Received: April 3, 2026. Revised: June 21, 2026.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

A Feature Fusion CNN Approach for Fake Educational Credential Detection and Forgery Region Localization

Sarala M[1]* **Muralidhara B. L[1]** **Suresh R[2]** **Rashmi Siddalingappa[3]**

[1]*Department of Computer Science and Applications, Bangalore University, Bengaluru, India*
[2]*Department of Statistics, Bangalore University, Bengaluru, India*
[3]*Department of Computer and Data Science, York St John University, London, United Kingdom*
* Corresponding author's Email: sarala5111@gmail.com

Abstract: Educational institutions are implementing tamper-proof techniques to prevent fraudulent degrees. In today's digital world, the creation of fake educational credentials has become prevalent, facilitated by document editing tools, which are often used to secure admissions and job placements. This study focused on identifying fake educational credentials across five categories: Handwritten, Computerized, Choice-Based Credit System (CBCS), Degree Certificate, and Non-CBCS using a machine learning (ML) approach. The entropy features from the preprocessed educational credentials dataset were extracted using Shannon entropy. We have chosen one hundred samples from the training data of each category to choose the best distance value using the Knee point method. The test dataset includes seven hundred test samples, which contain one hundred held-out normal samples, and forty fake samples as Out of Sample (OOS) in each category. It was clustered into either 'genuine' or 'fake' using Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method, providing an accuracy of 99.14%. Further, a customized feature fusion Convolutional Neural Network (CNN) model was designed with an attention mechanism and an Atrous Spatial Pyramid Pooling (ASPP) layer to locate fake regions in fake credentials. This proposed model shows significantly better results in locating fake regions of both image and text-based regions.

Keywords: Shannon-entropy, DBSCAN, ASPP, CNN, Attention, Locate-region.

1. Introduction

Identifying anomalies and forgeries is an important task across various sectors, including banking, industry, healthcare, networking, and departments handling documents and credentials. We focused towards forged activities in the education sector, in that recent news reports have highlighted the creation of fake mark cards and degree certificates [1-5]. Educational institutions have begun employing tamper-proof techniques such as barcodes, QR codes, holograms, and watermarking to avoid the creation of fake mark cards. Even though fraudulent degrees persist because of advancements in document editing software. As a result, institutions and companies demand the genuineness of certificates to offer admission to higher studies and employment opportunities. It leads to authenticating the multiple categories of educational credentials of the academic institutions, which range from initial to recent ones. Additionally, the authentication process requires human resources and time to verify the credentials.

In image forgery detection (IMD), copy-move forgery (CMF) is detected using wavelet-based and edge-based feature extraction methods; splicing forgery is detected using colour filter array (CFA) and noise detection methods. However, these traditional methods have issues in accurately detecting forged regions. Hence, DL techniques are being used to overcome the drawbacks of traditional methods. The pre-trained CNN models such as VGG16, VGG19, DenseNet, ResNet50, U-Net, ResNet101, HR-Net, and FCN have been employed to detect fake images in datasets like CASIA, CASIA v2, CoMOFoD, USCISI, Columbia, NIST16, and COVERAGE [6].

A comparative study was conducted using traditional and DL methods for IMD, and different techniques based on pixels, image formats, physical environment, and geometry were explored for detecting passive image forgery [7].

To overcome the issue of manual authentication process, we detected fake credentials using ML techniques, and located the fake regions across multiple categories of credentials using deep learning (DL) techniques. The main contribution of this study is to detect fake educational credentials across five categories, and locate the fake regions of both image, and text-based regions in such credentials using a feature fusion CNN model. Further, we proved the model's prediction using the statistical method t-test. In the remaining sections, literature studies explore fake detection using traditional and DL-based techniques in Section 2. Section 3 describes the methodology for detecting fake educational credentials, and locating fake regions in fake credentials. The performance of detecting fake educational credentials, and locating fake regions in fake credentials is analyzed in Section 4. Finally, findings of the empirical study and future work are discussed in Section 5.

2. Literature review

Anomaly detection algorithms are selected depending on the nature of the datasets. These algorithms utilize machine learning (ML) techniques for structured datasets to identify anomalies, while deep learning (DL) methods are applied to high-dimensional datasets such as images and videos. To locate the fake region, traditional methods extract the features based on wavelet analysis and edges. In contrast, the deep learning technique uses the feature fusion CNN model integrated with attention mechanisms and an ASSP layer. Moreover, Explainable Artificial Intelligence (XAI) methods, such as Gradient-weighted Class Activation mapping (GradCAM), GradCAM++, and Local Interpretable Model-agnostic Explanations (LIME), are utilized to locate high-intensity regions.

To understand the CMF, we referred the work [8] to gain insights on forgery detection with methods used, limitations, and solutions for CMF. Further, we referred the work [9] to detect image forgeries using ML and deep learning methods along with merits and limitations. The forged images were detected using different colour channels such as grayscale, RGB, and HSV, and ML algorithms MLP, SVM, and KNN [10]. An extended U-Net model was developed to locate the fake regions across four types of sources: medical images, natural images, scanned documents, and identity documents [11]. A hybrid method was proposed by integrating traditional methods such as block-based method, and keypoint-based method; and deep learning methods to locate fake regions [12].

Document forgery detection is identified by various methods such as edge detection methods, OCR extraction, printer model, ink type, ML, and deep learning methods. Ranveer et al. reviewed the document forgery detection with ML and deep learning methods, limitations, and merits [13]. The DCLNet was proposed to detect document forgery with various desensitization and tampering operations [14]. The textual features were extracted from scanned passport images using OCR, and LBP method; image regions were extracted using ORB method; the CNN model was used to authenticate the document [15]. Zengqi et al. evaluated fourteen methods for document forgery detection on eight datasets [16]. The Edge concatenation (EC) layer extracts the features from document image using Sobel filter, and the edge attention layer (EA) extracts the edge features for enhancing the boundary information. The EA-EC method improved the performance by identifying manipulations along with text boundaries [17]. The fake regions in seven types of certificate images were located using attention mechanism [18]. This work proved that fake regions in certificates can be located using a feature fusion CNN model.

The feature fusion CNN models are designed with an attention mechanism for locating fake regions in the image dataset. A CNN model was incorporated with the attention mechanism to detect geographical location [19], highlight the features in CIFAR-10 and STL-10 datasets [20], and find the machinery faults [21]. A multi-scale model with attention mechanism was proposed to locate fake regions in CoMoFoD, CASIA v2.0, and MICC-F220 [22]. The fusion model with spatial attention mechanism was proposed to detect fake face images [23]. The CNN model was designed with attention mechanism to detect and locate image forgery, with a probability map generated using an ASPP layer [24]. An encoder-decoder model with ASSP was proposed with a fusion approach to find the fake region in infrared images [25], to locate fake regions in KITTI-360 2D, and Cityscapes datasets [26].

The following pre-trained CNN models are used to locate fake regions in the secondary dataset: The PRRNet was designed to detect forgery regions in face image datasets of celeb-DF, Deepfake, and faceforensics++. In this model, the pixel region module separates the original and manipulated regions [27].

Mantra-Net handled images with arbitrary sizes, and forgery types of CMF, splicing, and enhancement. It has limitations in detecting regenerated forged images, noisy images, and images manipulated with different types [28]. ObjectFormer focuses on detecting and locating image manipulation, trained on the synthesized datasets including FakeParis, FakeCOCO, and Pristine image, and tested using CASIA, Columbia, NIST Nimble, Carvalho, and IMD20 datasets, and evaluated with PSCCNet, MantraNet, and SPAN [29]. MVSS-Net was proposed to detect image manipulation, proving effective for images that are compressed, blurred, or recaptured screenshots [30]. EMT-Net identified the fake images across six datasets, including CASIA, Columbia, NIST, COVER, DEFACTO, and CoMoFoD. It extracted the features using transformers, and located the regions using CNN. However, this model has limitations in detecting unlearned manipulated images [31].

The traditional methods are highly focused on copy-move and splicing forgery. The CNN models were trained to identify CMF, and employed XAI methods to locate the fake regions. To overcome the limitations of these conventional methods, researchers are training a feature fusion CNN model with an attention mechanism to locate the regions. Some pre-trained CNN models [30-34] are available to locate fake regions in the secondary dataset, but those models would not predict fake regions in other datasets, as they were trained on an attention mechanism. Due to the limitation in certificate image forgery detection, we considered the related works for comparing different studies, and it is represented in Table 1.

