

**A MAJOR PROJECT-II REPORT**  
**ON**  
**ENHANCED DATA HIDING BY SEGMENTATION FOR**  
**MEDICAL IMAGES**

Submitted in Partial Fulfilment of the Requirements for the  
Award of the Degree of

**MASTER OF TECHNOLOGY**  
**IN**  
**COMPUTER SCIENCE & ENGINEERING**

Submitted By:

**RASIKA GUPTA**

**2K22/CSE/29**

Under the Supervision of

**Dr. RAJEEV KUMAR**

**Assistant Professor**



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**  
**DELHI TECHNOLOGICAL UNIVERSITY**

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

JUNE, 2024



**DELHI TECHNOLOGICAL UNIVERSITY**  
(Formerly Delhi College of Engineering)  
Shahbad Daultapur, Main Bawana Road, Delhi - 110042

## **CANDIDATE'S DECLARATION**

I, Rasika Gupta (2K22/CSE/29) hereby certify that the work which is being presented in the Major Project-II entitled “**Enhanced Data Hiding by Segmentation for Medical Images**” in partial fulfilment of the requirements for the award of the Degree of Master of Technology, submitted to the Department of Computer Science and Engineering, Delhi Technological University is an authentic record of my own work carried out under the supervision of Dr. Rajeev Kumar.

The matter presented in the Major Project-II report has not been submitted by me for the award of any other degree of this or any other Institute.

Place: Delhi Technological University, Delhi

Date: June, 2024

Rasika Gupta

(2K22/CSE/29)



## DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daultapur, Main Bawana Road, Delhi - 110042

### CERTIFICATE

Certified that, **Rasika Gupta** (2K22/CSE/29) has carried out the research work presented in this Major Project II report entitled “**Enhanced Data Hiding by Segmentation for Medical Images**” for the award of **Master of Technology** from Department of Computer Science and Technology, Delhi Technological University, Delhi, under my supervision. The project embodies results of original work, and studies have been carried out by the student herself. The contents of the report do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other Institution.

Place: Delhi Technological University, Delhi

Date: June, 2024

Dr. Rajeev Kumar

Assistant Professor

Department of CSE

Delhi Technological University

## ABSTRACT

Due to the large amount of data flowing across internet, there is a need to safeguard the transmission of sensitive data, specifically in the field of utmost security and privacy. Data hiding significantly contributes to protect the sensitive information by embedding the data inside a cover media in such a way that is undetectable to human eyes. Due to the low contrast of medical images, it often suffers from distortion after embedding the data. In order to resolve this, most of the researchers incorporate image segmentation to segment the medical image into Region Of Interest (ROI) and NROI (Non-Region Of Interest). This helps in performing the selective embedding based on the features of different regions. However, existing methods suffer from accurately separating the ROI from the rest of the medical image, which can impact the data hiding performance.

Therefore, in order to solve this problem, this report introduces an improved UNet 3+ architecture, which is based on UNet framework. It overcomes the discrepancies faced by the base UNet model with the ground truth segmented mask. This enhanced network integrates residual blocks at both encoder and decoder to handle gradient degradation problem in deep neural network, incorporates attention module to focus on relevant features and reduces redundant skip connections to reduce computational complexity. The improved model has been evaluated on publicly available PH2 skin lesion dataset and achieved an accuracy of 96.79 %, dice score of 93.72 % and jaccard index of 96.97 % on test data. The model surpasses other state-of-the-art UNet models for skin lesion segmentation.

After segmenting the ROI, a conventional data hiding method has been used for selective embedding based on the local complexity of the pixels. This local complexity helps in embedding only in smoother region of the ROI to reduce the distortion. The embedding capacity is kept low to avoid affecting the most critical region but the visual quality of stego image has been enhanced with PSNR of 33.08 dB and SSIM value of 0.9893.

## ACKNOWLEDGEMENT

It is my profound privilege to convey my sincere thanks and gratitude to Professor Dr. Vinod Kumar, Head, Department of Computer Science and Engineering, Delhi Technological University, Delhi for the kind opportunity to work on this M.Tech. Project. He extended necessary facilities and resources of the department for doing this project appropriately and submit the same timely.

I express sincere heartfelt thanks to my supervisor Dr. Rajeev Kumar, Assistant Professor, Department of Computer Science and Engineering, Delhi Technological University, Delhi for his kind guidance, motivation and supervision in undertaking the project work in a planned systematic manner right from its inception to completion. He spared his precious time to extend continuous support and cooperation in carrying out different activities of the project as well as in preparation of the report.

Last but not the least, I owe my special thanks to my parents without whose blessings, I may not have been able to do this work. They have supported, encouraged and invigorated me in every possible manner.

Place: Delhi Technological University, Delhi  
Date: June, 2024

Rasika Gupta  
(2K22/CSE/29)

## TABLE OF CONTENTS

<i>Candidate's Declaration</i> .....	<i>ii</i>
<i>Certificate</i> .....	<i>iii</i>
<i>Abstract</i> .....	<i>iv</i>
<i>Acknowledgement</i> .....	<i>v</i>
 <i>List of Tables</i> .....	 <i>viii</i>
<i>List of Figures</i> .....	<i>ix</i>
<i>List of Abbreviations</i> .....	<i>x</i>
 Chapter 1: Introduction.....	 1-14
1.1 Brief Overview.....	1
1.2 Background.....	2
1.2.1 Image Segmentation .....	2
1.2.2 Data Hiding.....	10
1.2.3 Application of Segmentation based Data Hiding.....	12
1.3 Quality and Security Concerns in Medical Images.....	12
1.4 Problem Statement.....	13
1.5 Objectives of the Project.....	14
Chapter 2: Literature Review.....	16-26
2.1: Related work on Skin Lesion Segmentation .....	17
2.2: Related work on Segmentation based Data Hiding .....	18
Chapter 3: Proposed Methodology.....	27-40
3.1 Model Architecture of Proposed Segmentation Model.....	28
3.2 ROI Embedding .....	39
Chapter 4: Experimental Results and Analysis .....	41-52
4.1 Tools and Libraries used .....	41
4.2 Dataset.....	43
4.3 Pre-Processing .....	43
4.4 Evaluation Metrics .....	44
4.5. Implementation Details .....	46

4.6 Results of the Proposed System.....	47
4.6.1 Ablation Study.....	47
4.6.2 Qualitative Performance.....	48
4.6.3 Quantitative Performance.....	50
4.6.4 Comparison with the State-of-the-art.....	52
Chapter 5: Conclusions and Future Works.....	53-54
5.1 Conclusions .....	53
5.2 Future Works.....	54
References.....	55-60
List of Publications.....	61

## LIST OF TABLES

Table No.	Title	Page No.
Table 2.1	Summary of Literature Review	23-26
Table 3.1	Model Parameters	33-38
Table 4.1	Data Augmentation Technique	43
Table 4.2	Ablation Study on Different Modules	47
Table 4.3	Effect of different batch sizes and learning rate on model over PH2 dataset	47
Table 4.4	Proposed Model Performance	51
Table 4.5	Comparison of Proposed Model with other Models	51
Table 4.6	Performance of data hiding method	52
Table 4.7	Comparison of Proposed Segmentation Model with State-of-the-Art	52



## LIST OF FIGURES

Figure No.	Title	Page No.
Fig. 1.1	Types of Image Segmentation	2
Fig. 1.2	Classification of Image Segmentation Techniques	3
Fig. 1.3	GrabCut Segmentation	5
Fig. 1.4	SVM Classifier	6
Fig. 1.5	FCN-AlexNet	8
Fig. 1.6	SegNet Architecture	9
Fig. 1.7	UNet Architecture	9
Fig. 1.8	GAN Basic Model	10
Fig. 1.9	Data Hiding	10
Fig. 3.1	Flowchart of the Proposed System	27
Fig. 3.2	Proposed Segmentation Model: Enhanced UNet 3+ Architecture	28
Fig. 3.3	Convolutional Block Attention Module (CBAM)	32
Fig. 3.4	Cross and Dot Set	39
Fig. 4.1	Image Augmentation Techniques on Skin lesion Images	43
Fig. 4.2	The left graph shows the proposed model's training and validation loss, while the right graph shows the proposed model's training and validation accuracy on the PH2 dataset.	46
Fig. 4.3	Qualitative Performance of the Proposed Segmentation Model	48
Fig. 4.4	Skin lesion images alongside their ground truth in first and second column respectively. The predicted segmented mask by UNet by O. Ronneberger et al. (2015) and the Proposed Model in third and fourth column followed by ROI segmented image of the proposed model in fifth column	49
Fig. 4.5	Visual Comparison between Original image and Embedded Image	50

## LIST OF ABBREVIATIONS

S.No.	Abbreviation	Full Form
1	ROI	Region Of Interest
2	NROI	Non- Region Of Interest
3	FCN	Fully Convolutional Network
4	CNN	Convolutional Neural Network
5	SVM	Support Vector Machine
6	ATD	Adaptive Threshold Detector
7	CRF	Conditional Random Field
8	MRF	Markov Random Field
9	ReLu	Rectified Linear Unit
10	MRI	Magnetic Resonance Imaging
11	CT	Computed Tomography
12	API	Application Programming Interface
13	GPU	Graphics Processing Unit
14	TPU	Tensor Processing Units
15	GAN	Generative Adversarial Network
16	PSNR	Peak Signal to Noise Ratio
17	SSIM	Structural Similarity Index Measure
18	BN	Batch Normalization
19	DWT	Discrete Wavelet Transform
20	SVD	Singular Value Decomposition
21	LSB	Least Significant Bit
22	RDH	Reversible Data Hiding
23	IARC	International Agency for Research on Cancer
24	WHO	World Health Organization
25	FCRN	Fully Convolutional Residual Network
26	CBAM	Convolutional Block Attention Module

# **CHAPTER 1**

## **INTRODUCTION**

In this chapter, a brief overview has been presented to introduce the significance of image segmentation and data hiding in medical images, followed by a detailed explanation of different image segmentation and data hiding techniques. After the theoretical explanation, the background details of the existing image quality and security problem have been discussed before presenting the problem statement of the proposed project work. The chapter concluded by highlighting the objectives of the project and the project organization details.

### **1.1 BRIEF OVERVIEW**

In the current digital era, it is vital to safeguard sensitive and significant data, especially in the medical sciences where patient confidentiality is of utmost importance. In this regard, data hiding has emerged as a powerful technique that addresses this need by embedding information within a host media in a way that is invisible to the naked eyes. Since the current work aims at data hiding in low contrast medical images with an intention of improving medical data security and privacy without distorting the image much as it is mostly applied in the medical field, it is important to improve its quality.

<sup>2</sup> The primary purpose of using the medical images as host media is to protect the large and sensitive medical information of patients. This is very dangerous as private information of patient can be exposed, hence, lead to high risk of invasion of patient's privacy and reduced trust.

<sup>3</sup> The current work has explored the significance of image segmentation in the field of data hiding, techniques to increase segmentation accuracy, and methods to effectively implement data hiding for medical imaging. This research seeks to contribute to the ongoing efforts in the domain of medical data security and confidentiality as explored the significance of image segmentation in data hiding field, techniques to increase the segmentation accuracy and method to implement hiding of data for the medical imaging effectively.

## 1.2 BACKGROUND

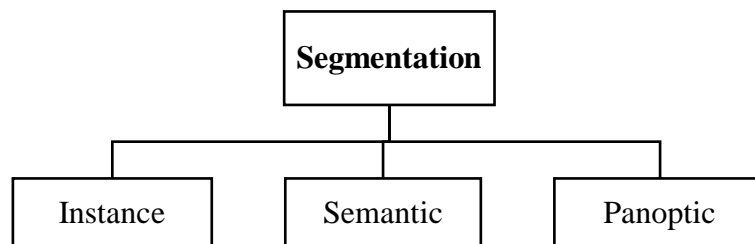
In this section, a comprehensive discussion on image segmentation is provided, detailing its various types and techniques. The section also offers an overview of the concept of data hiding, including its different types. Finally, the applications of segmentation-based data hiding are explored in various contexts.

### 1.2.1 IMAGE SEGMENTATION

Image segmentation is an important aspect of image processing in which an image is divided into homogeneous regions with an objective to get the ROI from its surrounding background (NROI). It is the process of identifying specific regions within an image that are considered important for analysis. These regions could include objects, structures, and areas of high importance in the context of a particular application. In the context of biomedical imaging, segmentation helps in proper diagnosis and timely treatment planning by isolating the infected area from irrelevant background details.

#### • TYPES OF SEGMENTATION

Image segmentation is divided into three broad categories as shown Fig. 1.1 based on the type and amount of information that are extracted from the image:



**Fig. 1.1: Types of Image Segmentation**

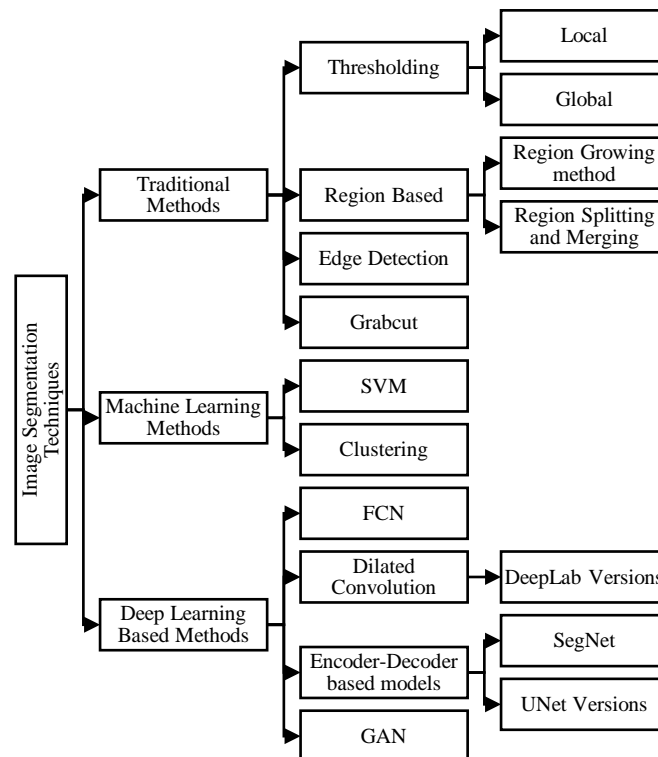
**Instance Segmentation** detects and segments each object in an image. It works like object detection with the goal of segmenting the boundaries of different objects. The algorithm has no idea about the class of the region, it's just separates the overlapping objects. **Semantic segmentation** is the process of assigning a class label to every pixel in an image. It includes classification, which makes prediction for whole input. Then, localization, which provide different classes with the additional information regarding

the spatial location of those classes. At last, fine-grained inference is achieved by making dense predictions labels for every pixel. The **Panoptic Segmentation** is a grouping of both. It gives class label to each pixel and then identify each object instance in the image. This type of segmentation is used in computer vision applications where the model need to detect and interact with the different objects in the surrounding world like autonomous robots.

The selection of which type of segmentation will be used depend upon the application. For example, for diagnosing disease in a medical image, precise identification and localization of abnormalities is needed, which can be achieved by using semantic segmentation.

## • IMAGE SEGMENTATION TECHNIQUES

Various image segmentation methods has been displayed in Fig. 1.2, segregated based on the underlying approach.



**Fig. 1.2: Classification of Image Segmentation Techniques**

Over a period of time, different techniques for each type of image segmentation, beginning with the classic traditional methods which primarily concentrated on extracting information from images, often required human intervention. But extracting

high-level semantic information from images proved to be challenging. Hence, it was followed with usage of Machine Learning (ML) models that helped in preserving semantic information. But these algorithms were unable to provide good results with large datasets. In modern years, Deep Learning (DL) techniques have gained popularity due to large size image datasets and provided an opportunity to deal with such challenges.

## • **TRADITIONAL METHODS**

These segmentation techniques depend upon two essential aspects within images, i.e. discontinuity and similarity. The different types of traditional techniques have been briefly described below:

### **1) Thresholding**

It is one of the primary method for image segmentation. A threshold value is used to divide an image into distinct regions.

1. *Global Thresholding*- In this method, only one threshold value is used, as the difference between the foreground and background pixel is very large.
2. *Local Thresholding*-In this method, different regions of an image are segmented by using different threshold values.

### **2) Region Based Segmentation**

This method works by comparing pixels with its neighbouring pixels on the basis of some parameters like color, texture, intensity and then divides the region based on similarities.

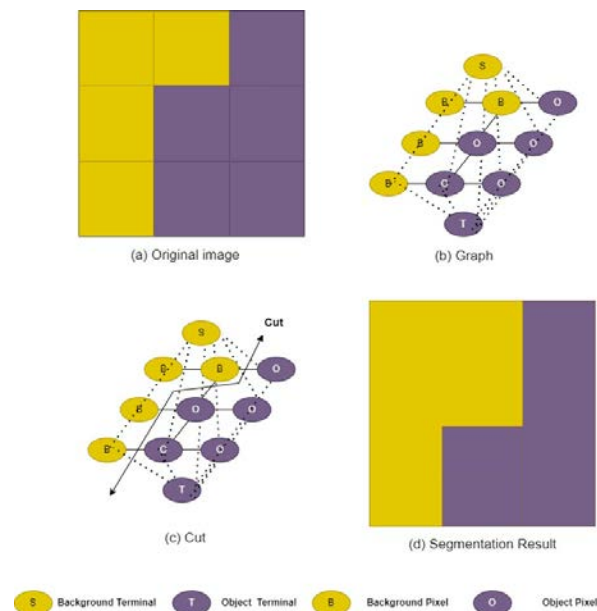
1. *Region Growing Method*- In this method, some pixels are selected as a seed particles. Then, the seed particles are assigned to the same region, if they have the same characteristics. This process continues until all the seed particles which has the same value assigned to a same region.
2. *Region Splitting and Merging*- Quad-tree based segmentation is an example of this method, where the region is partitioned into four quadrants and then merging is done until no further splitting is possible.

### 3) Edge Detection

Objects in an image are outlined with clear edges and corners, making sure each object has its own space separate from others, with noticeable lines or borders in between. The objective of edge detection method is to find these edges. Edge detection method identifies and separates the edges of an image. It marks the boundaries of the objects. Some of the edge detector used are Canny edge detector, Sobel edge detector, and Shen-Castan detector.

### 4) GrabCut

GrabCut is a widely utilized graph-based image segmentation technique renowned for its simplicity and accuracy in implementation. Its fundamental concept involves representing an image as a network diagram (Graph) as described in Fig. 1.3.



