# A Deep Mathematical Morphological Neural Network for the Classification of Periapical Radiographs in the Diagnosis of Dental Diseases

**Grace Tam-Nurseman**
Department of Computer Science,
Lead City University, Ibadan, Nigeria

**Wilson Sakpere**
Department of Computer Science,
Lead City University, Ibadan, Nigeria

**Akinola S.o**
Department of Computer Science,
University of Ibadan, Nigeria

**Philip Achimugu**
Department of Computer Science,
Air Force Institute of Technology Kaduna, Nigeria.

*Abstract*

*The importance of medical imaging cannot be overemphasized as it is one of the best ways to diagnose a disease in medical practice objectively. In dentistry, no imaging means no objective diagnosis for it is with imaging dentists find hidden dental structure, bone loss, malignant or benign masses, and other dental diseases that cannot be discovered or examined during a visual examination. The use of dental radiographs also helps dentists to detect hidden dental diseases early. This model was developed integrating mathematical morphology (MM) operations (dilation, erosion, opening and closing) in the convolution layer of convolutional neural network (CNN), for data preprocessing and quality feature extraction. With its high sense of intelligence (artificial) obtained during training, the system receives dental images and analyses them automatically for various clinical findings with which 6 dental disease problems were solved. With an achieved accuracy of 99.78%, it can be established that this system can be used in dental clinics with high confidence giving very little or no-error-diagnosis. To make this system more scalable and robust, more dental diseases be added through other MM based theory like lattice, topology and random functions other than set theory-based MM used in this study.*

*Keywords: Mathematical Morphology (MM), Dilation, Erosion, Opening, Closing, Convolutional Neural Network (CNN)*

## 1.    Introduction

A deep mathematical morphological neural network for the classification of periapical radiographs in the diagnosis of dental diseases is a continuation of a previous work. The previous study background was on diagnosing and prognosing dental diseases using symptoms as parameters[1]. The system was developed using Bayes rule as a mathematical model. It was an exploratory and experimental work that exploited the knowledge of dental experts to develop an expert system that could diagnose 20 common dental diseases given their various symptoms. The study design and development were on simple IF-THEN rules.

Artificial neural network (ANN) algorithms for medical diagnosis and treatment today are relying heavily on CNN, state-of-the-art technology such as deep learning to assist medical doctors confirm and reconfirm their subjective diagnosis, hence this study.

This study adopted the fundamental question concerning the difference between biological neural network and artificial neural network (ANN) which seeks to know "whether the strength of the electric potential of a signal traveling along an axon is the result of multiplicative process and the mechanism of the postsynaptic membrane of a neuron adds the various potentials of the electric impulses OR an additive process and the mechanism of postsynaptic membranes only accept signals of certain maximum strength?"[2] The state-of-the-art technology CNN, is a produce of the answer to the first phase of the question where ANNs are developed from the multiplicative process and the addition of the various potentials of the electric impulses of the mechanism of the postsynaptic membrane of a neuron.

Three key objectives make up this study; Number one is to design and develop a mathematical morphological neural network. Number two is the use of dental periapical radiographs as basic dataset thereby developing a dental diagnostic model. This study in a bit to answering the adopted question framed its research question to seek a way to apply CNN differently yet get expected or better output. The study experimented on the additive process of an ANN and accepting the signals of certain maximum strength of the mechanism of the postsynaptic membranes thereby employing the services of the operations of mathematical morphology which are basically dilation and erosion. In other words, this experimental study shifted from the regular or conventional approach of CNN by replacing the convolution layer of CNN from which the CNN derived it name with MM operations. With the outburst of CNN, ANN model developers and designers are so carried away especially in feature extraction such that little or no attention is given to ancient technologies like mathematical morphology. MM is more mentioned in literatures for segmentation and filtering applications than feature extraction[3]. The presence of CNN has drifted developers' minds away from realizing the powerful feature extraction ability in MM. Most literature articles on automated medical diagnosis are on the application of CNN.

The second objective upon which this system is developed is the dental disease diagnosis. People who suffer from one ailment or the other are best described as patients. Dental patients are people who suffered from dental disease(s). One best way to diagnose dental disease(s) is the use of medical radiological equipment such as X-rays, MRI, CT scan to mention but a few. The use of radiological equipment is advantageous to doctors as they help in the detection of diseases which are not symptomatic and beyond what the ordinary eye can see during physical examination. The radiological equipment such as X-rays help dentists to detect

hidden dental diseases early. A "disease when detected at the point when it is not symptomatic certainly gives good prognosis because the chances of altering its natural course with adequate interventions are very high which in turn halts the progression of the disease"[4].

Automated disease diagnosis models abound but not much is done on dental diseases. Mulrenan *et al* reviewed 23 articles on the application of AI for automated medical diagnosis. 11 out of these reviewed papers were on chest X-ray for covid-19 diagnosis, 12 were on CT scan for same disease[5]. Rogers *et al* in decades past reviewed 58 articles bothering on different diseases among which non was dental disease[6]. Among the several articles on automated disease diagnosis reviewed by Kumar *et al*, is void of dental disease[7].

## 2.      Related Works

Several literatures have been reviewed to know what exists in scholarly literatures related to this study.   The reviewed literatures informed this study some gaps and areas that are presently making rounds. Among the reviewed literatures, some gaps that are observed discussed in this section.

LeNet was the first application of CNN by Yann LeCun and his team in 1998. The model was developed to recognize handwritten digits and was named LeNet after the principal author. The system was developed with only a few layers and few filters due to limitations of computational resources at that time. A major shortcoming of LeNet is that it was designed for a specific size input and it didn't do well in colour image recognition[8].

Kumar *et al* also reviewed various literatures on different disease diagnosis with various AI methods among which were Naïve Bayes, Generative Adversarial Networks (GAN), Convolutional neural networks, without a mention of MM[9].

A paper by Lakhani e*t al* provided a high-level overview of how to build a deep neural network for medical image classification, and codes that can help those new to the field begin their informatics projects. The aim of the tutorial was to spark interest and provide a basic starting point for those interested in machine learning in regard to medical imaging. "The tutorial assumed basic understanding of CNNs, some Python programming language (Python 3.6, Python Software Foundation, Wilmington DE), and is more of a practical introduction to using the libraries and frameworks. This work had space constraints and therefore did not cover everything in full detail[10].

MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications is a work by Howard *et al*. The authors demonstrated vision application embedded MobileNets for mobile devices based on depthwise separable convolutions.  In this work, a reasonable amount of accuracy was traded off for size and latency. The work was limited to mobile devices[11].

An Introduction to Deep Morphological Networks is a work by Keiller *et al*, where the convolution filters were replaced with non-linear filters of mathematical morphology. The authors failed to subject same data set to implement a CNN model in order to compare performance between the morphological network with that of CNN[12].

Wu *et al* reviewed Text Classification Methods based on Deep Learning (Convolutional Neural Network-Based (CNN-Based), Recurrent Neural Network-Based (RNN-based, Attention Mechanisms-Based and so on). In their review, many studies proved that text classification methods based on deep

learning outperform the traditional methods. This is because text classification methods based on deep learning is void of cumbersome feature extraction process and have higher prediction accuracy for a large set of unstructured data.[13]

Arunnehru *et al* in a paper titled "Human Action Recognition using 3D Convolutional Neural Networks with 3D Motion Cuboid Surveillance Videos used an advanced approach to propose for suspicious action recognition in intelligent video surveillance. They used 3D-CNN with 3D motion cuboid for action detection and recognition in real-time surveillance video to stop crimes. The experiments were conducted on KTH and Weizmann dataset and the outcome is a tremendous performance[14].

Human Activity Recognition with Convolutional Neural Networks is a paper by Bevilacqua *et al* in which the authors used CNN to classify human activities. The research work used raw data obtained from a set of inertial sensors. They got a promising result[15].

Ankita *et al* proposed a model that combined convolutional layers with long short-term memory (LSTM), along with the deep learning neural network for human activities recognition. The model extracts feature and categorizes them with some model attributes automatically. They used a dataset of UCI-HAR for Samsung Galaxy S2 in the proposed architecture for various human activities. The work achieved an accuracy of 97.89%for activity detection capability than traditional algorithms, which is a very good one[16].

Zeng *et al* developed a work based on CNN, to automatically extract discriminative features for activity recognition. They carried out their "experiments on three public datasets, Skoda (assembly line activities), Opportunity (activities in kitchen), Actitracker (jogging, walking, etc.) the work when applied in real life situation achieved higher accuracy than existing state-of-the-art methods of that time[17].

Krizhevsky *et al* developed a model that could classify 1.2 million high-resolution images into different classes. The neural network was made up of 60 million parameters and 650,000 neurons, consisted of five convolution layers some of which were followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. The model was trained with ReLU as an activation function to add non linearity to the model. AlexNet was the first GPU based CNN model and was 4 times faster than previous models. Though the success of this work brought about a revolution in computer vision it has its drawbacks such as the model depth compared with later CNN architectures like VGGNet, GoogleNet, etc. AlexNet convolution filters were 5 by 5 which "is quite large as large filters were discouraged shortly after" [18].

Besides the reviewed articles above, Raju and Modi proposed feature extraction technique for dental X-ray image using Fourier descriptor features for shape analysis and Gray Level co-occurrence Matrix (GLCM) properties such as Energy, Contrast, Correlation and Homogeneity to describe the texture of the extracted Tooth[19]. In the algorithm, both shape and texture analysis of the extracted tooth are explored as features. However, the algorithm fails when two images have similar texture and shape properties." Future work will involve finding a novel feature extraction technique which explores the geometry of teeth which remains inherently unique to each individual.

A Survey: "Segmentation in Dental radiographs for Diagnosis of Dental Caries" by Krithigaand Lakshmi is a paper which reviewed some influential papers in the dental x-ray segmentation

literature[20]. According to this paper, among the methods proposed only very few are fully automatic. Some methods were not accurate in segmenting caries due to improper teeth segmentations and some of them due to noise, low contrast and unusual arrangement of teeth. The future work will focus on retrieving dental radiographs using content-based retrieval systems.

