Received: 11 May 2026 | Revised: 28 July 2026 | Accepted: 14 August 2026 | Published online: 29 August 2026

RESEARCH ARTICLE

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–19
DOI: 10.47852/bonviewAIA620210360

BON VIEW PUBLISHING

Groupwise-Spatial Attention-Based Dense Neural Network for Image Multi-Forgery Detection

S. Aruna[1,*] and Surabhi Narayan[1]

[1] *Department of Computer Science and Engineering, PES University, India*

Abstract: Digital images have become the primary source of information exchange on digital platforms, making digital image forgery detection important and crucial in the field of digital forensics. Among various illegal manipulation techniques, copy-move and splicing manipulation are the two most common types of forgery. Copy-move forgery involves duplicating the areas within the same image, whereas splicing forgery involves taking out areas or objects from one or more images and combining them together to produce a spliced image. Detecting various forgeries in an image has become a crucial requirement due to the presence of subtle artifacts and minimal inconsistencies during the forgery process. Hence, a multi-forgery detection framework based on Groupwise-Spatial Attention-based Dense Neural Network (GSA-DNN) is proposed to detect copy-move and splicing forgery. Error level analysis (ELA) technique is used to detect image manipulation by figuring out the compression differences between authentic and forged images, followed by a generic standardizing procedure for all images based on dimensions and pixel values so as to improve final model convergence. InceptionNet V3 is employed to extract the multiscale features through convolutional kernels, which enables the learning of global and local feature representation. The proposed GSA-DNN method obtains high detection accuracy of 99.58%, 97.77%, and 94.59% on benchmark datasets such as CASIA V2, CoMoFoD, and Multiple Image Splicing Datasets compared to existing methods such as the Siamese neural network.

Keywords: error level analysis, groupwise-spatial attention-based dense neural network, image forgery, InceptionNet V3, Siamese neural network

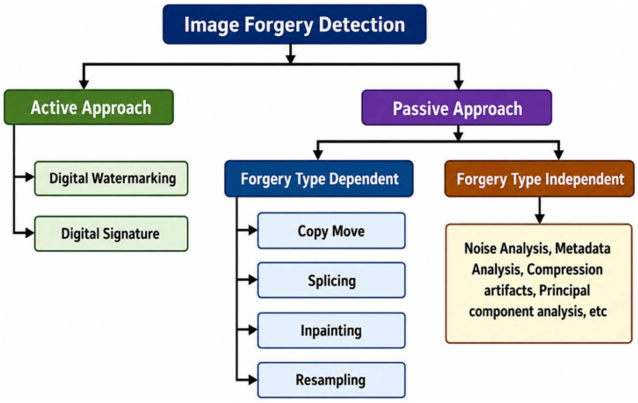
1. Introduction

Image forgery refers to the malicious alteration of digital images, and its detection falls under digital image forensics, involving the use of various tools to verify image authenticity. Forgery detection methods are broadly categorized into active and passive approaches [1], as illustrated in Figure 1. In the active approach, additional information about the image is embedded in advance. Watermarks integrated into images are employed for authentication. Compared to conventional watermarks, digital watermarks [2] have manipulations like compression and resizing. In a passive approach, no prior information is integrated into the image, but the images are analyzed for visual, statistical, and structural inconsistencies to identify anomalies in the image. Passive forgery detection is categorized into forgery-type dependent and forgery-type independent methods. The forgery-type dependent method identifies manipulation such as copy-move, splicing, image retouching [3], resampling, and in-painting. Among these, splicing and copy-move are commonly employed in image forgery detection [4]. The forgery-type independent method employs techniques based on metadata analysis, noise analysis, principal component analysis, and compression artifacts. Copy-move forgery [5] involves copying a region from an image and pasting it into another location in the same image [6].

Copy-move forgery involves post-processing operations like scaling, rotation, and smooth blending of the duplicated regions to make detection more challenging [7]. Splicing forgery involves copying a region of one image and inserting it into another, typically integrating images captured under diverse lighting conditions and obtaining varying noise levels [8]. To achieve a seamless blend with the target image, the spliced region typically undergoes multiple post-processing adjustments [9]. As a result, the primary objective of copy-move forgery detection is to identify duplicated regions within the same image, whereas splicing detection focuses on identifying manipulated forged regions using advanced approaches with a vast proliferation of fake images [10]. Other types of image forgery include retouching, resampling [11], and object removal [12]. Retouching involves altering visual attributes like color, brightness, and similar visual attributes to represent particular features [13, 14]. Object removal is the act of removing an object entirely from an image [15]. Image forensics in recent times has gained significant research attention, leading to the development of different forensic methods for identifying manipulations in forged images [16, 17]. Diverse methods are developed to identify and localize image forgeries [18] to ensure the authenticity of digital information. The deep learning (DL) [19] method is chosen for detecting multi-forgery

***Corresponding author:** S. Aruna, Department of Computer Science and Engineering, PES University, India. Email: PES1PD20CS001@stu.pes.edu

Figure 1
Overview of digital image forgery detection methods and manipulation categories



because it automatically learns the intricate feature representation when compared to machine learning models. This ensures DL [20] method to capture subtle patterns across multiple types of manipulations [21]. Hence, DL provides high robustness in multi-forgery detection [22].

Existing methods face challenges in detecting subtle manipulation, small forged regions, and intricate multi-forgery scenarios. These challenges demonstrate the need for a robust method to preserve the integrity and authenticity of digital images via accurate forgery detection.

Therefore, Groupwise-Spatial Attention-based Dense Neural Network (GSA-DNN) is proposed for multi-forgery detection. Groupwise-spatial attention (GSA) mechanism divides feature representations into diverse groups and then learns spatial attention for each group independently instead of generating a single spatial attention map for all feature channels as in conventional attention mechanisms. This process enables the model to extract diverse forgery characteristics at different spatial locations, which preserves fine-grained manipulations. Moreover, the integration of the error level analysis (ELA) method, multi-level feature extraction via InceptionV3, and GSA-based dense learning leads to the development of a unified model for multi-forgery detection.

The main contributions of the proposed GSA-DNN are presented as follows:

1) GSA-DNN is used to detect multiple forgeries by dividing feature maps into groups and learning independent spatial attention for each group. The GSA mechanism focuses on manipulation-related regions by minimizing irrelevant information, which enables the model to capture subtle artifacts and enhance localization and detection performance.
2) InceptionNet V3 is used for effective feature extraction with the help of multiscale convolutional filters, effectively capturing diverse patterns.
3) ELA identifies the forgery by highlighting the differences in compression values between the authentic and forged regions of an image.

The remaining part of the paper is organized as follows. Section 2 discusses in detail the literature review of the existing works. Section 3 describes in detail the proposed methodology. Section 4 presents the results obtained from various experiments performed. Section 5 provides the conclusion.

2. Literature Survey

The existing works on multi-forgery detection using the CASIA V2, CoMoFoD, and Multiple Image Splicing Datasets (MISD) datasets are discussed in this section.

Ramirez-Rodriguez et al. [23] presented a Siamese neural network to find the manipulated regions in images. This work identifies a single feature representation of specific parts of an image, and by doing a feature-based comparison, it identified the forged areas effectively. The Siamese network used here has two identical branches with shared weights and features to process image blocks and identify forged areas. Also, the K-means algorithm was employed to determine similar centroids, facilitating accurate localization of repeated regions within the image. However, the presented method faces challenges in accurately identifying subtle manipulations as it focused on feature similarity rather than pixel information. Nazir et al. [24] suggested an enhanced Mask Region-based Convolutional Neural Network (Mask RCNN) for detecting forged images. DenseNet-41 was employed to extract keypoint features and then performed localization using Mask RCNN to recognize the manipulated areas. Mask RCNN identifies manipulated regions by generating a semantic mask with a classification score, which minimizes computational complexity. Qazi et al. [25] introduced a ResNet50v2 to detect image forgery. ResNet50v2 employed a residual layer to enhance the detection rate of the tampered regions. Transfer learning with pre-trained You Only Look Once-CNN weights generates feature initialization that minimizes complexity and improves detection performance. Shinde et al. [26] developed a Graph Convolution Network (GCN) to detect and categorize image forgeries. The GCN was utilized to improve the feature extraction process through structural and spatial dependencies. GCN represents an image as a graph, where pixel values are represented as nodes and edges are treated as spatial and structural relationships. Later, a support vector machine (SVM) was used for classifying the images. However, GCN struggles to accurately detect multiple forgeries due to its limited ability to capture fine-grained spatial relationships.

Ganguly et al. [27] presented a Local Tetra Pattern (LTrP) method to localize tampered regions via block-level image comparison. Initially, input images were partitioned into overlapping blocks, and then LTrP features were extracted from each block to produce a feature vector. LTrP method employed a shift-vector-based outlier removal method to enhance model performance in forged region localization. Nevertheless, LTrP relies only on local texture, which affects the model's ability to differentiate between tampered and genuine areas. Jaiswal and Srivastava [28] introduced a Convolutional Neural Network (CNN) to detect image splicing. The feature maps were extracted from convolutional layers and then upsampled in the decoder block. A sigmoid function was used to categorize pixels as forged or non-forged regions. Nevertheless, varying input scales led to inconsistent performance because different scales generated distinct feature representations that caused the model to focus on irrelevant information. Kadam et al. [29] established a Mask Region CNN with MobileNetV1 to identify image splicing forgery. MobileNetV1 was used to extract features from forged regions, whereas the region of interest (ROI) produced feature maps with the help of the Region Proposal Network. Then, the ROI was used to effectively handle precise spatial locations. The bounding box locations and corresponding segmentation masks were generated by utilizing a Fully Convolutional Network. Jalab et al. [30] presented a pixel's fractional mean (PFM) method to detect image splicing. Histogram of Gradient

(HoG) was used to evaluate the local pixel intensity distribution in each region of the image, while the Gray-Level Co-occurrence Matrix characterized image texture based on the spatial relationship between pixels to extract relevant features. PFM was used to adjust each pixel individually depending on its frequency, and these extracted features were fed into SVM to categorize the image as authentic or spliced image. Nevertheless, the PFM method obtains less effectiveness in detecting intricate splicing regions with subtle manipulations due to pixel-level inconsistencies. Ding et al. [31] developed a dual-channel U-Net to detect image splicing forgery by using a high-pass filter that generates a residual image having edge information. A standard dual-channel encoder and decoder approach is used to build the prediction of forged areas in images. Dual-channel U-Net enhanced feature extraction performance by leveraging multiscale context data and combining complementary image channels. However, dual-channel U-Net struggled to adapt to varying spatial resolutions and intricate image features, which led to blurred boundaries.

El Biach et al. [32] established a CNN based on an encoder and decoder to determine the manipulated regions. The encoder employed ResNet50 to extract discriminative features between non-manipulated and manipulated regions. To perform prediction, the decoder effectively learns the mapping from low-resolution feature maps. However, CNN struggled to detect subtle dependencies because of its limited receptive field, which minimized model performance. Walia et al. [33] introduced a Block-Quaternion Discrete Cosine Transform (QDCT) method for detecting image forgery. Quaternions represented the image as a vector field, while discrete cosine transform coefficients were extracted from the color components. Then, the Markov-based features were used to identify complex alterations introduced in the forgery process. Nevertheless, QDCT minimizes model performance due to its sensitivity to compression artifacts, which affects the proper detection of manipulated regions. Sabeena and Abraham [34] established a Convolutional Block Attention Mechanism (CBAM) to detect image forgery. CBAM integrates channel and spatial attention for capturing texture and contextual information, which enhances feature representation. Then, Atrous Spatial Pyramid Pooling was used to concatenate the scaled correlation maps, which result in a coarse mask. At last, bilinear upsampling was applied for resizing the predicted output to match the original image size. CBAM enhanced feature representation through an attention mechanism based on both channel and spatial dimensions, which enabled the network to stay focused on the significant parts of the input. The CBAM struggled with images that had undergone substantial noise or compression, thereby reducing its effectiveness in distinguishing forged areas.

