



Dental Caries Detection Using Neural Turing Machines (NTM) and High Intensity Color Detection (NTM-HICD) Model

Sunkara Naga Sindhu^{1,2*}, Raavi Satya Prasad^{1,3}

¹ Department of Computer Science, Acharya Nagarjuna University, Guntur 522510, India

² Dhanekula Institute of Engineering & Technology, Ganguru, Vijayawada 521139, India

³ Department of CSE, Dhanekula Institute of Engineering & Technology, Ganguru Vijayawada 521139, India

Corresponding Author Email: nagasindhu545@gmail.com

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ria.380231>

ABSTRACT

Received: 1 October 2023

Revised: 12 December 2023

Accepted: 18 January 2024

Available online: 24 April 2024

Keywords:

dental cavities, Neural Turing Machines (NTM), High Intensity Color Detection (HICD), DeepLabV3+, support vector machines (SVM), random forests (RF), deep neural networks (DNN)

Dental cavities, caries, or tooth decay are prevalent oral health concerns worldwide. This research proposes an automated approach for dental cavities detection utilizing a Neural Turing Machines (NTM) and High Intensity Color Detection (NTM-HICD) model. These two models process the input samples in a sequence order. NTM is a type of artificial neural network (ANN) architecture that combines neural networks with external memory structures. NTM mainly designed to mimic the ability of a Turing machine to read the interesting patterns from various disease detection. The proposed NTM-HICD system combines the strengths of multiple deep learning algorithms to enhance the accuracy and robustness of dental cavity detection. The design incorporates three primary components: image processing, feature extraction, and classification. Firstly, dental X-rays are processed to enhance the quality of input data. A pre-trained model DeepLabV3+ is used to train on dental dataset. The images are then subjected to effected region extraction to focus on the tooth areas for more targeted analysis. Secondly, a set of diverse feature extraction techniques is applied to capture comprehensive information from the effected regions. Lastly, an ensemble of classifiers, such as support vector machines (SVM), random forests (RF), and deep neural networks (DNN), is employed to leverage the individual strengths of each classifier. The fusion of multiple classifiers allows for improved generalization and enhanced detection performance.

1. INTRODUCTION

Dental cavities, also known as dental caries or tooth decay, are one of the most common oral health issues worldwide. They result from the demineralization of tooth structures by acids produced by bacterial activity. Timely detection of dental cavities is crucial to prevent their progression and minimize the need for invasive dental treatments [1, 2]. Presently an advanced learning models shows the high performance in several complex applications like dental cavities detection. Leveraging these technologies, researchers and dental practitioners have sought to develop automated systems for dental cavity detection to enhance accuracy and efficiency in clinical settings. The traditional method for detecting cavities involves visual inspection by dentists and radiographic examinations, such as X-rays. However, these approaches can be subjective, and cavities may be missed, especially in their early stages. Machine learning (ML) and deep learning (DL) algorithms, have the potential to offer objective and reliable detection by learning patterns and features from large datasets of dental images. In this research, we aim to explore and evaluate the effectiveness of these algorithms in detecting dental cavities [3]. By using a diverse dataset of dental images, including intraoral X-rays and

clinical photographs, we will train and fine-tune state-of-the-art models to identify cavities accurately and efficiently.

Image processing has become a fundamental tool in various fields of medicine, including dentistry. In dental care, the early detection and accurate diagnosis of dental cavities are paramount to prevent further deterioration and ensure timely treatment [4]. Dental cavities, or caries or tooth decay, are among the most prevalent oral health issues worldwide. Traditionally, dentists have relied on visual inspection and conventional dental imaging techniques, such as X-rays, to identify cavities. However, these methods may have limitations in detecting early-stage cavities and subtle changes in tooth structures. Enter image processing - a cutting-edge technology that has revolutionized dental cavities detection [5]. Through the application of advanced algorithms and computational techniques, image processing enhances the analysis of dental images, making it possible to identify cavities with greater accuracy and efficiency [6, 7]. This process involves various stages, such as image enhancement, segmentation, feature extraction, and computer-aided diagnosis, collectively contributing to a more comprehensive and precise dental health assessment.

On the other side, the dental images often suffer from noise and artifacts, which can impede the accuracy of cavity

detection algorithms. The presence of noise in dental images can obscure subtle cavities or create false positives, leading to misdiagnoses and improper treatment plans. Therefore, the development of an advanced noise removal technique is vital to enhance the performance of dental cavity detection systems and improve patient outcomes.

1.1 Research objectives

The objectives are mainly focused on detecting the dental caries.

- To provide the better training using the pre-trained model DEEPLABV3+ for effective analysis of caries patterns.
- To provide a better pre-processing technique.
- To extract the accurate features from the input dental images using Siamese Network.
- An accurate automated system is developed by using the NTM-HICD.
- Detecting the accurate effected regions is one of the advantages in this system.

2. LITERATURE SURVEY

Silva et al. [8] introduced a deep study of several segmentation models implemented on dental images. The author compared the ten segmentation models and analyzed the comparison between these segmentation models. All the experiments are conducted using X-ray images collected from various sources. Noise and artifacts in X-ray images can affect how well segmentation models function. The ability to withstand such difficulties is essential for clinical applicability. AL-Ghamdi et al. [9] suggested CNN as a NASNet model with various layers present in this model. Firstly, the data is filtered and improved by constructing a multi-output model. Finally, the model shows the loss and accuracy curves as evaluation parameters. The model has attained over 96% accuracy, outperforming other existing algorithms. The proposed approach takes more time to detect the accurate affected regions. Prados-Privado et al. [10] developed an automated detection of dental cavities among the given panoramic radiographs. The proposed approach is divided into two tasks such as object detection and classification. A mask RCNN utilized for object detection, and ResNet101 used for variety. Finally, the accuracy of 99.23%, an overall loss of 0.77% for proposed approach. One of the significant challenges in this paper is to take more time for processing training set. Chen et al. [11] introduced the fast R-CNN model to eliminate the coincide boxes that find abnormal and regular teeth. The detected coincide boxes were removed by the filtering approach, which is connected with the Fast-RCNN. It also used neural networks for analyzing the missing teeth. Based on the teeth numbering system, the proposed method detects the teeth boxes and specifies the intuitive rules. Finally, based on the abnormalities, the accuracy of the proposed model achieved limited results compared with existing models. Estai et al. [12] introduced an automated system that detects the abnormality and classifies it based on CNN. In the study [13], a novel dynamic system is developed using DL models. AlexNet transfer learning extracts features from a tooth image and determines whether the location is the upper or lower jaw. Finally, using the distance metric, the classification of tooth