Table 1. Comparative study of document forgery detection

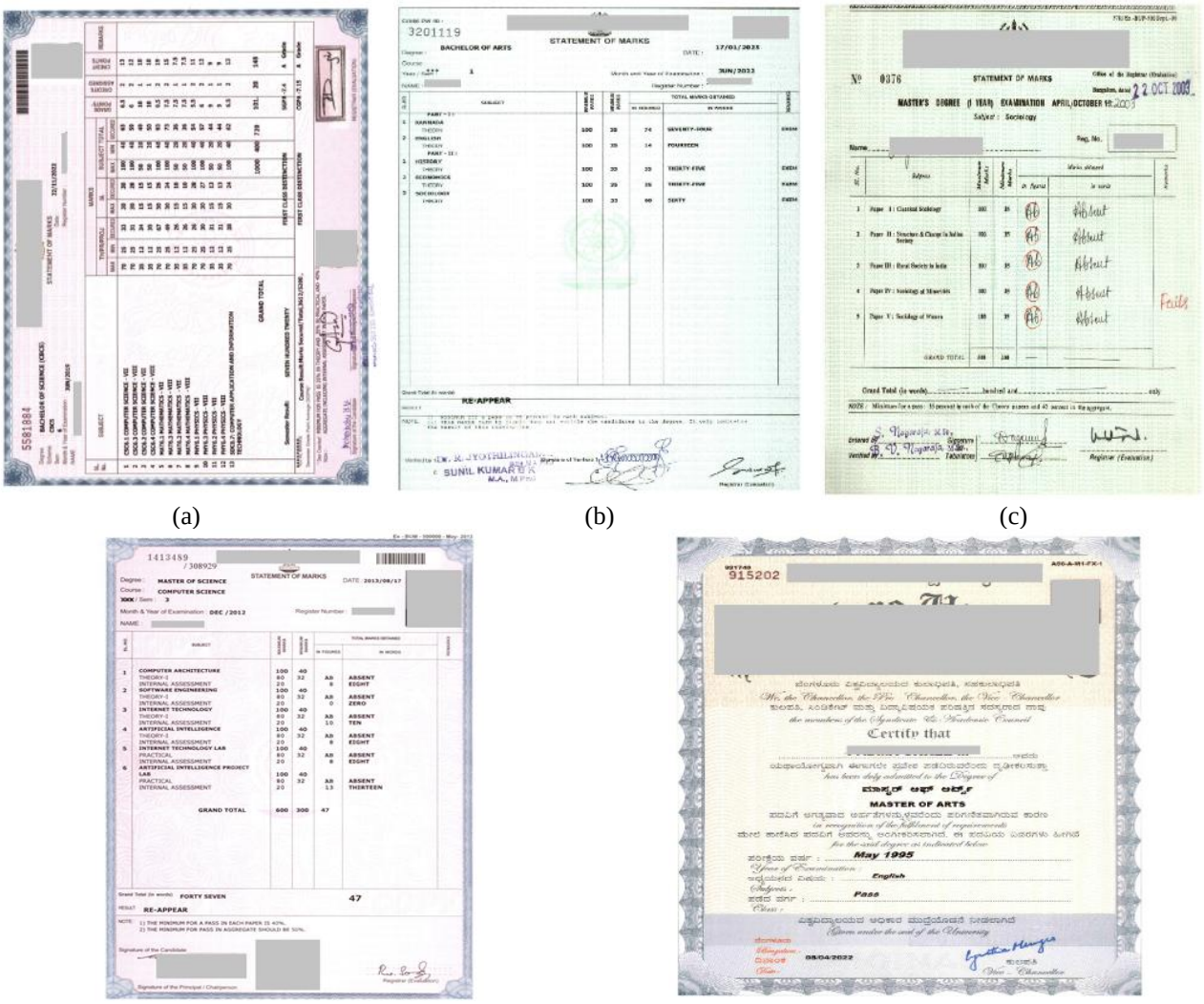
Sl. No.	Reference No.	Methods used	Dataset	Key findings	Limitations/ Future work
1.	[14]	DCLNet: Convolutional encoder-decoder attention-based network	DocTD (primary dataset)	Differentiated the desensitization from tampering operations.	Does not identify text characteristics.
2.	[15]	• OCR, and LBP to extract the text features. • ORB to extract the images. • CNN model to authenticate.	MIDV-500 (passport images)	Focused separately on textual, image, and hologram features to improve the performance.	Computationally expensive.
3.	[16]	Fourteen pre-trained models	T-SROIE, Real-TextManipulation, Tampered-IC13, Receipt Forgery, MixTamper, FSTS-1.5k, FantasyID, catalogued	Evaluated using fourteen methods, and eight datasets.	Does not identify LLM-based document forgeries
4.	[17]	Proposed edge-focused CNN model.	DocTamper, Receipt,MIDV-202	Conducted comprehensive experiments for validating the results.	• Computationally expensive for real-time systems. • Not evaluated on all types of forgery detection methods.
5.	[18]	Multi-level features attention	TIANCHI (certificates)	• Detected fake regions on multiple types of certificates. • Evaluated on different datasets.	• IoU score is 0.6360. Does not focus on hand-written template.
6.	Ours	Feature fusion and Attention mechanism	Primary dataset (certificates)	• Detected fake regions on multiple types of certificates including hand-written.	Does not focus on other university datasets due to confidential data.

Drawing from the literature review, we designed a feature fusion CNN model with an attention mechanism to locate fake regions in the primary dataset.

3. Materials and methods

3.1 Materials

The Bangalore University's original mark cards have four categories, namely CBCS, Computerized, Hand-written, Non-CBCS, and degree certificates were scanned using the FUJITSU scanner with 300 dpi (dots per inch) in Joint Photographic Experts Group (JPEG) file format. It includes six hundred data samples in each category shown in Fig. 1. The images were resized to 256 ^{x} 256, and normalized by dividing the pixel intensities by 255 to reduce the computational time. The training, validation, and testing data samples from each category include four hundred, one hundred, and one hundred respectively in the ratio of 4:1:1. Additionally, fabricated educational credentials were created by template duplication using Adobe Photoshop to test proposed model. These fabricated education credential data samples are called out-of-sample (OOS). The test dataset consists of 140 samples per category, comprising one hundred standard held-out unseen normal samples, and forty abnormal OOS data samples. These forty abnormal samples are divided into two equal groups: 20 samples are kept in their original, unmodified form, while the remaining twenty are synthetic variations created by altering combinations of the template's font size, logo size, border colour, border thickness, background colour, and texture.



(a) (b) (c)

(d) (e)

Figure. 1 Educational Credentials: (a) CBCS, (b) Computerized, (c) Handwritten, (d) Non-CBCS, and (e) Degree certificate

3.2 Methodology

The empirical study conducted for detecting fake educational credentials, and locating fake regions in such credentials is illustrated in Fig. 2. The Shannon entropy method was employed to extract the entropy feature of preprocessed image datasets, and is stored in the database. These extracted features were clustered as 'genuine' or 'fake' using the DBSCAN clustering method, and cluster labels were updated in the database. The customized feature fusion CNN model with ASPP layer was used to locate fake regions in fake educational credentials.

3.2.1. Detection of fake educational credentials

In the detection of fake educational credentials, the Shannon entropy (SE) method was applied to extract the entropy feature from both genuine and fake credentials as represented in Eq. (1). These extracted features were stored in the database and then clustered either as 'genuine' or 'fake' using the DBSCAN method, which automatically clusters based on the maximum distance between two samples ($\boldsymbol{\varepsilon}$) in a cluster. The entire procedure to cluster the sample is defined in Algorithm-1. To determine the optimal $\boldsymbol{\varepsilon}$ and number of neighbours $\boldsymbol{k}$ value, we randomly selected one hundred normal samples from each category. The ε value was chosen as 0.13 using the normalized Knee point method shown in Fig. 3. In this method, we have chosen the number of neighbours (k) as ten to select the optimal $\boldsymbol{\varepsilon}$ value based on the empirical study. The clustering of all categories is shown in Fig. 4. The predicted cluster labels were updated in the database to analyze the performance of the detection of fake educational credentials, and it is discussed in Section 4.1. We explored the detection of fake educational credentials by locating the fake regions using a feature fusion CNN model in the next Section.

$$\mathrm{SE} = -\sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} \quad (1)$$

where, n represents the total number of pixels, i.e, (256 × 256 × 3), and pi represents the pixel at the ith position.

