# Dental Caries Detection using NIR Images with Encoder Modified U-Net and Feature Pyramid Network

R. Raja[1], S. Regina[2], N. Ezhilmathi[3] & C. Murugan[4]

[1]Professor & [2,3,4]Associate Professor

[1&2]Department of Electronics and Communication Engineering

[3]Department of Electrical and Electronics Engineering

[4]Department of Mechanical Engineering

Pandian Saraswathi Yadav Engineering College, Sivagangai, Tamilnadu, India.

rajamdu2013@gmail.com, regina.shali@gmail.com, ezhilmemiste@gmail.com, cmn.muru@gmail.com

**Abstract**: Automated dental caries detection has garnered increased attention due to technological advancements in machine learning methods. This stands as a crucial concern in dentistry, particularly in the accurate identification of caries, as they can lead to severe health complications. This study endeavors to precisely segment and identify dental diseases. In this work, a pioneering technique employing the Encoder Modified U-Net in conjunction with the Feature Pyramid Network is proposed for segmentation purposes. The U-Net model is widely recognized and utilized in medical image segmentation due to its encoder-decoder architecture and skip-connection capability, allowing for the capture of multi-scale information in medical images. The initial stage of segmentation produces a preliminary result, which is utilized to extract the Region of Interest (ROI). This ROI is subsequently fed into the second stage of the U-Net model. Maintaining the original resolution as much as possible for the input image in the second stage significantly enhances the segmentation performance. This proposed model has been implemented using MATLAB and rigorously compared against existing algorithms in terms of accuracy, F-score, precision, and recall rates to validate its efficacy and performance.

**Keywords: Dental Caries, Encoder, U-Net, Segmentation, MATLAB**

## 1. Introduction

The system's design [1] incorporates intensity-modulated light and relies on detecting early-stage caries by analyzing the altered thermal-wave field. This change is a result of increased light absorption at sites affected by caries. Consequently, this altered thermal field is captured by an IR camera through the emission of infrared radiation. To ensure cost-effectiveness and clinical applicability, a low-cost LWIR (8-14 µm) camera has been seamlessly integrated into the thermophotonic imaging system. The IR camera utilized [2] an uncooled microbolometer LWIR detector. This choice is deliberate, as these detectors are highly suitable for implementation in clinically viable imaging systems. A straightforward approach utilizing laser-induced fluorescence spectrum with enhanced backscattering is proposed for the early detection of initial caries in

vitro teeth. Fluorescence spectra corresponding to carious erosion at different stages were provided for analysis. During the process, the radiation emitted from backscattering, reflection, and autofluorescence of teeth was concurrently recorded. Through an in-depth analysis of the characteristics of autofluorescence and anti-Stokes fluorescence, both chemical changes and structural or morphological alterations can be assessed. To detect early-stage dental caries using digital color images, facilitating easier and more effective treatment. This classification method is also well-suited for tele-dentistry, which serves as an informatics-based oral health care system. The CNN model [3] aims to analyze digital color images of dental conditions, particularly targeting the early detection of caries. This technology not only aids in accurate diagnosis but also holds potential for remote dental consultations and treatments through tele-dentistry platforms. The automatic diagnosis of dental caries using periapical images. To achieve this, a multi-input deep convolutional neural network [4] ensemble (MI-DCNNE) model was used. The process of score fusion was executed within the Softmax layer of the multi-input CNN architecture was studied. K-means clustering and a threshold method were applied [5] for the segmentation process, utilizing Computed Tomography (CT) images to build a three-dimensional representation of the carious lesion. This three-dimensional view plays a crucial role in diagnosing dental cavities. The results indicated that the average Peak Signal-to-Noise Ratio (PSNR) achieved was 19.98 and 20 using the K-means clustering and threshold method, respectively. The algorithm is designed [6] to perform self-attention at the super-pixel level and integrate the local-to-global features of these super-pixels. This fusion is guided by the robust principles of deep learning, allowing for a comprehensive integration of diverse and complementary information across different scales simultaneously. A hybrid method [7] combining deep learning and machine learning techniques to identify noticeable dental conditions such as evident dental caries/periapical infection, changes in periodontal bone height, and third molar impactions in panoramic dental radiographs. The study focuses on automating the detection [8] of dental cavities using transfer learning techniques. Specifically, two transfer learning algorithms, ResNet50 and MobileNet, were employed. An alternative method was implemented [9] where the determination of these causal weights was accomplished through the real-coded genetic algorithm (RCGA) and historical data. An imaging system [10] utilizing a near-infrared (NIR) light-emitting diode (LED) with a wavelength of 850 nm, coupled with an intraoral camera, was developed for swift dental assessments. This NIR system was employed to capture images of the teeth of ten consenting human subjects