A study was conducted by Na`am *et al* to overcome the difficulties in identifying proximal caries through image processing of panoramic dental x-ray image[21]. The image processing method used in this study is Multiple Morphological Gradients. These image processing results can be used to appropriately identify proximal caries and the level of severity. Based on the results of this study, further research is needed to compare the results of Panoramic Dental X-Ray processed by multiple morphological gradients (mMG) with the results of bitewing Dental X-Ray from any proximal caries Patients.

## 3. Research Methodology
## 3.1 Research Approach

This study achieved its research objectives and was able to answer the research question through a mixed research method of quantitative and qualitative approach. The approach allowed for the collection and analysis of the various data types that were needed to design and implement the system. The data collection and analysis methods were approached in both methods. The 6 common types of dental diseases to diagnose in this study were decided through a closed-ended questionnaire where several dental diseases were listed for the participant to tick the 6 most common dental diseases from his/her professional experience (see appendix A). Catherine Dawson[22] in her book, "Practical Research Methods" when explaining the advantages and disadvantages of open and close questionnaire said close questions are easy for participants to tick their opinion and by that, the likelihood of responses to all questions are increased. The results of this questionnaire were analyzed with the use of Microsoft office Excel. This quantitative method was easier as respondents didn't have to explain opinions rather, just a tick. From Hajic and Simonett comparisons of qualitative and quantitave image analysis this study depended more on quantitative methods since the study involved the use of probabilistic statements such as error rates associated with identification and classification.

## 3.2 Research Design

This research is both experimental and exploratory. It is experimental because it attempts to experiment on integrating MM operations into the convolution layer of CNN thereby replacing the mathematics of convolution with that of set theory. The system was developed with dental periapical radiographs as the core dataset. At this point of using dental radiographs as training and testing dataset, the experiences of dental experts were explored; what they look at for in dental radiographs and the steps taken to planning a treatment.

The study is designed to reveal the possibility of shifting from the regular application of CNN algorithm on image and pattern recognition.

## 3.3 Data Acquisition

In every research work, data collection plays an important role. To get substantive findings in research work, necessary information is gathered through some form of data gathering methods. Usually, these data collection methods in most cases are guided by the research question formed at the onset of the research work. It is a key process of the work without which the research work remains a mere idea. There are different styles and methods of data collection but the result is the same; data is collected. The data that is acquired for this study are periapical dental radiographs and treatment planning for six common dental diseases. We chose to use periapical radiographs because being a result of an intraoral x-ray, it is more detailed and common than the extra oral

X-rays[23].

To achieve the objectives of this research work, data (secondary) was collected. Benchmark dental dataset was collected from some repositories online. The benchmark dataset was used for both training and testing of the model. Besides the benchmark dataset, random dataset was also collected in renowned hospitals.

In this work, we chose to use interviews and questionnaires as our key research instruments because it is one of the ways we talked to knowledgeable people one-on-one in the discipline (dentistry) we have chosen to work in. we drafted a list of questions to guide us in asking questions during the interview. Rubin and Rubin in Qualitative Interviewing said "the researcher is looking for rich and detailed information, not for yes-or-no, agree-or-disagree response" [24]. In the light of this, we listened to knowledgeable people in renowned hospitals which certainly made us explore thoroughly dental diseases, which can be transformed into an abstract machine to help experts confirm their subjective diagnosis, upcoming dental doctors learn and the society at large to benefit from.

Besides in-depth qualitative interviewing and questionnaires, we also got data through participation. This was to enable us to get familiarized with the various people (dental doctors, dental assistants, dental technologists, etc.) involved in dentistry, the language, and allow future interviewees to get to know us before we start asking them questions. Another excellent reason is building trust and a comfortable interview environment.[25]

Participant observation as mentioned earlier is another major data collection instrument that we used besides other non-significant methods. The Oxford English dictionary defines observation as the ability to notice things, especially significant details. Observation in this context is a way to gather data by watching people and activities in their natural settings. Here we established a rapport so that people will go about their business as usual when we show up. The above-stated methods of data collection for this study gave us a clue into how dental doctors analyze dental radiographs and how these dental radiographs are mapped to dental diseases. Figure1displays some of the dental images collected for this study.

**Figure 1 Dental images of Digital Radiographs.**

## 3.4 System Framework

Figure 2 is a summary of the frame work of the model. The system uses dental Periapical radiographs as its training data. The data is in the ratio of 70%:30 % where 70% is used to train the model and 30% for testing.
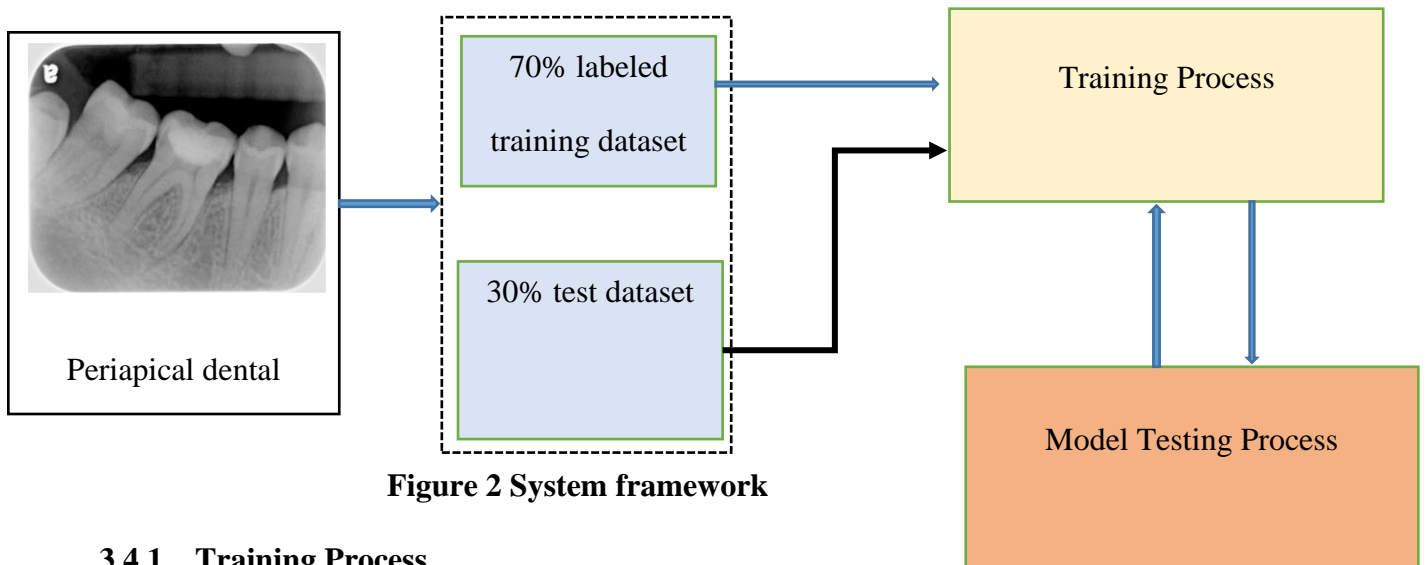


**Figure 2 System framework**

### 3.4.1 Training Process

Figure 3 is a graphical representation of the training process. The model is trained with 70% of collected data. A preprocessor is designed to receive the raw data to cleanse the data of any noise for easy training. Denoised data is passed on to a feature extractor known as MM-NN

extractor to extract features of the various images for learning purpose. Extracted feature map is downsampled at the pooling layer. Downsampled feature map is flattened to a vector as input and fed to the fully connected layer through a softmax activation function for final output classification.
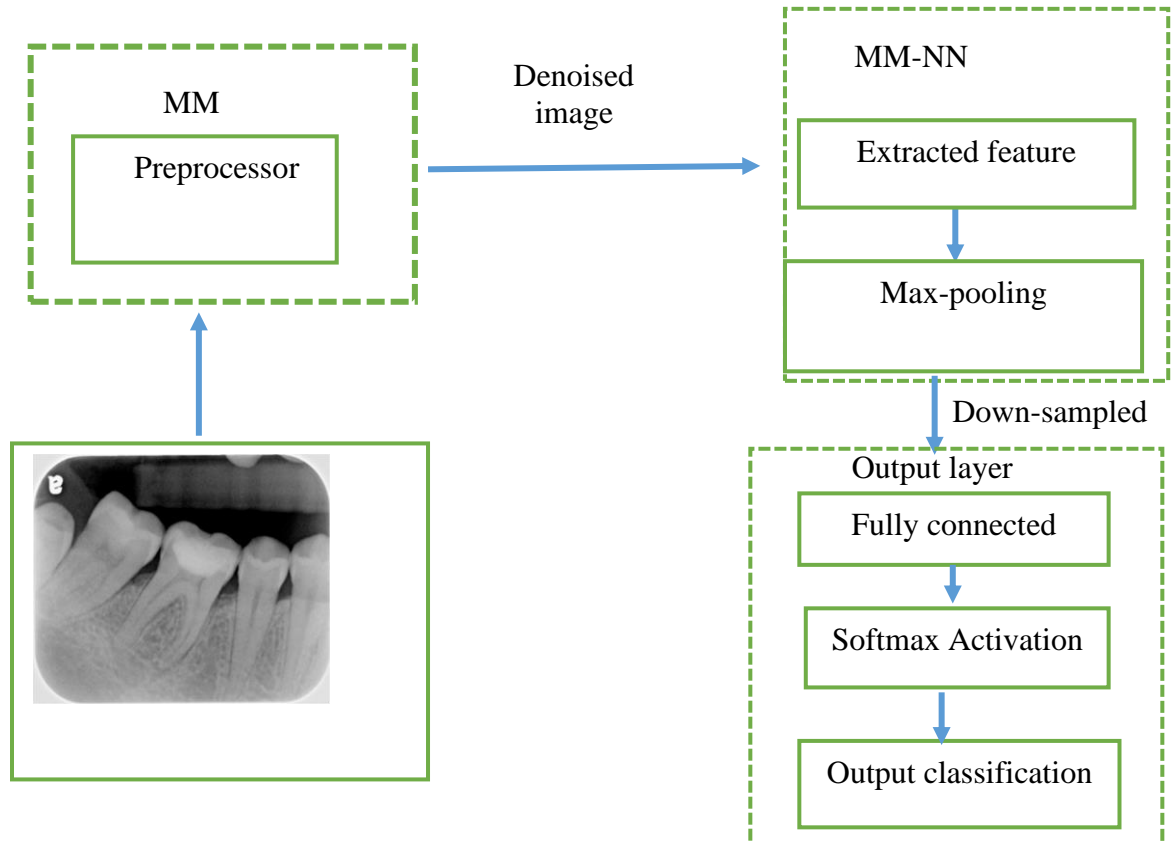


