IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 3, 2023

Detecting Image Forgery Through the Integration of Lightweight Deep Learning Models

Mrs. Ethakula Avyaktha[1]., P .Sahithi Reddy[2]

1 Assistant Professor, Department of CSE, Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

2, B.Tech CSE (20RG1A05G2),

Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

ABSTRACT

Various sectors, including judicial processes, social media platforms, and insurance fraud investigations, are increasingly using images and videos as convincing kinds of evidence in this digital era. There is reason to doubt the veracity of digital picture editing tools due to their inherent flexibility, especially when no obvious signs of manipulation are present. Authorities in the field of picture forensics are tasked with developing new technologies that can detect image fraud. So far, research has focused on three primary types of approaches for detecting modifications or forgeries: features descriptors, uneven shadows, and double JPEG compression. Many online information systems, social media, and real-time applications have image manipulation detection as a major challenge. Conventional detection approaches have their limitations because to long-held assumptions about things like hand-crafted characteristics, size, and contrast, which are used to identify indications of photo alterations. This study introduces a fusion-based judgment procedure for the detection of picture fraud. A trio of lightweight deep learning models—SqueezeNet, MobileNetV2, and ShuffleNet—form the basis of the decision fusion. A biphasic approach is used to perform the fusion decision system. To begin determining whether the images are real, we use the current weights of the effective deep learning models. In addition, the results of image forgeries are compared to the pre-existing models using the improved weights. The fusion-based decision strategy achieves better accuracy than the state-of-the-art methods according to the experimental data.

Image fusion, support vector machines, detection, deep learning, lightweight models, and light fusion are all related terms.

1. INTRODUCTION

In today's digital age, visual evidence such as photos and videos is becoming important in many areas, such as trial testimony, insurance fraud, social media, etc. Concerns over their veracity arise due to the versatility of digital image alteration technologies, particularly in the absence of obvious signs of tampering. Experts in the field of picture forensics are tasked with creating new technologies that can identify photo frauds. Feature descriptors, inconsistent shadows, and, finally, double JPEG compression are the three main categories of manipulation or forgery detectors that have been studied so far.

It is simple to manipulate the image's contents to influence others' views using advanced software. Splicing and copy-move are the two main types of image faking methods. A copy-move forgery involves tracing and smearing sections of the image's content region within another image, while a splicing forgery involves splicing parts of the image's content from other photographs. In recent years, several methods for detecting image forgeries have been suggested, with the goal of restoring confidence in visual content. Past research has attempted to detect fabricated regions by extracting several picture attributes, such as lighting, shadows, sensor element noise, and camera reflections, in an effort to detect copy-paste or splicing.

Scientists established the image's veracity wherever it was identified as genuine or fake. The artifacts left behind by various JPEG compression and other picture editing techniques may be used by several current methods to detect faked areas. We have also looked at camera-based methods where detection

is based on demosaicing regularity or sensing element pattern noise, where anomalies are found by extracting and comparing sensing element pattern abnormalities. False or doctored images have the potential to cause harm and even death. In place of manually extracting features or engineering them, this work seeks to automate the feature extraction procedure in order to detect the altered images. Using deep learning, we may take into account clustered native connections by making use of nearby pixels that are strongly connected.

Reasons to employ lightweight models include their ability to learn richer representations, ease of deployment on hardware with limited resources, and prevention of convolutional neural network (CNN) overfitting. When it comes to tiny network efficiency, ShuffleNet is crucial since it produces more feature map channels for a given computation complexity budget, allowing for more information to be encoded. Using deep-separable convolutions, MobileNet achieves state-of-the-art results and proved its efficacy across a variety of workloads. With fewer parameters than AlexNet while maintaining standard accuracy, SqueezeNe optimizes the architecture for a CNN system with high processing speed using 50×. The models are small and efficient, allowing them to learn richer representations and run well on hardware with limited resources.

2. LITERATURE SURVEY

By investigating the methods for revealing and localizing a single or double JPEG compression using convolutional neural networks (CNNs), Amerini et al. presented a step forward in this area. Various experiments have been conducted to attempt to identify possible concerns that need more investigation, and other types of input to the CNN have also been considered.

Xiao et al. presented a two-pronged approach to splicing forgery detection using diluted adaptive clustering and a coarse-to-refined convolutional neural network (C2RNet). By combining a coarse convolutional neural network (C-CNN) with a refined convolutional neural network (R-CNN), the suggested C2RNet is able to recover the scale-dependent variations in picture attributes between untouched and altered areas. In addition, C2RNet now uses an image-level CNN instead of a patch-level CNN, which significantly reduces the computational complexity. In order to provide consistent detection performance, the suggested technique learns the differences of different picture attributes, and the computational time required by the image-level CNN is significantly reduced.