**Fig. 1.3: GrabCut Segmentation**

Initially, the user is required to designate foreground and background samples. The method then calculates the distance between image points based on the user-supplied foreground and background pixels, as well as the distance between the image points themselves. Further, the energy weight values for each of the edge of a graph is calculated by combining these distances linearly. Then, the minimum energy value for splitting the edges within the graph is find out to get the resultant segmented graph. As

a result of this methodology, GrabCut tends to produce segmentation outcomes that are more comprehensive and exhibit natural boundaries compared to other methods.

## • MACHINE LEARNING METHODS

The machine learning algorithm is capable of learning from training data and automatically extracting meaningful features. The two mostly used models Support Vector Machine (SVM) and Clustering have gained importance in the field of image segmentation. SVM is a supervised learning algorithm which requires labelled data whereas Clustering is an unsupervised learning algorithm which identifies patterns and groups within unlabelled dataset.

### 1) Support Vector Machine (SVM)

As demonstrated in Fig. 1.4, a SVM classifier, can be used for image segmentation. Class 1 can be the object in an image and Class 2 can be the background of the image. Support Vectors are the points that are deemed to be nearest to hyperplane, act as a separator between the two classes. They are used for finding the best hyperplane that would be used to classify data in one class from the other in the feature space. SVM is trained using features vectors like gradient information, texture descriptors, color histogram of images and prepared a labelled dataset with corresponding ground truth. A suitable kernel function is selected based on the data characteristics. For every pixel, SVM classifier determine whether it is the foreground / background pixel. Once trained, this classifier can be used to determine segmentation masks for new images.

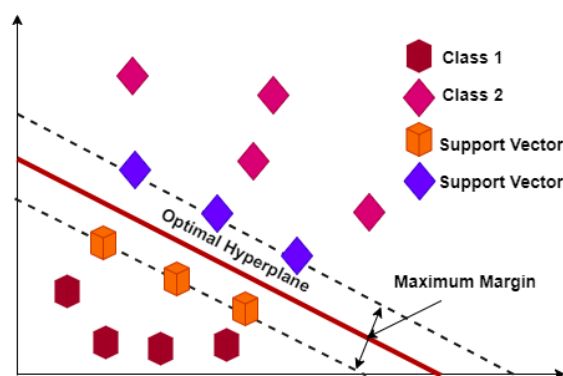


Fig. 1.4: SVM Classifier



## **2) Clustering**

Clustering is a technique where entities are grouped together into clusters based on their shared characteristics. Each cluster is made up of data points that are more similar to one another than those in other clusters. Most commonly used clustering technique used for segmentation is Fuzzy based clustering. It combines geometric analysis with fuzzy logic. Radial lines are drawn from a central point within an image. Along these lines, the significant changes in image characteristics such as intensity, color or texture are find out using the membership function of fuzzy logic, which assign a degree of belonging to each point based on these characteristics. These changes help in separating boundaries or Region of Interest. This approach enhances partition compactness in segmentation results by ensuring contiguous regions based on spatial relationships and similarity to characteristics features.

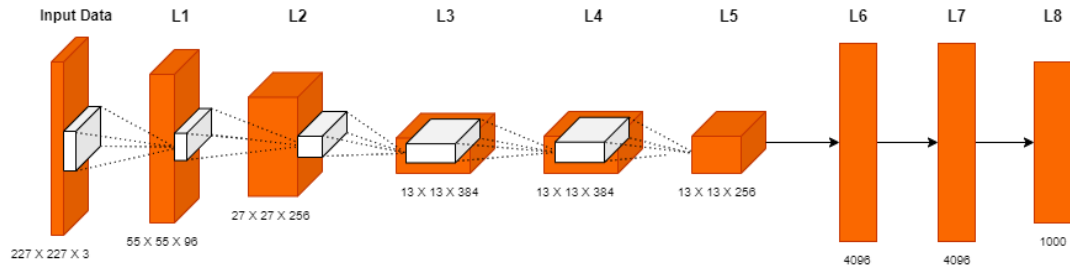
## **• DEEP LEARNING METHODS**

Image segmentation techniques have been greatly impacted by the use of Deep Learning (DL) models in the past few years and these techniques are powered with high accuracy. It uses training of neural networks for image segmentation so that they can identify which features are important in an image. These technique do not rely on the customized function like in traditional methods. These methods provide good results in terms of accuracy and time. Deep learning methods work exceptionally well with large datasets. This section discusses about some of the most effective DL models for image segmentation.

### **1) Fully Convolutional Network (FCN)**

FCN's convolutional layers generate a segmentation map that matches the input image of the same size. It is used as a baseline model to develop other neural networks for segmentation based on deep learning. The most common FCN network used for image segmentation is AlexNet displayed in Fig. 1.5. The first two layer L1 and L2 are the convolution, ReLu, Max Pooling and normalization. The third L3 and fourth L4 layer performed only convolution and ReLu operations. L5 is same as L1 except it does not perform normalization. It also convert the output to the long vector for the L6 and L7

layers, which are the fully connected layer. The Final layer is L8 or Softmax layer which gives a label for each pixel of the input image.



**Fig. 1.5: FCN-AlexNet**

## 2) Dilated Convolutional Segmentation model

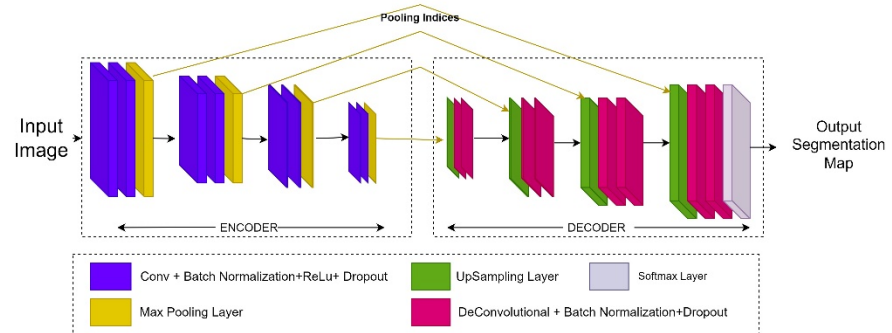
In this model, a dilation rate gets added to the convolutional layer. This increases the receptive field size without any increase in computational cost. One of the most prominent dilated convolutional model used for image segmentation is DeepLab family. These models are combination of Convolutional Network and probabilistic models like CRFs and MRFs.

## 3) Encoder-Decoder Models

These are the convolution network consisting of two parts- i) encoder part, made up of convolution and pooling layer which extract high-dimensional features maps containing the semantic information and ii) decoder part generates segmentation masks using deconvolutional and unpooling layers. Two most commonly used encoder-decoder models are as follows:

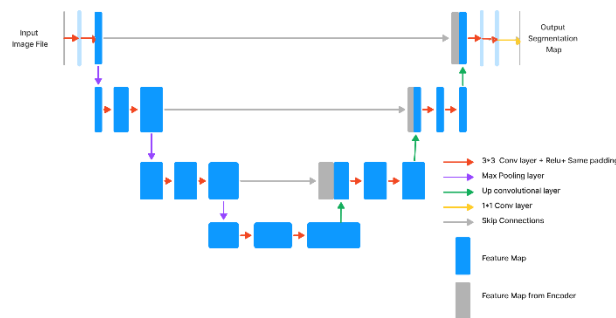
1. **SegNet**- The convolutional encoder-decoder model is employed for semantic segmentation at pixel level. SegNet as shown in Fig. 1.6, consist of encoder network, made up of simple convolution, followed by BN and ReLu activation function. Dropout Layer has been used in the model to prevent overfitting. Each encoder block has a corresponding decoder block. A Softmax layer is employed at the network's end for final pixel-wise classification to generate the segmentation mask. MaxPooling layer maintains translation invariance with respect to small shifts in the spatial position of features within the input image. During the UpSampling, the output of

MaxPooling layer from the corresponding encoder block are used but it ignores neighboring information during decoding.



**Fig. 1.6: SegNet Architecture**

2. **UNet**- To enhance the accuracy of pixel positioning, skip connections were introduced in encoder-decoder model. As deep neural networks grow in depth, their performance tends to decrease. UNet is a novel approach called long skip connections model that transmit and cascade features from encoder to decoder layer that corresponds them, enabling the capture of fine-grained image details as illustrated as Fig. 1.7.



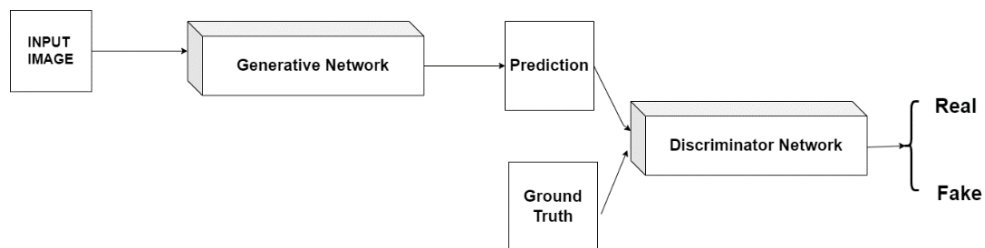
**Fig. 1.7: UNet Architecture**

UNet has become widely adopted in medical image segmentation research. The U-Shaped structure help in capturing fine details at low-level spatial dimension. U-Net can be used for low training data, because the skip connections allow the use of feature maps from contracting path directly into the expansive path.

#### 4) Generative Adversarial Neural Models

Adversarial Neural Models consists of two networks as presented in Fig. 1.8, a Generator and a Discriminator, which contend with one another in order to get better

their performance. The generator which is the input produces data by taking it and synthetic data samples while the discriminator or the binary classifier in the model is able to distinguish the real data sample with fake data samples. During training, generator tries to fool discriminator by providing indistinguishable samples, while the discriminator tries its ability to discern between authentic and counterfeit samples. GAN proved as a powerful framework to work with realistic data.



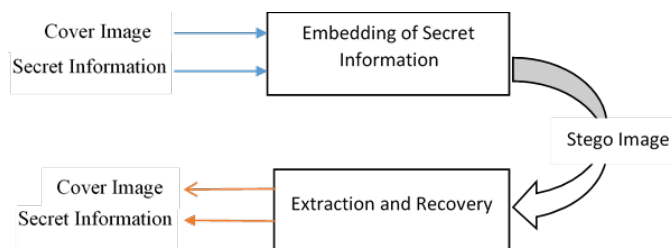
**Fig. 1.8: GAN Basic Model**

These Generative and Discriminator network can be built on any suitable SOTA architecture depending on the application and performance.

The application and the degree to which a given segmentation technique aids in precise image segmentation determine which technique is best.

## 1.2.2 DATA HIDING

It involves concealing information in a host medium to prevent uninvited recipients from finding it. The host media can be restored without any data loss after extracting the embedded data from Data Hiding. The conceptual diagram has been shown in Fig. 1.9.



**Fig. 1.9: Data Hiding**

It is applied to numerous fields, including law, medical, and military imagery. Two key characteristics of Data Hiding —its ability to embed data and its imperceptibility—are used to gauge its effectiveness. Imperceptibility is a measure of how closely the cover and stego (image after extraction) images match. The payload or embedded capacity describes the amount of information that can be hidden in the cover image. Capacity and imperceptibility are frequently traded off; extra embedding may raise the chance of detection or host medium deterioration. The main goal is to maintain this trade-off to ensure robust data hiding.

The following are the three techniques used for hiding data:

- **Steganography**

The technique of embedding data in pictures, sounds, or videos is used to hide its existence. The most widely used technique is called Least Significant Bit (LSB), and it involves changing an image's pixel values' least significant bit.

- **Digital Watermarking**

This method is used to certify authenticity or ownership of digital media by embedding information (watermark) into it. The watermark should be resistant to different kinds of processing and attacks and can be either visible or invisible.

- **Reversible data hiding (RDH)**

RDH makes it possible to fully restore the original host medium following the extraction of the embedded data. This is important because original data integrity is critical in fields like medical imaging. Prediction Error Expansion given by Sachnev et al. [48]. is the most popular technique for reversible data hiding. It calculate the difference between the predicted value calculated using correlation of neighbouring pixels and the actual pixel selected for embedding. They are used for shifting the prediction error histogram pixels and this difference is known as the prediction error. It is done in this manner that the original picture and hidden message can be retrieved without any loss.

### **1.2.3 APPLICATION OF SEGMENTATION BASED DATA HIDING**

The Segmentation based Data Hiding approach is applicable in numerous areas of practice since it is capable of hiding data in particular areas of interest without the necessity of covering the entire image. Below are the some of the listed applications:

- **Medical Imaging**

Purpose: It may be necessary to hide some crucial clinical information such as patient's data or any other related data.

Application: It involves the process of placing the patient profile or other details of additional diagnosis into the segmented picture without distorting the overall image.

- **Forensic Imaging**

Purpose: There are some areas in an image which are more informative or significant compared to other areas of an image, for instance facial regions/marks of an object.

Application: Any data within the ROI but also concealing other information connected to investigation or any data within the ROI while still maintaining the integrity of the complete forensic image.

- **Satellite and Remote Sensing**

Purpose: The satellite images or images captured using remote sensing instruments may contain some areas of interest like geographical features.

Application: The other possible information such as the geospatial coordinates or region details to the ROI of the satellite images. This is especially helpful if one needs the privilege to be able to pull up data or analyse data that was saved in the database at some point in the past.

### **1.3 QUALITY AND SECURITY CONCERNS IN MEDICAL IMAGES**

The healthcare sector requires the storage, analysis, and retrieval of extensive historical medical imaging data over prolonged periods. It enhances the accessibility, sharing and coordination among the healthcare providers. Most healthcare fields maintain comprehensive individual clinical systems within localized data repositories encompassing directly attached storage or storage through network. It consolidates a

patient's medical history, diagnosis, treatment plans and other important information into a centralized form. This revolutionized healthcare record-keeping leads to more efficient and patient-centric care. This can help to overcome the risk of duplication of tests and treatments.

The urgent necessity for extensive and accumulating storage of copies of the primary and secondary images is causing a significant surge in "Big Data" within imaging. As a notable advancement, the healthcare sector is transitioning on the way to standardized methods for keeping and retrieving medical images, leading to a unified enterprise archive cutting across various disciplines. Recently, due to the large amount of storage, the medical industries store their data on the cloud. But because of privacy and security concerns, a lot of users are hesitant to use the cloud. Due to this reason, there is a growing emphasis on protecting the information from unauthorized access. In this regard, data hiding techniques have emerged as a potential way out for the protection of sensitive and critical information.

Several data-hiding schemes have been developed over the past two decades in order to overcome the security challenges, hidden information's imperceptibility, embedding ability and visual quality of images. These methods of data hiding sometimes provide poor performance on low contrast images.

## **1.4 PROBLEM STATEMENT**

Most of the medical images have low contrast than that of ordinary images, which has been considered as a problem in the diagnosis process due to poor visual quality. Therefore, an effective method needs to be developed that can segment the area that contains the infected/organ part (ROI) from the background area (NROI), followed by enhancing the ROI through effective embedding or contrast enhancement. The remaining data can be embedded inside the NROI part, if necessary. However, there are some challenges in accurate segmentation of the ROI from background particularly in the medical images with twisted contours. The accuracy of the segmentation process needs to be increased. Additionally, it is essential to evaluate the suitability of ROI for embedding of data. Effective data hiding is feasible only if the ROI contains smooth

pixels, which are less likely to distort the embedded information and the overall image quality. By focusing on these critical aspects—precise segmentation and thorough local complexity prediction—a more efficient data hiding method for medical images can be developed.

## 1.5 OBJECTIVES OF THE PROJECT

The primary objective of this project is to evaluate and compare various segmentation techniques to determine their effectiveness in data hiding within biomedical images through extensive review.

- **Comprehensive Review and Comparison:** To conduct comprehensive review and comparison so as to select an optimal technique for segmentation of medical image.
- **Enhancement of Segmentation Accuracy:** To identify the research gap in the selected segmentation technique and explore the methods to refine the technique for accurately segmenting the ROI of medical image from background region.
- **Application and Performance Evaluation:** To test the refined model performance on publicly available medical images in order to ensure its effectiveness and reliability in real-world situations.
- **Local Complexity Prediction:** To incorporate the statistical technique for computing local complexity of ROI pixels to determine the smooth pixels of the image for effective data hiding with less distortion.
- **Data Embedding:** To embed secret data inside the smooth pixels using well-known conventional data hiding technique in a way that is undetectable to human eyes.

By enhancing segmentation accuracy and streamlining data hiding procedures in biomedical images, the project is expected to make progress in the field of medical data security.

The project report has been divided into five chapters: Chapter 1 has discussed about the essence of segmentation in data hiding followed by the detailed explanation of



various segmentation and data hiding techniques. It also presented the problem statement of the project preceded by the existing security concerns in health sector, concluded by highlighting the main objectives of the proposed work. Chapter 2 presented the related work on skin lesion segmentation and on segmentation based data hiding. It presented a detailed summary on the existing technique, their drawback and the concluding paragraph on the proposed work to overcome the drawbacks. In Chapter 3, the detailed methodology of the proposed work has been discussed focussing on the architectural diagram, flowcharts and techniques used. It has been followed by Chapter 4, which discussed about the qualitative and quantitative experimental results and their analysis with the help of graphs, tables and figures. It also analysed the performance of the proposed model by comparing it with the existing methods. The conclusion and future work of the project has been discussed in Chapter 5 followed by the references.

## **CHAPTER 2**

### **LITERATURE REVIEW**

This chapter discusses about the state-of-the-art methods for skin lesion segmentation and data hiding and has been divided into two sections. Section 2.1 provides a brief overview of the existing technique for skin lesion segmentation with their performance on different datasets. It helps in finding the potential gap needed to design an efficient segmentation model for accurate segmentation of ROI. In the Section 2.2, the related work on segmentation based data hiding has been discussed to review the existing techniques and its shortcomings.

The purpose of using skin lesion medical images because Skin cancer is a serious public health concern across the world and is a deadly disease if not diagnosed at an early stage. It is often diagnosed through dermoscopic images that usually have low contrast, irregular boundaries and contains irrelevant artifacts. Due to these factors, the segmentation of Region of Interest (ROI) for effective diagnosis of this disease becomes a challenging task. According to IARC and WHO, there will be over 1.5 million new skin cancer cases worldwide in 2020. Thus, the data indicates a significant number of disease cases with a rising trend. Hence, it became very important to study skin cancer very well for its appropriate timely diagnosis and treatment. The scientists have recently predicted that more than 5,00,000 new cases of melanoma skin cancer per year may happen and around 1,00,000 deaths due to melanoma can be expected worldwide by 2040. The American Cancer Society predicted in 2023 that cancer survival rate would drop from over 99% to 32% when it is detected at an advanced stage. Early diagnosis of this cancer is critical for good prognosis. Dermatologists uses a dermatoscope to capture dermoscopic images for the detection of skin cancer, if any. However, the accuracy of the diagnostic process depends on the clinician's level of expertise and quality of medical image. Therefore, to reduce the diagnostic error rate due to difficulty in visual interpretation and tedious diagnosis process, Computer Aided Diagnosis (CAD) systems have been developed. This helps in automation of medical imaging tasks such as image segmentation. In this regard, various segmentation methodologies have been developed in the past years, beginning with

the traditional methods like Otsu's thresholding, region growing, GrabCut to the Machine Learning (ML) based techniques, including unsupervised clustering, and supervised support vector machines as discussed in Chapter 1. These techniques provide less accuracy for skin lesion images, as these are not able to handle variability in imaging conditions such as low contrast and blurred boundaries. Artifacts such as hairs and air bubbles also affect the segmentation accuracy. Deep Learning (DL) techniques have emerged as an effective tool in recent years, extensively outperforming traditional and machine learning methods. These techniques have made great progress in image analysis.

