Detecting and Investigating Digital Image Forgeries

M RamyaSree  
*Computer Science Engineering (AI)*  
*Amrita Vishwa Vidyapeetham*Bangalore, India  
bl.en.u4aie21072@bl.students.amrita.edu

Dr. Nidhin Prabhakar T V  
*Department of Computer Science*   
*Amrita Vishwa Vidyapeetham*Bangalore, India

Mettu Siddhartha  
*Computer Science Engineering (AI)*  
*Amrita Vishwa Vidyapeetham*Bangalore, India  
bl.en.u4aie21080@bl.students.amrita.edu

Dr. Rimjhim Padam Singh  
*Department of Computer Science*   
*Amrita Vishwa Vidyapeetham*Bangalore, India

Poli Vamsi Vardhan Reddy  
*Computer Science Engineering (AI)*  
*Amrita Vishwa Vidyapeetham*Bangalore, India  
bl.en.u4aie21101@bl.students.amrita.edu

***Abstract*— In today's digital era, the ease of manipulating images using various software tools has given rise to a growing concern regarding the authenticity of digital visual content. Image forgeries, whether for malicious intent or innocent retouching, pose significant challenges to trustworthiness in digital media. This project explores the field of image forensics, a specialized domain of digital forensics, focusing primarily on detecting and identifying image forgeries. The main objective of image forensics is to investigate and determine whether a digital image has been manipulated or forged. This project delves into the techniques and methodologies employed to achieve this goal, including both classical and state-of-the-art approaches, covering common manipulations like copy-move, splicing, and retouching, and advanced techniques such as deepfake generation. The project emphasizes the importance of understanding the underlying principles and characteristics of digital images to detect these manipulations effectively.**

***Keywords—****Image Forgery Detection, Image Forensics, Copy-Move, Splicing, Deep Fake, Deep Learning (Attention Mechanism, DenseNet, Cross Model, Few-Short).*

1. Introduction

Images have become essential in today's digital age across a wide range of industries, including news media, sports, education, digital forensics, medical, and scientific research. They are important in many facets of our lives and act as key information sources. However, it is now much simpler to produce fake photographs because of the widespread availability of image editing programs like Photoshop, GIMP, and Coral Draw as well as smartphone apps like photo hacker. The ease with which photos may be altered raises serious questions regarding their veracity, particularly when those images are utilized as evidence in court.

Picture editing, sometimes referred to as picture manipulation, is the broad category of operations performed with software on digital images. Image forgery is the process of changing an image's content to lie about historical events. Image splicing is a specific type of image tampering in which new material is added to a picture, either from a separate image or duplicated from the same image (copy-move tampering). Given that such tampering may frequently go undetected to the human eye, detecting these modifications has become a crucial task.

Image manipulation detection approaches can be categorized into two primary types: active and passive. Active approaches involve the embedding of additional information, such as digital watermarks, during image acquisition or later stages, which is then utilized for manipulation detection. In contrast, passive approaches, often referred to as "blind approaches," do not rely on embedded information but instead extract image features for forgery detection. Passive techniques can address various types of forgeries, including compression and resampling.

Within passive forgery type-dependent approaches, two major categories are copy-move and splicing. Detecting these manipulations can be particularly challenging, as they are designed to be inconspicuous to human observers. Therefore, it is essential to develop effective techniques for detecting these forms of forgery, which are invaluable in the field of digital image forensics.

1. Literature survey

The rapid advancement of digital imaging technologies has made it increasingly easy to manipulate and alter images. This has led to the proliferation of forged images, which can be used for malicious purposes such as spreading misinformation, propaganda, and defaming individuals. As a result, image forgery detection has become an important area of research.

Image forgery detection is an important and growing field of research. With the increasing sophistication of image manipulation techniques, the need for effective detection methods is more important than ever.There are two main types of image forgery detection methods: active and passive. Active methods require additional data to be attached to the image, such as a watermark or fingerprint. Passive methods do not require any additional data to be attached to the image.Active methods are generally more effective than passive methods, but they are also more likely to be detected by forgers. Passive methods are less effective than active methods, but they are more difficult to detect [1].

Deep learning-based methods have emerged as powerful tools for image forgery detection, significantly surpassing traditional techniques. They offer high accuracy, robustness, and versatility in handling diverse manipulation types. Deep learning outperforms traditional methods but faces challenges like limited datasets and explainability [2].