Rao et al. [35] presented a self-supervised domain adaptation with a Siamese model and a Compression Approximation Network (ComNet) to localize image forgery. ComNet was employed to approximate the compression process through supervised learning, which enhances model performance. Then, the backbone network was trained with a domain adaptation mechanism to localize the tampered regions and boundaries with Joint Photographic Expert Group (JPEG) images. Nevertheless, the Siamese approach suffers from high variability in forgery styles due to its dependence on learned similarity measures, resulting in limited ability to capture intricate features. Diwan et al. [36] developed a Multiscale Detector Neural Network to localize copy-move forgery. The multiscale detector employed a sophisticated function to accurately detect and localize the copy-move manipulation. Nevertheless, the multiscale detector struggled to accurately identify forgery instances, which makes it difficult to differentiate between forged and non-forged images.

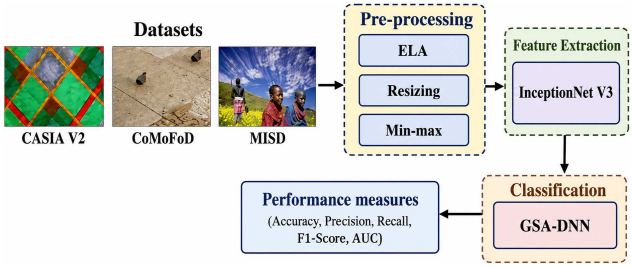
Diwan et al. [37] suggested a CenSure keypoint detection and CNN to identify and localize copy-move forgery. By combining key points with CNN features, the suggested method improved detection performance, even in the presence of attacks like geometrical transformations, scaling, and rotation. Also, the suggested method managed post-processing operations like noise, JPEG compression, color reduction, image blur, contrast adjustment, and brightness changes. However, the CenSure keypoint with CNN missed smooth regions as it focused on high-contrast key points, which limits learning features from forged regions. Diwan et al. [38] evaluated different methods used in image forensics like authentication, tampering detection, and source device identification. The proposed method obtains better performance in detecting and categorizing image manipulations. Mohamed Abdelmaksoud et al. [39] developed a Vision Transformer (ViT) with SVM to detect forgery. ViT was employed for extracting the features by dividing images into uniform patches and linearly transformed as patch embeddings. The embeddings involved positional encodings to retain spatial data by utilizing a multi-layer perceptron and multi-head self-attention. SVM was used to detect and classify the forgery by learning a decision boundary that separates authentic and forged data depending on extracted features. Anjani Kumar Rai and Subodh Srivastava et al. [40] presented an Enhanced Center Mask Regional Convolutional Neural Network (ECMNet) for copy-move forgery localization. A MobileNetV2 and convolutional block attention module were used to capture fine-grained feature representation. In ECMNet, a fully convolutional one-stage object detector and spatial attention-guided mask head were utilized to detect and classify forged instances with the help of a bounding box. The reliance on bounding box-based localization limits the precise detection of uneven-shaped forged regions. Vanishri. V. Sataraddi and Nethravathi [41] suggested an attention-guided DL method to detect copy-move forgery with precise localization. The suggested method employs hierarchical convolution feature maps to capture representational saliency via spatial and channel attention mechanisms. The transformer block and the multi-attention mechanisms increase computational complexity and training time and do not generalize across various forgery types.

From the overall evaluation, it is obvious that the existing methods have the following shortcomings: They struggle in detecting fine-grained manipulations, improper classification of manipulated images due to subtle artifacts, difficulties while dealing with low-quality images or highly manipulated regions, struggles to adapt to varying spatial resolutions and complex image features, and challenges with images subjected to significant compression or noise. In order to address the aforementioned issues, this research proposes a GSA-DNN for multi-forgery detection and classification using spatial attention, with a focus on fine-grained information. This enables the model to detect subtle artifacts in manipulated regions and capture intricate features, thereby improving the overall detection performance.

3. Proposed Methodology

The GSA-DNN method is proposed for detecting multi-forgeries using CASIA V2, CoMoFoD, MISD, and CG-1050. During preprocessing, ELA is used to detect traces of manipulation, image resizing standardizes image dimensions, and min–max normalization scales pixel values into a consistent range. Then, Inception V3 is utilized to extract the features, while GSA-DNN detects and localizes the extracted features. Figure 2 depicts the overview process of the GSA-DNN method for multi-forgery detection.

Figure 2
Working process of the proposed GSA-DNN method for multi-forgery detection



3.1. Dataset

This research utilizes the CASIA V2, CoMoFoD, MISD, and CG-1050 datasets to determine the model's performance on multi-forgery detection, which are explained in detail below. The datasets are divided into training and testing sets using a stratified 80:20 split, where 80% of the data is utilized for training and 20% is used for testing. The stratified 80:20 split is performed using a fixed random seed (42) to ensure reproducibility. From the training data, a subset is employed for validation during training to analyze the model's performance.

CASIA V2: It is a commonly used dataset for image tampering detection research. It contains 7492 authentic images, with 5123 ground truth images that present pixel-level annotation. Additionally, it has 5123 tampered images exposed to manipulations such as splicing and copy-move. Figure 3 shows a representative sample image of the CASIA V2 dataset.

CoMoFoD: It is a copy-move forgery dataset that contains 10,400 images with a resolution of 512 × 512 pixels, organized into 200 sets of images. CoMoFoD dataset has employed five transformations on images like distortion, rotation, scaling, translation, and combination, each with 40 sets of images. The samples involve six kinds of post-processing attacks like intensity variations, blurring, contrast adjustment, compression, noise, and chrominance changes. Figure 4 denotes sample images from the CoMoFoD dataset, specifying the ground truth and tampered images.

Figure 3
Sample from the CASIA V2 dataset illustrating tampering and ground truth images

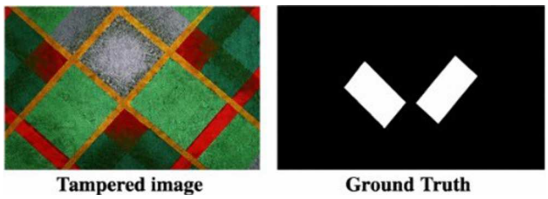
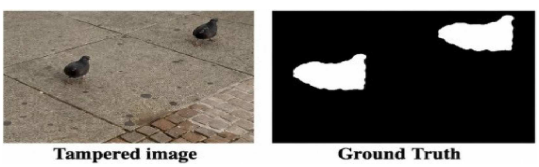
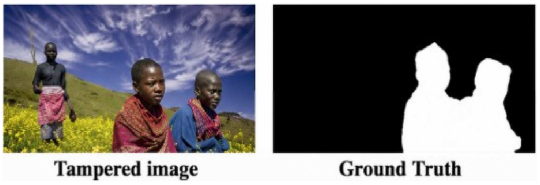


Figure 4
Sample from the CoMoFoD dataset representing a tampered and ground truth image



MISD: The MISD dataset contains a total of 620 authentic images, which act as an original unaltered set of images, consisting of 302 tampered and 731 ground truth masks with detailed annotation of the manipulated areas. The MISD dataset contains multiple types of forgeries such as copy-move and splicing. Figure 5 shows a sample image from the MISD dataset, indicating both the tampered and ground truth image.

Figure 5
Sample from the MISD dataset indicating tampered and ground truth image



CG-1050: It contains 2530 images for training and testing including 100 natural images, 1380 masks, and 1050 tampered images. Every image has undergone 10–11 manipulations, such as copy-move, cut-paste, colorizing, and retouching. The dataset is structured into four directories: original images, tampered images, mask images, and a description file. Table 1 shows a summary of forgery types involved in each benchmark dataset.

3.2. Preprocessing

After image acquisition, ELA, resizing, and min–max normalization are applied to detect inconsistencies, standardize dimensions, and scale pixel values for further processing. The CASIA V2 dataset lacks uniformity in image dimensions, which creates bias. Hence, a subset of images with less compression is chosen in this stage by utilizing ELA.

ELA: It is a normalized technique used to identify image regions with varying compression levels, indicating image tampering. ELA [42] is particularly effective for JPEG (.jpg) images because authentic regions have similar compression characteristics, whereas manipulated regions contain different compression levels. JPEG is a widely employed lossy image compression standard that minimizes image size by removing a portion of the image information while maintaining visual quality. Typically, JPEG employs a compression ratio of nearly 10:1, which provides a balance between image quality and compression efficiency. The JPEG compression process operates on images using 8 × 8 pixel block. Block sizes larger than 8 × 8 are less practical for standard JPEG compression, whereas smaller block sizes do not capture sufficient information for reliable compression. Hence, the ELA method employs 8 × 8 pixel block to preserve original JPEG compression characteristics and accurately determine inconsistency associated with digital tampering.

During JPEG compression, each 8 × 8 pixel block performs the same compression procedure, resulting in relatively uniform compression artifacts across authentic images. In an unaltered image, the error level is consistent when the image is recompressed. Compared with nearby regions, tampered regions have diverse compression levels, leading to inconsistent error levels. As a result, ELA recognizes these inconsistencies and detects the manipulated regions. In this research, the input image F is recompressed using a JPEG quality factor of 95% to obtain compressed

Table 1
Overview of image forgery types, authentic samples, ground truth masks, and tampered images across standard evaluation benchmark datasets

Datasets	Types of forgeries	Authentic	Ground truth	Tampered
CASIA V2	Copy-move and splicing	7492	5123	5123
CoMoFoD	Copy-move	2080	6240	2080
MISD	Splicing	620	731	302
CG-1050	Cut-paste, copy-move, colorizing, retouching	100	1380	1050

image F_{comp}. Then, the pixel-wise difference between the original and recompressed images is calculated using Equation (1):

$$F_{diff} = F - F_{comp} \tag{1}$$

where F_{diff} denotes the ELA difference image. Regions with higher error values correspond to inconsistent compression levels, representing image manipulation, whereas authentic regions denote relatively uniform error levels. Figure 6 represents the sample images from the CASIA V2, CoMoFoD, and MISD datasets after applying the ELA method. The generated ELA image is fed as input to the further process of the proposed framework.

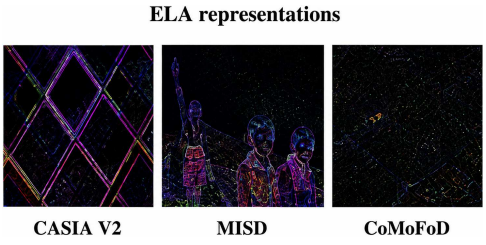
Resizing: The image is resized to 256 × 256 pixels to minimize complexity and ensure consistent performance. Also, resizing improves the model's ability to concentrate on tampering-related artifacts.

Min–max normalization: This method is used to scale pixel values to a range (0, 1), which ensures uniformity. Also, min–max normalization minimizes bias created by varying pixel intensities, and its mathematical formula [43] is expressed in Equation (2):

$$f(x, y) = \frac{f(x, y) - Z_{min}}{Z_{max} - Z_{min}} \tag{2}$$

where $f(x, y)$ denotes the normalized image and Z_{min} and Z_{max} refer to the minimum and maximum pixel values.

Figure 6
Sample images from CASIA V2, CoMoFoD, and MISD datasets applying the ELA method



3.3. Feature extraction

Inception V3 [44] is used to extract features and capture multiscale patterns by convolutional paths for detecting subtle image artifacts. ELA represents regions with varying error levels which shows potential image tampering, whereas InceptionNet V3 effectively extracts features by capturing low-level and high-level features. Inception V3 adopts a sparse structure while increasing the network's depth and width. Then, it converts similar and non-uniform sparse data into a dense representation to enhance accuracy. Figure 7 demonstrates the architecture of Inception V3 for feature extraction.

