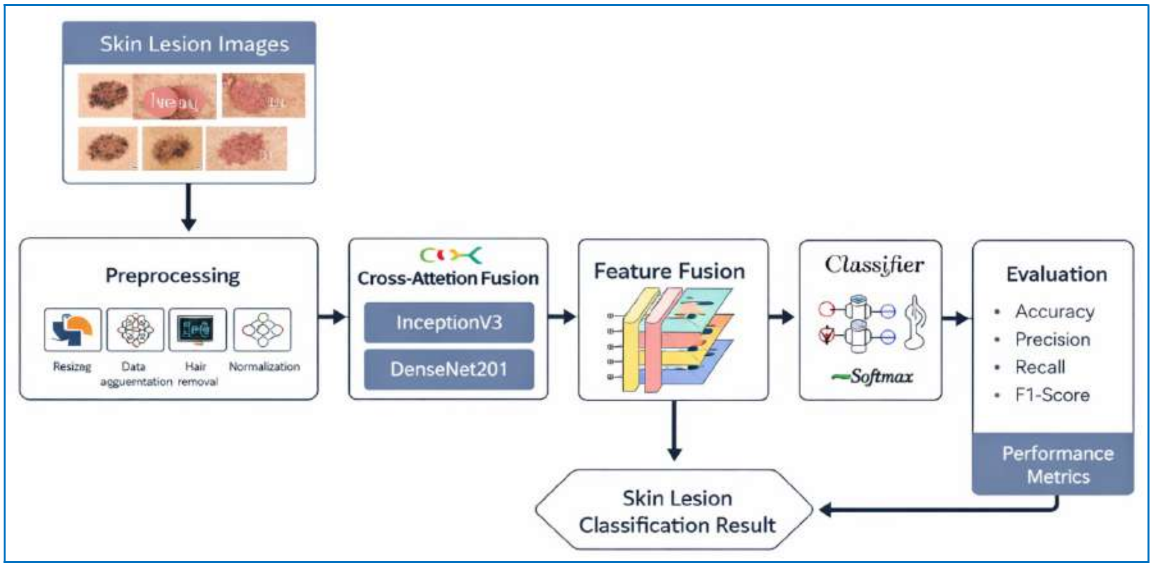

Figure 2: Overall Workflow of Skin Lesion Classification System

2. Problem Statement

Skin lesion classification is a critical task in medical image analysis for the early detection of skin cancer and other dermatological diseases. Accurate classification of skin lesions is challenging due to the high similarity between different lesion types, irregular lesion boundaries, variations in color and texture, presence of noise such as hair and shadows, and differences in image acquisition conditions. These challenges make manual diagnosis difficult and increase the chances of misclassification, which may lead to delayed treatment and increased mortality rates. Several deep learning models have been proposed for skin lesion and plant disease classification using convolutional neural networks (CNNs). Traditional CNN-based models such as standard deep CNN architectures provide good feature extraction but often suffer from overfitting and limited generalization when trained on small datasets **(Bansal et al., 2021; Liu et al., 2018).** Transfer learning models such as DenseNet and EfficientNet improve classification accuracy, but they still rely on single-network feature extraction, which may not capture all relevant features from the image **(Srinidhi et al., 2021).**

Recent studies introduced feature fusion and hybrid deep learning models to improve classification performance by combining features from multiple networks. Feature fusion techniques help improve classification accuracy by combining low-level and high-level features, but many existing fusion methods use simple concatenation or averaging methods, which may lead to redundant feature learning and inefficient feature representation **(Fan et al., 2022).** Additionally, multi-scale and lightweight CNN models were proposed to improve performance in complex background conditions, but these models often increase computational complexity **(Gao et al., 2023; Liu et al., 2024).**

Attention mechanisms and cross-attention models have recently been introduced to improve feature selection and focus on important regions in the image. Cross-attention fusion allows interaction between feature maps extracted from different deep learning models, improving classification performance **(Wei et al., 2025; Xie & Zhao, 2024).** However, most existing models do not effectively integrate cross-attention mechanisms with hybrid CNN architectures such as InceptionV3 and DenseNet201 for skin lesion classification.