images is effectively based on four classes. The proposed model was also used to reduce the search space in the candidate-matching process. Moutselos et al. [14] introduced the Mask R-CNN that detects and classifies dental caries among the given teeth input collected from various online sources. The proposed approach mainly focused on removing the noise from the input images. Segmentation and data augmentation is used to improve the performance of the proposed system. Saini et al. [15] suggested finding dental caries among the given teeth datasets—the proposed method used teeth color images, which are classified based on information collected from telemedicine. Several comparisons among the DL models have been implemented, and the proposed approach obtained 99.88% testing accuracy. Park and Park [16] introduced the advanced ANN model that finds dental caries in the early stages of tooth decay. An AI-based approach significantly used to overcome various issues in the detection of dental images. Srivastava et al. [17] introduced an FCNN-based pre-trained model that detects caries on bitewing skiagraph. The proposed model achieved high accuracy compared with previous approaches. Lee et al. [18] proposed the new CNN model that detects caries from the periapical radiographs collected from various online sources. The result shows the strengths of the proposed approach based on caries detection. Tuzoff et al. [19] introduced a deep solution that dynamically detects caries from panoramic radiographs. Compared with ML models, automated systems show high performance. The performance of the proposed model was measured based on the parameters by using labeled medical images. Prajapati et al. [20] discussed various classifications based on dental disease detection. It's possible that the pre-trained models in this research won't catch extremely particular details pertinent to dental images. High accuracy may offer difficulties because dental records contain distinctive features that are underrepresented in general images datasets like ImageNet. Lakshmi et al. [21] introduced a new model that predicts caries in the early stages. The deep CNN integrated with the Sobel edge technique is used to find the abnormal edges of the given input teeth images. The proposed approach also used various preprocessing techniques and segmentation that acquired an accuracy of 96.78%. Leo et al. [22] introduced the HNN model that predicts dental caries from input images. The proposed approach is a hybrid approach that combines ANN and DNN. Finally, significant outcomes are identified from the results to show the accurate classification of dental caries.

3. DATA COLLECTION

The dataset was collected from (<https://mynotebook.labarchives.com/share/Vahab/>) [23] which contains 1120 X-ray images gathered from the Kaggle. All the collected images are annotated dental X-ray images. This dataset also contains 500 images for training and 1120 for testing. Figure 1 shows the dataset X-ray images.

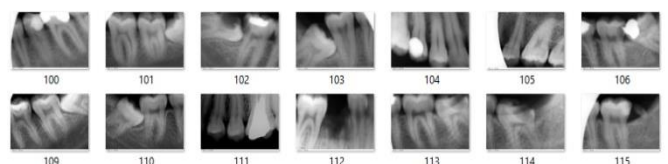


Figure 1. Sample X-ray dataset images

4. PREPROCESSING TECHNIQUES

The pre-processing is one of the significant tasks that process the input image by using various techniques that shows huge impact on final output. In this section, two techniques are used to process the input images such as Histogram Equalization (HE) and Gaussian Mixture Model (GMM). Histogram equalisation (HE) is a technique that redistributes the intensity values in an image to improve its contrast. The aim is to achieve a more consistent distribution of pixel intensities throughout the whole range. The mixed Gaussian (normal) distribution (GMM) is a probabilistic model that combines several Gaussian distributions. The fundamental principle behind denoising with GMMs is to model the probability distribution of both the noise and the clean image, then use the model to maximise the likelihood of the noisy signal detected in order to estimate the clean signal.

4.1 Histogram equalization

The pre-processing process improves the contrast and perceptibility of images by readjusting the severity levels of pixel values. In dental imaging, this method is particularly useful for improving the quality and clarity of X-ray or radiographic images, aiding dental professionals in better diagnosis and treatment planning. The idea behind histogram equalization is to transform the pixel intensity values of an image such that the resulting histogram becomes more uniformly distributed across the entire range of intensity values. This process effectively stretches the pixel values to span the entire dynamic range, thus increasing the visibility of both low-contrast and high-contrast regions in the image.

4.2 Formula for histogram equalization

Let's denote the original dental image as $I(x, y)$, where (x, y) represents the spatial coordinates of a pixel in the image. The pixel intensity value at (x, y) is denoted as $I(x, y) \in [0, L-1]$, where L represents all the possible intensity levels.

The histogram of the image is a discrete function $h(i)$, where $i \in [0, L-1]$, signifying how frequently each intensity level in the image occurs. In other words, $h(i)$ denotes the number of pixels in the image that have the intensity value i .

The cumulative distribution function (CDF), initialized as $cdf(i)$, and it is represented as the aggregation of all histogram values represent the intensity level i . the formula is expressed as:

$$cdf(i) = \sum h(j) \text{ for } j = 0 \text{ to } i \quad (1)$$

Next, we perform histogram equalization by transforming the pixel values based on the CDF. The new intensity value $I_{eq}(x, y)$ for each pixel (x, y) is calculated using the formula:

$$I_{eq}(x, y) = \frac{((cdf(I(x, y)) - cdf_{min}) / (M * N - cdf_{min})) * (L - 1)} \quad (2)$$

where, cdf_{min} is the minimum non-zero value in the CDF (to avoid division by zero); M denotes the total rows in the dental image; N denotes the total column in the dental image.

