




Identifying Suitable Deep Learning Approaches for Dental Caries Detection Using Smartphone Imaging

Leena Rohan Mehta^{1*}, Megha S. Borse¹, Meenal Tejan², Janhavi Shah²

¹ Department of Electronics and Telecommunication, MKSSS's Cummins College of Engineering for Women, Pune 411052, India

² Department of Oral Medicine & Radiology, Bharati Vidyapeeth's Dental College, Pune 411052, India

Corresponding Author Email: leena.mehta@cumminscollege.in

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

<https://doi.org/10.18280/ijcmem.120306>

ABSTRACT

Received: 9 August 2024

Revised: 10 September 2024

Accepted: 27 September 2024

Available online: 30 September 2024

Keywords:

deep learning, dental caries, ResNet50V2, ResNet101V2, ResNet152, DenseNet169, DenseNet201, dental imaging

This study aims to identify the most suitable deep learning model for early detection of dental caries in a new database of dental diseases. The study compares the performance of residual and dense networks using standard performance metrics. Dental caries is categorized into four classes based on dental practitioner recommendations. A novel database consisting of 1064 intraoral digital RGB images from 194 patients was collected in collaboration with Bharati Vidyapeeth's Dental College, Pune. These images were cropped to obtain a total of 987 single-tooth images, which were divided into 888 training, 45 testing, and 54 validation images. In Phase I experimentation, ResNet50V2, ResNet101V2, ResNet152, DenseNet169, and DenseNet201 were utilized. Phase II focused on ResNet50V2, DenseNet169, and DenseNet201, while Phase III concentrated on DenseNet169 and DenseNet201. For Phase I experimentation, the overall accuracy of dental caries classification ranged from 0.55 to 0.84, with DenseNet exhibiting superior performance. In Phase II, the overall accuracy varied from 0.72 to 0.78, with DenseNet achieving the highest accuracy of 0.78. Similarly, in Phase III, DenseNet201 surpassed other models with an overall accuracy of 0.93. The DenseNet201 algorithm shows promise for detecting and classifying dental caries in digital RGB images. This finding is significant for the future development of automated mobile applications based on dental photographs, which could assist dental practitioners during examinations. Additionally, it could enhance patient understanding of dental caries severity, thereby promoting dental health awareness.

1. INTRODUCTION

“Dental caries”, commonly referred as “tooth decay or cavities”, is recognized as one of the most prevalent chronic diseases globally. “Dental caries” is one such disease that affects individuals across all age groups. As per the reports by “World Health Organization (WHO)”, approximately 60-90% of young children and about all of adult population are affected by dental caries, making it a widespread public health issue. Early childhood caries (ECC) typically initiates at approximately 7 months of age and can progress to affect permanent dentition [1]. According to “Centers for Disease Control and Prevention (CDC)” over 52% of children between 6 and 8 years old have had decay in their primary dentition, indicating the scale of the problem.

ECC not only affects dental health but can have broader impacts on a child's development, nutrition, and quality of life. As the condition advances, it causes pain, infection, and difficulties in eating, speaking, and sleeping, leading to poorer general health and delayed growth. ECC typically starts as “white-spot lesions” on the “gingival margin of the upper primary incisors” and, if untreated, can result in severe destruction of the tooth crown [1, 2]. The rapid progression of

ECC is exacerbated by modern dietary habits, particularly a diet high in sugars, which accelerates bacterial activity in the mouth [3]. Studies show that sugary foods and drinks are the primary contributors to caries development, especially in young children, further emphasizing the need for preventive strategies and early intervention [4].

The global burden of ECC is alarmingly high, with reports of prevalence reaching up to 70% in preschool-aged children in some regions. Despite the severity of the problem, the disease often goes unnoticed in its early stages, and many parents fail to recognize its potential consequences, assuming that damage to temporary (primary) teeth is insignificant. However, untreated caries in primary teeth can lead to complications in permanent teeth, such as misalignment, enamel hypoplasia, and increased susceptibility to decay in the future [5]. Moreover, untreated ECC can progress to severe ECC (S-ECC), which affects the smooth surfaces of teeth and often requires more invasive treatments, including tooth extractions, under general anaesthesia in severe cases [6, 7]. Thus, ECC must be treated as a significant health issue rather than a minor inconvenience.

Beyond dental caries, other oral health problems, including gingivitis, periodontal disease, tooth sensitivity, and

malocclusion (misaligned teeth), are also common in children. These conditions, if left untreated, can lead to more serious dental and systemic health issues. Suboptimal oral hygiene practices and a delayed identification of oral diseases further exacerbate these problems, contributing to the worldwide impact of oral health. According to surveys conducted in different regions, many children, especially in developing countries, lack access to routine dental care, and parents often neglect the importance of dental hygiene in the early stages of life [8, 9]. A survey in Riyadh highlighted that awareness and knowledge about dental care, particularly in early childhood, are often lacking, leading to higher incidences of oral health problems [10]. Implementing preventive dental care and education, starting even before teeth emerge, has proven to greatly lower the occurrence of dental caries and other oral health issues.

Early detection of oral health issues is crucial for preventing further damage and reducing the burden on healthcare systems. While larger lesions are often visible during routine dental examinations, the initial stages of caries, such as white-spot lesions, are not easily detectable through visual inspection alone, even by experienced dental practitioners. Tools such as dental mirrors, light sources, and X-rays are commonly used in dental clinics to aid in diagnosis, but these methods are not always accessible in rural or underprivileged areas. The lack of effective screening and detection tools outside of dental offices, particularly in non-dental environments like schools, homes, or community health centers, poses a significant challenge to ensuring early intervention.

This gap underscores the urgent need for developing automated, non-invasive systems for early detection of dental diseases. Advances in technology, particularly in the fields of "artificial intelligence (AI) and deep learning", offer promising solutions. Automated systems can assist in screening for dental caries and other oral diseases using simple tools, such as smartphones or intraoral cameras, enabling timely identification of issues even in resource-limited environments. Such systems can be especially beneficial in rural or developing regions where access to professional dental care is limited. Moreover, these technologies have the potential to assist healthcare providers and non-professionals alike in identifying early signs of dental diseases, ultimately improving patient outcomes and reducing the time and costs associated with late-stage treatments.

Larger lesions in teeth are visible to the naked eye, but initial stages of dental caries or other dental diseases are not easily detected through visual examination alone, even by trained dental practitioners. They may use tools such as light sources, dental mirrors, or X-rays for a more thorough assessment.

There is a notable absence of effective screening and detection methods for dental diseases, particularly in non-dental environments like schools or homes, especially in rural areas of developing nations. Consequently, there is a pressing need for an automated system using simple tools to facilitate the timely identification of dental diseases. The outcomes of such a system could assist dentists and physicians in conducting thorough oral health examinations and save valuable time. Moreover, these automated systems could mitigate the limitations posed by the lack of training among non-professionals. The organization of the paper employs five major sections. Section 1 covers "Introduction" which offers general overview of the topic followed by bibliographic analysis covered in Section 2. Section 3 provides the details of

"Materials and Methods" used for "dental caries" classification. Section 4 is dedicated to the "Results and Discussion", while Section 5 presents the "Conclusion".