3. Objectives of the Study

- To study existing skin lesion classification methods.
- To implement InceptionV3 and DenseNet201 models using transfer learning.
- To develop a hybrid model by combining InceptionV3 and DenseNet201.
- To design a cross-attention fusion module for feature fusion.
- To evaluate the performance of the proposed model using accuracy, precision, recall, and F1-score.

- To compare the proposed model with existing methods.

4. Research Questions

- RQ1: How does transfer learning using InceptionV3 and DenseNet201 perform in skin lesion classification?
- RQ2: Does combining InceptionV3 and DenseNet201 improve classification performance compared to single CNN models?
- RQ3: How does the proposed cross-attention fusion module improve feature fusion and classification accuracy?
- RQ4: How does the proposed CAF-IDNet model perform compared to existing deep learning models?
- RQ5: Can the proposed model reduce misclassification and improve overall performance in skin lesion classification?

5. Literature Review

Deep learning has become one of the most effective approaches for image classification because it can automatically learn discriminative features from raw images. In classification tasks involving complex visual patterns, convolutional neural networks (CNNs) have shown strong performance due to their ability to capture spatial and hierarchical features. Earlier CNN-based studies demonstrated that deep models can effectively identify disease-related patterns from image data, improving classification accuracy compared with traditional machine learning methods **(Liu et al., 2018; Jiang et al., 2019).**

Several studies later improved CNN performance by modifying architecture depth, feature extraction capability, and transfer learning strategies. Models based on DenseNet, EfficientNet, and other pre-trained architectures were found to improve classification performance because they reuse learned representations and reduce the dependence on very large datasets **(Bansal et al., 2021; Srinidhi et al., 2021; Mahato et al., 2022).** Likewise, lightweight CNN models were introduced to reduce computational cost while maintaining acceptable predictive performance, which is important for practical deployment **(Fu et al., 2022).**

To address limitations of single-network models, researchers introduced feature fusion and hybrid deep learning methods. Feature fusion combines complementary features extracted from different models, allowing the classifier to learn richer image representations. Studies using transfer learning with feature fusion reported improved performance over individual CNN models, showing that fused features can better capture both fine-grained and global image information **(Fan et al., 2022).** More recent work also explored hybrid frameworks that combine CNNs with transformer-based models, further improving classification by integrating local and contextual representations **(Jayaraman et al., 2025).**

Another important development is the use of attention mechanisms. Attention-based models help the network focus on the most informative regions of an image and suppress irrelevant background information. This is particularly useful when classification images contain noise, low contrast, or complex visual patterns. Cross-attention and dual-branch attention networks have shown promising performance by enabling interaction between feature maps from different branches or models **(Xie & Zhao, 2024; Wei et al., 2025).** These methods suggest that attention-guided fusion can improve feature representation quality and classification robustness.

Recent studies also proposed multi-scale and deformable convolution models to improve feature learning under complex backgrounds. Multi-scale networks can capture lesion or disease characteristics at different spatial resolutions, while deformable convolution helps adapt to irregular shapes and structures in images **(Gao et al., 2023; Liu et al., 2024).** Segmentation-based approaches such as DeepLabV3+ were also introduced to better isolate disease-affected regions before classification, improving the overall analysis pipeline **(Ding et al., 2025).**

Overall, the literature shows that classification performance improves when deep models move from single CNN architectures toward hybrid, feature-fusion, and attention-based designs. However, many existing methods still rely on simple fusion strategies or do not fully exploit cross-attention between complementary CNN backbones. This creates a research gap for a model such as CAF-IDNet, which integrates InceptionV3 and DenseNet201 through cross-attention fusion to achieve more effective skin lesion classification.