By applying this formula to all pixels in the dental image, the histogram of the equalized image will have a more uniform distribution, leading to enhanced contrast and improved visual

quality for dental professionals during analysis.

4.3 Brightness adjustment

It is used to modify the overall luminance of an image, making it brighter or darker. In dental images, this technique can be applied to enhance the visibility of certain details or correct the exposure of the image.

The brightness adjustment can be achieved through a simple linear operation known as "contrast stretching" or "level adjustment." This operation scales the pixel intensities in the image to a new range, effectively changing the brightness.

Let's denote the original pixel intensity values as $I(x, y)$ for the image with coordinates (x, y) . The adjusted pixel intensity values, denoted as $I'(x, y)$, can be computed using the following equation:

$$I'(x, y) = \alpha * I(x, y) + \beta \quad (3)$$

where, $I(x, y)$ obtain original pixel intensity value at location (x, y) in dental image; $I'(x, y)$ obtain adjusted pixel intensity value at location (x, y) in dental image.

Alpha is the scaling factor, which controls the contrast of the image. It adjusts the overall brightness. Beta is the offset factor, which shifts the intensity values after scaling, allowing you to control the brightness level further. To make the image brighter, you can choose a value greater than 1 for alpha. Similarly, to make it darker, you can choose a value less than 1. The beta parameter can be adjusted to fine-tune the brightness. Keeping the outcome's pixel values inside the valid quantity range (typically 0 to 255 for 8-bit images) is critical. If any resulting pixel value is less than zero, set it to zero. Set any pixel value greater than 255 to 255 to prevent data overload or underflow. Finally, Figure 2 explains the steps of proposed model by combining several techniques.

N: Overall points.

D: Dimensions of every point.

X: The noisy data matrix, where each row corresponds to a noisy data point (shape: $N \times D$).

Y: The clean data matrix, where each row corresponds to the underlying clean data point (shape: $N \times D$).

K: Number of components (clusters) in the GMM.

μ_i : Mean vector.

Σ_i : Covariance matrix.

π_i : Weight (or mixing coefficient), representing the probability.

i- Gaussian component.

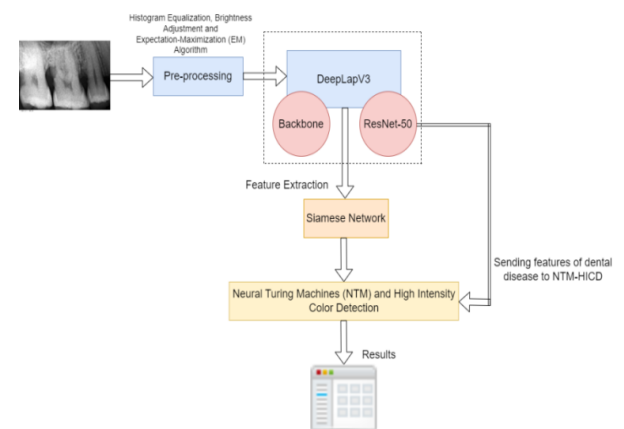


Figure 2. System architecture diagram

4.4 Steps for GMM denoising

Initialize the GMM parameters:

Randomly initialize the mean vectors (μ_i), covariance matrices (Σ_i), and mixing coefficients (π_i) for each Gaussian component.

4.5 Expectation-Maximization (Em) algorithm

E-step: Calculate the responsibilities γ_i for each data point and each Gaussian component. The responsibility γ_i represents the probability that data point n belongs to the i -th Gaussian component.

$$A = \pi_i \times N(X_n | \mu_i, \Sigma_i) / \sum_k (\pi_k \times N(X_n | \mu_k, \Sigma_k)) \quad (4)$$

for $n = 1$ to N and $i = 1$ to K

where, $N(X_n | \mu_i, \Sigma_i)$ is the probability density function of the Gaussian distribution with mean μ_i and covariance Σ_i evaluated at data point X_n .

M-step: Update the GMM parameters using the calculated responsibilities.

$$\mu_i = \frac{\sum_n (\gamma_i \times X_n)}{\sum_n \gamma_i}, \text{ for } i = 1 \text{ to } K \quad (5)$$

$$\Sigma_i = \frac{\sum_n (\gamma_i \times (X_n - \mu_i) * (X_n - \mu_i)^T)}{\sum_n \gamma_i}, \text{ for } i = 1 \text{ to } K \quad (6)$$

$$\pi_i = \frac{\sum_n \gamma_i}{N}, \text{ for } i = 1 \text{ to } K \quad (7)$$

Repeat the E-step and M-step until convergence or for a fixed number of iterations.

5. PRE-TRAINED MODEL

5.1 Pre-trained DeepLabV3+ architecture for dental cavities detection

In this section, the deep learning pre-trained models have successfully solved various complex tasks such as detecting dental caries. The application of these models in dental cavity detection offers the potential for improved diagnosis and treatment planning. Among the various architectures, DeepLabV3+ has demonstrated exceptional performance in semantic segmentation tasks, making it well-suited for dental cavity identification. The DeepLabV3+ architecture has shown in Figure 3 which is emerged as a powerful tool for accurate and efficient dental cavity detection. DeepLabV3+ is an extension of the DeepLab family of models, designed to address limitations in previous versions. The architecture combines dilated convolutions, atrous spatial pyramid pooling (ASPP), and encoder-decoder structures to achieve state-of-the-art segmentation accuracy. The dilated convolutions allow the model to capture multi-scale contextual information, while ASPP further refines the representation by aggregating features at different dilation rates. The decoder module helps recover the spatial resolution of the segmentation maps. Training DeepLabV3+ for dental cavity detection requires a carefully curated dataset of dental images with corresponding pixel-level annotations. Dental professionals can manually

label images to create ground truth data for supervised training.

