ISSN: 2093-5374 / E-ISSN: 2093-5382

Latent Fingerprint Enhancement and Segmentation Through Advanced Deep-Learning Techniques

Poornima E Gundgurti[1*], and Dr. Shrinivasrao B Kulkarni[2]

[1*] Department of Computer Science, Central University of Karnataka, Kadaganchi, Kalaburagi, Karnataka and Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India. poornimae25@gmail.com, https://orcid.org/0009-0009-1171-546X

[2]Department of Computer Science and Engineering, SDM College of Engineering and Technology, Dharwad, Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India. sbkulkarni_in@yahoo.com, https://orcid.org/0000-0001-5576-5076

Received: October 18, 2024; Revised: December 06, 2024; Accepted: January 06, 2025; Published: March 31, 2025

Abstract

Fingerprint identification in criminal investigations and biometric systems faces challenges because of poor image quality at crime scenes. This work addresses the complex task of latent fingerprint enhancement and segmentation through a Modified Deep Learning Model. The initial step involves normalisation to predefine mean and variance, mitigating grey-level volatility caused by ridges and valleys. Subsequently, a2 Edge Directional Total Variation model-based adaptable de-noise efficiently removes structured noise, enhancing latent fingerprint images. For segmentation, the Modified Mask R-CNN is proposed to identify critical features of overlapping latent fingerprints by assigning additional weight to neighboring borders, enhancing separation. The research introduces a strategy called Atrous-based Modified Mask RCNN with cascaded Atrous II-blocks, residual learning, and Instance Normalization, which proves accurate in improving and segmenting latent fingerprints. The result reveals the proposed technique utilises the Tsinghua Latent Overlapped Fingerprint Database for analysis, which attains a high accuracy of 0.99 compared with existing models.

Keywords: Latent Overlap Fingerprint, Mask RCNN, Edge Directional Total Variation Model (EDTV), Normalization, Segmentation.

1 Introduction

Fingerprints are highly reliable biometrics because of their ability to identify individuals individually. Latent fingerprints are extremely helpful in forensics and law enforcement applications (Al-Rifaee et al., 2023). They are derived from the accidental finger skin imprints left at the crime scene. A typical example of a poor-quality fingerprint is a latent fingerprint (Prabakaran & Pillay, 2021). Random noise that overlaps and has inadequate ridge structure are common problems with latent fingerprints, in contrast to plain and rolled fingerprints obtained under controlled conditions (Shreya & Chatterjee, 2024). Automated Fingerprint Identification Systems (AFIS) have been frequently used to identify fingerprints. For plain and rolled fingerprints, AFIS offers promising accuracy; nevertheless, for latent fingerprint images, the performance is still unsatisfactory (Gibb & Riemen, 2023). Most popular feature

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 16, number: 1 (March), pp. 72-93. DOI: *10.58346/JOWUA.2025.I1.004*

*Corresponding author: Department of Computer Science, Central University of Karnataka, Kadaganchi, Kalaburagi, Karnataka and Affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India.

extraction methods frequently fail to reliably extract relevant features from these fingerprint images because of their low image quality (Kilinc et al., 2023). Enhancing fingerprint images is crucial to address these issues since it lowers noise, restores damaged areas and makes the ridge structure more visible (Dhaneshwar et al., 2021). Improved latent fingerprint images permit more accurate and efficient feature extraction, which improves fingerprint matching and recognition performance (Grosz & Jain, 2023; Kumar et al., 2023).

Traditional techniques for improving fingerprint images concentrate on identifying and eliminating noise from the critical ridge pattern (Qiu et al., 2024; Farhang & Rashidi, 2015). Utilizing data gathered from the frequency domain and orientation field makes it easier to improve the useable zone since noise has different orientation and frequency characteristics than the pattern of the ridge section (Jia et al., 2024). In the early stages of fingerprint recognition research, global alignment issues were solved via 2D minutiae point cloud matching (Trisiana, 2024; Bhilavade et al., 2024). The conventional techniques, however, needed more resilience against distortions and missing details and were computationally costly (Duan et al., 2022). To overcome these difficulties, local descriptor-based fingerprint-matching methods were created (Kumar, 2024). These techniques analyze images' similarity corresponding to rolling and slap fingerprint imprints by utilizing spatial coordinates and angle information (Öztürk et al., 2022). They performed worse for latent fingerprint images despite being efficient for sensor images (Chhabra et al., 2020). Ridge counts and orientation maps are two characteristics beyond minutiae in the research that suggest ways to enhance local descriptor-based techniques in latent fingerprint recognition (Valdes Ramirez, 2021). A deformation-tolerant extension of local descriptor methods was recently presented to address non-linear distortion in latent images. This method involves iteratively grouping matched minutiae (Artan & Semiz, 2024). Previous techniques for creating local descriptor generators were intended to use a $\ell 2$ difference between latent and sensor embeddings for fixed-size image patches derived around minutiae to differentiate between them (Wahab et al., 2024).

In recognition of images, Deep Neural Networks (DNNs) have demonstrated revolutionary performance and comprehension applications during the past several decades, leading to a massive revolution in artificial intelligence (Deshpande et al., 2020; Deshpande & Malemath, 2021). DNNs are often utilized for computer vision applications, including categorization, super-resolution images, and object segmentation. Image enhancement and other image processing applications have extensively used convolutional neural networks, or CNNs (Chai et al., 2021). Research on automatic latent fingerprint identification with deep learning techniques has exploded, with approaches presented for subtasks such as orientation field estimation, latent image minutia extraction, latent image quality evaluation, and minutia descriptor in the synthesis process (Grosz et al., 2022; Alkishri et al., 2023). However, techniques for distinguishing overlapping fingerprints primarily rely on orientation field data, which provides information about the alignment of ridges within the immediate fingerprint environment (Alkishri et al., 2024). Since these separation techniques can be sensitive to background image noise, removing them before using them is essential. Low contrast, noise, and uneven illumination are common problems for latent fingerprints that reduce clarity and hide ridge features (Farhang & Rashidi, 2015). Furthermore, fingerprint features are further distorted by environmental conditions and skin deformations, making minutiae extraction more difficult. Additionally, ridge visibility may be hampered by backdrop patterns created by textured surfaces. It becomes considerably more difficult to isolate distinct ridge patterns when several overlapping fingerprints are present. Hence, there is a need to develop a framework for improving the overlap of latent fingerprint images (Thanaraj et al., 2023).

This work is motivated by the ongoing challenges in fingerprint identification, particularly in cases involving low-quality latent fingerprints from crime scenes. Enhancing and segmenting these

fingerprints is crucial for accurate identification, but complex backgrounds, structured noise, and poor image quality make this task difficult. Traditional down-sampling layers in CNNs, while effective at capturing abstract semantic information, reduce the spatial resolution, which can harm segmentation accuracy.

The work's primary contribution is highlighted as follows:

- The research employs a Modified Deep Learning Model to enhance latent fingerprint segmentation and address the intricate problem of separating the fingerprint area of interest from the complex background, especially in low-quality images and when complex background noise is present.
- The first stage involves normalisation depending on lengthy ridges and valleys, limiting the grey level values. The second step uses an Edge Directional Total Variation model-based adaptable de-noise to efficiently remove latent fingerprint image noise.
- The proposed Atrous-based Modified Mask R-CNN separates overlapping objects by giving additional weight to neighbouring borders, which segment the latent fingerprint image and determine its valuable properties (Kustiawan, 2018). The proposed edge adaptive direction model reaches accurate latent fingerprint enhancement and segmentation.

The remaining portions of the article are organized as follows: Section 1 presents the introduction, Section 2 highlights the previous research, Section 3 analyzes the proposed approach, Section 4 presents the results of the proposed methodology, and Section 5 concludes the study.

2 Related Works

Several algorithms for fingerprint segmentation are available in the literature and are discussed below.

Yoo et al., (2020) offered a deep learning algorithm-based end-to-end overlapping fingerprint separation technique. An image of an overlapping fingerprint is sent into the FinSNet neural network, which is taught to extract the individual fingerprint straight from the input. The scheme achieved better computational efficiency than standard approaches that need background removal independently with a single assumption to remove redundant fingerprints and background images. One drawback is the difficulty in identifying overlapping fingerprints, which can make it difficult to accurately identify someone since it can be challenging to differentiate one print from another.

Bao et al., (2023) provided a unified approach for correcting latent fingerprints taken at a crime scene. The technique is a coarse-to-fine method that combines deep learning network accuracy with the strength of classical pattern recognition. The combined method's drawback is that its performance is contingent upon earlier manual processing, such as image pretreatment and fingerprint extraction.

Pan et al., (2024) proposed a dense minutia descriptor (DMD) for latent fingerprint matching that is based on deep learning. The fingerprint patch aligned by its centre minutia was extracted to produce a DMD, providing both texture and detailed minutia information. With two dimensions linked to the original image plane and the other representing the intangible characteristics, the dense descriptor is defined in three dimensions. The extraction procedure also produces the fingerprint segmentation map, guaranteeing that the descriptor is only helpful in the foreground area. The overlapping regions of two descriptors are matched, and a score normalization technique is used to lessen the significance of disparities outside the valid area. Numerous latent fingerprints with intricate background patterns impair the system's performance, leading to inaccurate minutiae extraction and occasionally insufficient minutiae extraction for successful matching.

