A Hybrid CNN–GNN Architecture for Automated Dental Disease Classification from X-ray Images

1st Akshai Maran M
Student
Department of Information Technology
St Joseph's Institute of Technology, Chennai, India
maranakshai3@gmail.com

2nd Gokulraj K
Student
Department of Information Technology
St Joseph's Institute of Technology, Chennai, India
kgrtamizhan@gmail.com

3rd Josepha Menandas J
Assistant Professor
Department of Information Technology
St Joseph's Institute of Technology, Chennai, India
josepha82@gmail.com

Abstract - Dental radiographs are relevant in diagnosis of oral conditions which include caries, periodontal bone loss and periapical infections, but manual interpretation is time-consuming and prone to inter-observer variations. In this paper, the author suggests a hybrid Convolutional Neural Network-Graph Neural Network (CNN-GNN) architecture to be used in automated dental disease classification on the basis of X-ray imaging. The CNN learns discriminative tooth-level visual representation whereas the GNN learns the anatomical relationships between nearby teeth in the form of a graph. The given approach is more accurate in classifications and it minimizes false predictions by combining the local visual information with the global structural context. However, the experimental evidence on public datasets of dental X-rays is established to represent better precision, recall and F1-score than the traditional CNN-based alternatives, indicating the suitability of the graph-based learning in computer-aided medical diagnosis of dental cases.

Keywords: Dental X-ray, CNN, GNN, Deep Learning, Medical Imaging, Disease Classification, Graph Learning, Computer-Aided Diagnosis.

I. INTRODUCTION

Dental caries, periodontal bone loss, and periapical infections are the most common oral health issues, and dental X-ray imaging is critical in diagnosing oral health issues because it provides elaborate visualization of structures of teeth and bone around them. But the dental radiographs also have a problem with manual analysis of the radiographs because of the low contrast, noise, overlapping of anatomy and image quality variation. Moreover, dental experts mostly rely on experience in the diagnosis, which may result in subjectivity and unreliability. These drawbacks make automated and smart diagnostic systems that can deliver reliable and precise diagnosis of dental diseases based on X-ray images a necessity. The latest developments in AI, specifically deep learning have demonstrated a lot of success in the analysis of medical images. Convolutional Neural Network has been shown to be extremely efficient with respect to extracting complex information on the eye like the existence of a cavity, bone loss and structure abnormalities even without the special feature engineering. The traditional CNN models, however, process the teeth separately and do not capture anatomical features between the adjacent teeth. To solve this shortcoming, this paper suggests a hybrid deep learning model, which is based on CNN and Graph Neural Networks. CNN facilitates the efficient extraction of tooth level features, whereas GNN models depict inter-tooth connectivity via the graph, hence achieve better classification efficiency and dependable and contextual milieu to a computer-aided clinical diagnosis and future medical use of the data.

II. LITERATURE SURVEY

The focus on radiographic imaging has been a significant area of dental disease diagnosis research because of the increased demand that requires accuracy and the need to detect the disease early. The traditional dental diagnosis makes use of manual interpretation of X-rays that is subjective and consumes time. The emerging possibilities of an artificial intelligence system have led to the popularity of using deep learning methods in the analysis of medical images, including dental imaging. According to surveys like De Berardinis et al., [6], convolutional neural networks have great potential to identify dental caries, bone loss and structural abnormalities. Nevertheless, there are issues that include low contrast, superimposition of structure, and imaging condition change which still hamper performance and facilitate more sophisticated learning strategies.

According to the recent literature, automated healthcare dental image analysis should not be limited to the pixel level of classification to become clinically relevant. Zhang et al. [9] found out that dental professionals have better diagnostic consistency and accuracy when using clinically applicable AI systems. Likewise, Lee et al. [2] demonstrated that deep learning models are superior to the conventional ones in detecting dental caries. These advances notwithstanding, most current systems consider the individual teeth as analyzed other than taking in consideration the anatomical relationships. Dental diseases are usually spatial dependent and therefore a failure to consider their spatial relationships may lead to misclassification as it is shown that models including the extraction of features and the modeling of structure are required.