Using a Stacked Autoencoder model to learn the complex feature for each individual patch, Zhang et al. analyzed the first step. The second step of this paper's detection process involves properly integrating the contextual information of each patch.

In their quantitative investigation on picture tampering, Goh et al. suggested a hybrid evolutionary framework to assess all aspects and determine the optimal collection of attributes. The classification method has to be fine-tuned for optimal performance after feature assessment and selection. Thus, not only can the hybrid framework find the best features for a classifier, but it can also find the best multiple classifier ensembles, all while attaining the greatest classification performance for picture tampering detection in terms of low complexity and high accuracy.

By applying the Markovian rake transform on the luminance component of a picture, the statistical characteristics provided by Sutthiwan et al. may be created. If you take 2-dimensional arrays with different block sizes and apply the quantized block discrete cosine transform to them, you'll get difference arrays that can be transformed using the Markovian rake transform. After addressing several pertinent concerns and making related adjustments, the effectiveness of the features that were developed have been validated on a newly created large-scale picture dataset that was intended for tamper detection. The first results from applying the classifiers that were developed in this way to certain real-life manipulated photographs that are accessible online demonstrate both the potential of the suggested characteristics and the difficulty that researchers have in detecting instances of image manipulation.

In order to identify this particular artifact, he and colleagues proposed a Markov-based method. In order to capture both the intra-block and inter-block correlation between block DCT coefficients, the

initial Markov features developed by Shi et al. from the transition probability matrices in the DCT domain are enlarged. To further define the three types of interdependence among wavelet coefficients across locations, scales, and orientations, additional features are built in the DWT domain. Then, to make the computational cost more reasonable, the feature selection technique SVM-RFE is used to accomplish the feature reduction job. Last but not least, the legitimate and spliced pictures are classified using the final dimensionality-reduced feature vector by using support vector machine (SVM).

Among the many successful methods for image modification, Change et al. presented a new forgery detection system for identifying altered inpainting pictures. There are two main steps in the suggested algorithm: detecting suspicious regions and identifying forged regions. To identify potentially malicious areas in a picture, suspicious region identification employs a similarity vector field to weed out uniform area-based false positives and examines the image's similarity blocks for relevant regions. To distinguish between legitimate and fraudulent areas, forged region detection uses a novel approach called multi-region relation (MRR). Even for photographs with a homogeneous backdrop, the suggested method can successfully detect whether the image is fabricated and locate the fabricated areas. In addition, to enhance calculation performance, this research suggested a weight-transformation-based two-stage searching method.

The brief feature vector introduced by Rhee et al. consists of three distinct groups of features. According to the variation, the first set is the three-dimensional length of the gradient difference between the intensity values of the neighboring row and column line pairs in the picture, and the second set is the same. The fluctuation in the coefficient difference of the Fourier transform, which is the three-dimensional length in the neighboring line pairs, defines the second set. The third set is defined by the three-dimensional length, which is also the residual image between the original and reconstructed images, as determined by the gradient obtained from solving Poisson's equation. Three sets are obtained: two from the picture's spatial domain and one from its spectral domain. The last set is derived from the residual image. Once the 9-dimensional feature vector is complete, the support vector machine classifier is used to train MFD.

A method for detecting picture duplication was suggested by Lamba et al., and it relies on discrete fractional wavelet transforms. Overlapping picture blocks with predetermined size are used to divide the test image. In order to extract features from each picture block, a discrete fractional wavelet transform is used. Following the lexicographical systematization of all feature vectors, the repeated blocks, if any, are obtained by block matching and block filtering. Both single and numerous duplicated areas may be effectively detected using the suggested technique.

To find altered photos, Lin et al. suggested looking for the discrete cosine transform (DCT) coefficients, which conceal a twofold quantization effect. Among its many benefits, this paper is unique in its ability to automatically detect tampered regions; it also has fine-grained detection at the scale of DCT blocks, is fast, and is insensitive to various forgery methods (including simple image cut/paste, inpainting, and alpha matting). Additionally, it can work without completely decompressing the JPEG images. Promising experimental outcomes have been seen using JPEG pictures.

