How to Cite:
Khan, F., Raza, M., Ullah, N., Khisrow, K., Ali, I., & Jamil, A. (2023). Dental caries and non caries detection using deep learning. *International Journal of Health Sciences*, *7*(S1), 790–804. https://doi.org/10.53730/ijhs.v7nS1.14277

Dental caries and non caries detection using deep learning

Dr. Farzeen Khan
Assistant professor, Department of Community and preventive dentistry, Peshawar Dental College,Peshawar, Pakistan
Corresponding author email: farzeen.khan7@gmail.com

Mr. Muhammad Raza
Department of ICS/IT, The University of Agriculture Peshawar, Pakistan

Mr. Naveed Ullah
Department of ICS/IT, The University of Agriculture Peshawar, Pakistan

Mr. Khalid Khisrow
Department of ICS/IT, The University of Agriculture Peshawar, Pakistan

Mr. Iqtidar Ali
Lecturer, Department of ICS/IT, The University of Agriculture Peshawar, Pakistan

Asif Jamil
Department of Rehabilitation and Health Sciences, Iqra University Islamabad Campus.

***Abstract*---**Cavities are the most common indication of dental caries, a contagious condition that leads to the deterioration of the tooth's structure. Dental caries has been identified as one of the most common oral health issue. This research has been conducted to identify them early, owing to the discomfort and high expense of treatment. Artificial intelligence has been utilized in recent years to create models which can forecast the risk of dental caries due to restrictions in medical research in oral healthcare, such as the high costs and lengthy requirements. Data for our study were collected from Khyber College of Dentistry and Hospital. On this data, a number of Deep Learning algorithms were implemented, and their performances were evaluated using recall, precision, F1-score, and accuracy. In comparison to CNN, LeNet and AlexNet deep learning techniques, VGG16 has the best performance, scoring accuracy of 98.99%, F1-score of 0.96% with precision of 0.95%, and a recall of 0.97%. This suggested paper demonstrated that DL is strongly advised

Manuscript submitted: 27 Jan 2023, Manuscript revised: 09 March 2023, Accepted for publication: 18 April 2023

for dental professionals to use in helping them make decisions for the early diagnosis and treatment of dental caries.

***Keywords*---**Cavities, dental caries, deep learning.

Introduction

Dental caries is a prevalent and persistent oral infection that impacts individuals of all ages globally. Given that there are over 1.8 billion caries-related sickness cases reported each year [1-3], a reduction in oral health will eventually tends to enhance in the number of cases of tooth loss being documented. The prevalence of dental cavities continuously increases from birth through age. Periodontal disease rises after puberty and has an epidemiologic feature that eventually remains through adolescence [4]. Because untreated children caries can develop into permanent tooth caries and oral diseases, dental caries must be treated in infancy. As a result, it is acknowledged as a public health concern rather than a personal issue because it significant ly affects the overall quality of life [5-7]. Dental caries can be treated via restorative and preventive therapy if it is discovered early. Although, long-term care such as thorough, root canal therapy, prosthetic treatment and tooth cleaning are necessary after tooth extraction if long-term neglect affects the dentin or pulp [8]. Everyone should be conscious of their dental health and schedule routine checkups with their dentist for evaluation and treatment. Unfortunately, many people neglect preventative care or inspection before noticing symptoms because of financial limitations or a lack of awareness of the importance of oral health [9]. The decayed-missing-filled dental (DMFT) index, which is frequently regarded as the main population-based measure of caries-related problems internationally (decayed, missing, or filled), is used to determine the number of permanent teeth afflicted by caries. Eating habits of people, economic circumstances, use of dental care and improper oral health management behaviors are all known to have an impact on oral health [10]. Thus, identifying DMFT and its related traits as well as predicting irreversible dental damage due to dental caries might be a crucial starting point for creating a personalized oral prevention strategy.

The development of automated solutions for identifying tooth caries illness and non-caries has recently been made possible b, machine learning (ML) computer vision (CV) technology and advances in artificial intelligence (AI). Without requiring human intervention, these methods can rapidly, effectively, and reliably identify dental disorders. Recently, numerous deep learning architectures have been proposed for dental disease classification. The most well-known technique is the convolutional neural network (CNN). The CNNs are supervised by deep learning models inspired by the biological nervous system and visual system and perform better than other models. Compared to artificial neural networks (ANNs), CNNs require a small number of neurons and multi-layered convolutional layers to learn their function, but training requires a large dataset. The tooth caries and non caries dataset has a total of 3518 images containing 1710 images for images of healthy dentals, 1808 images for images of dental caries disease. It is divided into 80% for training, 10% for validation, and 10% for testing. This research also compares the performance of pre-trained models to that of basic CNN models developed from scratch. With Computer Vision CNN, VGG16, LeNet, and AlexNet, we would use pre-trained models. We created a Convolutional Neural Network (CNN) from scratch in Python using TensorFlow and Keras, the former being an open-source software framework used for machine learning applications such as neural networks and the latter being a high-level neural network library. We worked on Google Colab as a development environment.