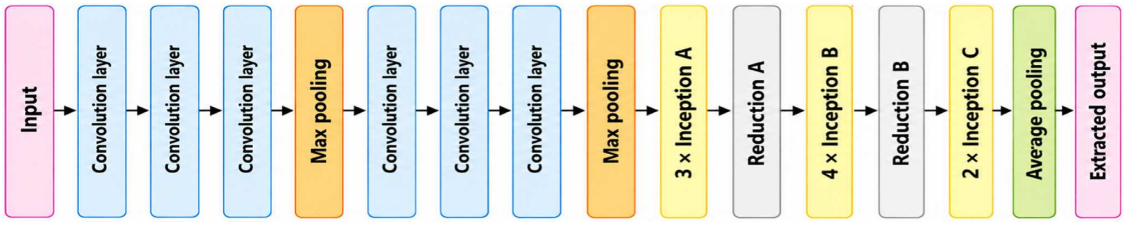
To perform the convolution process, a large number of filters are employed for extracting fine-grained features. Before the convolution process, nearby filters are utilized to reduce the dimensions that conserve appropriate information and reduce data loss. This process increases the model's depth and width by providing efficient resource utilization. Overall, Inception V3 extracts 2048 features that are passed to GSA-DNN to detect and classify the multi-forgeries.

3.4. Detection and classification

After feature extraction, GSA-DNN detects and classifies the forged and non-forged regions by focusing on appropriate features. This ensures to selectively weigh the significance of different spatial areas and enhance accuracy. Compared to CNN, a Dense Neural Network (DNN) is used because it learns intricate and high-level feature representations from different forgery patterns. CNN concentrates on local spatial features, whereas DNN effectively captures complex relationships among multiple manipulations, enhancing detection performance across images with varying textures. DNN layers contain input, hidden, and output layers [45], where hidden layers have a convolutional layer, pooling layer, and transposed convolutional layer.

Convolutional layer: It has a learnable set of filters with small receptive fields that are applied across the entire input for extracting local spatial features. The primary role is to map the input feature map by computing the dot product. The convolutional

Figure 7
Architectural flow of Inception V3 emphasizing multiscale feature learning through convolutional paths to extract the features



layer parameters include the convolution kernel size, padding, and stride. These three parameters determine the output feature map and are considered hyperparameters of the CNN. When a single convolutional kernel of size 2 × 2 is utilized, the output node value is calculated using Equation (3).

$$c_{ij} = f\left(\sum_{m=1}^{2}\sum_{k=1}^{2} w_{k,m}.x_{i+k-1,j+m-1}\right) \tag{3}$$

where $w_{k,m}$ represents the node value in the convolution kernel, $x_{i+k-1,j+m-1}$ indicates the input layer value associated with the convolution kernel, and f denotes the activation function.

Pooling layer: Also known as the downsampling layer, the pooling layer considers only the average and maximum values at the corresponding positions. It minimizes the number of parameters, enhances the computation speed, and avoids overfitting. If the pooling layer window size is 2 × 2 and the stride is 1, the output node value is determined using Equation (4).

$$P_{ij} = \max(c_{ij}, c_{i+1,j}, c_{i,j+1}, c_{i+1,j+1}) \tag{4}$$

Transposed convolution layer: The transposed convolution layer is a special forward convolution process that enhances the size of the input matrix by inserting zeros and then accomplishing forward convolution based on the convolution kernel. If the convolution kernel size is 2 × 2, the padding is 1, and the input array is 2 × 2, then the input size is computed as shown in Equation (5):

$$i_s = i + (s-1)(i-1) + p \tag{5}$$

where i, s, and p define input size, stride, and padding. The transposed convolution output size is computed using Equation (6):

$$o = s(i-1) + 2p - k + 2 \tag{6}$$

where k indicates the convolution kernel size.

GSA: The GSA mechanism splits feature representation into diverse channel groups and then learns a spatial attention map for each group instead of relying on a single attention map. This process ensures that the model concentrates on localized forgery and preserves subtle manipulations over spatial regions. Furthermore, the DNN employs these attention-refined feature representations to enhance feature discrimination and minimize data loss. Therefore, the GSA with dense learning enables the model to capture fine-grained spatial inconsistencies and contextual relationships that improve the multi-forgery detection performance.

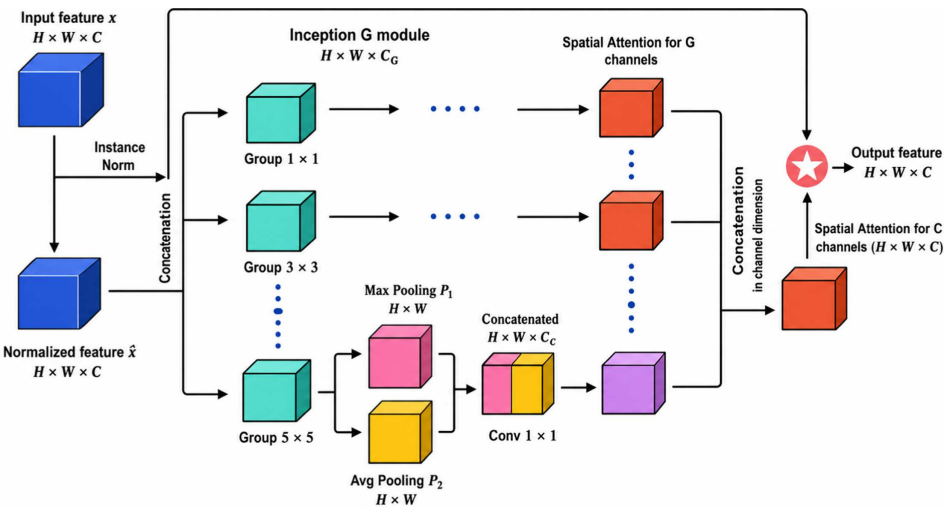
It divides feature representations into multiple channel groups and learns a spatial attention map for each group rather than relying on a single global attention map. This group-specific attention enables the network to selectively focus on localized forgery while preserving subtle manipulation artifacts across different spatial regions. Moreover, DNN effectively utilizes these attention-refined feature representations to improve feature discrimination and reduce information loss. Hence, the integration of GSA with dense feature learning enables the proposed GSA-DNN to simultaneously capture contextual relationships and fine-grained spatial inconsistencies, which enhances the detection of multiple forgery types under challenging imaging conditions.

Spatial attention plays a primary role in DNN with a focus on the significant convolutional feature spatial regions. Generally, these approaches involve multiplying the input feature dimension $H \times W \times C$ via the matrix of spatial attention $H \times W$. Meanwhile, the pathological patterns of various regions exhibit varying shapes and sizes, and so a uniform matrix of single-channel spatial attention is not suitable for all channels. The GSA processes each feature group separately, which makes the attention model focus on unique features and refines the convolutional features effectively. Figure 8 illustrates the proposed GSA model structure, where input features are split into groups, spatial attention is computed for each group, and the outputs are aggregated for enhanced representation.

Initially, an input feature of shape $H \times W \times C$ is normalized to $\tilde{x}$ using an instance normalization layer to rescale single-channel spatial components. Then, the normalized features $\tilde{x}$ are split into G groups, each containing $\frac{C}{G}$ channels. The i^{th} feature group is denoted as $\tilde{x}_1$, and the entire set of groups is indicated as $\tilde{x} = [\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_G]$. The group-wise features are squeezed with the channel dimension through the average pooling P_A and maximum pooling P_M layers. Traditional spatial attention mechanisms generate a single attention map across all feature channels, considering that every channel contributes equally to forgery detection. The forged regions have heterogeneous characteristics like texture variations, boundary inconsistencies, illumination changes, and

Figure 8

Architecture of the GSA mechanism, where the input features are partitioned into groups, spatial attention is determined for each group, and the outputs are aggregated for enhanced representation



compression artifacts, which are distributed across different feature channel and spatial locations. Using only a single spatial attention map minimizes discriminative features across different spatial regions. To solve this issue, the GSA mechanism is used, which splits the feature representation into diverse channel groups and learns spatial attention for each group independently. A 7×7 convolutional layer ($Conv_i$) is applied to the same pooled features, which generates the spatial attention matrix $A(\tilde{x}_1)$ for the i^{th} group using Equation (7).

$$A(\tilde{x}_1) = Conv_i([P_A(\tilde{x}_1), P_M(\tilde{x}_1)]) \tag{7}$$

Then, the G spatial attention matrices are concatenated along the channel dimension. The attention matrix for every G group is expressed using Equation (8).

$$A(\tilde{x}) = [A(\tilde{x}_1), A(\tilde{x}_2), \ldots, A(\tilde{x}_G)] \tag{8}$$

As every spatial attention matrix is associated with $\frac{C}{G} = 32$ channels, the resultant $A(\tilde{x})$ attention size matrix $H \times W \times C$ is extended with a channel dimension, which creates the last spatial attention map channel C. Finally, the output features are obtained by multiplying the spatial attention map channel C with the input feature x. The proposed GSA has the ability to learn group-specific spatial attention rather than relying on a single global attention representation. GSA enables various channel groups to focus on various suspicious image regions. This grouped attention strategy enhances feature discrimination, improves localization of small tampered regions, and preserves fine-grained manipulation traces that are overlooked through global attention mechanisms. The proposed model is most effective for identifying subtle copy-move and splicing forgeries under challenging imaging conditions. GSA-DNN improves feature representation by focusing on the most appropriate spatial regions, which assist in capturing fine-grained dependencies and subtle artifacts that enhance model performance.

Based on grid search, hyperparameter values are selected to ensure better performance. The cross-entropy loss function is used to reduce the difference between ground truth labels and predicted probabilities. Training is performed for 50 epochs with a batch size of 64 to achieve stable convergence by maintaining computational efficiency. The Adam optimizer is used with a learning rate of 0.0001 to offer adaptive parameter updates during training. The Softmax activation function is used in the output layer to convert the network output into normalized class probabilities, which enable accurate classification. Algorithm 1 shows a pseudo-code for the overall process.

Algorithm 1:

1. Input: CASIA V2, CoMoFoD, MISD, CG-1050 datasets
2. Data preprocessing:
   Start Loop
   For each image I in the dataset:
   Convert I to ELA format
   Resize I to 256 ×256 pixels
   Normalize pixel values of I to [0, 1] range
   End Loop
3. Feature Extraction:
   Start Loop
   For each preprocessed image:
   Apply preprocessed image to the Inception V3 model
   Extract deep features F from an intermediate layer
   End Loop
4. Classification:
   Start Loop
   For each feature vector F:
   Employ GSA mechanism
   Divide feature vector into groups
   Start Inner Loop
   For each group:
   Compute spatial attention scores
   Multiply attention scores with feature vector
   End Inner Loop
   Pass the attention-enhanced feature vector into DNN
   Input layer: Flatten feature vector
   Hidden layer: Fully Connected (FC) layer with Rectified Linear Unit activation function
   Output layer: SoftMax activation for classification
   End Loop
5. Train DNN using labelled data and binary-cross-entropy loss function
6. Output: Predicted labels indicating whether each image is forged or not

4. Experimental Results

The GSA-DNN is simulated on Python 3.10.12 with Anaconda Navigator 3.5.2.0 (64-bit). The experiments are evaluated on an Intel Core i5 processor, 8GB random-access memory, and TensorFlow, Keras, Transformers, and scikit-learn libraries. To validate model performance, the GSA-DNN method is evaluated on the CASIA V2, MISD, and CoMoFoD datasets. Also, the CG-1050 dataset is employed to analyze generalization ability on unseen data. Accuracy, F1-score, precision, and recall are used to measure the proposed method's performance, which are formulated in Equations (9)–(12). To ensure fair comparison, the baseline methods are reimplemented using identical settings like an 80:20 data split, preprocessing method, feature extraction, hyperparameters, and hardware/software environment. In Section 4.3, existing method values are directly adopted from respective papers.