Guan et al., (2023) presented a rectification technique that directly estimates the deformed fingerprint's dense distortion field rather than relying on a low-dimensional representation using a self-

reference-based network. With various finger positions and distortion patterns, this technique may produce realistic distortion fields of deformed fingerprints. The drawback is the preprocessing of input fingerprints, which mitigates most of the impacts of low image quality but occasionally still has preprocessing mistakes. Furthermore, it has a higher sensitivity to local textures and has the potential to misidentify some uncommon and particular fingerprint patterns.

Deshpande et al., (2022) presented two methods based on minutiae's proximity. It presented the latent minutiae similarity (LMS) approach to address the patterned boundaries separating the nearest pairs of sites encompassing a minutia. It presented some clustered latent minutiae pattern (CLMP) technique, founded on geometric configurations of a collection of Latent minutiae patterns surrounding a minutia. The construction of a deep neural network for an end-to-end matching framework might be complex owing to large amounts of data and computing resources, and the primary concern is the possibility of testing techniques for database scalability, as it can result in longer processing times and decreased efficiency.

Megha et al., (2021) presented a novel strategy for segmentation and classification problems based on dependable post-feature extraction and effective pre-feature extraction. To identify the structure of interest, it channels less input for patch-based classification in favour of saliency and colour-map-based information. The proposed work provided an enhanced identification and segmentation of latent fingerprints by resolving the requirements for intelligent structure differentiation, patch size impact, and the best selection of features and using a stacked convolutional autoencoder for object identification. With less overfitting, the system automatically selects the best features and detected objects. The segmentation method was more successful and efficient because of clever parameter standardization.

Seidlitz et al., (2021) created Generative Adversarial Networks (GANs) to transform privacy-sensitive latent fingerprint images into synthetic fingerprint datasets that are friendly to privacy. Unlinkable synthetic images are created using three GAN architectures (ProgressiveGAN, StyleGAN, and StyleGAN2), with the true fingerprint data source being NIST Special Database 27. Following StyleGAN2 and StyleGAN, the technique shows that ProgressiveGAN provides the highest anonymity, with 63.53% of images having low Bozorth3 scores (indicating strong privacy protection). However, the speed with which StyleGAN generates more high-quality photos is superior. Limitations include reliance on particular databases and findings that are not entirely generalizable across datasets, even though it effectively creates unlinkable fingerprints.

Zhu et al., (2023) presented a novel approach that uses a generative adversarial network (GAN) architecture to define latent fingerprint enhancement as a restricted fingerprint generation issue. The suggested network is called FingerGAN. It may ensure that the enhanced latent fingerprint it produces, with its orientation domain established by the FMFF model and its fingerprint skeleton map weighted by minutia positions, is compatible with the corresponding actual occurrence. Since the fingerprint's skeleton map may be used to get minutia, an essential feature of fingerprint identification, it provided a comprehensive structure that can improve latent fingerprints in surroundings that are immediately enhanced minutia data. This will significantly improve the performance of latent fingerprint recognition.

Garg et al., (2024) created a model that tackles the problem of small fingerprint dataset sizes by utilizing VGG16 and VGG19 convolutional neural networks with inversion and multi-augmentation approaches. Fingerprint patterns are used to categorize the classes in the FVC2000_DB4 dataset. Sample variety is increased by generating extra images for every feature map through inversion and augmentation. With VGG16 attaining up to 97% and VGG19 reaching 93% on the enhanced dataset, the method outperformed models such as ResNet50 and InceptionV3, demonstrating its great accuracy. Nevertheless, disadvantages include the requirement for more research using alternative datasets and

fusion approaches, such as PCA or LDA, to enhance generalization and efficiency across many biometric modalities (Al-Rifaee et al., 2024).

However, these methods have drawbacks regarding segmentation and image augmentation, especially regarding feature recognition and general performance. Latent fingerprints often have low contrast, noise, and uneven illumination, reducing clarity and obscuring ridge features. Skin deformations and environmental influences can further distort features, making it more difficult to extract minute details, while textured surfaces can generate background patterns that make it difficult to see the ridges. Furthermore, when there are several overlapping fingerprints, it becomes difficult to distinguish individual ridge patterns, requiring sophisticated methods to preserve distinctive details because overlapping regions may obscure important identification characteristics and reduce matching precision. This research proposes a system architecture designed to enhance and segment overlapping latent fingerprints to address these challenges. Table 1 illustrates the Analysis of Deep Learning Approaches for Fingerprint Identification.

Table 1: Analysis of Deep Learning Approaches for Fingerprint Identification

S.No.	Author(s)	Technique Proposed	Result of the Proposed Model	Limitation(s)
1.	Yoo et al., (2020)	Deep learning-based end-to-end overlapping fingerprint separation using FinSNet neural network	Achieved a 75.5% true acceptance rate (TAR) on the Tsinghua OLF database and 84.5% TAR on the Tsinghua SOF database	Difficulty in accurately identifying overlapping fingerprints, challenging to distinguish one print from another
2.	Bao et al., (2023)	Coarse-to-fine latent fingerprint correction, combining deep learning with classical pattern recognition	The cumulative match rate (CMR) improved from 0.65 (original), outperforming other approaches, which achieved 0.85	Performance depends on prior manual processing, such as image pretreatment.
3.	Pan et al., (2024)	Dense Minutia Descriptor (DMD) for latent fingerprint matching	Achieved 79.07% Rank-1 accuracy and 80.23% TAR@FAR=0.1% on NIST SD27; 52.68% Rank-1 accuracy and 51.73% TAR@FAR=0.1% on N2N Latent dataset	Complex backgrounds impair minutia extraction, leading to inaccuracies in matching
4.	Guan et al., (2023)	Self-reference-based rectification technique for estimating fingerprint-dense distortion field	Achieved MRE/AP of 6.34/33.77 on FVC2004 DB1 full subset; Rank-1 accuracy of 92.2% and Rank-5 accuracy of 94.3% on TDF V2 T subset	Sensitive to local textures, prone to misidentification of rare fingerprint patterns
5.	Deshpande et al., (2022)	Latent minutiae similarity (LMS) and Clustered Latent Minutiae Pattern (CLMP) techniques based on minutiae proximity	Achieved Rank-1 accuracy: 97.5% (LMS) and 100% (CLMP) on FVC2004; 88.8% (LMS) and 93.8% (CLMP) on NIST SD27	Complexity and high resource requirements for large databases; potential inefficiency
6.	Megha et al., (2021)	Segmentation and classification using stacked convolutional autoencoder for latent fingerprints	Achieve accuracy of 98.45%	Limited success with non-ideal patch sizes and feature selections
7.	Seidlitz et al., (2021)	Generative Adversarial Networks (ProgressiveGAN, StyleGAN, StyleGAN2) for privacy-friendly synthetic fingerprint data	ProgressiveGAN achieved 63.53% anonymous fingerprints, StyleGAN 49.06% anonymous fingerprints, and StyleGAN2 55.5% anonymous	Reliant on specific databases; limited generalizability
8.	Zhu et al., (2023)	FingerGAN for latent fingerprint enhancement using FMFF model and fingerprint skeleton map	Achieved Rank-1 accuracy: 76.36% with FingerGAN,	Complex GAN structure requires precise skeleton mapping
9.	Garg et al., (2024)	VGG16 and VGG19 CNNs with inversion and multi-augmentation for fingerprint classification	High accuracy achieved (up to 97% with VGG16); improves sample diversity through augmentation	Dependence on specific augmentation techniques needs further exploration with other datasets and fusion methods

3 Proposed Method

The research proposed a Modified Deep Learning Model-Based Latent Fingerprint Enhancement and Segmentation to improve low Image quality and fingerprint accuracy. The approach involves enhancing latent fingerprint images by removing various types of structured noise. Normalisation is the first stage, requiring predefinition of mean and variance. The second step includes an Edge Directional Total Variation model-based adaptable de-noise, efficiently removing latent fingerprint image noise. The improved image output serves as the next step's input, which uses a modified deep-learning model to segment the latent fingerprint image and determine its valuable features. Instance segmentation has improved border and object detection in AI, particularly in fingerprint recognition. However, edge identification between nearby objects still needs to be solved due to the tendency for objects to mix in a single image. The proposed Modified Mask R-CNN successfully separates overlapping objects by giving additional weight to neighbouring borders. Modern CNNs typically include down-sampling layers to widen the receptive area and gather intellectual semantic data, which can minimize the feature maps' spatial dimension. An alternative is Atrous convolution, which retains the feature maps' spatial dimension while expanding the receptive field. A new proposed Atrous-based Modified Mask RCNN with cascaded Atrous II-blocks, residual learning, and Instance Normalization is proposed. Instance Normalizing is used as the normalization approach, and residual learning is incorporated to address gradient vanishing/exploding concerns and facilitate backpropagation. Softmax at the pixel level is used to convert network outputs to probabilities, and Intersection over Union (IoU) is employed to assess segmentation accuracy. The proposed model achieves adaptive, effective, and accurate latent fingerprint enhancement and segmentation. Figure 1 represents the block schematic of the proposed technique. The Pre-processing procedure will be discussed in the section below.