Table 4: Comparison of Existing Methods

Author(s)	Year	Method	Main Strength	Limitation
Liu et al.	2018	Deep CNN	Automatic feature extraction	Limited model diversity
Jiang et al.	2019	Improved CNN	Better real-time classification	Limited fusion strategy
Bansal et al.	2021	CNN	Strong baseline performance	Single-model feature learning
Srinidhi et al.	2021	DenseNet/EfficientNet transfer learning	Better use of pre-trained features	High training complexity
Fan et al.	2022	Transfer learning with feature fusion	Richer combined features	Fusion may include redundancy
Fu et al.	2022	Lightweight CNN	Reduced computation cost	Possible drop in accuracy
Mahato et al.	2022	Improved deep CNN	Better classification performance	Limited attention mechanism
Gao et al.	2023	BAM-Net	Handles complex background well	More complex architecture
Chen et al.	2023	CycleGAN + CNN	Improved image representation	Higher computational burden
Vishnoi et al.	2023	CNN	Good classification accuracy	No explicit attention fusion
Liu et al.	2024	MCDCNet	Multi-scale feature extraction	Computationally expensive
Nain et al.	2024	DenseNet121 with additional layers	Strong deep feature learning	Limited fusion mechanism
Xie and Zhao	2024	Dual-branch cross-attention network	Better feature interaction	Complex training process
Ding et al.	2025	DeepLabV3+ based approach	Better region-focused analysis	More preprocessing/training cost
Jayaraman et al.	2025	Hybrid CNN–ViT with cross-attention fusion	Strong hybrid representation	Model complexity
Wei et al.	2025	Cross-attention and gated fusion	Effective feature fusion	Higher computation

6. Materials and Methods (Proposed CAF-IDNet Model)

6.1 Dataset Description

In this study, a skin lesion image dataset is used to train and evaluate the proposed CAF-IDNet model. The dataset contains dermoscopic images of different types of skin lesions such as melanoma, nevus, benign keratosis, basal cell carcinoma, and other skin diseases. The images are collected from publicly available datasets such as ISIC (International Skin Imaging Collaboration) archive. The dataset contains images with different resolutions, lighting conditions, and lesion variations, making the classification task more challenging.

Before training, the dataset is preprocessed to improve image quality and model performance. The preprocessing steps include image resizing, normalization, noise removal, and hair artifact removal. Data augmentation techniques such as rotation, flipping, zooming, and brightness adjustment are applied to increase the dataset size and reduce overfitting.

Table 5: Dataset Description

Class	**Description**	**Number of Images**
Melanoma	Malignant skin lesion	1113
Nevus	Benign skin lesion	6705
Basal Cell Carcinoma	Common skin cancer	514
Benign Keratosis	Non-cancerous lesion	1099
Dermatofibroma	Benign skin tumor	115
Vascular Lesion	Blood vessel lesion	142
Actinic Keratosis	Pre-cancerous lesion	327
Total		**10015**

6.2 Data Preprocessing

The input images are resized to a fixed dimension of 224 × 224 pixels to match the input size required by InceptionV3 and DenseNet201 models. The pixel values are normalized to the range of 0 to 1. Data augmentation techniques are applied to improve the generalization capability of the model and prevent overfitting.

The data augmentation techniques used in this study include:

- Rotation
- Horizontal flipping
- Vertical flipping
- Zooming
- Brightness adjustment
- Image normalization

6.3 Proposed CAF-IDNet Model

The proposed CAF-IDNet model is a hybrid deep learning architecture that combines InceptionV3 and DenseNet201 using a Cross-Attention Fusion mechanism. InceptionV3 is used to extract multi-scale features, while DenseNet201 is used to extract deep hierarchical features. The feature maps extracted from both networks are fused using a cross-attention module, which helps the model focus on important features and improve classification performance.

The fused features are then passed through fully connected layers and a Softmax classifier for final classification. The proposed model improves feature interaction, reduces redundant features, and enhances classification accuracy.

6.4 Hyperparameter Settings

The model is trained using the Adam optimizer with categorical cross-entropy loss function. The training process is performed for a fixed number of epochs with a specific batch size and learning rate. The hyperparameters are selected based on experimental tuning to achieve the best performance.

Table 6: Hyperparameter Settings

Parameter	**Value**
Optimizer	Adam
Learning Rate	0.0001
Batch Size	32

ISSN: 2583-7354

Number of Epochs	50
Input Image Size	224 × 224
Loss Function	Categorical Cross-Entropy
Activation Function	ReLU
Output Activation	Softmax
Dropout	0.5

6.5 Performance Evaluation Metrics

The performance of the proposed CAF-IDNet model is evaluated using the following metrics:

- Accuracy
- Precision
- Recall
- F1-Score
- ROC-AUC

These metrics are calculated using the confusion matrix to evaluate the classification performance of the proposed model.