The GSA-DNN method is evaluated based on primary and secondary validation. Primary validation is performed using a stratified 80:20 train–test split where 80% of the data is utilized for training and 20% is used for testing. The stratified 80:20 split is performed using a fixed random seed (42) to ensure reproducibility and consistency across experiments. From the training set, a subset is employed for validation during training to analyze the model's performance. The secondary validation is performed using 5-fold cross-validation on different test folds.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \tag{12}$$

where TP depicts true positive, FN demonstrates false negative, TN represents true positive, and FP indicates false positive.

4.1. Performance evaluation

Table 2 denotes a primary evaluation of different feature extraction methods for multi-forgery detection. The baseline techniques like VGG-16 [46], DenseNet 201 [47], and Xception [48] are compared with Inception V3 [44]. Although VGG-16, DenseNet201, and Xception references present improved methodologies, only the baseline methods are considered and implemented for comparison. When compared to these techniques, the Inception V3 obtains a high accuracy of 99.58%, 94.59%, 97.77%, and 97.45% using the CASIA V2, MISD, CoMoFoD, and CG-1050 datasets. Because the proposed Inception V3 employs multiple convolutional filter sizes within the same module, which allows it to capture features at various scales effectively.

CASIA V2 contains both copy-move and splicing forgeries, involving a combination of duplicated and composited image regions. CoMoFoD includes only copy-move forgeries, featuring duplicated parts within the same image. MISD has splicing forgeries, where elements from diverse images are combined to generate a tampered image. CG-1050 is a challenging dataset that has different forgery types like cut-paste, copy-move, colorizing, and retouching. The cut-paste operation introduces clear boundary inconsistencies to identify spatial artifacts. Copy-move forgeries assist to analyze the model's sensitivity to duplicated patterns within the same image. Colorizing involves subtle alterations to evaluate unnatural color transitions, whereas retouching involves fine-grained alterations to detect minimal edits that ensure high robustness. Overall, using these four operations represents the model's ability to solve forgery challenges. To analyze generalization, the CG-1050 dataset is utilized in this research.

The primary analysis of diverse classification methods is shown in Table 3. The baseline techniques like Artificial Neural Network (ANN) [49], ResNet-101 [50], and CNN [37] are

Table 2
Primary analysis of different feature extraction methods for multi-forgery detection on diverse datasets

Methods	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
VGG-16 [46]	CASIA V2	86.37	84.15	87.89	85.99
	MISD	80.56	78.12	81.98	79.95
	CoMoFoD	82.91	81.43	83.76	82.59
	CG-1050	88.34	89.12	88.76	88.94
DenseNet 201 [47]	CASIA V2	84.29	82.67	85.54	83.96
	MISD	79.22	77.54	80.35	78.42
	CoMoFoD	81.67	80.12	82.45	81.29
	CG-1050	90.67	91.21	90.13	90.67
Xception [48]	CASIA V2	81.67	80.12	82.45	81.29
	MISD	77.89	75.54	78.92	76.94
	CoMoFoD	80.12	78.34	81.19	79.76
	CG-1050	91.78	92.04	91.23	91.63
InceptionV3 [44]	CASIA V2	99.58	99.70	99.43	99.56
	MISD	94.59	95.08	92.43	93.60
	CoMoFoD	97.77	98.17	98.23	98.20
	CG-1050	97.45	97.47	97.45	97.45

Table 3
Primary analysis of different classification methods for multi-forgery detection on benchmark datasets

Methods	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
ANN [49]	CASIA V2	84.32	82.47	85.11	83.78
	MISD	78.56	76.22	79.91	77.82
	CoMoFoD	80.89	79.34	81.43	80.36
	CG-1050	88.34	89.12	88.76	88.94
ResNet-101 [50]	CASIA V2	87.21	85.12	88.44	86.77
	MISD	81.04	79.65	82.83	80.73
	CoMoFoD	83.79	82.34	84.47	83.39
	CG-1050	90.67	91.21	90.13	90.67
CNN [37]	CASIA V2	85.67	83.90	86.78	85.31
	MISD	79.91	77.54	80.67	78.96
	CoMoFoD	82.12	80.87	83.58	82.21
	CG-1050	91.78	92.04	91.23	91.63
GSA-DNN	CASIA V2	99.58	99.70	99.43	99.56
	MISD	94.59	95.08	92.43	93.60
	CoMoFoD	97.77	98.17	98.23	98.20
	CG-1050	97.45	97.47	97.45	97.45

compared with GSA-DNN. Although ResNet-101 and CNN references present improved methodologies, only the baseline methods are considered and implemented for comparison. GSA-DNN obtains an accuracy of 99.58%, 94.59%, 97.77%, and 97.45% on different datasets as it learns complex and high-level feature representation. Also, the deeper architecture enables efficient differentiation between authentic and manipulated regions, which improves the model's performance when compared to other methods.

The primary analysis of different attention mechanisms is presented in Table 4. Compared to baseline methods, the proposed GSA-DNN obtains a high accuracy of 99.58%, 94.59%, 97.77%, and 97.45% on different datasets as its feature representations are partitioned into groups, which assist in learning different features in each group. By employing an attention mechanism across diverse spatial and feature groups, GSA-DNN captures local and global inconsistencies. Later, DNN refines the extracted features to learn discriminative representations, which improve forgery detection performance.

Table 5 presents a primary evaluation of different computational complexities. Training time represents the total duration required for training the model, whereas inference time denotes the time taken to analyze a single image after the model is trained. The GSA-DNN reduces computational complexity through feature reuse via dense connectivity. Moreover, GSA-DNN is used to minimize unwanted computation by focusing on relevant spatial regions compared to existing methods. Unlike deeper methods like GSA-ResNet and GSA-CNN, the GSA-DNN avoids excessive parameterization that makes rapid processing possible. Also, its architecture balances between computational efficiency and

Table 4
Primary analysis of different attention mechanisms for multi-forgery detection on standard datasets

Methods	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Global Attention-DNN	CASIA V2	91.84	90.57	89.92	90.24
	MISD	85.42	84.77	86.33	85.54
	CoMoFoD	94.27	92.89	93.12	93.00
	CG-1050	90.67	89.58	90.01	89.79
Multi-Head Attention-DNN	CASIA V2	93.65	92.30	91.88	92.09
	MISD	89.23	88.45	87.61	88.02
	CoMoFoD	92.89	91.33	92.12	91.72
	CG-1050	92.81	91.75	92.08	91.91
Local Attention-DNN	CASIA V2	89.74	88.56	88.04	88.30
	MISD	83.41	84.19	82.65	83.41
	CoMoFoD	91.45	90.63	89.77	90.20
	CG-1050	88.45	87.90	88.11	88.00
GSA-DNN	CASIA V2	99.58	99.70	99.43	99.56
	MISD	94.59	95.08	92.43	93.60
	CoMoFoD	97.77	98.17	98.23	98.20
	CG-1050	97.45	97.47	97.45	97.45

Table 5
Primary analysis of computational complexity with baseline and proposed GSA-DNN method

Methods	Total images for training and testing	Datasets	Memory usage (MB)	Training time per epoch (s)	Inference time (s)
GSA-ANN	17738	CASIA V2	12.34	0.112	2.4831
	1653	MISD	9.85	0.034	0.8912
	10400	CoMoFoD	10.76	0.068	1.2054
	2530	CG-1050	11.24	0.091	2.0153
GSA-ResNet	17738	CASIA V2	15.91	0.145	2.7654
	1653	MISD	12.43	0.058	1.1029
	10400	CoMoFoD	13.77	0.081	1.3973
	2530	CG-1050	14.62	0.099	2.1426
GSA-CNN	17738	CASIA V2	10.52	0.096	2.0123
	1653	MISD	7.78	0.039	0.7591
	10400	CoMoFoD	9.12	0.063	1.0085
	2530	CG-1050	9.89	0.084	1.9674
GSA-DNN	17738	CASIA V2	7.98	0.080	1.7264
	1653	MISD	4.92	0.021	0.4998
	10400	CoMoFoD	6.11	0.056	0.8096
	2530	CG-1050	6.98	0.074	1.8506

accuracy, leading to better performance in multi-forgery detection. The eight dense layers are used for extracting and refining intricate features to detect image forgeries. Fewer dense layers such as 4–7 affect the model's ability to capture deeper patterns, while higher dense layers such as 12, result in overfitting and high complexity. Therefore, eight dense layers are used, which provides sufficient structural depth without increasing the complexity.

4.1.1. Epoch vs accuracy

The primary analysis of accuracy vs epoch for GSA-DNN is shown in Figure 9 on benchmark datasets. The training and validation accuracy increase rapidly near 0.0096 on the CASIA V2 dataset, representing efficient model learning. In the CoMoFoD and MISD datasets, accuracy increases rapidly to 0.9, which shows effective learning across epochs. Also, validation accuracy performs better than training accuracy in the CG-1050 dataset that enhances the model's generalization capability on unseen data. As a result, GSA-DNN achieves enhanced detection performance and better generalization on different datasets.

4.1.2. Epoch vs loss

Figure 10 demonstrates the primary evaluation of loss vs epochs for the GSA-DNN method using different datasets. For CASIA V2, the figure shows smooth convergence with minimal

Figure 9
Primary analysis of accuracy variations with respect to epochs for GSA-DNN on different datasets for multi-forgery detection

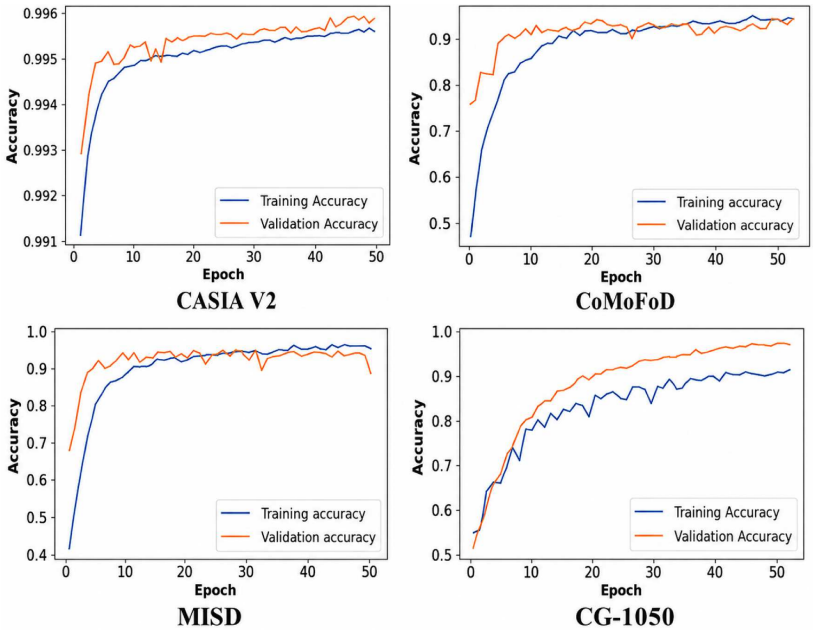
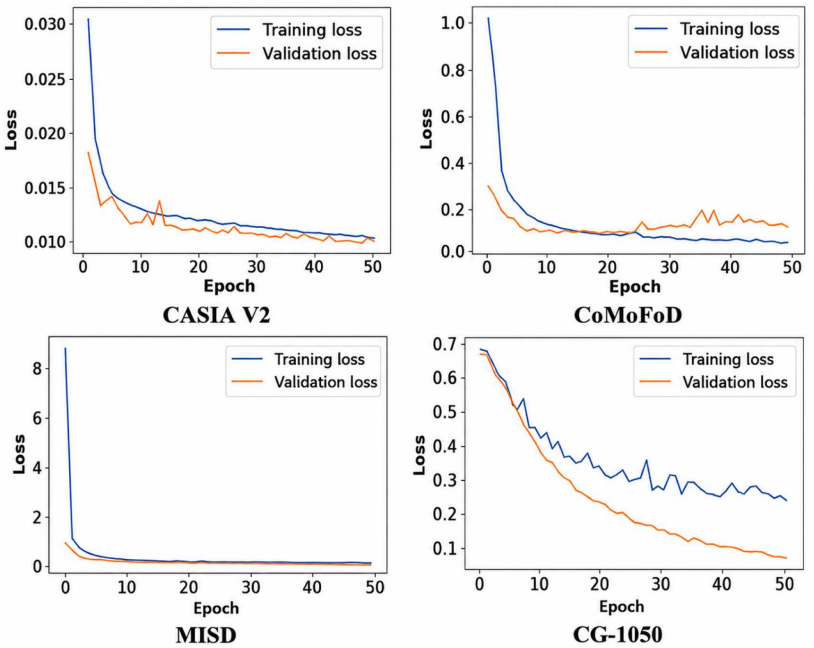


Figure 10
Primary analysis of loss variations with respect to epochs during training for GSA-DNN on benchmark datasets



gap between training and validation loss, which indicates better performance. With respect to validation loss, the CoMoFoD dataset exhibits certain fluctuations, whereas the MISD dataset begins with a higher initial loss, but validation and training losses stabilize effectively, indicating efficient model learning.