Algorithm 1: Clustering of data samples using the DBSCAN method

Input: i: number of categories, ε: maximum distance between two samples, k: number of neighbours, SE: Shannon Entropy, DB: database Output: return cluster labels of test_data

Step 1: Repeat i for all the categories
Step 2: Repeat until data samples exist in the $normal\_dataset_i$
$img \leftarrow$ read the image from the $normal\_dataset_i$
$img \leftarrow \frac{img}{255}$
$SE\_fea \leftarrow SE([img])$
Store the SE_fea in DB
Step 3: End repeat
Step 4: End repeat i
// Selection of optimal k, and ε from SE_fea stored in DB to cluster
Step 5: Repeat until no SE_fea cluster as outlier
selection of k
selection of ε using the normalized knee point method
Step 6: End repeat
Step 7: Repeat steps from 2 to 7 for test_dataset
Step 8: Cluster the features using optimal k, and ε
Step 9: Update the cluster_labels in the DB_2

3.2.2. Locating fake region

The methodology for detecting fake educational credentials is discussed in section 3.2.1. This section outlines locating the fake regions in fake credentials using a feature fusion CNN model incorporating an ASPP layer. The layer diagram of this proposed model is shown in Fig. 5. The five unique CNN models were trained using an RGB image dataset, with each category consisting of images sized 256 × 256. These models were trained with the following parameters, which were chosen based on empirical studies: epochs = 70, learning rate = 0.001, batch size =40, and optimizer ='adam', Loss = Mean Square Error(MSE) is mentioned in Eq.2. The model will return the predicted value as mentioned in Eq.3. In proposed feature fusion CNN model, the first four convolutional layers ($\mathbf{Conv_{(m,f)}(fea)}$) have 64, 32, 32 and 16 filter sizes. In these layers, convolution operation was performed between feature map and weight matrix as represented in Eq.4. The 'ReLU' activation was applied in all Conv2D layers as mentioned in Eq.5. The average pooling with the size of 2 × 2 was followed after the first two Conv2D layers as calculated in Eq.6, and is illustrated in Fig. 6. The ASPP layer was applied on the feature map extracted from the fourth layer with convolutional dilation rates 2, 4, 6, 8 and 10 as defined in Eq. (7). The features from the first three layers, and the ASPP layer were resized to 64 × 64, to ensure uniformity for concatenation as mentioned in Eq. (8).

Received: April 3, 2026. Revised: June 21, 2026.