3. PROPOSED SYSTEM

The suggested decision fusion architecture relies on the lightweight deep learning models shown in Figure 1. The selected deep learning models are SqueezeNet, MobileNetV2, and ShuffleNet, all of which are lightweight. There are two stages to implementing the suggested system: first, using pre-trained deep learning models; and second, refining those models. When using the pre-trained model, regularization is omitted and the pre-trained weights are used. However, when applying regularization to the fine-tuned model, picture counterfeiting is detected. Three steps make up each stage: pre-processing the data, classification, and fusion. The query picture is pre-processed at the data pre-processing step according to the dimensions needed by the deep learning models. To determine

International Journal of Information Technology & Computer Engineering
ISSN 2347-3657
Volume 11, Issue 3, 2023

whether a picture is fabricated or not, support vector machines (SVMs) are used. We start with a brief overview of lightweight deep learning models, and then go on to detail the regularization technique.

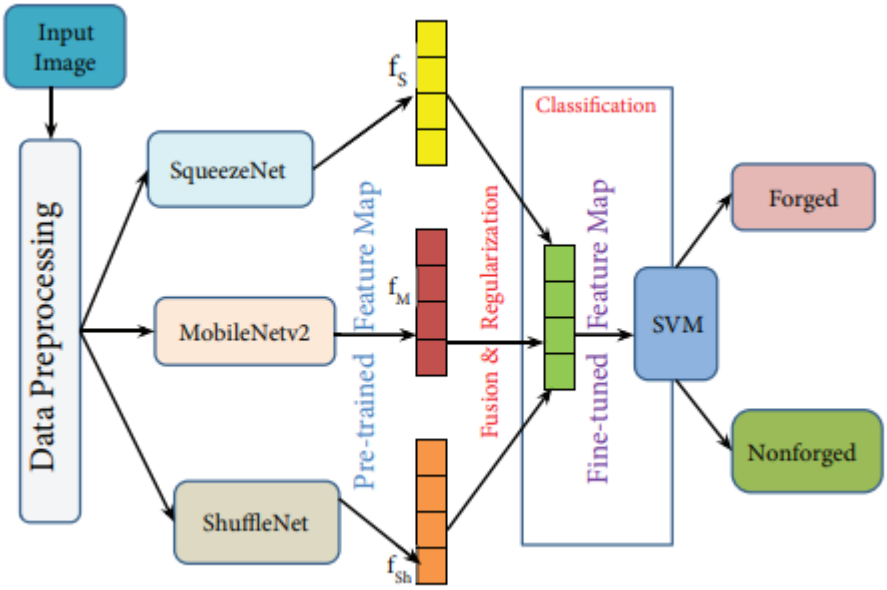

Figure 1: Fusion based decision model for forgery detection.

Data preprocessing

The preprocessing phase involves determining if a picture in a query is fabricated or not. For SqueezeNet to work, the picture dimensions must be 227×227. The picture dimensions needed for MobileNetV2 are 224×224. For ShuffleNet to work, the picture dimensions must be 224×224. Prior to processing the input picture, it is pre-processed according to the dimensions needed by each model. After that, each model uses the input picture to generate a feature vector.

ShuffleNet

This CNN can categorize photos into up to a thousand different categories; it was trained on the ImageNet dataset and has 50 layers of depth. Lightweight deep learning model parameters (Table 1). Here, "parameter" denotes the overall number of learnable parameters in each layer, "image input size" denotes the necessary input picture size, and "depth" is the maximum number of consecutive convolutional or fully connected layers along a route from the input layer to the output layer.

Table 1. Models description.

Models	Depth	Parameter (millions)	Image input size
SqueezeNet	18	1.24	227 × 227
MobileNetV2	53	3.5	224 × 224
ShuffleNet	50	1.4	224 × 224

International Journal of Information Technology & Computer Engineering
ISSN 2347-3657
Volume 11, Issue 3, 2023

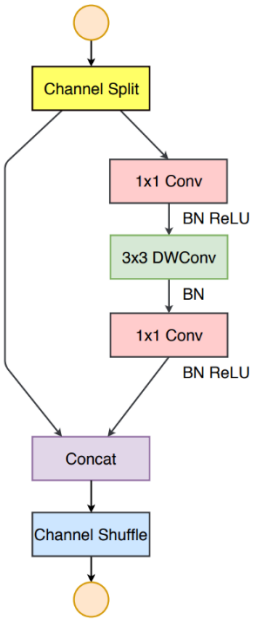

Figure 2: ShuffleNet.

