National University of Computer and Emerging Sciences, Lahore

Image Forgery Detection

Ruman Malik l21-5159 BS(CS)

Talha Dogar l21-7559 BS(CS)

Abdullah Nawaz i21-1393 BS(CS)

DIP Project

November 26, 2024

Chapter 1 Introduction

1.1 Intro to Domain

As digital media continues to expand rapidly, ensuring the authenticity of images and videos has become increasingly important. The rise of sophisticated manipulation techniques, such as splicing, copy-move, and deepfakes, has raised significant concerns about the credibility of visual content. Image forensics is a specialized field focused on detecting these manipulations, playing a crucial role in maintaining trust in digital media. With the advent of deep learning technologies, there have been substantial advancements in detecting image forgeries, providing powerful tools to identify even subtle alterations. However, the challenge persists in creating systems that can reliably detect forgeries across various manipulation methods while balancing high accuracy with low computational costs.

1.2 Problem Statement

The primary difficulty in detecting image forgery stems from the wide array of manipulation techniques and their increasing sophistication. Techniques such as splicing, copy-move, inpainting, and deepfake generation can modify images in ways that traditional detection methods often struggle to identify. Additionally, variations in image formats, resolutions, and post-processing transformations complicate the detection process. Many existing systems find it challenging to detect tampered images when they undergo common post-processing techniques like compression, resizing, or noise addition. This highlights the urgent need for more robust, scalable, and efficient forgery detection models capable of addressing these challenges in real-time applications.

1.3 Proposed Solution

To tackle these issues, this paper proposes leveraging advanced deep learning techniques—specifically Convolutional Neural Networks (CNNs), Hybrid Models, and self-supervised learning frameworks—to detect and localize image forgeries effectively. Our approach harnesses the detailed feature extraction capabilities of CNNs alongside the robust generalization strengths of newer models like Capsule Networks, Diffusion Models, and Recurrent Neural Networks (RNNs). Furthermore, we investigate incorporating preprocessing techniques such as Wiener filtering for noise reduction and segmentation algorithms for localized forgery detection. By integrating these advanced methodologies, we aim to improve both the accuracy and efficiency of forgery detection systems, making them more adaptable to various manipulation types and image conditions.

1.4 Major Contributions

This research presents several significant contributions to the field of image forgery detection:

- Thorough Review of Existing Techniques: We conduct an extensive analysis of current deep learning and traditional methods for detecting image forgery, highlighting their strengths and weaknesses.
- Introduction of Hybrid Deep Learning Models: We propose a hybrid model that combines CNNs with Capsule Networks and self-supervised learning techniques to achieve higher accuracy in forgery detection, even when faced with noise and image degradation.
- Evaluation Using Diverse Datasets: Our study assesses the proposed model using multiple diverse datasets (e.g., CASIA, MICC) and employs rigorous performance metrics such as accuracy, precision, recall, and F1-score to evaluate its effectiveness.
- Insights into Practical Implementation: The paper offers practical insights into the computational challenges and resource requirements involved in implementing deep learning-based forgery detection models in real-world scenarios, especially when handling large-scale datasets.

1.5 Organization of the Article

- Introduction to the problem domain along with a review of existing research.
- A comprehensive literature review discussing recent advancements in image forgery detection techniques, datasets utilized, and evaluation metrics.
- Presentation of our proposed hybrid deep learning model, detailing its architecture, feature extraction methods, and preprocessing techniques.
- Description of our experimental setup, dataset details, and performance evaluation based on various metrics including accuracy, precision, recall, and F1-score.
- Discussion on practical implementation challenges faced by our proposed model along with limitations and future research directions.
- Conclusion summarizing the major findings from our study.

Chapter 2 Literature Review / Related Work

This section comprises comprehensive literature reviews of research conducted by researchers for image forgery detection, delving into the methodologies and algorithms they used to solve this problem. Additionally, we will present a summarizing table at the conclusion of this section, offering a brief overview of the findings from each paper.

2.1 Detailed Literature Review

This section provides detailed information about the research that has been done by different researchers on Image Forgery Detection that will be helpful for our research.

Deep Learning-Based Approaches

2.1.1 M. Zanardelli et al.

2.1.1.1 Summary of the Research Item

The work [1]presents the current solutions of deep learning techniques for detecting picture forgeries, focusing on copy move masquerade and splicing. Since the DL approaches outperformed the classical approaches in numerous benchmark situations, the paper also provides emerging importance of DL approaches. However, a large deal of effort is used to explain copy-move and splicing assaults, while DeepFake generated content is only briefly explained.

The paper analyzing different DL architectures to identify pivotal models that are used for forgery detection. It also enumerates all of the useful datasets including the MICC-F2000, CASIA1 and CASIA2 which were used in training and testing diverse DL models. The best performing algorithms achieves over 95

2.1.1.2 Critical Analysis of the Research Item

More particularly useful for detecting then current state of the art advancements in DL based detection, this survey [1] provides a comprehensive and recent survey of different methods related to detection image forgery. The paper has a lot of strength because in the proposed DL model, it pays a lot of attention to several attacks like copy move, splicing and DeepFake for demonstrating the versatility of DL models.

However, in one sense, the study does not exactly avoid this drawback which is that no experimental comparison was made using a single platform to compare the efficiency of each of the approaches.

Further, while the survey incorporates several methods, it does not always perform comprehensive performance evaluation of all the methods so it cannot be considered as a simple study to measure the performance of each of the approach separately.

2.1.1.3 Relationship to the Proposed Research Work

The work reported in this paper is directly relevant to our recent activity on image forgery detection with deep learning. [1]. It gives us an idea of various advantages and disadvantages of difference DL approaches in detection of splicing and copy-move forgery either with or without any prior estimation of the angles. We also pay special attention to the fact that we have used appropriate datasets including MICC-F2000 and CASIA2 in the survey.

2.1.2 Younis Abdalla et al.

2.1.2.1 Summary of the Research Item

In [2], this work offers a research study to recognize and find copymove forgery through a new deep learning method. For classification and localization of the forgeries, the authors design a fusion model based on GAN and CNN. End to end training is done on MICC-F600 and we achieve nearly 95% accuracy. The architecture is divided into two branches: The first one employs GAN to identify forged images and the other CNN to identify the fake area.

The article then compares the effectiveness of fully implementating the model for forgery recovery in several different datasets and this does indeed work quite well for recall as well as precision. Our results show that the described above suggested model can detect forged verifications with F1 score, ranging between 0.49 to 6 and 0.88 to 35 for various datasets.

2.1.2.2 Critical Analysis of the Research Item

In [2], A new method based on a combination of CNN and GAN is proposed since it is a significant step forward in the field of picture forgery detection. The main strength of this method is that by generating the SSM and using the distance measure to identify forged region in photos, this method gives a complete solution to the copy-move forgery detection problem. Besides, to improve the model's more generalisable, other datasets, including MICC-F600 are introduced.

Nevertheless, the study may be bolstered, however, by comparing the proposed model with other state-of-art deep learning approaches. Indeed, positive results are expected for the work, however, further research will be required to determine the computing cost of the model, as well as required time for

training the model with large scale datasets.

2.1.2.3 Relationship to the Proposed Research Work

This paper [2] is related earlier in connection with image forgery detection and can be considered as directly related to our research. This study proposes the use of GAN and CNN architectures to establish a basis for their application in our project. The datasets and the measures adopted in this work can be ensured to fit the testing of our deep learning models and the high accuracy level attributed by the authors is sufficient indication of the applicability of the approach for forgery detection.

2.1.3 Tahira Nazir et al.

2.1.3.1 Summary of the Research Item

The paper [3] is aimed at detecting and segmenting copy-move forgery with an enhanced Mask Region-Based Convolutional Neural Network (Mask-RCNN). The authors present their own variant of Mask-RCNN model, where DenseNet-41 is used as the base network for deep feature extraction to improve the accuracy of the forgery detection. The procedure is effective against those post-processing attacks like noise, blurring and changes in brightness levels. Their system was tested on three standard datasets: A comparison is made of the proposed method with the state-of-the-art methods, including CoMoFoD, MICC-F2000, and CASIA-v2, with a precision of 98.12%, 99.02%, and 83.41% respectively which indicates the ability to work under variety of image transformations and post processing conditions. The results highlighted the efficiency of the proposed Mask-RCNN based on DenseNet-41 for forensic applications.