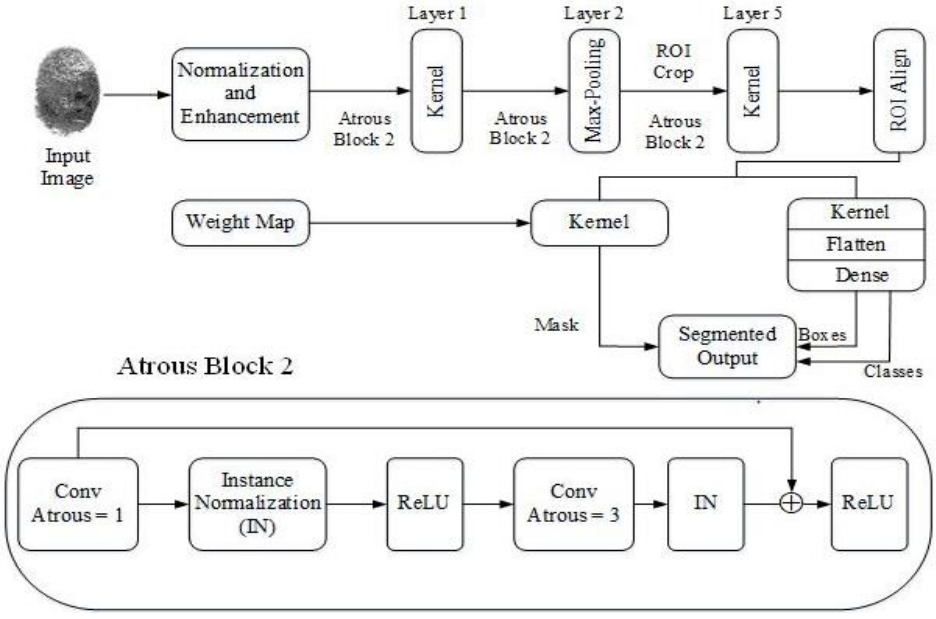

Figure 1: Schematic Representation of the Proposed Approach

1) Pre-Processing

The suggested approach consists of improving latent fingerprint images, which is necessary to remove various structured noises based on the EDTV model to enhance image quality.

- **Normalization**

The technique of processing latent fingerprint image enhancement involves presenting I as NxN, the grey level of the latent fingerprint image (I (i,j)), which indicates the pixel's intensity at I, j for each row and column. Therefore, the following equation (1, 2) defines and illustrates the mean $M(I)$ and variance $Var(I)$ in the grey level of the latent fingerprint Image, I.

$$M(I) = \frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} I(i,j) \quad (1)$$

$M(I)$represent the mean grey level of the fingerprint image, calculated as the average of all pixel intensities, N denotes the dimension of the matrix, indicating that the image is N x N, and Var(I) denotes the variance of the grey levels in the fingerprint image, measuring the spread of pixel values around the mean.

$$Var(I) = \frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(I(i,j) - M(I))^2 \quad (2)$$

As demonstrated by equations (1) and (2), normalization is the initial stage of latent fingerprint enhancement, where I (i, j) is denoted as the grey-level value pixel (i, j). Mean and variance must be predefined to prevent long-term ridges and valleys from causing fluctuations in the grey-level values. Conversely, M and VAR are the predicted mean and variance of I, while G (i, j) represents the normalized grey-level value at pixel (i, j). The following equation (3) is used to determine the normalization of an image.

$$G(i,j) = \begin{cases} M_0 + \sqrt{\frac{Var_0(\mathrm{I(i,j)-M})^2}{Var}}, if\ \ I(i,j) > M \\ M_0 - \sqrt{\frac{Var_0(\mathrm{I(i,j)-M})^2}{Var}},\ \ otherwise \end{cases} \quad (3)$$

The clarity of the ridge and furrow patterns is unchanged despite the pixel operation, where M_0 and Var_0 represent the intended mean and variance values for the normalisation process.

- **Edge Directional Total Variation Model-based Adaptable De-noise**

The noise of the latent fingerprint images was successfully removed using it. The model can handle Images with many dominating orientations and is adaptable to the image's edge direction. The total variation model was suggested to extract an image f (a, b) from a noisy Image i (a, b). This model may be expressed mathematically as follows as equation (4)

$$E_{TV}(f) = \lambda \int \Omega|\nabla f|d\Omega + {}^1\!/_2 \int \Omega(f-1)^2 d\Omega \quad (4)$$

$E_{TV}(f)$denotes the total variation energy of the image f, which measures the smoothness of the image while preserving edges, λ specifies the regularization parameter that controls the trade-off between fidelity to the data (the noisy image) and the smoothness of the reconstructed image, Ω describes the domain of the image where the integral is calculated, representing the area over which the image is defined, ∇fdenotes the magnitude of the gradient of the image **f**, which indicates the strength of the edges within the image and $(f-1)^2$illustrates the squared difference between the image f and the constant value 1, measuring the fidelity to the data in the denoising process.

A potential model for the Image denoising process is the directional total variation model (DTV). Let $\lambda DTV_{\alpha,\theta(f)}$illustrate the directional total variation term that adapts to the dominant orientation θ of edges in the image, with α representing additional parameters that may influence the model's sensitivity to

direction, and $E_{DTV}(f)$ denotes the directional total variation energy of the image f, which accounts for the preferred orientation in edge preservation. Equation (5) is given below.

$$E_{DTV}(f) = \lambda DTV_{\alpha,\theta(f)} + {}^{1}/_{2} \int \Omega(f-1)^2 d\Omega \qquad (5)$$

When the dominating direction in an image corresponds with the direction θ, the DTV model is needed to improve diffusion together with that direction. When the image has many dominating orientations, it is essential to have a parameter of θ that is spatially distinct over the whole image. As a result, presented the EDTV, which allows for spatially altering θ (a, b) according to the image's edge direction. Latent fingerprint image noise is efficiently removed with this technique. The segmentation process will be discussed in the upcoming section.

2) Modified Mask R-CNN

The two main components of Mask R-CNN network topologies are instance segmentation and feature extraction. The backbone and feature pyramid network (FPN) architecture are used for increased accuracy and processing speed. The ROI is detected and classified based on the produced ROI once the model has identified the ROI in the network architecture block. Finally, forecasting the fully convolutional mask, for instance, segmentation, is constructed using these frameworks. Based on the success of the Mask R-CNN, we now suggest a modified Mask R-CNN, specifically made to consider pre-established weights for the primary objective function. The primary goal of this technique is to accurately identify the boundaries between several samples in the framework of segmenting instances.

In summary, the Mask RCNN's tasks primarily accomplish three objectives: (1) segmenting instances, (2) identifying bounding boxes, and (3) categorizing class labels. The model generates feature maps by running scaled images through the CNN. The Mask R-CNN's novelty aids in advancing earlier image recognition models in this process. The Mask R-CNN implements pixel-wise binary classification, decoupling mask prediction from both bounding box detection and category classification while constructing the objective function. Notably, the binary classification approach has benefits in terms of lower computing costs. Targeting exact mask alignment, the ROI alignment approximates ground truth regions.

The suggested goal task for the modified Mask R-CNN is as follows as equation (6) and (7). Let Ldenote the total loss for the modified Mask R-CNN, L_{cls}specifies the classification loss, L_{box}illustrates the bounding box regression loss that measures the accuracy of the predicted bounding boxes, w_{mod}denotes the weighting parameter that adjusts the importance of the mask loss in the overall objective function, and L_{mask}defines the mask loss.

$$L = L_{cls} + L_{box} + w_{mod} L_{mask} \qquad (6)$$

$$\mathrm{L} = \frac{1}{N_{cls}} \sum_i \{-p_i^* . \log p_i - (1-p_i^*). \log(1-p_i)\} + \frac{\lambda}{N_{box}} \sum_i p_i^* . L_1^{smooth}(t_i - t_i^*) - w_{mod} . \frac{1}{m^2} \sum_{1 \le i,j \le m} \{b_{ij} . log\hat{y}_{ij}^k + (1-b_{ij}). \log(1-\hat{b}_{ij}^k)\} \qquad (7)$$

Where t_i is the expected four coordinates with parameters,t_i^* is the coordinates for the actual truth,N_{cls}specifies the total number of class labels, L_1^{smooth}denotes the smooth loss, $\frac{1}{m^2}$ defines the number of pixel instances being evaluated, b_{ij} denotes the binary indicator for the presence of object j at pixel i,$log\hat{y}_{ij}^k$ represent the predicted probability that pixel i belongs to class k,$p_i$represents the anticipated probability that anchor i is an object, and p_i^* is the binary actual label indicating whether or not anchor i is an object. To ensure that the weights of the $L_{cls}$and L_{box} terms are equal, the balance

parameter λ is between 0 and 10. The number of ground truth classes is k, and the weight matrix allocated to pixel instances is $w_{ood}$and N_{cls}, N_{box}, and N_{boxes} are the normalisation terms setting the total amount of anchor sites (0~2400), to mini-batch size (0~256) locations and the equation (8) is given below. Let(a) represent the difference between the predicted and actual coordinates, and $L_1^{smooth}(a)$ defines the smooth L_1 loss function to mitigate issues with gradient vanishing/exploding during training.

$$L_1^{smooth}(a) = \begin{cases} 0.5a^2, & if\ |a| \leq 1, \\ |a| - 0.5, & otherwise \end{cases} \quad (8)$$

Moreover, it incorporates image representations and previous information about adjacency into the model. This weight was developed with the U-Net as inspiration, creating a substantial separation among samples as borders approach. Theoretically, the weight increases as the barrier gets closer. Equations (9) and (10) are given below

$$w(a) = w_c(a) + w_0 . e^{-(d_1(a)+d_2(a))^2/2\sigma^2} \quad (9)$$

From equation (9), $w(a)$denotes the computed weight for each pixel based on its proximity to boundaries, $w_c(a)$ illustrates the weight map that balances class frequencies across the image, $w_0$specifies the base weight that contributes to the overall weight calculation, and σ^2 denotes the weight adjusting parameter that emphasizes the importance of object boundaries.

$$w = 1 + \frac{w(a) - \min(w(a))}{\max(w(x)) - \min(w(x))} . \delta \quad (10)$$

Where d_1: Ω → R indicates the distance that exists from the boundary of the nearest cell, d_2: Ω → R indicates the distance to the second closest cell's boundary, and δ specifies the weight adjusting parameter, respectively. The $\min(w(a))$ denotes the minimum weight value, the max specifies the maximum weight value in the weight map, and δ denotes the weight adjusting parameter, emphasizing the importance of object boundaries. The weight map w_c: Ω → R balances the class frequencies. W(x) essentially depends on the size, shape, and distance of the objects in an image. Scale each weight mapped separately to the interval from 0 to 1 to accommodate variability. Next, to get the weight matrix's power, consider the parameter δ. The weight parameter δ can be used to emphasize an object's border more, mainly when there is little space between them, making boundaries challenging to see. Then, this weight matrix is applied to the mask's objective function using element-by-element computing. The image is segmented from the overlapping images, and the Atrous-based modified mask RCNN will be discussed in the section below.