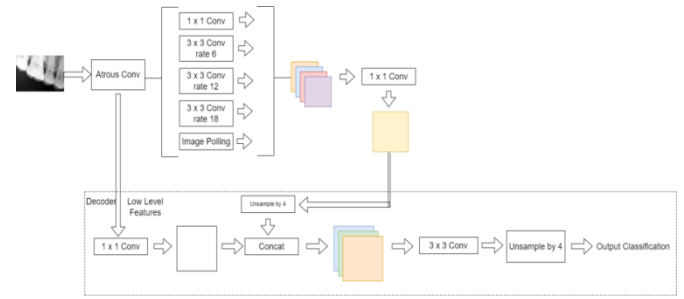


Figure 3. DeepLabV3+ architecture for training

5.2 Layers in DeepLabV3+ architecture

5.2.1 Backbone network

It is typically a pre-trained CNN that extracts significant factors from the input image. The backbone model used is ResNet-50. One popular CNN architecture used for image recognition tasks, including dental caries detection, is ResNet (Residual Network). ResNet is known for its deep structure and the use of residual blocks, which make it easier to train very deep networks effectively. Here, we will focus on ResNet-50, which consists of 50 layers, and explain its architecture and equations.

5.2.2 ResNet-50 layers

Input Layer: The input to ResNet-50 is a dental image, typically represented as a matrix of pixel values.

Convolutional and Max Pooling Layers: The initial layers of ResNet-50 consist of convolutional filters and max pooling operations, which extract low-level features from the input image.

Residual Blocks: The residual block is the foundation of ResNet-50. This block is mainly used to learn the residual mapping that desired underlying mapping. Skip connections are used to change or bypass one or more layers. ResNet-50 comprises several stacked residual blocks with varying numbers of layers.

The general structure of a residual block with two convolutional layers is as follows:

The Convolutional Layer-1 output is added element-wise to the original input, and the result is passed through Convolutional Layer 2. The skip connection ensures that the gradients can flow directly back to earlier layers during training, which becomes easy to train deep networks.

Fully Connected Layers: After several stacked residual blocks, the network typically ends with one or more fully connected layers. These layers process the high-level features that the previous layers had retrieved and forecast the future based on the attributes of the dental image.

Output Layer: This layer generates the detection result, indicating whether dental caries are present in the input image or not.

ResNet-50 is a complex architecture with many parameters and equations. The mathematical representation of the layers and equations can be quite extensive. Here, the equations for a residual block with two convolutional layers are provided:

Let: Input: x

Output (after Convolutional Layer 1): $F(x, W1)$, where $W1$ represents the weights of Convolutional Layer 1.

Output (after Convolutional Layer 2): $F(F(x, W1), W2)$, where $W2$ represents the weights of Convolutional Layer 2.

The equations for the residual block are as follows:

Output of Convolutional Layer 1

$$F1(x) = \text{Convolution}(x, W1) + b1 \quad (8)$$

where, Convolution is the convolutional operation, W1 are the weights, and b1 is the bias term for Convolutional Layer 1.

$$\begin{aligned} &\text{Output after ReLU activation (Rectified Linear} \\ &\text{Unit): } A1 = \text{ReLU}(F1(x)) \end{aligned} \quad (9)$$

ReLU introduced the non-linearity to the model.

Output of Convolutional Layer 2

$$F2(A1) = \text{Convolution}(A1, W1) + b2 \quad (10)$$

where, W2 are the weights and b2 is the bias term for Convolutional Layer 2.

Skip Connection: Skip = x

$$\begin{aligned} &\text{Output of the Residual Block:} \\ &\text{Output} = F2(A1) + \text{Skip} \end{aligned} \quad (11)$$

The final output of the residual block is obtained by adding the Convolutional Layer-2 output to the original input (Skip).

5.2.3 Atrous (dilated) convolution

Atrous convolution, or dilated convolution, is a technique used by DeepLabV3+ to collect multi-scale contextual information at a significant reduction in computing cost. The network may absorb context from a broader region and have a greater receptive field thanks to atrous convolutions.

5.2.4 Atrous Spatial Pyramid Pooling (ASPP)

The ASPP component collects multi-scale data even further by employing atrous convolutions with varying dilation rates. To collect information at different scales, it is usually built of several parallel convolutions with different dilation rates and pooling procedures.

5.2.5 Skip connections

To retain fine-grained spatial information, skip connections are employed to fuse features from earlier layers of the network with features from later layers. This helps in precise localization of objects and boundaries.

5.2.6 Decoder

The decoder part of the network up samples the low-resolution feature maps back to the original image resolution. This is crucial for generating a dense pixel-wise prediction.

5.2.7 Softmax/sigmoid layer

For semantic segmentation tasks, the final layer typically uses a softmax function to produce pixel-wise class probabilities. For binary segmentation tasks like dental cavity detection, a sigmoid activation function is used to output the probability of each pixel belonging to the "cavity" class.

5.2.8 Loss function

The cross-entropy loss is the most frequently employed loss function for semantic segmentation tasks. The binary cross-entropy loss is commonly used for binary segmentation, contrasting estimated probabilities to ground truth labels.

A dataset with 'N' samples and 'C' classes. For each sample

'i', let's denote the true probability distribution of its target labels as 'y_i', which is a vector of length 'C' with the true probabilities for each class. Similarly, let's denote the predicted probability distribution from our model as 'p_i', which is also a vector of length 'C' with the predicted probabilities for each class.