2.1.3.2 Critical Analysis of the Research Item

The work by Nazir et al. [3] was an insightful work that integrating Mask-RCNN with DenseNet-41 for accurate copy-move forgery detection. We, therefore, salute the great strengths of this approach which include the efficiency ofpost-processing attacks like noise and scaling which are traditionally a weakness for common forgery detection techniques. Second, the evaluation on three datasets (CoMoFoD, MICC-F2000, CASIA-v2) also shows the effectiveness of the proposed model under different image manipulation circumstances. However, the method's computational demands are not covered in detail especially in large scale utilization of the method. In the same way, further comparison with other deep learning techniques could have provided a better assessment of the performance of the proposed model.

2.1.3.3 Relationship to the Proposed Research Work

This paper [3] directly supports our research in image forgery detection. The use of DenseNet-41 for deep feature extraction and Mask-RCNN for segmentation provides a valuable reference for developing a similarly robust model in our project. The evaluation techniques and datasets used by the authors can serve as a benchmark for testing and validating our own deep learning models. Furthermore, the high accuracy results validate the effectiveness of deep learning in forgery detection, making this approach highly relevant to our research objectives.

2.1.4 Thakur and Rohilla

2.1.4.1 Summary of the Research Item

The research paper [4] This paper therefore provides a detailed survey of the recent advances in digital image manipulation detection specifically under the deep learning. The authors also explain why it is crucial to identify image splicing in an environment where manipulation is the new norm. With the help of the paper, important techniques for identifying different kinds of forgery such as copy-move, splicing, and JPEG compression are described. The paper also compares other approaches like CNN-based techniques while also stressing on the usage of synthetic datasets and deep learning for model training in forgery detection.

The paper highlights important results, such as achieving over 95% accuracy in detecting various tampering techniques using datasets like CASIA, Dresden, and MICC-F600. Moreover, specific methods like stack autoencoders, multi-scale CNNs, and deep Siamese networks are discussed, achieving accuracies ranging from 79% to 98% depending on the method and dataset. The article further explores techniques for detecting more complex manipulations such as median filtering and JPEG compression artifacts, with accuracy rates up to 99.10%.

2.1.4.2 Critical Analysis of the Research Item

This research [4] provides comprehensive insights into different approaches in image forgery detection and shows that deep learning models are successful in detecting the forged parts of images. One of its unique attributes is multi-type manipulation detection and with the use of multiple datasets increases its validation capabilities. Also, the paper offers a clear contrast between the traditional hand-crafted feature extraction method and the new available deep leaning based approaches and shows where the drawback of using hand crafted methods lies and why deep learning based methods are better.

However, the research could be extended to handle more system and training time comparisons in regards

to computational cost of each model particularly in real-time applications. Furthermore, despite high accuracy, the paper lacks elaboration on false positives, and the model's performance when working with highly compressed or low-quality images.

2.1.4.3 Relationship to the Proposed Research Work

This paper [4] is most closely related to the topic of our paper detecting image forgery using deep learning. The richness of datasets alongside the investigation of both general and specific approaches of methods include a beneficial groundwork for our project as a high accuracy outcome on datasets like CASIA and MICC-F600. The practical application of deeper learning models both in CNN and stack autoencoders is perfectly in line with the concept we have developed of having more reliable tool for detecting forgery within images.

2.1.5 Ferreira et al.

2.1.5.1 Summary of the Research Item

The research paper [5] We present an inclusive survey of digital image forgery, that covers the active and passive techniques and also acknowledges the latest innovations of deep learning methods. This paper presents methods including steganography, watermarking, and source camera identification and reviews the most recent approach in detecting forgeries which include copy-move forgery and resampling detection.

The need to recognize the message in the image, especially where there has been a copy-move manipulation, explains why the review highlights the efficiency of CNNs. The paper describes several deep learning structures that deployed high accuracy on CASIA, MICC-F600 and Dresden datasets while the most accurate architectures that were developed reached an accuracy of 98.5%. In addition, the paper analyses performance indicators, including precision, recall, and F1-score, and the results of some models are F1-score 0.97 of resampling detection tasks.

Other approaches are also described by Ferreira et al., including feature-based and block-based methods, and the authors utilize SIFT and SURF techniques to describe keypoint extraction, which are invariant to geometry and scale. For example, SIFT and SURF methods are described and, according to the author, detection based on the SURF algorithm is faster and more efficient than detection based on the SIFT algorithm.

2.1.5.2 Critical Analysis of the Research Item

This paper [5] offers a general insight into classical and contemporary approaches to DIF, with a focus on CNNs in image forgery detection. A strong point of the paper is the comprehensive comparison of methods on the large scale datasets CASIA and Dresden, where deep learning generated high levels of accuracy. (up to 98.5%) and its ability in detecting image manipulation. The paper consequently presents feature-based methods such as the SIFT method and block-based methods also proves the benefits of integrating feature- based methods with deep learning methods.

Yet, this paper delivers vast evidence regarding the efficiency of CNNs with few details comparing training times and computational complexity between various architectures, especially for real-time applications. Furthermore, the paper also points out that the decision of forgeries in Deep learning models is good, although in low quality highly compressed images this problem still remains unsolved.

2.1.5.3 Relationship to the Proposed Research Work

This paper [5] is very much related to the problem that we are investigating in the area of image forgery detection using deep learning. The part that discuss CNN-based methods, the superior performance in actually existing datasets such as MICC-F600 and CASIA, and their use in copy-move forgery detection are fully aligned with our project. This paper's concentration on deep learning architectures, especially CNNs is suitable for the creation of our image forgery detection tool.

2.1.6 Davide Cozzolino and Luisa Verdoliva

2.1.6.1 Summary of the Research Item

The research paper [6] presents a new approach called 'Noiseprint', a CNN based camera model fingerprint for digital image forensics. The authors put forward a system that rises the camera model artifacts, and diminshes the high level scene content thereby providing good results for the forgery localization. Extracted features are called Noiseprint and it is obtained by using a Siamese network with pairs of patches which belong to the same camera (label +1) or to different cameras (label -1). This experiment was conducted with 125 end-user cameras from 70 models and 19 brands and involves separate training and validation sets including 100 and 25 cameras, correspondingly. The dataset and experiment on which the performance of the model is analysed include DSO-1, VIPP, FaceSwap, NIMBLE, and others. In particular, the Noiseprint-based method of proposed attains the mean MCC score of 0.403 surpassing other methods such as Splicebuster and EXIF-SC. On average largest F1-score of 0.78 is achieved on the DSO-1 set with large splicing forgeries and without compressed images. Cross-validation of the data

also shows that the proposed Noiseprint methodology is accurate and stands first or second highest in most of the datasets making the tool reliable for forgery detection.

2.1.6.2 Critical Analysis of the Research Item

This study [6] contributes to the field of image forgery detection by suggesting the Noiseprint method and outperforms most methods which focus on suppressing high level image content and strengthening camera related peculiarities. The approach based on the generation of the noiseprints using Siamese networks allows the extraction of the exact model of the camera for precise identification of the crime scene, and thus, makes the method versatile for different kinds of forensics. One of the major advantages of this work is the ability to achieve a high level of success when tested on various forgeries examples, which show that the method performs well for different cases of forgery. Moreover, the experimental results show that the proposed Noiseprint-based method offers significant enhancement over that of traditional PRNU-based and JPEG artifact based methods. A limitation of this method is that they need several networks optimized for different QF JPEG classes, and it is not easy to implement. Further, the accuracy of the raw images such as Korus dataset is not that high and it shows potential problem in processing the uncompressed images. Possible future research targets include the use of Noiseprint methodology in combination with other forensics to improve the detection rate of the method and the existing issues.