# 2. Methodology

The proposed model consists of three main steps: initial preprocessing utilizing a median filter, segmentation employing a cascaded U-Net approach, and subsequent performance analysis as illustrated in Figure 1.
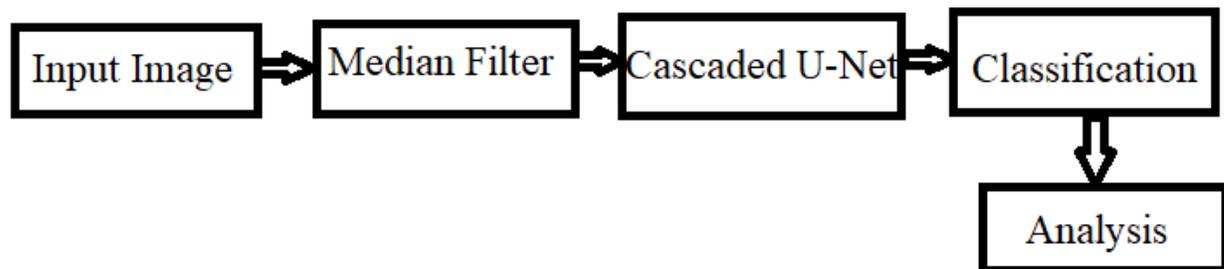


**Figure 1. Block diagram of Proposed system**

## Mean filtering:

The median filter operates by examining each pixel within an image and assessing its similarity to its neighboring pixels. Rather than substituting the pixel value with the average of neighboring pixel values, it replaces it with the median value of those pixels. To calculate the median, the filter sorts all pixel values from the adjacent neighborhood into numerical order. Subsequently, the pixel under consideration is replaced with the middle pixel value. In cases where the neighborhood contains an even number of pixels, the median is determined by averaging the two middle pixel values. This technique helps in minimizing noise or irregularities in the image by considering the central value within the surrounding pixel group.

## Segmentation:

In proposed approach, a two-stage cascaded U-Net architecture is utilized for segmentation tasks (as depicted in Figure 1). Initially, images sized at $4\times128\times128\times128$ are processed through the first stage U-Net, generating an initial segmentation map. This coarse segmentation map is combined with the original images and fed into the second stage U-Net. The aim of the second stage is to refine and improve the segmentation map accuracy, utilizing a larger set of network parameters. The cascaded network is trained end-to-end. For segmentation purposes, the data undergoes resizing to a consistent size and is fed into the first U-Net. The resulting segmentation map from the initial stage is used to isolate the Region of Interest (ROI), which is then fed into the second U-Net. The input image for the second stage aims to maintain the original resolution as much as possible, enhancing segmentation performance. In classification tasks,

segmentation is used as an auxiliary task to enhance classification performance. The U-Net architecture is popular in medical image segmentation due to its encoder-decoder structure and skip connections, enabling it to capture multi-scale information in medical images. U-Net I is designed to locate nodules and estimate their size and shape. U-Net II refines the predictions made by U-Net I to achieve more accurate nodule boundaries. During the training of U-Net II, the segmentation outcomes of U-Net I are utilized as input. Pseudo segmentation results of U-Net I are generated from labels, aiming to provide an approximate location and size information of thyroid nodules. However, in training U-Net II, pseudo segmentation results should not closely resemble the labels to prevent U-Net II from mimicking U-Net I's outputs instead of learning from the original images. To prevent this, techniques such as ellipse fitting, geometric transformation (including small-scale adjustments, translations, and random 180° rotations), and cutout are applied to the U-Net I results, erasing detailed information and encouraging U-Net II to learn from the original image data.