2. LITERATURE SURVEY

The existing literature predominantly emphasizes the utilization of 'deep learning algorithms' within the dental field. Numerous systems have been devised for diagnosing and prognosticating various dental diseases. Numerous studies have explored the application of various "machine learning models" for detecting different dental conditions, such as "caries, gingivitis, and other oral health issues", using digital images. Most approaches have focused on leveraging deep learning models like "convolutional neural networks (CNNs)", with significant progress reported in accuracy and diagnostic capabilities.

Patil et al. [11] proposed an "Adaptive Dragonfly Algorithm (DA) & Neural Network (NN) classifier" for the classification of 120 "digital X-ray images", achieving an accuracy of 93%. This enlightens the potential of hybrid models combining optimization algorithms with neural networks. While many of the studies focus on the potential of deep learning and machine learning in dental disease diagnosis, certain limitations persist. Sun et al. [12] provided a comprehensive review on the application of "machine learning" in dentistry, encompassing areas such as "oral cancer, periodontitis, dental caries, diseases of dental pulp and periapical lesion, dental implants, and orthodontics". For example, Stratigaki et al. [13] evaluated the use of "near-infrared light transillumination (NILT)" alongside bitewing radiography (BWR) for diagnosing dental conditions. Their results indicated that while NILT could be useful for routine examinations, it was not reliable enough to replace BWR for critical treatment decisions. This points to the ongoing challenge of finding non-invasive diagnostic methods that can match the reliability of traditional imaging techniques.

Similarly, Divakaran et al. [14] demonstrated that utilizing "GLCM features, SVM, KNN, and ANN classifiers" can effectively differentiate between decayed teeth and healthy ones. These studies underscore the versatility of machine learning models in handling different types of dental images.

Another area of research involves using machine learning for predictive analysis. Park and Choi [15] introduced the concept of "decayed occupied teeth" (DOT) to assess the relationship between feeding practices and cavity development in infants. Their study used logistic regression to find a significant correlation between feeding practices and early cavity development, revealing lower instances of cavities in children who consumed external foods compared to those exclusively breastfed. Similarly, Hung et al. [16] analyzed a large dataset of 5135 samples using SVM, XGBoost, random forest, KNN, and logistic regression classifiers to predict root caries based on patient age. They achieved 95% accuracy with SVM, indicating that machine learning models can be highly effective for predictive analysis in dentistry.

A notable trend is the use of "CNNs" for diagnosing dental diseases. The studies [17-19] demonstrated the superiority of "CNNs" for dental applications, particularly in detecting caries. "CNNs" have proven to be highly effective in processing dental images. In terms of future research directions, Chen et al. [20] emphasized the importance of collaboration between clinicians, researchers, and engineers to

advance AI integration into dentistry. They argued that interdisciplinary efforts are necessary to ensure that AI tools not only achieve high accuracy but are also practical and user-friendly in clinical settings. Javid et al. [21] employed ResNet50 to detect enamel decay from digital photographs with a 95% accuracy rate. However, while CNNs show great promise, their performance often depends on the quality and type of data used. For instance, Leo and Reddy [22] proposed a “hybrid neural network (HNN)” combining “artificial neural networks (ANN)” and “deep neural networks (DNN)” for dental caries classification on 480 digital radiographs. Their model outperformed traditional CNNs, suggesting that hybrid models may offer advantages over standalone deep learning methods in certain contexts. Additionally, Myint et al. [23] identified a gap in dental caries and gingivitis detection, stressing the need for more comprehensive models that consider bacterial levels and oral hygiene habits alongside image-based data. Uoshima et al. [24] emphasized the importance of a comprehensive skill set, including technical and non-technical skills, in dental education. They suggested integrating artificial intelligence (AI) to enhance the dental education system.

Beyond CNNs, other deep learning architectures have been explored. For example, Verma et al. [25] combined CNN with SVM for image classification, applied to 250 digital radiographs, achieving better performance than conventional CNN approaches. This suggests that hybrid deep learning models may provide more robust solutions for dental disease detection. Kumar et al. [26] conducted a comprehensive review of dental image fractionation and modalities utilized in dental image analysis. Similarly, Chen et al. [27] introduced a stage-wise detection approach for dental image analysis, utilizing a neural network to detect missing teeth and apply a numbering system, but their dataset was limited to 1250 digital X-rays. Tuzoff et al. [28] applied Faster R-CNN to 1594 panoramic dental radiographs, concluding that their proposed method could effectively update digital dental records in practice. Reyes et al. [29] identified the potential of machine learning in various dental subfields but noted the challenge of generalizing machine learning methods across different applications. Additionally, Musri et al. [30] reviewed the use of “deep learning convolutional neural networks (DLCNNs)” for identifying dental problems, concluding that DLCNNs have shown promising results, particularly in detecting dental caries.

A key limitation in many studies is the restricted scope of datasets used. For instance, Zhang et al. [31] developed a multistage “deep learning model” using SSD MobilenetV2 for cavity detection from RGB images, but their study was limited to front teeth, leaving out other areas of the mouth. While these models show high precision and recall, the restricted dataset scope limits their applicability to broader, real-world dental scenarios.

Some researchers have attempted to tackle the challenge of limited datasets by integrating different types of imaging techniques or by expanding the size of their datasets. For example, Rashid et al. [32] used a mixed dataset of 936 digital radiographs and 90 digital photographs to develop a hybrid Mask RCNN model for automated dental caries detection, achieving accuracy levels ranging from 0.78 to 0.92. This demonstrates the potential of combining different image types to improve the robustness of deep learning models in detecting dental diseases across varied conditions. In many other studies [33-38] specialised models for dedicated dental caries

detection were designed.

Additionally, real-time applications of deep learning in clinical settings are beginning to emerge. Hung et al. [39] suggested that a real-time online clinical tool could significantly enhance diagnostic precision for dentists. While tools like CNNs and hybrid models have shown success, the generalization of these methods in clinical practice remains an obstacle due to variability in dental conditions and imaging techniques.

Data from the Indian Dental Association's National Oral Health Programme survey indicates a concerning shortage of dentists relative to the rural population. Dental health in India is further hampered by socioeconomic factors such as limited education, awareness, and economic constraints, leading to severe oral health issues. Dental treatments are often financially prohibitive for disadvantaged families. Therefore, there is a critical need for easily accessible early diagnosis of dental problems. An automated, easily accessible, and cost-effective system for early detection and prognosis of dental issues is essential. Such a system would facilitate timely intervention and treatment, ultimately improving oral healthcare outcomes. Additionally, it would enhance precision in diagnosis and optimize time utilization for dental practitioners. This approach has the potential to significantly enhance oral healthcare accessibility and affordability for underserved communities.

In conclusion, while significant progress has been made in dental healthcare using machine learning models, several gaps remain. Many studies are limited by small datasets, specific imaging techniques, and challenges in generalizing models to clinical environments. More research is needed to validate models in real-world settings and develop AI systems for non-clinical environments, such as rural areas. Additionally, there is a lack of research on automation using digital RGB images, which this study addresses by evaluating deep learning models on intra-oral photographs and comparing residual and dense networks to find the best-performing algorithm.

3. MATERIALS AND METHOD

The literature review identified a notable challenge in dental research: the absence of labeled dental databases containing digital RGB images. This motivated us to create a new database specifically targeting dental diseases. For an overview of the methodology used in our proposed model, please refer to Figure 1.