2.1.6.3 Relationship to the Proposed Research Work

This paper [6] is directly related to our research field of detecting forged images as this work proposes a novel methodology for extracting camera model fingerprints based on CNNs. Noiseprint is a great approach in order to get a better idea of what kind of models could be used in the MoC in order to reliably detect manipulations of digital images in the described manner. This research forms a solid ground for benchmarking forgery deep learning models by using multiple datasets and other evaluation metrics such as MCC and F1-score. The integration of similar techniques and the evaluation of proposed systems will in turn improve on the proposed system accuracy and reliability.

2.1.7 Mohammed Fakhrulddin Abdulqader et al.

2.1.7.1 Summary of the Research Item

The paper by Abdulqader et al. [7] concerning the study of some features of tamper forgery detection in digital security images, specifically it concerns passive (blind) detection such as copy-move forgery

and image splicing. Due to availability of better editing software's required to manipulate the images, the authors therefore focus at the need to come up with better measures to detect the faked pictures. The fundamental forgery detection methods considered are copy-move, where a small segment of an image is shifted to a new position within the same image, and image splicing, where portions of different images are edited together in a way that does not allow a viewer to notice that it has been tampered with.

In their presented study, the authors reconsider the methods for detection and databases, as CASIA v1.0, CASIA v2.0, and Columbia while focusing on the statistical and machine learning as the classification tools. One of the dominant approaches highlighted is the extraction of hand-made features including HoG, LTE, DWT, and LBP; in conjunction with logistic regression classifiers, yields 99.5% accuracy on CASIA v2.0. This paper also proposes new methods, for example the multi-task FCNs used in splicing and which proved to provide better results than previous methods in tasks of coarse localization.

2.1.7.2 Critical Analysis of the Research Item

The research by Abdulqader et al. [7] gives significant contribution to the area of image forgeries as it contains passive as well as active detection methodologies. The area where this paper shines, or can be easily strengthened is the classification and especially the distinction of the copy-move and splicing forgery. There is also a good discussion of the dataset assessment which gives the reader the heads up on the features of large and easy available datasets like CASIA and CoMoFoD that can be used in the validation of models. Mainly, they use the classical/methodical logistic regression and the CNN-based models demonstrating how machine learning methods can be applied with deep learning ones to obtain better performance.

However, future work can include a detailed discussion of scalability, mainly concerning large applications of the paper in the forensic fields. Despite achieving high accuracy when trained with CASIA datasets (overall accuracy of 99.5% for CASIA v2.0), the specific paper is rather reserved when it comes to the question of the actual time and resources needed to train such models. Moreover, the comparison of deep learning based models with other advanced forgery detection schemes has not been done in a comprehensive manner, which can provide a valuable idea, whether deep learning based solutions are more effective in terms of real time or less hardware intensive environment.

2.1.7.3 Relationship to the Proposed Research Work

This paper [7] relates directly to our project of deep learning based image forgery detection in the sense that. The knowledge regarding copy-move and splicing techniques is specifically significant, as these represent the main kinds of forgery that should be identified using CNNs. Using CASIA and CoMoFoD to setting up this paper has helps to weighs up our selected datasets and the outcomes of the

models. Also, the multi-task learning approach described for splicing detection using edge and surface probability maps can be considered as the avenue for improving the accuracy of our model in detecting slight manipulation.

Furthermore, the literature review in this research covers related methods, such as logistic regression and CNN-based, which form a suitable background for our study. That signals high performance of similar models on standard datasets means that similar models can be adopted or enhanced in this study to achieve the objective aimed at achieving when developing a dependable and robust forgery detection tool. The evaluation criteria suggested in this paper such as accuracy, precision and F1-score will also fit into our performance evaluation framework.

2.1.8 Ritu Agarwal and Om Prakash Verma

2.1.8.1 Summary of the Research Item

The research article [8] This paper presents a novel copy-move forgery detection system utilizing feature extraction from deep learning models and matchers. It is proffered that the authors propose a method which uses VGGNet to obtain multiple scales of feature maps of the tampered image and utilizes Adaptive Patch Matching (APM) algorithm to detect matched area. The method is implemented on the MICC-F220 dataset, which includes 220 images, and yields the outperforming accuracy of 95% compared to other available techniques. As demonstrated in the proposed works, it exhibits high performance against a range of image manipulation including rotation, scaling and noise. This paper reveals TNR = 0.971, FNR = 0.092, FPR = 0,550, Precision = 98,0 26, Recall = 89,583, F1-Score = 0,922, and E-measure for various conditions. The computational experiments carried out in this study showed that the proposed model is superior to the traditional procedures, such as the SWT and spatial features based methods.

2.1.8.2 Critical Analysis of the Research Item

This study [8] provides an important contribution to the area of image forgery detection by employing feature learning on the use of deep learning and adaptive matching for forgery detection. One of the major advantages of this work is the ability to incorporate multi-scale features of VGGNet, in order to detect forgeries under different transformations. Besides, the APM algorithm methodology for forging area detection is efficient. However, a weakness of the approach put forth for the study is that it cannot cope with multi-cloned attacks since there is confusion in matching multiple altered patches. A limitation of this work is that only one dataset has been used (MICC-F220); this limits the model's applicability for other types of forgery. These can be expanded to future works that analyze multi-cloned

attack detection and validation on a larger dataset to establish the expansibility of the study.

2.1.8.3 Relationship to the Proposed Research Work

This paper [8] is also relevant to our work on the detection of image forgery. From the method of extracting multi-scale features based on VGGNet by the authors and APM for the matching the regions, we learned few important things about designing our CNN architectures. Furthermore, the chosen experimental results and evaluation metrics applied throughout this study define the performance of the developed models. Since we are also using similar techniques and performance indices, our is project is poised to enhance the detection and localization of image forgeries as shown in this research.

Traditional and Hybrid Methods

2.1.9 Priyanka et al.

2.1.9.1 Summary of the Research Item

The research paper [9] introduces a method for copy-move forgery detection based on the use of DCT and SVD for the compression of images. The method is targeted to solve the problem of copy-move forgery detection in digital images including in conditions where images may be subjected to post-processing such as rotation, scaling, noise addition, and JPEG compression. Copy-move forgery takes place when a particular region in the image is removed and then placed at another position to either conceal or manipulate knowledge. By using DCT for feature extraction together with SVD for dimensionality reduction, the detection becomes invariant to geometric transformations. The process includes converting of the images to blocks and using the DCT to come up with the feature vectors. These vectors still undergo SVD and then a Simple Classifier such as the Support Vector Machine (SVM) is applied. Localization of the forged regions is then obtained by applying K-means clustering.

The authors tested the system on two datasets: the CMFD benchmark dataset (48 images) and the Columbia image splicing dataset (1,845 images). The proposed method classified 93% of substances in CMFD benchmark dataset and 90% in Columbia dataset. It also did well in the post-processing conditions, it had higher accuracy, precision, recall and F1 score than other state-of the-art methods. The paper under discussion demonstrates that the proposed method outperforms the previous methods based on such features as Zernike moments, SIFT, SURF for forgery detection, showing the method's efficiency in respect of transformations.

2.1.9.2 Critical Analysis of the Research Item

This study [9] outperforms the previous copy-move forgery detection methods by implementing DCT and SVD for modifying features and reducing the dimensions. A particularly noteworthy strength of the paper is the capability for dealing with the geometric transformations in real forgeries, including rotation, scaling and noise addition. The authors have chosen SVM for classification and K-means clustering for localization which proves to be efficient, and the results obtained are highly accurate which makes it suitable for its use (Accuracy of 93 % on CMFD benchmark data set). Also, the study indicates that the outcomes are more accurate than the previous methods like the Zernike moments, SIFT and SURF in terms of precision, recall and F1 measure.

The paper does not, however, discuss the possibility of using other deep learning models like CNNs to enhance performance when indeed deep learning practices are becoming popular in image analysis. An area of limitation is that while the model achieves a fairly good recall rate under certain post processing conditions, there is still potential for increased improvement on it. To improve the model's robustness, there is a potential to increase the dataset selection of more diverse forgeries and using the ensemble methodology.

2.1.9.3 Relationship to the Proposed Research Work

This research [9] is closely related to our project in using deep learning to identify cases of image forgery. Despite the use of the conventional methods such as DCT and SNV, the paper clearly explains the significance of proper feature extraction as well as the part played by classification models like SVMs in forgery detection. The corpus of this paper builds on this study to use Convolutional Neural Networks (CNNs) in order to yield slightly higher accuracy in large datasets and for more complex forgery methods.