Methodologies

A. CNN Model Architecture

The CNN model, which received a score of 99.7%, used hyper parameters that were tested repeatedly until a stable model was created that was neither under fitting nor overfitting. The outcomes of the user study might have been different if the machine learning models had been applied differently or if a different machine learning model had been chosen. It is impossible to say with certainty that changing the model's hyper parameters or by using different model of classification would make any difference in results of user's study. Moreover, the concluded results depicted the opinion of participants which relays on certain factors previously discussed.

.

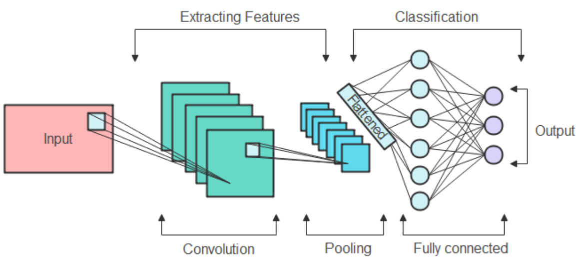


Fig: 1.1 B. VGG16

ConvNets is another term for the artificial neural network type convolutional neural networks. There are three layers in a convolutional neural network: input, output, and hidden. The Convolutional Neural Network (CNN) family, is considered as one of the best vision models of computer, including VGG16 model. Using tiny (3 by 3) convolution filters, this model greatly outperformed setup in regardance of increasing depth. The depth was increased by 16 to 19 weight layers, (or around 138 trainable parameters). VGG16 (an object identification and classification algorithm) has an accuracy rate of 92.7% when it categorize the 1000 images into 1000 individual categories. And considered as an effective method for categorization of photos. A 224x224 RGB fixed input image is supplied into the 16-layer stack of convolution layers, which employs 3x3 filters. The input channels have also been made non-linear by using a 1x1 Conv. filter. In order to maintain the spatial resolution, the stride and padding are maintained to one pixel per filter. For spatial pooling, 5-2x2 max-pooling windows with stride 2 are introduced. Then, three fully connected layers are added to the stack: the first two contains 4096 channels each, the third layer does classification with 1000 channels for 1000 picture classes, and the last layer of the design is a soft-max layer.

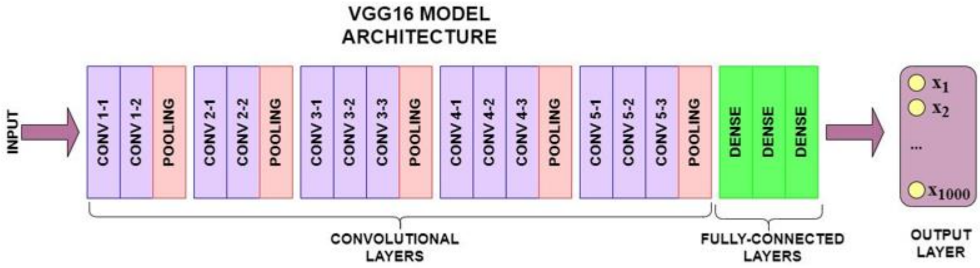


Fig: 1.2

B. LeNet:

The LeNet (CNN) architecture is developed by Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick Haffner in 1998. One of the earliest effective deep learning applications for picture identification was this one. Seven layers make up the LeNet model, including two pooling layers, three fully connected layers and two convolutional layers. A grayscale image, typically measuring 32x32 or 28x28 pixels, serves as the network's input. A convolutional layer with six 5x5 filters makes up the first layer. The second layer, (along with a max pooling layer), lowers the feature maps' spatial dimensions by a factor of two. Following another max pooling layer, the third layer is another convolutional layer with 16 5x5 filters. The 120-node first fully connected layer receives the output of this layer after flattening it into a vector. The last layer (a softmax layer) distribution of probability over the potential classes, and the second fully linked layer comprises 84 nodes. LeNet has been applied to a variety of image recognition problems since it was first developed to recognise handwritten numbers. LeNet was a pioneering model that opened the door for the creation of more advanced and potent deep learning architectures, despite its architecture being relatively simple in comparison to contemporary CNNs.