Figure 1. Sample images: (a) Anterior region, teeth in centric occlusion (b) Right posterior region, teeth in centric occlusion (c) Left posterior region, teeth in centric occlusion (d) Anterior region, teeth in edge-to-edge occlusion (e) Palatal/occlusal surface view of maxillary teeth (f) Palatal/occlusal surface view of mandibular teeth

3.1 Noninvasive data acquisition

The collected database comprises 1164 digital photographs of 194 adult patients in the age group 14 to 60 years who came to OPD (out patient's department) of 'Bharati Vidyapeeth's Dental College and Hospital, Pune'. The dentist will assess each patient clinically for the presence of disease or abnormal conditions. The general dental assessment commences from the moment the patient enters the room. The external appearance of the patient includes facial appearance, skin, mobility, smell, etc. General dental assessment is divided into extraoral and intraoral examinations. An extraoral examination is performed for head and neck posture, etc. It involves the symmetry of the face. The extraoral examination is followed by the intraoral examination.

For this procedure, each patient is asked to rinse their mouth with normal tap water. For opening of the mouth, winged cheek retractors are used. For the purpose of intraoral examination, diagnostic instruments including mirror, straight probe and explorer were used. With the help of One plus Nord 2T images were clicked. One plus Nord 2T is with 50-megapixel Sony IMX766 camera sensor with pixel size of 1.0 μm , lens quality of 6P, optical image stabilization, aperture of $f/1.8$, and ARM Mali-G77 MC9 GPU for post-processing of the image.

For examination and clicking of intraoral cavities images, the patient was asked to sit in a supine position on the dental chair with the light turned off. Camera flash was used when clicking pictures. The camera was adjusted and stabilized at a distance of 2 inches from the lips such that the camera is perpendicular to the plane of the teeth in focus for each image. The patient's face was adjusted such that the head is parallel to the plane on which the camera lens was adjusted and stabilized. Then on the basis of clinical examination, a diagnosis was made by the dentist.

Six images per patient are taken to cover the entire oral cavity. Sample images are given below. Please refer to Figure 1. For the first and second images, the central incisors are considered the object of focus and camera lens will be perpendicular to them. The patient faces directly upwards. The third and fourth images are taken with the first premolar as the focus on the respective sides. For these, the patient is asked to turn their head at 45° to the left side while taking image of the right side and the head will be turned to the right side while taking a picture of the left side. For the first 4 images, a winged cheek retractor will be used to ensure the required details are captured well and for stabilization. For the fifth and sixth images, an intraoral mirror will be placed for taking maxillary and mandibular occlusal surfaces, respectively. The focus will be adjusted to the mirror so that the occlusal surface of each arch is clearly visible.

i. Inclusion criteria

- Patients willing to participate in the study and those who provide their consent.

- Patients with hard tissue diseases including caries, stains, erosion, attrition, abrasion, abfraction, periodontal diseases including gingivitis and malocclusions.

- Patients with healthy teeth, ideal occlusion, and ideal periodontal conditions.

- Images with appropriate resolution taken with predetermined standardized method

ii. Exclusion Criteria

- Patients with incomplete clinical records, previous history of surgery or craniofacial anomalies, maxillofacial trauma

- Patients with a history of medical conditions and vulnerable patients

3.1.1 Database labeling

Image-wise labeling would have become tedious in this case and the model would also have become complicated, hence tooth-wise labeling is done. Labeling is done in consultation with the dentist. The standard FDI tooth notation system is used. This was also useful for detecting the location of the tooth. Dental caries is classified as Grade 0 (Healthy), Grade 1 (Pit and Fissure / Start of cavitation), Grade 2 (Deep Cavities, Structural damage, Occlusal) and Grade 3 (Total loss of tooth structure, root stumps) as suggested by dental practitioner.

According to the FDI notation system, the mouth is divided into four quadrants, each with its specific numbering: "the upper right" quadrant is numbered 11 to 18, "the upper left" 21 to 28, "the lower right" 41 to 48, and "the lower left" 31 to 38. These numbers assist dentists in identifying individual teeth. In our study, we adopt a labeling system that aligns with this numbering. For example, patient 1 is recorded with class 0 caries in tooth 11, while patient 34 has class 3 caries in tooth 45. Dental practitioners document these labels in an Excel sheet, where the first column shows patient identifiers and the other columns list tooth numbers, with caries class numbers noted in the intersecting cells. Images were cropped tooth-wise. They were labeled class-wise as per the excel sheet prepared.

3.1.2 Data augmentation

Consider the database 'D' as the collection of original RGB images of individual teeth. Initially, the database was organized into two primary folders namely "train and test". Each of these folders was further divided into subfolders corresponding to different classes, with images manually sorted according to the type of "dental caries" with the assistance of a dental practitioner. Upon conducting a statistical analysis, it was discovered that the database is significantly imbalanced, as class 0 contains nearly 4,000 images, while classes 1 and 2 have considerably fewer images. To address this imbalance, "data augmentation techniques" such as image scaling, zooming, flipping, and shearing were applied. This process also involved filling in missing values and encoding the dataset to prepare it for further processing.

Let $I_i(x,y)$ represents the images where $i=1,2,3,4$. Each image is located in the dataset 'D' either in train or test folders. The size of each image $I_i(x,y)$ is (150×150) . After this rescaling of the image was carried out as shown in Eq. (1).

1. Rescaling

$$I_r(x', y') = I\left(\frac{x}{s_x}, \frac{y}{s_y}\right) = I\left(\frac{x}{255}, \frac{y}{255}\right) \quad (1)$$

2. Shear transformation

$$I_s(x', y') = I_s[I_r(x', y')] = I(x' + ky', y') = I(x' + 0.2y', y') \quad (2)$$

3. Zoom transformation

Pixel values of the image x' and y' are updated in this phase of augmentation. It is scaled up by a factor of 20%. Nearest neighbour interpolation method is used.

$$I_z(x', y') = I_s\left(\frac{x'}{s_x}, \frac{y'}{s_y}\right) = I_s\left(\frac{x'}{0.2}, \frac{y'}{0.2}\right) \quad (3)$$

4. Horizontal flip

$$I_f(x', y') = I_z((W - x'), y') \quad (4)$$

where, W =width of the image.

5. Sequential operation for augmented output image

$$I_{augmented}(x', y') = I_r\left(\left(W - \frac{x' - ky'}{S_x}\right), \frac{y'}{S_y}\right) \quad (5)$$

This equation represents the final pixel value after applying the rescaling, shear, zoom, and horizontal flip operations sequentially. Figure 2 showcases sample augmented sample images.

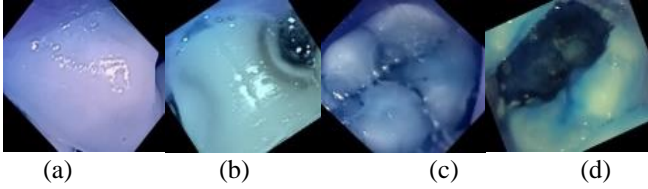


Figure 2. Augmented (a) Class 0 (b) Class 1 (c) Class 2 (d) Class 3

3.2 Methodology

3.2.1 Deep neural network models

“ResNet” and “DenseNet” variants were chosen for their proven performance in complex image classification tasks, including medical imaging, due to their ability to handle deep network training and feature extraction effectively. ResNet’s residual learning framework helps mitigate the vanishing gradient problem, enabling it to train very deep networks and extract features from intricate dental images with high accuracy. DenseNet’s feature reuse mechanism enhances gradient flow and feature propagation, which is advantageous for detecting subtle patterns in dental images. Both models are also known for their efficiency in training and inference, which is crucial given the computational resources available for dental datasets. Their previous success in medical and dental imaging tasks further validates their suitability for this study. The advanced feature extraction capabilities of ResNet and DenseNet, along with their scalability to various dataset sizes and image resolutions, align well with the diverse nature of dental images. Supported by existing literature, these models are confirmed to be effective for dental image analysis, ensuring they are well-suited to deliver high accuracy and reliable results for this research.