3) Atrous-Based Modified Mask RCNN

The Atrous-based Modified Mask R-CNN is a sophisticated variant of the Mask R-CNN architecture that uses atrous (or dilated) convolution layers to increase the receptive field of the model while preserving computing effectiveness and lowering the number of parameters. In particular, the design seeks to enhance feature extraction for intricate segmentation tasks. Atrous convolutions allow the model to cover a larger receptive field without adding more layers by introducing gaps (or dilation rates) within the kernel. This characteristic is essential in applications like segmentation and object identification, where it is necessary to identify global patterns over wide spatial extents. With the fewest possible convolutional layers, the model maximizes the receptive field by optimising the atrous rate configuration, also known asar_M. Each atrous layer uses a distinct dilation rate to strike a compromise between computational efficiency and field coverage.

Concentrate on upgrading the atrous rate configuration to obtain the biggest and most enclosed reception area with the fewest atrous convolutional layers. Put $f_{1D}$and F_{2D}in the 1D and 2D mapping of feature designations using a feature map as input. The m^{th} atrous convolutional layer with an atrous rate ar_m calculates $F_{2D}^{m-1}$of size $H \times W \times c_{m-1}$, an output feature map F_{2D}^{m} of size $H \times W \times c_m$, where$F_{2D}^{0} \in \mathbb{R}^{H \times W \times c_0}$, $F_{2D}^{m} \in \mathbb{R}^{H \times W \times c_m}$, $m \in [1, M] \cap \mathrm{M}$, and$ar = (ar_1, \dots, ar_M)^T \in \mathrm{M}^M$, whereM $\in \mathbb{Z}_+$ is the overall amount of convolutional layers in the atrous. The feature height, in this case, is represented by $H \in$ M, and the feature width by $W \in$ M; The input Image is F_{2D}^{0}, and the channel number for feature maps is$c_0 = (c_0, \dots, c_M)^T \in \mathrm{M}^{M+1}$. A comparable 2D atrous convolution, with kernel v^m indexed by$t \in Z$, might produce a backward propagation from F_{2D}^{m} to F_{2D}^{m-1}, which can be broken down into two 1D atrous convolutions by omitting the non-linear modules, or relu, and the biases. The equation (11) is given as

$$v_t^m(k, ar_m) = \sum_{u=-(k-1)/2}^{(k-1)/2} w_u^m . 1(t - u\, ar_m) \tag{11}$$

Here, the kernel size is denoted by an odd integer, k, which might be 3, 5, or 7. The index of a pixel is t.$v_t^m(k, ar_m)$ denotes the convolutional kernel of M^{th} layer, and $1(t - u, ar_m)$ denotes the indicator function that equals one if t=u ar_motherwise 0.Every entry in the weight matrix$w^m \in \mathbb{R}^k$, denoted as w_u^m, is a trainable variable. The definition of the indicator function $1(t)$: $\mathbb{Z} \longrightarrow \{0,1\}$ is given as equation (12):

$$1(t) \coloneqq \begin{cases} 1 & t = 0 \\ 0 & t \neq 0 \end{cases} \tag{12}$$

$1(t)$denotes the indicator function, which outputs 1 when t=0 and 0 otherwise. This function controls the inclusion of specific pixels based on their positions. Vectorsf_{1D}^{0}, f_{1D}^{m} represent the m^{th} 1D feature map and the 1D input Image and are indexed by t, and $v^1 * \cdots * v^m$denotes the sequence of kernels for each convolutional layer up to the M^{th}layer.The calculation of f_{1D}^{0} from f_{1D}^{m}may be done as follows given as equation (13):

$$f_{1D}^{0} = v^1 * \cdots * v^m * f_{1D}^{m} \tag{13}$$

Compute the expression $d^m(k, ar) \coloneqq f_{1D}^{0}(f_{1D}^{m} = 1(t))$ where $f_{1D}^{m} = 1(t)$ denotes that only the centre pixel of f_{1D}^{m} has a non zero value of one. $d^m(k, ar)$Illustrates the result of convolving a sequence of kernels $v^1 * \cdots * v^m$ with the indicator function $1(t)$. It is computed as follows is given as equation (14):

$$d^m(k, ar) \coloneqq v^1 * \cdots * v^m * 1(t) \tag{14}$$

The link number from $f_{1D_0}^{m}$to the pixel or node of the input image is determined by setting$w^n = (1)_k, \forall m$, vectors that have one, then$d_t^m \in \mathbb{N}$, the defined component by$t \in \mathbb{Z}$. Consequently, the non-zero element number in vector d^m may indicate the receptive field coverage of $f_{1D_0}^{m}$, signified by d^m. The equation (15) is given as

$$\|d^m\|_0 \coloneqq \sum_t (1 - 1(d_t^m)) \tag{15}$$

Where $\|d^m\|_0$denotes the count of non-zero elements in the vectord^m, representing the coverage of the receptive field. $1(d_t^m)$ defines the indicator function applied to (d_t^m) to check if (d_t^m) is non-zero. The$s^m \in \mathbb{N}$receptive field size is computed as follows given as equation (16).

$$s^m(k, ar) = 1 + (k - 1) \sum_{n=1}^{m} ar_m \tag{16}$$

Next, determine the proportion of receptive field area of $f_{1D_0}^{m}$, represented by $\rho^m \in \mathbb{R}_+$, as follows is given as equation (17):

$$\rho^m(k, ar) \coloneqq \frac{\|d^m\|_0}{s^m} \quad (17)$$

From equation (17), $\rho^m(k, ar)$denotes the receptive field coverage ratio for the Mthlayer, $\|d^m\|_0$ illustrates the non-zero count ind^m, representing receptive field coverage and s^mdenotes the receptive field size for the Mth layer. The aim is to maximize the receptive field dimensionss^Mwhile keeping the coverage ratio$\rho^M = 1$, constrained to guarantee an entirely enclosed area for reception, which is given as equation (18). The set of possible atrous rates across M layers is denotedasM^M.

$$\underset{ar \in M^M}{max}\{s^M : \rho^M = 1\} \quad (18)$$

Maximize the count of non-zero elements ind^M, the coverage metric, while ensuring it equals the expression for the receptive field size. Upon replacing equations (16) and (17) with (18), the optimization problem may be transformed into equation (19):

$$\underset{ar \in M^M}{max}\{\|d^M\|_0 : \|d^M\|_0 = 1 + (k-1)\sum_{n=1}^{m} ar_m\} \quad (19)$$

The following represents the total number of links from $f_{1D_0}^m$ to f_{1D}^0 is given as equation (20):

$$\|d^m\|_1 = \sum_t d_t^m = (k)^m \quad (20)$$

Where an exponent computation is represented by $(k)^m$. The reason it is the upper bound of $\|d\|_1$ is given as equation (21) and (22):

$$\|d\|_0 \leq \|d\|_1, \forall d_t \in M, \forall t \in \mathbb{Z} \quad (21)$$

The non-zero element count $\|d\|_0$ is always less than or equal to the total count $\|d\|_1$, ensuring consistent coverage within bounds. The condition $\|d\|_0 = \|d\|_1$ holds if d_t values are binary (0 or 1), representing full coverage with no partial activations.

$$\|d\|_0 = \|d\|_1, \Leftrightarrow d_t \in \{0,1\}, \forall t \in \mathbb{Z} \quad (22)$$

Assumed to be true is that (22). Putting this in place of the limitation in (19). The receptive field size with M layers and cumulative atrous rate ar_m matches the cumulative kernel exponent, ensuring proportional scaling. The equation (23) is given below:

$$1 + (k-1)\sum_{n=1}^{M} ar_m = (k)^M \quad (23)$$

Acquiring a single solution for this geometric progression is possible, as given below in equation (24).