The J. -L. Zhong reviews existing methods [3] for image copy-move forgery detection. These methods can be broadly categorized into two groups: block-based methods and keypoint-based methods. Block-based methods divide the image into small blocks and compare the similarity between blocks. Keypoint-based methods extract key points from the image and compare the similarity between key points.The author proposes a novel method for image copy-move forgery detection called Dense-InceptionNet. Dense-InceptionNet is a deep convolutional neural network that can learn features from the image that are useful for detecting copy-move forgeries. The author shows that Dense-InceptionNetcan achieve state-of-the-art performance on a variety of image copy-move forgery detection datasets.[3]

The authors. Zhang and J. Ni [4] propose a novel method for image forgery detection called Dense-U-Net with Cross-Layer Intersection. Dense-U-Net with Cross-Layer Intersection is a deep convolutional neural network that can learn features from the image that are useful for detecting forged regions. The author [4] shows that Dense-U-Net with Cross-Layer Intersection can achieve state-of-the-art performance on a variety of image forgery detection datasets.

The paper [5] surveys existing methods for deep fake image detection. These methods can be broadly categorized into two groups: convolutional neural network (CNN)-based methods and Siamese network-based methods. CNN-based methods use CNNs to extract features from the image that are useful for detecting deepfakesfor which author proposes a deep learning method called DeepFD for deep fake image detection. DeepFD is based on a modified minimized Xception Net and DenseNet. The paper shows that DeepFDcan achieve state-of-the-art performance on a variety of deep fake image detection datasets.

The paper proposes a novel deep learning method called D-CNN for deepfake image detection. D-CNN is based on a modified dense convolutional neural network architecture. The paper shows that D-CNN can achieve state-of-the-art performance on a variety of deepfake image detection datasets [6].

The author [7] proposed a method that can be broadly categorized into two groups: handcrafted feature-based methods and deep learning-based methods.Handcrafted feature-based methods rely on manually designed features to capture specific aspects of the image that are indicative of forgery. These features can be based on color, texture, statistics, or other image properties.Deep learning-based methods automatically learn features from the image data using deep convolutional neural networks.

The proposed method by the author [8] utilizes Mask R-CNN, a popular object detection and instance segmentation architecture, with MobileNet V1 as the backbone network. MobileNet V1 is a lightweight convolutional neural network (CNN) that offers good performance while being computationally efficient. This makes it suitable for real-world applications where processing power and resource constraints are important the proposed model achieves performance on several image splicing detection benchmarks, outperforming other techniques based on ResNet backbones.

The author proposes [9] a method for detecting copy-move forgeries in digital images using a Deep Convolutional Neural Network (DCNN) with a ResNet-50 learning algorithm. The paper effectively combines the strengths of two powerful techniques:Deep Convolutional Neural Networks (DCNNs): DCNNs are a type of artificial neural network that excel at extracting features from images. They have achieved state-of-the-art performance in various image recognition tasks.ResNet-50: ResNet-50 is a specific DCNN architecture known for its high accuracy and efficiency. It incorporates residual connections that help to address the vanishing gradient problem, which can hinder the performance of deep neural networks.By combining these two techniques, the proposed method aims to achieve accurate and efficient copy-move forgery detection.  
The authors [10] revolutionizes hyperspectral image classification with a deep learning architecture. It uses RBMs, powerful models capturing complex relationships, to extract higher-level features from each band. Additionally, a novel non-linear diffusion technique enhances image quality by boosting class separation and reducing noise. The combined approach outperforms traditional methods, as demonstrated by experiments using real hyperspectral data.  
Existing real-time change detection methods struggle with complexity or limited efficiency. The authors [11] proposes a lightweight hybrid algorithm using RGB features and local binary patterns to balance accuracy and speed. Evaluation on the 2014 CDnet dataset highlights its potential as a promising solution for practical change detection tasks.

Especially CNNs, are revolutionizing cancer detection through multiclass classification of cancers from histopathological images. This project leverages CNNs to predict cancer types from images, building a user-friendly tool for early detection and potentially improving patient outcomes[12].  
Feature extraction is crucial for classification, but extracting good features can be hard. Convolutional neural networks (CNNs) automate this well, but are slow to train. The authors[13] proposes a new classifier architecture that groups similar classes, making it faster to train even on large datasets without needing powerful GPUs. The new model achieves similar accuracy to other classifiers, making it a good choice when computers are slow or limited.