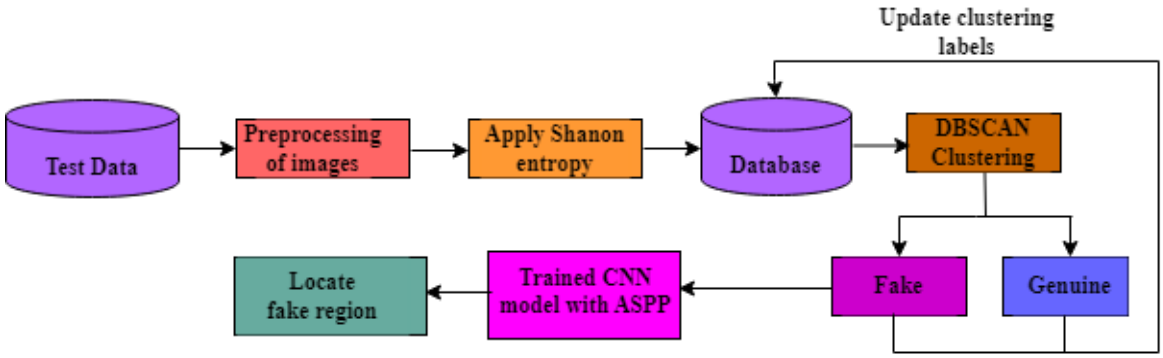


Figure. 2 Proposed methodology

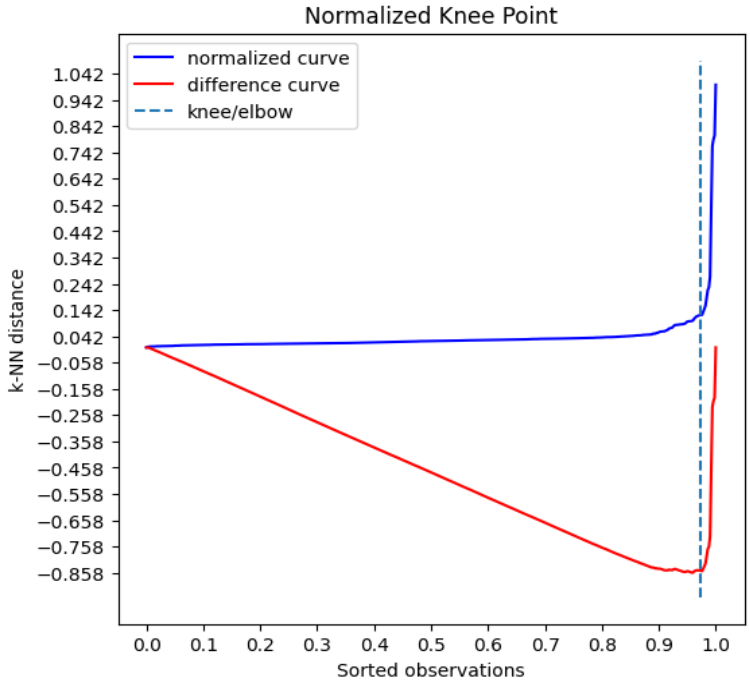


Figure. 3 Knee point method

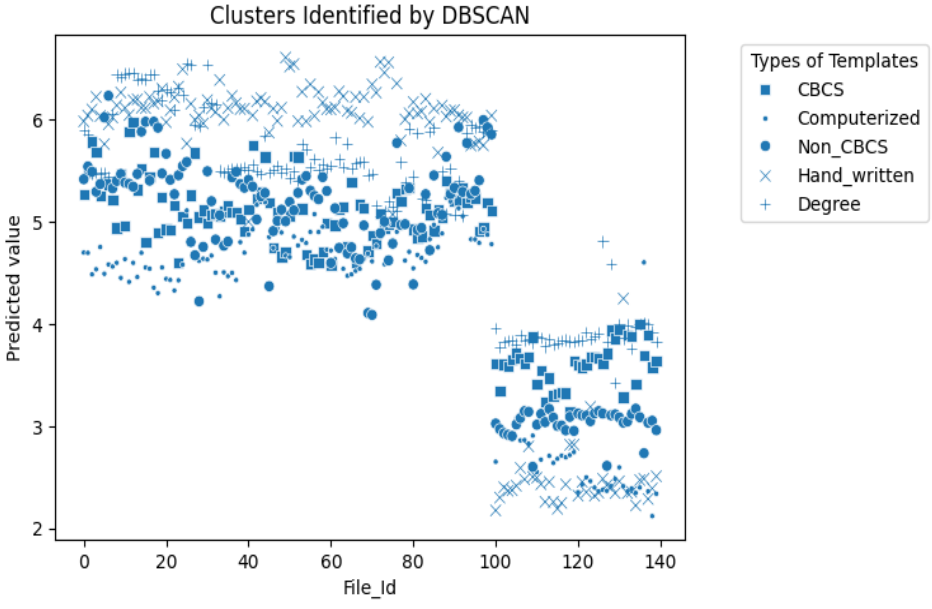


Figure. 4 Clustering of all categories

 DOI: 10.22266/ijies2026.0831.41

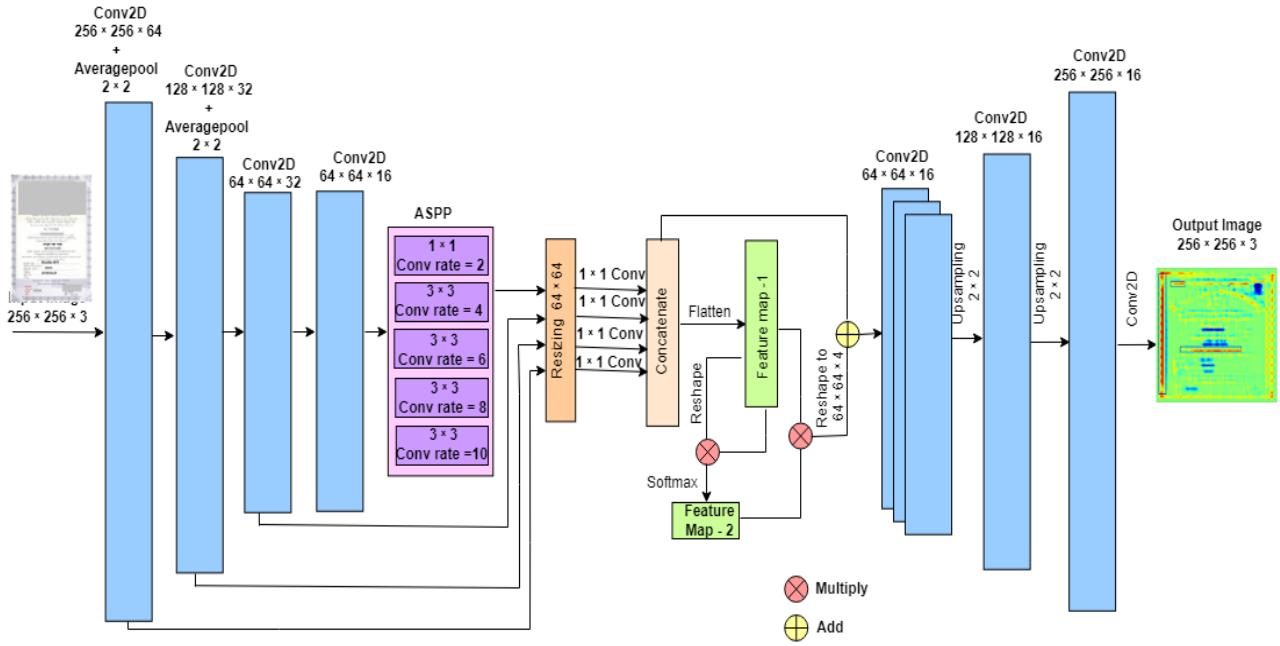


Figure. 5 Layer Diagram of the proposed feature fusion CNN model

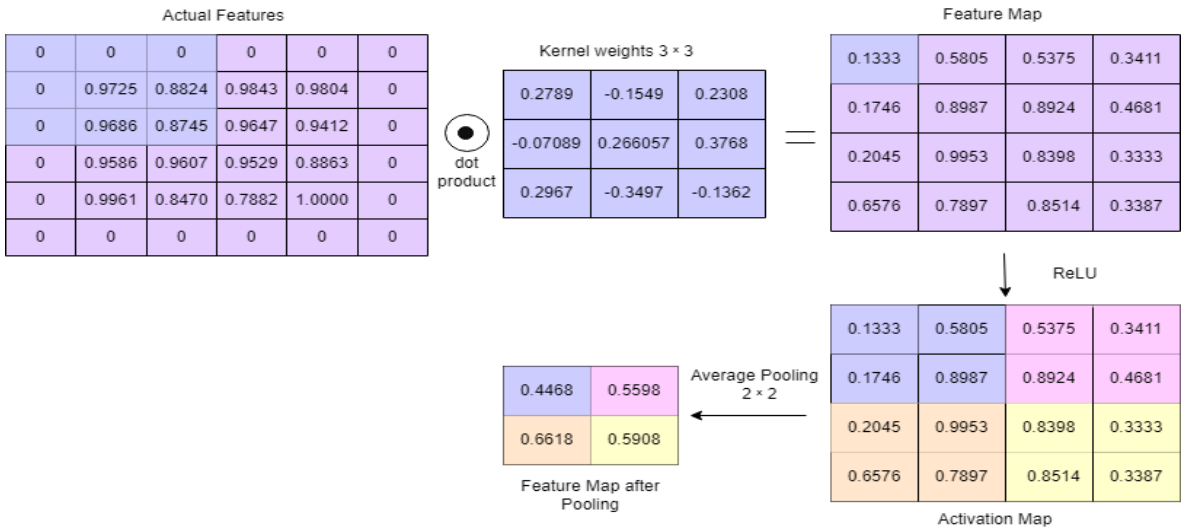


Figure. 6 Extraction of features using Average Pooling

These features were then concatenated following a 1 × 1 convolutional operation. This extracted feature map was transformed into a vector using flattening operation, referred to as 'Feature map-1' as represented in Eq. (9). To emphasize the essential features, these features were recalibrated by multiplication, addition, softmax, and reshape operations as shown in Fig. 5, and is defined in Eqs. (10) and (11). Finally, the image was reconstructed using upsampling with a size of 2 × 2 to get the predicted mask as mentioned in Eq. (12). Here, we applied the bicubic interpolation method to upsampling the image as represented in Eq. (13). The bicubic method works better than the bilinear method by considering the closest 16 pixels.

The entire process to predict the reconstructed mask is defined in Algorithm 2.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{2}$$

where y_i and $\hat{y}_i$ represent the actual and predicted value of features, and N represent the total number of samples in a batch.

$$\hat{y}_i = \mathrm{U}_{2\times2}\left(\mathrm{U}_{2\times2}\left(\mathrm{F}_{fus}\left(\mathrm{F}_{aspp}\left(\mathrm{Conv}_{(\mathrm{m,f})}(\mathrm{fea})\right)\right)\right)\right) \tag{3}$$

where, U represent the upsampling of the feature map

using the bicubic interpolation method with an upsampling window size of $U_{2\times2}$. F_{fus} denotes the feature map extracted using a multiplication and addition attention mechanism to emphasize essential features. F_{aspp} represents feature map extracted after applying the ASPP layer. $\mathrm{Conv}_{(m,f)}$ performs a convolution operation between features (fea) and the weight matrix in the f[th] filter of the m[th] convolutional layer.

$$\mathrm{Conv}_{(m,f)}(fea) = \sum_{j=1}^{k\_h \times k\_w} w_j x_j + b \quad (4)$$

where k_h and k_w represent height and width of kernel weight matrix. The w_j represents the kernel weight in the kernel matrix at the j[th] position. This weight matrix will stride around the entire feature map to perform the convolution operation. The features in feature map at j[th] position are represented by x_j , and b represents the bias value.

$$\sigma\left(\mathrm{Conv}_{(m,f)}(fea)\right) = \mathrm{ReLU}\left(\mathrm{Conv}_{(m,f)}(fea)\right) = \max\left(0, \mathrm{Conv}_{(m,f)}(fea)\right) \quad (5)$$

where fea represents the features generated after the convolution operation. It retains the same feature if it is greater than zero; otherwise, it changes the negative features to zero.

$$\mathrm{avg}_{pool} = p\_w \sum_{i=1}^{p\_h}\sum_{j=1}^{p\_w} \sigma\left(\mathrm{Conv}_{(m,f)}\left(fea_{ij}\right)\right) \quad (6)$$

where, p_h and p_w represent the height and width of the pooling window, fea_{ij} denotes feature at the [i,j] position in the pooling window.

$$aspp = \mathrm{Conv}_{(4,f)}(fea)^{rate} \quad (7)$$

Here, the aspp layer was used to extract the features in the 4[th] convolutional layer with different convolutional rates.

$$R_m = \mathrm{Concat}(\mathrm{resize}(\mathrm{Conv}_{(m,f)}(fea), aspp, (64 \times 64)) \quad (8)$$

where, R_m represents the concatenated features from all convolutional layers except the 4th convolutional layer, and aspp layer, which is all resized to 64×64.

$$F\_map_1 = \underset{R_m}{\rightarrow} \quad (9)$$

where, F_map1 convert the feature map of R_m to the vectors using the flatten operation.

$$F\_map_2 = (\mathrm{softmax}(F\_map_1))^T \times F\_map_1 \quad (10)$$

$$F\_map_3 = \mathrm{Reshape}(F\_map_1 \times F\_map_2) \quad (11)$$

$$F\_map_4 = U_{2\times2}(U_{2\times2}(\sigma\left(\mathrm{Conv}_{(out,3)}\left(F_{map_3}\right)\right))) \quad (12)$$

where, F_map2 multiply the transposed features of F_map1, which is activated using softmax, and the features of F_map_1.The F_map_3 reshape the feature map, which is received after applying the product operation between the features of F_map_1 , and F_map_2 .Finally, the predicted mask F_map_4 is achieved by applying upsampling with the bicubic interpolation method. $\mathrm{Conv}_{(out,3)}$ represents the output layer with three channels.

$$F_{bicu(x,y)} = \sum_{i=-1}^{2}\sum_{j=-1}^{2} F_{map_{3_{ij}}} . (x - x_0)^i . (y - y_0)^j \quad (13)$$

where, x and y are the coordinates of the point being interpolated, x_0 and y_0 are the coordinates of the top-left pixel in the 4×4 neighbourhood. They are the interpolation coefficients derived from the 16 surrounding pixels.

The predicted mask from the model emphasizes well of image-based features such as logo, and patterns of the templates' border in fake educational credentials, since the pixel intensities are high in image regions. To focus on text-based features with small font-size, the channel-wise multiplication operation was performed between the predicted mask and the actual mask as mentioned in Eq. (14). The actual mask refers to the actual template of the respective category. It will fetch the actual template based on the category since the experimental study focus on automation.

$$\mathrm{Locate\_region} = \prod_{i=1}^{3} P_i - \prod_{i=1}^{3} A_i \quad (14)$$

where, P_i and A_i represent the feature map of each channel in the predicted mask and actual mask, respectively.