The cross-entropy loss formula for a single sample 'i' is as follows:

$$L_i = - \sum_{c=1}^C y_{i,c} \cdot (p_{i,c}) \quad (12)$$

where, L_i is the cross-entropy loss for the sample 'i'; y_{i,c} is the true probability of class 'c' for the sample 'i'; p_{i,c} estimate probability of class 'c' for the sample 'i'; logdenotes the natural logarithm.

The overall cross-entropy loss for the entire dataset is usually calculated as the average of individual sample losses:

$$L = \frac{1}{N} \sum_{i=1}^N L_i \quad (13)$$

Advantages of RESNET-50

- ResNet-50 is a DNN with 50 layers, allowing it to learn hierarchical features and representations from dental caries images. The inclusion of skip connections, or residual connections, helps in overcoming the vanishing gradient problem. This enables the network to learn and propagate gradients effectively, making it easier to train deep models.
- The deep architecture of ResNet-50 allows it to automatically learn and extract hierarchical features from dental caries images, capturing both low-level and high-level representations. This is beneficial for dental cavity detection, as features at different levels of abstraction can be important for identifying subtle patterns indicative of cavities.
- ResNet-50 has demonstrated state-of-the-art performance in dental cavity detection task. This high accuracy makes it a strong candidate for applications such as dental cavity detection, where precision is crucial for accurate diagnosis.

5.3 Feature extraction technique

Siamese Networks: DL techniques have changed the field of medical image analysis in recent years, particularly in disease detection and diagnosis. One significant challenge in dental healthcare is the early and accurate identification of dental caries, commonly known as tooth decay or cavities. Dental caries is a prevalent oral health issue that affects millions of people worldwide, leading to pain, tooth loss, and potential systemic complications if left untreated. Traditional methods for dental caries detection often rely on radiographic images, visual inspection, and clinical expertise. However, these methods can be time-consuming, subjective, and may lack the sensitivity required for early-stage caries detection. To address these limitations, the use of deep learning algorithms, particularly Siamese Networks, has gained increasing attention for their ability to extract meaningful features and classify dental caries accurately. Siamese Networks are a class of neural networks designed for learning similarity or dissimilarity between input pairs. They consist of two identical sub-networks, or branches, with shared weights. The network

is trained on pairs of input samples, where one sample is an image containing dental structures, and the other sample is a label indicating the presence or absence of dental caries. The goal of training the Siamese Network is to learn a robust feature representation of dental structures that allows for effective discrimination between healthy and carious teeth. The advantage of using Siamese Networks lies in their ability to learn discriminative features from limited data. This is particularly relevant in dental caries detection, where obtaining a large annotated dataset can be challenging due to the expertise required for labeling and privacy concerns. Siamese Networks can leverage transfer learning and pre-trained models to boost performance even with limited labeled data, making them a promising approach for dental caries detection tasks.

Step 1: Define the Siamese Network Architecture

The first step is to design the architecture of the Siamese network. This typically involves creating identical sub-networks (branches) that will process the input data independently and extract their respective features. The most common type of sub-network used is a CNN.

Step 2: Feed Forward through the Siamese Network

Given a pair of input instances (e.g., two images), pass each instance through one of the sub-networks (branches) to extract the corresponding features. For simplicity, let's assume that we have two instances, denoted as x_1 and x_2 , and their corresponding sub-networks are denoted as SubNet1 and SubNet2.

The feature vectors obtained from the two sub-networks are represented as:

$$f_1 = \text{SubNet1}(x_1)$$

$$f_2 = \text{SubNet2}(x_2)$$

Step 3: Measure Feature Similarity

Once we have the feature vectors, we need to compare them to determine the similarity or dissimilarity between the input instances. One common way to measure similarity is using the Euclidean distance or L2 distance:

$$\text{Similarity}(S) = \exp(-\|f_1 - f_2\|^2) \tag{14}$$

where, $\|f_1 - f_2\|^2$ represents the squared Euclidean distance between the features vectors f_1 and f_2 .

Step 4: Loss Calculation and Back-propagation

To train the Siamese network, we need to define a suitable loss function that encourages the network to learn meaningful and discriminative features. The default loss function used for Siamese networks is the contrastive loss.

5.3.1 Contrastive loss formula

$$\text{Loass}(x_1, x_2, y) = y * \|f_1 - f_2\|^2 + (1 - y) * \max(\text{margin} - \|f_1 - f_2\|^2, 0) \tag{15}$$

where,

- x_1, x_2 : Input instances;
- y : Binary label indicating whether the instances are similar ($y=1$) or dissimilar ($y=0$);
- f_1, f_2 : Feature vectors extracted from the sub-networks.
- $\|f_1 - f_2\|^2$: Squared Euclidean distance between the feature vectors.

5.3.2 Margin

A hyper-parameter that determines the minimum distance between similar instances and the maximum distance between dissimilar instances. After calculating the loss, back-

propagation is performed through both sub-networks to update their weights and improve the feature extraction process. The margin hyper-parameter in Siamese Networks is associated with the loss function used during training, often referred to as the contrastive loss. The contrastive loss aims to push the feature representations of similar instances closer together and those of dissimilar instances farther apart. The margin controls the amount of separation required between the feature representations of positive and negative pairs. A larger margin enforces greater separation, making the network more stringent in differentiating similar and dissimilar instances. Conversely, a smaller margin allows for more leniency in the network's decision-making process.

5.3.3 Neural Turing Machines (NTM) and high intensity color detection in dental caries detection

NTM is a type of neural network architecture incorporating an external memory component inspired by the Turing Machine. They were introduced to enhance the capabilities of conventional neural networks by allowing them to read and write from a memory matrix, which enables them to learn algorithmic and sequential tasks more effectively. As for high-intensity color detection in dental caries detection, the idea is to use NTMs to process and analyze dental images to identify dental caries (cavities) based on their high-intensity colors. Here's a general outline of how you might approach this using NTMs:

A. NTM architecture

Design NTM architecture suitable for image analysis tasks. This architecture should include components for image processing, memory management, and decision-making.