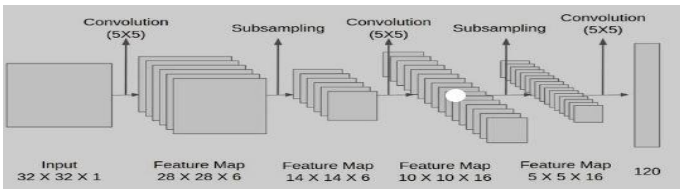


Fig: 1.3

C. Alex Net:

It is a deep convolutional neural network architecture which was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton in 2012. It was the victor of the 2012 ImageNet Large Scale Visual Recognition Competition (ILSVRC), which tested computer vision models' abilities to recognise objects in pictures. AlexNet is made up of three fully connected layers, five max pooling layers, and five convolutional layers. The network receives an RGB image with a 224x224 pixel input. A convolutional layer having 96 filters in the first layer, each measuring 11 by 11, and a stride of 4 pixels. A convolutional layer with 256 5x5 filters and a 1 pixel stride makes up the second layer as well. Convolutional layers with 384, 384, and 256 filters of size 3x3 make up the third, fourth, and fifth layers. Each convolutional layer's output is run via a ReLU activation function. The max pooling layers minimise the feature maps' spatial dimensionality after the convolutional layers. After being flattened into a vector and passing through three fully connected layers with 4096, 4096, and 1000 nodes, the output of the final max pooling layer is finally output. The probability distribution over the potential classes is produced by the softmax layer (the last layer). The development of AlexNet, which showed the effectiveness of deep CNN for image classification problems, represented a major advance in deep learning. Its performance on the ILSVRC served as a catalyst for deep learning's quick development in computer vision and other fields.

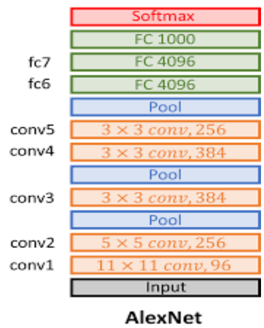


Fig: 1.4 Classification

A. Forward Propagation

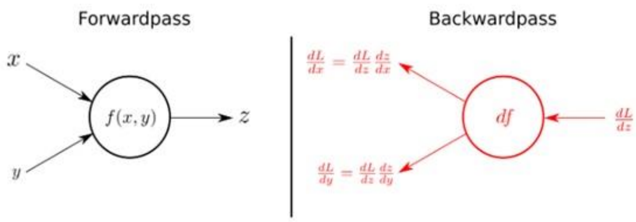
In this, the network is fed with input data which moves forwards through many layers. After inserting, the data is accepted and processed by each hidden layer and then transfer it to another layer vice versa.

B. Backward Propagation

The foundation of neural network training is back-propagation. In this procedure, a neural network's weights are modified based on the mistake rate, also known as loss, that was recorded in the iteration before it. Proper weight adjustment ensures lower mistake rates, expanding the model's range of applications and enhancing its dependability.

Here are the overall steps:

- Data travels across the network at the forward propagate stage to obtain outputs.
- Using the loss function, the overall error is calculated.
- The backward propagation algorithm is then utilized to calculate the gradient of the loss function depending on each bias and weight
- Gradient descent is used in the last step to update the biases and weights for each and every layer.
- The above steps are repeated to minimize the neural network's total error.



Forward difference formula

$$f'(x_i) \approx \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$$

Backward difference formula

$$f'(x_i) \approx \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}$$

Results

1. CNN

A total of 3518 photos make up the dataset, of which 70% were used to train the model and 30% to analyse it. It utilised the ReLU activation function. The Runed model has 20 maximum no. of epochs. Performance metrics like F-measured, recall, accuracy, precision, and MSE losses were utilized to assess the model's performance.