## **2.1 RELATED WORK ON SKIN LESION SEGMENTATION**

The journey of semantic segmentation using deep learning methods initiated with the arrival of FCN [1]. Bi et al. [2] have used FCN with multi-scale integration that extracts low-level information using FCN and multi scale integration that provides precise boundaries of skin lesions. But this method failed to segment lesion that were not found in training dataset. Yu et al. [3] have used FCRN for segmentation of skin lesion and to overcome the inaccuracy problem of fully convolutional network, which results in achieving the accuracy of 0.949 on ISIC 2016 dataset. Yuan et al. [4] have proposed Deep Fully Convolutional Networks with jaccard distance as a loss function to optimize the training process and achieved an accuracy of 0.963 on ISIC 2016, but not suitable for low contrast images. Lin et al. [5] have compared UNet with clustering approach for skin lesion segmentation and the result showed that UNet model is superior to latter. SkinNet by Vesal et al. [6] have replaced convolution layer of encoder block of UNet with dilated convolutions and achieved an accuracy of 0.932 on ISIC 2017 dataset. Tong et al. [7] have applied spatial and channel attention to UNet and thereby, the accuracy they received was 0.943 on PH2 dataset. Xu et al. [8] have used modified UNet3+ with CBAM and achieved an accuracy 0.9604 on ISIC 2018. This model was able to segment large size skin lesion images with distinct lesion boundaries but lacked transformation modelling capability of CNN. Recently, Nampelle et al. [9] have employed sequence of skip blocks, called as skip path between encoder and decoder blocks. It was tested on PH2 dataset and received an accuracy of

0.960. In March 2024, Kaur et al. [10] have introduced ARU-Net model with optimized loss function for automatic detection and localization of skin lesion. It achieved an accuracy rate of 0.96 on PH2 dataset but it is computationally expensive.

After reviewing the existing techniques, it can be noted that the model performance is largely affected due to the variations such as shape, size, angle and texture in skin lesion images and presence of artifacts. Some of these existing models are not suitable for low-contrast images. To overcome the limitations of existing segmentation problems, an optimized UNet 3+ based architecture with dilated convolution, residual block and attention model has been proposed. It uses an optimized loss function with less computational complexity.

## **2.2 RELATED WORK ON SEGMENTATION BASED DATA HIDING**

In recent years, several researches targeted to design a suitable approach to incorporate data hiding technique to enhance the security of the patient's health data in medical images such as digital health record system, online diagnosis [11][12]. Some of the existing techniques for data hiding are difference expansion that was used by Tian [13], histogram shifting as employed by Fallahpour et al. [14] and Huang et al. [15]. The flow of the review has been in accordance with the type of segmentation technique incorporated by the researcher.

Wu et al. [16] have proposed the first ROI based contrast enhancement in medical images followed by data hiding using histogram shifting. For proper segmentation of ROI from the images, Adaptive Threshold Detector [44] has been used, which was proposed by Pai et al. [17]. This ATD overcome the class invariance problem of Otsu thresholding method [18]. Yang et al. [19] have used ATD method to automatically segment ROI and NROI region. The embedding rate of data in an image by this method is higher than Gao et al. [20]. Gandhi and Kumar [21] have also utilized this method for precise segmentation of medical images. Despite of many advantages, it is not well-suited for intricate object boundaries. Oo et al. [22] have introduced region based algorithm for capturing the complex boundaries using the gradient information of the image. Parah et al. [23] have employed this region-based segmentation approach,

wherein the image is partitioned into blocks. Subsequently, each block is further divided into seed and non-seed pixels, with data hidden within the non-seed pixel regions. Lin and Li [24] have divided input image into variable size pixel blocks using Quad-Tree Segmentation, which is based on region splitting and merging approach, according to the maximal capacity for each block. An embedding capacity distribution that is more efficient is made possible by this hierarchical segmentation. In case of the data hiding, and contrary to many conventional approaches using single histogram obtained from the entire image, the author has used multiple local histograms which have been computed from the pixels of area sub-images that are locally adjacent. This approach has enabled the algorithm to exploit the simple local distribution of pixel intensities, leading to higher embedding capacity. Salman [25] has utilized quad-tree based segmentation for the segmentation process. Here the region is divided into four quadrants, or blocks in a way that every block has the maximum limit. The secret data and the partition information are placed within each block using the histogram shifting technique in the R. D. H, where minimum and maximum have been found from the peak pixel points and histogram shifting has been done to embed the secret data bits. Chaithanya and Srujana [26] have used Quad-tree segmentation. In this, the blocks are generated by simply partitioning the input image and placed within a quad-tree configuration. By applying the shifting of histogram technique, low contrast blocks have higher capacity for storing more secret data bits than blocks of high contrast that have less storage capacity for the secret data. Consequently, higher contrast blocks necessitate subdivision into smaller sub-blocks. These sub-blocks tend to exhibit smoother characteristics, collectively offering a larger hiding capacity. However, this segmentation can every so often leads to over segmentation and need human intervention for defining the makers. Lakshmanan and Rani [27] have used edge detector segmentation technique to find out ROI of image followed by difference expansion technique to hide data in NROI and ROI. Hiding more data in NROI region increases the payload capacity and also the lossless recovery of image and data is guaranteed. However, this method is not suitable for low contrast images and images with too many edges. GrabCut is a widely utilized graph-based image segmentation technique renowned for its simplicity and accuracy in implementation. Lu et al. [28] have used Gaussian Mixture Model combined with GrabCut in order to improve the

segmentation precision and speed. Wu et al. [16] have adopted GrabCut technique to extract the ROI region from image. For some of the images whose gray level between the ROI and background is close, this method outperforms Otsu's method [18]. The research has also highlighted the benefits of enhancing this ROI region to improve the visual effects. The high accuracy of GrabCut segmentation has the significant impact on the performance of data hiding. These methods perform well on high quality images but the dermoscopic medical images have irregular shape, complex boundaries, rough edges and complicated foreground information.

To handle these images, machine learning based segmentation techniques have come into existence. Rai and Singh [29] have employed SVM classifier to classify medical images into ROI and Non-ROI, so as to embed watermark data into the Non-ROI using DWT-SVD technique. The image's coefficient bands are divided by DWT, and then the watermark is included in the selective coefficients. SVD protects images from attacks, and minor variations/changes to an image's Singular Values have little impact on visual quality. But SVM struggles with imbalanced dataset, where one class is superior to others, leading to suboptimal segmentation results for minority class. Clustering based segmentation approach offer advantages over SVM in terms of unsupervised learning capabilities, scalability, and ability to handle multi-class segmentation. In Li et al. [30] paper, the original image has first been pre-processed using the fuzzy C-means clustering algorithm to produce a number of highly linked classes. After that, each class is split into two halves based on the appropriate threshold. One is encrypted using a stream cypher, while the other uses compressive sensing technology to simultaneously encrypt and compress data for simple data embedding by the information hider. With the data-hiding key and encryption key, the recipient can retrieve the original image and extract additional data. Machine learning algorithms often neglects the spatial relationship between neighboring pixels and this spatial information is important for preserving the image boundaries.

Nowadays, advanced deep learning models reveal many remarkable achievements in considerable many fields associated with medical image segmentation. One of these models is Fully Convolutional Network abbreviation: FCN, the fully connected neural networks are designed using convolution and pooling layers through which spatial

features can be preserved. Based on achieving the satisfactory result in the foreground object segmentation, Meng et al. [31] have proposed the utilization of the special type of segmentation mask known as Mask Region based Convolution Neural Network (R-CNN). The ‘Fully Convolutional Network’ in Mask R-CNN implies that the feature maps corresponding to the set of proposals of an image is fed to the network to make dense prediction at the pixel level. It is integrated into the pixel classification process, and in bit embedding, the stego image is generated using the LSB method. This is an image steganography and it carries secret information being concealed. For segmentation, Saidi et al. [32] have also employed the Mask R-CNN using the following formulation to determine the image’s ROI: Mask R-CNN is designed to draw boundaries of objects of interest and at the same time, identify particular pixels within an image. It is particularly effective in identifying areas of data embedment because of the said peculiarity. As a result of segmentation, the suggested model produced frequency components by changing the area of an image to the frequency region, and then utilized DCT to place data into the components. Some of the amendments are imperceptible to the naked human eye due to data-hiding methods. However, the result of segmentation appears quite subjective and less attentive to the finest lines in a picture. In the presented Dilated Convolutional Segmentation model, a dilation rate gets added to the convolutional layer. This increases the receptive field size without any increase in computational cost. One of the most prominent dilated convolutional model used for image segmentation is DeepLab family. Chen et al. [33] have discussed about DeepLabV1 and DeepLabV2 for image segmentation. These models are basically Convolutional Network integrated with probabilistic graphical models like CRFs and MRFs. Pan et al. [34] have used MobilenetV2, developed based on DeepLab CNN for pixel-level image segmentation. The secret information is embedded inside the semantic features. Secret information retrieval from the semantic level becomes more difficult for the attacker with varying neural network parameters. Dilated convolutional network preserves high level semantics. But, the actual pixel classification lies in low level semantic that is why it sometimes produce vague segmentation results. An encoder-decoder symmetric network called as SegNet has been proposed by Badrinarayanan et al. [35] to achieve end-to-end pixel level information using Softmax classifier. It retains the integrity of high frequency

information but during downsampling low resolution feature map, it ignores neighboring information. To enhance the accuracy of pixel positioning, skip connections were introduced in encoder-decoder model. As deep neural networks grow in depth, their performance tends to decrease.

UNet is a novel approach called long skip connections model that transmit and cascade features transferring encoder layers to the respective decoder layers, allowing for the capture of fine-grained image details. UNet by Ronneberger et al. [36] has become widely adopted in medical image segmentation research. The U-Shaped structure help in capturing fine details at low-level spatial dimension. U-Net can be used for low training data, because the skip connections allow the use of feature maps from encoder path directly into the decoder path. Gao et al. [37] have used enhanced UNet architecture by reducing redundant skip connections, named as UNet3+ by Huang et al. [38] with CBAM attention mechanism by Woo et al. [39] to segment the image into ROI and NROI. The secret data is hidden in ROI in an amplitude-sensitive manner by using the stretched ROI histogram. If after achieving the embedding limit of ROI, still there exists secret data to be embedded, it is then embedded into the Non-Region of Interest (NROI). Additionally, the side information necessary in image restoration is also incorporated into the NROI at the same time. Amrit et al. [40] have used customized UNet3+ architecture to segment cover image into ROI and NROI. Doctor's signature and MAC address have been embedded in ROI region after encrypting it. Other information like hospital logo and patient report have been embedded in NROI region. The watermark image has then been obtained by combining embedded ROI and embedded NROI images. As of late, Shi et al. [41] have developed a fresh structure called as J-Net which is derived from UNet and AlexNet models. In this way, it has been analysed that the UNet model can be developed by the AlexNet with a bridge layer in-between. Bridge layer is useful in modifying the dimensions of the output tensor from the UNet to be in the format that can be fed into the AlexNet input tensor. The AlexNet model uses ReLu activation function; however, the new and improved version has integrated LeakyRelu activation function so as to solve the issue of 'dying ReLu'. The ROI region has been encrypted using plaintext encryption to prevent it from unauthorized access and skipping hiding strategy has



been used to embed secret information into the NROI. Skipping hiding finds out the skipping position in the histogram of an image. The experimental results of this paper are significantly good and can be a break-through in the field of image segmentation based RDH. Luc et al. [42] have introduced an adversarial training method for semantic segmentation. This was how the Generative Adversarial Network (GAN) was trained: to decide if the segmentation maps it is given are the ground truth, or if they were produced by the segmentation network. The study indicates that the adversarial training does not succeed in increasing accuracy much on PASCAL VOC 2012 dataset. But after fine tuning the GAN models and changing their architecture, GAN improves the accuracy of image segmentation task. Annadurai et al. [43] have used convolutional GAN model for the segmentation and classification of MRI and CT watermarked images. The summarized review has been discussed in the below Table. 2.1.

**Table 2.1: Summary of Literature Review**

<b>Image Segmentation Techniques</b>	<b>Advantages</b>	<b>Limitations</b>	<b>Reversible Data Hiding Methods used</b>	<b>Performance Metrics evaluated for Reversible Data Hiding</b>
Thresholding [17]	Require low computational cost with less storage space.	1. Doesn't work well with uneven distribution of data. 2. Sensitive to noise	LSB substitution in Non-ROI and high-payload frequency based RDH in ROI part	Max Embedding Rate= 0.5 bpp  Mean PSNR value= 47.95 dB
Region Based [25]	1. Used for images with good characteristics features , where it is easy to define similarity criteria  2. Noise Resistant	1. Different seed points may give different results  2. High computational cost	Histogram Shifting	PSNR value $\geq 40$ dB

Edge Detection [27]	<p>1. Works well with images having prominent edges</p> <p>2. The boundaries of images with high contrast can be preserved</p>	<p>1. Human intervention is required</p> <p>2. not suitable for low contrast images and images with too many edges</p>	Difference Expansion Technique	Data Not Available
GrabCut [16]	<p>1. Iterative refinement improves segmentation results</p> <p>2. Works well with images having complex background</p>	Quality of segmentation result depends on the initial estimates provided by the user	Histogram Shifting	<p>Max Embedding Rate= 0.5 bpp</p> <p>PSNR value= 33.04 dB</p> <p>SSIM= 0.9711</p>
SVM	<p>1. Can work with complex boundaries images</p> <p>2. SVMs are inherently robust to noise and outliers in data</p>	Struggle with imbalanced dataset, where one class is superior to others, leading to suboptimal segmentation results for minority class.	Discrete Wavelet Transform – Singular Value Decomposition [29]	<p>PSNR value= 52.18 dB</p> <p>SSIM=0.9872</p>
			Integer Wavelet Transform [45]	<p>Max Embedding Rate= 0.44 bpp</p> <p>PSNR value = 64 dB</p> <p>SSIM=0.9802</p>
Clustering [46]	Fuzzy based clustering is better than k-means clustering as it does not use distance as the evaluation metric. It	<p>1. Determining membership function is difficult</p> <p>2. It can be affected by noise</p>	Discrete Wavelet Transform	<p>Max Embedding Rate= 1.25 bpp</p> <p>PSNR value=49.5 dB</p>

	uses membership function, which is more suitable for real world problems.			
Fully Convolutional Network	1. It preserves spatial information throughout the network.  2. Can operate on large datasets	1. Overfitting in the model  2. Huge Training data and time is required  3. Can Suffer from vanishing gradient problem	Least Significant Bit substitution [31]	PSNR value $\geq 50$ dB
			Discrete Cosine Transform [32]	Max Embedding Rate=0.50 bpp  PSNR value= 115.53 dB
Encoder-Decoder Based Model	1. Model can learn complex features.  2. Incorporate skip connections that preserves fine grained details by interfacing low level components of the encoder with the high level components of the decoder	1. Computational cost of these models is high.  2. Difficulty in long range dependencies.	Histogram Shifting [37]	Max Embedding Rate= 0.98 bpp  Mean PSNR value= 34.24 dB  SSIM = 0.8578
			Skipping Hiding [41]	Embedding Capacity= 0.6 bpp  PSNR value= 43dB
Generative Adversarial Network [40]	1. Improve segmentation quality by using the discriminator , which evaluates the	1. High computational resource requirement 2. Hyper tuning of model	Discrete Wavelet Transform	Mean PSNR value=60.8dB

	ground truth segmentation maps with the one produced by the generator	parameters is difficult.		
	2. Robust to noise due to its generative nature.			

From the above review, it can be concluded that ROI can be used to embed secret information as far as possible and to increase visual quality of an image by contrast enhancement. The more data can be embedded in the NROI, as it is the less critical region for diagnosis. The selective embedding modifies only the less important region, without much affecting the visually critical region. Without segmentation, the low contrast image has poor visual quality after data embedding. Therefore, image pre-processing operation like segmentation are combined with data embedding in order to enhance the quality of the image. This technique of hiding data creates a balance between embedding capacity, imperceptibility and security of images. However, performance of this technique depends on how accurate segmentation of ROI and NROI is being carried out.

To enhance the performance of data hiding, an optimized segmentation model discussed in section 2.1 has been utilized. The process of segmentation has been followed by local complexity computation by statistical method to get the smooth pixels for data hiding using Prediction Error Histogram technique.

## CHAPTER 3

### PROPOSED METHODOLOGY

The project has been divided into three main stages: segmentation of medical image, local complexity prediction and data embedding. The first step is the segmentation of a medical image using the appropriate segmentation technique to obtain the ROI of the image which has been followed by local complexity prediction of pixels of segmented ROI image to categorize the smooth pixels that are most suitable for data embedding. After predicting the local complexity, a conventional data hiding has been used to embed and extract the secret data.

The basic flow of the proposed methodology has been presented in Fig. 3.1 below. First, an original image is segmented into ROI and NROI, then convert it to grayscale prior to embedding. Statistical technique is used to calculate the local complexity of pixels of ROI to find out the number of smooth pixels, then the secret data is embedded inside the smooth pixels. Stego image is generated by combining the unaffected NROI and embedded ROI. During data extraction, first the stego image undergoes segmentation then secret data is extracted from embedded ROI, by reversing the operations used for embedding. Finally, the cover image is restored by combining ROI and NROI.