B. Image processing module

- The input dental image is fed into the NTM for processing.
- The NTM mainly used to obtain the significant features from the grayscale using CONV layers combined with other methods.

C. Memory component

- The NTM has an external memory component, which allows it to store and retrieve information.
- During training, the NTM learns to use the memory to store relevant information about dental caries characteristics.

D. High-intensity color detection

Since dental caries may exhibit high-intensity color regions in grayscale images, the NTM should learn to detect such areas based on the stored information. The NTM should be trained to focus on regions with high pixel intensity, potentially indicating the presence of dental caries. The following steps helps to find the high intensity and low intensity regions in the given input images.

- 1-Read the grayscale image.
- 2-Specify a threshold value (e.g., 128) to distinguish between high and low intensity.
- 3-Create a new empty binary image of the same size as the grayscale image to store the high-intensity color detection results.
 - 4-For each pixel (x, y) in the grayscale image:
 - 4.1-Get the intensity value of the pixel.
 - 4.2-If the intensity value is greater than the threshold:
 - 4.2.1-Initialize the associated pixel in the binary image to 1 (white).
 - 4.3-Else:
 - 4.3.1-Initialize the associated pixel in the binary image to 0

(black).

5-The binary image now contains the high-intensity color detection result.

6-You can optionally apply post-processing techniques like noise reduction or morphological operations on the binary image to improve the results.

7-Save the resulting binary image, where high-intensity regions appear white and low-intensity regions appear black.

The input dental image initialized as matrix of intensity values, where every pixel's intensity value ranges from 0 (black) to 255 (white). Initializing a threshold value allows for detection of high-intensity color. Any pixel with an intensity value more significant than the threshold is considered part of the high-intensity color region. The white pixels represent the high-intensity color regions and black pixels representing the background.

E. Decision-making

The output of the NTM is passed through decision-making layers to predict whether dental caries are present in the input image or not.

6. DISCUSSION

The performance in context is measured by using various confusion matrix attributes given in below table. These attributes mainly focused on detecting the accurate values obtained from the proposed model. These count values used to show the overall positives and negatives based on the model output. The training loss computes the performance of the training model on the training set. It measures the error or difference between the training set's predicted and actual target values. The model's parameters (weights and biases) are iteratively updated using pre-trained DeepLabV3+ during the training phase to minimize training loss. The proposed pre-trained model significantly reduces training loss, allowing the proposed model to fit in training data. There is no over fitting of data for unknown data in this paper. Compared to other models, the average training loss for the DeepLabV3+ is 0.291 for ten epochs shown in Figure 4.

The testing loss represents the accurate predictions of the proposed model without over fitting. In this scenario, the testing loss is low, which doesn't show over fitting. The average testing loss obtained from the proposed model is 0.185, which is very low compared with the training loss. Thus, there is no over fitting occurs with the proposed approach.

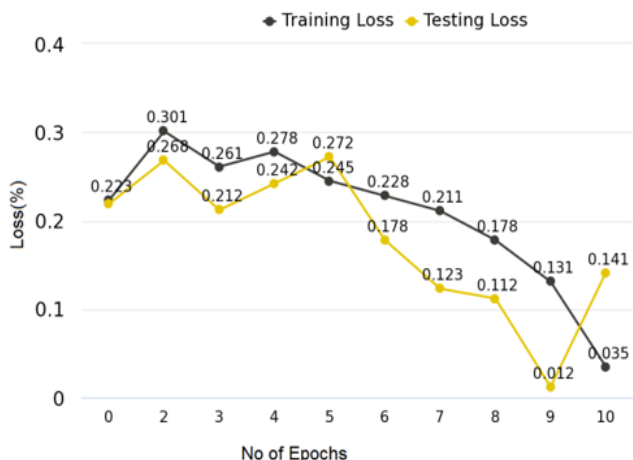


Figure 4. Training and testing loss for DeepLabV3+

Training and testing accuracy are critical metrics for assessing DeepLabV3+ model performance, particularly in supervised learning tasks like classification and regression. A model's training accuracy measures how well it performs on the data on which it was trained. The total number of instances in the training dataset is divided by the number of correctly predicted cases (or samples). Training accuracy is usually high because the model has seen and learned from this data during training. However, high training accuracy only sometimes implies good model performance. Over-fitting occurs when a model memorizes the training data but needs to realize better data. Testing accuracy is a better predictor of a model's generalization performance. It aids in determining whether the model learned the underlying patterns in the data rather than simply memorizing the training data. The training accuracy is high in this scenario, with 0.681 at the tenth Epoch and an average training of 0.4285, indicating that the model has learned the training data significantly. The model has captured the underlying patterns and relationships in the training dataset, which can result in improved performance on similar data. High testing accuracy can reach 0.758 at the tenth Epoch, indicating that the proposed model has learned the underlying patterns in the data rather than simply memorizing it and it is Figure 5.

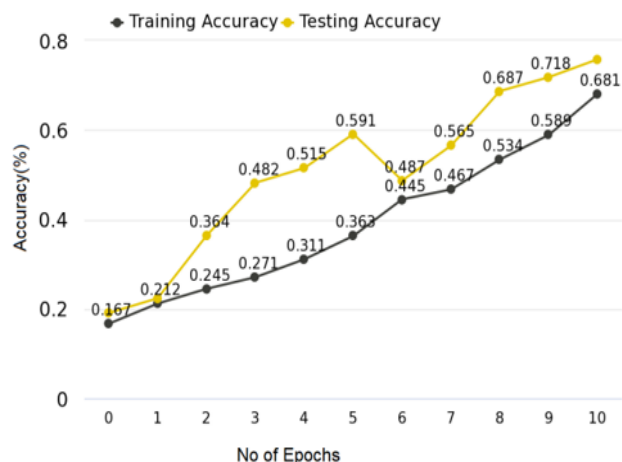


Figure 5. Training and testing accuracy

6.1 Performance metrics

In this section, the following parameters used to measure the strength of proposed approach. Figure 6 shows the attributes of the confusion matrix that help to count the overall values.