In CG-1050, the loss curve represents a consistent decrease in both validation and training loss as the number of epoch increases, which shows effective learning. Moreover, validation loss dropped lower than training loss, representing that the model performs efficiently on unseen data and avoids overfitting.

4.1.3. ROC curve

Figure 11 illustrates the primary evaluation of the receiver operating characteristic (ROC) curve for the proposed GSA-DNN. The ROC curve indicates the trade-off between true positive rate (TPR) and false positive rate (FPR). The CASIA V2 dataset achieves an area under the curve (AUC) of 1.00, representing better classification performance, whereas CoMoFoD obtains an AUC of 0.99, and the MISD dataset achieves 0.98. The ROC curve for CG-1050 shows a high TPR with an AUC of 0.97, representing better discrimination between forged and authentic images. These AUC values indicate that GSA-DNN performs effectively on different datasets.

4.1.4. Confusion matrix

Figure 12 represents a primary evaluation of the confusion matrix on different datasets for GSA-DNN. On CASIA V2, GSA-DNN correctly classified 1509 instances in class 1 and 383 instances in class 0, with 9 misclassifications (4 FP and 5 FN). For the MISD dataset, the classifier correctly predicted 50 instances of class 0 and 122 instances of class 1, with 13 misclassifications (9 FP and 4 FN). In the CoMoFoD dataset, the classifier obtained 1008 TN and 1026 TP with 46 errors (27 FP and 19 FN). The CG-1050 dataset indicates that GSA-DNN correctly classified 303 real and 309 fake images, with 5 FN and 11 FP, respectively. A low number of misclassifications reflects the model's precision and robustness, whereas a slight imbalance between FP and FN indicates fewer FN, which is critical in forgery detection. As a result, the proposed method obtains better performance with minor misclassifications. The FP occurred in images with intricate textures and repetitive structures that resembled manipulated regions, whereas FN occurred in images containing small tampered regions. These observations indicate that the proposed model remains sensitive to highly challenging forgery scenarios and provide directions for future improvement.

The confusion matrix reported in this research is calculated using the test set, which is separated from the training set. This ensures that the evaluation represents the model's generalization ability rather than memorization of the training samples. To solve the overfitting risk, especially to dataset-specific compression artifacts introduced by JPEG compression, the dataset is divided at the image level to avoid overlap of manipulation patterns between training and test sets. Moreover, the model is trained on samples with a range of compression levels and post-processing variations, which minimizes its reliance on fixed compression. The consistent performance achieved across multiple evaluation metrics shows that the model learns complex forgery-related features rather than superficial compression artifacts.

Based on the primary evaluation, robustness analysis is conducted for GSA-DNN against post-processing methods on different datasets in Table 6. Figure 13 depicts the robustness analysis of GSA-DNN against post-processing methods: (a) original images, (b) authentic and tampered images with Gaussian noise with σ = 20, (c) Gaussian blur with a kernel size of 20, and (d) JPEG compression with a quality level of 50, applied on the MISD dataset. In this analysis, Gaussian noise with σ = 20 was applied, as it introduced moderate graininess while keeping the image clearly recognizable. Lower levels, such as σ = 5, produced nearly invisible noise, while σ = 10 generated subtle noise that was

Figure 11
Primary evaluation of the ROC curve representing the trade-off between TPR and FPR for the GSA-DNN method

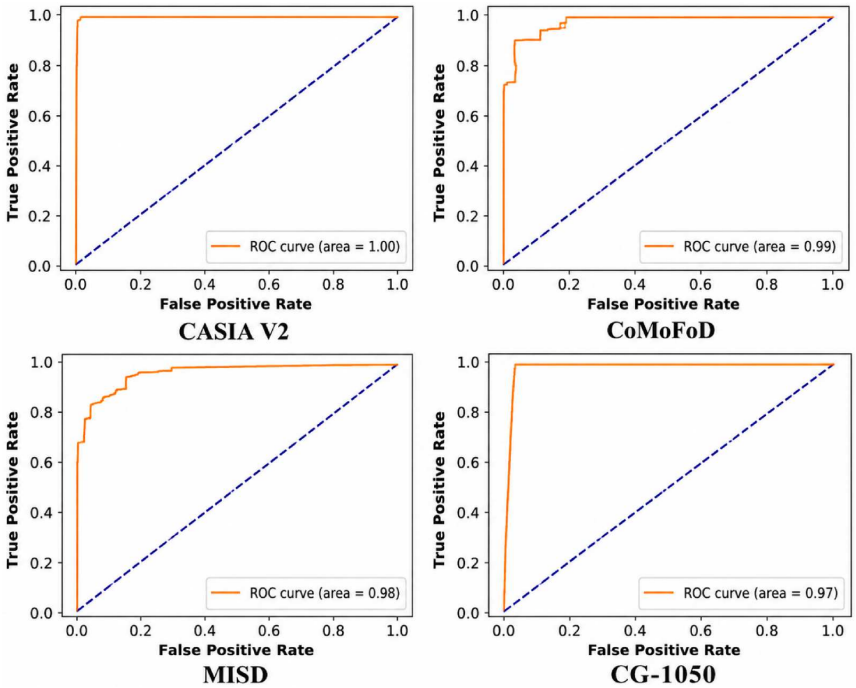


Figure 12
Primary evaluation of confusion matrix representing the classification performance across testing instances for CASIA V2 (1901), CoMoFoD (2080), MISD (185), and CG-1050 (628) datasets

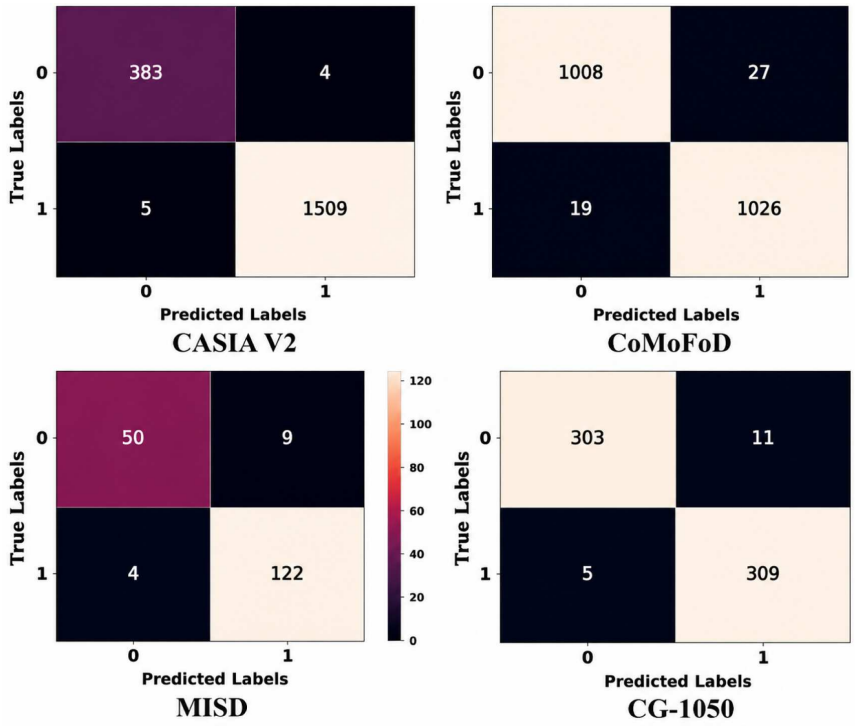


Table 6
Primary evaluation of robustness analysis against post-processing techniques on different datasets

Methods	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
GSA-DNN with Gaussian noise	CASIA V2	97.90	97.62	95.15	96.75
	MISD	93.81	92.25	90.30	91.17
	CoMoFoD	96.43	94.45	94.43	94.43
GSA-DNN with image blurring	CASIA V2	98.42	99.57	98.18	99.35
	MISD	92.85	93.18	91.72	92.51
	CoMoFoD	96.76	97.14	97.00	97.10
GSA-DNN with JPEG compression	CASIA V2	99.14	99.36	99.12	99.27
	MISD	93.75	94.00	92.90	93.35
	CoMoFoD	97.43	97.55	97.35	97.44
Proposed GSA-DNN	CASIA V2	99.58	99.70	99.43	99.56
	MISD	94.59	95.08	92.43	93.60
	CoMoFoD	97.77	98.17	98.23	98.20

barely noticeable, whereas higher noise levels like $\sigma = 40$ caused severe distortions that significantly impaired image recognizability. The JPEG quality was set to 50, which introduced moderate artifacts by conserving the overall structure with a compression ratio of nearly 33.7%. This provided appropriate degradation of the image for robustness analysis.

Higher quality levels produce minimal compression artifacts, whereas lower quality levels have degradation. Despite the presence of Gaussian noise, JPEG compression, and Gaussian blur, GSA-DNN achieves better performance in distinguishing between authentic and tampered images. Even in distorted conditions, the dense connectivity structure performs efficient feature reuse and improves the propagation of information across layers, contributing to better performance in multi-forgery detection.