Fusion model and regularization

Using pretrained weights for picture forgery detection, the proposed system is first constructed using lightweight deep learning models. Subsequently, the system is implemented as a combination of the lightweight models' decisions, as covered in the preceding section. To get the feature maps of the lightweight models, the input picture is first delivered to them. The feature maps obtained by SqueezeNet, MobileNetV2, and ShuffleNet are identified by the notations f_s, f_m, and f_sh, respectively. The output feature mapping f_p from the pretrained lightweight deep learning model is utilized in the fusion model. Equation (1) illustrates how the feature maps derived from the lightweight models are combined to create this feature map, f_p.

$$f_p = f_s + f_m + f_{sh} \quad (1)$$

In order to extract the image's features, the fusion model employs the feature map f_p as a local descriptor for an input patch. The function Y_fusion 〖=f〗 _((x)) represents the image for the fusion model, where x is the input picture patch. A sliding window of size p×p is used to calculate the local descriptor for a test picture of size m×n. Equation (2) illustrates how Y_fusion is calculated, with Y_1, Y_2, and Y_3 standing for the descriptors of the picture patches that the deep learning models have produced. It is obtained by concatenating all of the input patches (xi). The new image representation, f_fusion, is then used as the feature map for the SVM to classify as forged or nonforged. Equation (3) gives the new image representation, where s is the size of the stride used to transform the input patch.

$$Y_{fusion} = [Y_1 + Y_2 + \ldots + Y_T] \quad (2)$$

$$f_{fusion} = \frac{m-w}{s} + 1 * \frac{n-w}{s} + 1 \quad (3)$$

Equation (4) illustrates how the weight kernels are initialized in order to fine-tune the fusion model's parameters. The weights of the fusion model are represented by Wf in this equation, the SqueezeNet model by W_s, the MobileNetV2 model by W_m, and the ShuffleNet model by W_sh and W_sh. Equation (5) illustrates how the weight of the fusion model W_f is initialized. The weights' initialization serves as a regularization term, which makes it easier for the fusion model to learn the reliable characteristics of forgery detection rather than the intricate picture representations.

$$W_f = [W_{sj}\ W_{mj}\ W_{shj}]\ j = 1, 2, 3 \quad (4)$$

International Journal of Information Technology & Computer Engineering
ISSN 2347–3657
Volume 11, Issue 3, 2023

$$W_f = [W_f{}^{4k-2}\, W_m{}^{4k-2}\, W_{sh}{}^{4k}] \text{ where } k = [[j + 1] mod 11] + 1 \quad (5)$$

4. RESULTS AND DISCUSSION

Dataset

The experiment used the publicly available benchmark MICC-F220 dataset, which consists of 110 non-forged and 110 forged pictures with three channels. These images are color and range in size from 722 × 480 to 800 × 600 pixels. As can be seen in Figure 3, the non-forged picture is Figure 3k, while Figures 3a–3j are forged images using ten distinct combinations of geometrical and transformational assaults. 154 randomly selected photos are used for training from the dataset, with the remaining images being used for testing.

Figure 3: Dataset with 10 different combinations of geometrical and transformation attacks; (a–j), forged; (k), non-forged images.

Baseline modules

The following is a summary of the baseline models that are used to compare the fusion model.

1) Upload MICC-F220 Dataset: We will upload the dataset to the program using this module.
2) Pre-process Dataset: Using this module, all photos will be read, their pixel values will be normalized, and they will be resized to the same size.
3) The third module, Generate & Load Fusion Model, is used to train three algorithms: SqueezeNet, MobileNetV2, and ShuffleNet. Features are then extracted and used to train the fusion model. Test data will be used to determine the prediction accuracy of all algorithms.
4) The fourth module, Fine Tuned Features Map with SVM, is used to create a fusion model by extracting features from all three procedures. The fusion data is then trained using SVM, and the prediction accuracy is determined.

International Journal of Information Technology & Computer Engineering
ISSN 2347-3657
Volume 11, Issue 3, 2023

5) Run the Baseline SIFT Model: This module will be used to extract features from pictures using the SIFT approach, which will then be trained using SVM to determine its prediction accuracy.
6) Accuracy Comparison Graph: We will draw an accuracy graph for each method using this module.
7) Performance Table: We will show the performance tables for all methods using this module.