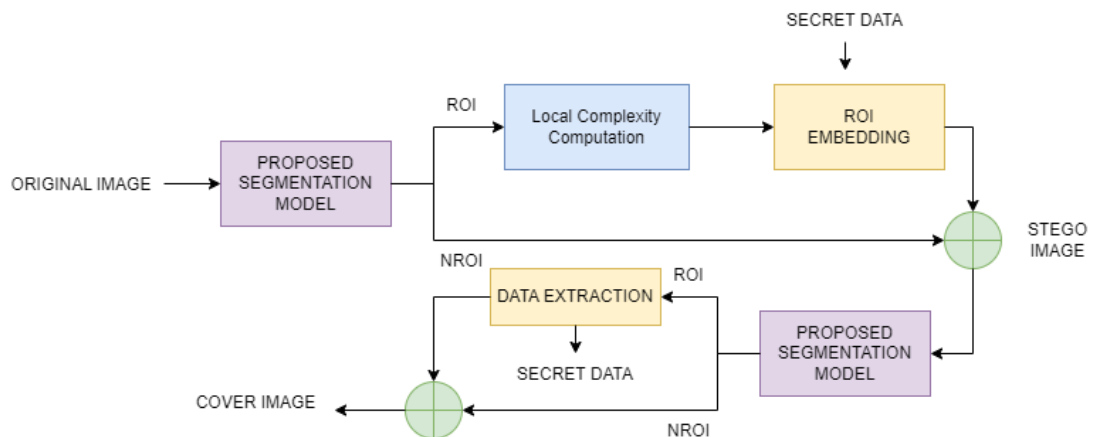
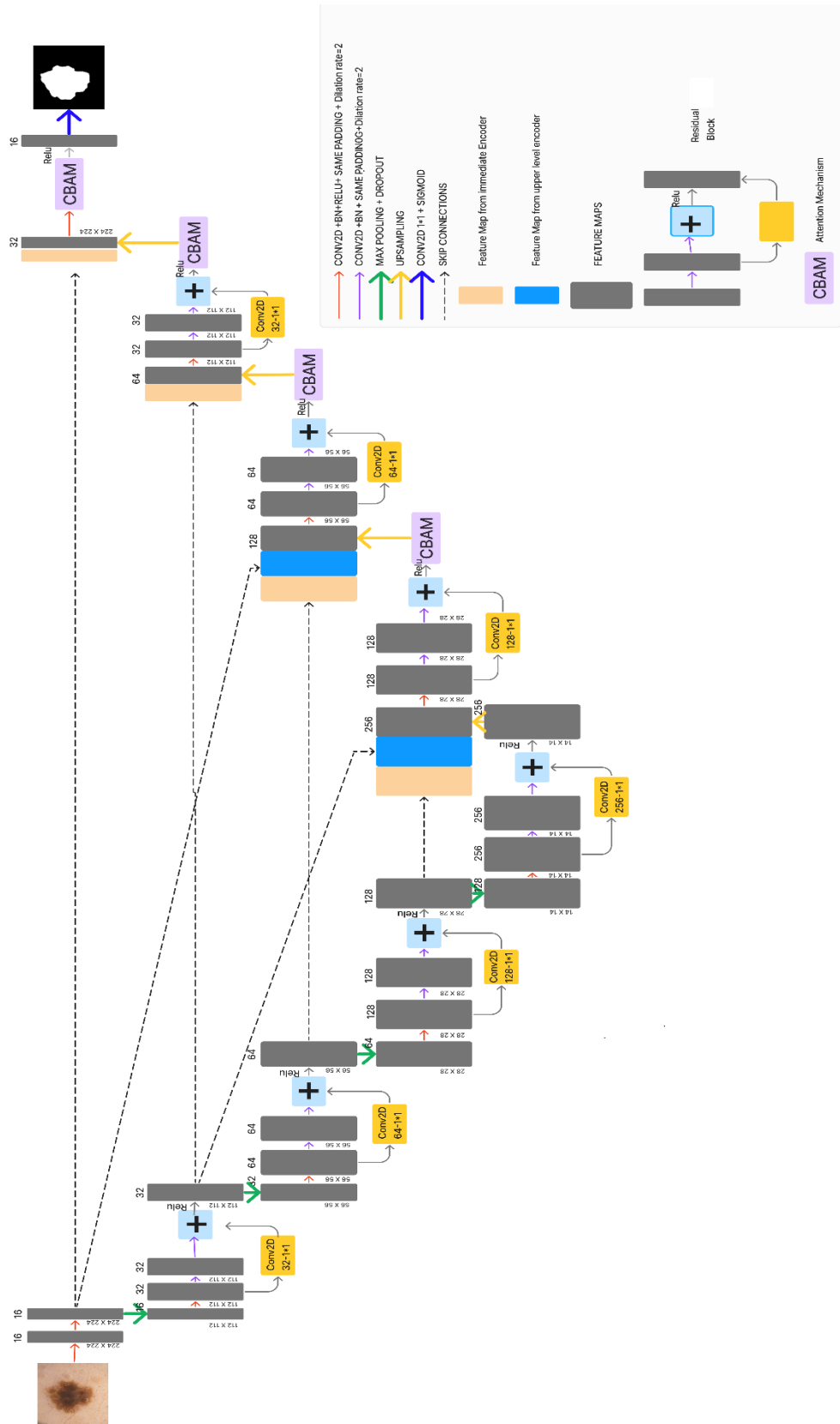


Fig. 3.1: Flowchart of the Proposed System



**Fig. 3.2: Proposed Segmentation Model: Enhanced UNet 3+ Architecture**

### **3.1 MODEL ARCHITECTURE OF PROPOSED SEGMENTATION MODEL**

The architectural diagram illustrating the proposed UNet 3+ segmentation model has been displayed in Fig. 3.2. It consists of encoder path, also called as contracting path made up of residual blocks, multi scale feature fusion using redesigned skip connections, a bottleneck bridge, and decoder path, also called as expanding path, consisting of residual and CBAM block.

An enhanced UNet3+ model has been designed to improve the semantic segmentation accuracy. Residual connections have been implemented in both the encoder and decoder parts to address the gradient degradation issue in a very deep neural network. Additionally, to focus on more specific detailed features, a CBAM has been included at each level in decoder part. Additionally, the decrease in resolution caused by the max pooling was addressed by replacing normal convolution operations with dilated convolution operations and helps in capturing larger contextual information. The redundant skip connections have been removed from the original UNet3+ architecture to reduce the computation cost and the training time.

The detailed explanation of modules used in above proposed model has been discussed in the subsequent points.

- **Modified Convolutional Layer**

In UNet 3+ architecture, low-level feature maps get concatenated with the decoder block during upsampling. However, due to the MaxPooling operation in the encoder block, these low-level feature maps suffer translation invariance and have low resolution. One possible way is to get rid of the pooling layer but UNet would not be able to learn holistic features in the image since convolution is a local operation. A potential solution to this problem is dilated convolution. By using an exponentially increasing receptive field, they can capture global context while maintaining the feature map's resolution. The global information can be shared between different layers without any resolution loss. Therefore, in proposed model, the convolutional block in both encoder and decoder contains 2D convolutional layer with dilation

rate ‘d’ followed by BN and ReLu activation function. The mathematical expression for dilated convolution is as follows-

$$B_{ij} = f\left(\sum_{m=0}^{k-1} \sum_{n=0}^{k-1} (x_{i+m,j+n})^{BN} * w_{mn}^d + b_{ij}\right) \quad (3.1)$$

Where,  $f(.)$  is the ReLu activation function,  $x$  represents input feature with dimension  $(H \times W \times C)$ ,  $H$ =Height,  $W$ =Width,  $C$ = No of channels in input image,  $k$  =kernel size,  $w_{mn}^d$  represents the weight of dilated convolution kernel,  $b_{ij}$  =bias,  $BN$  is the batch normalization,  $*$  is the 2D cross-correlation operator and  $B_{ij}$  = output feature map at position( $i,j$ ) with dimension  $(H_{out} \times W_{out} \times C_{out})$ .  $C_{out}$  depends upon number of filters used in the model and  $H_{out}$  and  $W_{out}$  can be computed using-

$$H_{out} = \frac{H + 2p - d \times (k - 1) - 1}{s} + 1 \quad (3.2)$$

$$W_{out} = \frac{W + 2p - d \times (k - 1) - 1}{s} + 1 \quad (3.3)$$

This  $H_{out}$  and  $W_{out}$  depends on input dimensions, padding ( $p$ ), dilation rate ( $d$ ), kernel size ( $k$ ) and stride ( $s$ ).

### • Residual Block in Encoder and Decoder Path

The structure of residual block employed in proposed model consists of two 2D convolutional layer with BN. Before going to the ReLu activation function, a shortcut connection between the output of the previous modified convolution layer and the output feature map of the second convolution layer of the residual block is added. This shortcut connection also include a two-dimensional convolution layer to reshape the shortcut path's result to match the dimension of the main path. Residual block has been incorporated to resolve network degradation problem during model training.



- **Redesigned Skip Connections**

In Unet3+, the segmentation map resulting from the adjacent multi-scale feature maps has the same contributions. This leads to excessive redundant computations. In order to solve this issue, skip connections are redesigned in the proposed model. The skip connection from adjacent encoder block to the decoder block are removed.

- **Encoder Network / Contracting Path**

The encoder network consist of four different convolutional blocks. First block consists of only two modified convolutional layer and the remaining three convolutional blocks consists of sequence of one modified convolutional layer and residual block. High-level features and context are extracted from the input skin lesion images, which are each  $224 \times 224 \times 3$  by the encoder. It focuses on “what” is in the image. Each convolutional block of encoder is followed by MaxPooling layer that reduces the spatial dimension of the feature maps, thereby minimizing the number of parameters in the network. To prevent model from overfitting, a dropout layer has been added after MaxPooling. The first convolutional block of encoder network has 16 feature map, each of  $224 \times 224$  size. As the network progresses deeper through its layer, it doubles the number of feature map and reduced the size of each feature map by half.

- **Bottleneck Bridge Block**

It is a convolutional block of modified convolutional layer and a residual structure having 256 feature maps, each of size  $14 \times 14$ . It is a bridge between the encoder and decoder network. The output of this block is passed to the first layer of the decoder network after upsampling by Kolarik et al. [47].

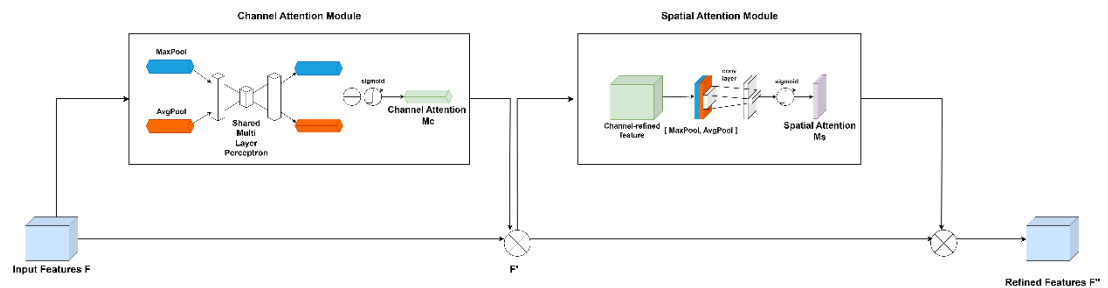
- **Decoder Network / Expanding Path**

The Decoder network of the proposed model consist of four blocks. Each block except the last, is made up of concatenated feature maps from corresponding encoder block,

bilinear upsampled feature map from lower level block and feature map from upper level encoder block using redesigned skip connections, followed by residual and CBAM block. The loss of spatial information during encoding is recover by the decoder. It produces the segmentation maps by performing the reverse operation of the encoder. It focuses on “where” is the object present in the image. The last block of the decoder network consist of concatenated feature maps followed by CBAM and a single 1 X 1 2D A convolutional layer that uses a sigmoid activation function.. This function generates the output in the range of 0 to 1. Value between 0 and 0.5 defines the background region and from 0.5 to 1 represents ROI (skin lesion). This numerical computation helps the last block of the decoder to generate the segmented map.

- **CBAM in the Decoder Path**

Attention module is capable of focussing on important features of the input while ignoring the irrelevant details. CBAM is a lightweight attention module used in the proposed model to focus on skin lesion pixels while ignoring the irrelevant background pixels. It also contains channel and spatial attention module to obtain the enhanced feature map as depicted in Fig. 3.3. The input for the CBAM module is the intermediate feature map  $F$  in shape of  $H \times W \times C$  from the preceding residual block. It sequentially computes 1D channel attention map  $MC \in 1 \times 1 \times C$  and a channel refined map  $F'$ . This  $F'$  is pass to the spatial attention module that compute the 2D spatial attention map  $Ms \in H \times W \times 1$  and a final refined feature map  $F''$ .



**Fig. 3.3: Convolutional Block Attention Module (CBAM)**

These two attention modules whether be in a sequential or parallel can be used interchangeably. But, Woo et al. [39], have proved that this technique of architectural

design in which two or more types of modules are attached one after the other performs better than the technique in which they are delivered simultaneously.

- **Loss Function**

The segmentation model that has been proposed utilized Jaccard Distance as the loss function to measure the difference between the segmentation mask that has been predicted and the ground truth of the image. Of these, the Jaccard index or the Jaccard similarity coefficient is obtained by subtracting 1 from the Jaccard distance as defined in the equation below. 3. 4. Jaccard index is calculated by the ratio of the overlapped area between the ground truth map  $Y$  and the predicted segmentation map  $Z$ , to that of the combined area of both ground truth and segmentation maps across all classes.

$$d_j(Y, Z) = 1 - J(Y, Z) = 1 - \frac{|Y \cap Z|}{|Y| + |Z| - |Y \cap Z|} \quad (3.4)$$

where,  $J(Y, Z)$  represents Jaccard Index and  $d_j(Y, Z)$  is the Jaccard distance. But this  $d_j(R, S)$  itself is not differentiable, so it cannot be apply directly into backpropagation to train the model. Therefore, the above function is changed to-

$$L_{dJ} = 1 - \frac{\sum_{i,j}(t_{ij}p_{ij})}{\sum_{i,j}t_{ij}^2 + \sum_{i,j}p_{ij}^2 - \sum_{i,j}(t_{ij}p_{ij})} \quad (3.5)$$

Here,  $t$  represents true (ground truth) and  $p$  represents predicted. Now, this loss function is differentiable and is efficient to use for model training.

Table 3.1 summarises an overview of the number of layers applied, type of each layer, input and output shape, parameter values, as well as how each layer is related to one another. It shows how many parameters are trainable or non-trainable during the model training. It helps in understanding the flow of the network.

**Table 3.1: Model Parameters**

LAYER(TYPE)	OUTPUT IMAGE SIZE	PARAMETERS	CONNECTED TO
input_1	224*224*3	0	[]
conv2d	224*224*16	448	['input_1[0][0]']
batch_normalization	224*224*16	64	['conv2d[0][0]']

Table 3.1(Continued)

conv2d_1	224*224*16	2320	['batch_normalization[0][0]']
batch_normalization_1	224*224*16	64	['conv2d_1[0][0]']
max_pooling2d	112*112*16	0	['batch_normalization_1[0][0]']
dropout	112*112*16	0	['max_pooling2d[0][0]']
conv2d_2	112*112*32	544	['dropout[0][0]']
batch_normalization_2	112*112*32	128	['conv2d_2[0][0]']
activation(ReLu)	112*112*32	0	['batch_normalization_2[0][0]']
conv2d_3	112*112*32	9284	['activation[0][0]']
conv2d_4	112*112*32	544	['dropout[0][0]']
batch_normalization_3	112*112*32	128	['conv2d_3[0][0]']
batch_normalization_4	112*112*32	128	['conv2d_4[0][0]']
add	112*112*32	0	['batch_normalization_3[0][0]', 'batch_normlization_4[0][0]']
activation_1 (ReLu)	112*112*32	0	['add[0][0]']
max_pooling2d_1	56*56*32	0	['activation_1[0][0]']
dropout_1	56*56*32	0	['max_pooling2d_1[0][0]']
conv2d_5	56*56*64	2112	['dropout_1[0][0]']
batch_normalization_5	56*56*64	256	['conv2d_5[0][0]']
activation_2(ReLu)	56*56*64	0	['batch_normalization_5[0][0]']
conv2d_6	56*56*64	36928	['activation_2[0][0]']
conv2d_7	56*56*64	2112	['dropout_1[0][0]']
batch_normalization_6	56*56*64	256	['conv2d_6[0][0]']
batch_normalization_7	56*56*64	256	['conv2d_7[0][0]']
add_1	56*56*64	0	['batch_normalization_6[0][0]', 'batch_normlization_7[0][0]']
activation_3 (ReLu)	56*56*64	0	['add_1[0][0]']
max_pooling2d_2	28*28*64	0	['activation_3[0][0]']
dropout_2	28*28*64	0	['max_pooling2d_2[0][0]']
conv2d_8	28*28*128	8320	['dropout_2[0][0]']
batch_normalization_8	28*28*128	512	['conv2d_8[0][0]']
activation_4(ReLu)	28*28*128	0	['batch_normalization_8[0][0]']
conv2d_9	28*28*128	147584	['activation_4[0][0]']
conv2d_10	28*28*128	8320	['dropout_2[0][0]']
batch_normalization_9	28*28*128	512	['conv2d_9[0][0]']
batch_normalization_10	28*28*128	512	['conv2d_10[0][0]']
add_2	28*28*128	0	['batch_normalization_9[0][0]', 'batch_normlization_10[0][0]']
activation_5 (ReLu)	28*28*128	0	['add_2[0][0]']
max_pooling2d_3	14*14*128	0	['activation_5[0][0]']
dropout_3	14*14*128	0	['max_pooling2d_3[0][0]']
conv2d_11	14*14*256	33024	['dropout_3[0][0]']
batch_normalization_11	14*14*256	1024	['conv2d_11[0][0]']
activation_6(ReLu)	14*14*256	0	['batch_normalization_11[0][0]']
conv2d_12	14*14*256	590080	['activation_6[0][0]']