**Figure 3 Training section of system framework**

## 3.4.2   Testing Process

The Testing process as depicted in Figure 4 is a process to check how much the model has learned to identify and classify trained data. Unlike the training process where data are labeled, here unlabeled data is passed to the model. If the output is not desirable, output loss function is calculated and learnable parameters are adjusted and retraining continues until loss function is minimized drastically.
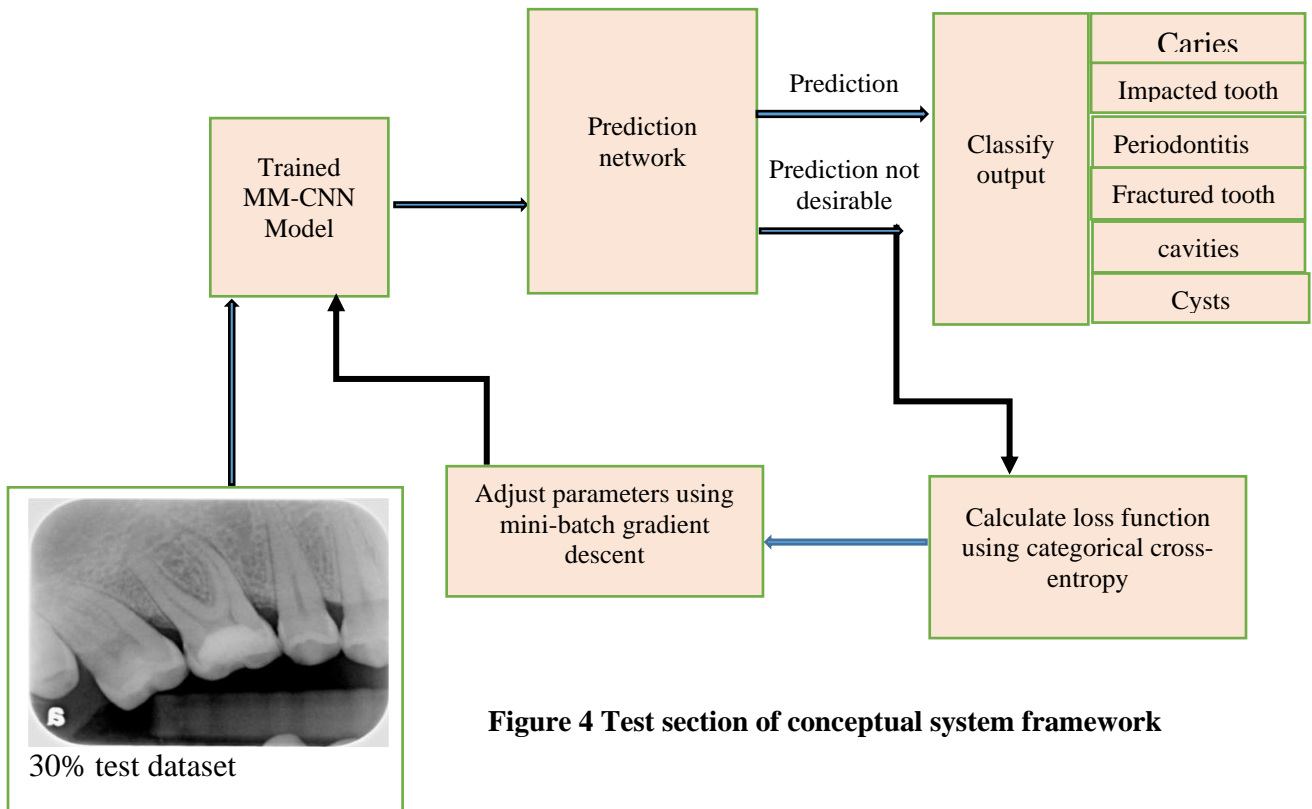
**Figure 4 Test section of conceptual system framework**

## 3.5  System Evaluation.
### 3.5.1  Performance Evaluation Metrics

In this experimental study, a CNN model was developed alongside the MM-NN model with a total of 864 images. At the end of training the models, the models were tested with several random data including the 30% test data. Accuracy, precision, recall and error rate were the performance metrics used to evaluate the models. The outcomes are expressed in Figures 5, 6, 7 and 8.

**Accuracy:** Accuracy tells us how often our model's classifier is correct.
It is calculated as the sum of all true values divided by total values as demonstrated in Equation1 below.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

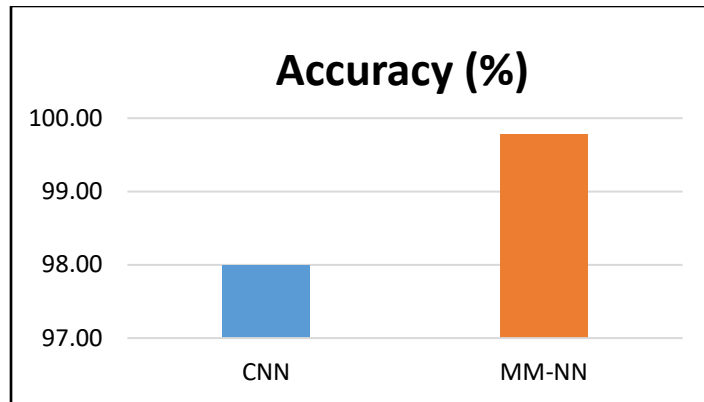Figure 5 shows an accuracy of 99.78% MM-NN against 97.99 CNN.

**Figure 5 System's accuracy bar chart**

**Precision**: How often is the model's predicted positive classification right? It is calculated as the true positives divided by total number of predicted positive values as seen in Equation 2.

$$Precision = \frac{TP}{TP+FP}$$ …………………………...(2)

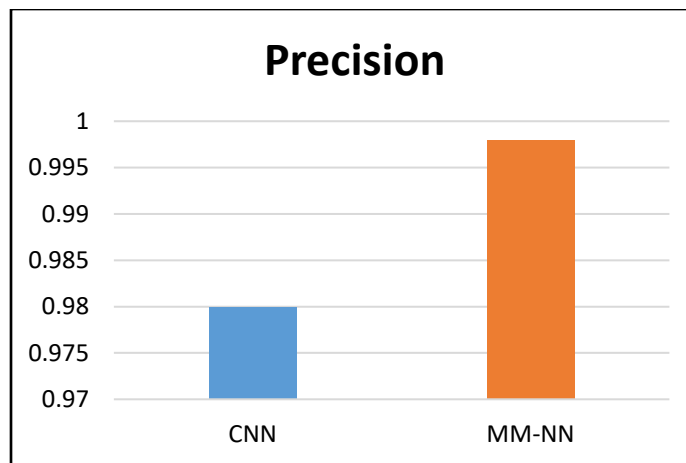 As seen in Figure 6, MM-NN has a precision value of 0.998 against CNN with precision value 0.979



**Figure 6 System's precision bar chart**

**Recall/Sensitivity**: The model's ability to predict the positive values correctly as shown in Equation 3. It is calculated as the true positives divided by total number of real positive values

$$Recall = \frac{TP}{TP+FN}$$ …………………………………(3)

Figure 7 shows a recall value of 0. 997 MM-NN against 0.979 CNN
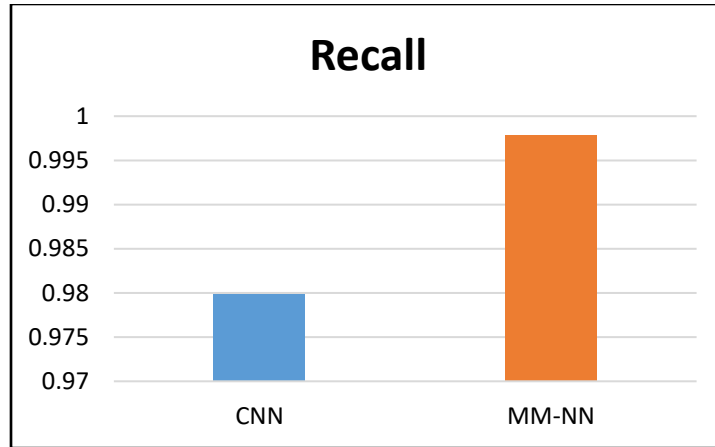
**Figure 7. System's recall bar chart**

**Specificity:** Specificity as seen in Equation 4 corresponds to the true negative rate of the considered class. It is calculated as true negative divided by the total number of all the negative values

Specificity $= \frac{TN}{TN+FP}$ …………………………(4)