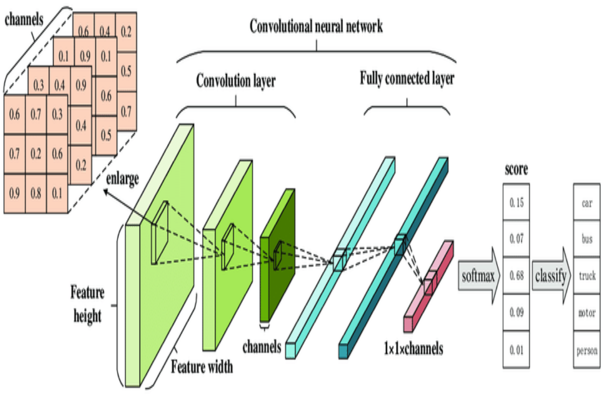


Fig: 3.1

In the initial step, the directory contained all of the photos. Different layers were present in the CNN model. The sigmoid activation function was connected to the topmost connection layer There were several layers in the CNN model. The final layer of connection was related with the sigmoid activation function. The CNN model that distinguishes between teeth with caries and teeth without caries was trained and validated in the final session.

Figure 3.2 showed the model training and testing accuracy while training and testing accuracy is started from 51 and the model are improving himself and at last epoch the model give us 87 accuracy. We run total of 20 epochs. While Validation accuracy are started from 48 slowly model improved and at last it goes to 87 accuracy.

While in figure 3.3 showed model training and validation loss is started from 85 and the model started learning and training loss dropped to 31. Validation lost is started from 69 and dropped to 31.

Fig: 3.2

Fig: 3.3

Table 1: Performance Parameter for CNN Model

	Precision	Recall	F1_score
Non Caries	0.86	0.89	0.88
Caries	0.88	0.85	0.86

Table 1 showed the precision, recall, and F1 score are performance metrics used to assess the effectiveness of classification algorithms such as Convolutional Neural Network (CNN) models.

A ratio of true positives (properly predicted positives) to total predicted positives is called precision. Precision for non-caries instances in the current context is 0.86, meaning that 86% of all projected non-caries cases were truly non-caries. Similarly, the precision for caries cases is 0.88, meaning that 88% of all projected caries cases actually had caries.

Recall (the ratio of true positives to all other positives) for non-caries cases in this instance is 0.89, meaning that 89% of all non-caries cases were properly recognised by the model as such. According to the recall for caries instances, which is 0.85, 85% of the real caries cases were properly diagnosed as such by the model.

F1 score (the harmonic mean of precision and recall) provides a balanced perspective of precision and recall and is a valuable metric whenever the dataset is imbalanced. In this situation, the F1 score for non-caries patients is 0.88, while for caries cases, it is 0.86.

These performance features are significant because they provide insight into the classification model's advantages and disadvantages and can be used to evaluate other models or various iterations of the same model. However, before making decisions about the model's performance, it is critical to evaluate other elements i.e., dataset, evaluation technique, and real-world implications.

2. VGG 16
A total of 3518 photos make up the dataset, of which 70% were used to train the model and 30% to analyse it. It utilised the ReLU activation function. The maximum number of epochs for the Runed model was 50. Performance metrics like Accuracy, Recall, Precision, F1-score, and MSE losses were utilized to assess the model's performance.

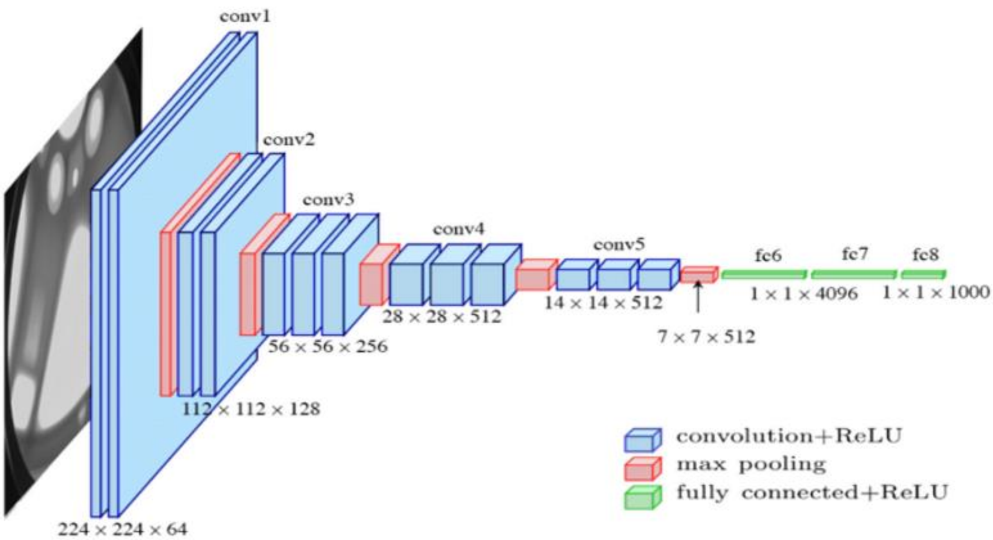


Fig: 3.4

In the first step, all the images were found in the directory. The VGG-16 model consists of different pre-defined layers. sigmoid activation function is connected by the final layer. In the last step, validate and train the CNN modal is trained and validated that tends to distinguish an infected caries from a non caries dental.