**EM-U Net-FPN:**

The U-Net architecture is a fully convolutional network comprised of a contracting path (Down sampling) and an expansive path (Up sampling), forming a U-shaped design. In the contracting path, successive convolutional layers, max-pooling operations, and rectified linear unit (ReLU) activations are applied to extract features and reduce spatial information. The purpose of these operations is to increase the feature information for later use in the expansive path. The expansive path involves concatenating features from the contracting path, employing deconvolution methods to up sample the features, and combining spatial information for better segmentation results. The model eventually produces a segmentation map, which is compared with the ground truth. The modification in the U-Net architecture replaces pooling operators in the fully convolutional network (FCN) with up sampling operators, utilizing a higher number of feature channels. This alteration maintains feature information in high-resolution layers, establishing a symmetrical structure resembling the letter 'U'. The Encoder Modified U-Net and Feature Pyramid Network (EM-UNet-FPN) is an advanced neural network architecture that merges two popular models: U-Net and Feature Pyramid Network (FPN). This fusion aims to achieve superior performance in semantic segmentation tasks. Semantic segmentation involves assigning a class label to each pixel in an image. The U-Net architecture, comprising an encoder network for down sampling and feature extraction, and a decoder network for up sampling and mask generation, is widely used for this task. On the other hand, FPN utilizes a pyramid structure of feature maps with various scales to enhance segmentation accuracy. The combination of U-Net and FPN in the EM-U Net-FPN architecture leverages their respective strengths to improve the precision of semantic segmentation tasks.

## 3. Software Description

The DSP System Toolbox offers a comprehensive set of algorithms, applications, and visualization tools designed for creating, simulating, and analyzing signal processing systems within MATLAB and Simulink environments. This toolbox enables the modeling of real-time DSP (Digital Signal Processing) systems for various applications such as communications, radar, audio processing, medical devices, IoT (Internet of Things), and more. Within the DSP System Toolbox, users can design and evaluate various types of filters including Finite Impulse Response (FIR), Infinite Impulse Response (IIR), multi-rate, multistage, and adaptive filters. It allows for streaming signals from diverse sources like variables, data files, and network devices, facilitating system development and verification.

Dynamic visualization and measurement of streaming signals can be performed using tools like the Time Scope, Spectrum Analyzer, and Logic Analyzer, providing insights into signal behavior in real-time. For both desktop prototyping and deployment on embedded processors, especially ARM Cortex architectures, the toolbox supports C/C++ code generation. Additionally, it offers support for bit-accurate fixed-point modeling and HDL (Hardware Description Language) code generation derived from algorithms like filters, Fast Fourier Transform (FFT), Inverse FFT (IFFT), and others. Algorithms are available in various forms such as MATLAB functions, System objects, and Simulink blocks, ensuring versatility and ease of use across different stages of signal processing system design and implementation.

## 4. Results and Discussion

### Confusion Matrix:

The Confusion Matrix, as its name implies, provides a matrix-based output that comprehensively illustrates the performance of a model in Table 1.

**Table 1. Performance of a model**

| n=165 | Predicted: No | Predicted: Yes |
|---|---|---|
| **Actual: No** | 50 | 10 |
| **Actual: Yes** | 5 | 100 |

Using the confusion matrix, the four terms are important to note.

- **True Positives :** The cases in which we predicted YES and the actual output was also YES.
- **True Negatives :** The cases in which we predicted NO and the actual output was NO.
- **False Positives :** The cases in which we predicted YES and the actual output was NO.
- **False Negatives :** The cases in which we predicted NO and the actual output was YES.

Confusion Matrix forms the basis for the other types of metrics. The proposed work is specificity recall and precision rate measurements. These measurements can be defined as:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

$$Specificity = (TN) / (TN + FP)$$

$$Recall = (TP) / (TP + FN')$$

$$Precision = TP / (TP + FP)$$

Accuracy for the matrix can be calculated by taking average of the values lying across confusion matrix forms the basis for the other types of metrics. The proposed model is compared to other models in terms of specificity, recall and precision rate measurements. These measurements can be