One of the first CNN-based dental caries detection systems modeling on the use of X-ray image was proposed by Abdelhafiz et al. [1]. They found that through their work, CNNs can be able to automatically learn discriminative features related to caries without the need of manually engineering features and have a higher accuracy than the traditional machine learning methods. Nonetheless, this model was rather regional as it did not assume relationships between adjacent teeth. This weakness limits its capability to capture disease transmission patterns at complex dental cases, although the study made CNNs a solid base towards automated dental disease identification.

Lee et al. [2] also discussed CNN based methods of detecting dental caries and reported better accuracy with less human error. Their study however showed performance to be influenced by differences in image quality and dental anatomy. Likewise, Jeon et al. [8] devoted their attention to the detection of periodontal bone loss with the help of CNNs and obtained encouraging findings. Still, they have noted that the effect of neighboring teeth was not explicitly represented. These results show that CNNs are successful in feature extraction, but not at intra-tooth dependencies which are needed to make a correct diagnosis of the tooth.

Segmentation is an effective tool of dental image analysis because it allows easy isolation of tooth parts. Ronneberger et al. [7] proposed the U-Net architecture that has become a paradigm in the field of biomedical image segmentation because of its encoder-decoder design and skip connections. Oktay et al. [3] boosted the performance of segmentation with Attention U-Net that enabled networks to concentrate on the pertinent areas. In Liu et al. [13], the dental panoramic image was segmented using deep learning and the tooth boundary was better detected. Although segmentation improves the quality of features, this approach cannot learn relational reasoning between teeth on its own.

Outside dental imaging, general progress in deep learning has had major impacts on the processing of medical images. Krizhevsky et al. [11] proved the usefulness of deep CNNs in ImageNet classification making CNNs a hegemony of visual tasks. To optimize model scaling to achieve a better performance with a fewer number of parameters, Tan and Le [12] came up with EfficientNet. Wang et al. [10] proposed non-local neural networks, which apply long-range dependencies. Although they are known to speed feature learning, such architectures accept only grid-based representation, and explicit adjacency relationships are not represented.

Graph-based learning techniques have become useful techniques to overcome the grid-based constraints. Graph Convolutional Networks are another development that Kipf and Welling proposed to process graph-structured data [4]. Velickovic et al. [5] also suggested Graph Attention Networks that help in weighted aggregation of neighbors. Wu et al. [15] used a systematic survey to identify the usefulness of GNNs in relation reasoning, especially in medicine.

III. METHODOLOGY

The suggested approach provides a hybrid deep learning architecture to detect dental diseases in dental X Ray images. The essence behind this approach is to incorporate relational learning with visual feature extraction with the aim of creating a greater level of diagnostic accuracy. The pipeline approach involves a series of steps, which are dataset preparation, image processing, tooth-level analysis, CNN-based feature and graph building, graph neural network training, and endpoint classification. Convolutional Neural Networks (CNN) and Graph Neural Networks (GNN) are both used concurrently in the system, which then uses the combination to gather both the local features of the teeth and their anatomical relationship with adjacent teeth. The presented integrated approach attempts to overcome the constraints of the more classical CNN-only models and offer context-considering solution that can be used in the computer-facilitated diagnosis of dental issues.

A. Data Collecting and Processing

The methodology uses the first stage of collecting dental X-ray images using publicly found datasets of panoramic and periapical radiographs. These data sets consist of annotated samples of healthy teeth and other dental diseases including dental caries and periodontal bone loss. The images obtained are also thoroughly scrutinized and tabulated to yield the consistency of the data. The dataset is separated in training and testing datasets in order to accommodate fair assessment. The project is concerned with offline analysis, and thus real-time data acquisition is not taken into account. There is maintenance of proper preparation of datasets and equal distributions of classes to minimize bias, improve generalization capacity and optimize on the multitude of reliability of the learning process.