$$ar' = (1 \ldots (k)^{m-1} \ldots (k)^{M-1})^T \quad (24)$$

From equation (24), ar'specifies the specific atrous rate configuration where the atrous rate scales in a geometric progression, achieving maximal coverage with optimal settings. It fulfils a receptive field that is evenly covered as equation (25):

$$d_t^M(k, ar') = \begin{cases} 0 & t \notin \mathbb{S} \\ 1 & t \in \mathbb{S} \end{cases}, where\ \mathbb{S} \coloneqq [-\frac{s^M-1}{2}, \frac{s^M-1}{2}] \cap \mathbb{Z} \quad (25)$$

in 1D and the similar in 2D, which fulfils the corresponding need in (21) and therefore resolves (19). The most extensive and fully covered receptive field, considering the same quantity of atrous convolutional layers, might thus result from the atrous rate setting of $(k)^{m-1}$ at the m^{th} atrous convolutional layer. Let y and x be the two border pixels, where $y \leq 0$; $x \geq 0$, and $x - y = H - 1$. Hence, we might rewrite equation (13) as equation (26), where $\varepsilon(t) :=$. The one-dimensional path number's broad concept formula is found in equation (27) when equation (10) is substituted into equation (25).

$$d^m(y,k,ar) \coloneqq \underbrace{(\dots[(o * v^m).\varepsilon(t-y).\varepsilon(x-t)]\dots * v^1).\varepsilon(t-y).\varepsilon(x-t)}_{m-fold\ atrous\ convolution\ with\ truncation} \quad (26)$$

$$d_t^m = \sum_{u_m=max\left(-\frac{k-1}{2},\left\lceil x/ar_m\right\rceil\right)}^{min\left(-\frac{k-1}{2},\left\lfloor x/ar_m\right\rfloor\right)} \left\{ \dots \sum_{u_1=max\left(-\frac{k-1}{2},\left\lceil y-\sum_{n=2}^{m} u_n ar_n/ar_1\right\rceil\right)}^{min\left(-\frac{k-1}{2},\left\lfloor x-\sum_{n=2}^{m} u_n ar_n/ar_1\right\rfloor\right)} \{1(t-\sum_{n=1}^{m} u_n ar_n)\} \dots \right\} \quad (27)$$

With the proof, a block of M atrous convolutional layers might reach a receptive field of $(k)^M$. $d^m(y,k,ar)$defines the resulting feature map after mmm-fold atrous convolution with truncation, y,x specifies the boundary pixels in the receptive field, and $\varepsilon(t-y)$ defines the truncation function restricting values to a specific range. In the receptive field, every node is equally connected. According to the parameters employed, the atrous convolutional layers in this research have a kernel size of 3. The receptive field of a block of M atrous convolutional layers is $(3)^M$. Referring to this block as the atrous block and the one with M atrous convolutional layers as the M-block, the Roman numeral M is used here. Large N values result in a strong truncation effect and fewer trainable parameters, even if they also represent more extensive receptive fields. To extend the receptive field linearly by $(3)^2$, the suggested AMMRCNN is built into several cascaded atrous II blocks as a trade-off. $\frac{RF-1}{(3)^2}$. blocks are required to attain an RF receptive field. Residual learning is incorporated while Instance Normalization (IN) is utilised as the normalization technique, where IN demonstrated more excellent performance than another normalization method to solve the gradient vanishing/exploding difficulties and facilitate backpropagation. The intended receptive field determines the residual II-blocks, or$-\frac{(RF-1)}{8}$.The Atrous-based Modified Mask R-CNN integrates the EDTV (Enhanced Dual-Task Variational) model, which improves spatial feature alignment by refining the atrous configuration and introducing atrous II blocks. The proposed architecture offers enhanced receptive field management by adapting atrous rates efficiently and implementing residual connections. The EDTV model improves the network's capacity to extract temporal features, especially for video segmentation tasks where temporal coherence is crucial. It collects motion dynamics and offers a contextual understanding of the objects across frame sequences. This temporal awareness uses information from neighbouring frames to increase segmentation accuracy. The cascaded layers of atrous convolutions known as the Atrous II blocks are intended to efficiently increase the receptive field while regulating the model's parameter count. Multiple atrous convolutional layers, with variable dilation rates, make up each atrous II block. This enables the model to gradually collect data from more extensive areas of the input feature maps. These blocks enable the Atrous-based Modified Mask R-CNN to better segment images by handling their complicated structures and different scales. Better contextual information collecting is made possible by the receptive field's linear extension through these blocks. Finally, a pixel-level softmax layer converts the network's outputs to probability distributions, optimizing segmentation accuracy by treating each pixel independently. The output probability map thus provides a more precise delineation of object boundaries. The outcome of the proposed scheme will be discussed in the below section. Algorithm 1 for the proposed technique is given below.

Algorithm 1: Modified Deep Learning Model-Based Latent Fingerprint Enhancement and Segmentation

Step 1:	**Input:** Overlap latent fingerprint image I (i, j)
Step 2:	**Normalization:**
	Calculate the mean and variance intensity value of the entire fingerprint image I (i, j)
	The normalisation process adjusts each pixel value I (i, j) based on the original image's mean and variance, enhancing the contrast and clarity.
Step 3:	**EDTV model-based adaptable de-noise**
	Initialize the parameters: λ: Regularization parameter, α: Parameter controlling the directional total variation and θ (a, b): Edge direction map.
	Compute the value using Eq. (4)
	for each iteration
	Compute the DTV using Eq. (5)
	Update θ (a, b) based on the current estimate f (a, b)
	end for
	return denoised_image
Step 4:	**Atrous-based Modified Mask RCNN**
	It adds extra weight to neighbouring borders using Eq. (6), which helps separate overlapping objects within the fingerprint image.
	Detect and classify ROIs in the feature maps.
	Implement pixel-wise binary classification to predict masks for each instance.
	Define Atrous Convolutional Layers
	Compute Receptive Field using Eq. (16)
	Optimize Atrous Rate using Eq. (19)
	Use multiple II blocks for a sizeable receptive field.
	Incorporate residual learning and instance normalization.
	Softmax at Pixel Level, which converts network outputs into probabilities, enabling precise segmentation of each pixel in the image
Step 5:	**Output:** Final Segmented Image

4 Result and Discussion

The performance indicators, comparison analysis, and results of the proposed approach are presented in this section. The Python Tensorflow tool developed the Modified Deep Learning Model-Based Latent Fingerprint Enhancement and Segmentation. Tsinghua Latent Overlapped Fingerprint Database carried out the analysis. There are 12 simple fingerprints and 100 latent overlapping fingerprints in this database. There are 100 overlapping fingerprints in the overlap folder within the subdirectory, and binary mask images of these overlapped fingerprints (2 masks per image) may be found in the mask folder. The template folder has 12 template fingerprints. It is employed to evaluate methods for fingerprint overlap separation. In both datasets, an image consists of two overlapping fingerprints with distinct rotations or spatial displacements. Separation may be done double for distinct target fingerprints in a single overlapping image since each image contains two fingerprints. Consequently, 100 overlapping images from each database may be used to evaluate the method. To match fingerprints, the resolutions of the divided images were changed to 500 pixels per inch. The fingerprint corresponding to the separation goal in each overlapping fingerprint image is vertically aligned and centred. The dataset was split up into training and testing sets. For 25 epochs, the proposed technique is trained using 80% of the dataset and tested using 20% of it. Without any decay methods, the networks were trained using the Adam optimiser with a batch size of 32 and a learning rate of 0.0001. The input image has 256 × 256-pixel dimensions. Adam optimizer was used to decrease the loss, computed using binary cross entropy, a loss function frequently used in deep learning.

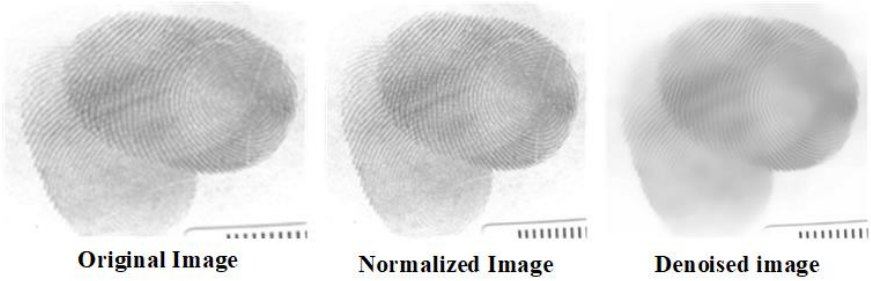

Figure 2: Pre-Processing Output

Figure 2 depicts the Pre-processing for the proposed method. Using the mean and variance from the original image, the normalization process averages out the pixel values in an image. The EDTV model selectively targets noise without compromising image details because it considers edge-directed features and Image structure. It enhances image quality by lowering noise while maintaining structural and directional information by combining normalization with the EDTV model.

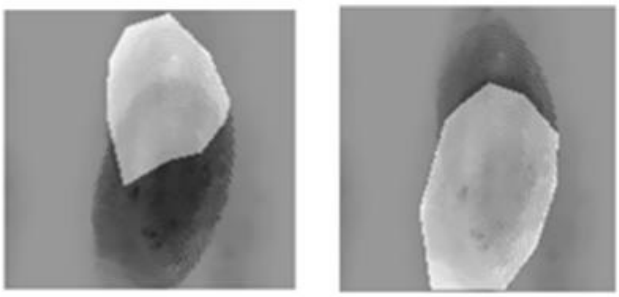

Figure 3. Segmentation Results

Figure 3 shows the segmentation results for the proposed Modified Mask RCNN. For instance, in segmentation, the model employs a comprehensive pixel-by-pixel mask that enables accurate delineation of Image borders in complex forms or closely spaced images by adding the weight function in the Mask RCNN. Its exceptional ability to capture minute details and distinguish different occurrences within the same image enhances the quality.