Skin cancer, rising fast, demands new solutions. Deep learning, powered by AI, could revolutionize screening. CNNs analyze images, spotting subtle cancer signs the human eye misses. This has the potential to save lives through early detection and better outcomes. Research is key to refine these models and integrate them into clinical practice, ultimately changing the fight against skin cancer[14].  
The authors[15] proposes automating WBC classification for early disease detection. To achieve this, the authors use several feature fusion techniques with pre-trained neural networks like DenseNet201 and VGG16. They combine features extracted from these networks and feed them to a CNN classifier. This approach achieved an impressive accuracy of 89.75%, demonstrating the potential of feature fusion for automated WBC classification.

1. IMPLEMENTATION

**Datasets**

The paper also discusses several different datasets that are available for evaluating image forgery detection methods.

Total Images: There are 12,616 images in the dataset, of which 7,492 are genuine (authentic) and 5,124 are fabricated (manipulated). Models are guaranteed to learn to differentiate between actual and altered information the dataset's equitable distribution of real and fake photos. These datasets include:

* The CASIA image forgery dataset: This dataset contains over 10,000 images, including both real and fake images.

Real Images (Authentic): The 7,492 real images serve as the authentic baseline, providing a diverse set of content for the model to understand and learn from.

Fake Images (Manipulated): The 5,124 fake images are intentionally manipulated to simulate various types of forgeries, such as copy-move, splicing, and other tampering techniques.

Real image:

|  |  |
| --- | --- |
|  |  |

Fake image:

|  |  |
| --- | --- |
|  |  |

* The Celeba dataset contains over 200,000 images with patterns, and places, including a small number of fake images.

<https://www.kaggle.com/datasets/sophatvathana/casia-dataset>

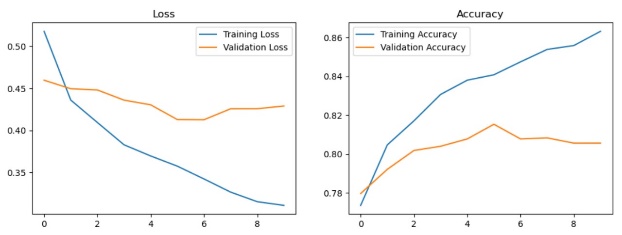
**Deep learning methods**

Deep learning methods are a more recent approach to image forgery detection. Deep learning methods can automatically learn features from images, without the need for handcrafting. This makes them more robust to a wider variety of image forgery techniques.

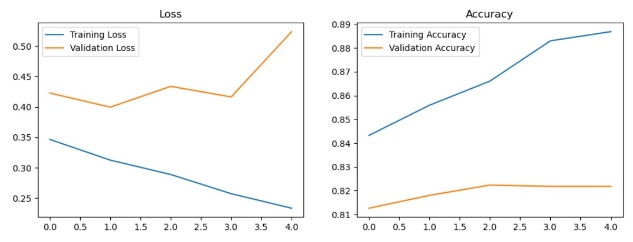
**Implementation techniques**

The paper discusses several different implementation techniques for deep learning-based image forgery detection. These techniques include:

* **DenseNet:**DenseNet's densely connected design facilitates the efficient learning of complex spatial connections and feature reuse, which improves DenseNet's capacity to identify minute discrepancies in modified images. By using pre-trained models on a variety of datasets, including ImageNet, its transfer learning capabilities combined with the concatenation of features from several layers give it robustness for forgery detection applications.



*Fig 1 DenseNet*



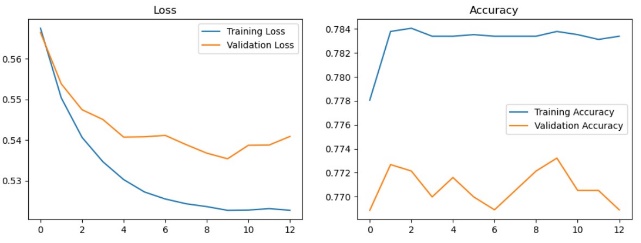
*Fig 2DenseNet- Fine Tuning*

In this paper we incorporated DenseNet201,DenseNet201 is a convolutional neural network (CNN) architecture that belongs to the DenseNet family. DenseNet (Densely Connected Convolutional Networks) is known for its unique structure where each layer receives direct input from all preceding layers. This connectivity pattern facilitates feature reuse and encourages the network to be more parameter-efficient.

