

Improving Digital Image Forgery Detection through Transfer Learning Techniques

Panchami B R, Dr Prasad G R

Department of CSE, BMS College of Engineering, Bengaluru, India
Professor, Department of CSE, BMS College of Engineering, Bengaluru, India

Abstract - In today's digital age, images shared on social media have become a primary means of communication. However, the rise of sophisticated malicious software capable of falsifying images has made it essential to identify such forgeries. Existing research in this field often focuses on detecting a single type of forgery, such as image splicing or copy-move, which may not be applicable to real-world scenarios. This paper introduces a novel method for enhancing digital image forgery detection by leveraging deep learning techniques and transfer learning. The proposed approach aims to identify two types of image forgery simultaneously by examining the compressed quality variations in forged regions compared to the rest of the image. The method involves creating a feature image from the difference between the original and compressed versions of the image. This feature image is then used to train a pre-existing deep learning model, which has had its original classifier removed and replaced with a new, fine-tuned one. The study evaluates eight different pre-trained models adapted for binary classification. The experimental findings demonstrate that the proposed method, when applied to these eight models, significantly outperforms current state-of-the-art techniques, as evidenced by various evaluation metrics, charts, and graphs. Notably, the DenseNet121 model achieved the highest detection accuracy (approximately 98%) with fewer training parameters, resulting in a quicker training process.

Key Words: Detect forgery, Deep learning, DenseNet121, MobileNetV2, Image Forgery, CNN, Transfer Learning etc.

1. INTRODUCTION

In the current digital era, image forgery has become a significant issue, particularly with the widespread use of social media and online platforms. The manipulation of images has become increasingly prevalent, leading to concerns about misinformation and the spread of fake news. This underscores the critical need for effective forgery detection methods to address these challenges. Advanced editing tools have made it easier to alter images, complicating the task of identifying various forms of forgery. As a result, developing robust detection techniques is essential for ensuring the integrity of visual information and combating the negative effects of digital deception.

The objectives section details the aims and targets of the research endeavor. It outlines key goals such as advancing the methods for detecting image forgery, enabling the simultaneous identification of various types of forgery, and applying deep learning techniques to boost accuracy.

Additionally, this section specifies the goal of comparing the performance of different pre-trained models to gauge their effectiveness in detecting forgeries. It also emphasizes the objective of demonstrating the benefits of incorporating these advanced techniques into forgery detection strategies.

2. RELATED WORK

In reference [1], the authors introduced an innovative method for detecting Copy-Move Forgeries (CMFD) by integrating deep learning techniques, specifically using a combination of Convolutional Neural Networks (CNNs) and Convolutional Long Short-Term Memory (ConvLSTM) networks. Their approach involves extracting features from images through a sequence of convolutional layers, ConvLSTM layers, and pooling operations, followed by a feature matching process to identify forgeries. The method was evaluated using four public datasets: MICC-F220, MICC-F2000, MICC-F600, and SATs-130, which were merged to form new datasets aimed at improving generalization and reducing overfitting. Additionally, the performance of a ConvLSTM-only model was assessed to compare with the hybrid ConvLSTM-CNN model. The results demonstrated that the proposed method achieved exceptional accuracy, reaching up to 100% on some datasets, with processing times as brief as 1 second, thereby surpassing the performance of previous methods.

In reference [2], the authors present a technique for identifying both copy-move and splicing image forgeries using Convolutional Neural Networks (CNNs) and three distinct models: Error Level Analysis (ELA), VGG16, and VGG19. Their method includes a pre-processing phase where images are compressed to a particular quality level before being used to train the models. Once trained, these models are employed to classify images as either genuine or manipulated. The experimental findings indicate that the proposed approach yields accuracy rates of 70.6% with Error Level Analysis (ELA), 71.6% with VGG16, and 72.9% with VGG19, when tested on images from the CASIA2.0 and NC2016 datasets.