In this study, the datasets available including the CMFD benchmark and Columbia image splicing datasets could be used to conduct comparison and evaluation in our project to compare how well the model performs in comparison with other developed models. What challenges could be avoided if M/mko and/or additional local data were incorporated within the context of what other related actions could be obtained applying CNNs for the detection of even higher quality forgeries?

2.1.10 Yıldız Aydın et al.

2.1.10.1 Summary of the Research Item

The research paper [10] suggests a new CMFD approach based on LIOP descriptors, more specifically on first-order local intensity patterns. The method unfolds makes an effort to handle some of the problems that post-processing attack brings about like JPEG compression, noise insertion and geometric transformation. The system initially invets the images into YCbCr color space, and the system then extracts the LIOP of the Y, Cb and Cr channels in order to using them in feature matching. The authors also employ Random Sample Consensus (RANSAC) test and a parallelism condition for eradicating false to make a better identification of the forged objects.

The performance of the model was tested on two datasets: CoMoFoD and GRIP. The CoMoFoD dataset consists of 200 manipulated images, where the images are of different types of manipulations on the face, whereas the GRIP dataset has 80 manipulated images, of different sizes and shapes. In context with the CoMoFoD dataset and color reduction attacks, the method has a precision of 99.37%, recall 88.27%, and F1-score of 93.49. Simple forgery detection achieved an F1-score of 1.0 on all the GRIP dataset.

2.1.10.2 Critical Analysis of the Research Item

This paper [10] contributes fairly to CMFD by using LIOP features in the YCbCr color space to improve on the identification of forgeries in the images that have been processed with post processing attacks. A strength of the paper can be seen in supplementing what the block-based as well as the keypoint-based approach cannot: homogenious image regions and postprocessing artefacts. Incorporation of RANSAC with a parallelism condition for filtering out the false matches, strengthens the general method, further.

Nevertheless, one of the major drawbacks of the present investigation is the absence of a comparison with more recent methodologies, as the deep learning based forgeries detection, like Convolutional Neural Networks (CNNs), which are currently the state of the art. Also, there are limitations with the proposed method to detect forgeries especially when it is applied in fairly homogeneous image regions since it is based on keypoint features. The analysis on a larger number of datasets, particularly datasets with splicing forgery, will strengthen the argument of the proposed method.

2.1.10.3 Relationship to the Proposed Research Work

The work presented in [10] is highly related to our study focused on image forgery detection based on deep learning methods. The feature extraction methods described in this paper might give direction for increasing the robustness of our system that deals with detecting various forgery types including splicing and copy-move based on CNN architectures. The study will also be useful to assess the proposed models using the utilized datasets, namely CoMoFoD and GRIP. If the results were to be integrated with keypoint-based methods as an additional to the deep learning approach it would give better detection of forgeries in the difficult regions of the image.

2.1.11 Neeti Taneja et al.

2.1.11.1 Summary of the Research Item

Taneja et al. [11] is highly related to our study focused on image forgery detection based on deep learning methods. The feature extraction methods described in this paper might give direction for increasing the robustness of our system that deals with detecting various forgery types including splicing and copy-move based on CNN architectures. The study will also be useful to assess the proposed models using the utilized datasets, namely CoMoFoD and GRIP. If the results were to be integrated with keypoint-based methods as an additional to the deep learning approach it would give better detection of forgeries in the difficult regions of the image.

2.1.11.2 Critical Analysis of the Research Item

This work [11] Delving deep into the realm of anti-forensics this study uncovers the ways in which modern forensic approaches are challenged by sophisticated manipulation tactics. By offering a detailed comparison of methods used to counter forensics especially focusing on JPEG compression contrast enhancement and median filtering this paper stands out for its significant insights illustrated through examples and findings from different forensic detection tools.

Despite the authors offering insightful observations their work falls short in providing an in-depth comparison to newer machine learning approaches for spotting forgeries particularly those that utilize cutting-edge deep learning structures such as Convolutional Neural Networks. With the advancement of anti-forensic methods it becomes imperative to explore the ways in which deep learning models might be refined or created anew to combat these evolving tactics. In addition, it would be helpful if the paper explored how to handle false positives that might pop up in scenarios where anti-forensic techniques are at play. This is crucial because even tiny changes in what is detected can lead to big problems in

situations that matter a lot, such as in court cases or when checking if media is real.

Even with its drawbacks the study shines by charting the obstacles digital image forensic experts are up against. It lays down a solid base for upcoming studies that want to craft tough methods to spot anti-forensic tactics. These methods could help fight identity theft and solve cyber crimes.

2.1.11.3 Relationship to the Proposed Research Work

This paper [11] the relevance of this work to our project is profound, particularly with respect to its exploration of JPEG compression and anti-forensic methodologies. In our pursuit of image forgery detection via Convolutional Neural Networks (CNNs), the revelations concerning anti-forensic strategies will be instrumental in enhancing our methodology for developing more resilient models. The findings pertaining to the mitigation of blocking and quantization artifacts are of particular significance, as we endeavor to improve our model's capacity to identify forgeries, even in the presence of advanced anti-forensic techniques. Moreover, the comparative analysis of datasets and the assessment of True Positive Rates (TPR) will be critical in shaping our evaluation metrics as we examine the efficacy of our CNN-based detection frameworks. The manuscript emphasizes the necessity of establishing robust detection mechanisms, which is in direct accordance with the objectives of our research endeavor.

2.1.12 S B G Tilak Babu et al.

2.1.12.1 Summary of the Research Item

The research paper [12] presents a holistic methodology for the detection of Copy-Move Forgery (CMFD) through the utilization of a refined technique predicated on statistical attributes. The authors advocate a dual-phase framework for the identification of forgeries. In the initial phase, the image under suspicion is categorized as either fabricated or genuine. The subsequent phase entails the localization of the counterfeit segment via a block-matching algorithm. To facilitate this process, the image undergoes decomposition into various orientations utilizing the Steerable Pyramid Transform (SPT), from which features are extracted using the Grey Level Co-occurrence Matrix (GLCM). These extracted features are subsequently employed to train an Optimized Support Vector Machine (OSVM), which executes the conclusive classification.

The model was evaluated on widely-used datasets such as CoMoFoD and CASIA, showing promising results. It remained resilient even when the images underwent common post-processing attacks, including JPEG compression, scaling, and rotation. With a True Positive Rate (TPR) of up to 99% and an overall accuracy of 99%, the model demonstrated its ability to robustly detect copy-move forgeries. The high accuracy in challenging scenarios highlights the effectiveness of combining SPT and GLCM

features for forgery localization.

2.1.12.2 Critical Analysis of the Research Item

The research endeavor put forth by Babu et al. [12] establishes a robust framework for the detection of copy-move forgery, emphasizing both the processes of detection and localization. A prominent strength inherent in this manuscript is the employment of an optimized machine learning model (OSVM) alongside GLCM-based features, which contribute to the efficiency and accuracy of the detection methodology. The model's robustness was convincingly illustrated through its evaluation on datasets such as CoMoFoD and CASIA, exhibiting consistent performance across a variety of image alterations. The incorporation of Steerable Pyramid Transform (SPT) for feature extraction significantly enhances the system's resilience against common transformations, thereby representing a noteworthy advantage.

Notwithstanding its merits, the study does present certain limitations regarding computational efficiency. The block-matching methodology utilized for localization may prove to be resource-intensive, particularly when confronted with large-scale or high-resolution images, which could potentially constrain its scalability. Furthermore, although the findings of the study are encouraging, the absence of comparative analysis with more sophisticated deep learning models, including Convolutional Neural Networks (CNNs) or hybrid frameworks, indicates potential avenues for enhancement. A thorough examination of the system's performance in terms of execution time and memory utilization on larger datasets would yield more profound insights into its practical applicability.