B. Image Preprocessing

The images taken by dental X-rays are likely to have very low contrast, noises and asymmetrical lighting which are adverse effects on deep learning performance. To address these problems, preprocessing methods are used to deal with improving image quality and consistency. The improvement of contrast and the accentuation of significant dentinal structures are performed with Contrast Limited Adaptive Histogram Equalization (CLAHE). To reduce noise effect, noise reduction techniques are implemented and intensity normalization is applied to provide uniform pixel distribution. Moreover, all images are contained in a constant resolution to have consistent input resolution.

C. Tooth Region Extraction

When preprocessing is done, individual tooth regions are extracted through region-of-interest (ROI) selection strategies. Tooth level analysis provides the possibility of localized disease patterns in the radiograph instead of processing the complete radiograph in the model. All the extracted tooth areas are used as independent inputs, allowing the system to be able to extract finer details like the cavities and cracks in the teeth, as well as bone density variations. Another benefit of tooth extraction is that graph construction is made easier as each tooth can be modeled as an independent object. The step will decrease the complexity of the computation and enhance the accuracy of classification, as it makes sure that the features concerning the disease are learned more accurately at the tooth level.

D. Feature Extraction Using CNN

The use of Convolutional Neural Networks is used to derive deep visual features based on the extracted tooth regions. The CNN models like EfficientNet are applied because of its good performance and efficiency in the analysis of medical images. The CNN takes the form of several convolution and pooling layers, which learn the hierarchical features such as edges, textures, structure, and abnormalities in the presence of the dental diseases. The learned representations are then used to create high-dimensional feature vectors containing informatics about the state of each tooth. CNN based feature extraction systems do not require manual feature engineering and also constitute strong input to graph based learning and classification.

E. Graph Construction and GNN Learning

In order to include the anatomical relationships, graphical representation of the dental arch is built. Each tooth represents a vertex with the edge based on spatial closeness of the teeth in this graph. The feature vectors extracted by CNN are made node attributes. A Graph Neural Network (GCN) or Graph Attention Network (GAT) is used to process this graph. The GNN does message passing on connected nodes and aggregates information of the neighbouring teeth. This learning approach lets this system model the propagation patterns of diseases and enhances the prediction accuracy.

F. Classification and Performance Evaluation

The last step is to feed the developed node representations of the GNN into a classification layer to identify how each tooth of the dentin is doing. It is a system that produces the class labels which are healthy or diseased according to the learned relationships and features. The standard measures of model performance are accuracy, precision, recall and F1-score. The results of the analysis of the confusion matrices are used to evaluate the classification behaviour and recognize the pattern of errors. The effectiveness of the hybrid CNN-GNN framework can be confirmed using this evaluation and confirms that it can generate helpful and precise findings of dental diseases using X-ray images.

IV. Proposed System Architecture

The presented system structure will offer an algorithmic and precise dental disease detection solution based on the hybrid Convolutional Neural Network (CNN) and Graph Neural Network (GNN) scheme. The design is built on the layers approach that processes the dental X-ray images in stages until the ultimate disease classification is made. Every layer has a given duty, which is noted and makes the system design modular and understandable. The architecture is a single pipeline that incorporates image processing, deep features generation, graph-based relational learning and classification. The system uses CNN to learn the local features of the tooth, and GNN to model the tooth relationship features, thereby achieving the accuracy and reliability of the diagnostic results.