In reference [3], the authors introduce a compact Convolutional Neural Network (CNN) designed for real-time splicing image forgery detection, combining high accuracy with a minimal number of parameters. The model features four convolutional layers and four max-pooling layers, making it well-suited for environments with limited resources. A thorough comparison with other models reveals that this approach achieves notable sensitivity and specificity. It recorded accuracies of 99.1% on the CASIA 1.0 dataset, 99.3% on the CASIA 2.0 dataset, and 100% on the CUISDE dataset. Its high performance and efficiency make it an effective tool for automated, real-time forgery detection.

In reference [4], the authors present an automated deep learning-based fusion model called DLFM-CMDFC for the detection and localization of copy-move forgeries. This approach integrates Generative Adversarial Networks (GANs) with Densely Connected Networks (DenseNets). The combined outputs are then processed by an Extreme Learning Machine (ELM) classifier, with its weight and bias parameters optimized through the Artificial Fish Swarm Algorithm (AFSA). This fusion technique is used to detect inconsistencies between the input and target regions within images. The model's effectiveness was tested on two benchmark datasets, demonstrating superior performance compared to recent methods, with a precision of 97.27%, recall of 96.46%, and F-score of 96.06%.

In reference [5], the authors develop a comprehensive fully convolutional neural network designed for detecting image forgery. This model integrates multi-resolution hybrid features from both RGB and noise channels to detect visual and compression inconsistencies in altered images. It includes a tamper-guided dual self-attention (TDSA) module that targets and segments tampered regions by identifying discrepancies between altered and unaltered areas. Extensive testing demonstrates that this approach provides enhanced pixel-level forgery localization and image-level detection accuracy compared to existing methods, exhibiting improved precision and robustness.

In reference [6], the authors introduce SD-Net, a novel approach aimed at overcoming the limitations of many CNN-based methods for copy-move forgery detection (CMFD), which often suffer from low accuracy. SD-Net employs super-BPD segmentation technology to enhance edge detection and leverages a Deep Convolutional Neural Network (DCNN) to bolster robustness. Experimental results reveal that SD-Net achieves precise edge localization and performs effectively in identifying large-scale forgeries. However, the addition of the segmentation module and dual-branch architecture adds complexity to the method. Future research is recommended to simplify this complexity while preserving accuracy and to further investigate forgery detection in cases where altered regions are similar to genuine ones.

In reference [7], the authors investigate two critical aspects of utilizing deep convolutional neural networks (CNNs) for detecting image forgeries. The first aspect focuses on analyzing various pre-processing techniques alongside different CNN architectures. The second aspect assesses the efficacy of transfer learning by fine-tuning pre-trained ImageNet models on the CASIA V2.0 dataset. The study involves implementing and evaluating several models and their combinations, highlighting the effects of pre-processing and transfer learning. Notably, the CNN model incorporating sharpening and Error Level Analysis (CNN_SharpEN_ELA) achieves a training accuracy of 97% with a minimal training loss of 0.1%. In contrast, the ResNet50 model attains a test accuracy of 95% with a low test loss of approximately 0.4%.

In reference [8], the authors present a digital image forgery detection system utilizing the ResNet50v2 deep learning architecture. This model processes image batches and integrates YOLO CNN weights within the ResNet50v2 framework to enhance performance. The system was tested on the CASIA_v1 and CASIA_v2 benchmark datasets, which contain both authentic and forged images, to detect image splicing. The datasets were divided with 80% allocated for training and 20% for testing. A comparison with other methods revealed that the model, when fine-tuned through transfer learning, achieved an accuracy of 99.3% on the more extensive CASIA_v2 dataset. In contrast, the accuracy dropped to 81% without the use of transfer learning. These findings highlight the effectiveness and advantages of the proposed system.

In reference [9], the authors introduce a streamlined model, combining Mask R-CNN with MobileNet V1, designed for detecting and identifying both copy-move and image splicing forgeries. The model was assessed using several standard datasets, including COVERAGE, CASIA 2.0, MICC F220, MICC F600, MICC F2000, COLUMBIA, and CASIA 1.0. Compared to ResNet-101, the proposed model achieved superior results, with an F1-score of 70% for copy-move forgery on the MICC F600 dataset and 64% for image splicing forgery on CASIA 1.0. It also recorded an average precision of 90% for copy-move forgery on both MICC F2000 and COVERAGE datasets, and 90% for image splicing on the COLUMBIA dataset. The model offers a more efficient computational performance than ResNet-101, providing a good balance between efficiency and computational cost. Additionally, it is capable of estimating the percentage of forgery present in different regions of an image.