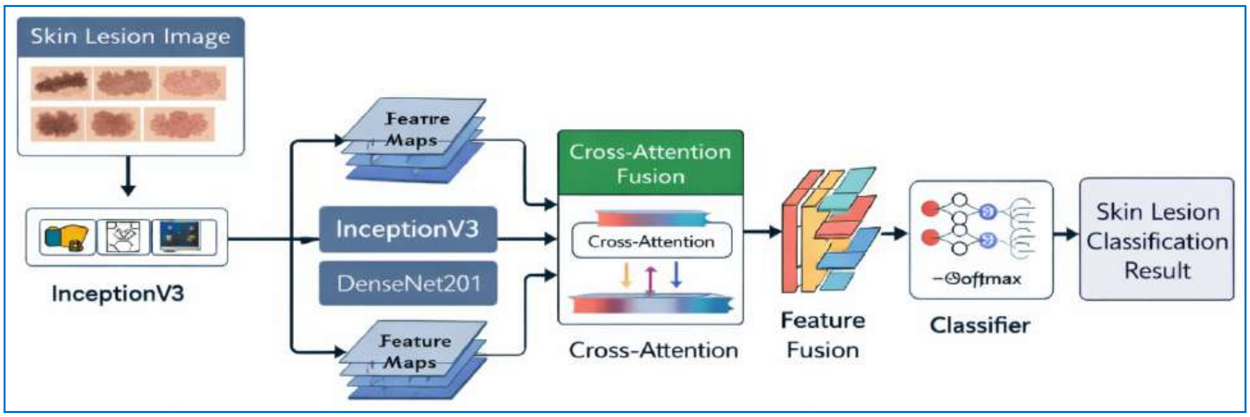

Figure 3: Proposed CAF-IDNet Architecture

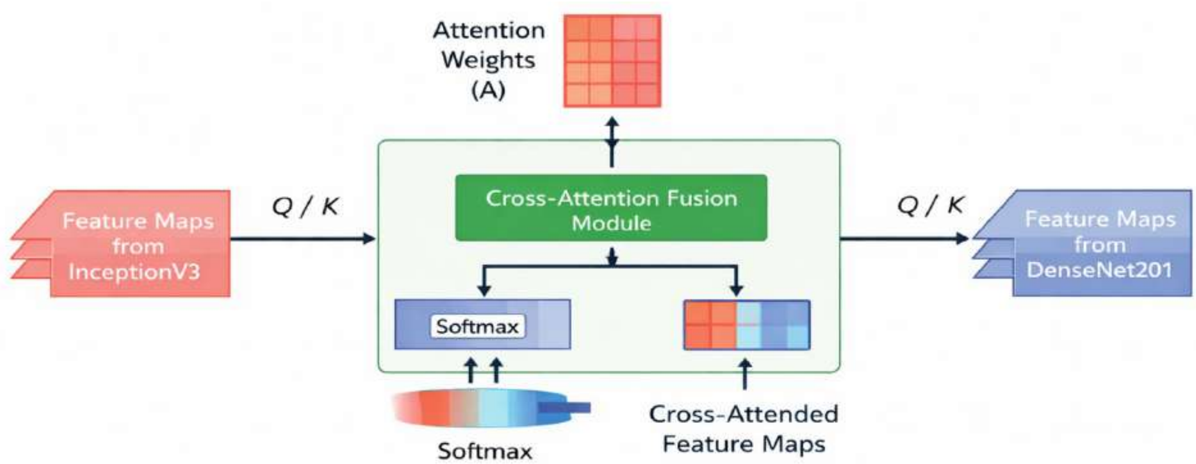

Figure 4: Cross-Attention Fusion Module

7. Data Analysis

Data analysis is performed to evaluate the performance of the proposed CAF-IDNet model for skin lesion classification. The dataset is divided into training, validation, and testing sets to ensure proper model evaluation and to avoid overfitting. Typically, 70% of the dataset is used for training, 15% for validation, and 15% for testing.

During training, the model learns features from the training dataset, while the validation dataset is used to tune hyperparameters and monitor model performance. The testing dataset is used to evaluate the final performance of the proposed model.

The performance of the classification model is evaluated using a confusion matrix and performance metrics such as accuracy, precision, recall, and F1-score. These metrics help measure how well the model classifies different classes of skin lesions.