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

$$Precision = TP / (TP + FP)$$

where FP is the false positive, FN is the false negative, TP is the true positive and TN is the true negative of the samples. The Figure 2, 3, 4, 5, 6, 7 & 8 displays the input image, median filtered image, layer calculation of U-net, segmentation output, classification output, validation output and performance analysis, respectively.
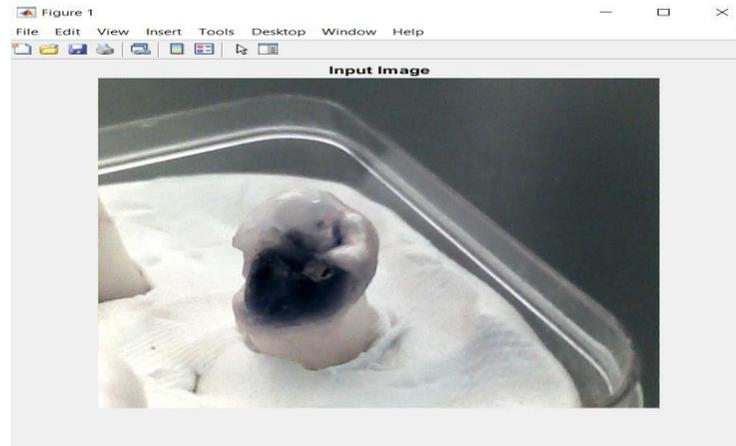
**Figure 2. Input Image**