Algorithm 2: Procedure for training the feature fusion CNN model

Input: $Train_{DS}$: training dataset, : Val_{DS} : validation dataset

Output: $PredMask$: reconstructed mask predicted by the model

Step 1: $l1 \leftarrow Conv2D(64,(3,3))$
Step 2: $l2 \leftarrow AveraagePooling2D((2,2))$
Step 3: $l3 \leftarrow Conv2D(32,(3,3))$
Step 4: $l4 \leftarrow AveraagePooling2D((2,2))$
Step 5: $l5 \leftarrow Conv2D(32,(3,3))$
Step 6: $l6 \leftarrow Conv2D(16,(3,3))$
// ASPP layer
Step 7: $d1 \leftarrow Conv2D(3(1,1), dilation\_rate = 2)\ (l6)$
Step 8: $d2 \leftarrow Conv2D(3(3,3), dilation\_rate = 4)\ (d1)$
Step 9: $d3 \leftarrow Conv2D(3(3,3), dilation\_rate = 6)\ (d2)$
Step 10: $d4 \leftarrow Conv2D(3(3,3), dilation\_rate = 8)\ (d3)$
Step 11: $d5 \leftarrow Conv2D(3(3,3), dilation\_rate = 10)\ (d4)$
Step 12: $a1 \leftarrow AveraagePooling2D((2,2))(d5)$
// Resizing
Step 13: $r1 \leftarrow Resize(l1,(64 \times 64))$
Step 14: $r2 \leftarrow Resize(l3,(64 \times 64))$
Step 15: $r3 \leftarrow Resize(l5,(64 \times 64))$
Step 16: $r4 \leftarrow Resize(a1,(64 \times 64))$
Step 17: Initialize $j \leftarrow 1$
//convert to one channel with 1×1 convolution
Step 18: Repeat j until j reaches 4
Convert rj with one channel, and perform
1×1 convolution
$j \leftarrow j + 1$
End j
Step 19: $fea\_fuse \leftarrow$ Concatenate()([$r1, r2, r3, r4$])
Step 20: $fea\_fuse' \leftarrow (\overrightarrow{fea\_fuse})$
Step 21: $fea\_fuse^T \leftarrow (fea\_fuse')^T$
Step 22: $fea\_mul \leftarrow fea\_fuse' \times fea\_fuse^T$
Step 23: $s \leftarrow Softmax(fea\_mul)$
Step 24: $att\_s \leftarrow s \times fea\_fuse'$
Step 25: $att\_s\_reshape \leftarrow Reshape((64,64,4)(att\_s))$
Step 26: $frm\_value \leftarrow [att\_s\_reshape] + [fea\_fuse]$
Step 27: $dec1 \leftarrow Conv2D(16,(3,3))(frm\_value)$
Step 28: $dec2 \leftarrow Conv2D(16,(3,3))(dec1)$
Step 29: $dec3 \leftarrow Conv2D(16,(3,3))(dec2)$
//upsampling
Step 30: $dec4 \leftarrow UpSampling2D\big((2,2)\big)(dec3)$
Step 31: $dec5 \leftarrow UpSampling2D\big((2,2)\big)(dec4)$
Step 32: $PredMask \leftarrow Conv2D(3)(dec5)$

3.3 Performance evaluation metrics

The performance of fake educational credential detection using the DBSCAN method was evaluated with the following statistical metrics: Precision, Recall, F1-Score, and Accuracy as mentioned in Eqs. (15)-(18).

$$\text{Precision} = \text{TP}/\ \text{TP} + \text{FP} \tag{15}$$

$$\text{Recall} = \text{TP}/\ \text{TP} + \text{FN} \tag{16}$$

$$\text{F1} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{Precision + Recall} \tag{17}$$

$$\text{Accuracy} = \text{TP} + \text{TN}/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{18}$$

where, TP represents true positive, FP denotes false positive, TN symbolizes true negative, and FN represents false negative.

Further, we evaluated the prediction of clustering method using following metrics: Rand Index (RI), which measures the similarity between two clusters by counting the pairs in the similar or different clusters in the actual and predicted clusters using Eq. (19). Adjusted Rank Index (ARI) is ensured to have a value near zero if all labels are independent of clusters, and reaches 1 when the clustering is identical, and it is measured using Eq. (20). Normalized Mutual Information Score (NMI) returns 0 when there is no mutual information, and 1 when there is a perfect correlation. Homogeneity indicates that each cluster has data points belonging to the same class label. Completeness shows that all data points from the same class are allotted to the same cluster. V-measure indicates the harmonic mean of homogeneity and completeness as seen in Eq. (21). Fowlkes-Mallows Index (FMI) calculates the geometric mean of recall and precision using Eq. (22).

$$\text{RI} = (\text{no. of agreeing pairs})/(\text{no. of pairs}) \tag{19}$$

$$\text{ARI} = (\text{RI} - \text{expected\_RI})/(\max(\text{RI}) - \text{expected\_RI}) \tag{20}$$

$$\text{V} - \text{measure} = \frac{(1+\beta)*\text{homogeneity}*\text{completeness}}{(\beta*\text{homogeneity}+\text{completeness})} \tag{21}$$

$$FMI = TP/\sqrt{(TP + FP) * (TP + FN)} \quad (22)$$

where , β represents the score to weightage the homogeneity vs. completeness.

If β is greater than one, it emphasizes completeness, otherwise it emphasizes homogeneity.

4. Results and discussion

The performance of detecting fake educational credentials using statistical metrics such as accuracy, F1-score, recall, precision, confusion matrix, and AUROC curve is discussed in Section 4.1.

Section 4.2 represents the localization of fake regions in fake credentials using the proposed feature fusion CNN model. Additionally, Section 4.3 compares the localization of fake regions using the proposed model with the pre-trained customized Autoencoder (AE) model, and the MantraNet model. Finally, Section 4.4 analyzes the model's prediction using statistical method t-test.

4.1 Results of fake detection

The performance of detecting fake educational credentials is analyzed by statistical metrics, and the results are presented in Table 2. The CBCS category has a higher score than the other categories, because this template has background textures with complex patterns, while the remaining categories have text-based background textures, which provided moderate performance. The Non-CBCS and Degree certificates were the least accurate among all categories because these templates have logos in various colours. Hence, the accuracy of these categories was less than that of other categories. The overall accuracy of all five categories using the proposed method in detecting fake educational credentials is 99.14%.

Further, the performance of detecting fake educational credentials is analyzed by the similarity score between two clusters using statistical metrics, and the results are presented in Table 3. The confusion matrix of each category and the consolidated prediction is depicted in Fig. 7(a) to (f). The AUROC curve is plotted for five categories in Fig. 8.

4.2 Results of locating fake region

The results of locating fake regions in fake

Table 2. Performance analysis of fake detection

Sl. No	Category	Precisio n	Recall	F1-Score	Accuracy
1.	CBCS	100	100	100	100.00
2.	Computerized	99.00	100	99.50	99.28
3.	Degree	98.00	100	99.00	98.57
4.	Hand-written	99.00	100	99.50	99.28
5.	Non-CBCS	98.00	100	99.00	98.57
6.	Overall	99.00	100	99.00	99.14

Table 3. Performance of Cluster metrics

Sl. No.	Metrics	Category					
		CBCS	Computerized	Degree	Handwritten	Non-CBCS	Overall
1.	Rand Index	0.98	0.97	0.96	0.97	0.96	0.96
2.	Adjusted Rank Index	0.97	0.94	0.91	0.94	0.91	0.94
3.	Normalized MIS	0.94	0.89	0.85	0.89	0.85	0.88
4.	Homogeneity	0.93	0.88	0.84	0.88	0.84	0.87
5.	Completeness	0.94	0.90	0.86	0.90	0.87	0.89
6.	V-measure	0.94	0.89	0.85	0.89	0.85	0.88
7.	Fowlkes_Mallows_Score	0.99	0.98	0.96	0.98	0.96	0.97

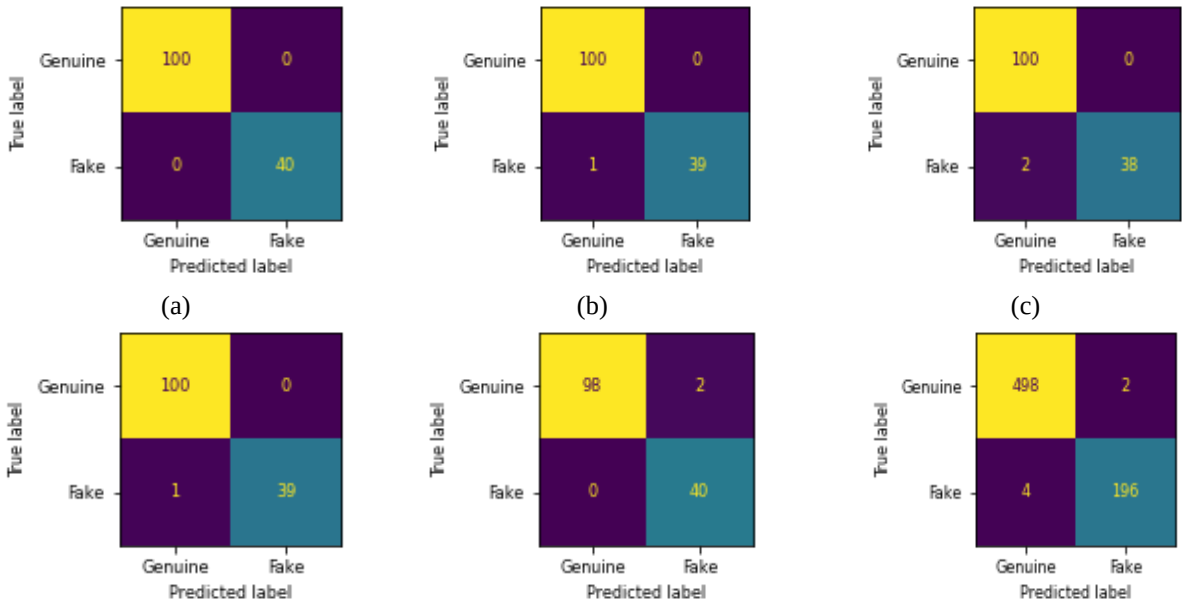


Figure. 7 Confusion matrix: (a) CBCS, (b) Computerized, (c) Degree,(d)Hand-written, (e) Non-CBCS, and (f) Overall