2.1.12.3 Relationship to the Proposed Research Work

This paper [12] the research is of significant importance to our investigation regarding image forgery detection, as it establishes a robust methodological framework. The implementation of GLCM and OSVM is pertinent to our endeavor, which similarly seeks to identify copy-move and various other types of image alterations. The selection of datasets, including CoMoFoD and CASIA, is consistent with our evaluative approach, thereby facilitating the comparison and assessment of our models' performance. Additionally, the documented accuracy and resilience in countering post-processing attacks provide a standard against which we can measure the efficacy of our own model.

Moreover, the paper's emphasis on localization techniques offers critical perspectives on the precise identification of forged regions, which constitutes a fundamental element of our research. While the execution of a comparable block-matching algorithm may entail significant computational costs, it could still prove beneficial for our system. Nonetheless, we may explore the integration of more sophisticated

deep learning techniques, such as Convolutional Neural Networks (CNNs), to enhance both efficiency and scalability in extensive applications. The limitations identified in this study serve as a crucial reminder to mitigate computational burdens and investigate optimization strategies within our project.

2.1.13 Ashgan et al.

2.1.13.1 Summary of the Research Item

This study [13] focused on detection of two types of image forgeries i.e copy-move and splicing. It addressed the limitations of traditional approaches that cater only one type of image forgery in their study. The researchers used approach to compare difference in quality of original and modified images in compressed form. Deep learning models were trained based on this approach. To increase the accuracy of models, transfer learning technique was applied in this study using the standard dataset of CASIA 2.0. The paper used eight deep learning architectures for their experimentation and showed that MobileNet was best performing with accuracy of about 95%.

2.1.13.2 Critical Analysis of the Research Item

This study takes into account two image forgery techniques in a single system and provide comparison among different deep learning models. It provides insight that how different models can perform for same dataset. Further, they acquired high accuracy of MobileNet that can be used in systems that are constrained to computation limitations. The problem of localization of forged regions in images is not discussed in the study. Localization has fundamental role in detecting image forgery. The study claimed to have above 90 percent accuracy but it is not generalized well and there are problems of overfitting in it.

2.1.13.3 Relationship to the Proposed Research Work

This study is related highly to our approach of identification of modified images. The comparative study shows about performance of different models that is helpful for our research. Transfer learning approach can be used for achieving higher accuracy with limited resources in our research as applied in our study.

2.1.14 Yang et al.

2.1.14.1 Summary of the Research Item

The research paper [14] introduces a novel copy-move forgery detection (CMFD) algorithm using a two-stage filtering (TSF) technique. The TSF consists of a Grid-Based Filter (GBF) and a Clustering-Based

Filter (CBF), which work together to reduce false positive keypoint matches. The algorithm is evaluated on three datasets: IMD, CoMoFoD, and CMHD. Keypoint extraction is performed using SIFT, SURF, and ORB methods. The study achieves an F1-score of 82% on the CMHD dataset and up to 63.8% on the CoMoFoD dataset.

In comparison to existing methods, the proposed TSF algorithm significantly improves the precision and F1 scores for detecting forgery regions in images, especially under geometric and post-processing operations. The datasets used include IMD (1344 forged images), CoMoFoD (4800 forged images), and CMHD (108 forged images). The results show that TSF-SIFT performs the best, achieving superior precision and recall across different attack scenarios.

2.1.14.2 Critical Analysis of the Research Item

The research [14] presents an innovative two-stage filtering algorithm for reducing false positives in CMFD. One major strength of this approach is the integration of two complementary filters that enhance detection accuracy. The use of multiple datasets, such as IMD, CoMoFoD, and CMHD, demonstrates the algorithm's robustness against different types of forgeries, including scaling and JPEG compression.

However, the performance of the algorithm is highly dependent on the keypoint extraction method used. TSF-SIFT consistently outperforms TSF-SURF and TSF-ORB, raising questions about the generalizability of the algorithm with other feature extraction techniques. Additionally, the study does not provide a detailed computational cost analysis, which would be crucial for practical real-time applications.

2.1.14.3 Relationship to the Proposed Research Work

This paper [14] is highly relevant to our project on detecting image forgery using deep learning techniques. The focus on copy-move forgery detection and the use of keypoint-based methods align with our objectives. Moreover, the datasets and evaluation methods used in this paper provide a solid benchmark for testing the effectiveness of our proposed deep learning models for forgery detection.

2.1.15 Fatemeh Zare Mehrjardi et al.

2.1.15.1 Summary of the Research Item

The paper [15]This paper conducts a comprehensive review of methodologies grounded in deep learning for the purpose of image forgery detection. It categorizes image forgeries into distinct classifications, including splicing, copy-move, inpainting, and object removal, while elucidating the application of deep learning frameworks, notably Convolutional Neural Networks (CNNs), Recurrent Neural

Networks (RNNs), and Region-based Convolutional Neural Networks (R-CNNs), in the detection and localization of such forgeries. The review encompasses an array of datasets, featuring prominent examples such as MICC, CASIA, and CoMoFoD, which are routinely employed for the benchmarking of forgery detection systems. The authors present a meticulous assessment of these datasets, addressing their respective limitations and deliberating on the transformations and post-processing methodologies that frequently hinder detection efficacy. Furthermore, the review accentuates the importance of evaluation metrics, including accuracy, precision, recall, and F1-score, as critical indicators for assessing the reliability and performance of the models deployed in forgery detection endeavors.

Additionally, the manuscript highlights both advancements and obstacles within the domain of deep learning-based forgery detection. The authors offer an exhaustive comparison between deep learning strategies and traditional techniques, underscoring the resilience of deep learning models in managing intricate manipulations, such as geometric transformations and the introduction of noise. The paper also examines a variety of preprocessing strategies (e.g., resizing, color space conversions) and post-processing techniques (e.g., morphological operations) that are vital for the enhancement of forgery detection algorithms. By concentrating on the progress in deep learning, the investigation illustrates the potential for forthcoming advancements in this domain, advocating for increased focus on hybrid methodologies that amalgamate traditional and deep learning strategies to achieve enhanced performance.

2.1.15.2 Critical Analysis of the Research Item

The research by Mehrjardi et al. [15]The manuscript makes a significant contribution by delivering a comprehensive examination of contemporary deep learning techniques employed for the detection of image forgery. A particularly notable advantage of this survey is its extensive breadth, encompassing a diverse array of datasets, types of forgery, and architectural frameworks, thereby providing an all-encompassing overview for scholars entering this domain. The emphasis on Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), alongside their proficiency in addressing various forgery modalities, furnishes the audience with an understanding of the merits of deep learning relative to conventional methodologies, including enhanced resilience to image post-processing operations such as compression and the introduction of noise. Moreover, the thorough examination of dataset inadequacies—such as the absence of ground truth in certain datasets like MICC—illuminates the obstacles that researchers encounter in the training and validation of forgery detection systems.

Nonetheless, the investigation could be augmented by incorporating additional quantitative performance comparisons among the disparate architectures discussed. For example, while the manuscript underscores the robustness of CNNs and R-CNNs, it fails to present a comparative table detailing metrics

such as training duration, memory utilization, and accuracy across various architectures. In addition, the document lacks a critical discourse on computational resources, particularly regarding the scalability of these deep learning frameworks when deployed on large-scale, real-world datasets. Addressing these considerations would yield a more comprehensive perspective on the practicality of implementing these models in real-world applications, particularly for organizations that may encounter hardware constraints.

2.1.15.3 Relationship to the Proposed Research Work

This paper [15] aligns closely with our project on image forgery detection using deep learning. Its review of CNN-based methods provides a foundational reference for our model development. The benchmark datasets discussed, such as MICC and CASIA, are relevant for our work, and the evaluation metrics proposed in the survey—accuracy, precision, recall, and F1-score—will be used to assess our system's performance. The insights into different forgery detection techniques, such as splicing and copy-move, will be useful in the design of our tool.