In reference [10], the authors investigate the application of two deep learning models, Smaller VGGNet and MobileNetV2, for detecting copy-move image forgery, with a focus on post-processed attacks. These models are engineered to be both time-efficient and resource-conservative, making them ideal for deployment on embedded systems. Through comprehensive analysis, it was determined that a modified version of MobileNetV2 is particularly effective in detecting copy-move forgeries and

managing post-forgery alterations such as changes in brightness, blurring, noise addition, cropping, and rotation. Experimental results demonstrate that the MobileNetV2-based model achieves an 84% True Positive Rate (TPR) and a 14.35% False Positive Rate (FPR) when identifying digital image forgeries after undergoing multiple post-processing attacks.

3. PROPOSED SYSTEM

3.1 Importing Required Libraries

The script starts by importing key libraries necessary for data processing, visualization, and model building. This includes modules like os, random, glob, numpy, matplotlib, PIL, sklearn and tensorflow.keras.

3.2 Loading the Dataset

The dataset is located in the directory 'data/CASIA2.0_revised/data_new'. The script loads the contents from this directory and tallies the number of images present in each of its subdirectories.

3.3 Exploratory Data Analysis

Some exploratory data analysis is performed, which includes:

- **Showing Sample Images:** Randomly display images from the dataset along with their metadata, including details such as file path, dimensions, color mode, and format.
- **Visualizing Image Distribution:** Generate bar charts to depict the count of images in each category and pie charts to represent the proportional distribution of images among various categories.

3.4 Splitting the Dataset

The split folders library is utilized to partition the dataset into training, validation, and test sets, allocating 70% of the data for training, 10% for validation, and 20% for testing.

3.5 Importing Required Libraries

Image data undergoes preprocessing and augmentation using ImageDataGenerator. This process involves transformations such as rotation, shifting, brightness adjustment, zooming, and horizontal flipping. Furthermore, the data is normalized with the preprocess_input function from tensorflow.keras.applications.densenet.

3.6 Model Construction

The DenseNet121 model is used as the foundational base for feature extraction, incorporating pre-trained weights from ImageNet. The base model's layers are kept frozen to ensure they are not updated during the fine-tuning process. Additional layers are added to the model, including:

- A flattening layer

- Batch normalization layers
- Dense layers with ReLU activation functions
- Dropout layers for regularization
- A final Dense layer with a softmax activation function for classification

3.7 Model Compilation

The model is compiled with the Adam optimizer and configured to use categorical cross-entropy as the loss function. It evaluates performance using metrics such as categorical accuracy, precision, and recall.

3.8 Training the Model

A learning rate scheduler is set up to dynamically adjust the learning rate throughout the training process. The model is trained on the training dataset, incorporating class weights to address class imbalance, and is validated using the validation dataset.

3.9 Evaluating the Model

The model's performance is evaluated on the test set, with key metrics including categorical accuracy, loss, precision, and recall being reported. Additionally, a confusion matrix and a classification report are created and visualized to provide further insight.

3.10 Saving the Model

The weights and architecture of the trained model are saved for future reference and use.

3.11 Loading the Model

The model is loaded from the file densenet121.hdf5 for making predictions or conducting additional evaluations.