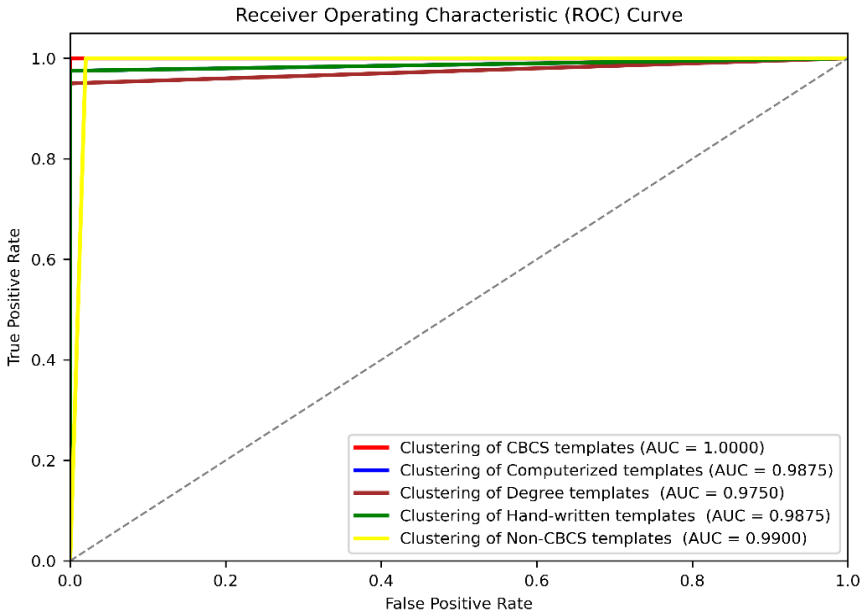


Figure. 8 AUROC of Clustering model

educational credentials are shown in Fig. 9 using the proposed feature fusion CNN model.

The row represents the five types of credentials: Hand-written, Computerized, CBCS, Degree, and Non-CBCS. The first three columns represent the original credentials, the fake credentials, and the ground-truth mask for the fake credentials, respectively. The ground truth mask is generated manually from the noted fake regions in the template. The predicted mask of fake regions using the proposed model is located in the fourth column. The drawback of this prediction is that it locates the image regions well, but does not consider the text regions.

Hence, the channel-wise multiplication was applied between the predicted and actual masks, and is located in the fifth column, which locates both image and text-based fake regions better than other methods.

4.3 Comparative results of locating a fake region

In this section, the localization of fake regions in the fake credentials is illustrated in Fig. 10 using theproposed model, pre-trained customized Autoencoder (AE) model which contains four layers with filter sizes of 64,32,16, and 8 at encoder part; four layers with filter sizes of 8,16,32, and 64 at

decoder part. The upsampling was done with bicubic interpolation method. The MantraNet displays only some standard features in the templates, as it was trained on synthetic datasets with copy-move and splicing forgery.

Here, we calculate the t-statistic value for the test data samples in 'Fake' and 'Genuine' groups using the proposed fusion model as mentioned in Eq. (24). The predicted value is stored in a list to find the t-statistic value. We chose the $\alpha = 0.05$ (5% level of significance) with the significance level of 5%. Hence, conclude that, there is a significant difference between the two groups if the p-value is less than 0.05; otherwise, no difference. The t-statistic value and p-value for all five categories are represented in Table 4, and illustrated in Fig. 11. Even though the proposed fusion model highlights the fake regions, we proved the model prediction by a t-test.