DenseNet201 is a deeper variant of DenseNet, containing 201 layers.It uses densely connected blocks, where each layer receives feature maps from all previous layers in a block.

DenseNet201 is a powerful architecture known for its efficiency in training and parameter sharing. It is often used in various computer vision tasks, including image classification.

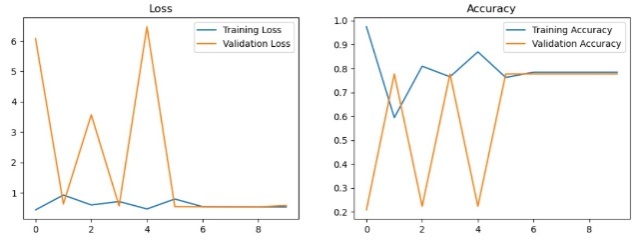
* **Attention Mechanisms**: Deep learning models' attention mechanisms—which draw inspiration from human visual attention—arecrucial in improving picture fraud detection effectiveness. By using these methods, models may dynamically focus on particular areas of a picture, which makes it possible for them to spot minute discrepancies brought about by forgeries.



*Fig 3DenseNet-Attention mechanism*

Utilizing attention mechanisms like the Transformer architecture helps the model focus on specific image regions with potential forgery, enhancing its ability to detect subtle inconsistencies.

* **Cross-Modal techniques**:Cross-modal algorithms combine information from several modalities, such as text and image data. Combining text and image data in cross-modal learning enables verification of image content authenticity by identifying inconsistencies between textual descriptions and visual content.



*Fig 4DenseNet-cross model*

By incorporating cross-modal techniques, one can gain a deeper comprehension of the context around a picture and improve the identification of forgeries that could incorporate both visual and non-visual information. The combination of these modalities results in an image forgery detection system that is more resilient and adaptable, able to handle a variety of manipulation approaches.

* **EfficientNet:** Compound scaling allows for the efficient construction of an EfficientNet, which guarantees a fair trade-off between computational efficiency and model accuracy. Because its scalability enables customization according to available resources, it can be used in real-time image forgery detection applications even in contexts with limited resources.

****

*Fig 5EfficienNet*

We implemented an image classification model using the EfficientNetB0 architecture to distinguish between real and fake images from the dataset.

The EfficientNetB0 model is created with pre-trained weights from ImageNet.Additional layers are added to the top of the model, including Global Average Pooling, Dense (with ReLU activation), Dropout, and a Dense output layer with softmax activation for binary classification.

* **ResNet:**ResNet stands for residual network, is a type of deep neural network architecture that was introduced to address the challenges of training very deep networks.ResNet's architecture incorporates skip connections or shortcut connections, allowing input from one layer to bypass others and be added to the output of a later layer. This facilitates the learning of identity mappings, simplifying the optimization of deep networks.

****

*Fig 6ResNet*

In summary, ResNet's architecture, developed for deep network training, can be repurposed in image forensics to create models capable of detecting manipulated or forged images by learning patterns associated with such tampering.

This section presents the implementation details of a ResNet50-based model for the classification of real and fake images from the CASIA2 dataset. The model is trained to distinguish between authentic (Au) and tampered (Tp) images. The ResNet50 architecture is utilized as a feature extractor, and additional layers are added to adapt the network for the binary classification task.

ResNet50 is employed as the base model with pre-trained weights from ImageNet.The input shape is set to (224, 224, 3) to match the model's requirements.

* **Xception:** Xception is a deep learning architecture for image classification, it is based on the Inception architecture, but with a key modification involving depthwise separable convolutions.Depthwise separable convolutions are used in Xception as an alternative to traditional convolutions. In a depthwise separable convolution, a regular convolution is split into two separate layers: a depthwise convolution and a pointwise convolution.

****

*Fig 7 Xception*

The effectiveness of models in image forensics is intricately tied to the quality and diversity of the training data. No singular model can assure the detection of all types of forgeries, emphasizing the need for ongoing research to enhance the robustness and generalization capabilities of deep learning models in this field.  
Xception is chosen as the base model with pre-trained weights from ImageNet.

The input shape is set to (299, 299, 3) to match Xception's default input size.