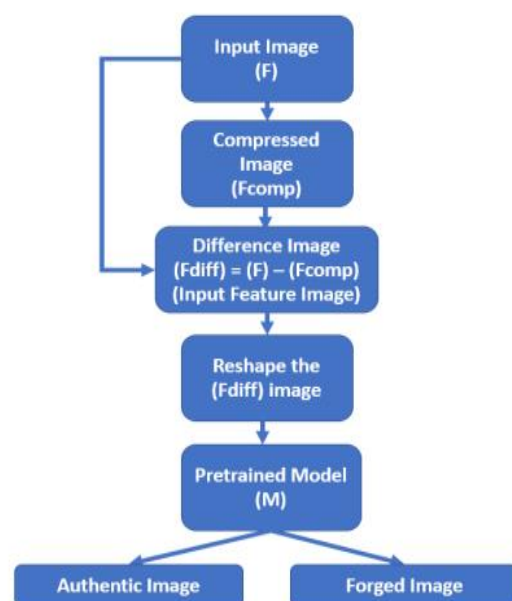


Figure -1: Proposed architecture model

The proposed system aims to overcome the limitations of current forgery detection methods by employing advanced deep learning techniques, particularly convolutional neural networks (CNNs) and transfer learning. Key aspects of the system include:

- 1. Deep Learning Framework:** The system leverages deep learning architectures, specifically CNNs, for both feature extraction and classification. By utilizing CNNs' ability to learn hierarchical representations, the system can automatically extract pertinent features from images, minimizing the need for manual feature extraction.
- 2. Transfer Learning:** The system utilizes transfer learning by adapting pre-trained CNN models, such as VGG16, VGG19, ResNet, and MobileNet, which have been trained on extensive image datasets like ImageNet. Fine-tuning these models for forgery detection tasks allows the system to leverage previously acquired knowledge, improving its ability to generalize and perform effectively.
- 3. Advanced Forgery Detection Methods:** The system incorporates innovative techniques for detecting various types of image forgeries, such as copy-move and splicing forgeries. By examining differences in compression quality, it identifies inconsistencies between genuine and altered image areas, thereby enhancing detection accuracy.
- 4. Model Evaluation and Comparison:** A thorough comparative analysis of different pre-trained models is conducted to assess their performance in forgery detection. Metrics such as accuracy, precision, recall, and F1-score are used to evaluate the effectiveness of each model in identifying image forgeries.

4. RESULTS

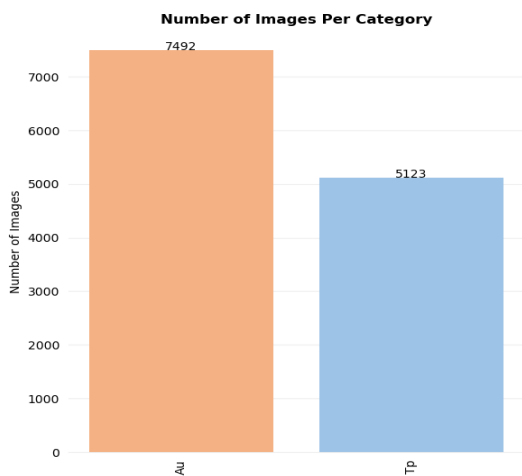
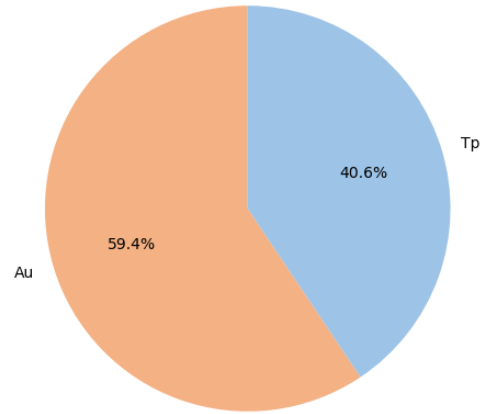


Fig -1: Name of the figure

Percent Distribution of Images Across Categories



Model: "sequential"