Figure 3.5 depicts the testing accuracy and model training. While, training and testing accuracy is started from 87 and the model are improving himself and at last epoch the model give us 99 accuracy. We run total of 50 epochs. While Validation accuracy are started from 87.3 slowly model improved and it goes to 98.99 accuracy at last epoch.

While in figure 3.6 showed model training and validation loss is started from 36 and the model started learning and training loss dropped to 4.3. Validation lost is started from 23 and dropped to 6.2.

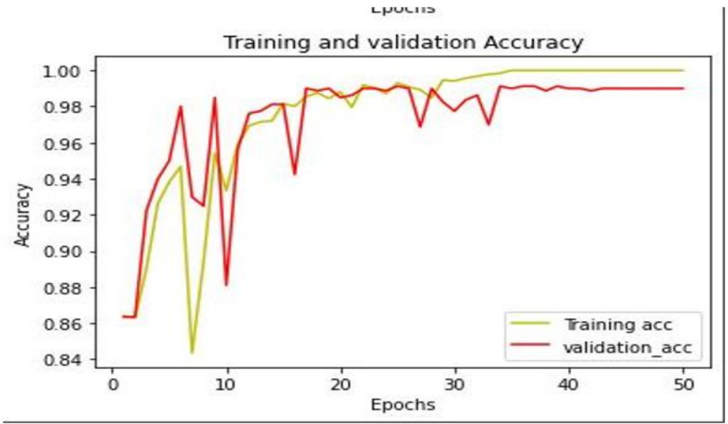


Fig: 3.5

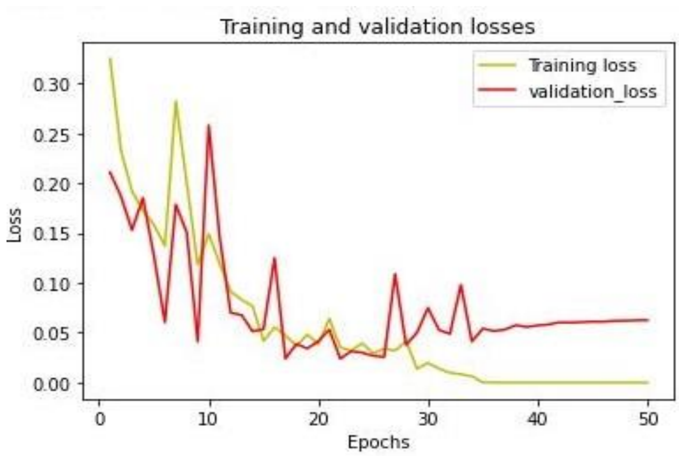


Fig: 3.6

Table 2: Model Performance Parameter for the VGG16

	Precision	Recall	F1_score
Non Caries	0.95	0.97	0.96
Caries	1.00	0.99	0.99

Table 2 showed the precision (defined ratio of true positives to all predicted positives), Recall (percentage of true positives to the total number of actual positives), and the F1 score (the harmonic mean of recall and precision). These performance metrics are used to evaluate how well classification models like the VGG16 perform.

In this instance, it appears that the VGG16 model has good precision, recall, and F1 scores for both cases of caries and non-caries. for non caries Precision are 0.95, recall are 0.97, and F1 score are 0.96. For caries Precision are 1.00, recall are 0.99, and F1 score are 0.99. These results indicate that the VGG16 model performs quite well at recognising both caries and non-caries instances.

3. LeNet

3518 photos total are included in the dataset, of which 70% were used to train the model and 30% to analyse it. It was activated using ReLU. The model for Runes has a maximum of 50 epochs. Accuracy, Precision, Recall, F1-score, and MSE losses were some of the metrics that tends to analyze the activity of models.