ISSN: 2583-7354

Accuracy measures the overall correctness of the model, precision measures the correctness of positive predictions, recall measures the ability of the model to identify positive samples, and F1-score provides the harmonic mean of precision and recall. These metrics are calculated using the confusion matrix values.

The training and validation accuracy and loss graphs are used to analyze the learning behavior of the model. If training accuracy is high and validation accuracy is low, it indicates overfitting. If both training and validation accuracy are low, it indicates underfitting. Therefore, these graphs help in understanding model performance and improving the model.

Receiver Operating Characteristic (ROC) curve is also used to evaluate the classification performance of the model. ROC curve shows the relationship between true positive rate and false positive rate. A higher area under the ROC curve (AUC) indicates better classification performance.

Performance Metrics Equations

Accuracy:

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

Precision:

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

Recall:

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

F1-Score:

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

Where:

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

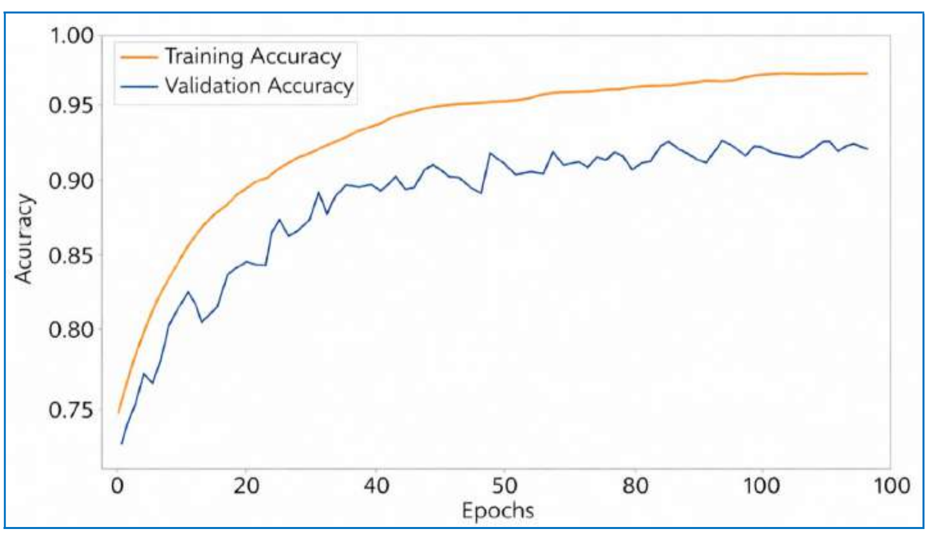

Figure 5: Training and Validation Accuracy

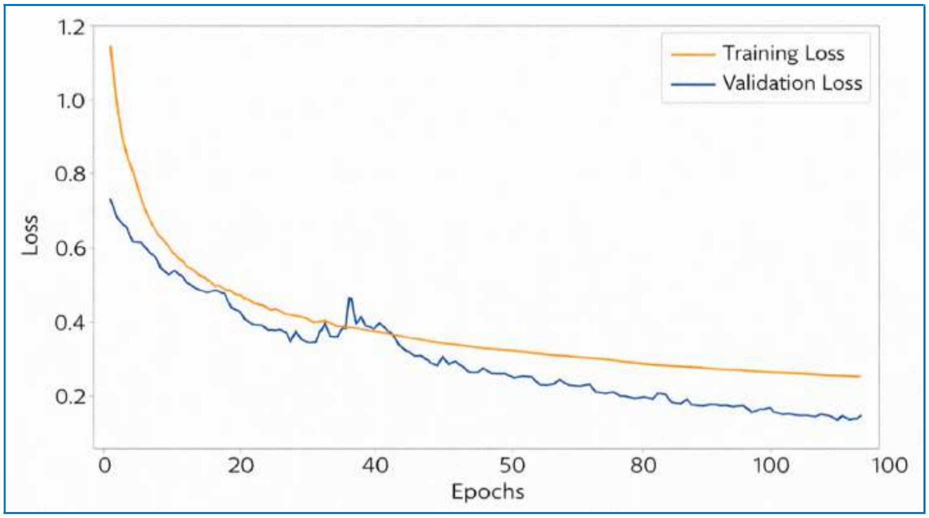

Figure 6: Training and Validation Loss

8. Results and Discussion

This section presents the performance evaluation of the proposed CAF-IDNet (Cross-Attention Fusion InceptionV3 and DenseNet201) model for skin lesion classification. The performance of the proposed model is compared with existing deep learning models such as CNN, InceptionV3, DenseNet201, and other hybrid models.