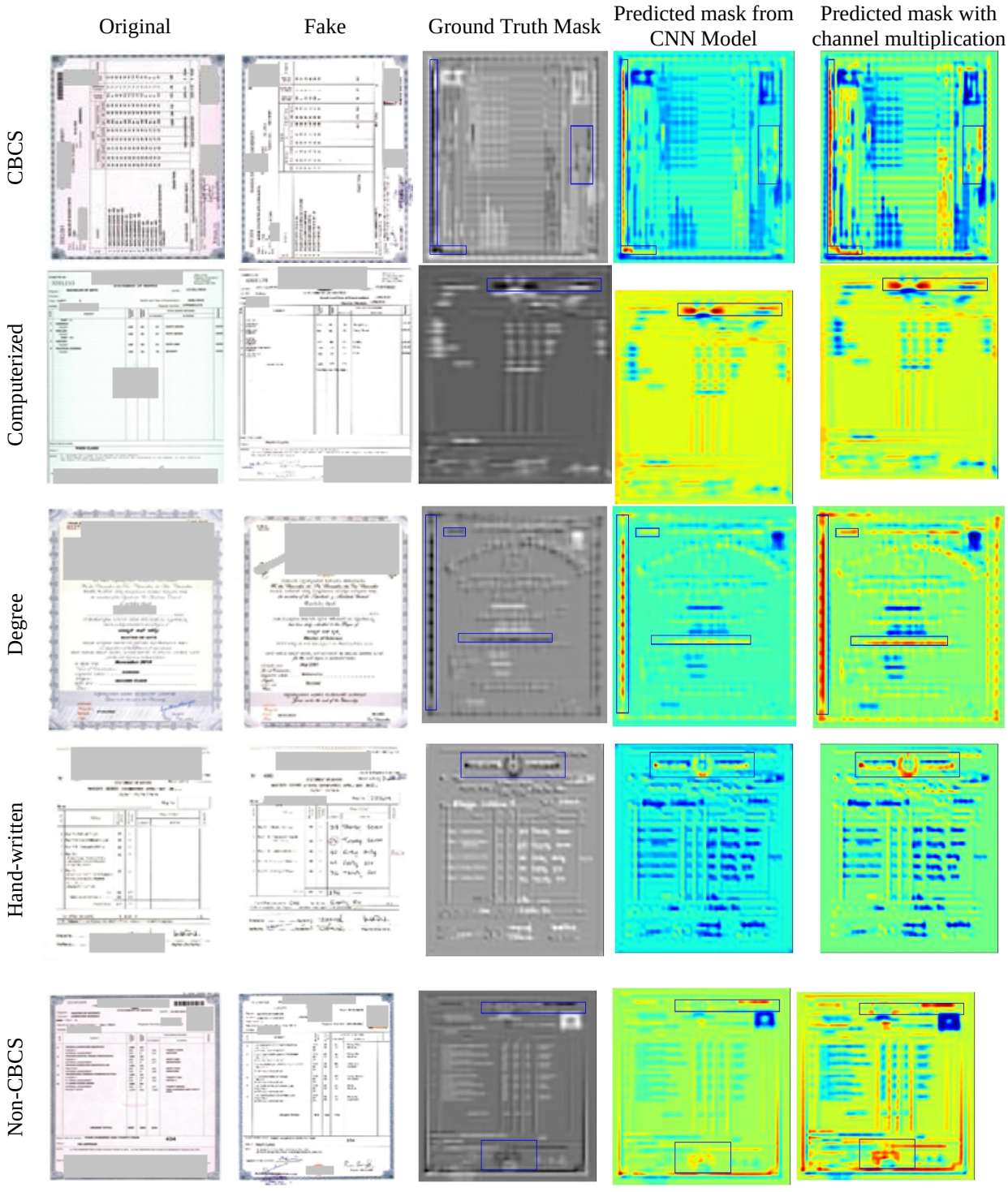


Figure. 9 Locating a fake region

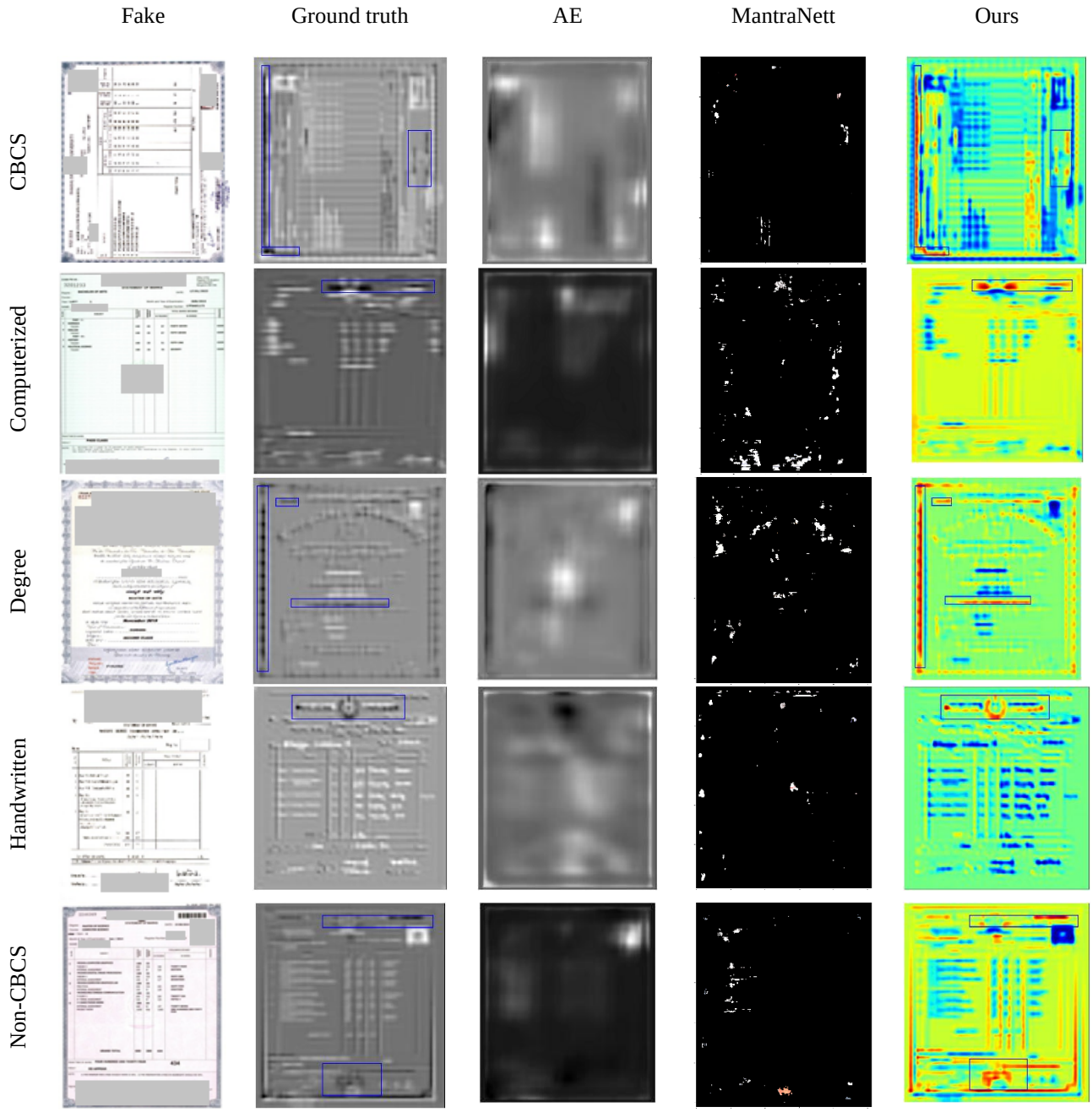


Figure. 10 Comparative results of locating fake regions

Table 4. Statistical analysis of model prediction using t-test

Sl. No.	Comparison of two groups	Category	t_statistic	p_value
1.	Fake vs Genuine	CBCS	5.0769	1.219e-6
2.	Fake vs Genuine	Computerized	5.7220	6.277e-12
3.	Fake vs Genuine	Degree	3.1868	2.40e-3
4.	Fake vs Genuine	Hand-written	7.6571	2.995e-12
5.	Fake vs Genuine	Non-CBCS	20.747	2.945e-44

$$t = (\bar{x}_1 - \bar{x}_2) / \sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}} \quad (23)$$

where, $\bar{\boldsymbol{x}}_{\boldsymbol{1}}$ and $\bar{\boldsymbol{x}}_{\boldsymbol{2}}$ are the mean of two groups, ${s_1}^2$ and ${s_2}^2$ are the variances of two groups, and n_1 and $\boldsymbol{n_2}$ are the number of samples in group1 and group2 respectively.

$$\text{sample\_score} = \mu(\hat{y}_i) \quad (24)$$

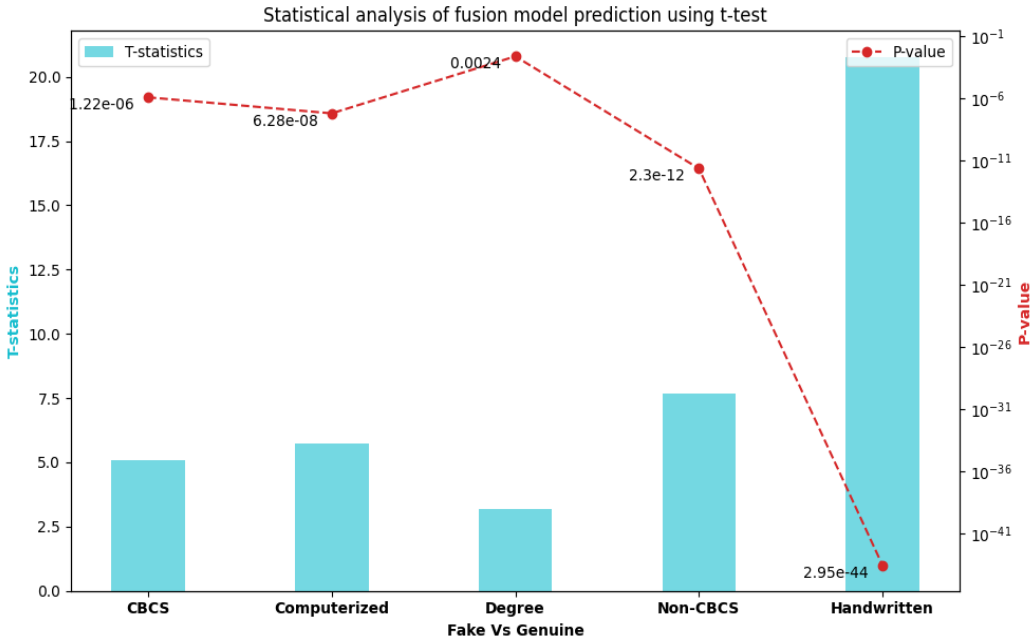


Figure. 11 Statistical analysis of proposed fusion model prediction

where, μ represents the mean, and $\hat{\mathbf{y}}_i$ is the predicted heatmap using the fusion model.

5. Conclusion

The fake educational credentials across five categories were detected using Shannon entropy, and the DBSCAN clustering method provided an accuracy of 99.14%. This approach is highly suitable for educational credentials with complex texture patterns. To assess the predictability of detecting fake educational credentials, we designed a feature fusion CNN model, which integrates an ASPP layer and an attention mechanism to locate the fake regions in fake credentials. While this model effectively predicted the image-based fake regions, it faces difficulties with text-based regions, especially when the font size is small. To address this, we applied channel-wise multiplication between actual and predicted masks to accurately locate text-based regions. Further, we evaluated the localization capabilities of our proposed model against those generated by a pre-trained customized AE model, and the MantraNet model. To validate the predictions made by the feature fusion model, we demonstrated a statistically significant difference between the 'fake' and 'genuine' groups using the t-test.

This approach enhances the automated authentication of educational credentials, and effectively locates fake regions. Additionally, it can be used in other fields, such as the identifying disease-affected areas in medical imaging in healthcare, and detecting forged documents in e-governance. Future work will focus on detection of fake educational credentials and locating fake regions based on content using Natural Language Processing techniques.

Acknowledgment

The authors thank Bangalore University for providing the scanned datasets of mark cards and degree certificates for this research work. The author Sarala M acknowledges the Department of Science and Technology (DST), Karnataka Science and Technology Promotion Society (KSTePS), Government of Karnataka, India, for receiving the fellowship support (Award letter No.: DST/KSTePS/Ph.D. Fellowship/PHY-06: 2022 – 23/472) for research work.

Data Availability

The dataset used in this study contains sensitive institutional information collected from multiple departments of the university and therefore cannot be publicly released. To support reproducibility, the authors will make available the source code, trained model weights, preprocessing procedures, experimental configurations, hyperparameter settings, and evaluation scripts upon reasonable request. Researchers interested in reproducing the study may contact the corresponding author. The dataset itself may be shared only when institutional regulations and ethical considerations permit.