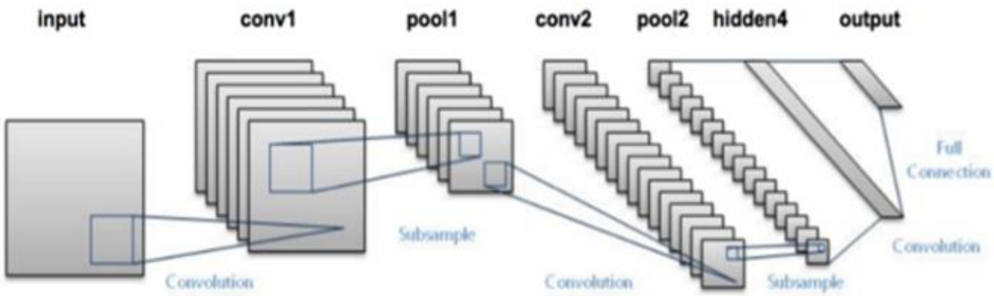


Fig: 3.7

The images were uniformly resized to dimensions of 150 x 150 before being utilized for both training and testing. Following the completion of training, the model accurately classified the dataset into two categories: Caries class and Non-Caries class.

Figure 3.7 described the testing accuracy and model training. While, training and testing accuracy is started from 72 and the model are improving himself and at last epoch the model give us 98 accuracy. We run total of 50 epochs. While Validation accuracy are started from 86 and then at 26 epoch accuracy are dropeed to 80 and then slowly model improved and it goes to 98.37 accuracy at last epoch.

While in figure 3.8 showed model training and validation loss is started from 68 and the model started learning and training loss dropped to 7.3. Validation lost is started from 24 and at 26 epochs it goes to 80 and the model improved himself and dropped to 6.2.

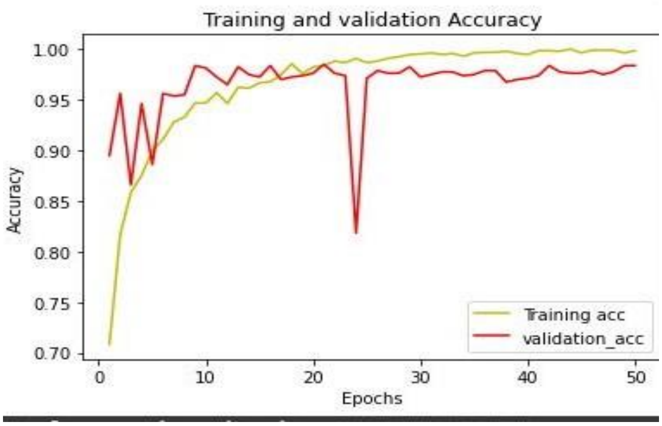


Fig: 3.8

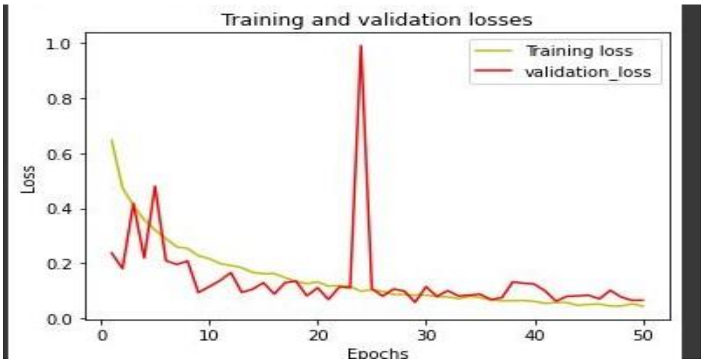


Fig: 3.9

Table 3: Performance Parameter for the LeNet Model

	Precision	Recall	F1_score
Non Caries	0.92	0.96	0.94
Caries	0.99	0.99	0.99

Table 3 showed the performance metrics like precision, recall, and F1 score are used for analysing how effectively classification models like the LeNet Model are performing. The LeNet Model appears to have excellent precision, recall, and F1 scores in this instance for both caries and non-caries instances. For non-caries cases, the precision value is 0.92, recall value is 0.96, and F1 score value is 0.94. For caries cases, the precision is 0.99, recall is 0.99, and F1 score value is 0.99.

These results indicate that the LeNet Model performs effectively in identifying both caries and non-caries images.

4. AlexNet

A total of 3518 photos make up the dataset, of which 70% were used to train the model and 30% to analyse it. It utilised the ReLU activation function. The maximum number of epochs for the Runed model was 50. Performance metrics like Accuracy, Precision, Recall, F1-score, and MSE losses were utilized to assess this model's activity.