2.1.16 Rupesh D.Sushir, Dinkar Govindrao Wakde AND Sarita S.Bhutada

2.1.16.1 Summary of the Research Item

This study [16] also provide comprehensive approach for image forging detection using deep learning method. It also addresses the issue of low quality image that lead to poor performance of model. It uses Wiener filtering to remove noise and improve sharpness in image for enhancement. For deep learning model, researchers used capsule network and autoencoder framework named as Hybrid DCCAE. In addition, researchers used Hybrid DTCWT and DCT alongside VGGNET to extract features from images. For localization, segmentation was performed using Adaptive Density-Based Fuzzy Clustering to separate forged and non forged pixels. This paper used three different types of dataset (CASIA V1, Coverage and GRIP) that ensures diversified data and got different accuracy of 99.23%, 98.75% and 98.07% respectively.

2.1.16.2 Critical Analysis of the Research Item

In the present investigation, a hybrid model is employed for the classification task with the objective of enhancing the performance of classification models. This study implemented image preprocessing techniques on the dataset and employed various methodologies to extract features from the images. Despite the dataset being diversified, the limited size of the data has led to potential overfitting issues, as indicated by the authors. In the present investigation, a hybrid model is employed for the classification

task with the objective of enhancing the performance of classification models. This study implemented image preprocessing techniques on the dataset and employed various methodologies to extract features from the images. Despite the dataset being diversified, the limited size of the data has led to potential overfitting issues, as indicated by the authors.

2.1.16.3 Relationship to the Proposed Research Work

The different image processing techniques and feature extraction approaches used in this study can be helpful for our study. Although hybrid model architecture is used in this study that is not related directly for our work but it showed that how deep learning models other than commonly used architectures can be used.

2.1.17 Luisa Verdoliva et al.

2.1.17.1 Summary of the Research Item

This paper [17] in detail, presents all methodologies concerning the identification of forged images in various media. Subsequently, it has expanded on all the various techniques employed for manipulating images in media, as well as the strategies utilized for the examination of the media images and the detection of deepfake contents. Image manipulation techniques through splicing, copy-move, inpainting, and post-processing are employed, whereas contemporary practices involve GANs, autoencoders, and domain translation methods to synthesize counterfeit images or videos. For the discussion of deep learning methods, supervised CNN models are used to detect image forgery while the deepfake is detected based on the frequency domain analysis along with deep learning approaches. Detection of facial manipulation is done using RNNs and LSTM models. It emphasizes further that any generalized model would require a diversified dataset for its development.

2.1.17.2 Critical Analysis of the Research Item

In the study authors provide the general overview of image forgery techniques used in media and techniques used for its detection. It also provides information of changing methods for changing images with advancement in technology. This study discussed different areas that are related to improving detection systems for identification and classification of forged images. Any practical implementation is not provided in the study.

2.1.17.3 Relationship to the Proposed Research Work

This study is really helpful in understanding of image forgery domain in the light of deep learning. As mentioned earlier, it does not provide any implementation details but it must be helpful for clear understanding that can reinforce our research in image forgery detection.

2.1.18 DOHYUN KIM , WONHYUK AHN AND HEUNG-KYU LEE

2.1.18.1 Summary of the Research Item

This paper [18] focused on investigating and enhancing JPEG and double JPEG (DJPEG) detection in digital images, with special consideration to anti-forensic techniques. JPEG compression is the most widely used technique for storage and transmission of images. This kind of technique generates many identifiable artifacts that may be useful in forensic analysis. The study under consideration addresses two deficiencies that methods of anti-forensics using existing techniques suffer from: lack of control over non-aligned cases, and low-quality images. To defeat these challenges, the authors proposed a novel approach via CNN that not only improves the quality of an image but also has the added advantage of being able to deceive detectors, both for JPEG and DJPEG. They evaluate their algorithm for approach using data like BossBase 1.01 and BOWS2, where they measure detection accuracy along with minimum decision error.

2.1.18.2 Critical Analysis of the Research Item

In this study hybrid model is used for classification task intended to increase performance of classification models. This research applied image pre processing techniques on data and used different approaches to extract features from images. Although dataset used was diversified but size of data is limited that resulted in overfitting possibilities as mentioned by authors.

2.1.18.3 Relationship to the Proposed Research Work

2.1.19 Zeqin et al.

2.1.19.1 Summary of the Research Item

This paper [19] introduces a novel approach to image forgery detection and localization through a two-stage self-supervised framework that integrates the denoising diffusion paradigm with a fine-tuning phase. In the initial stage, a pre-trained encoder is employed to extract general, macroscopic features from the image, while the decoder focuses on learning the microscopic representation. This methodol-

ogy directs the model's attention to mesoscopic features, which are essential for identifying tampering. Additionally, a fine-tuning phase incorporates a specialized Edge Cue Enhancement Module, facilitating improved detection in areas affected by tampering. Experiments conducted using publicly available datasets, including both artificially and AI-generated manipulated images, demonstrate superior accuracy and robustness in comparison to existing state-of-the-art models.

2.1.19.2 Critical Analysis of the Research Item

This paper is the incorporation of the macroscopic and microscopic features through diffusion, thereby intensifying the model for the detection of subtle traces due to any possible tampering process. Addition of ECEM imparts a fine grain detection capability further refining the results. It mainly relies on pre-trained encoders, which may be limiting in terms of flexibility when applied to novel forms of image manipulations not present in the training datasets. Additionally, whereas the robustness of the model is a major advantage, the computational overhead of the two-stage process and the complexity of the diffusion paradigm might make it fairly challenging for certain real-time applications.

2.1.19.3 Relationship to the Proposed Research Work

In image forgery detection, it employs diffusion models in this paper to match most modern approaches that depend on learning fine-grained features. Unlike the traditional deep learning algorithms and CNNs used to detect image forgeries or artifacts due to compression at the pixel level, DiffForensics advances on these by learning mesoscopic features across between global and local analysis. That enhances detection as well as localization. It strengthens image forgery detection using deep learning algorithms.

2.1.20 Chothmal Kumawat and Vinod Pankajakshan

2.1.20.1 Summary of the Research Item

This paper [20] examines the detection of JPEG compression in images through forensic analysis aimed at authenticating images. The study highlights the statistical differences in the discrete cosine transform coefficients between JPEG images and their uncompressed counterparts, particularly focusing on the subbands of the AC component. A proposed JPEG compression detection method incorporates calibration techniques to enhance resilience against anti-forensic and post-processing attacks. This approach is grounded in a solid theoretical framework, and experiments conducted across various image sizes and compression factors validate the effectiveness of the method. The dataset utilized includes thousands of uncompressed and JPEG images, subjected to various post-processing techniques such as filtering and

noise addition, to assess the performance of the detector.

2.1.20.2 Critical Analysis of the Research Item

This paper makes a significant contribution to the field of image forensics, particularly concerning images that have undergone post-processing or anti-forensically altered JPEG-U formats. The proposed method is grounded in both theoretical and empirical research, enhancing its reliability. However, the study's performance seems somewhat constrained by the JPEG quantization settings and does not generalize effectively to images that have experienced substantial transformations, such as cropping or rotation. While experiments have been carried out on a variety of datasets, there remains potential for further enhancement of the machine learning models' outputs through the application of post-processing techniques.

2.1.20.3 Relationship to the Proposed Research Work

Deep learning techniques like CNNs can identify subtle and diminished image characteristics, such as compression artifacts, that may help in detecting tampering. Moreover, deep learning allows it to be more adaptive to different manipulation techniques, including splicing and resampling. With statistical methods implemented within this paper and a fusion of the deep learning, detection can be expected to be more robust with the emergence of new manipulation forms.