This implementation highlights the successful integration of the Xception architecture for image classification, specifically for detecting tampered images. The use of pre-trained weights, data augmentation, and appropriate callbacks contributes to the robustness and efficiency of the model. The reported results indicate its potential applicability in real-world scenarios

**Evaluation**

The paper discusses several different metrics that can be used to evaluate image forgery detection methods. These metrics include:

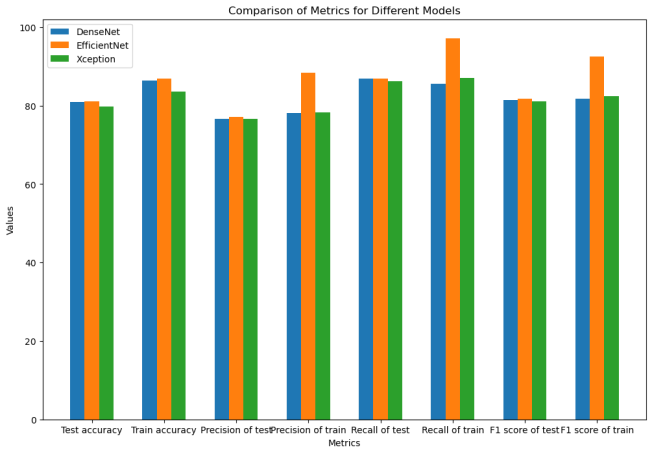
* Accuracy: Accuracy is the percentage of images that are correctly classified as either real or fake.Accuracy measures the overall correctness of the model by considering both true positives and true negatives.
* Precision: Precision is the percentage of images that are classified as fake that are fake.Precision focuses on the accuracy of positive predictions, indicating the proportion of correctly predicted positives among all predicted positives.
* Recall: Recall is the percentage of fake images that are correctly classified as fake.Recall measures the ability of the model to correctly identify all relevant instances, showing the proportion of actual positives correctly predicted.
* F1-score: The F1-score is a harmonic mean of precision and recall.F1 score is the harmonic mean of precision and recall, providing a balanced measure that considers both false positives and false negatives. It is particularly useful when there is an imbalance between positive and negative classes.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Test Accuracy | Train Accuracy | Validation Accuracy | Validation Loss |
| DenseNet | 0.8087 | 0.863 | 0.819 | 0.4290 |
| Fine Tuning | 0.819 | 0.887 | 0.822 | 0.524 |
| Attention Mechanism | 0.229 | 0.7834 | 0.7689 | 0.5409 |
| Cross Model | 0.7705 | 0.7842 | 0.7769 | 0.5784 |
| EfficientNet | 0.8152 | 0.8626 | 0.8164 | 0.4360 |
| Xception | 0.7969 | 0.8363 | 0.7981 | 0.4654 |
| ResNet | 0.7683 | 0.7834 | 0.7736 | 0.5310 |

*Fig 8.1 Metrics Table*

|  |  |  |  |
| --- | --- | --- | --- |
|  | DenseNet | EfficientNet | Xception |
| Accuracy | 0.8632 | 0.8682 | 0.8363 |
| Precision | 0.7816 | 0.8835 | 0.7820 |
| Recall | 0.8556 | 0.9715 | 0.8716 |
| F1 score | 0.8169 | 0.9254 | 0.8237 |

*Fig 8.2 Result Table*



*Fig 9Comparison of metrics for different models*

|  |
| --- |
|  |
|  |
|  |

*Fig 10.1 Model Testing For Real Image*

|  |
| --- |
|  |
|  |
|  |

*Fig 10.2 Model Testing for Fake Image*

Conclusion & Future Work

In the field of digital forensics use of Deep learning has shown promising results.Detection of image forgeries plays a critical role in forensic investigation, criminal investigation, and intelligence systems projects aimed at detecting and investigating digital image forgeries using deep learning models DenseNet and EfficientNet which have shown promising results.The project's significance lies in providing a reliable method to verify the authenticity of digital images, making it valuable for various applications, including forensic investigations and intelligence systems.

In future work, potential research directions include the exploration of advanced deep learning models, particularly convolutional neural networks (CNNs), for enhanced feature extraction and classification in the realm of image forgery detection.

* **Deep Learning Enhancements:**

Further advancements in deep learning models, especially convolutional neural networks (CNNs), can be explored for better feature extraction and classification of forged images.