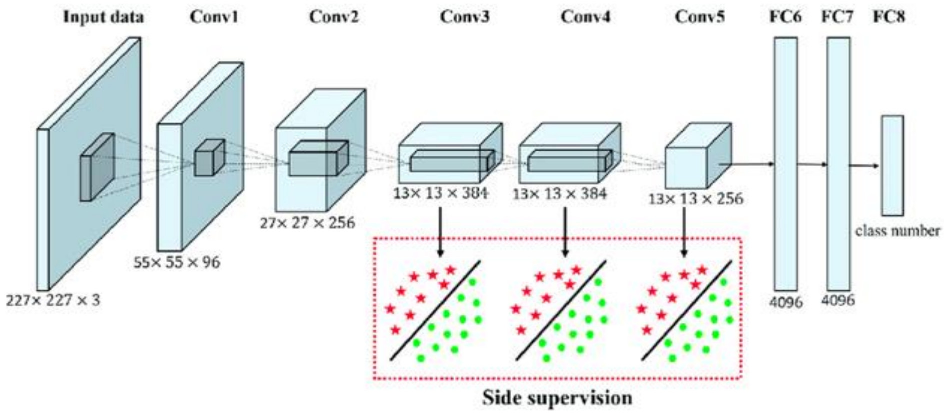


Fig 3.10

Figure 3.10 shows the working process of training and testing. Firstly, all the infected images are placed in caries folder while healthy and normal images are placed in non caries folder. The Alex Net model has many different pre-defined layers. The final layer is connected by activation function..lastl, the CNN model is valideated and trained which tends to distinguish the caries from the non caries cases.

In Figure 3.11 the model training and testing accuracy were shown. While, training and testing accuracy is started from 93 and the model are improving himself and at last epoch the model give us 98.3 accuracy. We run total of 50 epochs. While Validation accuracy are started from 64 and then slowly model improved and it goes to 97.61 accuracy at last epoch.

While in figure 3.12 showed model training and validation loss is started from 19 and the model started learning and training loss dropped to 7.2. Validation lost is started from 92 and then model improved himself and dropped to 6.1.

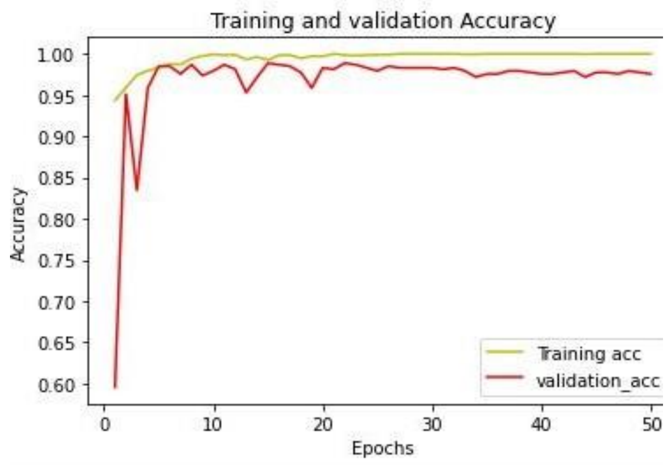


Fig: 3.11

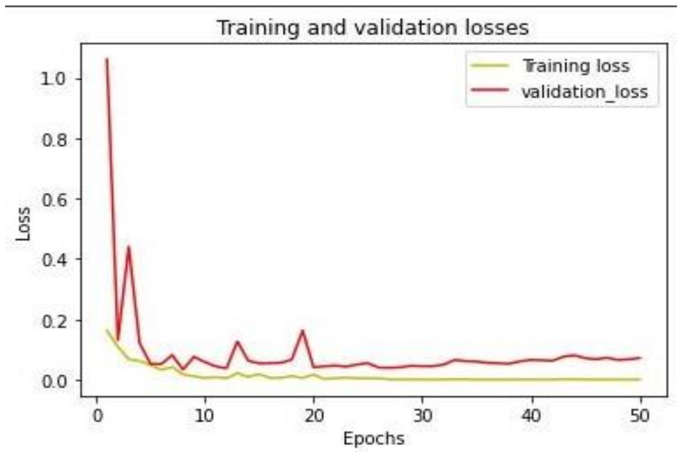


Fig 3.12

Table 4: Performance Parameter for the AlexNet Model

	Precision	Recall	F1_score
Non Caries	0.92	0.91	0.91
Caries	0.99	0.99	0.99

Table 4 showed the metrics performance such as recall, F1 score and precision which are used for estimating how effectively classification models like the AlexNet Model are carrying out. The AlexNet Model appears to have excellent precision, recall, and F1 scores in this instance for both caries and non-caries instances. For non-caries cases, the value of precision is 0.92, recall is 0.91, and F1 score value is 0.91. For caries cases, the precision value is 0.99, recall value is 0.99, and F1 scoring value is 0.99. These results indicate that the AlexNet Model performs excellently in detecting both caries and non-caries images.