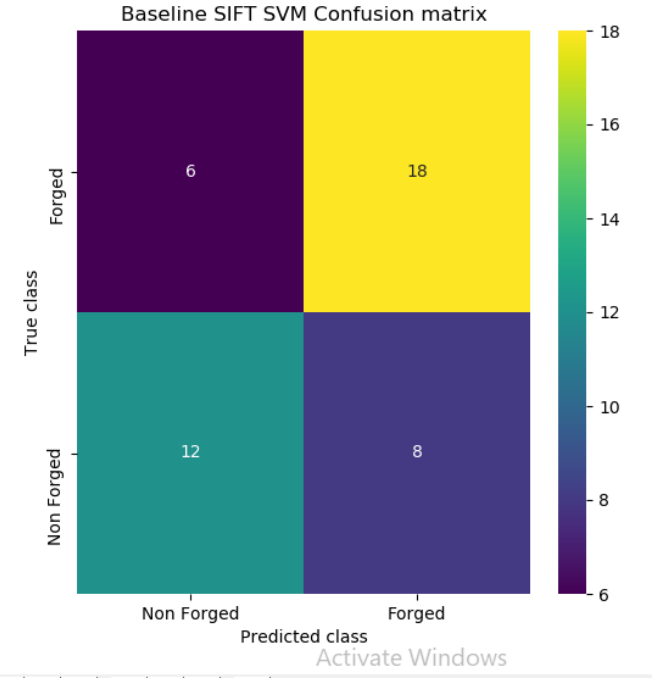

Figure 4: Confusion matrixes of fusion model and baseline SIFT SVM.

Table 2: Performance comparision.

Method	Accuracy	Precision	Recall	FSCORE
Existing SIFT SVM	68.1	67.9	67.5	67.5
Only SqueezeNet	79.5	81.1	79.5	79.2
Only ShuffleNet	56.8	62.7	56.8	51.1
Only MobileNetV2	81.8	82.9	81.8	81.6
Proposed Fusion Model SVM	95.4	95	96.1	95.3

5. CONCLUSION

Identifying altered or false pictures is what image forgery detection is all about. Image forgery detection in this study is achieved by the decision fusion of lightweight models based on deep learning. The original plan was to decide whether or not a picture was fake by combining three lightweight deep learning models: SqueezeNet, MobileNetV2, and ShuffleNet. The pretrained models' weights are regularized in order to get a forging conclusion. The results of the trials show that compared to the state-of-the-art methods, the fusion-based strategy provides better accuracy. Future work on picture fraud detection might make use of other weight initialization procedures to refine the fusion choice.

REFERENCES

[1] Amerini, T. Uricchio, L. Ballan, and R. Caldelli, "Localization of JPEG Double Compression Through Multi-Domain Convolutional Neural Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1865-1871, doi: 10.1109/CVPRW.2017.233.

[2] B Xiao, Y Wei, X Bi, W Li, J Ma. “Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering”, Information Sciences, Volume 511, Pages 172-191, 2020, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2019.09.038.

[3] Zhang Y, Goh J, Win LL, Thing VL. Image region forgery detection: a deep learning approach. SG-CRC 2016; 2016: 1-11.

[4] Goh J, Thing VL. A hybrid evolutionary algorithm for feature and ensemble selection in image tampering detection. International Journal of Electronic Security and Digital Forensics 2015; 7 (1): 76-104

[5] Sutthiwan P, Shi YQ, Zhao H, Ng TT, Su W. Markovian rake transform for digital image tampering detection. In: Shi YQ, Emmanuel S, Kankanhalli MS, Chang S-F, Radhakrishnan R (editors). Transactions on Data Hiding and Multimedia Security VI. Lecture Notes in Computer Science, Vol. 6730. Berlin, Germany: Springer; 2011, pp. 1-17

[6] He Z, Lu W, Sun W, Huang J. Digital image splicing detection based on Markov features in DCT and DWT domain. Pattern Recognition 2012; 45 (12): 4292-4299.

[7] Chang IC, Yu JC, Chang CC. A forgery detection algorithm for exemplar-based inpainting images using multi-region relation. Image and Vision Computing 2013; 31 (1): 57-71.

[8] Rhee KH. Median filtering detection based on variations and residuals in image forensics. Turkish Journal of Electrical Engineering & Computer Science 2017; 25 (5): 3811-3826.

[9] Lamba AK, Jindal N, Sharma S. Digital image copy-move forgery detection based on discrete fractional wavelet transform. Turkish Journal of Electrical Engineering & Computer Science 2018; 26 (3): 1261-1277.

[10] Lin Z, He J, Tang X, Tang CK. Fast, automatic, and fine-grained tampered JPEG image detection via DCT coefficient analysis. Pattern Recognition 2009; 42 (11): 2492-2501.