The study was conducted in three phases, each utilizing different numbers of images across all classes. Phase I involved 240 images, Phase II utilized 800 images, and Phase III used 967 images. During each phase, the performance of residual networks and dense networks was compared for dental caries classification on the novel dataset. Standard architectures of Residual networks and DenseNet were adapted to accommodate the customized dataset, with certain layers modified as needed. Hyperparameter tuning was conducted to enhance performance. In total, three models of residual networks and three models of dense networks were implemented. Fine-tuning and adjustments were made to these models to improve accuracy and training. The following architectures were employed. Python was used for implementation of the models. Libraries like “Numpy, Pandas, Matplotlib, Sklearn, Seaborn, Tensorflow, Keras” etc. were used.

A. Modified residual networks

The residual networks comprise skip connections, due to which the issue of vanishing gradients is resolved up to great extent in backward propagation. Three different layer models were implemented, namely “ResNet50, ResNet101 and ResNet152”. The number in the name of the model indicates the depth of the model. For example, ResNet50 is 50 layers deep and so on. In the model implemented, we changed the pooling function from Maxpooling to Average pooling. Figure 3 shows the implementation of “ResNet50V2” for the collected database.

The implemented ResNet model comprises functional layers such as “average pooling, batch normalization, dropout, and dense layers”. The input image size used was (256×256). It incorporates skip connections, enabling direct connections from input to output to mitigate the issue of vanishing gradients. The network is pretrained on the large-scale “Imagenet” database with millions of images. Base model of the residual network is frozen and top layers were added to deal with novel database. The model employs the “categorical cross-entropy loss function” and the “Adam optimizer” with a standard learning rate, along with early stopping. A batch size of 32 was employed for training. Dropout layer with a factor of 0.2 was added to avoid overfitting of the model. The ‘Relu’ activation function was employed, while the “softmax” activation function was used in the last dense layer to reduce data dimensionality from 2048 to 256 in this study. Similar configurations and layers are utilized for ResNet101 and ResNet152 models.

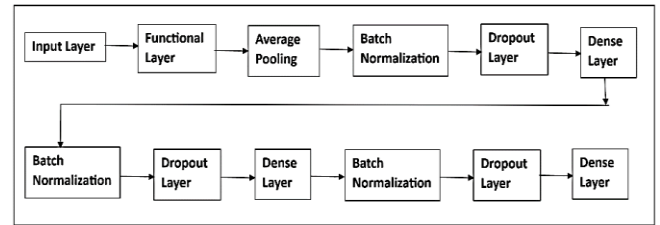


Figure 3. Modified ResNet50v2 architecture

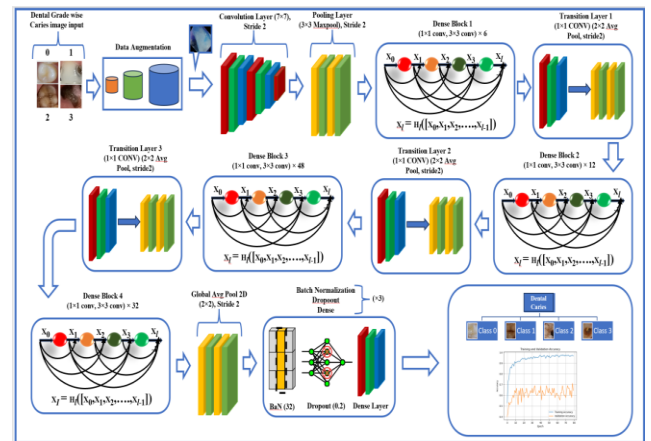


Figure 4. Modified DenseNet201 model architecture

B. Modified DenseNet

DenseNet is a parametrically efficient model that has been pre-trained on large datasets such as “ImageNet.” In our approach, we utilized transfer learning by leveraging the pretrained DenseNet model. Unlike residual networks, DenseNet exhibits strong connectivity between all previous and future layers. This dense connectivity allows even smaller

features from the initial layer to influence the final feature maps. This connectivity proves advantageous, especially in smaller object databases like dental images. Additionally, DenseNet is known to perform better in mitigating the vanishing gradient problem compared to other models. Figure 4 presents the proposed DenseNet201 model for the classification of dental diseases.

1. Base model (DenseNet201)

The primary base model employed in this study utilizes "DenseNet201", which features four dense blocks. Each dense block incorporates a bottleneck layer with a (1×1) convolutional filter, preceding a (3×3) "convolutional layer". This design is computationally efficient to reduce the size of the feature maps maintaining high feature extraction quality. A transfer learning approach was employed, leveraging the "DenseNet201 architecture" pre-trained on the "ImageNet dataset", which includes millions of labelled images across thousands of classes. The pre-trained weights provide a robust foundation for feature extraction. During training, the weights of the "DenseNet201 base model" were frozen. Thus, only top layer weights were updated resulting adaptation of the model for the specific task of dental caries classification without altering the foundational features learned. This selective training approach contrasts with the standard DenseNet201 model, where all weights, including those in the base architecture, are typically trainable by default unless specified otherwise.

2. Top layers of the model

Following the DenseNet201 base model, the proposed algorithm integrates several additional layers designed to tailor the model's learned features to the specific task of dental caries classification across four classes. These top layers include "Global Average Pooling, Batch Normalization, Dropout, and Dense layers". The "Global Average Pooling layer (GlobalAvgPool2D ())" replaces traditional flattening layers, reducing spatial dimensions by averaging each feature map, thereby capturing global spatial information while minimizing overfitting risks. Furthermore, to decrease the dimensionality of the feature matrix and reduce the number of trainable parameters, a (2×2) average pooling with a stride of 2 was applied.

"Batch Normalization" is implemented to stabilize the output by normalizing it to have a "zero mean and unit variance," which speeds up training and minimizes internal covariate shifts. The "Dropout layer" randomly set 20% of the input units to zero during training updates that prevent overfitting. The architecture included two "fully connected Dense layers" each with 128 units, utilizing "ReLU (Rectified Linear Unit) activation function" to introduce "non-linearity" and identify complex patterns within the data. The combination of these layers effectively transforms the 4D tensor output from DenseNet201 into a 1D tensor suitable for the final classification task.

3. Output Layer

The proposed model culminates in an output layer specifically designed for the classification of dental caries into four classes. It features a "dense layer" with a number of units equal to the four classes, using the "Softmax activation function" to produce a "probability distribution" over the classes, making it ideal for multi-class classification tasks. Unlike the standard DenseNet architecture, which typically concludes with classification layers following four dense blocks, this study modifies the architecture by incorporating a combination of global average 2D pooling and three BDD

layers ("Batch Normalization, Dropout, and Dense layers"). Image batch size used was of 32 and dropout factor of 0.2 was utilized. The final "Dense layer" applies the "Softmax activation function" to output a flattened vector representing the model's confidence in each class. The model is trained using the "Adam optimizer with a standard learning rate", ensuring efficient and effective convergence.