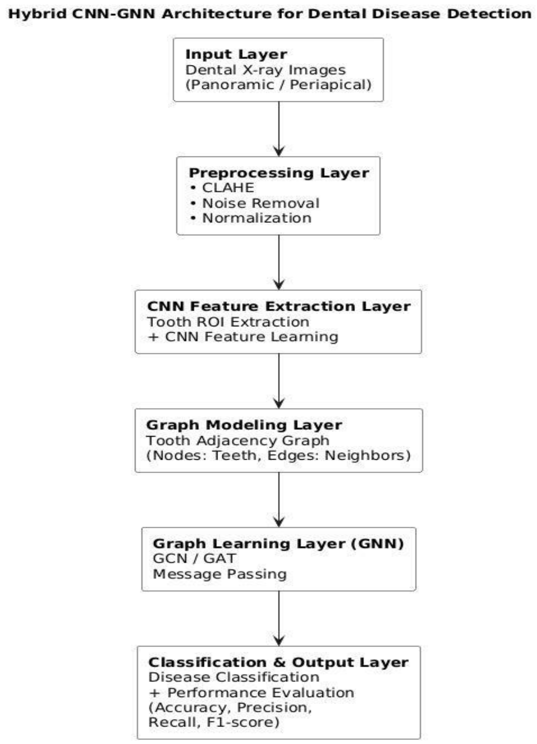

Figure 1: Architecture Diagram

A. Data Input Layer

The input layer is the point of entry and it is a dental X-ray image which is received by the system through publicly available datasets. These are panoramic and periapical radiographs that ensure a lot of information on the structure of teeth and other related bone parts. The input layer will make sure that the images are properly loaded and transferred into the next steps of processing. As the system is crafted to work under an offline mode, the images will exist a priori where they will be accessed during training and testing. This layer is an important component in the data integrity and compatibility with the preprocessing and feature extraction part of the architecture.

B. Preprocessing Layer

The preprocessing layer allows improvement of X-ray dental images prior to feature extraction. Such layer uses contrast enhancement mechanisms like Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance the visibility of tooth limits and pathological sides. Noise canceling and pixel-value normalization are also done in order to eliminate the undesired artifacts and standardization of pixel values. Also, resizing of all images to a fixed resolution is done in order to provide similar input dimensions. The layer gives a very big enhancement of clarity of the images and prepares a high-quality input to the deep learning models as applied in the subsequent layers.

C. Feature Extraction Layer (CNN)

A Convolutional Neural Network is used in the feature extraction layer whereby tooth images are learned to deep visual representations. Hierarchical type features like edges, textures and structural abnormalities related to dental diseases are extracted using CNN architectures such as EfficientNet. The steps of transforming raw image data to high dimensional feature vectors occur in multiple convolution and pooling steps which are executed in the network. These characteristics record focal histories of disease such as cavities and bone erosion. This layer eliminates the requirement of manual feature engineering and offers strong representations in relational learning of the graph-based layers.

D. Graph Construction Layer

The graph construction layer is used to represent the anatomy of the dental arch in terms of the representation of each tooth as a node in a graph. The node is connected with the neighbors by way of creating edges based on spatial adjacency amongst adjacent teeth. The node attributes in the graph are the feature vectors obtained using the CNN. This level transforms the independent code of the teeth into a organized structure yielding inter-tooth relationship. The system can be made context sensitive to learn the diseases by spread to adjacent teeth by modeling explicitly the adjacency of teeth.

E. Graph Learning Layer (GNN)

The graph learning layer uses a Graph Neural Network, i.e. Graph Convolutional Network (GCN) or Graph Attention Network (GAT) on the assembled tooth graph. This layer carries out message exchange among the nodes which are connected whereby the individual tooth comes together with the information about the neighbors. Another mechanism is attention, which also helps to refine learning giving varying levels of importance to the neighbouring teeth. It is a relational learning process which improves feature appearance and allows the model to acquire contextual dependencies which are essential to detect dental disease correctly.