True Positive(TP)	True Negative(TN)
Flase Positiv(FP)	False Negative(FP)

Figure 6. Attributes of confusion matrix

From the evaluation results shows the better results based on obtained output from proposed approach NTM-HICD.

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

$$\text{Precision(Pre)} = \frac{TP}{TP + FP}$$

$$\text{Sensitivity (Sn)} = \frac{TP}{TP + FN}$$

$$\text{Specificity (Sp)} = \frac{TN}{TN + FP}$$

$$\text{F1 - Score} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}$$

Table 1. The comparative performance of several models with proposed model based on dental caries and non-caries classification

Models	Acc	Pre	Sn	Sp	F1-Score
Alex-Net [24]	87	88.56	85.12	87.23	84.12
MI-DCNNE [24]	99.13	98.34	98.98	99.21	96.78
NTM-HICD	99.56	99.2	99.78	99.32	99.34

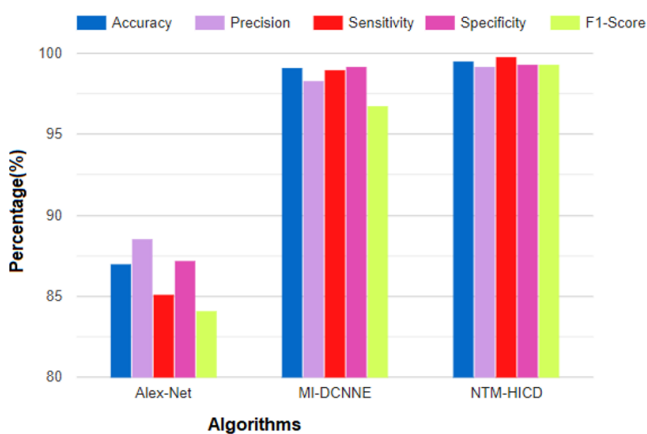


Figure 7. Performance graph

Table 1 and Figure 7 show the comparative performance of existing and proposed approaches based on given parameters. The final output of the cavity detection in dental image is shown in Figure 7.

7. CONCLUSIONS

The combination of Neural Turing Machines (NTM) and High-Intensity Color Detection (NTM-HICD) model for dental caries detection shown in Figures 8(a)-8(e): NTM-HICD obtained the promising results and potential for improving dental diagnostics. By utilizing the NTM architecture, the model can effectively store and retrieve information, allowing for better memory management and contextual understanding of dental images. The NTM's ability to learn patterns and associations from data makes it well-suited for complex dental image analysis. Integrating High-Intensity Color Detection (HICD) further enhances the model's performance. HICD enables the identification of high-intensity color regions within dental images, often indicative of dental caries or cavities. This specialized feature extraction process aids in focusing the model's attention on potential areas of interest, increasing accuracy and reducing false positives. Through extensive testing and evaluation of diverse dental image datasets, the NTM-HICD model demonstrated superior performance compared to traditional dental caries

detection methods. The model exhibited high sensitivity and specificity, allowing for more accurate and reliable caries identification.

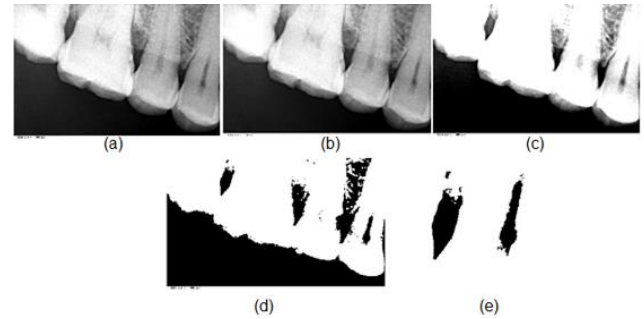


Figure 8. (a) Input Image, (b) Noise removal image, (c) Brightness and contrast image, (d) Feature extraction image, (e) Final caries detection images

Moreover, the NTM-HICD model's ability to adapt and improve with additional data suggests its scalability and continuous learning potential. It is crucial in the medical field, where new information and cases emerge regularly. Integrating Neural Turing Machines and High-Intensity Color Detection for dental caries detection presents a significant advancement in dental diagnostics. The model's ability to accurately detect dental caries has the potential to enhance early diagnosis, leading to timely interventions and improved patient outcomes. As the technology progresses, it anticipates the NTM-HICD model will be crucial in supporting dental professionals and revolutionizing how dental caries are detected and managed. However, further research and clinical validation are necessary before widespread implementation in detection of dental caries in the early stages based on abnormalities in the teeth.

REFERENCES

- [1] Oztekin, F., Katar, O., Sadak, F., Yildirim, M., Cakar, H., Aydogan, M., Ozpolat, Z., Yildirim, T.T., Yildirim, O., Faust, O., Acharya, U.R. (2023). An explainable deep learning model to prediction dental caries using panoramic radiograph images. *Diagnostics*, 13(2): 226. <https://doi.org/10.3390/diagnostics13020226>
- [2] Talpur, S., Azim, F., Rashid, M., Syed, S.A., Talpur, B.A., Khan, S.J. (2022). Uses of different machine learning algorithms for diagnosis of dental caries. *Journal of Healthcare Engineering*, 2022: 5032435. <https://doi.org/10.1155/2022/5032435>
- [3] Mohammad-Rahimi, H., Motamedian, S.R., Rohban, M., Krois, J., Uribe, S.E., Mahmoudinia, E., Rokhshad, R., Nadimi, M., Schwendicke, F. (2022). Deep learning for caries detection: A systematic review. *Journal of Dentistry*, 122: 104115. <https://doi.org/10.1016/j.jdent.2022.104115>
- [4] Huang, Y.P., Lee, S.Y. (2021). Deep learning for caries detection using optical coherence tomography. *medRxiv*.
- [5] Bhan, A., Goyal, A., Chauhan, H.N., Wang, C.W. (2016). Feature line profile based automatic detection of dental caries in bitewing radiography. In 2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE), Ghaziabad, India, 2016, pp. 635-640.