This proposed model's architecture, with its customized top layers and output configuration, presents a novel approach for the task of dental caries detection, capitalizing on the strength of DenseNet201's feature extraction capabilities while adapting the model to the specific needs of this classification problem.

C. Modified DenseNet201's architectural benefits for dental caries classification

1. Leveraging DenseNet201 for feature extraction

a. Dense connections: It allows reuse of all features reducing the need for redundant parameters resulting in more compact and efficient representations. "DenseNet" architecture inherently mitigates the "vanishing gradient problem", maintaining high efficiency and reduced computational costs.

b. Pre-trained Weights: Since the proposed model was already pre-trained on the "ImageNet dataset", it is beneficial in scenarios with limited data availability, enabling the model to generalize more effectively on smaller, domain-specific datasets.

c. Adaptability: In the proposed model, the weights of the DenseNet201 base layers are frozen, allowing the network to retain the features learned from "ImageNet" while concentrating on fine-tuning the newly added dense layers.

2. Efficient dimensionality reduction

a. Dimensionality Reduction: "Global Average Pooling (GAP)" reduces each feature map to a single value, transforming high-dimensional tensors into lower-dimensional vectors without losing spatial information. This technique not only reduces the risk of overfitting, particularly in complex models, but also maintains translational invariance, making the model robust to different spatial configurations.

b. Contrast with Flattening: Traditional flattening methods can lead to a high number of parameters, thereby increasing the risk of overfitting. GAP addresses this issue by minimizing the number of parameters while preserving crucial features.

3. Robust Regularization with Batch Normalization and Dropout

a. Combination of "Batch Normalization and Dropout: Integrating "Batch Normalization and Dropout" between layers is a novel approach that ensures model stability during training. Batch Normalization reduces internal covariate shifts by normalizing inputs, accelerating training and improving gradient flow through the network. Dropout, as a regularization technique, randomly drops neurons during training, preventing co-adaptation and improving the model's generalization.

b. Effect on Convergence: The sequential application of Batch Normalization and Dropout enhances convergence speed and stability, often leading to better accuracy and robustness against noise in the input data.

D. Mathematical Model of modified DenseNet201

Let 'D' be the image database containing the images of "class 0, class 1, class 2 and class 3". As explained in section 3.1, Data augmentation was carried out for increasing the number of images and balancing the class-wise dataset. Following mathematical operations were carried out on the

images of the dataset.

1. Base Model Transformation

Each input image $X^{(i)} = I_{augmented}(x', y')$ from the database is transformed into feature maps $X_{Base}^{(i)}$ using the Modified DenseNet201 architecture.

$$X_{Base}^{(i)} = DenseNet201(X^{(i)}) \quad (6)$$

where, $X^{(i)}$ =ith Image in the database 'D'; $X_{Base}^{(i)}$ =feature map with shape (4,4,1920).

2. Global Average Pooling 2D

The feature maps $X_{Base}^{(i)}$ are reduced to a 1D vector $X_{Pooled}^{(i)}$ by applying global average pooling. Eq. (7) represents standard equation for global average pooling and Eq. (8) shows Global average pooling implemented in Modified DenseNet201 model.

$$GAP(f) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W f_{i,j} \quad (7)$$

$$X_{Pooled}^{(i)} = \frac{1}{4 \times 4} \sum_{x=1}^4 \sum_{y=1}^4 X_{Base}^i(x, y, c) \quad (8)$$

For $c = 1, 2, 3, \dots, 1920$

where, $X_{Pooled}^{(i)}$ has shape (1920); c =Number of channels.

3. Batch normalization Layer 1

Batch normalization is applied to the pooled feature map $X_{Pooled}^{(i)}$. Eq. (9) shows standard form of Batch Normalization and Eq. (10) shows implemented one.

$$BN(x) = \gamma \left(\frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}} \right) + \beta \quad (9)$$

$$X_{Normalized1}^{(i)} = \gamma_1 \frac{X_{Pooled}^i - \mu_1}{\sqrt{\sigma_1^2 + \epsilon}} + \beta_1 \quad (10)$$

where, μ_1 and σ_1^2 =Mean and variance of X_{Pooled}^i within the batch of 32 images respectively; γ_1 and β_1 =learnable parameters

4. Dropout Layer 1

Dropout is applied to normalized feature map $X_{Normalized1}^{(i)}$.

$$X_{Dropped1}^{(i)} = Dropout(X_{Normalized1}^{(i)}, p=0.2) \quad (11)$$

where, p =Dropout rate=20% of the neurons are randomly set to zero.

5. Dense Layer 1

A fully connected dense layer with ReLU activation was employed.

$$X_{Dense1}^{(i)} = ReLU(W1 \cdot X_{Dropped1}^{(i)} + B1) \quad (12)$$

where, $W1$ =Weight Matrix of the dense layer; $B1$ =Bias Vector.

6. Batch normalization Layer 2

$$X_{Normalized2}^{(i)} = \gamma_2 \frac{X_{Dense1}^i - \mu_2}{\sqrt{\sigma_2^2 + \epsilon}} + \beta_2 \quad (13)$$

where, μ_2 and σ_2^2 =Mean and variance of X_{Dense1}^i within the batch of respectively; γ_2 and β_2 =learnable parameters.

7. Dropout Layer 2

Dropout is applied to normalized feature map $X_{Normalized2}^{(i)}$.

$$X_{Dropped2}^{(i)} = Dropout(X_{Normalized2}^{(i)}, p=0.2) \quad (14)$$

where, p =Dropout rate = 20% of the neurons are randomly set to zero.

8. Dense Layer 2

A fully connected dense layer with ReLU activation was employed.

$$X_{Dense2}^{(i)} = ReLU(W_2 \cdot X_{Dropped2}^{(i)} + B_2) \quad (15)$$

where, W_2 = Weight Matrix of the dense layer; B_2 =Bias Vector.

9. Batch normalization Layer 3

$$X_{Normalized3}^{(i)} = \gamma_3 \frac{X_{Dense2}^i - \mu_3}{\sqrt{\sigma_3^2 + \epsilon}} + \beta_3 \quad (16)$$

where, μ_3 and σ_3^2 =Mean and variance of X_{Dense2}^i within the batch of respectively; γ_3 and β_3 =learnable parameters.

10. Dropout Layer 3

Dropout is applied to normalized feature map $X_{Normalized3}^{(i)}$.

$$X_{Dropped3}^{(i)} = Dropout(X_{Normalized3}^{(i)}, p=0.2) \quad (17)$$

where, p =Dropout rate=20% of the neurons are randomly set to zero.

11. Output Layer

Softmax function is given by the Eq. (18).

$$\phi(x_i) = \frac{e^{x_i}}{\sum_{j=1}^k e^{x_j}} \quad (18)$$

$$Y_{Output}^{(i)} = Softmax(W_{Output} \cdot X_{Dropped3}^i + B_{Output}) \quad (19)$$

where, W_{Output} =Weight Matrix of the output layer; B_{Output} =Bias Vector.

The aggregate count of parameters in the model is 18,593,604, with 267,268 trainable parameters and 18,326,336 non-trainable parameters.