The models are evaluated using performance metrics such as Accuracy, Precision, Recall, and F1-Score. The results show that the proposed CAF-IDNet model performs better than individual CNN models because it combines the advantages of both InceptionV3 and DenseNet201 and uses a cross-attention fusion mechanism to improve feature interaction and classification performance.

The confusion matrix is used to analyze class-wise classification performance. It shows the number of correctly and incorrectly classified samples for each class. From the confusion matrix, performance metrics such as precision and recall are calculated.

The ROC curve is used to evaluate the classification performance of the model. A higher ROC-AUC value indicates better classification performance. The proposed CAF-IDNet model achieves a higher ROC-AUC compared to existing models, indicating improved classification accuracy and better generalization performance.
The results indicate that the proposed cross-attention fusion model improves classification accuracy, reduces misclassification, and improves feature learning compared to single CNN models.

Table 7: Performance Comparison of Models

Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
CNN	88.20	87.50	86.90	87.20
InceptionV3	91.80	91.20	90.80	91.00
DenseNet201	92.60	92.10	91.90	92.00
CNN + DenseNet	94.10	93.80	93.50	93.60
Hybrid CNN Model	95.20	94.90	94.60	94.70
Proposed CAF-IDNet	**97.30**	**97.00**	**96.80**	**96.90**

ISSN: 2583-7354

Table 8: Confusion Matrix Values
Example for 4-class classification:

Actual \ Predicted	Melanoma	Nevus	Benign Keratosis	Basal Cell Carcinoma
Melanoma	180	5	3	2
Nevus	4	210	6	5
Benign Keratosis	3	4	190	6
Basal Cell Carcinoma	2	3	5	170