Table 7 demonstrates the evaluation of different DL methods that are trained on the CASIA V2 dataset and tested on the MISD dataset. The cross-dataset validation is performed to analyze the generalization ability on unseen data. Although CASIA V2 involves compression artifacts, this research focused on distinguishing genuine areas from dataset-specific artifacts by doing cross-dataset validation. CASIA V2 is chosen for training as it contains a large number of images, whereas MISD is selected for testing as it contains fewer images, which shows effectiveness

Figure 13
Robustness analysis based on primary analysis of GSA-DNN against post-processing methods: (a) original images, (b) authentic and tampered images with Gaussian noise with σ = 20, (c) Gaussian blur with kernel size of 20, and (d) JPEG compression with quality level 50 on the MIST dataset

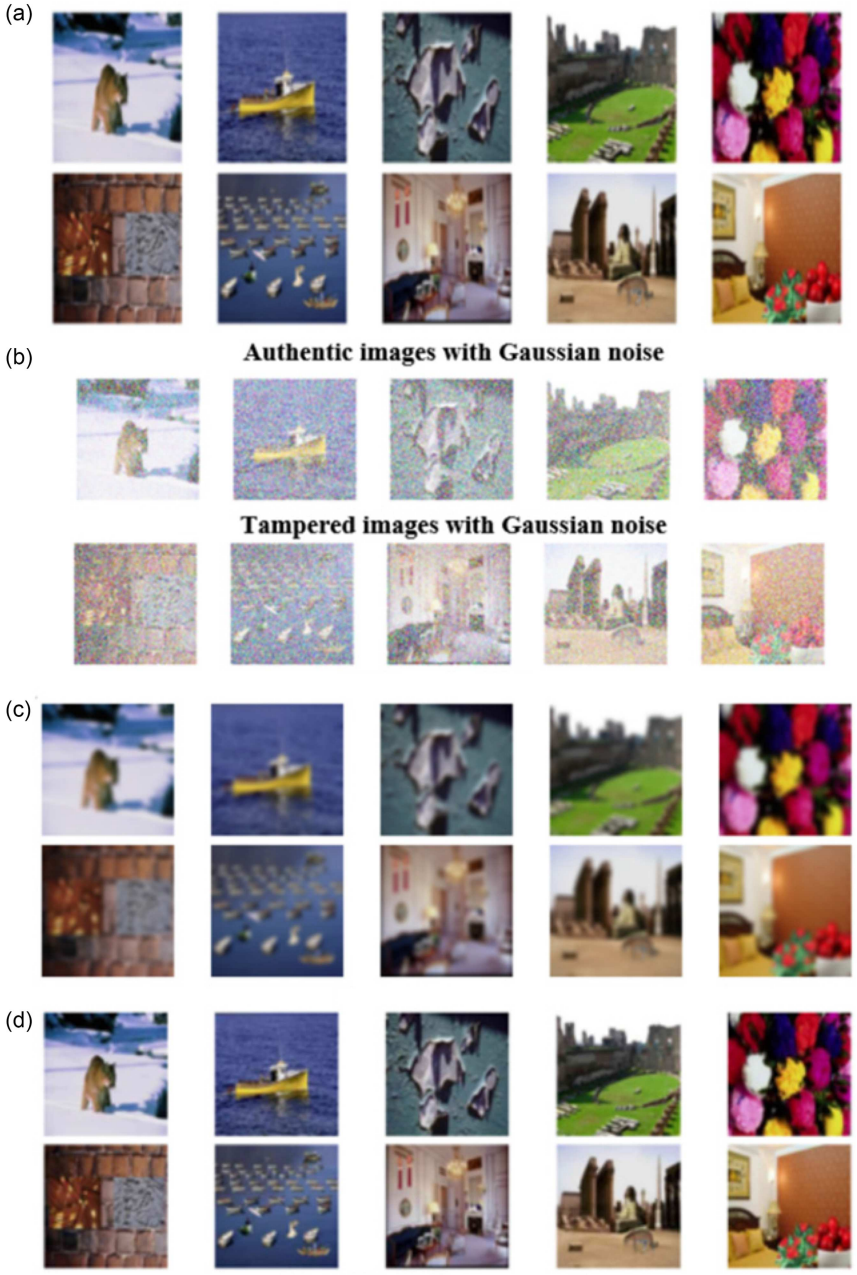


Table 7
Performance analysis of different DL methods trained on CASIA V2 and tested on the MISD dataset

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
GSA-ANN	74.25	73.80	72.10	72.94
GSA-ResNet	76.10	75.55	74.00	74.77
GSA-CNN	75.3	74.65	73.20	73.92
GSA-DNN	90.81	90.25	88.30	89.17

across varied data distributions. The outcomes show that the proposed GSA-DNN obtains better performance as it captures intricate and subtle forgeries compared to baseline methods like GSA-ANN, GSA-ResNet, and GSA-CNN. These findings confirmed that the proposed GSA-DNN generalized well from training data and proved to be highly suitable for image forgery detection tasks.

Table 8 shows the primary evaluation of statistical analysis for the proposed method using five independent runs with a fixed random seed (42). The 95% confidence interval (CI) quantifies the uncertainty related to the mean performance, with narrower intervals indicating greater precision as well as stability. Standard deviation measures the variability of metrics across five runs, which shows the consistency of the proposed GSA-DNN method. Together, these statistics represent the reliability of the proposed method in multi-forgery detection.

Table 8
Statistical analysis based on primary evaluation for GSA-DNN across all four datasets

Dataset	Accuracy mean (%)	Standard deviation	95% CI
CASIA V2	99.58	±0.21	[99.37, 99.79]
MISD	94.59	±0.48	[94.11, 95.07]
CoMoFoD	97.77	±0.33	[97.44, 98.10]
CG-1050	97.45	±0.41	[97.04, 97.86]

Figure 14 depicts visual examples of baseline methods and the proposed GSA-DNN using different datasets to identify subtle manipulations: (a) CASIA V2 sample—Copy-move + Hide Add manipulation, (b) CASIA V2 sample—Animal + Nature, (c) MISD sample—Nature + Architecture + Animal, (d) MISD sample—Copy-move + Hide Add manipulation, (e) CoMoFoD sample 1, and (f) CoMoFoD sample 2. Traditional methods such as ANN, CNN, and DNN face challenges in accurately detecting forgeries because of limited capability in capturing localized tampering artifacts. For example, in MISD sample 2, the ground truth mask (white region represents manipulated areas) and ELA output (manipulation region) show the difference between the manipulated region and detection result. Traditional methods determine overly broad tampered regions, which leads to inaccurate localization of fine-grained manipulation. ANN faces challenges in learning intricate hierarchical features, which leads to suboptimal performance. Due to a uniform convolution process across the entire image, CNN fails to identify subtle manipulations when the artifacts are very minimal. DNN does not focus on discriminative regions, leading to suboptimal performance and high FP rates on different datasets. In subtle manipulations, the proposed GSA-DNN effectively detects forgeries as it employs ELA, which suppresses the noise and enhances only the tampering inconsistencies and compression artifacts. GSA focuses on very relevant regions by grouping the spatial features of tampered areas, whereas DNN preserves fine-grained dependencies, facilitating accurate localization of tampered regions on different datasets.

4.2. Ablation study

Table 9 shows a primary analysis of the ablation study to evaluate the individual components of GSA-DNN. Compared to other variants, the combination of ELA + Inception V3 + GSA + DNN achieved high accuracies of 99.58%, 94.59%, 97.77%, and 97.45% on different datasets due to the complementary strengths of each component. ELA effectively highlighted subtle tampering artifacts in images, thereby improving feature visibility. Since the CASIA V2 dataset lacks uniformity in image quality, it introduced bias into the analysis.

To mitigate this, a subset of images with minimal compression was selected during the preprocessing stage using ELA, addressing concerns related to compression fingerprints. Inception V3 is used to extract multiscale features by capturing local and global patterns. GSA refined these extracted features by focusing on relevant spatial region that improves model performance. The GSA-DNN incorporates GSA-guided features into DNN, which allows it to leverage enriched feature representation for classification and obtain better performance on different datasets.

Table 10 shows secondary analysis of 5-fold cross-validation on different test folds for detecting multi-forgeries. In this analysis, each dataset is partitioned into five equal folds. During each iteration, one fold is employed as the test set, whereas the remaining four folds are used for training. This process is repeated five times so that each fold acts as the test set once. The GSA-DNN performance is analyzed for each fold, and the mean values across five folds are computed to evaluate the generalization on unseen data.

4.3. Comparative analysis

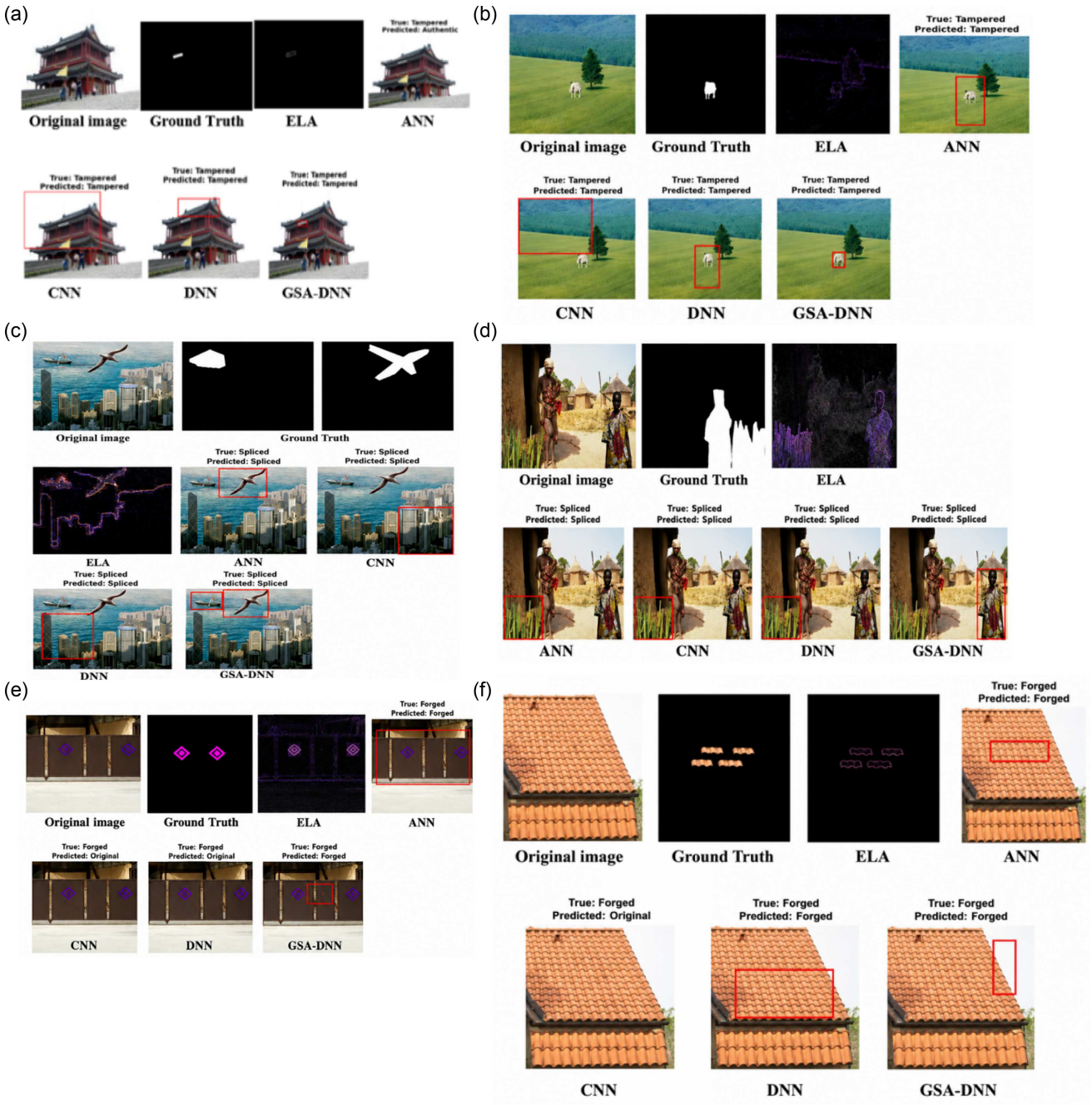
Tables 11, 12, and 13 illustrate a comparative analysis of existing methods on the CASIA V2, CoMoFoD, and MISD datasets. Compared to existing methods, GSA-DNN achieves high accuracy as the employed GSA mechanism enables the model to concentrate on localized forgery patterns while minimizing irrelevant background information. Moreover, DNN employs attention-refined features to enhance feature discrimination and preserve subtle manipulation patterns. Hence, the proposed GSA-DNN method detects multi-forgery types more accurately than existing methods.