4. RESULTS AND DISCUSSION

A total of 1064 intraoral images from 194 patients were initially captured and cropped tooth by tooth. Among these, 200 images were specifically selected for a performance analysis study. The patients' ages ranged from 14 to 70 years, with an average age of 26 years. Both male and female samples were included in the dataset, and patient demographics such as age and sex were kept undisclosed. Various types of teeth (molars, premolars, canines, etc.) were considered, while teeth with root canal caps and implants were excluded from the study. The primary objective was to classify dental caries into four classes to aid in early detection.

To enhance the model's robustness, data enrichment techniques were employed. The research primarily centers on early dental caries detection and was conducted in phases. Phase I involved training on 200 images and testing on 40 images. Phase II expanded the dataset to 640 training images and 160 testing images. Finally, in Phase III, the dataset further increased to 888 training images and 79 testing images, with each phase building upon the results of the previous one.

In Phase I, five different models—ResNet50V2, ResNet101V2, ResNet152, DenseNet169, and

DenseNet201—were implemented and meticulously evaluated for performance. Phase II focused on ResNet50V2, DenseNet169, and DenseNet201 models, while Phase III compared DenseNet201 and DenseNet169 models. The performance parameters considered for the study include precision, recall, F1 score, and overall accuracy. Additionally, a confusion matrix is presented in the results to justify the suitability of the chosen model.

4.1 Phase I experimental results

The accuracy levels ranged from 0.66 to 0.86 when utilizing

50 images per class for training, calculated across random samples encompassing all classes collectively. For class 0, precision varied from 0.69 to 0.90 across different models. Notably, ResNet101 exhibited superior performance in identifying healthy teeth, whereas DenseNet169 demonstrated greater efficacy in early dental caries detection, achieving a precision of 0.86 for class 1. However, it showed lower precision (0.77) in detecting class 3 caries. DenseNet201 also yielded a precision of 0.83 for class 1 and achieved 100% precision for class 3. Overall, DenseNet201 emerged as a consistently reliable solution among the models under consideration.

Table 1. Performance matrix for Phase I experimentation

Sr. No.	Model	Number of Epochs	Class 0	Class 1	Class 2	Class 3	Accuracy	Recall	F1 Score
1	ResNet50V2	50	0.69	0.50	0.83	0.77	0.70	0.75	0.68
2	ResNet101V2	50	0.90	0.80	0.71	0.82	0.81	0.8	0.78
3	ResNet152V2	50	0.86	0.74	0.38	0.50	0.66	0.57	0.55
4	DenseNet169	50	0.80	0.86	0.75	0.77	0.81	0.8	0.79
5	DenseNet201	50	0.87	0.83	0.77	1.00	0.86	0.85	0.84

Upon conducting comparative performance analysis, DenseNet models exhibited superior performance. The detailed performance results are tabulated in Table 1, while graphical representations of accuracy, class-wise precision, and Comparative F1 scores are depicted in Figures 5(a), 5(b), and 5(c) respectively. Figures 6-20 depict confusion matrices, as well as plots illustrating "training and validation accuracy" and "training and validation loss," for all models employed in the study.

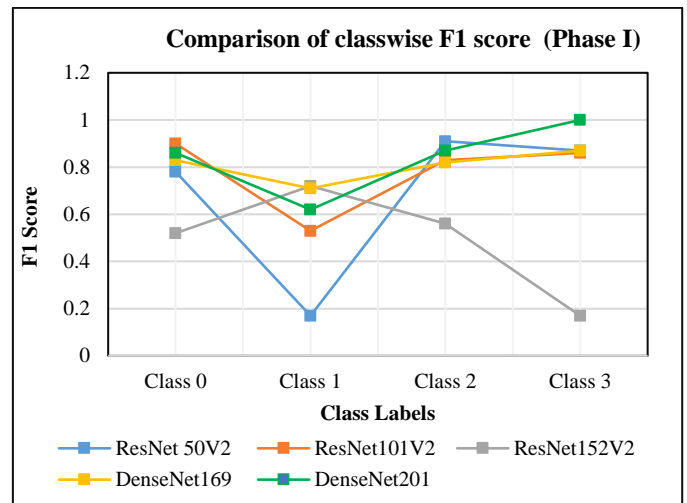
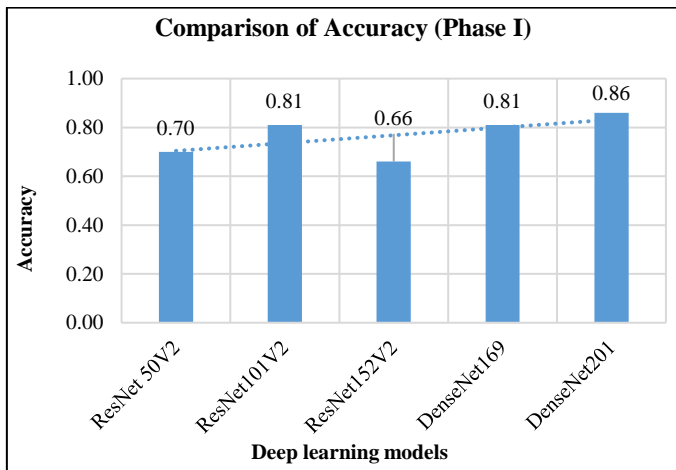


Figure 5. (a) Comparison of accuracy (Phase I) (b) Comparison of class-wise precision (Phase I) (c) Comparison of class-wise F1 score (Phase I)

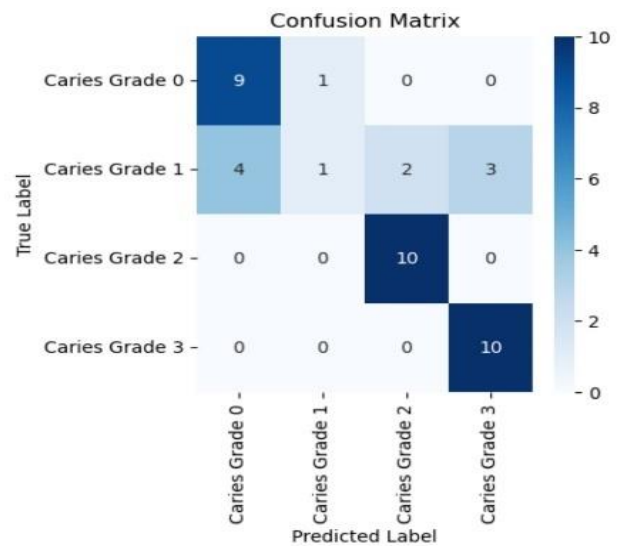
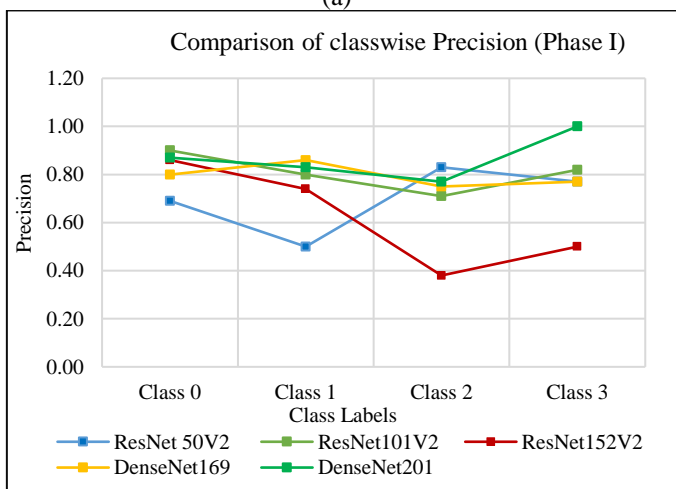