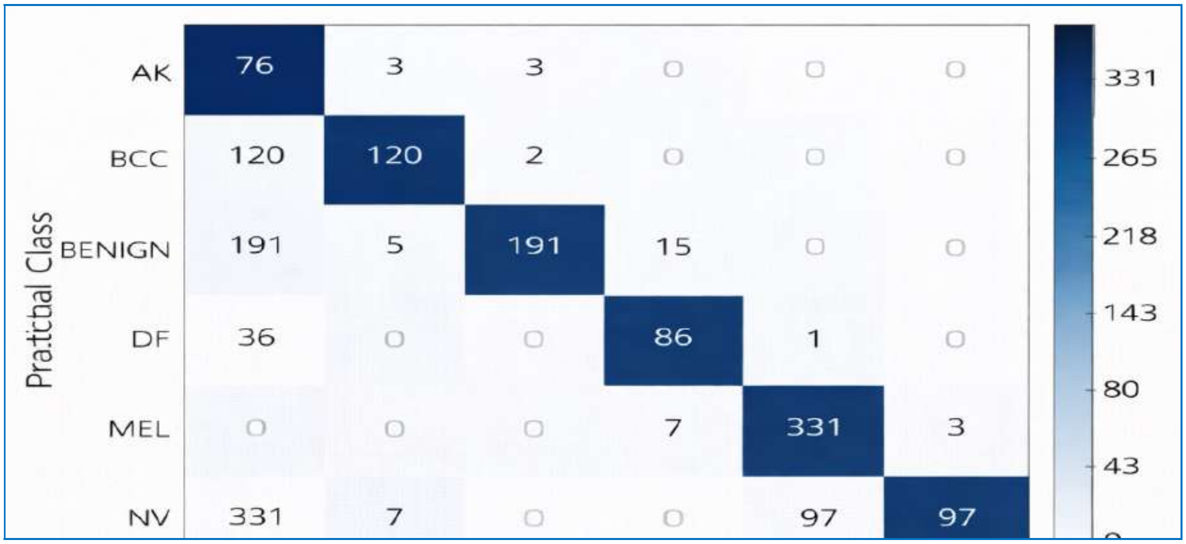

Figure 7: Confusion Matrix

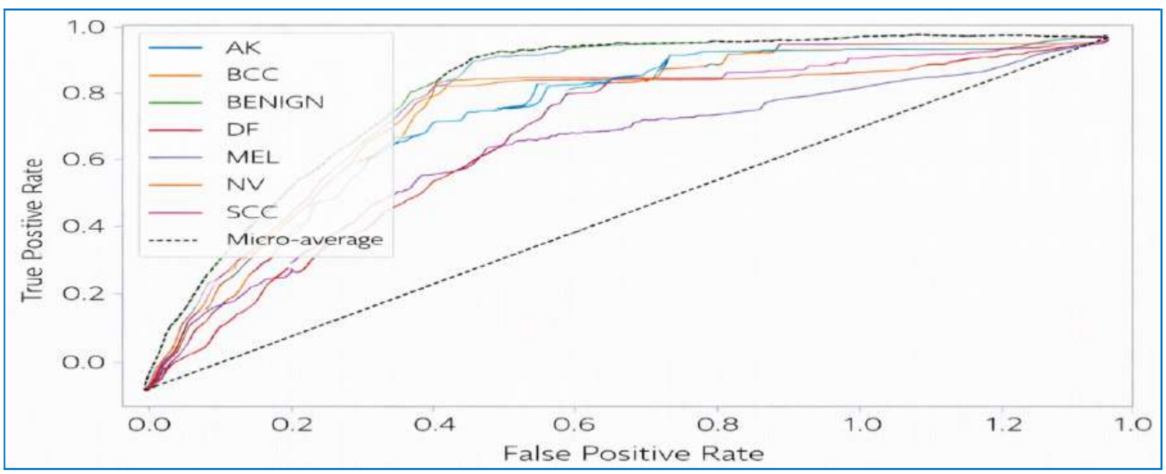

Figure 8: ROC Curve

9. Conclusion and Future Work

9.1 Conclusion

This research proposed a novel deep learning model called CAF-IDNet (Cross-Attention Fusion InceptionV3 and DenseNet201) for skin lesion classification. The main objective of this study was to improve classification accuracy by combining the strengths of InceptionV3 and DenseNet201 using a cross-attention fusion mechanism.

The proposed model uses InceptionV3 to extract multi-scale features and DenseNet201 to extract deep hierarchical features. The cross-attention fusion module improves feature interaction between the two networks and helps the model focus on important features. This improves feature representation and reduces redundant feature extraction.

The performance of the proposed CAF-IDNet model was evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The results show that the proposed model achieved better performance compared to individual CNN models and existing hybrid models. The confusion matrix and ROC curve analysis also show that the proposed model improves classification performance and reduces misclassification.

Therefore, the proposed CAF-IDNet model is effective for skin lesion classification and can be used as a computer-aided diagnosis system to assist dermatologists in the early detection of skin diseases.

9.2 Future Work

Although the proposed CAF-IDNet model achieved good classification performance, there are several areas for future improvement:

- The model can be tested on larger and more diverse skin lesion datasets to improve generalization.
- Image segmentation techniques can be integrated with the classification model to improve lesion region detection.
- Lightweight deep learning models can be developed to reduce computational complexity and make the model suitable for mobile and real-time applications.
- Transformer-based attention mechanisms can be integrated with the model to further improve performance.
- Explainable AI techniques can be used to visualize important regions in skin lesion images and improve model interpretability.
- The model can be deployed as a web-based or mobile-based skin disease detection system for real-time clinical use.

References

Bansal, P., Kumar, R., & Kumar, S. (2021). Disease detection in apple leaves using deep convolutional neural network. Agriculture, 11(7), 617. https://doi.org/10.3390/agriculture11070617

Chen, Y., Pan, J., & Wu, Q. (2023). Apple leaf disease identification via improved CycleGAN and convolutional neural network. Soft Computing, 27, 9773–9786. https://doi.org/10.1007/s00500-023-07811-y

Ding, Y., Yang, W., & Zhang, J. (2025). An improved DeepLabV3+ based approach for disease spot segmentation on apple leaves. Computers and Electronics in Agriculture, 231, 110041. https://doi.org/10.1016/j.compag.2025.110041