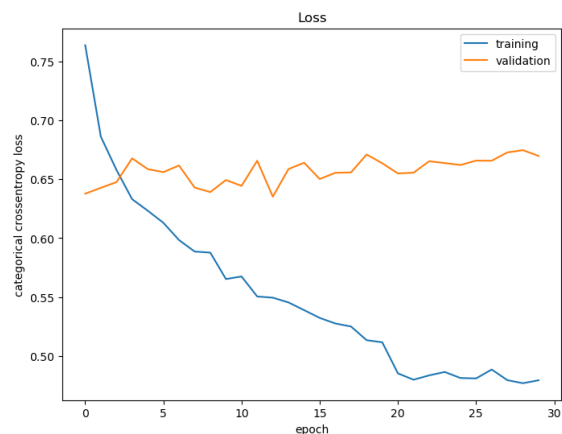
Layer (type)	Output Shape	Param #
densenet121 (Functional)	?	7,037,504
flatten (Flatten)	?	0 (unbuilt)
batch_normalization (BatchNormalization)	?	0 (unbuilt)
dense (Dense)	?	0 (unbuilt)
dropout (Dropout)	?	0
batch_normalization_1 (BatchNormalization)	?	0 (unbuilt)
dense_1 (Dense)	?	0 (unbuilt)
dropout_1 (Dropout)	?	0
batch_normalization_2 (BatchNormalization)	?	0 (unbuilt)
dense_2 (Dense)	?	0 (unbuilt)
dense_3 (Dense)	?	0 (unbuilt)

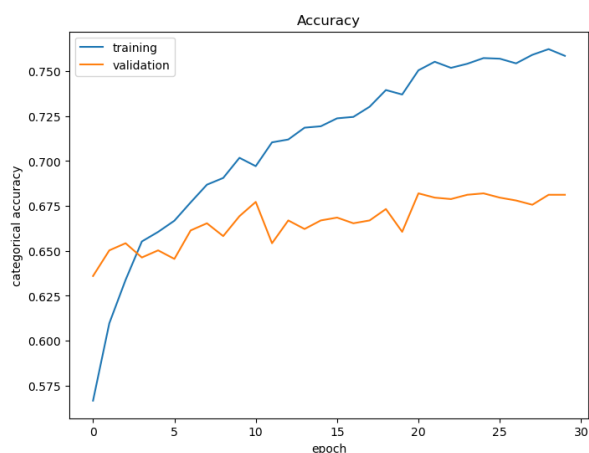
Total params: 7,037,504 (26.85 MB)

Trainable params: 0 (0.00 B)

Non-trainable params: 7,037,504 (26.85 MB)

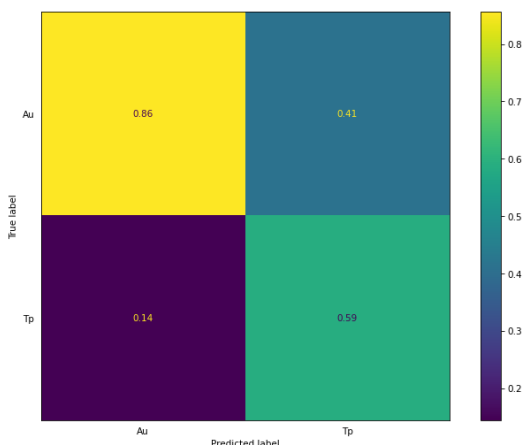
```
Epoch 27/30
552/552 ----- 220s 398ms/step - categorical_accuracy: 0.7535 - loss: 0.4904 - precision: 0.7535 - recall: 0
.7535 - val_categorical_accuracy: 0.6780 - val_loss: 0.6656 - val_precision: 0.6780 - val_recall: 0.6780 - learning_rate:
1.0000e-05
Epoch 28/30
552/552 ----- 219s 396ms/step - categorical_accuracy: 0.7592 - loss: 0.4829 - precision: 0.7592 - recall: 0
.7592 - val_categorical_accuracy: 0.6757 - val_loss: 0.6727 - val_precision: 0.6757 - val_recall: 0.6757 - learning_rate:
1.0000e-05
Epoch 29/30
552/552 ----- 219s 396ms/step - categorical_accuracy: 0.7668 - loss: 0.4746 - precision: 0.7668 - recall: 0
.7668 - val_categorical_accuracy: 0.6812 - val_loss: 0.6746 - val_precision: 0.6812 - val_recall: 0.6812 - learning_rate:
1.0000e-05
Epoch 30/30
552/552 ----- 221s 399ms/step - categorical_accuracy: 0.7571 - loss: 0.4789 - precision: 0.7571 - recall: 0
.7571 - val_categorical_accuracy: 0.6812 - val_loss: 0.6697 - val_precision: 0.6812 - val_recall: 0.6812 - learning_rate:
1.0000e-05
```