Table 3.1 (Continued)

conv2d_13	14*14*256	33024	['dropout_3[0][0]']
conv2d_15	112*112*32	9248	['activation_1[0][0]']
batch_normalization_1_2	14*14*256	1024	['conv2d_12[0][0]']
batch_normalization_1_3	14*14*256	1024	['conv2d_13[0][0]']
conv2d_14	28*28*128	147584	['activation_5[0][0]']
batch_normalization_1_5	112*112*32	128	['conv2d_15[0][0]']
add_3	14*14*256	0	['batch_normalization_12[0][0]', 'batch_normlization_13[0][0]']
batch_normalization_1_4	28*28*128	512	['conv2d_14[0][0]']
tf.nn.relu_1	112*112*32	0	['batch_normalization_15[0][0]']
activation_7 (ReLU)	14*14*256	0	['add_3[0][0]']
tf.nn.relu	28*28*128	0	['batch_normalization_14[0][0]']
max_pooling2d_4	28*28*32	0	['tf.nn.relu_1[0][0]']
up_sampling2d	28*28*256	0	['activation_7[0][0]']
concatenate	28*28*416	0	['tf.nn.relu[0][0]', 'max_pooling2d_4[0][0]', 'up_sampling2d[0][0]']
dropout_4	28*28*416	0	['concatenate[0][0]']
conv2d_16	28*28*128	53376	['dropout_4[0][0]']
batch_normalization_1_6	28*28*128	512	['conv2d_16[0][0]']
activation_8(ReLU)	28*28*128	0	['batch_normalization_16[0][0]']
conv2d_17	28*28*128	147584	['activation_8[0][0]']
conv2d_18	28*28*128	53376	['dropout_4[0][0]']
batch_normalization_1_7	28*28*128	512	['conv2d_17[0][0]']
batch_normalization_1_8	28*28*128	512	['conv2d_18[0][0]']
add_4	28*28*128	0	['batch_normalization_17[0][0]', 'batch_normlization_18[0][0]']
activation_9(ReLU)	28*28*128	0	['add_4[0][0]']
tf.math.reduce_mean	1*1*128	0	['activation_9[0][0]']
dense	1*1*16	2064	['tf.math.reduce_mean[0][0]']
dense_1	1*1*128	2176	['dense[0][0]']
tf.math.multiply	28*28*128	0	['activation_9[0][0]', 'dense_1[0][0]']
tf.math.reduce_mean_1	28*28*1	0	['tf.math.multiply[0][0]']
tf.math.reduce_max_1	28*28*1	0	['tf.math.multiply[0][0]']
tf.concat	28*28*2	0	['tf.math.reduce_mean_1[0][0]', 'tf.math.reduce_max_1[0][0]']
conv2d_20	224*224*16	2320	['batch_normalization_1[0][0]']
dense_2	28*28*1	3	['tf.concat[0][0]']
conv2d_19	56*56*64	36928	['activation_3[0][0]']

Table 3.1 (Continued)

batch_normalization_20	224*224*16	64	['conv2d_20[0][0]']
activation_10	28*28*1	0	['dense_2[0][0]']
batch_normalization_19	56*56*64	256	['conv2d_19[0][0]']
tf.nn.relu_3	224*224*16	0	['batch_normalization_20[0][0]']
tf.math.multiply_1	28*28*128	0	['tf.math.multiply[0][0]', 'activation_10[0][0]']
tf.nn.relu_2	56*56*64	0	['batch_normalization_19[0][0]']
max_pooling2d_5	56*56*16	0	['tf.nn.relu_3[0][0]']
up_sampling2d_1	56*56*128	0	['tf.math.multiply_1[0][0]']
concatenate_1	56*56*208	0	['tf.nn.relu_2[0][0]', 'max_pooling2d_5[0][0]', 'up_sampling2d_1[0][0]']
dropout_5	56*56*208	0	['concatenate_1[0][0]']
conv2d_21	56*56*64	13376	['dropout_5[0][0]']
batch_normalization_21	56*56*64	256	['conv2d_21[0][0]']
activation_11(ReLU)	56*56*64	0	['batch_normalization_21[0][0]']
conv2d_22	56*56*64	36928	['activation_11[0][0]']
conv2d_23	56*56*64	13376	['dropout_5[0][0]']
batch_normalization_22	56*56*64	256	['conv2d_22[0][0]']
batch_normalization_23	56*56*64	256	['conv2d_23[0][0]']
add_5	56*56*64	0	['batch_normalization_22[0][0]', 'batch_normalization_23[0][0]']
activation_12(ReLU)	56*56*64	0	['add_5[0][0]']
tf.math.reduce_mean_2	1*1*64	0	['activation_12[0][0]']
dense_3	1*1*8	520	['tf.math.reduce_mean_2[0][0]']
dense_4	1*1*64	576	['dense_3[0][0]']
tf.math.multiply_2	56*56*64	0	['activation_12[0][0]', 'dense_4[0][0]']
tf.math.reduce_mean_3	56*56*1	0	['tf.math.multiply_2[0][0]']
tf.math.reduce_max_3	56*56*1	0	['tf.math.multiply_2[0][0]']
tf.concat_1	56*56*2	0	['tf.math.reduce_mean_3[0][0]', 'tf.math.reduce_max_3[0][0]']
dense_5	56*56*1	3	['tf.concat_1[0][0]']
conv2d_24	112*112*32	9248	['activation_1[0][0]']
activation_13(ReLU)	56*56*1	0	['dense_5[0][0]']
batch_normalization_24	112*112*32	128	['conv2d_24[0][0]']
tf.math.multiply_3	56*56*64	0	['tf.math.multiply_2[0][0]', 'activation_13[0][0]']
tf.nn.relu_4	112*112*32	0	['batch_normalization_24[0][0]']

Table 3.1 (Continued)

up_sampling2d_2	112*112*64	0	['tf.math.multiply_3[0][0]']
concatenate_2	112*112*96	0	['tf.nn.relu_4[0][0]', 'up_sampling2d_2[0][0]']
dropout_6	112*112*96	0	['concatenate_2[0][0]']
conv2d_25	112*112*32	3104	['dropout_6[0][0]']
batch_normalization_25	112*112*32	128	['conv2d_25[0][0]']
activation_14(ReLU)	112*112*32	0	['batch_normalization_25[0][0]']
conv2d_26	112*112*32	9248	['activation_14[0][0]']
conv2d_27	112*112*32	3104	['dropout_6[0][0]']
batch_normalization_26	112*112*32	128	['conv2d_26[0][0]']
batch_normalization_27	112*112*32	128	['conv2d_27[0][0]']
add_6	112*112*32	0	['batch_normalization_26[0][0]', 'batch_normalization_27[0][0]']
activation_15(ReLU)	112*112*32	0	['add_6[0][0]']
tf.math.reduce_mean_4	1*1*32	0	['activation_15[0][0]']
dense_6	1*1*4	132	['tf.math.reduce_mean_4[0][0]']
dense_7	1*1*32	160	['dense_6[0][0]']
tf.math.multiply_4	112*112*32	0	['activation_15[0][0]', 'dense_7[0][0]']
tf.math.reduce_mean_5	112*112*1	0	['tf.math.multiply_4[0][0]']
tf.math.reduce_max_5	112*112*1	0	['tf.math.multiply_4[0][0]']
tf.concat_2	112*112*2	0	['tf.math.reduce_mean_5[0][0]', 'tf.math.reduce_max_5[0][0]']
dense_8	112*112*1	3	['tf.concat_2[0][0]']
conv2d_28	224*224*16	2320	['batch_normalization_1[0][0]']
activation_16	112*112*1	0	['dense_8[0][0]']
batch_normalization_28	224*224*16	64	['conv2d_28[0][0]']
tf.math.multiply_5	112*112*32	0	['tf.math.multiply_4[0][0]', 'activation_16[0][0]']
tf.nn.relu_5	224*224*64	0	['batch_normalization_28[0][0]']
up_sampling2d_3	224*224*32	0	['tf.math.multiply_5[0][0]']
concatenate_3	224*224*48	0	['tf.nn.relu_5[0][0]', 'up_sampling2d_3[0][0]']
dropout_7	224*224*48	0	['concatenate_3[0][0]']
conv2d_29	224*224*16	784	['dropout_7[0][0]']
batch_normalization_29	224*224*16	64	['conv2d_29[0][0]']
activation_17(ReLU)	224*224*16	0	['batch_normalization_29[0][0]']
conv2d_30	224*224*16	2320	['activation_17[0][0]']
conv2d_31	224*224*16	784	['dropout_7[0][0]']

Table 3.1 (Continued)

batch_normalization_30	224*224*16	64	['conv2d_30[0][0]']
batch_normalization_31	224*224*16	64	['conv2d_31[0][0]']
add_7	224*224*16	0	['batch_normalization_30[0][0]', 'batch_normlization_31[0][0]']
activation_18(ReLu)	224*224*16	0	['add_7[0][0]']
tf.math.reduce_mean_6	1*1*16	0	['activation_18[0][0]']
dense_9	1*1*2	34	['tf.math.reduce_mean_6[0][0]']
dense_10	1*1*16	48	['dense_9[0][0]']
tf.math.multiply_6	224*224*16	0	['activation_18[0][0]', 'dense_10[0][0]']
tf.math.reduce_mean_7	224*224*1	0	['tf.math.multiply_6[0][0]']
tf.math.reduce_max_7	224*224*1	0	['tf.math.multiply_6[0][0]']
tf.concat_3	224*224*2	0	['tf.math.reduce_mean_7[0][0]', 'tf.math.reduce_max_7[0][0]']
dense_11	224*224*1	3	['tf.concat_3[0][0]']
activation_19(ReLu)	224*224*1	0	['dense_11[0][0]']
tf.math.multiply_7	224*224*16	0	['tf.math.multiply_6[0][0]', 'activation_19[0][0]']
conv2d_32	224*224*1	17	['tf.math.multiply_7[0][0]']
reshape	224*224	0	['conv2d_32[0][0]']
<p style="text-align: center;"> <b>Total Parameters: 1435275 (5.48 MB)</b>  <b>Trainable Parameters: 1430315 (5.46 MB)</b>  <b>Non- Trainable Parameters: 4960 (19.38 KB)</b> </p>			

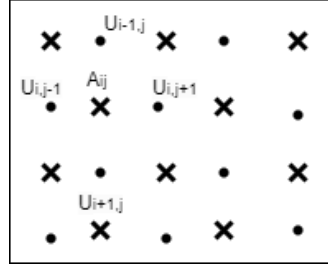
## • Post-Processing

In order to further improve the segmentation result, a post processing technique has been used. The outputs generated directly from the proposed model exhibit haziness due to the pixel values being constrained within the range of 0 to 1 because of sigmoid function in the last layer. To enhance the visibility of the edges, a binary thresholding have been applied by rounding the pixel values to either 0 or 1. Pixels with values exceeding 0.5 are rounded up to 1, while those below 0.5 are rounded down to 0.



### 3.2 ROI EMBEDDING

First, a ROI segmented image has been converted to grayscale image for embedding the data. Then, it has been divided into two set: cross and dot set as shown in Fig. 3.4. Cross set has been used for embedding and dot set for computing the prediction error values.



**Fig. 3.4: Cross and Dot Set**

For each cross set pixel, local complexity has been computed by incorporating the statistical measures. The mean/complexity and variance between the target pixel for embedding and its neighboring pixels has been calculated using the eqn. 3.6.

$$variance = \frac{1}{4} \sum_{k=1}^4 (complexity - v_k)^2 \quad (3.6)$$

Where,

$$v_1 = |U_{i,j-1} - U_{i-1,j}|,$$

$$v_2 = |U_{i,-1,j} - U_{i,j+1}|,$$

$$v_3 = |U_{i,j+1} - U_{i+1,j}|,$$

$$v_4 = |U_{i+1,j} - U_{i,j-1}| \text{ and}$$

$$complexity = \frac{(v_1 + v_2 + v_3 + v_4)}{4}$$

In the proposed system, pixels with a variance below the threshold value of 0.5 were considered to be smooth pixels. Low variance pixels are less likely to cause much distortion. These pixels are more embeddable than high variance pixels.

For embedding of secret data inside the smooth pixels of ROI, calculated using above discussed technique, a conventional Prediction Error Expansion Reversible Data Hiding (PEE-RDH) technique by Sachnev et al. [48] has been employed in the

proposed system. First, the predicted value  $P_{ij}$  for each pixel  $A_{ij}$  has been calculated from four neighbouring pixel  $U_{i,j-1}$  ,  $U_{i+1,j}$  ,  $U_{i,j+1}$  and  $U_{i-1,j}$  by using the eqn 3.7.

$$P_{ij} = \frac{U_{i,j-1} + U_{i+1,j} + U_{i,j+1} + U_{i-1,j}}{4} \quad (3.7)$$

The prediction error can be calculated as

$$d_{ij} = P_{ij} - A_{ij} \quad (3.8)$$

The data can be embedded inside each pixel using the following equations 3.9

$$\left. \begin{aligned} A'_{ij} &= P_{ij} - b \text{ if } d_{ij} == 0 \\ A'_{ij} &= P_{ij} + b \text{ if } d_{ij} == 1 \\ A'_{ij} &= P_{ij} + 1 \text{ if } d_{ij} < 0 \\ A'_{ij} &= P_{ij} - 1 \text{ if } d_{ij} > 0 \end{aligned} \right\} \quad (3.9)$$

where,  $A'_{ij}$  is the modified pixel after embedding, b is the message bit. The data extraction process is exactly the reverse operation of above described process. The dot set pixels has not been modified, so the value of prediction error value will remain same during extraction process. The embedding bit and original pixels can be recovered easily.

## **CHAPTER 4**

### **EXPERIMENTAL RESULTS AND ANALYSIS**

The flow of this chapter starts with the tools and libraries used to carry out the implementation of the project followed by short overview on dataset used. Afterwards it discusses about the pre-processing technique along with its visual outcomes. In the next section, various evaluation metrics have been discussed along with the implementation details in the next section. Further, the detailed analysis of results have been presented by comparing it with the state-of-the-art models.

#### **4.1 TOOLS AND LIBRARIES USED**

A set of powerful tools and libraries were used to develop this project. The proposed system's performance was enhanced by the combined use of these tools. The detailed explanation of the used tools and libraries are as follows:

- **Colab Notebook**

Google offers an online Jupyter notebook environment known as Colab Notebook that permits the writing and execution of Python code. It provides a collaborative workspace to perform data analysis, machine learning and deep learning tasks. Colab provides users with free GPU and TPU resources for accelerated computing, as well as a smooth integration with Google Drive that makes it simple to share projects and access data from drive. The entire project has been developed on Colab notebook with T4 GPU.

- **Tensorflow Framework**

It is a framework that is freely available for machine learning and deep learning applications. It consists of several libraries that aids in creation, training and deployment of deep neural network tasks. Its user friendly keras API helped in developing neural network models and provides extension set of tools for training models. In this project, keras has been used for pre-processing the data, creating layers

of proposed model, model learning through activation functions and optimizers for training of proposed model.

- **Scikit Learn (Sklearn) Library**

It is a python library used for data analysis and mining. It has been used for splitting of the dataset in this project. Scikit image, which is based on Sklearn, is widely used for image analysis tasks. It is used to determine the structural similarity of images during the data hiding task.

- **CV2**

It is an open-source OpenCV python library that specializes in image processing tasks. It provides ample amount of image processing tools that are convenient to use. The CV2 library offers an extensive range of functionalities that streamline the execution of intricate image processing tasks, allowing users to create sophisticated applications with comparatively less effort. CV2 has been used in image augmentation task and mapping of image with segmented mask in this project.

- **Matplotlib**

Matplotlib is a python library used for creating interactive visualizations. It is used to design customizable plots, graphs, scatter plots etc. Matplotlib turns out to be a priceless tool, helping the users to convey results and derive useful conclusions from their research. In this project, Matplotlib has been deployed for visual interpretation of images, qualitative and quantitative analysis of proposed system.

- **Figma**

Figma is a powerful editing tool used for designing precise graphical illustrations. It can create interactive diagrams to simulate flows and interactions, essential for presenting designs. The architectural diagrams of the existing and proposed segmentation model has been created using Figma. It offers user friendly interface, powerful set of plugins and functionalities.

## 4.2 DATASET

A PH2 dataset, a benchmark for segmentation of skin lesions, has been used to evaluate the suggested segmentation model. PH2 has been released by Mendonca et al. [49] and has been the first public dataset for skin lesion region based images with segmentation masks. It consists of 200 dermoscopic images of melanocytic lesions with their corresponding ground truth mask. The image size vary from 553 to 769. For memory optimization, the images has been resized to dimensions of  $224 \times 224 \times 3$ .


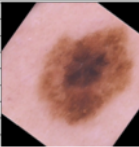
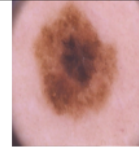
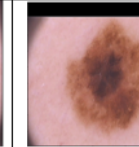
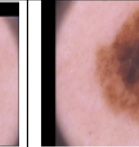
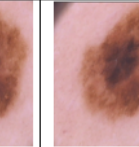






## 4.3 PRE-PROCESSING

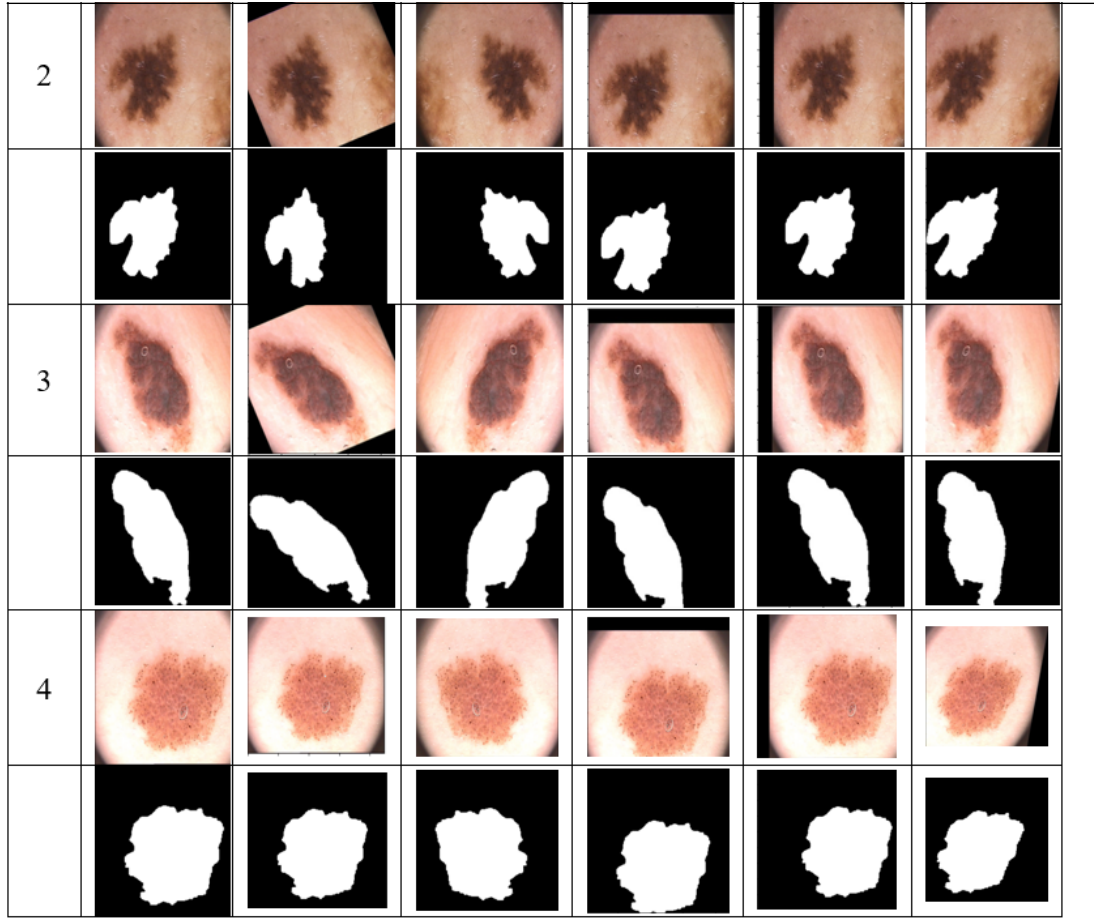
As the UNet 3+ is a deep learning model, it gives the best results when a large amount of data is fed to the model. As a result, some data augmentation have been applied to expand the quantity of data used in the training process. The parameters of the augmentation operations are shown in Table 4.1 together with the name of each operation.