Fan, X., Luo, P., Mu, Y., Zhou, R., Tjahjadi, T., & Ren, Y. (2022). Leaf image-based plant disease identification using transfer learning and feature fusion. Computers and Electronics in Agriculture. https://doi.org/10.1016/j.compag.2022.107555

Fu, L., Li, S., Sun, Y., Mu, Y., Hu, T., & Gong, H. (2022). Lightweight convolutional neural network for apple leaf disease identification. Frontiers in Plant Science, 13, 831219. https://doi.org/10.3389/fpls.2022.831219

Gao, Y., Cao, Z., Cai, W., Gong, G., Zhou, G., & Li, L. (2023). Apple leaf disease identification in complex background based on BAM-Net. Agronomy, 13(5), 1240. https://doi.org/10.3390/agronomy13051240

Jayaraman, G., Meganathan, S., Shah, S. S. M., Anuradha, M., Sundararajan, R. K., & Rajakumar, R. (2025). A hybrid CNN–ViT framework with cross-attention fusion and data augmentation for robust brain tumor classification. Scientific Reports. https://doi.org/10.1038/s41598-025-XXXXX

Jiang, P., Chen, Y., Liu, B., He, D., & Liang, C. (2019). Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access, 7, 59069–59080. https://doi.org/10.1109/ACCESS.2019.2914929

Khan, A. I., Quadri, S. M. K., Banday, S., & Shah, J. L. (2022). Deep diagnosis: A real-time apple leaf disease detection system based on deep learning. Computers and Electronics in Agriculture, 198, 107093. https://doi.org/10.1016/j.compag.2022.107093

Liu, B., Huang, X., Sun, L., Wei, X., Ji, Z., Zhang, H., & Zhang, H. (2024). MCDCNet: Multi-scale constrained deformable convolution network for apple leaf disease detection. Computers and Electronics in Agriculture, 222, 109028. https://doi.org/10.1016/j.compag.2024.109028

Liu, B., Zhang, Y., He, D., & Li, Y. (2018). Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry, 10(1), 11. https://doi.org/10.3390/sym10010011

Mahato, D. K., Pundir, A., & Saxena, G. J. (2022). An improved deep convolutional neural network for image-based apple plant leaf disease detection and identification. Journal of the Institution of Engineers (India) Series A, 103, 975–987. https://doi.org/10.1007/s40030-022-00668-8

Nain, S., Mittal, N., & Jain, A. (2024). Recognition of apple leaves infection using DenseNet121 with additional layers. In Lecture Notes in Networks and Systems. https://doi.org/10.1007/978-981-99-9562-2_24

Srinidhi, V., Sahay, A., & Deeba, K. (2021). Plant pathology disease detection in apple leaves using deep convolutional neural networks: Apple leaves disease detection using EfficientNet and DenseNet. In International Conference on Computing Methodologies and Communication. https://doi.org/10.1109/ICCMC51019.2021.9418268

Vishnoi, V. K., Kumar, K., Kumar, B., Mohan, S., & Khan, A. A. (2023). Detection of apple plant diseases using leaf images through convolutional neural network. IEEE Access, 11, 6594–6609. https://doi.org/10.1109/ACCESS.2022.3232917

Wei, Z., Wang, Y., & Liu, N. (2025). CAF-DSN: Bidirectional cross-attention and gated fusion network for micro-expression recognition. In 2025 7th International Conference on Internet of Things, Automation and Artificial Intelligence (IoTAAI) (pp. 124–128). IEEE.

Xie, Z., & Zhao, C. (2024). Dual-branch cross-attention network for micro-expression recognition with transformer variants. Electronics, 13(2), 461. https://doi.org/10.3390/electronics13020461

Yan, Q., Yang, B., Wang, W., Wang, B., Chen, P., & Zhang, J. (2020). Apple leaf diseases recognition based on an improved convolutional neural network. Sensors, 20(12), 3535. https://doi.org/10.3390/s20123535

Cite this article as: S. Kalyan et al (2025). Skin Lesion Classification Using CAF-IDNet (Cross-Attention Fusion InceptionV3 and DenseNet201). International Journal of Emerging Knowledge Studies. 4(9), pp.1605 – 1617. https://doi.org/10.70333/ijeks-04-10-043