Figure 8 shows a less error value of mm-NN 0.0021 against CNN 0.0200



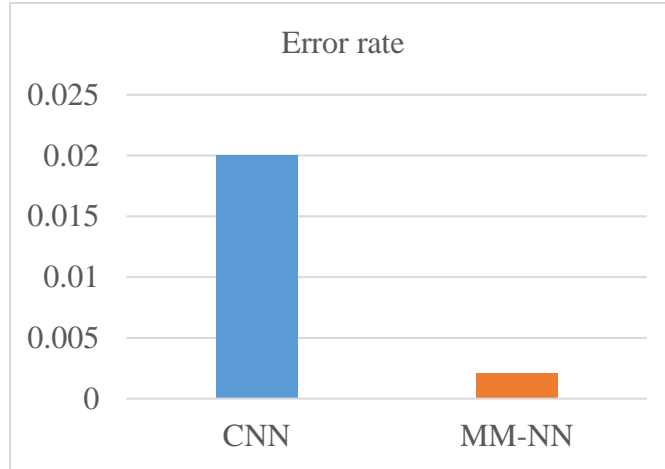**Figure 8 System's error rate bar chart**

*Table 1 Metric Table of the Model*

|  |  | Actual | | | | | |
|---|---|---|---|---|---|---|---|
|  |  | dA | dB | dC | dD | dE | dF |
|  | dA | $TP_A$ | $E_{AB}$ | $E_{AC}$ | $E_{AD}$ | $E_{AE}$ | $E_{AF}$ |
|  | dB | $E_{BA}$ | $TP_B$ | $E_{BC}$ | $E_{BD}$ | $E_{BE}$ | $E_{BF}$ |
| Predicted | dC | $E_{CA}$ | $E_{CB}$ | $TP_C$ | $E_{CD}$ | $E_{CE}$ | $E_{CF}$ |
|  | dD | $E_{DA}$ | $E_{DB}$ | $E_{DC}$ | $TP_D$ | $E_{DE}$ | $E_{DF}$ |
|  | dE | $E_{EA}$ | $E_{EB}$ | $E_{EC}$ | $E_{ED}$ | $TP_E$ | $E_{EF}$ |
|  | dF | $E_{FA}$ | $E_{FB}$ | $E_{FC}$ | $E_{FD}$ | $E_{FE}$ | $TP_F$ |

Representing caries, fractured tooth, dental cavities, dental cyst, imparted tooth, and periodontitis with dA, dB, dC, dD, dE, and dF respectively in table 1, the rows of the matrix correspond to a predicted class while the column of the matrix correspond to an actual class as seen in table 3.2.

*Table 2 Illustration of the Application of Accuracy metric*



FP (A)
Predicted as 'Positive A'
but actually 'Negative A'

"TN (A)
Predicted as 'Negative A'
and actually 'Negative A'

FN (A)
Predicted as 'Negative A'
but actually 'Positive A'

The accuracy metric of the model as shown in table 2 can be interpreted thus;

- TP: Red texted diagonal values are TP values for all the classes of  diseases dA-dF.
- FP:  Blue coloured row 'A' apart from the red texted TP is FP for disease A. Likewise other rows, every row apart from the TP value is FP for the disease in that row.
- FN:  Green coloured column  apart from the TP on that column is FN for disease A. Likewise other columns for the other diseases.
- TN: Content of gold coloured broken lines is TN for disease A. TN values are the values of that disease apart from the row and coloumn that contain that disease.

Above explained metric parameters and their respective equations (1-4) were calculated thus;

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

Where:
Total number of FP for a class = Sum of values in the corresponding row excluding the TP

Where FP for;
class A = $E_{BA}+E_{CA}+E_{DA}+E_{EA}+E_{FA}$
Class B = $E_{AB}+E_{CB}+E_{DB}+E_{EB}+E_{FB}$
Class C = $E_{AC}+E_{BC}+E_{DC}+E_{EC}+E_{FC}$
Class D = $E_{AD}+E_{BD}+E_{CD}+E_{ED}+E_{FD}$
Class E = $E_{AE}+E_{BE}+E_{CE}+E_{DE}+E_{FE}$
Class F = $E_{AF}+E_{BF}+E_{CF}+E_{DF}+E_{EF}$

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

Where:
Total number of FN for a class = Sum of values in the corresponding column excluding the TP

Where FN for;
class A = $E_{AB}+E_{AC}+E_{AD}+E_{AE}+E_{AF}$
Class B = $E_{BA}+E_{BC}+E_{BD}+E_{BE}+E_{BF}$
Class C = $E_{CA}+E_{CB}+E_{CD}+E_{CE}+E_{CF}$
Class D = $E_{DA}+E_{DB}+E_{DC}+E_{DE}+E_{DF}$
Class E = $E_{EA}+E_{EB}+E_{EC}+E_{ED}+E_{EF}$
Class F = $E_{FA}+E_{FB}+_{CCF}+E_{FD}+E_{FE}$

$$Specificity = \frac{TN}{TN + FP}$$

Where:

$Total\ number\ of\ TN\ for\ a\ class = Sum\ of\ all\ columns\ and\ rows\ excluding$
$that\ class's\ column\ and\ row$

*FP for the various classes is defined already in precision*

*Where TN for;*

*ClassA*
$$= TP_B + E_{BC} + E_{BD} + E_{BE} + E_{BF} + E_{CB} + TP_C + E_{CD} + E_{CE} + E_{CF} + E_{DB} + E_{DC} + TP_D + ED_E + E_{DF} + E_{EB} + E_{EC} + E_{ED} + TP_E + E_{EF} + E_{FB} + E_{FC} + E_{FD} + E_{FE} + TP_F.$$

*classB* $= TP_A + E_{AC} + E_{AD} + E_{AE} + E_{AF} + E_{CA} + TP_C + E_{CD} + E_{CE} + E_{CF} + E_{DA} + E_{DC} + TP_D + ED_E + E_{DF} + E_{EA} + E_{EC} + E_{ED} + TP_E + E_{EF} + E_{FA} + E_{FC} + E_{FD} + E_{FE} + TP_F.$

*classC* $= TP_A + E_{AB} + E_{AD} + E_{AE} + E_{AF} + E_{BA} + TP_B + E_{BD} + E_{BE} + E_{BF} + E_{DA} + E_{DB} + TP_D + ED_E + E_{DF} + E_{EA} + E_{EB} + E_{ED} + TP_E + E_{EF} + E_{FA} + E_{FB} + E_{FD} + E_{FE} + TP_F.$

*classD* $= TP_A + E_{AB} + E_{AC} + E_{AE} + E_{AF} + E_{BA} + TP_B + E_{BC} + E_{BE} + E_{BF} + E_{CA} + E_{CB} + TP_C + E_{CE} + E_{CF} + E_{EA} + E_{EB} + E_{EC} + TP_E + E_{EF} + E_{FA} + E_{FB} + E_{FC} + E_{FE} + TP_F.$

*class E* $= TP_A + E_{AB} + E_{AC} + E_{AD} + E_{AF} + E_{BA} + TP_B + E_{BC} + E_{BD} + E_{BF} + E_{CA} + E_{CB} + TP_C + E_{CD} + E_{CF} + E_{DA} + E_{DB} + E_{DC} + TP_D + E_{DF} + E_{FA} + E_{FB} + E_{FC} + E_{FD} + TP_F.$

*class F* $= TP_A + E_{AB} + E_{AC} + E_{AD} + E_{AE} + E_{BA} + TP_B + E_{BC} + E_{BD} + E_{BE} + E_{CA} + E_{CB} + TP_C + E_{CD} + E_{CE} + E_{DA} + E_{DB} + E_{DC} + TP_D + E_{DE} + E_{EA} + E_{EB} + E_{EC} + E_{ED} + TP_E.$

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

## 4    Experimental Setup

The model is developed to classify six dental diseases through periapical dental radiographs. Each of these six dental diseases has its peculiar pattern and is different from the others. To extract areas of interest from the images, the location of the diseased portion of the image, the size of the diseased portion, and the orientation of the diseased portion were key factors that were put into consideration.
First, all the dental images were resized to 256 x 256 binary images Every image passes through a denoising process with the application of MM operators for proper detection of areas of interest to be extracted.
we integrated MM into the CNN model replacing the conventional convolution layer as a morphological layer to effectively determine the shape of the teeth/gum due to the effect of disease, grouping same set as a particular disease.

The morphological feature extraction tool embedded in the CNN model was developed using dilation, erosion, and opening, and closing operators of mathematical morphology.

The objects of our images were represented as set points where each element of the set is a vector with x and y coordinates.

To remove the presence of internal noise from our images, we first applied dilation to the images. We started by defining a 3x3 mask (B), otherwise known as a structuring element with an origin t, with which we dilated the image X. The structuring element acted as a convolution mask.

Dilation operation was first applied to the images to expand some of the small areas thereby enlarging some noisy areas as shown in Figure 9.