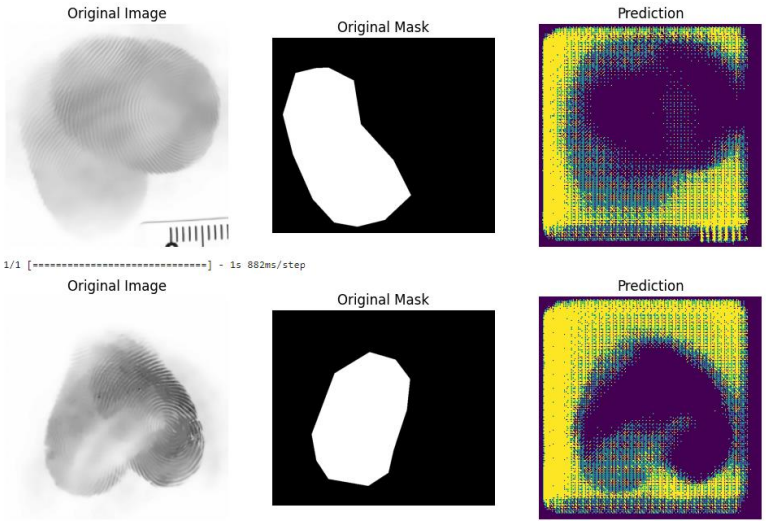

Figure 4: Prediction Results

Figure 4 portrays the proposed Atrous-based modified Mask RCNN prediction results. The outcomes reveal that the proposed technique utilizes the cascaded Atrous block and instance normalization with the softmax function, which calculates the probability of predicting the overlap image based on segmentation.

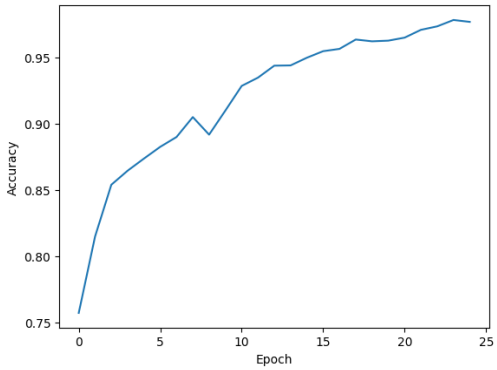

Figure 5: Accuracy Vs Epochs

The accuracy of the suggested strategy is exposed in Figure 5. The graph indicates that the proposed strategy used the Atrous-based MMRCNN with Adam optimizer and attained an accuracy of 0.99 for 25 epochs. The proposed method increased gradually and obtained high accuracy.

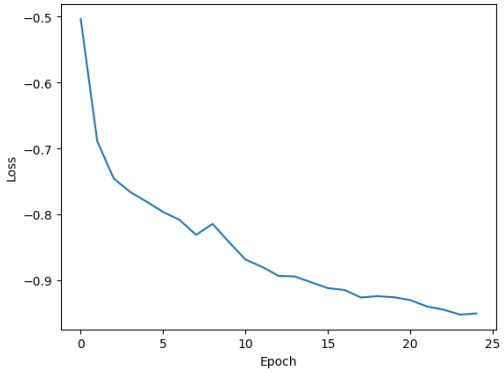

Figure 6: Loss Vs Epochs

Figure 6 illustrates the loss for the proposed approach. The graph indicates that the proposed method decreased gradually from -0.5 and attained a loss of -0.95. When the accuracy increases, the loss automatically decreases, and the error will be low.

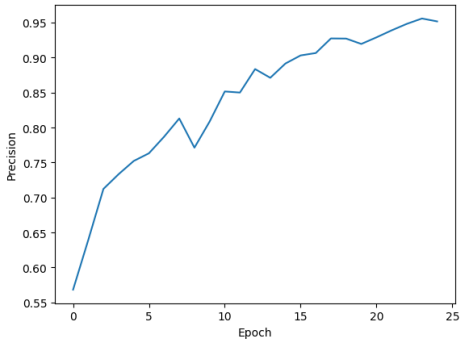

Figure 7: Precision vs Epochs

Figure 7 portrays the precision of the proposed architecture. The graph exposes the proposed architecture for precision vs epochs, which slightly increases and decreases up to 25 epochs. The proposed Modified deep learning model, which extracts the deep features based on Atrous-based modified mask RCNN, attained a precision of 0.58 for 0 epoch, 0.75 for 5 epochs, 0.85 for 10 epochs, 0.88 for 15 epochs, 0.90 for 20 epochs and 0.94 for 25 epochs.

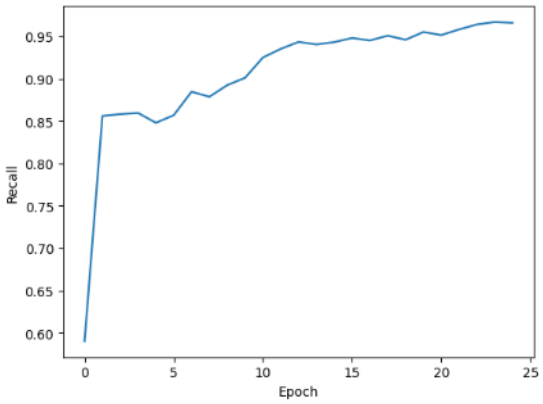

Figure 8: Recall vs Epochs

Figure 8 shows the recall for the proposed architecture. The graph exposes the proposed architecture for recall vs. epochs, which slightly increases and decreases up to 25 epochs. The proposed Modified deep learning model attained a recall of 0.59 for 0 epochs, 0.86 for 5 epochs, 0.91 for 10 epochs, 0.93 for 15 epochs, 0.95 for 20 epochs, and 0.96 for 25 epochs.

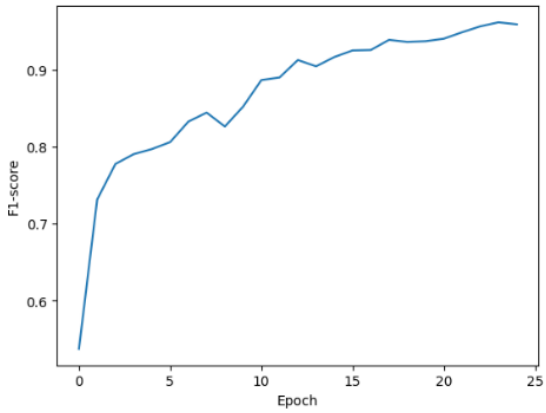

Figure 9: F1-Score vs Epochs

Figure 9 portrays the F1 score for the proposed architecture. The graph reveals that the proposed architecture for F1-score vs epochs slightly increases and decreases up to 25 epochs. The proposed Modified deep learning model attained the F1-score of 0.5 for 0 epochs, 0.8 for 5 epochs, 0.88 for 10 epochs, 0.90 for 15 epochs, 0.93 for 20 epochs and 0.95 for 25 epochs.

To estimate the performance metrics of the proposed modified deep learning model, such as Accuracy, Precision, Recall, and F1-score. The following describes the performance metrics.

Accuracy: A statistical metric assesses how closely a result is when the observed or anticipated value agrees with the actual or expected value. Accuracy is determined by multiplying the outcome by 100 and dividing the overall amount of predictions by the number of correct predictions (Equation 28).

$$Accuracy = \frac{TN+TP}{TP+TN+FP+FN} \quad (28)$$

Precision: It determines the percentage of correctly forecasted positive cases (true positives) out of all instances expected as positive (true positives plus false positives). "False Positives" are the percentage of situations where the model correctly forecasts a negative outcome, whereas "True Positives" are the percentage of situations where the model correctly forecasts a positive outcome. Precision provides insight into the model's capacity to prevent false positives. It is essential when the data exhibits a class imbalance or false positives, which have serious consequences (Equation 29).

$$Precision = \frac{TP}{TP+FP} \quad (29)$$

Recall: It calculates the proportion of all actual favourable occurrences accurately forecasted as positive (true positives). The number of instances falsely projected as negative when it should have been positive is known as "False Negatives." Recall is a helpful metric for assessing how well a model can identify or include positive instances throughout the population of actual positive cases. It is especially crucial when there is a requirement to maximize the identification of positive cases or severe repercussions from false negatives (missing positive occurrences) (Equation 30).

$$Recall = \frac{TP}{TP+FN} \quad (30)$$

F1-Score: Precision and Recall are weighted averages for the F1 Score. For this reason, erroneous positives and negatives are considered while determining this score. Accuracy operates better when the expenses associated with false positives and false negatives are almost equivalent. The F1 Score is the Harmonic Mean of recall and precision (Equation 31).

$$F1 - Score = 2 \times \frac{precision \times recall}{precision+recall} \quad (31)$$

Intersection over Union (IoU): Intersection over union is a frequently used measure in object identification algorithms to evaluate localization accuracy and compute localization faults. To calculate IoU, their union divides the gap between the expected annotations and the actual truth. The similarity between two sets is determined by splitting the point where two sets intersect by their union (Equation 32).

$$IoU = \frac{TP}{TP+FP+FN} \quad (32)$$

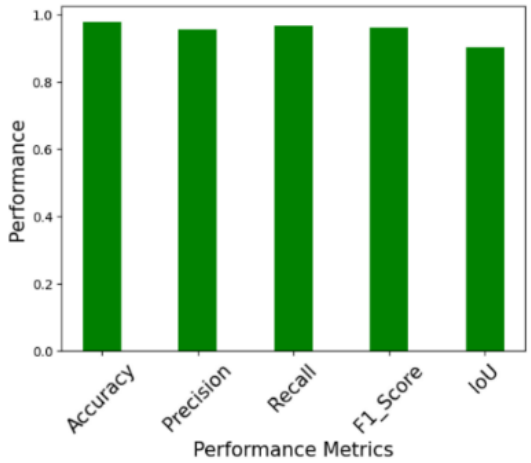

Figure 10: Performance Metrics Evaluation

Figure 10 depicts the performance evaluation of the proposed technique. The graph reveals that the proposed technique utilized the atrous-based modified mask RCNN, which segmented the overlap image and attained an accuracy of 0.99, precision of 0.94, recall of 0.96, and F1-Score of 0.95. Based on these performances, the latent overlap fingerprints are segmented and enhanced. Table 2 illustrates the accuracy comparison between different existing datasets.