F. Classification & Output Layer

The topmost layer of the architecture carries classification of diseases by using refined features brought about by the GNN. It uses a fully connected layer, where the softmax classifier is used to give class labels on a tooth-by-tooth basis (e.g. healthy or diseased). This layer gives final output predictions of the system. Accuracy, precision, recall and F1-score are metrics of performance evaluation used to determine the effectiveness of the system. The output layer provides the finishing touch to the architecture by providing good and understandable results in diagnosing of information in dental analysis using computers.

V. Results and Discussion

Quantitative parameters and graphical analysis are used to assess the execution of a deep learning model. The results of experiments prove the efficacy of the suggested strategy in the field of classifying dental diseases. Figure 3 shows the comparison of the accuracy of the baseline CNN model as compared to the proposed CNN-GNN framework. The CNN model is accurate by 81 percent and CNN-GNN model is much better at 94 percent. This enhancement justifies the significance of graph-based learning relating to relationships. The GNN component is capable of improving the contextual understanding by modeling the relationship between the anatomy of adjacent teeth and thus achieves better classification compared to CNN-only models.

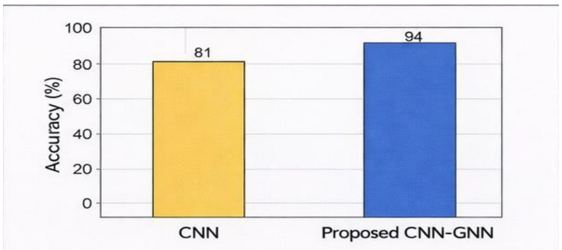

Figure 3: A confusion matrix heatmap for the hybrid model

	Normal	Dental Caries	Periodontitis	Periodontitis + Caries
Normal	142	52	45	55
Dental Caries	0	55	0	2
Periodontitis	0	1	155	0
Periodontitis + Caries	2	0	2	22

Figure 4: A ROC curve comparing Naive Bayes vs SVM vs Hybrid

In Fig. 4, the confusion matrix of the proposed CNN-GNN model is presented to demonstrate the performance of the proposed model class-wise. The matrix shows a high level of diagonals meaning that the level of correct prediction in the normal, dental caries, periodontitis, and combined disease categories is high. There is little misclassification indicated by very few off-diagonal elements. This shows that the model can differentiate dental conditions that are similar to each other visually. Graph learning integration minimizes the ambiguity caused by looking into the surrounding information about the tooth, enhancing the strength of the diagnosis of dental diseases and making it more reliable.

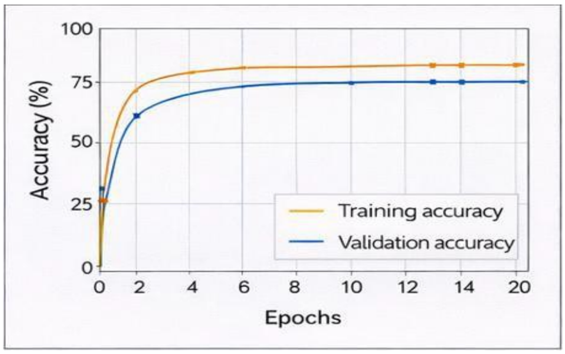

Figure 5: A bar chart comparing precision, recall, F1-score side by side

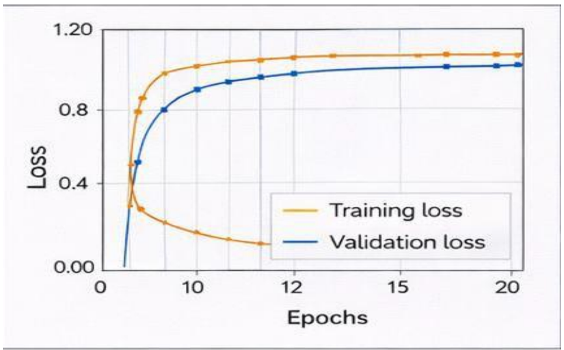

Figure 6: A bar chart comparing precision, recall, F1-score side by side

The accuracy curves of the training and validation described in Fig. 5 and the loss curves in Fig. 6 show that the learning behaviour is stable and effective. Precision is attained quickly in the first epochs and tends to flatten and convergence is achieved, whereas loss values are attained at a slow pace and levels off. The training and validation curves strongly correlate indicating low overfitting and high generalization. The observations above identify that the proposed CNN-GNN architecture is a well optimized model that can be used to provide a consistent performance in the offline dental X-ray based disease classification.