Dilation is defined as;

$$X \oplus B = \{P \in Z^2 \,|\, p = x + b,$$

$$x \in X, b \in B\}$$

P = summation of the Input image set points (X $\{x_1, x_2\}$) and the structuring element set points (B $\{b_1, b_2\}$) at particular points.
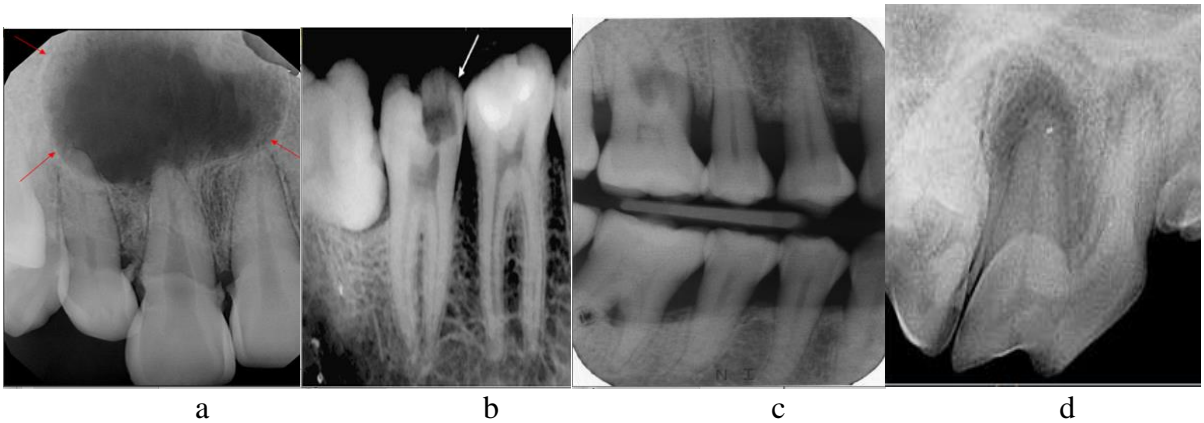


| a | b | c | d |

**Figure 9 Noisy dental images**



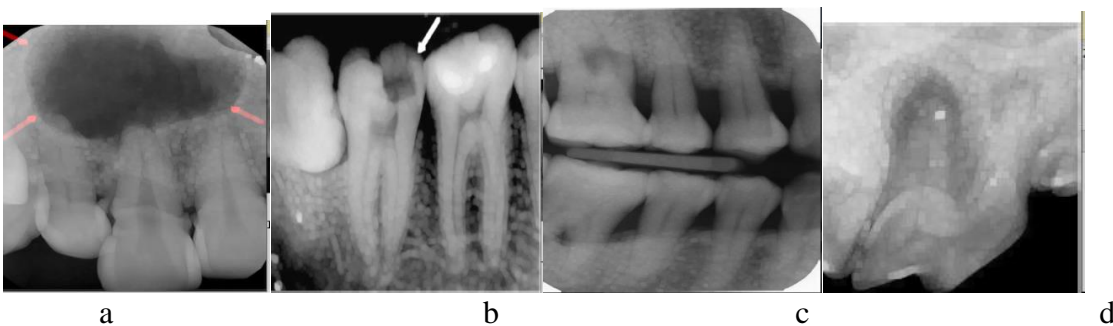| a | b | c | d |

**Figure 10 Dilated dental images**
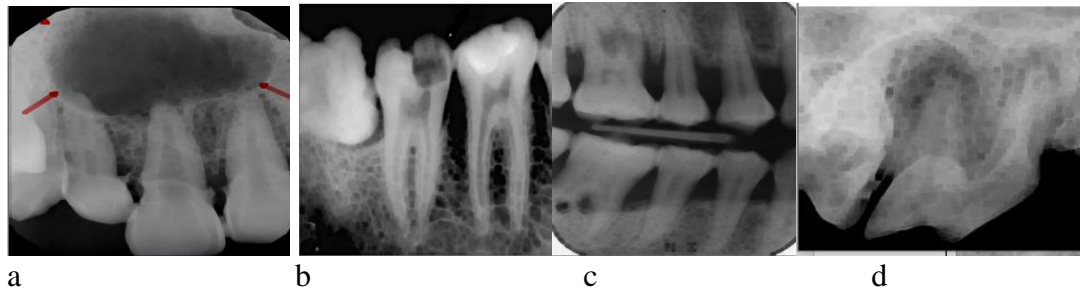
a          b          c          d

**Figure 11 Eroded dental images**

The images used in this study are represented as sets of points in two a dimension space $Z^2$, hence every pixel point in both our images and structuring element is represented by an ordered pair of pixel values X $\{x_1, x_2\}$ for the images and B $\{b_1, b_2\}$ for the structuring element. Consequent upon this, the application of MM is based on set theory.

Here, vector addition of the image set points and the structuring element set points were carried out. All the locus of points that satisfied the dilation equation is the dilated pixel sets.

The dilation process takes care of all internal noises as well as expanded the boundary of the image depicted in Figure 10. After the dilation operation on the input image, dilated images were eroded as shown in Figure 11 to clean the background by eliminating some noisy areas at the same time filtering out some boundary pixels added to the image during dilation.

Erosion is defined as:
$$X \ominus B = \{P \in Z^2 | p + b \in X$$
$$\text{For every } b \in B\}$$

Where;

X = set points of image

B = set points of structure ng element

P = point set of X

b = point set of B

p + b = Vector addition of input image set points (X $\{x_1, x_2\}$) and the structuring element set points (B $\{b_1, b_2\}$) at particular points.

P = Result of $p + b$

The erosion formula says that p + b must be within our image set points for every b $\in$**B**

Using the above formulae, the expanded boundaries of the dilated images eroded because those set points did not satisfy $p + b \in X$**.**

For contour smoothing of the images, a combined operation of erosion and dilation was applied. First, combination of erosion followed by dilation with the same structuring element, a combination termed morphological opening.

Morphological opening operation is the opening of set X by a structuring element B which is defined as;
$$X \cdot B = (X \ominus B) \oplus B.$$

Above formula is interpreted as morphological opening equals the erosion of X by B, followed by the dilation of the result by the same structuring element B.

With the defined structuring element, images were first eroded which shrank the images but removed all unwanted elements from the images. The shrunk images were followed by a dilation operation with the same structuring element. The dilation operation compensated for the shrinking thereby expanding the shrank images. It restored the shapes to their original sizes. Next, we applied dilation followed by erosion using the same structuring element, a situation known as morphological closing,

The closing operation is the closing of set X (set points of the image) by the structuring element B (set points of the structuring image), defined as;

$$X \cdot B = (X \oplus B) \ominus B$$

This is interpreted as morphological closing equals the dilation of X by B, followed by the erosion of the result by the same structuring element B.

The morphological opening is useful for removing small objects and thin lines from an image while preserving the shape and size of larger objects in the image.

In a summary, the morphological opening got rid of all unwanted objects from the images while preserving the shape and size of the useful objects. The morphological closing operation takes care of any internal noise present in any of our training and testing images' object regions.

The conventional convolution layer which is the brain behind CNN does multiplicative process of its input values with corresponding linear weights and takes an aggregate of the weighted input as depicted by the gold colour portions of Figure 12.
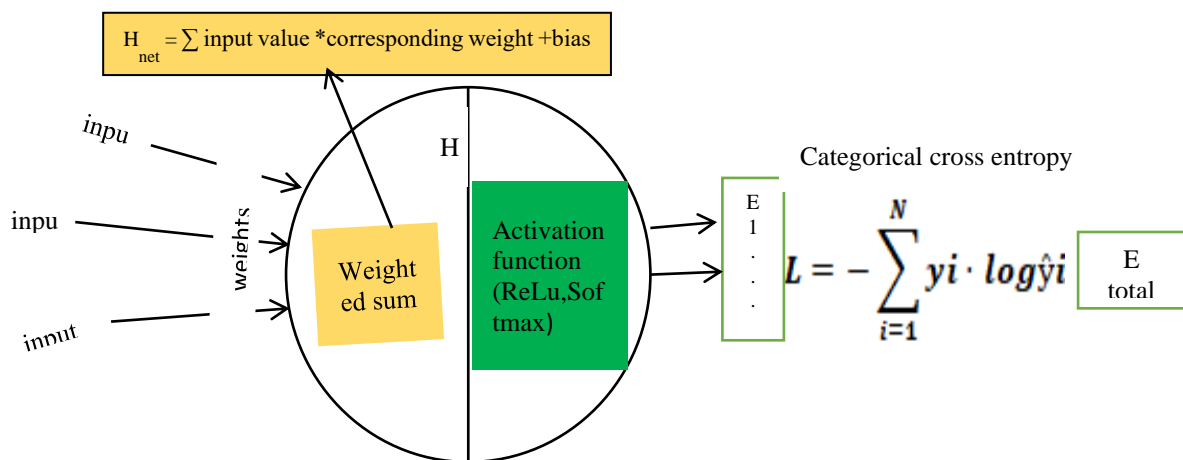


Figure 12 Graphical representation of CNN forward propagation

The weighted sum is passed through a non-linear activation layer depicted by the green-coloured portion of Figure 12 to get an output of the node H. In this layer, an element-wise operation is performed in which all negative pixels are set to zero (0).

f(x) = max (0, x) is applied to every value in the summed up weighted input data. If the function receives any negative input, zero (0) is returned; but if it receives any positive value x, it returns that value. As a result, the output has a range of 0 to infinite.

ReLU introduces non-linearity to the model's network and generates a rectified feature map. For this study, we choose ReLU over other activation functions because of its advantage over the others like sigmoid and hyperbolic tangent activation functions which have a vanishing gradient problem.[26] This step is carried out on every node in the network using the output of a layer as input to the next layer. At the output layer, the individual output neuron errors are calculated by taking a difference between the ideal output and actual output through the application of optimization function. These output errors are summed $E_1+E_2+E_3+E_4+\ldots En$ to give the total error of the model.
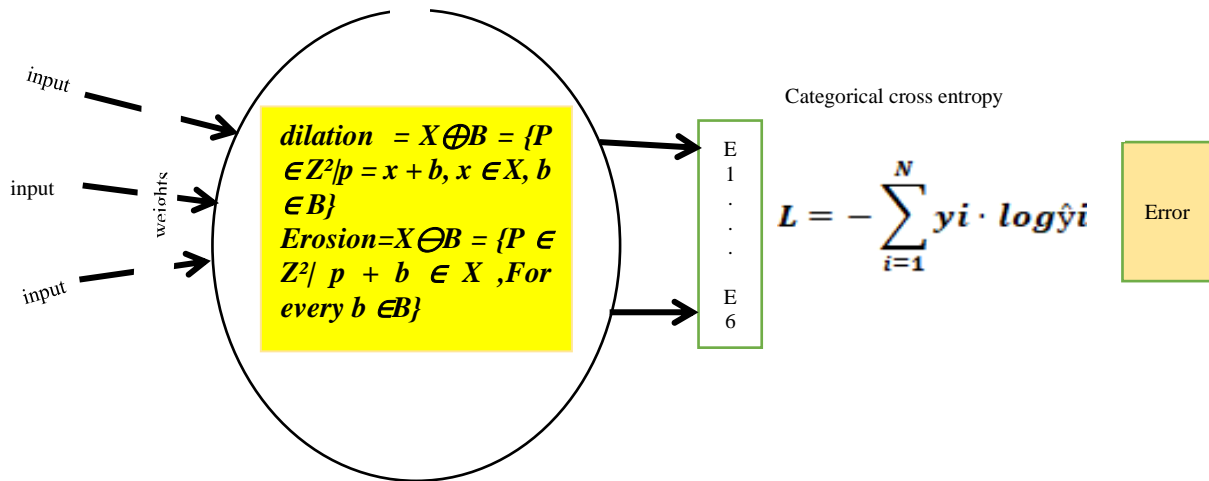


Figure 13 Graphical representation of MM forward propagation

The convolution process of CNN as depicted in Figure 13 is replaced with set-based MM operators; dilation, erosion and a combination of both (opening and closing).