2.2 Literature Review Summary Table

Deep Learning-Based Approaches

Author	Method	Results	Limitations
M. Zanardelli et al. [1]	Survey of deep-learning approaches for image forgery detection. Focuses on techniques for copy-move, splicing, and DeepFakes using CNN and GAN models.	Highlights key architectures and datasets used, such as MICC-F2000 and CASIA2. Accuracy rates exceeding 95% on standard benchmarks are reported.	No experimental performance comparison on a unified benchmark. Limited details on some performance metrics for discussed techniques.
Younis Abdalla et al. [2]	A hybrid model combining GAN and CNN for copy-move forgery detection and localization.	Achieved 95% accuracy on the MICC-F600 dataset with F1 score of 0.8835.	Does not compare performance with other recent deep learning approaches. Computational cost and scalability not discussed in detail.
Nazir et al. [3]	Custom Mask-RCNN with DenseNet-41 on CoMoFoD, MICC-F2000, CASIA-v2 datasets	Precision: 98.12% (CoMoFoD), 99.02% (MICC-F2000), 83.41% (CASIA-v2)	High computational cost not discussed, lacks comparison with other deep learning methods
Thakur and Rohilla [4]	Used CNN-based models, stack autoencoders, and multi-scale CNNs for detection.	Achieved accuracy up to 99.10% on synthesized datasets.	Computational cost and training time not discussed.
Ferreira et al. [5]	CNNs, SIFT, SURF, Block-based methods for copy-move forgery detection.	Achieved up to 98.5% accuracy on CASIA and MICC-F600 datasets, F1-scores up to 0.97 for resampling detection.	Computational costs not discussed in detail, difficulties in detecting low-quality and compressed images.
Agarwal et al. [8]	VGGNet with APM algorithm	Achieved 95% accuracy, 98.026 Precision, 0.971 TNR	Does not handle multi-cloned attack scenarios
Cozzolino and Verdoliva [6]	Noiseprint using Siamese CNN	Achieved MCC of 0.403, highest F1-score of 0.78	Performance on raw images slightly lower, requires multiple QF-specific networks
Mohammed Fakhrulddin Abdulqader et al. [7]	HoG, LTE, DWT, LBP, Logistic Regression, Interquartile Range (IQR), SIFT, Multi-domain CNN (RGB + DCT)	99.5% accuracy (CASIA v2.0), 97% (Columbia), 95% (JPEG compression)	Performance varies across datasets; lower accuracy on smaller datasets; not robust to all compression types

Traditional and Hybrid Methods

Author	Method	Results	Limitations
Priyanka et al. [9]	Used DCT and SVD-based technique to detect copy-move forgery. Applied SVM for classification and K-means clustering for localization.	Achieved 93% accuracy on CMFD benchmark dataset and 90% on Columbia dataset.	Method does not explore deep learning techniques like CNNs for potentially better performance.
Yıldız Aydın et al. [10]	Used LIOP features in YCbCr color space, combined with RANSAC and parallelism condition.	Achieved precision of 99.37%, recall of 88.27%, and F1-score of 93.49% on CoMoFoD dataset.	Lack of comparison with CNNs and challenges with homogeneous image areas.
Taneja et al.[11]	Reviewed JPEG, contrast enhancement, and median filtering anti-forensics, and proposed anti-forensic techniques to remove traces left by image processing.	Reported significant reduction in true positive rates for JPEG compression artifacts. For example, TPR dropped to 0.22 for a quality factor of 80.	Did not address modern deep learning models like CNNs for anti-forensic detection; potential false positives not thoroughly analyzed.
S B G Tilak Babu et al. [12]	Used OSVM with SPT and GLCM for detecting copy-move forgery and localization.	Achieved TPR of 99% and accuracy of 99% on CoMoFoD and CASIA datasets under no post-processing attacks.	Computationally intensive when dealing with high-resolution images and lacks comparison with CNN-based models.
Yang et al. [14]	Proposed a two-stage filtering method using GBF and CBF for CMFD.	Achieved an F1-score of 82% on CMHD and 63.8% on CoMoFoD.	Limited performance with SURF and ORB compared to SIFT.
Fatemeh Zare Mehrjardi et al. [15]	Survey on deep learning techniques including CNN, RNN, and R-CNN	Highlights several datasets like MICC, CASIA, and CoMoFoD	Lack of computational cost analysis and model scalability insights

2.3 Conclusion

The reviewed research papers offer significant insights into the expanding domain of image forgery detection, particularly in managing intricate manipulations such as copy-move, splicing, and DeepFakes. A majority of the methodologies utilize deep learning models, including CNNs, GANs, and hybrid architectures, which have demonstrated high accuracy on established datasets such as CoMoFoD, MICC-F600, and CASIA. Notable techniques like DenseNet-41, VGGNet, and Siamese CNN have proven effective, particularly in countering post-processing attacks such as noise and compression.

Nevertheless, several limitations are evident within the existing literature. Numerous studies do not provide thorough performance comparisons on a standardized benchmark, and the computational demands, especially concerning large-scale, high-resolution images, are frequently neglected. While traditional feature extraction methods (SIFT, SURF) continue to perform adequately in specific scenarios, they

tend to falter with low-quality images, where deep learning approaches demonstrate greater resilience.

These insights are directly pertinent to our research objectives, as we seek to develop a deep learning-based solution for forgery detection. By concentrating on model scalability, performance in post-processing scenarios, and computational efficiency, we can address the gaps identified in the current literature and enhance the robustness of forgery detection systems.

Chapter 3 Proposed Approach and Methodology

3.1 Methodology

This section describes the methodology used to develop the ViT-VAE Net model for detecting image forgeries, specifically by combining the Visual Transformer (ViT) with a Variational Autoencoder (VAE) network. The ViT-VAE Net model leverages the strengths of Visual Transformer encoders for effective feature extraction and VAE for deeper, nuanced representation of manipulated regions. This architecture also includes a Multi-Layer Perceptron (MLP) decoder for efficient segmentation of detected forged areas, as illustrated in Figure 3.1.

3.1.1 Visual Transformer Encoder

The Visual Transformer Encoder (ViT) is a key component of the model, designed to break down image data into meaningful patches and process each patch independently. This approach allows the model to capture details within the image without being constrained by specific input dimensions, which is crucial for forgery detection where image manipulations may occur at arbitrary resolutions. The ViT encoder operates through the following steps:

- **Input Data Transformation:** Unlike text data traditionally used with transformers, image data needs transformation into a compatible format. Here, the image is split into a series of flattened 2D patches, where each patch acts as a token analogous to words in natural language processing.
- **Patch Embedding:** Each patch is projected into a fixed-dimensional hidden vector space, denoted as D. This embedding process compresses each patch's information into a compact form, facilitating efficient processing and comparison across patches.
- **Special Class Embedding:** A learnable class embedding vector is introduced and appended to the set of patch embeddings. This vector serves as a global representation of the image, capturing high-level information that can be used for classification or further analysis.
- **Location Embeddings:** To preserve spatial information, location embeddings are added to each patch, ensuring that positional context is maintained throughout processing. These embeddings inform the model of each patch's position within the overall image.
- **Transformer Encoder Layers:** The main processing of the Visual Transformer Encoder is conducted through alternating layers of multi-head self-attention (MSA) and Multi-Layer Perceptron (MLP) blocks, both of which include layer normalization (LN) and residual connections. This structure allows the model to focus on relevant areas in the image independently and effectively

capture patterns and anomalies indicative of forgery.

Through this process, the ViT Encoder allows images to be processed at any resolution, avoiding resizing that could distort the locations of potential forgery artifacts.

3.1.2 Variational Autoencoder (VAE) Network

The Variational Autoencoder (VAE) Network further processes feature maps produced by the ViT Encoder. The VAE enables the model to retain critical details in the images, especially those that may have been manipulated. It consists of two primary parts:

- **Encoder:** The VAE encoder receives the feature maps and transforms them into a lower-dimensional latent space, capturing the input data's complexity and diversity within a condensed representation. During this process, the encoder calculates the mean and standard deviation of the input values and defines a Gaussian distribution from which latent vectors are sampled.
- **Decoder:** The decoder reconstructs the feature maps from the latent vectors. Using samples from the Gaussian distribution, the decoder attempts to produce output similar to the original input data. This reconstruction process allows the model to refine its understanding of the forgery artifacts and retain subtle details that may have been modified.

The VAE dynamically adjusts its input size based on the feature maps' dimensions from the ViT, allowing it to support images of varying sizes without requiring resizing, which is essential for accurate forgery detection.

3.1.3 MLP Decoder

The Multi-Layer Perceptron (MLP) Decoder is responsible for creating a binary mask that indicates regions of the image identified as manipulated. This lightweight decoder minimizes memory consumption while efficiently outputting a segmented mask in binary format. The MLP Decoder transforms the processed feature maps from the VAE into a final prediction with a dimension of $w \times h \times 1$, where w and h correspond to the width and height of the input image.