Table 2: Accuracy Comparison Between Different Existing Datasets

Dataset	Accuracy (%)		
	NIST SD 4 Gu et al., 2022)	NIST SD 14 Gu et al., 2022)	Tsinghua Latent Overlapped Fingerprint (proposed)
MSR (Gu et al., 2022)	99.27	99.81	-
Proposed	99.28	99.83	99.85

The table 2 compares the accuracy of existing methods such as MSR and three fingerprint datasets: NIST SD 4, NIST SD 14, and Tsinghua Latent Overlapped Fingerprint. The MSR method achieved 99.27% accuracy on NIST SD 4 and 99.81% on NIST SD 14 but did not report results for the Tsinghua dataset. In contrast, the Proposed Method slightly outperformed MSR with 99.28% on NIST SD 4, 99.83% on NIST SD 14, and demonstrated 99.85% accuracy on the Tsinghua dataset, indicating its effectiveness, especially in handling complex or latent fingerprints.

To compare the performance metrics as the accuracy of the proposed modified deep learning-based latent overlap fingerprint with another existing model such as CaoNet (Khan & Wani, 2019), VGGNet (Krizhevsky et al., 2012), AlexNet (Simonyan, 2014) and DNCNN (Deep Normalization Convolutional Neural Network) (Chegur et al., 2023). In existing techniques, edge identification between nearby objects is a complex problem, so the Atrous-based modified mask RCNN is proposed to segment the fingerprints from the latent overlap images. The proposed technique attained an accuracy of 0.99 in contrast to previous approaches, such as an accuracy of 0.86 for CaoNet, 0.85 for VGGNet, and 0.856 for AlexNet. The proposed modified deep learning model attained high accuracy when contrasted with existing techniques. Therefore, the proposed strategy performs superiorly and accurately segments the latent overlap fingerprints. Table 3 compares the proposed method for accuracy.

Table 3: Comparison of the Accuracy

Models	**Accuracy**	**Precision**	**Recall**	**F1 Score**
CaoNet (Khan & Wani, 2019)	0.86	-	-	-
VGGNet (Krizhevsky et al., 2012)	0.85	-	-	-
AlexNet (Simonyan, 2014)	0.85	-	-	-
DNCNN (Chegur et al., 2023)	0.88	0.85	0.85	0.85
Modified Proposed approach	0.99	0.94	0.96	0.95

The existing technique faces issues as extracting the overlap fingerprint is challenging. The proposed method utilized the Atrous-based modified Mask RCNN, which extracts the fingerprints from latent overlap images and predicts the segmented overlap fingerprint images attained a high accuracy of 0.99 when compared to existing models. Regarding accuracy, the proposed approach performs 15.12% better than CaoNet, 16.47% better than VGGNet, and 15.65% better than AlexNet.

According to the study above, the proposed strategy should have the following qualities: it should be precise, efficient, and flexible. The proposed edge adaptive direction model achieves accurate segmentation and enhancement of latent fingerprints.

5 Conclusion

The modified deep learning model utilized the Tsinghua Latent Overlapped Fingerprint Database for analysis. Normalization is the initial step, and latent fingerprint image noise is eliminated using an adjustable de-noise based on the EDTV model. The second step uses the improved Image output to segment the latent fingerprint image and identify its features using a modified deep-learning model. The tendency of items to blend in a single Image is a barrier for edge recognition between neighbouring objects, which is also addressed in this work. The suggested method uses Atrous convolution, which preserves the feature maps' spatial dimension while expanding the receptive field. With cascaded Atrous II blocks, residual learning, and instance normalization, the study introduces a novel Atrous-based Modified Mask RCNN that tackles gradient vanishing/exploding issues and promotes backpropagation. The proposed model achieved an accuracy of 0.99, a precision of 0.94, a recall of 0.96, and an F1-score of 0.95. Compared to existing models, the modified model demonstrated significant improvements, with an approximately 15.12% increase in accuracy over CaoNet, a 16.47% improvement over VGGNet, and a 15.65% enhancement over AlexNet. Future research could focus on data privacy with ethical norms, investigate practical uses of latent fingerprint segmentation, and improve computational efficiency for large-scale deployments. Furthermore, federated learning and other privacy-preserving methods could be incorporated to safeguard biometric processing.

References

[1] Alkishri, W., Widyarto, S., & Yousif, J. H. (2024). Evaluating the Effectiveness of a Gan Fingerprint Removal Approach in Fooling Deepfake Face Detection. *Journal of Internet Services and Information Security (JISIS)*, *14*(1), 85-103. https://doi.org/10.58346/JISIS.2024.I1.006

[2] Alkishri, W., Widyarto, S., Yousif, J. H., & Al-Bahri, M. (2023). Fake Face Detection Based on Colour Textual Analysis Using Deep Convolutional Neural Network. *Journal of Internet Services and Information Security*, *13*(3), 143-155. https://doi.org/10.58346/JISIS.2023.I3.009

[3] Al-Rifaee, Z. I. A., Ismaeel, T. Z., & Abood, S. I. (2024). Cryptography based on Fingerprint Bio Metrics. *Journal of Internet Services and Information Security*, *14*(4), 401-417. https://doi.org/10.58346/jisis.2024.i4.025

[4] Artan, Y., & Semiz, B. A. (2024). Fusion of Minutia Cylinder Codes and Minutia Patch Embeddings for Latent Fingerprint Recognition. https://doi.org/10.48550/arXiv.2403.16172

[5] Bao, Q., Wang, Y. G., Gao, C., Sha, L., & Lee, F. (2023). A Coarse-to-Fine Approach for Rectifying Distorted Latent Fingerprints from Crime Scenes. *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3344465

[6] Bhilavade, M. B., Shivaprakasha, K. S., Patil, M. R., & Admuthe, L. S. (2024). Fingerprint Reconstruction: Approaches to Improve Fingerprint Images. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *15*(1), 75-87. https://doi.org/10.58346/JOWUA.2024.I1.006

[7] Chai, J., Zeng, H., Li, A., & Ngai, E. W. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, *6*, 100134. https://doi.org/10.1016/j.mlwa.2021.100134

[8] Chegur, P., Patil, N., Doddamani, N., Thakkannavar, P., Pendari, N. T., Maralappanavar, M., & Ingalagaon, S. (2023, April). Separation of Overlapped Fingerprint Images using Deep Learning. In *2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS)* (pp. 233-238). IEEE. https://doi.org/10.1109/ICAECIS58353.2023.10169966

[9] Chhabra, M., Shukla, M. K., & Ravulakollu, K. K. (2020). State-of-the-art: A systematic literature review of image segmentation in latent fingerprint forensics. *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)*, *13*(6), 1115-1125. https://doi.org/10.2174/2213275912666190429153952

[10] Deshpande, U. U., & Malemath, V. S. (2021). MINU-EXTRACTNET: automatic latent fingerprint feature extraction system using deep convolutional neural network. In *Recent Trends in Image Processing and Pattern Recognition: Third International Conference, RTIP2R 2020, Aurangabad, India, January 3–4, 2020, Revised Selected Papers, Part I 3* (pp. 44-56). Springer Singapore. https://doi.org/10.1007/978-981-16-0507-9_5

[11] Deshpande, U. U., Malemath, V. S., Patil, S. M., & Chaugule, S. V. (2020). CNNAI: a convolution neural network-based latent fingerprint matching using the combination of nearest neighbor arrangement indexing. *Frontiers in Robotics and AI*, *7*, 113. https://doi.org/10.3389/frobt.2020.00113

[12] Deshpande, U. U., Malemath, V. S., Patil, S. M., & Chaugule, S. V. (2022). Automatic latent fingerprint identification system using scale and rotation invariant minutiae features. *International Journal of Information Technology*, *14*(2), 1025-1039. https://doi.org/10.1007/s41870-020-00508-7

[13] Dhaneshwar, R., Kaur, M., & Kaur, M. (2021). An investigation of latent fingerprinting techniques. *Egyptian Journal of Forensic Sciences*, *11*(1), 33. https://doi.org/10.1186/s41935-021-00252-4

[14] Duan, Y., He, K., Feng, J., Lu, J., & Zhou, J. (2022, March). Estimating 3D finger pose via 2D-3D fingerprint matching. In *Proceedings of the 27th International Conference on Intelligent User Interfaces* (pp. 459-469). https://doi.org/10.1145/3490099.3511123

[15] Farhang, A., & Rashidi, H. (2015). A modified fingerprint watermarking to improve security in wireless networks. *International Academic Journal of Science and Engineering*, *2*(2), 95–108.