MM operators, dilation and erosion as defined in Figure 9 depicts an image "X" under process being probed by a structuring element "B".

Where "X" and "B" are set points of the image and the structuring element respectively in $Z^2$ Cartesian grid as already explained. As depicted in Figure 9, MM-NN filters are non-linear, therefore have no need to pass through an activation layer or process like a CNN model. Dilation and erosion were used at the initial layers to identify simple structures while the combination of both (opening and closing) was used to identify complex structures.

The model is designed to extract areas of interest from the dataset using MM operations. First, we set a learning rate of 0.001 for the model. The learning rate is set small so as not to have over- fitting and then an objective function known as the loss function for the model was defined also. The MM-NN layer inputs were probed with 32 filters of 3x3 dimensions with a stride of 3 and valid padding. To avoid partial processing of the image, valid padding was introduced. Valid padding makes every pixel of the image valid so that the input image can get fully covered by the filter. Each MM-NN layer is followed by a pooling layer. At the pooling phase, max-

pooling was used in all the 3 layers. The feature maps obtained after the MM-NN convolution were max-pooled using a pooling stride of 2, 32 filters of 2x2 size to down-sample the MM-NN output. The down-sampling causes a drastic reduction in the image size thereby bringing about a reduction in spatial dimension.

The last pooling layer output is converted to a vector as input to a fully connected layer, then through a softmax activation layer to classify the images according to their classes in a one-hot-encoder.

## 5    Output Generation

This model is trained to use one-hot encoding; hence categorical cross-entropy is used as loss function.

The fully connected layer of this model receives input volume from the output of the last pooling layer. The fully connected layer of this study is a combination of different stages;

1.      The first stage receives the output of the last pooling layer as input and flattens it to a one-dimensional layer (1D).

2.      The next stage or process is another layer that received the flattened data and passed it through a linear function (also known as affine function) with an added bias as a constant. Mathematically expressed as;

$$y = f(Wx + b),$$

where;

$y$ = output,

$x$ = input,

$W$ = weights used in for the linear combination,

$b$ = bias

$f$ = activation function.

3.      The final output was then passed through a softmax activation layer for classification.

Input layer → MM-NN and pooling layers → Flattened layer → Linear function → Non-linear function → Softmax Activation → labels

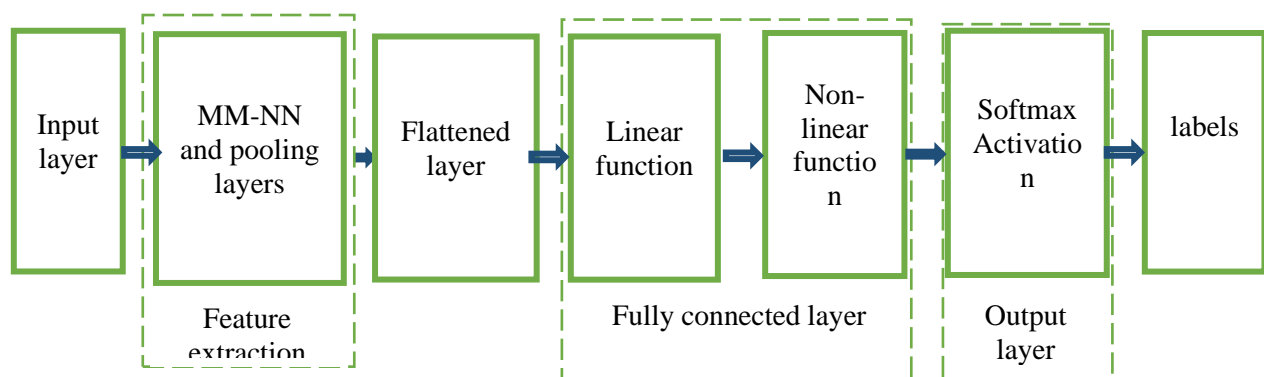Feature extraction | Fully connected layer | Output layer

**Figure 14 Data flow from input to output layer**

As demonstrated graphically in Figure 14, a summary of data flow from the input to the output and classification of the model. The model received dental images as input at the input layer. Areas of interest extracted for learning at the MM-NN layer. Extracted feature maps are spatially reduced at the pooling layer. Pooled feature maps flattened and passed through linear and non-linear functions at the fully connected layer. And finally, a non-linear activation function, softmax was used for the classification.

## 5.1    Model Parameter Adjustment in Backward Propagation

**1      CNN back propagation**: During the forward propagation, data was propagated in a batch size of 16 and from the outcome, the lost function which is the distance between real value Y and the predicted value y-hat was calculated using Categorical cross-entropy.

$$L = -\sum_{i=1}^{N} yi \cdot log\hat{y}i$$

Where;
L= Difference between system's predicted value, that is, outcome of softmax and expected value.
N= the number of classes; in this case, the number of classes is 6.
y = Expected value
$\hat{y}$  = predicted value which is the softmax outcome
$i$ = $i$-th sample in a set
Categorical cross-entropy is used to calculate the loss in this study as it is a multi-class classification problem where a given sample is classified into one of multiple classes. The model predicts the likelihood of input data belonging to one of a defined number of classes and then cross-entropy takes the output probability and measures the distance from the true value and gives the loss function.
 Having calculated the loss which is the error, the model is back-propagated by adjusting the key parameters, (the learnable weights and biases) of the model whose values were randomly initialized in the first layer of the model.
 Here the gradient of the loss function is calculated with respect to the weights and biases of the model using a mini-batch gradient descent algorithm.  The weights of the model are iteratively adjusted to reduce output error until the model reached a desired output. The relationship between the model's error and each of the learnable weights is a derivative $\frac{\partial E}{\partial W}$ that measures the extent or degree to which a slight difference in weight causes a change in the error. This is done when a l layer is given the derivative of its layer's error E with respect to its output Y ($\frac{\partial E}{\partial Y}$),  It gives back the derivative of error E with respect to its input X ($\frac{\partial E}{\partial X}$).

In order to update the parameters, the layers compute the derivative of the error E in respect to the weight ( $\frac{\partial E}{\partial W}$ ) and the derivative of error E with respect to the input X ( $\frac{\partial E}{\partial X}$).

2.  **MM-NN back propagation:** The back propagations of CNN and the MM-NN model are operated in the same fashion. As explained in the CNN back propagation, the MM-NN model's predicted outputs are compared with target output from which a loss function is calculated. The loss in MM-NN model is equally affected by the input "x" and structuring function "w" through its output "z".
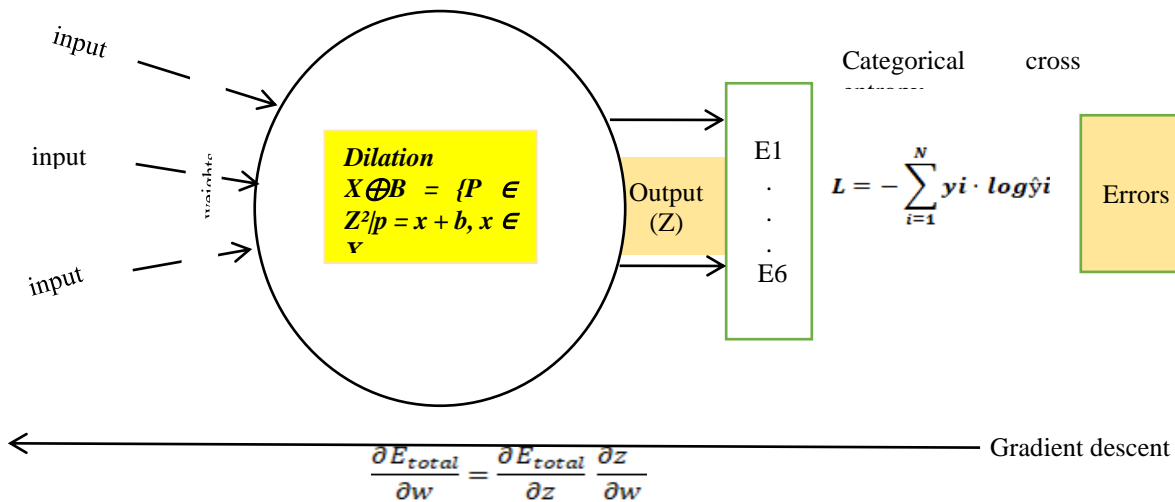


**Figure 15 Graphical representation of dilation back propagation**



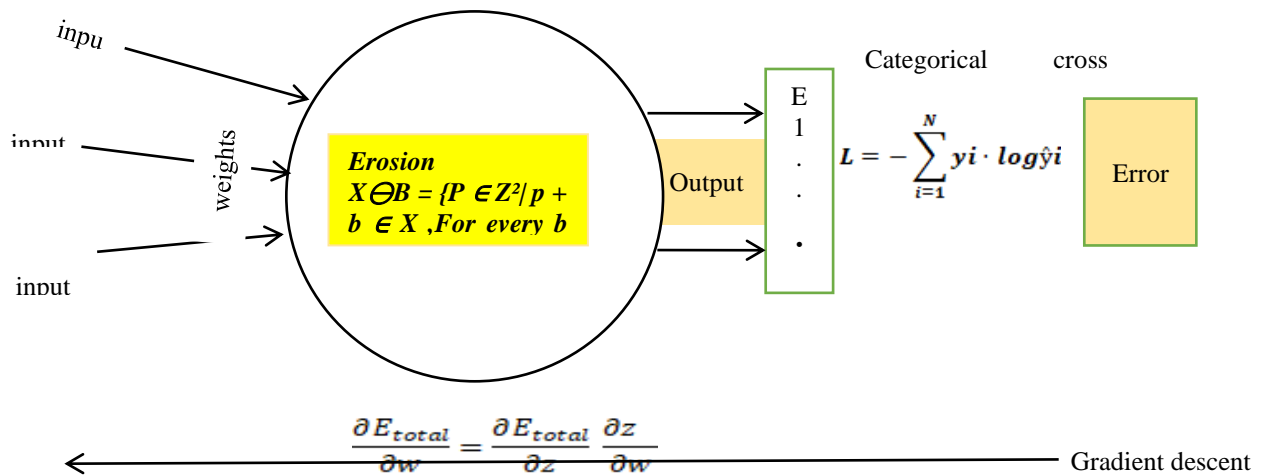**Figure 16 Graphical representation of erosion back propagation**

In Figures 15 and 16,
$\frac{\partial E_{total}}{\partial w}$ = Derivative of error with respect to structuring function w.