Conflicts of interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization: Sarala M, and Muralidhara B L; Methodology: Sarala M; Software: Sarala M; Validation: Sarala M and Muralidhara B L; formal analysis: Muralidhara B L; investigation: Muralidhara B L; writing—original draft preparation and editing: Sarala M; Review: Muralidhara B L, Suresh R, and Rashmi Siddalingappa; Supervision: Muralidhara B L.

References

[1] Scroll Staff, "Bengaluru police uncover fake mark sheet and degree certificate racket, says report", *Scroll.in*, 2017. [Online]. Available: https://scroll.in/latest/836710/bengaluru-police-uncover-fake-mark-sheet-and-degree-certificate-racket-says-report

[2] "Bengaluru: They offered fake degrees, even PhDs for a price", Deccan Chronicle, 2019. [Online]. Available: https://www.deccanchronicle.com/nation/crime/270118/bengaluru-they-offered-fake-degrees-even-phds-for-a-price.html

[3] H. M. C. Swamy, "Two involved in fake Bangalore University marks card racket arrested", *Deccan Herald*, 2022. [Online]. Available: https://www.deccanherald.com/india/karnataka/bengaluru/two-involved-in-fake-bangalore-university-marks-card-racket-arrested-1139063.html

[4] Express News Service, "Fake mark sheets of universities sold via WhatsApp, gang arrested in Bengaluru", *Indian Express*, 2022. [Online]. Available: https://indianexpress.com/article/cities/bangalore/fake-mark-sheets-universities-whatsapp-gang-bengaluru-arrest-8310271/

[5] Express News Service, "Bengaluru: Crime branch raids five organisations, seizes fake marksheets", *Indian Express*, 2023. [Online]. Available: https://indianexpress.com/article/cities/bangalore/bengaluru-crime-branch-raids-five-organisations-seizes-fake-marksheets-8407968/

[6] C. Shi, L. Chen, C. Wang, X. Zhou, and Z. Qin, "Review of image forensic techniques based on deep learning", *Mathematics*, Vol. 11, No. 14, p. 3134, 2023, doi: 10.3390/math11143134.

[7] P. Sharma, M. Kumar, and H. Sharma, "Comprehensive analyses of image forgery detection methods from traditional to deep learning approaches: an evaluation", *Multimedia Tools and Applications*, Vol. 82, No. 12, pp. 18117–18150, 2023, doi: 10.1007/s11042-022-13808-w.

[8] I. Shallal, L. R. Haddada, and N. E. B. Amara, "Image forgery detection with focus on copy-move: An overview, real world challenges and future directions", *Applied Sciences*, Vol. 15, No. 21, p. 11774, 2025, doi: 10.3390/app152111774.

[9] P. Deb, *et al*., "Image forgery detection techniques: Latest trends and key challenges", *IEEE Access*, Vol. 12, pp. 169452–169466, 2024, doi: 10.1109/ACCESS.2024.3498340.

[10] H. A. Raheem, M. A. Mohammed, and A. S. H. M. Ali, "Enhancing the Image Forgery Detection based Machine Learning Approach using Multiple Datasets", *Engineering, Technology & Applied Science Research*, Vol. 15, No. 3, pp. 22739–22745, 2025, doi: 10.48084/etasr.10151.

[11] A. K. Jaiswal and R. Srivastava, "Fake region identification in an image using deep learning segmentation model", *Multimedia Tools and Applications*, Vol. 82, No. 25, pp. 38901–38921, 2023, doi: 10.1007/s11042-023-15032-6.

[12] F. Z. Mehrjardi and M. S. Zarchi, "A hybrid model for image forgery detection using deep learning with block and keypoint methods", *Scientific Reports*, 2026, doi: 10.1038/s41598-026-41473-8.

[13] R. Sukhija, M. Kumar, and M. K. Jindal, "Document forgery detection: a comprehensive review", *International Journal of Data Science and Analytics*, Vol. 20, No. 5, pp. 4385–4407, 2025, doi: 10.1007/s41060-025-00723-0.

[14] W. Li, *et al*., "Document image forgery detection and localization in desensitization scenarios", *Signal Processing*, Vol. 238, p. 110123, 2026, doi: 10.1016/j.sigpro.2025.110123.

[15] M. Sirajudeen and R. Anitha, "Forgery document detection in information management

system using cognitive techniques", *Journal of Intelligent & Fuzzy Systems*, Vol. 39, No. 6, pp. 8057–8068, 2020, doi: 10.3233/JIFS-189128.

[16] Z. Zhao, *et al*., "DocForge-Bench: A comprehensive benchmark for document forgery detection and analysis", *arXiv Preprint*, arXiv:2603.01433, 2026.

[17] Y.-Y. Bae, D.-J. Cho, and K.-H. Jung, "Enhancing document forgery detection with edge-focused deep learning", *Symmetry*, Vol. 17, No. 8, p. 1208, 2025, doi: 10.3390/sym17081208.

[18] Y. Sun, R. Ni, and Y. Zhao, "MFAN: multi-level features attention network for fake certificate image detection", *Entropy*, Vol. 24, No. 1, p. 118, 2022, doi: 10.3390/e24010118.

[19] A. B. Kasgari, S. Safavi, M. Nouri, J. Hou, N. T. Sarshar, and R. Ranjbarzadeh, "Point-of-interest preference model using an attention mechanism in a convolutional neural network", *Bioengineering*, Vol. 10, No. 4, p. 495, 2023, doi: 10.3390/bioengineering10040495.

[20] Y. Wang, W. Wang, Y. Li, Y. Jia, Y. Xu, Y. Ling, and J. Ma, "An attention mechanism module with spatial perception and channel information interaction", *Complex & Intelligent Systems*, Vol. 10, No. 4, pp. 5427–5444, 2024, doi: 10.1007/s40747-024-01445-9.

[21] H. Lv, J. Chen, T. Pan, T. Zhang, Y. Feng, and S. Liu, "Attention mechanism in intelligent fault diagnosis of machinery: A review of technique and application", *Measurement*, Vol. 199, p. 111594, 2022, doi: 10.1016/j.measurement.2022.111594.

[22] V. V. Sataraddi and N. P. Nethravathi, "An Attention-Enhanced Multi-Scale Framework for Copy–Move Forgery Detection and Localization", *Engineering, Technology & Applied Science Research*, Vol. 16, No. 3, pp. 34927–34933, 2026, doi: 10.48084/etasr.18383.

[23] X. Wang, *et al*., "Supervised face tampering detection based on spatial channel attention mechanism", *Electronics*, Vol. 14, No. 3, p. 500, 2025, doi: 10.3390/electronics14030500.

[24] Y. Rao, J. Ni, and H. Xie, "Multi-semantic CRF-based attention model for image forgery detection and localization", *Signal Processing*, Vol. 183, p. 108051, 2021, doi: 10.1016/j.sigpro.2021.108051.

[25] C. Nie, D. Zhou, and R. Nie, "EDAfuse: A encoder-decoder with atrous spatial pyramid network for infrared and visible image fusion", *IET Image Processing*, Vol. 17, No. 1, pp. 132–143, 2023, doi: 10.1049/ipr2.12622.

[26] M.-H. Park, J.-H. Cho, and Y.-T. Kim, "CNN model with multilayer ASPP and two-step cross-stage for semantic segmentation", *Machines*, Vol. 11, No. 2, p. 126, 2023, doi: 10.3390/machines11020126.

[27] Z. Shang, H. Xie, Z. Zha, L. Yu, Y. Li, and Y. Zhang, "PRRNet: Pixel-Region relation network for face forgery detection", *Pattern Recognition*, Vol. 116, p. 107950, 2021, doi: 10.1016/j.patcog.2021.107950.

[28] Y. Wu, W. AbdAlmageed, and P. Natarajan, "Mantra-net: Manipulation tracing network for detection and localization of image forgeries with anomalous features", In: *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9543–9552, 2019, doi: 10.1109/CVPR.2019.00977.

[29] J. Wang, Z. Wu, J. Chen, X. Han, A. Shrivastava, S.-N. Lim, and Y.-G. Jiang, "Objectformer for image manipulation detection and localization", In: *Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2364–2373, 2022, doi: 10.1109/CVPR52688.2022.00240.

[30] C. Dong, X. Chen, R. Hu, J. Cao, and X. Li, "Mvss-net: Multi-view multi-scale supervised networks for image manipulation detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, No. 3, pp. 3539–3553, 2022, doi: 10.1109/TPAMI.2022.3180556.

[31] X. Lin, S. Wang, J. Deng, Y. Fu, X. Bai, X. Chen, X. Qu, and W. Tang, "Image manipulation detection by multiple tampering traces and edge artifact enhancement", *Pattern Recognition*, Vol. 133, p. 109026, 2023, doi: 10.1016/j.patcog.2022.109026.