Conclusion

By implementing four deep learning models, Convolutional Neural Networks (CNN), VGG-16, LeNet, And AlexNet on the disease detection of dental caries and non caries from the dataset and also evaluating the aforementioned model using the following Metrix: Accuracy the study shows that VGG16 model performs better than the CNN, LeNet And AlexNet models in the dental caries disease and non caries detection The single evaluation Metrix outperforms the VGG-16, LeNet, and AlexNet models. This study is able to obtain input from farmers on whether they trust the aforementioned AI models by conducting a user study. According to the findings of the user survey, farmers believe that AI forecasts and explanations are poor and, as a result, they do not trust the deployed tools for detecting tooth caries and non-caries. But, through extra feedback, the farmers indicate areas that could potentially assist enhance and trust the AI models.

Future Recommendation

1. Calling the model: While the current study was conducted using data from a single institution, future work could involve collecting data from multiple sources to create a larger dataset for model training and testing. This would help to evaluate the model's performance in more diverse populations and settings.
2. Exploring different deep learning architectures: Although the VGG16 model performed well in this study, future work could involve exploring other deep learning architectures such as ResNet, Inception, or DenseNet to see if they can achieve even better performance.
3. Investigating additional risk factors: The current study used a limited number of some factors i.e., age, gender, and oral cleanliness practices. Future work could involve exploring other risk factors such as diet, smoking, or genetics to see if they can improve the model's predictive power.
4. Development of a user-friendly application: Once the model has been further developed and validated, future work could involve developing a user-friendly application that can be utilized by dental professionals which tends to assist them in making decisions as well as identification and immediate dental therapy. This could potentially help to reduce the time and cost associated with traditional diagnostic methods.

References

[1] Dye, Bruce A. Oral health disparities as determined by selected healthy people 2020 oral health objectives for the United States, 2009-2010. No. 100. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics, 2012.
[2] Dye, Bruce A., et al. "Trends in oral health status; United States, 1988-1994 and 1999-2004." (2007).
[3] Vos, Theo, et al. "Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016." The Lancet 390.10100 (2017): 1211-1259.

[4] Casamassimo, Paul S., et al. "Beyond the dmft: the human and economic cost of early childhood caries." The Journal of the American Dental Association 140.6 (2009): 650-657.

[5] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT press, 2016.

[6] Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." Neural networks 61 (2015): 85-117.

[7] Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." nature 529.7587 (2016): 484-489.

[8] Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." Neural networks 61 (2015): 85-117.

[9] Chen, X., Guo, J., Ye, J., Zhang, M., & Liang, Y. (2022). Detection of Proximal Caries Lesions on Bitewing Radiographs Using Deep Learning Method. Caries Research, 56(5–6), 455–463. https://doi.org/10.1159/000527418

[10] Reyes, L. T., Knorst, J. K., Ortiz, F. R., & Ardenghi, T. M. (2022). Machine Learning in the Diagnosis and Prognostic Prediction of Dental Caries: A Systematic Review. *Caries Research*, *56*(3), 161–170. https://doi.org/10.1159/000524167

[11] Lussi, A., & Francescut, P. (2003). Performance of Conventional and New Methods for the Detection of Occlusal Caries in Deciduous Teeth. *Caries Research*, *37*(1), 2–7. https://doi.org/10.1159/000068226

[12] Malhotra, Y. (2018). AI, Machine Learning & Deep Learning Risk Management & Controls: Beyond Deep Learning and Generative Adversarial Networks: Model Risk Management in AI, Machine Learning & Deep Learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3193693

[13] Themes, U., & D. (2020, January 7). *Personalized Dental Caries Management in Children*. Personalized Dental Caries Management in Children | Pocket Dentistry. https://pocketdentistry.com/personalized-dental-caries-management-in-children/

[14] Kang, I. A., Ngnamsie Njimbouom, S., Lee, K. O., & Kim, J. D. (2022, March 16). DCP: Prediction of Dental Caries Using Machine Learning in Personalized Medicine. *Applied Sciences*, *12*(6), 3043. https://doi.org/10.3390/app12063043