[16] Garg, R., Singh, G., Singh, A., & Singh, M. P. (2024). Fingerprint recognition using convolution neural network with inversion and augmented techniques. *Systems and Soft Computing*, *6*, 200106. https://doi.org/10.1016/j.sasc.2024.200106

[17] Gibb, C., & Riemen, J. (2023). Toward better AFIS practice and process in the forensic fingerprint environment. *Forensic Science International: Synergy*, *7*, 100336. https://doi.org/10.1016/j.fsisyn.2023.100336

[18] Grosz, S. A., & Jain, A. K. (2023). Latent fingerprint recognition: Fusion of local and global embeddings. *IEEE Transactions on Information Forensics and Security*. https://doi.org/10.1109/TIFS.2023.3314207

[19] Grosz, S. A., Engelsma, J. J., Ranjan, R., Ramakrishnan, N., Aggarwal, M., Medioni, G. G., & Jain, A. K. (2022). Minutiae-guided fingerprint embeddings via vision transformers. https://doi.org/10.48550/arXiv.2210.13994

[20] Gu, S., Feng, J., Lu, J., & Zhou, J. (2022). Latent fingerprint indexing: Robust representation and adaptive candidate list. *IEEE Transactions on Information Forensics and Security*, *17*, 908–923. https://doi.org/10.1109/TIFS.2022.3154296

[21] Guan, X., Duan, Y., Feng, J., & Zhou, J. (2023). Regression of dense distortion field from a single fingerprint image. *IEEE Transactions on Information Forensics and Security*. https://doi.org/10.1109/TIFS.2023.3296310

[22] Jia, Z., Huang, C., Wang, Z., Fei, H., Wu, S., & Feng, J. (2024). Finger Recovery Transformer: Towards Better Incomplete Fingerprint Identification. *IEEE Transactions on Information Forensics and Security*. https://doi.org/10.1109/TIFS.2024.3419690

[23] Khan, A. I., & Wani, M. A. (2019). A common convolutional neural network model to classify plain, rolled and latent fingerprints. *International Journal of Biometrics*, *11*(3), 257-273. https://doi.org/10.1504/IJBM.2019.100843

[24] Kilinc, I., Artan, Y. O., & Başeski, E. (2023). A multistep fusion matcher approach for large scale latent fingerprint/palmprint recognition. *Turkish Journal of Electrical Engineering and Computer Sciences*, *31*(2), 412-430. https://doi.org/10.55730/1300-0632.3992

[25] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*.

[26] Kumar, S. S., Vinoth, R., Kumar, R., Kumar, M. J., Prasad, D. V. S. S. S. V., Prabhu, S., & Ram, S. (2024). Investigation of tempered AISI 420 SS under dry slided conditions for sustainable applications. In *E3S Web of Conferences* (Vol. 552, p. 01007). EDP Sciences. https://doi.org/10.1051/e3sconf/202455201007

[27] Kumar, T. S. (2024). Security Challenges and Solutions in RF-Based IoT Networks: A Comprehensive Review. *SCCTS Journal of Embedded Systems Design and Applications*, *1*(1), 16-19.

[28] Kustiawan, U. (2018). Character Value Education in Cirebon Mask. *International Academic Journal of Social Sciences*, *5*(2), 91–98. https://doi.org/10.9756/IAJSS/V5I2/18100029

[29] Megha, C., Kumar, S. M., & Kumar, R. K. (2021). Intelligent optimization of latent fingerprint image segmentation using stacked convolutional autoencoder. *International Journal of Performability Engineering*, *17*(4), 379. https://doi.org/10.23940/ijpe.21.04.p6.379393

[30] Öztürk, H. İ., Selbes, B., & Artan, Y. (2022). Minnet: Minutia patch embedding network for automated latent fingerprint recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1627-1635). https://doi.org/10.1109/cvprw56347.2022.00169

[31] Pan, Z., Duan, Y., Guan, X., Feng, J., & Zhou, J. (2024). Latent Fingerprint Matching via Dense Minutia Descriptor. https://doi.org/10.48550/arXiv.2405.01199

[32] Prabakaran, E., & Pillay, K. (2021). Nanomaterials for latent fingerprint detection: a review. *Journal of materials research and technology*, *12*, 1856-1885. https://doi.org/10.1016/j.jmrt.2021.03.110

[33] Qiu, Y., Chen, H., Dong, X., Lin, Z., Liao, I. Y., Tistarelli, M., & Jin, Z. (2024). Ifvit: Interpretable fixed-length representation for fingerprint matching via vision transformer. https://doi.org/10.48550/arXiv.2404.08237

[34] Seidlitz, S., Jürgens, K., Makrushin, A., Kraetzer, C., & Dittmann, J. (2021, February). Generation of Privacy-friendly Datasets of Latent Fingerprint Images using Generative Adversarial Networks. In *VISIGRAPP (4: VISAPP)* (pp. 345-352). https://doi.org/10.5220/0010251603450352

[35] Shreya, S., & Chatterjee, K. (2024). Gan-enable latent fingerprint enhancement model for human identification system. *Multimedia Tools and Applications*, *83*(9), 27565-27588. https://doi.org/10.1007/s11042-023-16510-7

[36] Simonyan, K. (2014). Very deep convolutional networks for large-scale image recognition.

[37] Thanaraj, M. S., Poomalai, R., Vibin, R., Kamalakar, V., Gupta, S., Perumal, M. P. R., & Vadivel, M. (2024). Sustainable energy integration for seawater desalination. *Journal of Environmental Protection and Ecology*, *25*(6), 2086–2096.

[38] Trisiana, A. (2024). A Sustainability-Driven Innovation and Management Policies through Technological Disruptions: Navigating Uncertainty in the Digital Era. *Global Perspectives in Management*, *2*(1), 22-32.

[39] Valdes Ramirez, D. (2021). A classifier-based fusion algorithm for latent fingerprint identification based on a neural network.

[40] Wahab, A., Khan, T. M., Iqbal, S., AlShammari, B., Alhaqbani, B., & Razzak, I. (2024). Latent fingerprint enhancement for accurate minutiae detection. *Procedia Computer Science*, *246*, 1558-1567. https://doi.org/10.1016/j.procs.2024.09.722

[41] Yoo, D., Cho, J., Lee, J., Chae, M., Lee, B., & Lee, B. (2020). FinsNet: end-to-end separation of overlapped fingerprints using deep learning. *IEEE Access*, *8*, 209020-209029. https://doi.org/10.1109/ACCESS.2020.3038707

[42] Zhu, Y., Yin, X., & Hu, J. (2023). Fingergan: a constrained fingerprint generation scheme for latent fingerprint enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(7), 8358-8371. https://doi.org/10.1109/TPAMI.2023.3236876

Authors Biography

Poornima E Gundgurti is an accomplished academic professional serving as an Assistant Professor in the Department of Computer Science at the Central University of Karnataka, India. With six years of dedicated teaching experience and made a significant contribution to educating both master's and undergraduate students, fostering their knowledge and skills in computer science. Currently in the second year of tenure at the Central University of Karnataka, Poornima E. Gundgurti continues to demonstrate a passion for research and teaching excellence. Before joining this institution, Poornima E. Gundgurti served as a Lecturer at the VTU Regional Centre, Kalaburagi, for four years. In addition to teaching, Poornima E. Gundgurti has a strong research background, having published two international conference papers and two international journal papers. These contributions reflect Poornima E. Gundgurti 's commitment to advancing the field of computer science and sharing innovative insights with the global academic community.

Dr. Shrinivasrao B Kulkarni is an Associate Professor in the department of Computer Science and Engineering at SDM College of Engineering and Technology, Dharwad, Karnataka, INDIA. Completed his undergraduate from BEC, Bagalkot, M.Tech. from BVBCET, Hubli and Ph.D. from Graphic Era University, Dehradun, Uttarakhanda. Four Research Scholars have completed Ph.D. under his guidance from Visvesvaraya Technological University (VTU) Belagavi and five are pursuing Ph.D. under VTU Belagavi. Guided many U.G. and PG. students. His are of Interest are Image Processing, Software Engineering. Medical Imaging, Context Aware, Social Media, Web, Networking and Communication. He has worked as HoD Computer Science and Engineering, Academic Council, BoS, DUGC, DPGC, PBSA, Web etc at SDMCET, Dharwad. BoS and BOE member at KSRDPRU, Gadag. BoE Member VTU, Belagavi. Research Advisory Committee member for research scholars in the affiliated colleges of VTU, Belagavi. He has published 52 papers in Journal and conferences at national and international level and 6 Patents. Reviewer for IUCE, VTU Journal, IEEE, RGD, ASTES, Evolving Systems-Springer, Machine Vision and Applications. Conference Chair for IEEE – International Conference for Emerging Technology (INCET – 2020). Presented Paper in ISMS2018 Conference at University of Malta, Europe. Conducted AICTE Sponsored Short Term Training programme on Research Trends in Computer Science and Engineering in two different phases. He is a life member of ISTE, IEI and International Association of Engineers and also member, IEEE and ACM in 2020. He has delivered 12 invited talks. Grant Sanctioned 2,53,167 -00 – AICTE Sponsored STTP– AQIS 2020 for the Title “Research Trends in computer Science and Engineering on 10-10-2020 and conducted Short Term Training Program in two phases at SDMCET, Dharwad during 2020-21. Published book on Neural Networks and Deep Learning, SIPH, Year 2022, He has received Best Teacher Award from Teacher’s Mitra Trust, Hubli, Dharwad, Karnataka.