$\frac{\partial E_{total}}{\partial z}$ = Derivative of error with respect to output z.

It is the back propagated gradient from subsequent node that acts as input to the prior node.

$\frac{\partial z}{\partial w}$ = Derivative of output with respect to structuring function w.

It is also the computed dilation or erosion at the various nodes.

Using chain rule, $\frac{\partial z}{\partial w}$ is calculated first with which result $\frac{\partial E_{total}}{\partial z}$ is calculated and finally $\frac{\partial E_{total}}{\partial w}$ which is the loss with respect to structuring function w. The parameter adjustment and updating of weights reiterates until the loss converges.

## 6 Results and Discussions

One of the key purposes of this study is to develop an intelligent system to diagnose dental radiographs through the experimentation of shifting from the conventional CNN algorithm to MM-NN algorithm. This shifting is the change from the result of a "multiplicative process" ($\sum$w*x +b) of the strength of the electric potential of a signal traveling along an axon on which CNN is built, to an "additive process" (maximum (F$\oplus$h(s)=max{f(n)+h(s-x)}) or minimum (F$\ominus$h(s)=min{f(s+n)-h(n)})) from which the mechanism of postsynaptic membranes only accepts signals of certain maximum strength. The system was successfully designed, developed and implemented from scratch through the employment of mathematical morphological tools to achieve the additive process of the postsynaptic membrane.

### 6.1 Outcome of Varying Filter Sizes

During the training phase of the model, the depth of each layer of both the CNN and the MM-NN were varied starting from 8, 16, 32 and 64. These depths represent the filter numbers applied to probe the training images. Table 3 displays the various filter depth outcomes. From the table, it is clear that layer depths 32 and 64 outputted similar accuracy results. Therefore, for reasons of training and computational time, depth 32 became an obvious depth to adopt; hence, 32 filters were used in all the layers of the model. Figure 17 is a graphical representation of table 3.

*Table 3 Accuracy Using Different Filter Sizes*

| Diseases | Filter Sizes | | | |
|---|---|---|---|---|
| | 64 | 32 | 16 | 8 |
| Dental caries | 0.9907985 | 0.9907984 | 0.9815465 | 0.9707984 |
| Dental Cysts | 0.9997654 | 0.9997654 | 0.9902356 | 0.9697654 |
| Fractured tooth | 0.9989684 | 0.9989684 | 0.9757938 | 0.9489684 |
| Periodontitis | 0.9999688 | 0.9999687 | 0.9688943 | 0.93997679 |
| Impacted tooth | 0.9987994 | 0.9987994 | 0.9889999 | 0.95657903 |

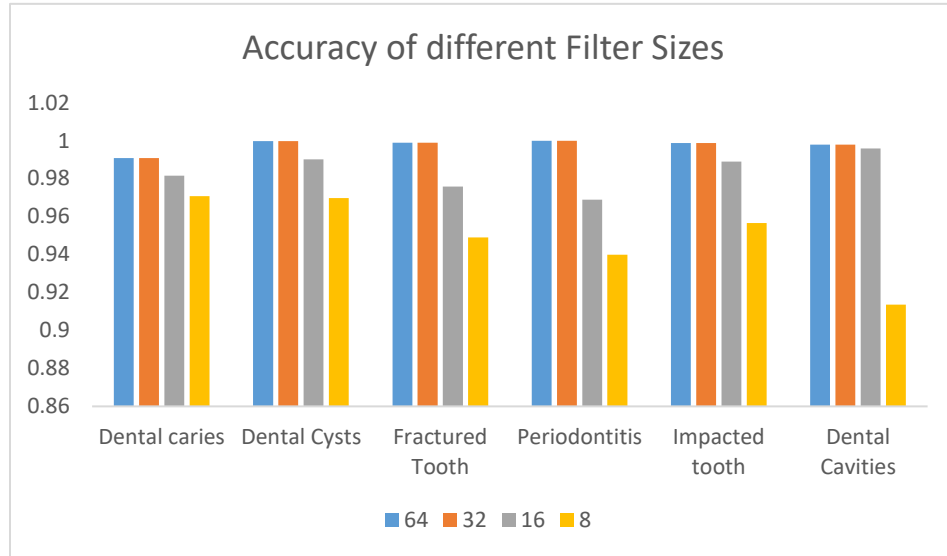| Dental cavities | 0.9979998 | 0.9979998 | 0.9959534 | 0.9135799 |
|---|---|---|---|---|



**Figure 17 Graphical representation of filter sizes outcome**

Error rates were also recorded and like accuracy rate, filter sizes 32 and 64 were the same and least error rates while filter size 8 had the highest error rate as depicted in table 4 and Figure 18

*Table 4 Error rates of different filter sizes*

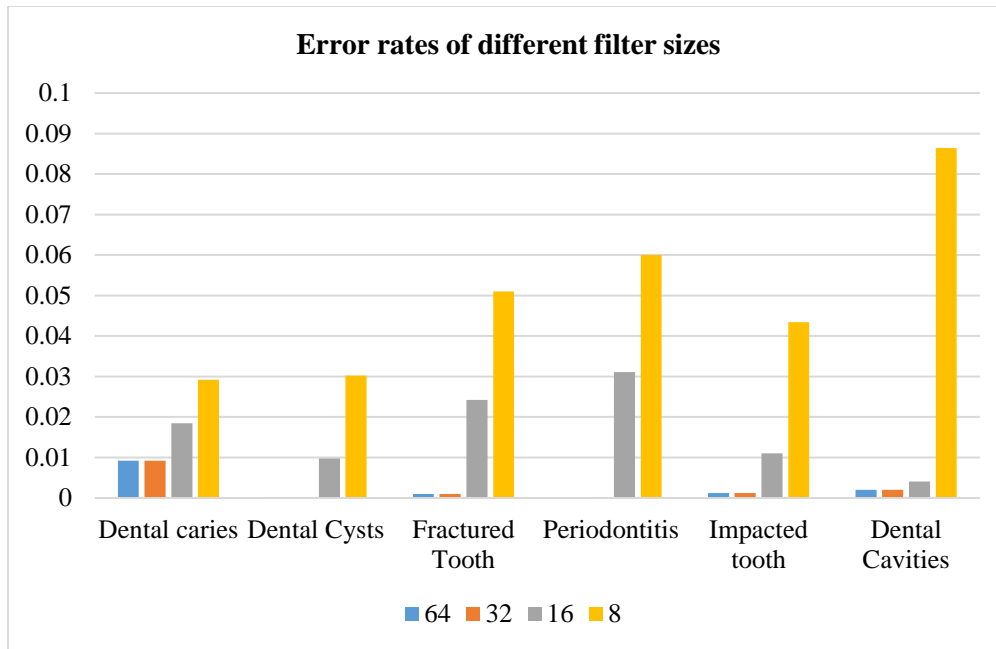| Diseases | Filter Sizes | | | |
|---|---|---|---|---|
| | 64 | 32 | 16 | 8 |
| Dental caries | 0.0092015 | 0.0092016 | 0.0184535 | 0.0292016 |
| Dental Cysts | 0.0002346 | 0.0002346 | 0.0097644 | 0.0302346 |
| Fractured Tooth | 0.0010316 | 0.0010316 | 0.0242062 | 0.0510316 |
| Periodontitis | 3.12E-05 | 3.13E-05 | 0.0311057 | 0.06002321 |
| Impacted tooth | 0.0012006 | 0.0012006 | 0.0110001 | 0.04342097 |
| Dental Cavities | 0.0020002 | 0.0020002 | 0.0040466 | 0.0864201 |

**Figure 18 Graphical representation of error rates of different filter sizes**

## 6.2 Training Time Difference between MM-NN and CNN

Table 5 illustrates a comparative difference between MM-NN and CNN training times in second. From the table, the MM-NN model and CNN model subjected under the same training data size shows an obvious difference time-wise. Table 5 is represented graphically in Figure 19

| Diseases | Filter Sizes | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 64 MM | 64 CNN | 32 MM | 32 CNN | 16 MM | 16 CNN | 8 MM | 8 CNN |
| Dental caries. | 274 | 337 | 263 | 218 | 250 | 298 | 239 | 287 |
| Dental Cysts. | 231 | 284 | 227 | 275 | 211 | 251 | 190 | 228 |
| Fractured tooth. | 227 | 279 | 218 | 264 | 200 | 238 | 188 | 226 |
| Periodontitis. | 259 | 319 | 248 | 300 | 231 | 275 | 213 | 256 |
| Impacted tooth. | 240 | 295 | 237 | 287 | 223 | 265 | 204 | 245 |
| Dental Cavities. | 267 | 328 | 250 | 303 | 243 | 289 | 228 | 274 |