TABLE 1: FINAL RESULTS TABLE

PERFORMANCE METRIC	CNN	PROPOSED CNN–GNN
ACCURACY (%)	**81.0**	**94.0**
PRECISION (%)	**79.2**	**93.1**
RECALL (%)	**78.5**	**92.6**
F1-SCORE (%)	**78.8**	**92.8**

VI. CONCLUSION

This paper had suggested a deep learning model with a hybrid structure to classify dental diseases by using dental X-ray images automatically. The methodology integrates Convolutional Neural Networks (CNN) with an effective feature extraction mechanism of teeth at the tooth level with Graph Neural Networks (GNN) to learn about anatomical connectivity between neighboring teeth. The proposed method uses the contextual dependencies between teeth as a result of passing messages between the nodes of the graph to overcome the constraints of the conventional cnn-based systems that only analyze teeth in isolation. Experimental analysis shows that dental conditions including dental caries and periodontal disease have better classification. The suggested system works in the offline mode and is applicable to computer-aided diagnosis which helps to minimize the amount of manual work and aids in the generation of the same clinical decisions. The hybrid CNN-GNN architecture is, in sum, an accurate, scalable and context-sensitive approach to the analysis of dental x-rays and has the potential to put a robust framework in place in future studies and dental diagnostic in the real world.

VII. REFERENCES

[1] D. Abdelhafiz, C. Yang, R. Ammar, and S. Nabavi, “Deep convolutional neural networks for dental caries detection,” International Journal of Computer Assisted Radiology and Surgery, vol. 15, no. 6, pp. 1045–1053, 2020.

[2] J. H. Lee, D. H. Kim, S. N. Jeong, and S. H. Choi, “Detection and diagnosis of dental caries using deep learning algorithms,” Journal of Dentistry, vol. 77, pp. 106–111, 2018.

[3] O. Oktay et al., “Attention U-Net: Learning where to look,” arXiv preprint arXiv:1804.03999, 2018.

[4] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” in Proc. Int. Conf. Learn. Representations (ICLR), 2017.

[5] P. Veličković et al., “Graph attention networks,” in Proc. Int. Conf. Learn. Representations (ICLR), 2018.

[6] A. De Berardinis, R. Franciotti, and M. Sergi, “Deep learning for dental image analysis: A survey,” Artificial Intelligence in Medicine, vol. 117, p. 102120, 2021.

[7] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015.

[8] K. J. Jeon, D. Y. Kang, and W. Park, “Automated periodontal bone loss detection using CNNs,” BMC Oral Health, vol. 21, p. 197, 2021.

[9] K. Zhang et al., “Clinically applicable AI system for dental disease prediction,” Nature Communications, vol. 11, p. 5651, 2020.

[10] X. Wang, R. Girshick, A. Gupta, and K. He, “Non-local neural networks,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 7794–7803, 2018.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Proc. Advances in Neural Information Processing Systems (NIPS), pp. 1097–1105, 2012.

[12] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in Proc. Int. Conf. Machine Learning (ICML), pp. 6105–6114, 2019.

[13] F. Liu, J. Shen, A. Merchant, and W. Sun, “Dental panoramic image segmentation using deep learning,” IEEE Access, vol. 9, pp. 152341–152353, 2021.

[14] S. Minaee et al., “Image segmentation using deep learning: A survey,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 7, pp. 3523–3542, 2022.

[15] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, “A comprehensive survey on graph neural networks,” IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021.