A linear activation function at the end of the MLP decoder produces an output image that matches the input dimensions. This process allows for a precise localization of the forged regions, enhancing the interpretability of the model's predictions.

3.1.4 Loss Calculation

The model's training objective is to minimize the Binary Cross Entropy (BCE) loss between the predicted mask and the ground truth mask. BCE loss is particularly suited for binary classification tasks like segmentation. The BCE loss L_{BCE} is defined as follows:

$$L_{\text{BCE}} = -\frac{1}{n}\sum_{i=1}^{n}\left(Y_i \cdot \log(\hat{Y}_i) + (1 - Y_i) \cdot \log(1 - \hat{Y}_i)\right), \tag{3.1}$$

where Y_i represents the ground truth pixel value, and $\hat{Y}_i$ represents the predicted pixel value by the model.

The total loss L for the training process is given by:

$$L = L_{\text{BCE}}(P, M), \tag{3.2}$$

where P and M denote the predicted mask and the ground truth mask, respectively. This loss function drives the model to produce accurate binary segmentation masks that align closely with the annotated ground truth.

3.1.5 Model Architecture Diagram

An overview of the ViT-VAE Net model architecture is illustrated in Figure 3.1. This diagram demonstrates the flow of data through the model's components: from the Visual Transformer Encoder, which divides the image into patches and generates feature embeddings, to the Variational Autoencoder, which further processes these embeddings to enhance detail retention, and finally to the MLP Decoder, which segments out forged regions with a binary mask.

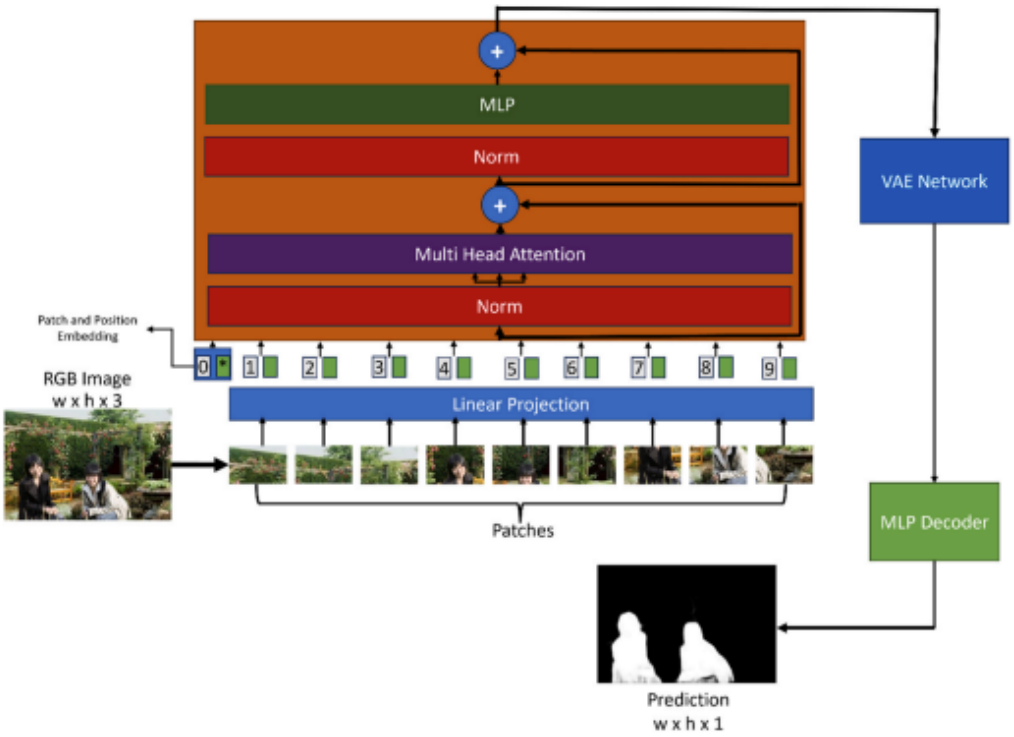

Figure 3.1: Overview of the ViT-VAE Net architecture, showing the Visual Transformer Encoder, Variational Autoencoder, and MLP Decoder for accurate localization of forged areas.

Bibliography

[1] M. Zanardelli, F. Guerrini, R. Leonardi, and N. Adami, "Image forgery detection: A survey of recent deep-learning approaches," *Multimedia Tools and Applications*, vol. 82, no. 11, pp. 17521–17566, 2023.

[2] M. T. I. Younis Abdalla and M. Shehata, "Copy-move forgery detection and localization using a generative adversarial network and convolutional neural-network," *Information*, vol. 10, no. 9, p. 286, 2019.

[3] M. M. A. J. Tahira Nazir, Marriam Nawaz, "Copymove forgery detection and segmentation using improvedmask region-based convolution network (rcnn)," *Applied Soft Computing*, vol. 131, p. 109778, 2022.

[4] R. Thakur and R. Rohilla, "Recent advances in digital image manipulation detection techniques: A brief review," *Forensic Science International*, vol. 312, p. 110311, 2020.

[5] W. D. Ferreira, C. B. Ferreira, G. da Cruz Júnior, and F. Soares, "A review of digital image forensics," *Computers & Electrical Engineering*, vol. 85, p. 106685, 2020.

[6] D. Cozzolino and L. Verdoliva, "Noiseprint: A cnn-based camera model fingerprint," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 144–159, 2020.

[7] M. F. Abdulqader, A. Y. Dawod, and A. Z. Ablahd, "Detection of tamper forgery image in security digital image," *Measurement: Sensors*, vol. 27, p. 100746, 2023.

[8] R. Agarwal and O. P. Verma, "An efficient copy move forgery detection using deep learning feature extraction and matching algorithm," *Multimedia Tools and Applications*, vol. 79, pp. 7355–7376, 2020.

[9] Priyanka, G. Singh, and K. Singh, "An improved block-based copy-move forgery detection technique," *Multimedia Tools and Applications*, vol. 79, pp. 13011–13035, 2020.

[10] Y. Aydın, "A new copy-move forgery detection method using liop," *Journal of Visual Communication and Image Representation*, vol. 89, p. 103661, 2022.

[11] D. B. A. T. Neeti Taneja, Vijendra Singh Bramhe, "Understanding digital image anti-forensics: an analytical review," *Multimedia Tools and Applications*, vol. 83, pp. 10445–10466, 2023.

[12] C. S. R. S B G Tilak Babu, "An optimized technique for copy–move forgery localization using statistical features," *ICT Express*, vol. 8, no. 2, pp. 244–249, 2022.

[13] A. H. Khalil, A. Z. Ghalwash, H. A.-G. Elsayed, G. I. Salama, and H. A. Ghalwash, "Enhancing digital image forgery detection using transfer learning," *IEEE Access*, vol. 11, pp. 91583–91594, 2023.

[14] J. Yang, Z. Liang, Y. Gan, and J. Zhong, "A novel copy-move forgery detection algorithm via two-stage filtering," *Digital Signal Processing*, vol. 113, p. 103032, 2021.

[15] F. Z. Mehrjardi, A. M. Latif, M. S. Zarchi, and R. Sheikhpour, "A survey on deep learning-based image forgery detection," *Pattern Recognition*, vol. 144, p. 109778, 2023.

[16] Rupesh D. Sushir and Dinkar Govindrao Wakde and Sarita S. Bhutada, "Enhanced blind image forgery detection using an accurate deep learning based hybrid DCCAE and ADFC," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 1725–1752, 2024.

[17] L. Verdoliva, "Media forensics and deepfakes: An overview," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 5, pp. 910–932, 2020.

[18] D. Kim, W. Ahn, and H.-K. Lee, "End-to-end anti-forensics network of single and double jpeg detection," *IEEE Access*, vol. 9, pp. 13390–13402, 2021.

[19] Z. Yu, J. Ni, Y. Lin, H. Deng, and B. Li, "Diffforensics: Leveraging diffusion prior to image forgery detection and localization," in *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12765–12774, IEEE Computer Society, 2024.

[20] C. Kumawat and V. Pankajakshan, "A robust jpeg compression detector for image forensics," *Signal Processing: Image Communication*, vol. 89, p. 116008, 2020.