**Table 4.1: Data Augmentation Technique**

Operation	Parameters
Rotation	Between -40 to 40 degree
Flipping	Across horizontal axis
Height Shift	10% in the vertical direction
Width Shift	10% in the horizontal direction
Shear	$10^0$ in the counter-clockwise direction along the x-axis

The size of the dataset approximately doubles after applying data augmentation techniques. The Fig 4.1 shows the sample of images and their segmentation mask after applying augmentation techniques.

S. No.	Original Image and Mask	Rotated Image and Mask	Flipped image and mask	Vertical Height shift image and mask	Horizontal width shift image and mask	Shear image and mask
1						
						



**Fig. 4.1:** Image Augmentation Techniques on Skin Lesion Images

#### 4.4 EVALUATION METRICS

The evaluation metrics below are utilized to assess the performance of the proposed segmented model:

- **Accuracy (ACC)** - It is characterized by the number of pixels that were correctly predicted compared to the total number of predictions. It is given by

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (4.1)$$

- **Dice Coefficient (DC)** - It calculates the number of positives and penalizes for the false positives predicted by the model. Its value ranges between 0 and 1.

$$Dice\ Coefficient = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (4.2)$$

- **Jaccard Index (JAC)** - It evaluates how much the predicted mask and the ground truth mask are alike or different.

$$Jaccard\ Index = \frac{TP}{TP + FP + FN} \quad (4.3)$$

- **Sensitivity/Recall (SE)** – It is the ratio of predicted skin lesion pixels to the actual skin lesion pixels. It can be computed using

$$Sensitivity = \frac{TP}{TP + FN} \quad (4.4)$$

Here, True Positive (TP) is the number of pixels correctly predicted by the model as skin lesion pixels. The proposed model predicts the number of background pixels as skin lesion pixels (FP= False Positive). The number of skin lesion pixels misclassified as background pixels is represented by FN = False Negative, while TN = True Negative represents the number of background pixels correctly classified as background pixels by the proposed model.

These metrics have been used to compare the performance of the proposed segmentation model with the already existing models to measure its efficiency and effectiveness on PH2 dataset. These metrics help in overall quantitative analysis of the proposed model and suggest the parameters to change if their values are not satisfactory.

To measure data hiding's performance, the PSNR value and SSIM value are calculated, as described below:

- **Peak Signal to Noise Ratio (PSNR)** - The difference in reconstructed image versus original image is measured by calculating how much noise or distortion was introduced.

$$PSNR = 10 * \log_{10} \frac{MAX^2}{MSE} \text{ dB} \quad (4.5)$$

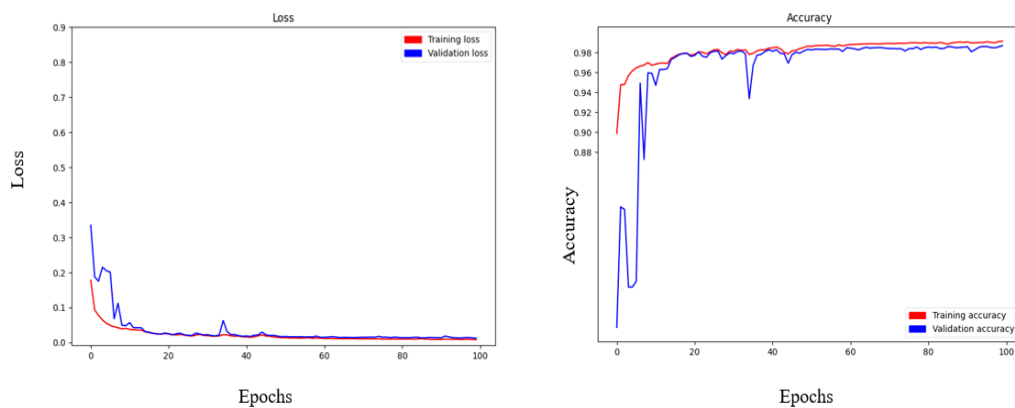
Where, the maximum possible value for the pixels in the image is MAX. MSE is the mean square error, dB= decibel.

- **Structural Similarity Index Measure (SSIM)** - Identify the average structural similarity score between the two images. It has been implemented using Skimage python library. Its value ranges between -1 to +1. When the two provided images have a value of +1, it means they are nearly identical, and when the value is -1, it means the two are significantly different. These numbers are often adjusted to be within the range of 0 to 1.

## 4.5 IMPLEMENTATION DETAILS

The PH2 dataset has been separated into a 75:25 division for the purpose of training and testing. The training data is further divided after augmentation into 80:20 training-validation split. The Python programming language and Keras framework were utilized to implement the proposed model on Google Colab Notebook. 100 epochs have been spent training the model with Jaccard distance as the loss parameter. An Adam optimizer was employed to improve the training process with a learning rate of 0.001. Check pointing has been incorporated during training to save the model's weights so as to resume training from the point where the best performance was achieved, in case any crash occurs. Fig. 4.2 illustrate the training and validation accuracy, as well as training and validation loss across 100 epochs.

In the Data hiding part, secret data has been converted from string to binary and then embed using the equations discussed in section 3.2 after calculating the local complexity of pixels. The size of the secret data has been calculated using the number of smooth pixels computed by local complexity method.



**Fig. 4.2:** The left graph shows the proposed model's training and validation loss, while the right graph shows the proposed model's training and validation accuracy on the PH2 dataset.



## 4.6 RESULTS OF THE PROPOSED SYSTEM

The proposed system's results are discussed here, with a focus on both qualitative and quantitative evaluation. The comparative performance analysis with other existing models has also been discussed in this section.

### 4.6.1 ABLATION STUDY

This study helps in determining the impact of different modules of the proposed segmentation model on its performance. Below Table 4.2 discusses about the contributions and significance of each component by methodically removing them and then monitoring the changes in performance that occur.

**Table 4.2: Ablation Study on Different Modules**

Method	Loss Function Used	Training Accuracy	Training Loss	Testing Accuracy	Testing Loss
UNet 3+	Binary Cross Entropy	0.9145	0.2206	0.9058	0.2397
UNet 3+ + CBAM	Binary Cross Entropy	0.9635	0.0945	0.9358	0.1529
UNet 3+ + CBAM + RESIDUAL BLOCK	Jaccard_distance	0.9659	0.0300	0.9518	0.0425
UNet 3+ + CBAM + RESIDUAL BLOCK + DILATION RATE	Jaccard_distance	0.9914	0.0083	0.9679	0.0308

From the above Table, it is clear that each component of the proposed segmentation model contributes to segmentation accuracy, and the proposed loss function reduces training and testing losses significantly.

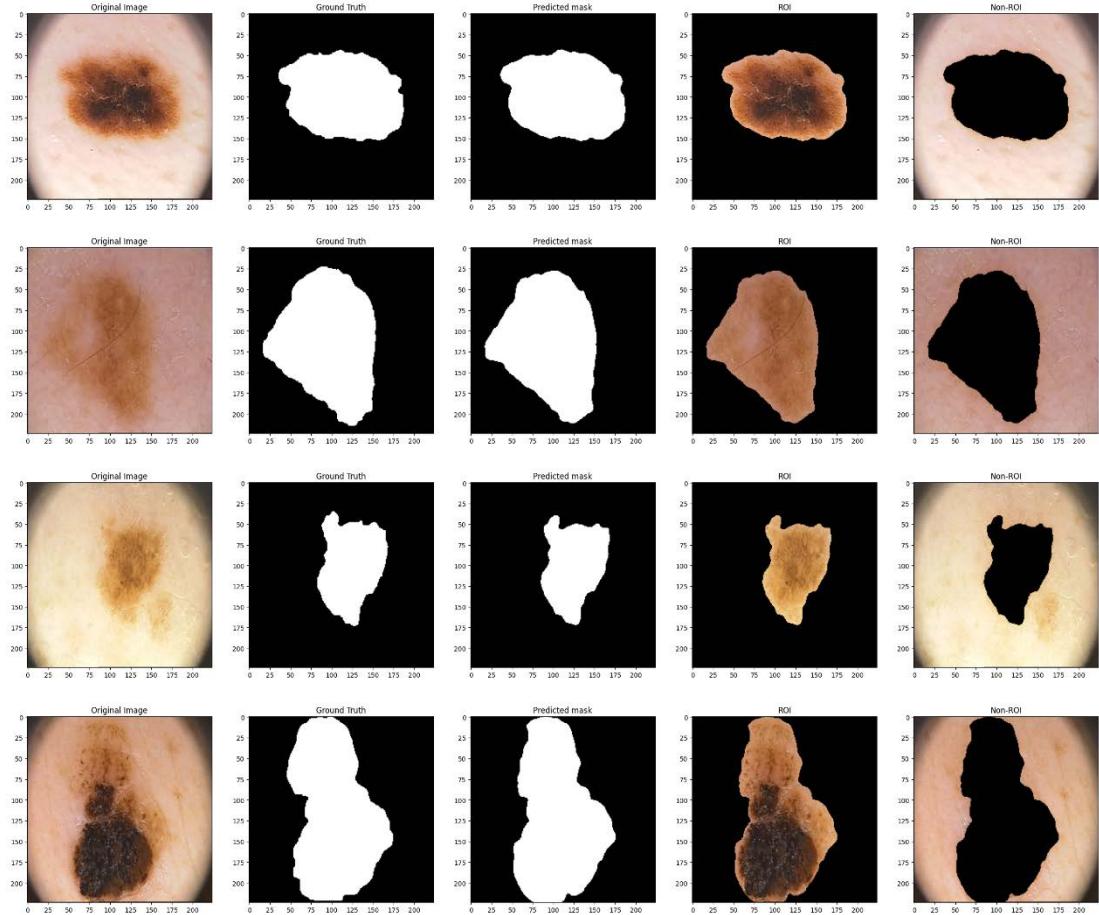
The other experiment done on this aspect involved the comparison of different batch sizes and learning rate as depicted in the Table. 4.3. The batch size of 32 and learning rate of 0.01 performed well on PH2 dataset has been determined by the experiment.

**Table 4.3: Effect of different batch sizes and learning rate on model over PH2 dataset**

Batch size Learning Rate	16			18			32		
	Acc	Jac	Dc	Acc	Jac	Dc	Acc	Jac	Dc
0.001	0.965 7	0.962 8	0.930 3	0.950 9	0.957 3	0.9137	<b>0.967</b> <b>9</b>	<b>0.969</b> <b>7</b>	0.9372
0.003	0.955 2	0.959 7	0.922 6	0.964 2	0.968 7	0.9387	0.964 4	0.968 2	<b>0.9427</b>

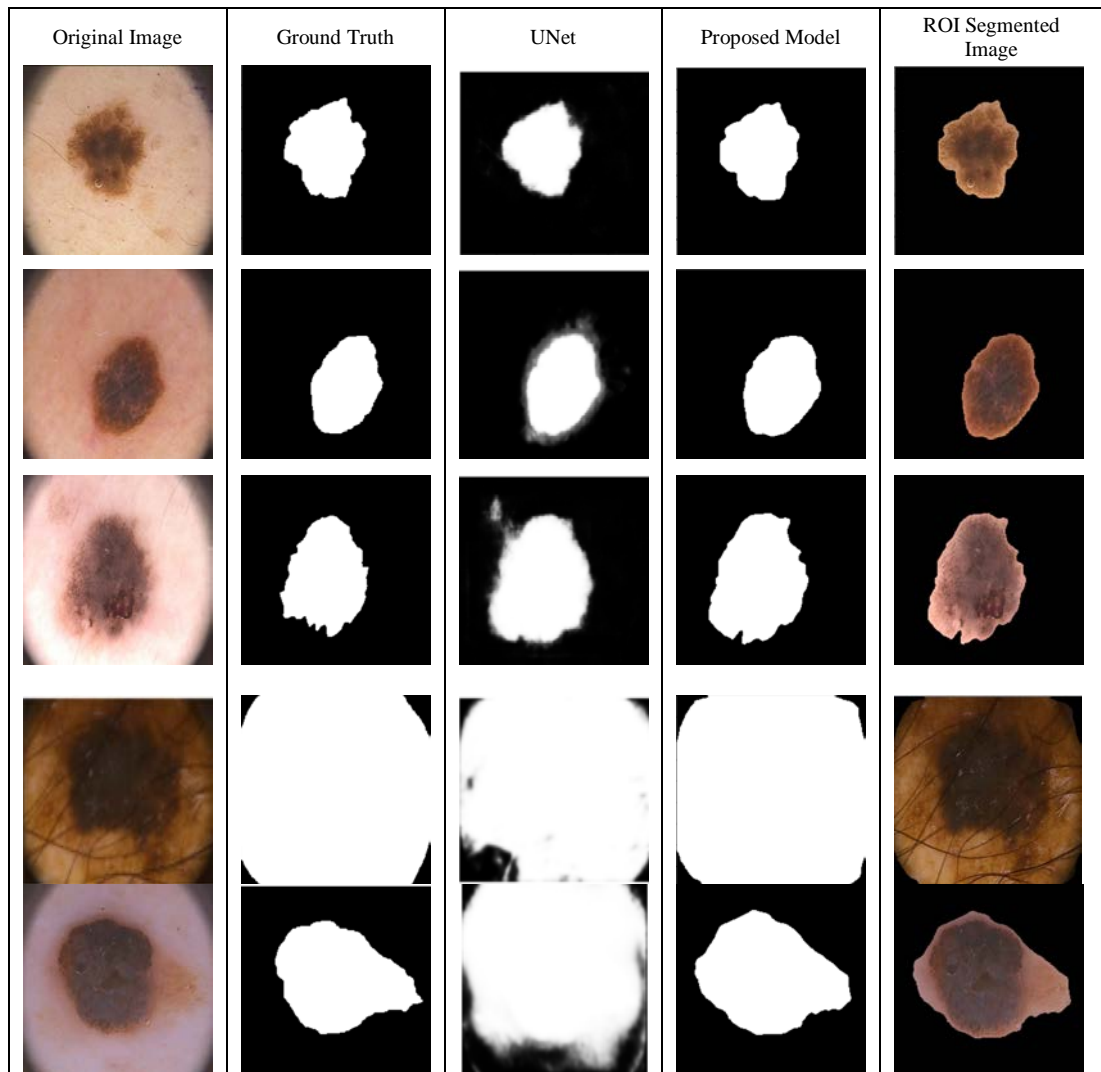
## 4.6.2 QUALITATIVE PERFORMANCE

The qualitative outcomes of the proposed segmentation model has been verified through the comparison of the quality of the segmented mask generated with the ground truth mask, along with the ROI and NROI segmented images, as shown in Fig. 4.3.



**Fig. 4.3:** Qualitative Performance of the Proposed Segmentation Model

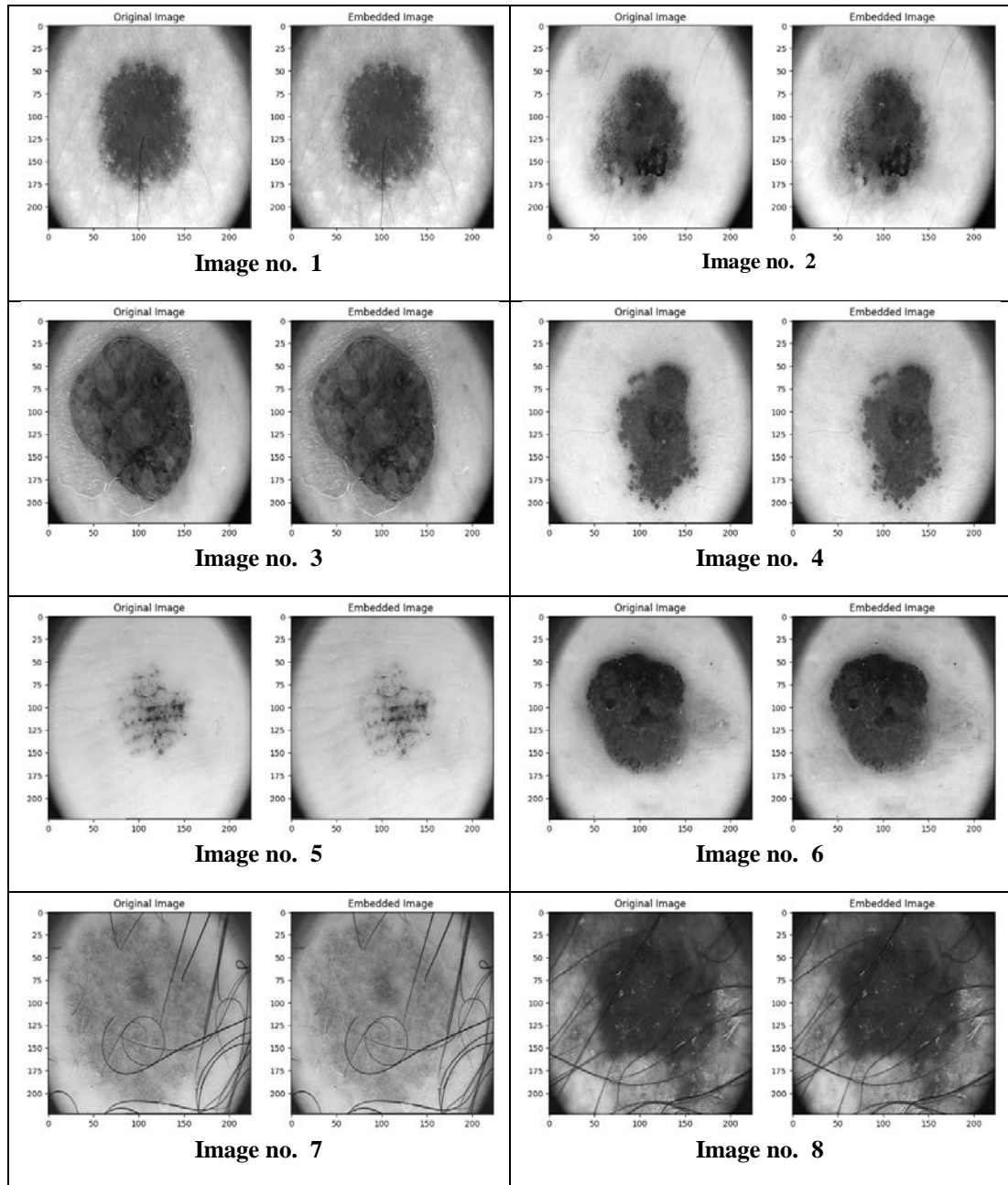
The effectiveness of the proposed model has been assessed by comparing to other existing models. Figure 4.4 shows that the UNet model misclassified background pixels as skin lesion pixels, resulting in white patches around the segmented ROI. The proposed model enhances the segmentation result closer to Ground Truth by incorporating both residual and attention mechanism.