4.4. Discussion

The advantages of the GSA-DNN method and the limitations of existing approaches are briefly discussed in this section. Existing methods such as the Siamese network [23] struggled to localize fine-grained manipulations due to the architecture's emphasis on feature similarity. Mask RCNN [24] showed limitations in accurately detecting small or subtle multi-forged regions. ResNet50v2 [25] exhibited performance degradation due to processing on low-quality images, which limits model effectiveness in forgery detection. The proposed GSA-DNN solves these issues by deploying GSA to concentrate on relevant spatial regions in an image. GSA-DNN captures fine-grained pixel-level information, which facilitates learning of subtle manipulations like texture variations, boundary inconsistencies, and compression artifacts that improve multi-forgery detection performance. Moreover, GSA-DNN identifies multi-forged regions by refining spatial attention and maintaining better performance even on low-quality images.

Although CASIA V2 has known compression artifacts, this research focuses on differentiating genuine forgery detection from dataset-specific artifacts through cross-dataset validation and statistical analysis. A *p*-test is evaluated to determine whether the observed difference in accuracy is statistically significant,

Figure 14
Visual instance of traditional methods and GSA-DNN based on primary analysis on different datasets to identify subtle manipulations: (a) CASIA V2 sample 1: Copy-move + Hide Add manipulation, (b) CASIA V2 sample 2: Animal + Nature, (c) MISD sample 1: Nature + Architecture + Animal, (d) MISD sample 2: Copy-move + Hide Add manipulation, (e) CoMoFoD sample 1, and (f) CoMoFoD sample 2



which shows the reliability of GSA-DNN. Compared to other methods, the statistical analysis shows that DNN achieves better performance as indicated by the lowest *p*-test values for different datasets. For example, DNN obtains the lowest *p*-test value with ELA of 0.020 for CASIA V2, 0.028 for MISD, 0.023 for CoMoFoD, and 0.024 for CG-1050 datasets. Without ELA, *p*-test value was enhanced to 0.022, 0.030, 0.025, and 0.026 for the respective datasets. Similar improvements are observed for ANN, ResNet, and CNN, where the inclusion of *p*-values improves detection performance. Moreover, the ablation study shows that each variant in the proposed GSA-DNN contributes to the final performance. As a result, experimental outcomes show that GSA-DNN achieves better performance compared to existing methods on different datasets.

Table 9
Primary analysis of the ablation study to evaluate the individual variant of the proposed GSA-DNN on different datasets

Methods	Datasets	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
ELA + Inception V3 + DNN	CASIA V2	98.31	98.54	97.89	98.21
	MISD	91.76	92.23	89.87	91.03
	CoMoFoD	95.83	96.2	96.12	96.16
	CG-1050	95.28	95.33	95.3	95.31
ELA + GSA + DNN	CASIA V2	97.62	97.85	96.98	97.41
	MISD	90.2	91.02	87.55	89.25
	CoMoFoD	94.35	94.9	94.75	94.82
	CG-1050	94.12	94.25	94.1	94.17
ELA + DNN	CASIA V2	96.8	96.91	95.87	96.39
	MISD	88.65	89.12	86.21	87.64
	CoMoFoD	92.98	93.43	93.3	93.36
	CG-1050	93.05	93.12	93	93.06
Inception V3 + GSA + DNN	CASIA V2	98.95	99.1	98.43	98.76
	MISD	92.84	93.22	90.76	91.97
	CoMoFoD	96.42	96.87	96.92	96.89
	CG-1050	96.07	96.15	96.1	96.12
Proposed method (ELA + Inception V3 + GSA + DNN)	CASIA V2	99.58	99.70	99.43	99.56
	MISD	94.59	95.08	92.43	93.60
	CoMoFoD	97.77	98.17	98.23	98.20
	CG-1050	97.45	97.47	97.45	97.45

Table 10
Secondary analysis of 5-fold cross-validation on different test folds for detecting multi-forgeries

Datasets	Folds (k = 5)	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CASIA V2	1	99.23	99.37	98.91	99.14
	2	99.67	99.73	99.43	99.58
	3	99.49	99.62	99.21	99.41
	4	99.74	99.82	99.53	99.67
	5	99.38	99.51	99.12	99.31
	Mean	99.50 ± 0.208	99.61 ± 0.177	99.24 ± 0.247	99.42 ± 0.211
MISD	1	93.73	94.11	91.83	92.96
	2	94.27	94.63	92.51	93.55
	3	94.89	95.22	93.31	94.26
	4	95.13	95.41	93.62	94.51
	5	94.04	94.32	92.23	93.25
	Mean	94.41 ± 0.451	94.73 ± 0.445	92.70 ± 0.575	93.70 ± 0.518
CoMoFoD	1	97.13	97.53	97.41	97.47
	2	97.87	98.12	98.03	98.07
	3	98.03	98.32	98.11	98.21
	4	97.62	97.92	98.01	97.96
	5	97.29	97.63	97.72	97.67
	Mean	97.58 ± 0.378	97.90 ± 0.329	97.85 ± 0.289	97.87 ± 0.301
CG-1050	1	96.91	97.02	96.83	96.92
	2	97.32	97.43	97.21	97.31
	3	97.77	97.82	97.63	97.72
	4	97.43	97.52	97.31	97.41
	5	97.06	97.11	96.92	97.02
	Mean	97.29 ± 0.259	97.38 ± 0.256	97.18 ± 0.256	97.27 ± 0.253

Table 11
Comparative evaluation of GSA-DNN and existing methods on the CASIA V2 dataset for multi-forgery detection

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Siamese neural network [23]	98.6	97.5	95	96
RCNN [24]	N/A	83.41	86.02	84.69
ResNet50v2 [25]	99.3	N/A	N/A	N/A
ViT + SVM [39]	98.92	N/A	N/A	N/A
Attention-guided DL method [41]	N/A	95	92	93
Proposed GSA-DNN	99.58	99.70	99.43	99.56

Table 12
Comparative evaluation of GSA-DNN and existing methods on the CoMoFoD dataset for multi-forgery detection

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
ResNet50v2 [25]	N/A	98.12	95.85	96.97
GCN [26]	78	N/A	N/A	78
LTrP [27]	N/A	96.33	91.81	93.46
CNN [28]	93.82	97.94	95.20	96.55
Attention-guided DL method [41]	N/A	95	93	94
Proposed GSA-DNN	97.77	98.17	98.23	98.20

Table 13
Comparative evaluation of GSA-DNN and existing methods on the MISD dataset for multi-forgery detection

Methods	Precision (%)	Recall (%)	F1-score (%)
Mask Region CNN [29]	73	62	67
Proposed GSA-DNN	95.08	92.43	93.60

5. Conclusion

The GSA-DNN is proposed to detect multi-forgeries using different datasets. GSA-DNN focuses on the most informative spatial regions, which enhance discriminative feature representation and detection performance. Inception V3 is used to capture intricate and fine-grained dependencies through multiscale convolution filters. ELA method helps to identify inconsistencies, and resizing standardizes the image dimension. Furthermore, min–max normalization scales the pixel intensities to a fixed range, which minimizes variations in feature magnitudes, enhances model convergence, and stabilizes gradient updates during training. Therefore, the GSA-DNN achieves a high accuracy of 99.58%, 97.77%, and 94.59% on the CASIA V2, CoMoFoD, and MISD datasets compared to existing methods such as Siamese neural network, ResNet50v2, and Mask RCNN. Also, GSA-DNN obtains high accuracy of 97.45% on the CG-1050 dataset for generalization analysis on unseen data compared to baseline methods. In future work, different DL methods and diverse datasets will be used to detect forgery, further enhancing the model's generalization capability and performance.

Acknowledgment

The authors gratefully acknowledge the assistance of Clue4Evidence Forensic Lab for their valuable contributions to this research, especially in offering expert technical support. Their expertise and resources greatly enhanced the quality and accuracy of the study.

Ethical Statement

The authors declare that this study did not require formal ethical approval because PES University, India, does not require Institutional Review Board or ethics committee approval for research conducted exclusively on publicly available, de-identified benchmark image datasets that involve no recruitment of, intervention on, or interaction with human participants. This study used only the publicly released CASIA V2, CoMoFoD, MISD, and CG-1050 datasets, which were collected and distributed by their original creators under their respective terms of use; no new data were collected from human participants by the authors. Any human faces or persons visible in the sample images reproduced in this manuscript originate from these public datasets and are shown solely for the purpose of illustrating forgery detection results. This exemption is based on the Institutional Research Ethics Policy of PES University, read with the National Ethical Guidelines for Biomedical and Health Research Involving Human Participants (ICMR, 2017), which exempt research on anonymous, publicly available data from ethics committee review.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in CASIA V2 at https://doi.org/10.3390/app14135545; in CoMoFoD at https://www.kaggle.com/datasets/tusharchauhan1898/comofod; in MISD at https://zenodo.org/records/5525829 and in CG-1050 at https://www.kaggle.com/datasets/saurabhshahane/cg1050.

Author Contribution Statement

S. Aruna: Conceptualization, Methodology, Formal analysis, Resources, Data curation, Writing – original draft. **Surabhi Narayan:** Software, Validation, Investigation, Writing – review & editing, Visualization, Supervision, Project administration.

References

[1] Rafique, R., Gantassi, R., Amin, R., Frnda, J., Mustapha, A., & Alshehri, A. H. (2023). Deep fake detection and classification using error-level analysis and deep learning. *Scientific Reports*, *13*(1), 7422. https://doi.org/10.1038/s41598-023-34629-3

[2] Alhaidery, M. M. A., Taherinia, A. H., & Shahadi, H. I. (2023). A robust detection and localization technique for copy-move forgery in digital images. *Journal of King Saud University Computer and Information Sciences*, *35*(1), 449–461. https://doi.org/10.1016/j.jksuci.2022.12.014

[3] Walia, S., Kumar, K., Agarwal, S., & Kim, H. (2022). Using XAI for deep learning-based image manipulation detection with Shapley additive explanation. *Symmetry*, *14*(8), 1–14. https://doi.org/10.3390/sym14081611

[4] Chen, H., Chang, C., Shi, Z., & Lyu, Y. (2022). Hybrid features and semantic reinforcement network for image forgery detection. *Multimedia Systems*, *28*(2), 363–374. https://doi.org/10.1007/s00530-021-00801-w

[5] Koul, S., Kumar, M., Khurana, S. S., Mushtaq, F., & Kumar, K. (2022). An efficient approach for copy-move image forgery detection using convolution neural network. *Multimedia Tools and Applications*, *81*(8), 11259–11277. https://doi.org/10.1007/s11042-022-11974-5

[6] Yue, G., Duan, Q., Liu, R., Peng, W., Liao, Y., & Liu, J. (2022). SMDAF: A novel keypoint based method for copy–move forgery detection. *IET Image Processing*, *16*(13), 3589–3602. https://doi.org/10.1049/ipr2.12578

[7] Fernandez-Castillo, E., Barbosa-Santillán, L. I., Falcon-Morales, L., & Sánchez-Escobar, J. J. (2022). Deep splicer: A CNN model for splice site prediction in genetic sequences. *Genes*, *13*(5), 907. https://doi.org/10.3390/genes13050907

[8] Lin, X., Wang, S., Deng, J., Fu, Y., Bai, X., Chen, X., . . . , & Tang, W. (2023). Image manipulation detection by multiple tampering traces and edge artifact enhancement. *Pattern Recognition*, *133*, 109026. https://doi.org/10.1016/j.patcog.2022.109026

[9] Wei, Y., Ma, J., Wang, Z., Xiao, B., & Zheng, W. (2022). Image splicing forgery detection by combining synthetic adversarial networks and hybrid dense U-net based on multiple spaces. *International Journal of Intelligent Systems*, *37*(11), 8291–8308. https://doi.org/10.1002/int.22939