* **Real-Time Forgery Detection in Social Media:**

Implementation of real-time image forgery detection systems for social media platforms to combat the rapid spread of manipulated content.

* **Multi-Modal Forgery Detection:**

Integration of multi-modal approaches that consider various types of data, such as metadata, sensor noise, and contextual information, to improve overall detection accuracy.

References

[1]Marcello Zanardelli, Fabrizio Guerrini, Riccardo Leonardi & Nicola Adami, “Image forgery detection: a survey of recent deep-learning approaches”,Multimedia Tools and Applications, 2022

.

[2] Zankhana J. Barad; Mukesh M. Goswami, “Image Forgery Detection using Deep Learning: A Survey”, 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)

[3] J. -L. Zhong and C. -M. Pun, "An End-to-End Dense-InceptionNet for Image Copy-Move Forgery Detection," in *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2134-2146, 2020, doi: 10.1109/TIFS.2019.2957693.

[4] R. Zhang and J. Ni, "A Dense U-Net with Cross-Layer Intersection for Detection and Localization of Image Forgery," *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, pp. 2982-2986, doi: 10.1109/ICASSP40776.2020.9054068.

[5] I. Sahib and T. A. A. AlAsady, "Deep fake Image Detection based on Modified minimized Xception Net and DenseNet," *2022 5th International Conference on Engineering Technology and its Applications (IICETA)*, Al-Najaf, Iraq, 2022, pp. 355-360, doi: 10.1109/IICETA54559.2022.9888278.

[6] Y. Patel *et al*., "An Improved Dense CNN Architecture for Deepfake Image Detection," in *IEEE Access*, vol. 11, pp. 22081-22095, 2023, doi: 10.1109/ACCESS.2023.3251417.

[7] S. Walia, K. Kumar, M. Kumar, and X. -Z. Gao, "Fusion of Handcrafted and Deep Features for Forgery Detection in Digital Images," in *IEEE Access*, vol. 9, pp. 99742-99755, 2021, doi: 10.1109/ACCESS.2021.3096240.

[8] K. Kadam, S. Ahirrao, K. Kotecha and S. Sahu, "Detection and Localization of Multiple Image Splicing Using MobileNet V1," in *IEEE Access*, vol. 9, pp. 162499-162519, 2021, doi: 10.1109/ACCESS.2021.3130342.

[9] V. Sharma and N. Singh, "Deep Convolutional Neural Network with ResNet-50 Learning algorithm for Copy-Move Forgery Detection," 2021 7th International Conference on Signal Processing and Communication (ICSC), Noida, India, 2021, pp. 146-150, doi: 10.1109/ICSC53193.2021.9673422.

[10] Midhun EM, Sarath R Nair, Nidhin PTV, Sachin Kumar S, Deep Model for Hyperspectral Image Classification using Restricted Boltzman Machines, International Conference on Interdisciplinary Advances in Applied Computing, ACM-Sig Conference, 2014 (Scopus)

[11] Rimjhim Padam Singh and Sharma, P., “Improving Change Detection Using Centre-Symmetric Local Binary Patterns”, in Pattern Recognition and Machine Intelligence, Cham, 2019.

[12] S. Chethana, S. S. Charan, V. Srihitha, A. Thakur and N. P. T V, "Multi-Class classification of Different Cancer Types Using CNN," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-5, doi: 10.1109/ICCCNT56998.2023.10306850.

[13] Neena, A., Geetha, M. (2018). Image Classification Using an Ensemble-Based Deep CNN. In: Sa, P., Bakshi, S., Hatzilygeroudis, I., Sahoo, M. (eds) Recent Findings in Intelligent Computing Techniques . Advances in Intelligent Systems and Computing, vol 709. Springer, Singapore. https://doi.org/10.1007/978-981-10-8633-5\_44

[14] M. C. Mannava, B. Tadigadapa, D. Anil, A. S. Dev, A. T and A. V. Vidyapeetham, "CNN Comparative Analysis for Skin Cancer Classification," 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2022, pp. 1-6, doi: 10.1109/ICCCNT54827.2022.9984324.

[15] S. Parayil and J. Aravinth, "Transfer Learning-based Feature Fusion of White Blood Cell Image Classification," 2022 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2022, pp. 1468-1474, doi: 10.1109/ICCES54183.2022.9835815.