Figure 6. Confusion matrix for ResNet50V2

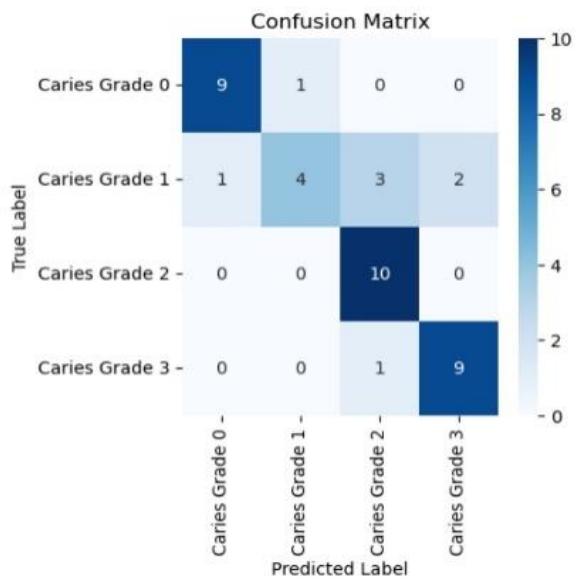


Figure 7. Confusion matrix for ResNet101V2

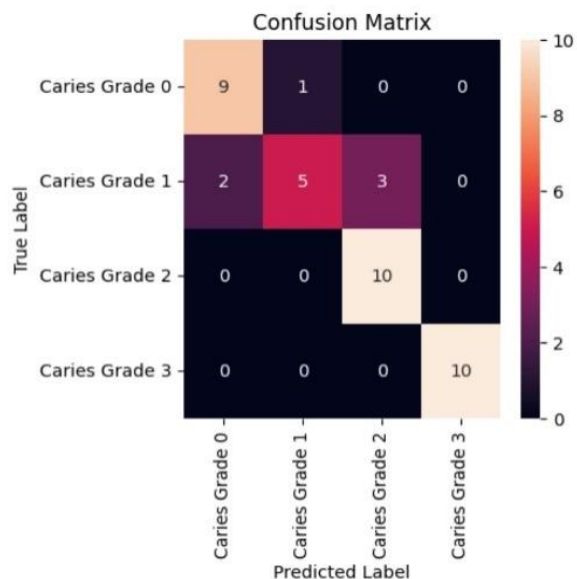


Figure 10. Confusion matrix for DenseNet201

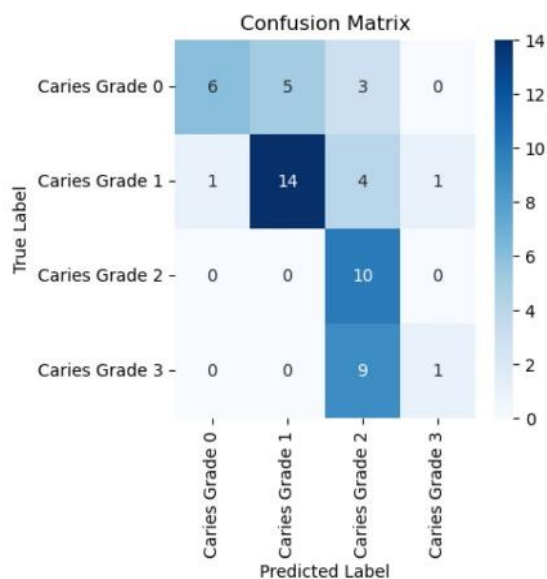


Figure 8. Confusion matrix for ResNet152

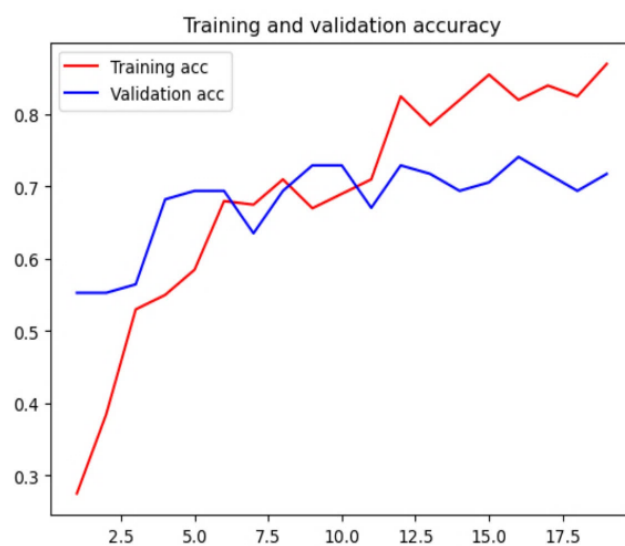


Figure 11. Training and validation accuracy for ResNet50V2

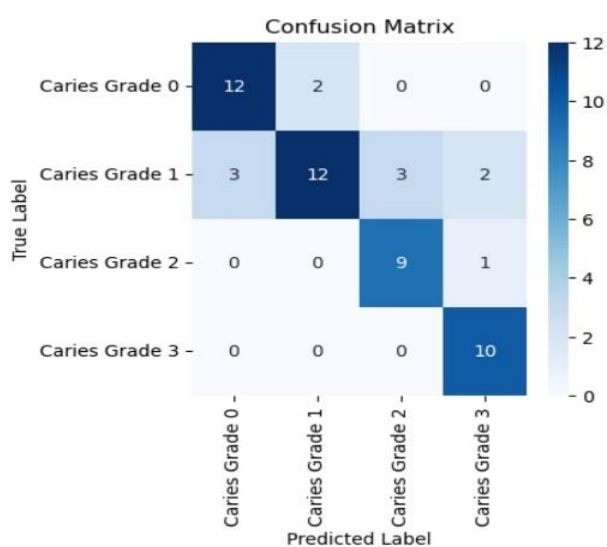


Figure 9. Confusion matrix for DenseNet169

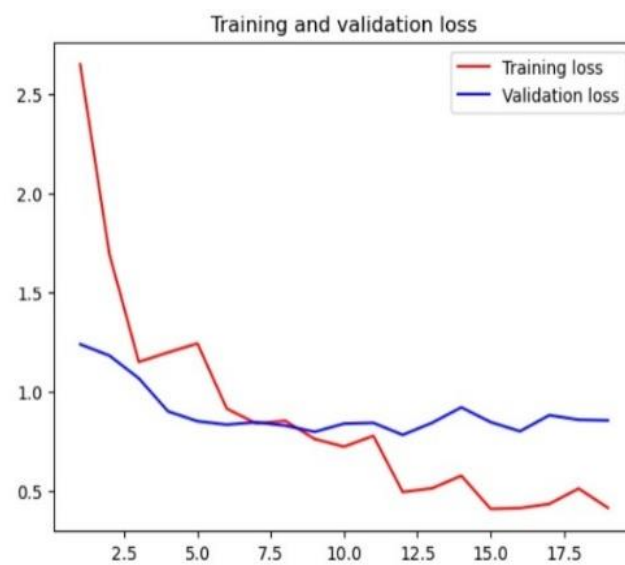


Figure 12. Training and validation loss for ResNet50V2



Figure 13. Training and validation accuracy for ResNet101V2



Figure 14. Training and validation loss for ResNet101V2

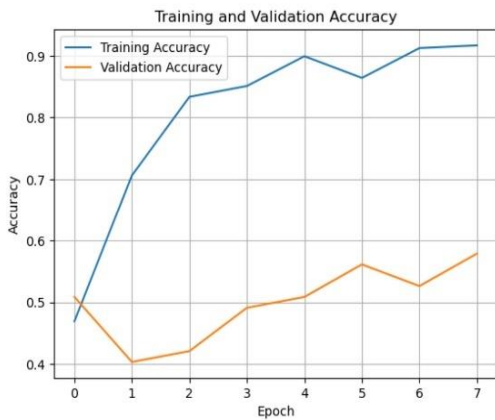


Figure 15. Training and validation accuracy for ResNet152



Figure 16. Training and validation loss for ResNet152

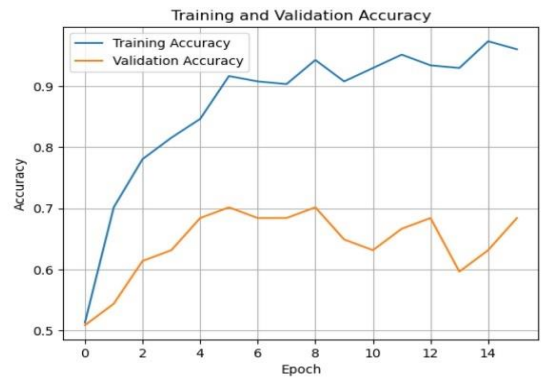


Figure 17. Training and validation accuracy for DenseNet169



Figure 18. Training and validation loss for DenseNet169



Figure 19. Training and validation accuracy for DenseNet201

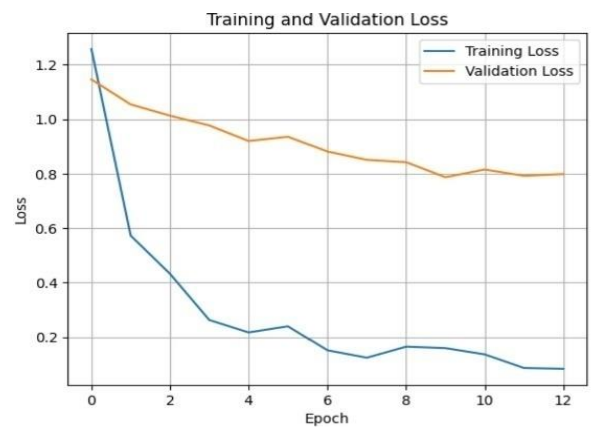
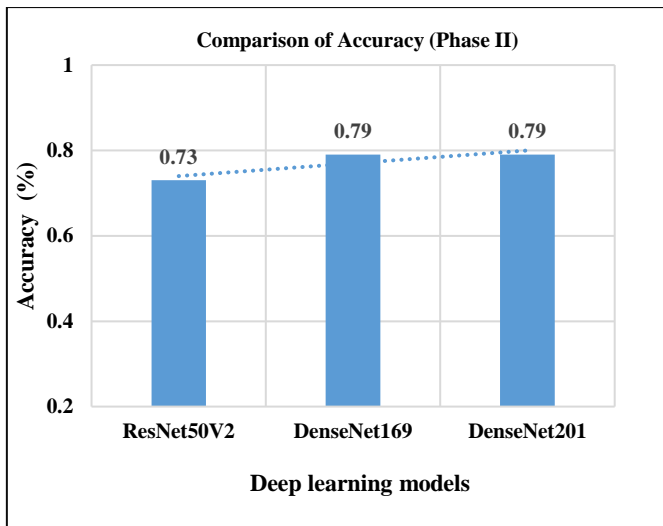


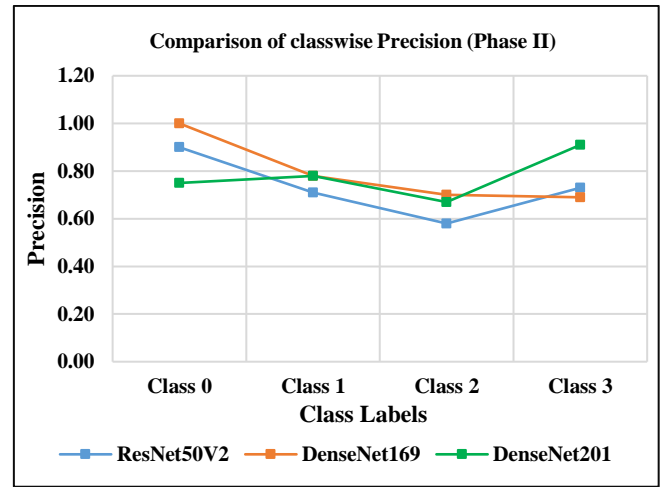
Figure 20. Training and validation loss for DenseNet201

4.2 Phase II experimental results