- <https://doi.org/10.1109/ICMETE.2016.59>
- [6] Zhu, H., Cao, Z., Lian, L., Ye, G., Gao, H., Wu, J. (2022). CariesNet: A deep learning approach for segmentation of multi-stage caries lesion from oral panoramic X-ray image. *Neural Computing and Applications*, 35: 16051-16059. <https://doi.org/10.1007/s00521-021-06684-2>
- [7] Majanga, V., Viriri, S. (2022). A survey of dental caries segmentation and detection techniques. *The Scientific World Journal*, 2022: 8415705. <https://doi.org/10.1155/2022/8415705>
- [8] Silva, G., Oliveira, L., Pithon, M. (2018). Automatic segmenting teeth in X-ray images: Trends, a novel data set, benchmarking and future perspectives. *Expert Systems with Applications*, 107: 15-31. <https://doi.org/10.1016/j.eswa.2018.04.001>
- [9] AL-Ghamdi, A.S., Ragab, M., AlGhamdi, S.A., Asseri, A.H., Mansour, R.F., Koundal, D. (2022). Detection of dental diseases through X-ray images using neural search architecture network. *Computational Intelligence and Neuroscience*, 2022: 3500552. <https://doi.org/10.1155/2022/3500552>
- [10] Prados-Privado, M., García Villalón, J., Blázquez Torres, A., Martínez-Martínez, C.H., Ivorra, C. (2021). A convolutional neural network for automatic tooth numbering in panoramic images. *BioMed Research International*, 2021: 3625386. <https://doi.org/10.1155/2021/3625386>
- [11] Chen, H., Zhang, K., Lyu, P., Li, H., Zhang, L., Wu, J., Lee, C.H. (2019). A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Scientific Reports*, 9(1): 3840. <https://doi.org/10.1038/s41598-019-40414-y>
- [12] Estai, M., Tennant, M., Gebauer, D., Brostek, A., Vignarajan, J., Mehdizadeh, M., Saha, S. (2022). Deep learning for automated detection and numbering of permanent teeth on panoramic images. *Dentomaxillofacial Radiology*, 51(2): 20210296. <https://doi.org/10.1259/dmfr.20210296>
- [13] Sathya, B., Neelaveni, R. (2020). Transfer learning based automatic human identification using dental traits-an aid to forensic odontology. *Journal of Forensic and Legal Medicine*, 76: 102066. <https://doi.org/10.1016/j.jflm.2020.102066>
- [14] Moutselos, K., Berdouses, E., Oulis, C., Maglogiannis, I. (2019). Recognizing occlusal caries in dental intraoral images using deep learning. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Berlin, Germany, pp. 1617-1620. <https://doi.org/10.1109/EMBC.2019.8856553>
- [15] Saini, D., Jain, R., Thakur, A. (2021). Dental caries early detection using convolutional neural network for tele dentistry. In 2021 7th International Conference on Advanced Computing and Communication Systems, Coimbatore, India, pp. 958-963. <https://doi.org/10.1109/ICACCS51430.2021.9442001>
- [16] Park, W.J., Park, J.B. (2018). History and application of artificial neural networks in dentistry. *European Journal of Dentistry*, 12(4): 594-601. https://doi.org/10.4103/ejd.ejd_325_18
- [17] Srivastava, M.M., Kumar, P., Pradhan, L., Varadarajan, S. (2017). Detection of tooth caries in bitewing radiographs using deep learning. *arXiv preprint arXiv:1711.07312*. <https://doi.org/10.48550/arXiv.1711.07312>
- [18] Lee, J.H., Kim, D.H., Jeong, S.N., Choi, S.H. (2018). Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. *Journal of Dentistry*, 77: 106-111. <https://doi.org/10.1016/j.jdent.2018.07.015>
- [19] Tuzoff, D.V., Tuzova, L. N., Bornstein, M.M., Krasnov, A.S., Kharchenko, M.A., Nikolenko, S.I., Sveshnikov, M.M., Bednenko, G.B. (2019). Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofacial Radiology*, 48(4): 20180051. <https://doi.org/10.1259/dmfr.20180051>
- [20] Prajapati, S.A., Nagaraj, R., Mitra, S. (2017). Classification of dental diseases using CNN and transfer learning. In 2017 5th International Symposium on Computational and Business Intelligence, Dubai, United Arab Emirates, pp. 70-74. <https://doi.org/10.1109/ISCBI.2017.8053547>
- [21] Lakshmi, M.M., Chitra, P. (2020). Classification of dental cavities from X-ray images using deep CNN algorithm. In Proceedings of the 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, pp. 774-779. <http://doi.org/10.1109/ICOEI48184.2020.9143013>
- [22] Leo, L.M., Reddy, T.K. (2021). Learning compact and discriminative hybrid neural network for dental caries classification. *Microprocessors and Microsystems*, 82: 103836. <https://doi.org/10.1016/j.micpro.2021.103836>
- [23] Umer Rashid. (2021). Dental Caries Dataset (Version 1). Zenodo.
- [24] Imak, A., Celebi, A., Siddique, K., Turkoglu, M., Sengur, A., Salam, I. (2022). Dental caries detection using score-based multi-input deep convolutional neural network. *IEEE Access*, 10: 18320-18329. <https://doi.org/10.1109/ACCESS.2022.3150358>