**Fig. 4.4: Skin lesion images alongside their ground truth in first and second column respectively. The predicted segmented mask by UNet by O. Ronneberger et al. (2015) and the Proposed Model in third and fourth column followed by ROI segmented image of the proposed model in fifth column.**

The performance of the data hiding method can be seen from the Fig. 4.5 below. It highlights the visual comparison between the original image and the embedded complete image (stego-image) after embedding the message in the ROI part. The original and stego-image are indistinguishable to the human eyes, ensuring the effectiveness of the embedding process. It became impossible for an attacker to detect any presence of secret data. It is interesting to note that the contrast of some of the

images has been enhanced due to embedding. Nevertheless, there is a need to conduct more studies to support this fact and the underlying mechanism.



**Fig. 4.5: Visual Comparison between Original image and Embedded Image**

### **4.6.3 QUANTITATIVE PERFORMANCE**

Table 4.4 provides a summary of the effectiveness of the proposed segmentation model for training, validation, and testing on the PH2 dataset using evaluation metrics defined

in section 4.4 and loss parameter. Based on the metrics value, it can be concluded that the model has not overfitted, and it did well on both the training and test sets.

**Table 4.4: Proposed Model Performance**

	<b>TRAINING</b>	<b>VALIDATION</b>	<b>TEST</b>
<b>ACC</b>	0.9922	0.9872	0.9679
<b>DC</b>	0.9781	0.9683	0.9372
<b>JAC</b>	0.9929	0.9884	0.9697
<b>SE</b>	0.9901	0.9805	0.9582
<b>LOSS</b>	0.0077	0.0122	0.0308

The performance of the proposed model has been evaluated against other UNet models which has been published previously on the PH2 dataset by using the standard metrics. UNet is the base model for skin lesion segmentation but it is not able to perform well. First, residual connections have been added to base UNet to overcome the degradation problem, which slightly increases its accuracy. Then attention gates were added to improve the segmentation results by highlighting only relevant region. Then, the Unet architecture has been redesigned to UNet 3+ by incorporating full scale skip connections. It enhanced the accuracy to 0.9579. But, CBAM+Ref UNet3+ surpasses this by reducing the redundant skip connections. However, it did not optimize the loss function and could not able to handle deformations around skin lesion boundaries. Therefore, the proposed model use dilated UNet3+ with optimized loss function to overcome these issues and to enhance the segmentation accuracy. Table 4.5 shows that the proposed model performs better than other UNet variants in every aspect.

**Table 4.5: Comparison of Proposed Model with other Models**

<b>Method</b>	<b>ACC</b>	<b>DC</b>	<b>JAC</b>	<b>SE</b>
U-Net [36]	0.9225	0.8761	0.7793	0.8165
ResU-Net [49]	0.9258	0.8856	0.8085	0.8213
AU-Net [50]	0.9389	0.9012	0.8445	0.8787
UNet3+ [38]	0.9579	0.8762	0.7976	0.8990
CBAM+Ref-UNet3+ [8]	0.9604	0.8848	0.8136	0.9010
ARU-Net MD [10]	0.9656	0.9265	0.8857	0.9161
<b>PROPOSED</b>	<b>0.9679</b>	<b>0.9372</b>	<b>0.9697</b>	<b>0.9582</b>

The following Table 4.6 summarizes the performance of data hiding method by listing the total number of bits embedded, PSNR value of the stego-image and SSIM

comparison between original image and embedded image. It can be noted that images with hairs in the ROI part has less embedding rate than other images.

**Table 4.6: Performance of data hiding method**

Image No.	Number of bits embedded	PSNR value	SSIM
1	18163	29.50	0.9877
2	18215	31.34	0.9884
3	14999	30.71	0.9895
4	18878	30.99	0.9925
5	21981	31.86	0.9881
6	18008	31.96	0.9871
7	8876	31.28	0.9842
8	905	46.98	0.9975

#### 4.6.4 COMPARISON WITH THE STATE-OF-THE-ART MODEL

The proposed model has been compared with the current state-of-the-art models for the skin lesion segmentation in below Table 4.7. The comparison shows that the proposed model outperformed state-of-the-art models across all metrics.

**Table 4.7: Comparison of Proposed Segmentation Model with State-of-the-Art**

Method	ACC	DC	JAC	SE
FCN [1]	0.9282	0.8903	0.8022	0.9030
UNet [36]	0.9255	0.8761	0.7793	0.8165
SegNet [35]	0.9336	0.8936	0.8077	0.8653
UNet++ [51]	0.9535	0.9281	0.8711	0.9484
UNet 3+ [38]	0.9570	0.8762	0.7976	0.8990
<b>Proposed</b>	<b>0.9679</b>	<b>0.9372</b>	<b>0.9697</b>	<b>0.9582</b>

The mean PSNR of the proposed system is 33.08 dB and the mean SSIM is 0.9893 while the method used by Gao et al. [37] has mean PSNR value is 34.24 dB and mean SSIM is 0.8578.

## **CHAPTER 5**

### **CONCLUSIONS AND FUTURE WORK**

#### **5.1 CONCLUSIONS**

The proposed system has introduced a novel and robust skin lesion segmentation module based on baseline U-Net architecture. By incorporating residual connections, an attention module and employing dilation rate, significant improvements over the baseline model have been achieved. The proposed model has been trained and tested on PH2 dataset and achieved excellent performance with accuracy of 0.9679. It has been able to surpass the state-of-the-art model for skin lesion segmentation.

The proposed segmentation model has been able to deal with complex skin lesion images with artifacts, hairs and bubbles, without any pre-processing of images. Some of the predicted segmented masks have been more accurate than their corresponding ground truth due to the efficient learning of deep neural network model. Some of the unwanted contours that the model has not been able to detect has been removed using predefined python library. The overall result produced by the segmentation model mark the significant advancement in the field of image processing.

Some amount of data can be embedded within the ROI of the medical image using the proposed system. The SSIM value indicates that the image is not distorted much. It recover the secret data without any loss. Thus, maintaining the security and integrity of the data. However, the embedding capacity of the proposed system is still low. Moreover, some more pre-processing technique like removal of hairs can be deployed to improve the embedding capacity of images with hairs. The proposed segmentation model can be integrated with high performance data hiding model to get the robust security system for medical data.

## 5.2 FUTURE WORKS

The following are the possible future research that can further enhance the performance and capabilities if the proposed system.

- For the future advancement, the model can be trained on different body organs dataset to measure its versatility and performance across various medical imaging.
- Efforts can be made to overcome the problem of unwanted contours by the model itself, without using the CV2 library. Refining Model so that it should be able to segment ROI with continuous irregular boundaries effectively.
- There is a scope of integrating machine learning or deep learning models for computing the local complexity of the pixels to automate the task. This may prove to be effective for large size images.
- The proposed system only hides data within the smooth pixels of ROI. To improve the embedding capacity, more data can be hidden in the NROI part.
- Since the PSNR value of the proposed system is not that much satisfactory, there is a room for improvement so as to improve the image quality.
- Data encryption technique can also be incorporated to add an additional layer of security to the data against unauthorized access.



## REFERENCES

- [1] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431–3440, 2015.
- [2] L. Bi, J. Kim, E. Ahn, D. Feng, and M. Fulham, “Semi-automatic skin lesion segmentation via fully convolutional networks,” in 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), *IEEE*, pp. 561–564, 2017.
- [3] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, “Automated melanoma recognition in dermoscopy images via very deep residual networks,” *IEEE transactions on medical imaging*, vol. 36, no. 4, pp. 994–1004, 2016.
- [4] Y. Yuan, M. Chao, and Y.-C. Lo, “Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance,” *IEEE transactions on medical imaging*, vol. 36, no. 9, pp. 1876–1886, 2017.
- [5] B. S. Lin, K. Michael, S. Kalra, and H. R. Tizhoosh, “Skin lesion segmentation: U-nets versus clustering,” in 2017 IEEE Symposium Series on Computational Intelligence (SSCI), *IEEE*, pp. 1–7, 2017.
- [6] S. Vesal, N. Ravikumar, and A. Maier, “Skinnet: A deep learning framework for skin lesion segmentation,” in 2018 IEEE nuclear science symposium and medical imaging conference proceedings (NSS/MIC), *IEEE*, pp. 1–3, 2018.
- [7] X. Tong, J. Wei, B. Sun, S. Su, Z. Zuo, and P. Wu, “Ascu-net: Attention gate, spatial and channel attention u-net for skin lesion segmentation,” *Diagnostics*, vol. 11, no. 3, p. 501, 2021.
- [8] Y. Xu, S. Hou, X. Wang, D. Li, and L. Lu, “A medical image segmentation method based on improved unet 3+ network,” *Diagnostics*, vol. 13, no. 3, p. 576, 2023.
- [9] K. B. Nampalle, A. Pundhir, P. R. Jupudi, and B. Raman, “Towards improved u-net for efficient skin lesion segmentation,” *Multimedia Tools and Applications*, pp. 1–18, 2024.

- [10] R. Kaur and S. Kaur, "Automatic skin lesion segmentation using attention residual u-net with improved encoder-decoder architecture," *Multimedia Tools and Applications*, pp. 1–27, 2024.
- [11] P. S. Govind and M. Judy, "A secure framework for remote diagnosis in health care: A high capacity reversible data hiding technique for medical images," *Computers & Electrical Engineering*, vol. 89, p. 106933, 2021.
- [12] C. Sekar, V. R. Falmari, and M. Brindha, "Secure IoT-enabled sharing of digital medical records: An integrated approach with reversible data hiding, symmetric cryptosystem, and IPFS," *Internet of Things*, vol. 24, p. 100958, 2023.
- [13] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 890–896, 2003.
- [14] M. Fallahpour, D. Megias, M. Ghanbari, "High capacity, reversible data hiding in medical images," 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, pp. 4241–4244, 2009.
- [15] L.-C. Huang, L.-Y. Tseng, and M.-S. Hwang, "A reversible data hiding method by histogram shifting in high quality medical images," *Journal of Systems and Software*, vol. 86, no. 3, pp. 716–727, 2013.
- [16] H.-T. Wu, Q. Huang, Y. Cheung, L. Xu, and S. Tang, "Reversible contrast enhancement for medical images with background segmentation," *IET Image Processing*, vol. 14, no. 2, pp. 327–336, 2020.
- [17] P.-Y. Pai, C.-C. Chang, Y.-K. Chan, and C.-M. Liu, "An ROI-based medical image hiding method," *International Journal of Innovative Computing, Information, and Control*, vol. 1, no. 8, pp. 4521–4533, 2012.
- [18] N. Ostu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, p. 62, 1979.
- [19] Y. Yang, W. Zhang, D. Liang, and N. Yu, "A ROI-based high capacity reversible data hiding scheme with contrast enhancement for medical images," *Multimedia Tools and Applications*, vol. 77, pp. 18043–18065, 2018.

- [20] G. Gao, X. Wan, S. Yao, Z. Cui, C. Zhou, and X. Sun, "Reversible data hiding with contrast enhancement and tamper localization for medical images," *Information Sciences*, vol. 385, pp. 250–265, 2017.
- [21] S. Gandhi and R. Kumar, "A high-capacity reversible data hiding with contrast enhancement and brightness preservation for medical images," *Multimedia Tools and Applications*, pp. 1–26, 2024.
- [22] S.Z. Oo, et al., "Brain tumor detection and segmentation using watershed segmentation and morphological operation", *International Journal of Research in Engineering and Technology*, vol.3, pp. 367-374, 2014.
- [23] S. A. Parah, F. Ahad, J. A. Sheikh, N. A. Loan, and G. M. Bhat, "A new reversible and high capacity data hiding technique for E-healthcare applications," *Multimedia Tools and Applications*, vol. 76, pp. 3943–3975, 2017.
- [24] Y.-C. Lin and T.-S. Li, "Reversible Image Data Hiding Using Quad-tree Segmentation and Histogram Shifting.," *Journal of Multimedia*, vol. 6, no. 4, pp. 349–358, 2011.
- [25] A. G. Salman , "Reversible Data Hiding Technique on Jpeg Image by Quad-Tree based Segmentation and Histogram Shifting Method Based on Android," *Procedia Computer Science*, vol. 59, pp. 530–539, 2015.
- [26] P. Chaithanya and V. Srujana, "Reversible Image Data Hiding using Quad Tree Segmentation," *International Journal of Emerging Technologies and Innovation Research*, vol. 7, no. 7, pp. 521-523, 2022.
- [27] S. Lakshmanan and M. Rani, "Reversible data hiding in medical images using edge detection and difference expansion technique," *Journal of Computational and Theoretical Nanoscience*, vol. 15, no. 6–7, pp. 2400–2404, 2018.
- [28] Y.W. Lu, et al. "Segmentation method for medical image based on improved GrabCut", *International Journal of Imaging Systems and Technology*, vol. 27, pp. 383-390, 2017.
- [29] A. Rai and H. V. Singh, "SVM based robust watermarking for enhanced medical image security," *Multimedia Tools and Applications*, vol. 76, no. 18, pp. 18605–18618, 2017.

- [30] M. Li, L. Wang, and H. Fan, "Privacy-preserved data hiding using compressive sensing and fuzzy C-means clustering," *International Journal of Distributed Sensor Networks*, vol. 16, no. 2, p. 1550147720908748, 2020.
- [31] R. Meng, Q. Cui, Z. Zhoul, C. Yuan, and X. Sun, "A Novel Steganography Algorithm Based on Instance Segmentation," *Computers, Materials & Continua*, vol. 63, no. 1, 2020.
- [32] H. Saidi, O. Tibermacine, and A. Elhadad, "High-capacity data hiding for medical images based on the mask-RCNN model," *Scientific Reports*, vol. 14, no. 1, p. 7166, 2024.
- [33] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.
- [34] N. Pan, J. Qin, Y. Tan, X. Xiang, and G. Hou, "A video coverless information hiding algorithm based on semantic segmentation," *EURASIP Journal on Image and Video Processing*, vol. 2020, pp. 1–18, 2020.
- [35] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [36] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234–241, 2015.
- [37] G. Gao, H. Zhang, Z. Xia, X. Luo, and Y.-Q. Shi, "Reversible data hiding-based contrast enhancement with multi-group stretching for ROI of medical image," *IEEE Transactions on Multimedia*, vol. 26, pp. 3909–3923, 2023.
- [38] H. Huang et al., "Unet 3+: A full-scale connected unet for medical image segmentation," in *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 1055–1059, 2020.

- [39] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in Proceedings of the European conference on computer vision (ECCV), pp. 3–19, 2018.
- [40] P. Amrit, K. Singh, N. Baranwal, A. Singh, J. Singh, and H. Zhou, "Deep learning-based segmentation for medical data hiding with Galois field," *Neural Computing and Applications*, pp. 1–16, 2023.
- [41] H. Shi, Z. Zhou, J. Qin, H. Sun, and Y. Ren, "A separable privacy-preserving technique based on reversible medical data hiding in plaintext encrypted images using neural network," *Multimedia Tools and Applications*, pp. 1–26, 2024.
- [42] P. Luc, C. Couprie, S. Chintala, and J. Verbeek, "Semantic segmentation using adversarial networks," *arXiv preprint arXiv:1611.08408*, 2016.
- [43] C. Annadurai, I. Nelson, K. N. Devi, R. Manikandan, and A. H. Gandomi, "Image Watermarking Based Data Hiding by Discrete Wavelet Transform Quantization Model with Convolutional Generative Adversarial Architectures," *Applied Sciences*, vol. 13, no. 2, p. 804, 2023.
- [44] P.-Y. Pai, C.-C. Chang, Y.-K. Chan, and M.-H. Tsai, "An adaptable threshold detector," *Information Sciences*, vol. 181, no. 8, pp. 1463–1483, 2011.
- [45] P. Chowdhuri, P. Pal, and T. Si, "A novel steganographic technique for medical image using SVM and IWT," *Multimedia Tools and Applications*, vol. 82, no. 13, pp. 20497–20516, 2023.
- [46] K. Balasamy and S. Suganyadevi, "A fuzzy based ROI selection for encryption and watermarking in medical image using DWT and SVD," *Multimedia tools and applications*, vol. 80, no. 5, pp. 7167–7186, 2021.
- [47] M. Kolarik, R. Burget, and K. Riha, "Upsampling algorithms for autoencoder segmentation neural networks: A comparison study," in 2019 11th International Congress on Ultra-Modern Telecommunications and Control Systems and Workshops (ICUMT), *IEEE*, pp. 1–5, 2019.

- [48] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y. Q. Shi, “Reversible watermarking algorithm using sorting and prediction,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 989–999, 2009.
- [49] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal, and J. Rozeira, “PH 2-A dermoscopic image database for research and benchmarking,” in 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2013, pp. 5437–5440.
- [50] O. Oktay, J. Schlemper, L. L. Folgoc, et al., “Attention u-net: Learning where to look for the pancreas,” *arXiv preprint arXiv: 1804.03999*, 2018.
- [51] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: A nested u-net architecture for medical image segmentation,” in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, pp. 3–11, 2018.

# LIST OF PUBLICATIONS

## **Paper 1**

**Title:** An Overview on Image Segmentation Techniques for Reversible Data Hiding

**Abstract:** The fields of image processing and computer vision have witnessed significant growth due to the proliferation of digital images across diverse domains. Image Segmentation is the fundamental task in digital image processing, finding applications in pivotal areas such as medical imaging, covert communication, autonomous driving, satellite imaging, among others. One particularly intriguing application of image segmentation lies in Reversible Data Hiding, where the delineation of the main Region of Interest (ROI) and Non-Region of Interest (NROI) using segmentation plays a crucial role. Over the last two decades, various studies focussed on developing an effective data hiding approach, which can embed secret data within ROI and NROI part of image while ensuring its quality. A detailed survey has been conducted that meticulously examines different segmentation techniques, along with its usage in Reversible Data Hiding. The surveyed techniques have been categorized systematically into three main classes: i) Traditional segmentation techniques, encompassing a spectrum of approaches like thresholding, region-based and edge detection based ii) Machine Learning based approach consisting of Clustering, Support Vector Machine (SVM) and iii) Deep learning based technique, propelled by Convolutional Neural Networks (CNNs) that have emerged as a transformative paradigm, revolutionizing segmentation tasks with their ability to learn complex images.