[10] Tallapragada, V. S., Reddy, D. V., & Kumar, G. P. (2024). Blind forgery detection using enhanced mask-region convolutional neural network. *Multimedia Tools and Applications*, *83*(37), 84975–84998. https://doi.org/10.1007/s11042-024-19347-w

[11] Sun, Y., Ni, R., & Zhao, Y. (2022). ET: Edge-enhanced transformer for image splicing detection. *IEEE Signal Processing Letters*, *29*, 1232–1236. https://doi.org/10.1109/LSP.2022.3172617

[12] Huang, Y., Bian, S., Li, H., Wang, C., & Li, K. (2022). DS-UNet: A dual streams UNet for refined image forgery localization. *Information Sciences*, *610*, 73–89. https://doi.org/10.1016/j.ins.2022.08.005

[13] Gu, A. R., Nam, J. H., & Lee, S. C. (2022). FBI-Net: Frequency-based image forgery localization via multitask learning with self-attention. *IEEE Access*, *10*, 62751–62762. https://doi.org/10.1109/ACCESS.2022.3182024

[14] Liu, X., Liu, Y., Chen, J., & Liu, X. (2022). PSCC-Net: Progressive spatio-channel correlation network for image manipulation detection and localization. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(11), 7505–7517. https://doi.org/10.1109/TCSVT.2022.3189545

[15] Wu, H., Zhou, J., Tian, J., Liu, J., & Qiao, Y. (2022). Robust image forgery detection against transmission over online social networks. *IEEE Transactions on Information Forensics and Security*, *17*, 443–456. https://doi.org/10.1109/TIFS.2022.3144878

[16] Akram, A., Jaffar, M. A., Rashid, J., Boulaaras, S. M., & Faheem, M. (2025). CMV2U-net: Au-shaped network with edge-weighted features for detecting and localizing image splicing. *Journal of Forensic Sciences*, *70*(3), 1026–1043. https://doi.org/10.1111/1556-4029.70033

[17] Eltoukhy, M. M., Alsubaei, F. S., Mortda, A. M., & Hosny, K. M. (2025). An efficient convolution neural network method for copy-move video forgery detection. *Alexandria Engineering Journal*, *110*, 429–437. https://doi.org/10.1016/j.aej.2024.10.030

[18] Ravula, J., & Singh, N. (2025). GAN-ViT-CMFD: A novel framework integrating generative adversarial networks and vision transformers for enhanced copy-move forgery detection and classification with spectral clustering. *Intelligent Systems with Applications*, *26*, 200524. https://doi.org/10.1016/j.iswa.2025.200524

[19] Alfraihi, H., Alzaidi, M. S. A., Alqahtani, H., Darem, A. A., Al-Sharafi, A. M., Alzahrani, A. A., . . . , & Alkharashi, A. (2025). A multi-model feature fusion based transfer learning with heuristic search for copy-move video forgery detection. *Scientific Reports*, *15*(1), 4738. https://doi.org/10.1038/s41598-025-88592-2

[20] Üstübıoğlu, A. (2025). Robust audio copy-move forgery detection using deep keypoint features on high-resolution texture images. *IEEE Access*, *13*, 202228–202252. https://doi.org/10.1109/ACCESS.2025.3636816

[21] Basavarajappa, S., & Sarapady Venkatramanayya, S. (2026). Biorthogonal wavelet transform-based adaptive segmentation and statistical keypoint matching approach for image forgery detection. *ETRI Journal*, *48*(3), 513–531. https://doi.org/10.4218/etrij.2025-0025

[22] Chang, J., Yang, J., Gan, Z., & Guo, L. (2026). Multibranch collaboration and segmented training network for image forgery comprehensive detection. *IET Biometrics*, *2026*(1), 1–16. https://doi.org/10.1049/bme2/1389267

[23] Ramirez-Rodriguez, A. E., Arevalo-Ancona, R. E., Perez-Meana, H., Cedillo-Hernandez, M., & Nakano-Miyatake, M. (2024). AISMSNet: Advanced image splicing manipulation identification based on Siamese networks. *Applied Sciences*, *14*(13), 5545. https://doi.org/10.3390/app14135545

[24] Nazir, T., Nawaz, M., Masood, M., & Javed, A. (2022). Copy move forgery detection and segmentation using improved mask region-based convolution network (RCNN). *Applied Soft Computing*, *131*, 109778. https://doi.org/10.1016/j.asoc.2022.109778

[25] Qazi, E. U. H., Zia, T., & Almorjan, A. (2022). Deep learning-based digital image forgery detection system. *Applied Sciences*, *12*(6), 2851. https://doi.org/10.3390/app12062851

[26] Shinde, V., Dhanawat, V., Almogren, A., Biswas, A., Bilal, M., Naqvi, R. A., & Rehman, A. U. (2024). Copy-move forgery detection technique using graph convolutional networks feature extraction. *IEEE Access*, *12*, 121675–121687. https://doi.org/10.1109/ACCESS.2024.3452609

[27] Ganguly, S., Mandal, S., Malakar, S., & Sarkar, R. (2023). Copy-move forgery detection using local tetra pattern based texture descriptor. *Multimedia Tools and Applications*, *82*(13), 19621–19642. https://doi.org/10.1007/s11042-022-14287-9

[28] Jaiswal, A. K., & Srivastava, R. (2022). Detection of copy-move forgery in digital image using multi-scale, multi-stage deep learning model. *Neural Processing Letters*, *54*(1), 75–100. https://doi.org/10.1007/s11063-021-10620-9

[29] Kadam, K., Ahirrao, S., Kotecha, K., & Sahu, S. (2021). Detection and localization of multiple image splicing using MobileNet V1. *IEEE Access*, *9*, 162499–162519. https://doi.org/10.1109/ACCESS.2021.3130342

[30] Jalab, H. A., Alqarni, M. A., Ibrahim, R. W., & Almazroi, A. A. (2022). A novel pixel's fractional mean-based image enhancement algorithm for better image splicing detection. *Journal of King Saud University-Science*, *34*(2), 101805. https://doi.org/10.1016/j.jksus.2021.101805

[31] Ding, H., Chen, L., Tao, Q., Fu, Z., Dong, L., & Cui, X. (2023). DCU-Net: A dual-channel U-shaped network for image splicing forgery detection. *Neural Computing and Applications*, *35*(7), 5015–5031. https://doi.org/10.1007/s00521-021-06329-4

[32] El Biach, F. Z., Iala, I., Laanaya, H., & Minaoui, K. (2022). Encoder-decoder based convolutional neural networks for image forgery detection. *Multimedia Tools and Applications*, *81*(16), 22611–22628. https://doi.org/10.1007/s11042-020-10158-3

[33] Walia, S., Kumar, K., & Kumar, M. (2023). Unveiling digital image forgeries using Markov based quaternions in frequency domain and fusion of machine learning algorithms. *Multimedia Tools and Applications*, *82*(3), 4517–4532. https://doi.org/10.1007/s11042-022-13610-8

[34] Sabeena, M., & Abraham, L. (2024). Convolutional block attention based network for copy-move image forgery detection. *Multimedia Tools and Applications*, *83*(1), 2383–2405. https://doi.org/10.1007/s11042-023-15649-7

[35] Rao, Y., Ni, J., Zhang, W., & Huang, J. (2022). Towards JPEG-resistant image forgery detection and localization via self-supervised domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *47*(5), 3285–3297. https://doi.org/10.1109/TPAMI.2022.3210379

[36] Diwan, A., Mahadeva, R., & Gupta, V. (2024). Advancing copy-move manipulation detection in complex image scenarios through multiscale detector. *IEEE Access*, *12*, 64736–64753. https://doi.org/10.1109/ACCESS.2024.3397466

[37] Diwan, A., & Roy, A. K. (2024). Cnn-keypoint based two-stage hybrid approach for copy-move forgery detection. *IEEE Access*, *12*, 43809–43826. https://doi.org/10.1109/ACCESS.2024.3380460

[38] Diwan, A., & Sonkar, U. (2024). Visualizing the truth: A survey of multimedia forensic analysis. *Multimedia Tools and Applications*, *83*(16), 47979–48006. https://doi.org/10.1007/s11042-023-17475-3

[39] Abdelmaksoud, M., Youssef, B., Wassif, K., & El-Khoribi, R. A. (2025). Hybrid framework for image forgery detection and robustness against adversarial attacks using vision transformer and SVM. *Scientific Reports*, *15*(1), 40371. https://doi.org/10.1038/s41598-025-25436-z

[40] Rai, A. K., & Srivastava, S. (2025). ECMNet: A deep framework for copy move forgery localization detection and classification. *Results in Engineering*, *29*, 1–18. https://doi.org/10.1016/j.rineng.2025.108627

[41] Sataraddi, V. V., & Nethravathi, N. P. (2026). An attention-enhanced multi-scale framework for copy–move forgery detection and localization. *Engineering, Technology & Applied Science Research*, *16*(3), 34927–34933. https://doi.org/10.48084/etasr.18383

[42] Khalil, A. H., Ghalwash, A. Z., Elsayed, H. A. G., Salama, G. I., & Ghalwash, H. A. (2023). Enhancing digital image forgery detection using transfer learning. *IEEE Access*, *11*, 91583–91594. https://doi.org/10.1109/ACCESS.2023.3307357

[43] Nijaguna, G. S., Babu, J. A., Parameshachari, B. D., de Prado, R. P., & Frnda, J. (2023). Quantum fruit fly algorithm and ResNet50-VGG16 for medical diagnosis. *Applied Soft Computing*, *136*, 110055. https://doi.org/10.1016/j.asoc.2023.110055

[44] Meena, G., Mohbey, K. K., & Kumar, S. (2023). Sentiment analysis on images using convolutional neural networks based Inception-V3 transfer learning approach. *International Journal of Information Management Data Insights*, *3*(1), 1–13. https://doi.org/10.1016/j.jjimei.2023.100174

[45] Huang, K., Li, S., Deng, W., Yu, Z., & Ma, L. (2022). Structure inference of networked system with the synergy of deep residual network and fully connected layer network. *Neural Networks*, *145*, 288–299. https://doi.org/10.1016/j.neunet.2021.10.016

[46] Bevinamarad, P., & Unki, P. (2026). ForgNetwork: An adaptive forgery detection and localization framework using stationary wavelet transform and dilated adaptive VGG16 with optimization strategy. *Information Security Journal: A Global Perspective*, *35*(2), 330–350. https://doi.org/10.1080/19393555.2025.2570170

[47] Pandey, A., Lal, N., & Chakraverti, A. K. (2025). Hybrid dense net 201 with CBPN based image pixel enhancement and optimized ADBGAN for image copy-move forgery detection. *Signal, Image and Video Processing*, *19*(6), 467. https://doi.org/10.1007/s11760-025-03972-5

[48] Ganguly, S., Ganguly, A., Mohiuddin, S., Malakar, S., & Sarkar, R. (2022). ViXNet: Vision transformer with Xception network for deepfakes based video and image forgery detection. *Expert Systems with Applications*, *210*, 118423. https://doi.org/10.1016/j.eswa.2022.118423

[49] Nath, S., & Naskar, R. (2021). Automated image splicing detection using deep CNN-learned features and ANN-based classifier. *Signal, Image and Video Processing*, *15*(7), 1601–1608. https://doi.org/10.1007/s11760-021-01895-5

[50] Vaishali, S., & Neetu, S. (2024). Enhanced copy-move forgery detection using deep convolutional neural network (DCNN) employing the ResNet-101 transfer learning model. *Multimedia Tools and Applications*, *83*(4), 10839–10863. https://doi.org/10.1007/s11042-023-15724-z

How to Cite: Aruna, S., & Narayan, S. (2026). Groupwise-Spatial Attention-Based Dense Neural Network for Image Multi-Forgery Detection. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA620210360