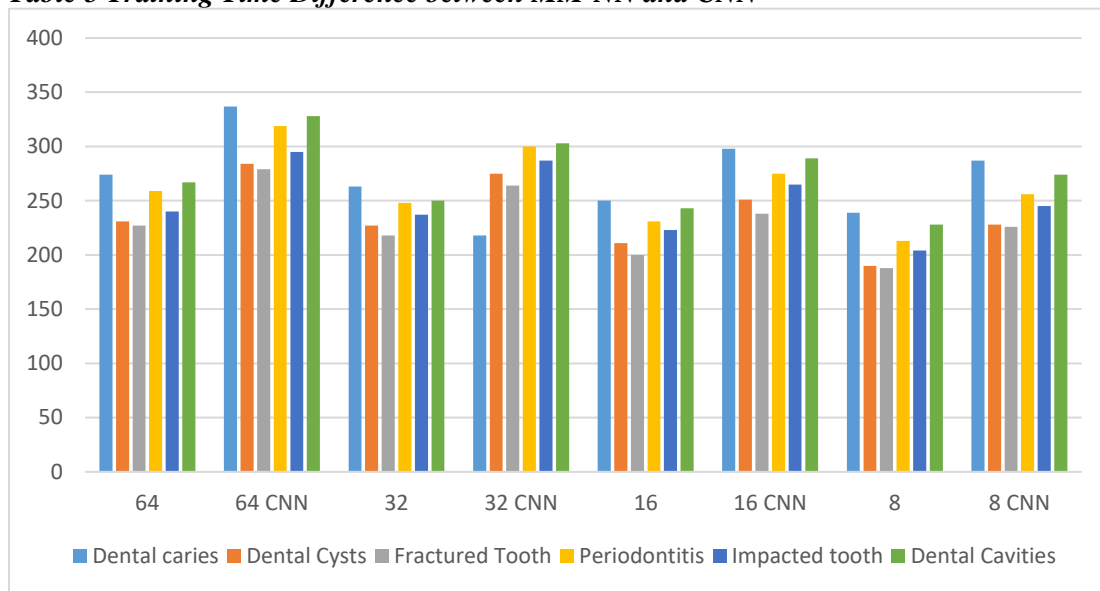*Table 5 Training Time Difference between MM-NN and CNN*



**Figure 19 Graphical representation of table 5**

## 7 Conclusion

This research presented a framework through the development of a computer system with artificial intelligence to handle dental medical diagnostic processes through periapical radiographs while ensuring high precision and accuracy.

The research methodology engaged in this work was an experimental method. Secondary data was used to implement the system after which random data from hospitals was used to test the efficiency of the model.

The findings of this research revealed that;

a) MM operations can be conveniently integrated into the convolution layer of a CNN algorithm to effectively extract image features.

b) The learning of the model was faster with the MM integrated convolution layer than the conventional CNN convolution layer.

c) In terms of the accuracy, precision, and recall metrics used, the MM-CNN model gave higher values which makes it better as demonstrated in accuracy, precision and recall charts in figure 10, figure 11, and figure 112.

d) MM-CNN recorded less errors.

e) In the face of lack of big dataset, MM-CNN algorithm can be implemented with available dataset.

f) The research findings also revealed that digital dental images are not available in the quantity needed to develop a CNN based algorithm as dental departments/clinic lack digital imaging equipment

## References

[1]Grace Tam-Nurseman, Philip Achimugu, Oluwatolani Achimugu, Hilary Kelechi Anabi, Sseggujja Husssein. "Expert System for the Diagnosis and Prognosis of Common Dental Diseases Using Bayes Network." *Journal of Biomedical Science and Engineering, Vol.14 No.11,* 14, no. 11 (November 2021): 361-370.

[2] Ritter, G.X, and Sussner, Peter. "An introduction to morphological neural networks." *Proceedings - International Conference on Pattern Recognition.* Vienna, Austria: IEEE, 1996/09/25. 709-717.

[3] Marcin Iwanowski, Sławomir Skoneczny, and Jarosław Szostakowski. "Image features extraction using mathematical morphology." 1997.

[4]Deep, Paul. "*Screening for Common Oral Diseases.*" ( **J Can Dent Assoc 2000; 66:298-9**) 2000.[Google Scholar]

[5] Mulrenan, Ciara, Kawal Rhode, and Barbara Malene Fischer. "*A Literature Review on the Use of Artificial Intelligence for the Diagnosis of COVID-19 on CT and Chest X-ray .*" **Diagnostics**, March 2022: 1-22.

[6] Rogers, W, B.Ryack, & G.Moeller. "*Computer-aided medical diagnosis: Literature review.*" **International Journal of Bio-Medical Computing** Volume 10, no. 4 August 1979: 267-289.

[7] Kumar, Yogesh, Koul Apeksha, Singla Ruchi, & Ijaz Muhammad Fazal. "*Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda.*" **Journal of Ambient Intelligence and Humanized Computing**, January 2022.

[8] Ayyar, Tejas Mohan. "LeNet." *A practical experiment for comparing LeNet, AlexNet, VGG and ResNet models with their advantages and disadvantages.* November 6, 2020. https://tejasmohanayyar.medium.com/a-practical-experiment-for-comparing-lenet-alexnet-vgg-and-resnet-models-with-their-advantages-d932fb7c7d17

[9] Kumar, Yogesh, Koul Apeksha, Singla Ruchi, & Ijaz Muhammad Fazal. "*Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework*

*and future research agenda.*" **Journal of Ambient Intelligence and Humanized Computing**, January 2022.

[10]Lakhani Paras, Gray, Daniel L., Pett, Carl R.,Nagy, Paul & Shih, George. *"Hello World Deep Learning in Medical Imaging*." **Journal of Digital Imaging** 31, no. 3 June 2018: 283-289.

[11] Howard G. Andrew, Menglong Zhu, Bo Chen , Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, & Hartwig Adam. "*MobileNets*: *Efficient Convolutional* Neural *Networks for Mobile Vision*." **arXiv**, 2017: 1-9.

[12] Keiller Nogueira, Jocelyn Chanussot, Mauro Dalla Mura & Jefersson A. Dos Santos. . "*An Introduction to Deep Morphological Networks.* ." **IEEE Access 9** , 2021: 114308-114324.

[13]Hongping Wu, Yuling Liu & Jingwen Wang. "*Review of Text Classification Methods on Deep Learning.*" **Computers, Materials and Continua (CMC)** 63, no. 3 February 2020: 1309-1321.

[14]Arunnehru, J.,Chamundeeswari G. & Bharathi S. Prasanna. "*Human Action Recognition using 3D Convolutional Neural Networks with 3D Motion Cuboids in Surveillance Videos.*" **Procedia Computer Science (Elsevier)** 133 January 2018: 471-477.

[15]Bevilacqua, Antonio, MacDonald Kyle, Rangarej Aamina, Widjaya Venessa, Caulfield Brian & Kechadi Tahar. "*Human Activity Recognition with Convolutional Neural Networks.*" **European Conference**, **ECML PKDD 2018**, *Dublin, Ireland.* Dublin, Ireland: RearchGate, 2018. 1-14.

[16]Ankita, Rani Shalli, Babbar Himanshi, Coleman Sonya, Singh Aman & Aljahdali Hani Moaiteq. "*An Efficient and Lightweight Deep Learning Model for Human Activity Recognition Using Smartphones*." **Sensors** 21, no. 11 (June 2021): 1-17.

[17]Ming Zeng, Le T. Nguyen, Bo Yu, Ole J. Mengshoel, Jiang Zhu, Pang Wu, & Joy Zhang. "*Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors.*" *2014 6th* **International Conference on Mobil Computing, Application and Services** *(MobiCase).* **Austin, TX, USA: IEEE**, 2014. 197-205.

[18]Krizhevsky, Alex, Sutskever Ilya & Hinton Geoffrey E. "ImageNet classification with deep convolutional neural networks." **Communications of the ACM** 60 (2012): 84-90.

[19]Raju, Jincy & Modi, Chintan. "*A Proposed Feature Extraction Technique for Dental X-Ray Images Based on Multiple Features."* **2011 International Conference on Communication Systems and Network Technologies**. **IEEE Xplore,** 2011. 545-549.

[20]Krithigaa, R.Rani & Lakshmia C. *A Survey:* Segmentation in Dental X-ray Images for s Diagnosis of Dental Caries. Vol. 09. **International Science Press**, 2016.

[21]Na`am, Jufriadif, Harlan, Johan & Wibowo, Eri Prasetyo. "*Image Processing of Panoramic Dental X-Ray for    Identifying Proximal Caries*." **TELKOMNIKA** 15, no. 2 June 2017: 702-708.

[22] Dawson, Catherine. "*Advantages and Disadvantages of Open and Close questiions*." In *Practical Research Methods: A User-friendly Guide to Mastering Research Techniques and Projects*, by Catherine Dawson, 88. Oxford, United Kingdom: **How To Books,** 2002.

[23]Colgate-Palmolive.  "*Types    of    X-rays*."    *Colgate-Palmolive.*    February    4,    2022. https://www.colgate.com/en-us/oral-health/x-rays/types-of-x-rays#

[24]Rubin, Herbert J., and Irene Rubin. *Qualitative Interviewing: The Art of Hearing Data.* Third Edition. Thousand Oaks,, California: **Sage Publications,** 2012.

[25]Sharon M. Ravitch, and Sharon M. Ravitch. *Qualitative Research:Bridging the Conceptual, Theoretical, and Methodological.* 2nd. **Thousand Oaks,** California: SAGE, 2016.

[26]Jason Brownslee on January 9, 2019. "*A Gentle Introduction to the Rectified Linear Unit (ReLU)." Machine Learning Mastery. January 9, 2019. https://machinelearningmastery.com*