**Figure 3. Median Filtered Image**
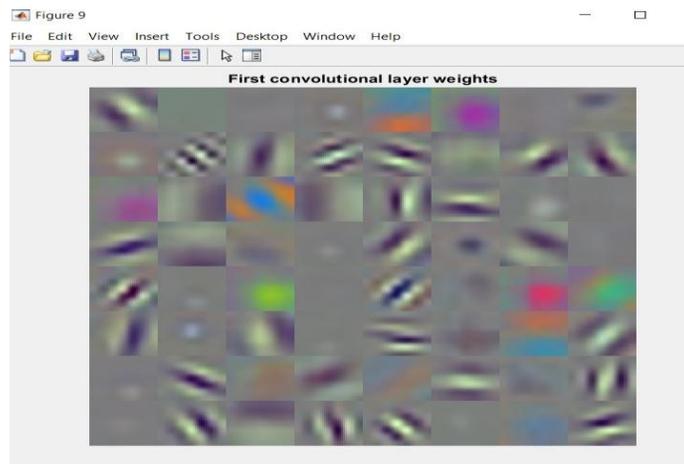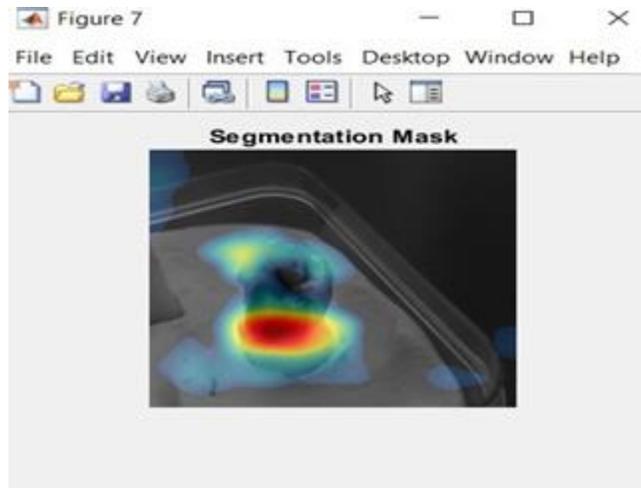


**Figure 4. Layer Calculation of U-Net**
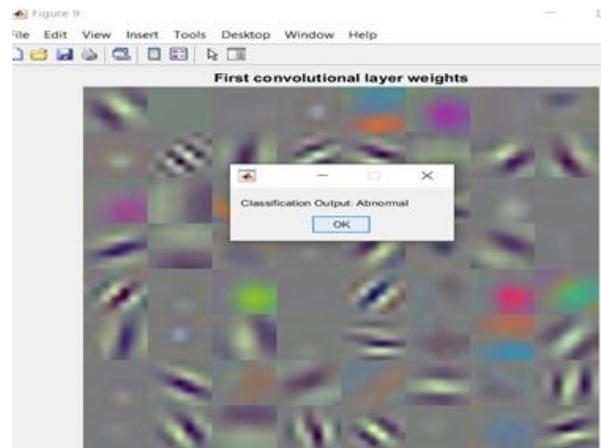
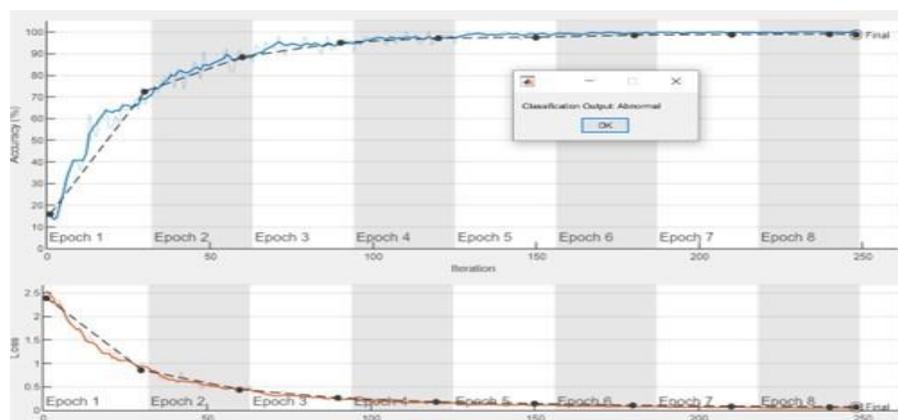**Figure 5. Segmentation Output**



**Figure 6. Classification Output**



**Figure 7. Validation Output**

**Figure 8. Performance Analysis**

## 5.  Conclusion

The proposed research introduces an innovative approach for identifying cavities in dental images. This method involves conventional image processing techniques, specifically segmentation following image enhancement, and outlining contours to achieve effective segmentation of teeth. Additionally, this research utilized EM-U Net-FPN to extract key features from dental images. These extracted data can be employed to derive measurements related to teeth, aiding in dental diagnostic systems. The primary focus of this methodology is to enable accurate classification or diagnosis of dental caries from images. The obtained results indicate a satisfactory level of accuracy in detecting cavities using the proposed technique.

## References

[1] Shih-Lun Chen et al, "Missing Teeth and Restoration Detection Using Dental Panoramic Radiography Based on Transfer Learning With CNNs", IEEE Access, vol.10, pp. 118654 – 118664, 2022.

[2] Shubhangi Babasaheb Goykar at al, "Dental Problem Classification Using Deep Learning and Image Processing", International Journal of Research Publication and Reviews, vol.4(5), pp. 6554-6559, 2023.

[3] Ana Marly Araujo Maia et al, "Evaluation of two imaging techniques: Near-infrared transillumination and dental radiographs for the detection of early approximal enamel caries", Dentomaxillofacial Radiology, vol. 40(7), pp. 429-33, 2011.

[4] Anuj Kumar et al, "Fuzzy Clustering with Level Set Segmentation for Detection of Dental Restoration area", International Conference on Advances in Computing and Communication Engineering, pp.322-326, 2018.

[5] Keith Angelino et al, "Near-infrared imaging for detecting caries and structural deformities in teeth", IEEE Journal of Translational Engineering in Health and Medicine, vol.5, pp.1-7, 2017.

[6] Arman Haghanifar et al, "Dental Caries Degree Detection Based on Fuzzy Cognitive Maps and Genetic Algorithm", Iranian Conference on Electrical Engineering, pp.976-981, 2018.

[7] Shuai Li et al, "Automatic Dental Plaque Segmentation Based on Local-to-Global Features Fused Self-Attention Network", IEEE Journal of Biomedical and Health Informatics, vol.26(5), pp. 2240 – 2251, 2022.

[8] Andac Imak et al, "Dental Caries Detection Using Score-Based Multi-Input Deep Convolutional Neural Network", IEEE Access, vol.10, pp. 18320 – 18329, 2022.

[9] Devesh Saini et al, "Dental Caries early detection using Convolutional Neural Network for Tele dentistry", International Conference on Advanced Computing and Communication Systems, pp. 958-963, 2021.

[10] Ashkan Ojaghi and Nima Tabatabaei, "Detection of Early Occlusal and Proximal Dental Caries using Longwavelength Infrared Thermophotonic Lock-In Imaging", Photonics North conference, pp. 1-3, 2016.