**Journal:** International Journal of Mathematics, Engineering and Management Sciences

**Indexing:** Impact Factor: 1.6 (Clarivate Analytics, Web of Science), Emerging Sources Citation Index (ESCI), Scopus, Google Scholar

**Status:** Accepted



Rasika Gupta &lt;rashika99@gmail.com&gt;

---

**Your manuscript is accepted - International Journal of Mathematical, Engineering and Management Sciences**

1 message

IJMEMS &lt;ijmems@ijmems.in&gt;

To: rashika99@gmail.com

Cc: noreply@ijmems.in

Tue, May 21, 2024 at 1:50 PM

---

**International Journal of Mathematical, Engineering and Management Sciences**

Dear **Ms. Rasika Gupta , WITH CC TO ALL AUTHORS,**

We are pleased to inform you that your article entitled An Overview on Image Segmentation Techniques for Reversible Data Hiding has been **accepted** for the publication in the International Journal of Mathematical, Engineering and Management Sciences

Very soon, you will receive the proof of your article.

**Manuscript ID:** R2-IJMEMS-24-0190

**Manuscript Title:** An Overview on Image Segmentation Techniques for Reversible Data Hiding

**Abstract:** The fields of image processing and computer vision have witnessed significant growth due to the proliferation of digital images across diverse domains. Image Segmentation is the fundamental task in digital image processing, finding applications in pivotal areas such as medical imaging, covert communication, autonomous driving, satellite imaging, among others. One particularly intriguing application of image segmentation lies in Reversible Data Hiding, where the delineation of the main Region of Interest (ROI) and Non-Region of Interest (NROI) using segmentation plays a crucial role. Over the last two decades, various studies focussed on developing an effective data hiding approach, which can embed secret data within ROI and NROI part of image while ensuring its quality. A detailed survey has been conducted that meticulously examines different segmentation techniques, along with its usage in Reversible Data Hiding. The surveyed techniques have been categorized systematically into three main classes: i) Traditional segmentation techniques, encompassing a spectrum of approaches like thresholding, region-based and edge detection based ii) Machine Learning based approach consisting of Clustering, Support Vector Machine (SVM) and iii) Deep learning based technique, propelled by Convolutional Neural Networks (CNNs) that have emerged as a transformative paradigm, revolutionizing segmentation tasks with their ability to learn complex images.

**Keywords:** Image Segmentation, Convolutional Neural Network, Encoder-decoder Model, Dilated Convolution Model, ROI segmentation, Reversible Data Hiding.

Kindly consider the International Journal of Mathematical, Engineering and Management Sciences for your future manuscripts.

Kindly cite International Journal of Mathematical, Engineering and Management Sciences published papers in your future research.

Regards,

Thank you for your cooperation and with regards.

Editor



## **Paper 2**

- Title:** An Improved U-Shaped Network for ROI Segmentation of Skin Lesion images
- Abstract:** Skin cancer, a serious public health concern across the world, is a deadly disease if not diagnosed at an early stage. It is often diagnosed through dermoscopic images that usually have low contrast, irregular boundaries and contains irrelevant artifacts. Due to these factors, the segmentation of Region of Interest (ROI) for effective diagnosis of this disease becomes a challenging task. However, UNet, a U-Shaped Network has proven to be an efficient segmentation model for medical image segmentation, but the result mostly suffers discrepancies with the ground truth segmented mask. Therefore, this paper introduces an improved UNet 3+ architecture with residual block at both encoder and decoder to handle gradient degradation in this deep network, incorporates attention module to focus on relevant features and reduces redundant skip connections. The improved model has been evaluated on publicly available PH2 dataset and achieved an accuracy of 96.62%, dice score of 93.84% and jaccard index of 96.79% on test data. The model surpasses other state-of-the-art UNet models for skin lesion segmentation.
- Conference :** IEEE 2<sup>nd</sup> World Conference on Communications and Computing (WCONF), 2024, Kalinga University, Raipur, Chhattisgarh, India.
- Indexing:** Scopus
- Status:** Accepted



Rasika Gupta &lt;rashika99@gmail.com&gt;

---

**ACCEPTANCE NOTIFICATION - 2nd IEEE WCONF 2024**

1 message

---

**Microsoft CMT** <email@msr-cmt.org>  
Reply-To: D k Gupta <deepak\_gupta@gibds.org>  
To: Rasika Gupta <rashika99@gmail.com>

Wed, May 29, 2024 at 1:03 PM

Dear Rasika Gupta

Paper ID / Submission ID : 2306

Title : An Improved U-Shaped Network for ROI Segmentation of Skin Lesion images

Greeting from WCONF 2024

We are pleased to inform you that your paper has been accepted for the Oral Presentation as a full paper for the- "IEEE 2nd 2024 World conference on communication and computing (WCONF), Raipur, India with following reviewers' comment.

All accepted and presented papers will be submitted to IEEE Xplore for the further publication and will be indexed by Ei Compendex and Scopus Indexing.

You should finish the registration before deadline, or you will be deemed to withdraw your paper:

Complete the Registration Process (The last date of payment Registration is 31 MAY 2024)

Payment Links

For Indian Authors: <https://rzp.io/l/fbbVUJABD>

For Foreign Authors: <https://in.explara.com/e/wconf-2024>  
(select stripe payment while doing payment )

Further steps like IEEE PDF xpress and E copyright will be given later once registration is over after the deadline.

Note :

1. Any changes with the Author name, Affiliation and content of paper will not be allowed after acceptance.
2. This is Hybrid Conference, both online and physical presentation mode is available,

Note : No any major revisions in paper.

The reviews are below

===== Review 1 =====

\*\*\* Relevance and timeliness: Rate the importance and timeliness of the topic addressed in the paper within its area of research.

Good (4)

\*\*\* Technical content and scientific rigour: Rate the technical content of the paper (e.g.: completeness of the

analysis or simulation study, thoroughness of the treatise, accuracy of the models, etc.), its soundness and scientific rigour.

Valid work but limited contribution. (4)

\*\*\* Novelty and originality: Rate the novelty and originality of the ideas or results presented in the paper.

Some interesting ideas and results on a subject well investigated. (3)

\*\*\* Quality of presentation: Rate the paper organization, the clearness of text and figures, the completeness and accuracy of references.

Well written. (4)

\*\*\* Strong aspects: Comments to the author: what are the strong aspects of the paper?

In this paper, a new application is proposed. In addition, the paper discusses the different aspects of technology in the field of Programming testing to guarantee programming quality

\*\*\* Weak aspects: Comments to the author: what are the weak aspects of the paper?

\*\*\* Recommended changes: Please indicate any changes that should be made to the paper if accepted.

The paper should be more result oriented.

===== Review 2 =====

\*\*\* Relevance and timeliness: Rate the importance and timeliness of the topic addressed in the paper within its area of research.

Acceptable (3)

\*\*\* Technical content and scientific rigour: Rate the technical content of the paper (e.g.: completeness of the

Analysis or simulation study, thoroughness of the treatise, accuracy of the models, etc.), its soundness and scientific rigour.

Valid work but limited contribution. (3)

\*\*\* Novelty and originality: Rate the novelty and originality of the ideas or results presented in the paper.

Some interesting ideas and results on a subject well investigated. (3)

\*\*\* Quality of presentation: Rate the paper organization, the clearness of text and figures, the completeness and accuracy of references.

Readable, but revision is needed in some parts. (3)

\*\*\* Strong aspects: Comments to the author: what are the strong aspects of the paper

The paper reviews different aspects of the concern

\*\*\* Weak aspects: Comments to the author: what are the weak aspects of the paper?

The future contribution of the paper should be mentioned in the paper

The presentation quality of the paper has to be improved significantly. The details of the author's contribution are too little.

\*\*\* Recommended changes: Recommended changes. Please indicate any changes that should be made to the paper if accepted...

The presentation quality of the paper has to be improved significantly

The detail of simulation/Result is too little.

===== Review 3 =====

\*\*\* Relevance and timeliness: Rate the importance and timeliness of the topic addressed in the paper within its area of research.

Good (4)

\*\*\* Technical content and scientific rigour: Rate the technical content of the paper (e.g.: completeness of the

analysis or simulation study, thoroughness of the treatise, accuracy of the models, etc.), its soundness and scientific rigour.

Marginal work and simple contribution. Some flaws. (2)

\*\*\* Novelty and originality: Rate the novelty and originality of the ideas or results presented in the paper.

Minor variations on a well investigated subject. (2)

\*\*\* Quality of presentation: Rate the paper organization, the clearness of text and figures, the completeness and accuracy of references.

Okay. Well Written

\*\*\* Strong aspects: Comments to the author: what are the strong aspects of the paper

In this research, a new scenario of the technology has been introduced; research may be interesting for readers in this area.

\*\*\* Weak aspects: Comments to the author: what are the weak aspects of the paper?

1. The presentation should be substantially improved.

2. The presentation of the proposed algorithm is too conceptual, and details of how to operate the proposed algorithm in practice should be clearly elaborated.

\*\*\* Recommended changes: Please indicate any changes that should be made to the paper if accepted.

NO ANY MAJOR REVISIONS

Thanks, and regards,  
Technical Program Committee Chair  
WCONF 2024  
[wconfer@gmail.com](mailto:wconfer@gmail.com)  
+91-7483158331

To stop receiving conference emails, you can check the 'Do not send me conference email' box from your User Profile.

Microsoft respects your privacy. To learn more, please read our [Privacy Statement](#).

Microsoft Corporation

ASIAN EDUCATIONAL MANAGEMENT & RESEARCH  
CONSULTANT

Invoicing and payments  
powered by

Payment Receipt Transaction Reference: pay\_OGitbLHXEKAkVZ

This is a payment receipt for your transaction on 2nd WCONF 2024

AMOUNT PAID ₹ 8,905.32

ISSUED TO  
rashika99@gmail.com  
+919968211685

PAID ON  
30 May 2024

Registered Author Name  
Rasika Gupta

DESCRIPTION	UNIT PRICE	QTY	AMOUNT
Non IEEE Member Full Paper Registration	₹ 8,700.00	1	₹ 8,700.00
Total			₹ 8,700.00
Amount Paid			₹ 8,905.32

NO REFUND POLICY

### **Paper 3**

<b>Title:</b>	Review on UNet Architecture for Image Segmentation
<b>Abstract:</b>	<p>In the rapidly evolving field of computer vision, Image Segmentation plays a significant role in various applications such as medical image analysis, image compression, satellite imaging, among others. Among the many segmentation approaches, the UNet models emerged as the most versatile technique for image segmentation. This review explores the UNet architecture, covering its variants, limitations and performance on benchmark datasets. Furthermore, it also highlights recent advancements and innovations in UNet architecture such as dense connections, attention mechanism, deformation convolution to increase the accuracy and efficiency of the model. Through this analysis, the author elucidates the potential future works that can be done in the field of image segmentation.</p>
<b>Conference :</b>	NATIONAL CONFERENCE ON ADVANCED COMPUTER SCIENCE AND INFORMATION TECHNOLOGY (NCACSI - 24), Guwahati, India.
<b>Indexing:</b>	Scopus
<b>Status:</b>	Accepted

**NATIONAL CONFERENCE ON ADVANCED COMPUTER  
SCIENCE AND INFORMATION TECHNOLOGY (NCACSI - 24)**

**19th May 2024 Guwahati, India**

## **Acceptance Letter**

**Authors Name:** Rasika Gupta

**Dear Authors,**

We are pleased to inform you that your paper has been accepted by the review committee for Oral / Poster Presentation at the **NATIONAL CONFERENCE ON ADVANCED COMPUTER SCIENCE AND INFORMATION TECHNOLOGY (NCACSI - 24)**

**Article Title:** **Review on UNet Architecture for Image Segmentation**

**Paper ID:** **National Conference\_2576439**

This conference will be held on **19th May 2024 in Guwahati, India**

Your paper will be published in the conference proceeding and Well reputed journal after registration.

Please register as soon as possible in order to secure your participation:

<https://www.nationalconference.in/event/registration.php?id=2481785>

You are requested to release the payment and mail us the screen of successful payment release with your name and title of paper to confirm your registration.

Sincerely,



**Dr. Tara Srivastava**  
National Conference




Rasika Gupta &lt;rashika99@gmail.com&gt;

## Transaction Acknowledgment - Successful Payment

1 message

**National Conference** <info@nationalconference.in>  
Reply-To: National Conference <info@nationalconference.in>  
To: Rasika Gupta <rashika99@gmail.com>

Sat, May 11, 2024 at 5:20 PM



**Dear Rasika Gupta,**

We are pleased to inform you that your online payment transaction has been successfully processed. The details of your payment are as follows:

**Order ID:** 663f5b2eb9923

**Reference ID:** 113284110759

**Amount:** 4725.00 INR

**Name:** Rasika Gupta

**Email ID:** [rashika99@gmail.com](mailto:rashika99@gmail.com)

**Phone Number:** 919968211685

Thank you for choosing **National Conference**. We appreciate your prompt payment.

For assistance or inquiries, feel free to reach out to us at [info@nationalconference.in](mailto:info@nationalconference.in) or **+91 9677007228**.

**Best Regards,**  
**National Conference**  
[info@nationalconference.in](mailto:info@nationalconference.in)  
**+91 9677007228**



PAPER NAME

**thesis\_plag.docx**

AUTHOR

**Rasika Gupta**

WORD COUNT

**12292 Words**

CHARACTER COUNT

**66672 Characters**

PAGE COUNT

**55 Pages**

FILE SIZE

**6.5MB**

SUBMISSION DATE

**May 29, 2024 11:52 AM GMT+5:30**

REPORT DATE

**May 29, 2024 11:53 AM GMT+5:30**

### ● 6% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 2% Internet database
- 4% Publications database
- Crossref database
- Crossref Posted Content database
- 3% Submitted Works database

### ● Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less than 8 words)

## ● 6% Overall Similarity

Top sources found in the following databases:

- 2% Internet database
- Crossref database
- 3% Submitted Works database
- 4% Publications database
- Crossref Posted Content database

### TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	<b>Rajdeep Kaur, Sukhjeet Kaur. "Automatic skin lesion segmentation usi...</b> Crossref	1%
2	<b>th-koeln on 2024-05-09</b> Submitted works	<1%
3	<b>mafiadoc.com</b> Internet	<1%
4	<b>repository.rice.edu</b> Internet	<1%
5	<b>rest.neptune-prod.its.unimelb.edu.au</b> Internet	<1%
6	<b>Higher Education Commission Pakistan on 2014-07-12</b> Submitted works	<1%
7	<b>erepository.uonbi.ac.ke</b> Internet	<1%
8	<b>Higher Education Commission Pakistan on 2020-09-03</b> Submitted works	<1%

9	<b>Lecture Notes in Computer Science, 2015.</b>	<1%
	Crossref	
10	<b>The University of Manchester on 2020-05-11</b>	<1%
	Submitted works	
11	<b>"Computer Vision, Pattern Recognition, Image Processing, and Graphic...</b>	<1%
	Crossref	
12	<b>Heriot-Watt University on 2023-12-12</b>	<1%
	Submitted works	
13	<b>Radboud Universiteit Nijmegen on 2020-08-20</b>	<1%
	Submitted works	
14	<b>Qihan Jiao, Zhi Liu, Linwei Ye, Yang Wang. "Weakly labeled fine-graine...</b>	<1%
	Crossref	
15	<b>frontiersin.org</b>	<1%
	Internet	
16	<b>eprints.qut.edu.au</b>	<1%
	Internet	
17	<b>"Intelligent Data Engineering and Analytics", Springer Science and Busi...</b>	<1%
	Crossref	
18	<b>Arficho, Tsegaye. "Artificial Intelligence and Deep Learning in the Early ...</b>	<1%
	Publication	
19	<b>research.libraries.wsu.edu</b>	<1%
	Internet	
20	<b>Agham, Vinit, and Tareek Pattewar. "Data hiding technique by using RG...</b>	<1%
	Crossref	

- 21 **Cuixiao Liang, Juntao Xiong, Zhenhui Zheng, Zhuo Zhong, Zhonghang ...** <1%  
Crossref
- 
- 22 **Gayatri Vidya Parishad College of Engineering (Autonomous) on 2019-...** <1%  
Submitted works
- 
- 23 **Visvesvaraya Technological University on 2016-01-11** <1%  
Submitted works
- 
- 24 **"CARS 2017—Computer Assisted Radiology and Surgery Proceedings ...** <1%  
Crossref
- 
- 25 **Haixing Li, Haibo Luo, Wang Huan, Zelin Shi, Chongnan Yan, Lanbo Wa...** <1%  
Crossref
- 
- 26 **Kush Shrivastava, Shishir Kumar, Deepak Kumar Jain. "An effective ap...** <1%  
Crossref
- 
- 27 **Muktar Bedaso, Antu Hussein, Chala Diriba. "Developing Sign Languag...** <1%  
Crossref posted content
- 
- 28 **RMIT University on 2023-04-02** <1%  
Submitted works
- 
- 29 **Showmick Guha Paul, Arpa Saha, Md Assaduzzaman. "A real-time dee...** <1%  
Crossref
- 
- 30 **Sobha Xavier P, Sathish P K, Raju G. "Advancing Brain Tumor Segmen...** <1%  
Crossref
- 
- 31 **University of Bristol on 2014-11-07** <1%  
Submitted works
- 
- 32 **downloads.hindawi.com** <1%  
Internet

- 33 "Digital TV and Wireless Multimedia Communication", Springer Scienc... <1%  
Crossref
- 
- 34 "Image Analysis and Processing – ICIAP 2005", Springer Science and B... <1%  
Crossref
- 
- 35 "Innovations in Bio-Inspired Computing and Applications", Springer Sci... <1%  
Crossref
- 
- 36 Arizona State University <1%  
Submitted works
- 
- 37 Hsien-Chu Hsia. "An effective image steganographic scheme based on... <1%  
Crossref
- 
- 38 Indian Institute of Technology, Kanpur on 2013-05-09 <1%  
Submitted works
- 
- 39 Jojo Moolayil. "Learn Keras for Deep Neural Networks", Springer Scien... <1%  
Crossref
- 
- 40 Nguyen, Duc Tuong. "Adaptive Multiple Access Schemes for Massive ... <1%  
Publication
- 
- 41 University of Leeds on 2020-09-10 <1%  
Submitted works
- 
- 42 Wan Dollah, Wan Ab. Kadir. "Determining the Effectiveness of Digital R... <1%  
Publication
- 
- 43 dergipark.org.tr <1%  
Internet