	precision	recall	f1-score	support
0	0.86	0.59	0.70	1499
1	0.59	0.86	0.70	1025
accuracy			0.70	2524
macro avg	0.72	0.72	0.70	2524
weighted avg	0.75	0.70	0.70	2524

Confusion Matrix



5. CONCLUSIONS AND FUTURE WORK

With the rise in accessibility to image editing tools capable of producing forgeries, image forgery detection techniques have become increasingly vital. This paper introduces a detection method leveraging deep learning, specifically using a pre-trained model and transfer learning. The proposed method examines the discrepancies between an original image and its compressed version to generate a feature map, which is then fed into a pre-trained model to enhance detection accuracy. The technique was evaluated across eight different pre-trained models adapted for binary classification and compared to state-of-the-art methods.

The experimental results indicate that utilizing pre-trained models can significantly improve detection accuracy compared to traditional CNN-based approaches. Among the

eight models tested, MobileNetV2 achieved the highest detection accuracy, around 95%, with fewer training parameters. This led to faster training times, reduced computational costs, and lower system complexity and memory usage. Consequently, MobileNetV2 is recommended as a robust backbone for image compression techniques, effectively detecting both image splicing and copy-move forgeries with promising results.

Despite its high performance, the proposed model has several limitations. It struggled with data generalization, performing well on training data but failing to generalize to new, unseen data. It also faced challenges in detecting certain types of image forgeries, including novel techniques. Additionally, the model did not address localization of the forged areas and required significant computational resources and time for both training and inference, which could be a drawback in resource-constrained environments.

REFERENCES

- 1] Deep learning-based algorithm (ConvLSTM) for Copy Move Forgery Detection Mohamed A. Elaskilya, Monagi H. Alkinanib, Ahmed Sedikc and Mohamed M. Dessouky March 2021
- 2] Copy Move and Splicing Image Forgery Detection using CNN Devjani Mallick, Mantasha Shaikh, Anuja Gulhane and Tabassum Maktum 2022
- 3] A New Method to Detect Splicing Image Forgery Using Convolutional Neural Network Khalid M. Hosny, Akram M. Mortda, Nabil A. Lashin and Mostafa M. Fouda January 2023
- 4] Design of Automated Deep Learning-Based Fusion Model for Copy-Move Image Forgery Detection N. Krishnaraj, B. Sivakumar, Ramya Kuppusamy, Yuvaraja Teekaraman and Amruth Ramesh Thelkar 2022
- 5] Image Forgery Detection Using Tamper-Guided Dual Self-Attention Network with Multiresolution Hybrid Feature Fengyong Li, Zhenjia Pei, Weimin Wei, Jing Li, and Chuan Qin 2022
- 6] Image copy-move forgery detection and localization based on super-BPD segmentation and DCNN Qianwen Li, Chengyou Wang, Xiao Zhou & Zhiliang Qin 2022
- 7] Image forgery detection using Deep Neural Network Anushka Singh, Jyotsna Singh Jan 2022
- 8] Deep Learning-Based Digital Image Forgery Detection System Emad Ul Haq Qazi, Tanveer Zia and Abdulrazaq Almorjan March 2022
- 9] Efficient Approach towards Detection and Identification of Copy Move and Image Splicing Forgeries Using Mask R-CNN with MobileNet V1 Kalyani Dhananjay Kadam, Swati Ahirrao, and Ketan Kotecha Jan 2022.

10] Lightweight Deep Learning Model for Detection of Copy-move Image Forgery with Post-processed Attacks
Muhammad Naveed Abbas, Mohammad Samar Ansari,
Mamoona Naveed Asghar, Nadia Kanwal, Terry O'NeBrian
Lee