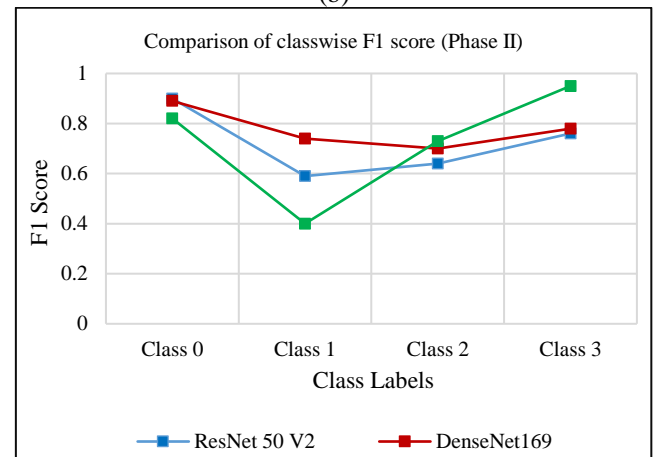
The accuracy ranged from 0.73 to 0.79 across the models. ResNet50v2 demonstrated a promising precision of 0.90 for class 0 (representing healthy teeth), but struggled to effectively identify and classify dental caries in other classes, particularly exhibiting poor performance with a precision of approximately 0.58 for class 2 caries. Overall, DenseNet models outperformed other models. However, DenseNet169 consistently struggled to accurately detect class 3 caries compared to DenseNet201. The performance metrics are summarized in Table 2. Graphical representations of accuracy, class-wise precision, and Comparative F1 scores can be found in Figures 21(a), 21(b), and 21(c) respectively. Confusion matrices, as well as plots illustrating training and validation accuracy, and training and validation loss for all models used in the study, are depicted in Figures 22-30.



(a)



(b)



(c)

Figure 21. (a) Comparison of accuracy (Phase II) (b) Comparison of class-wise precision (Phase II) (c) Comparison of class-wise F1 score (Phase II)

Table 2. Performance matrix for Phase II experimentation

Sr. No.	Model	Number of Epochs	Class 0	Class 1	Class 2	Class 3	Accuracy	Recall	F1 Score
1	ResNet50V2	30	0.90	0.71	0.58	0.73	0.73	0.72	0.72
2	DenseNet169	30	1.00	0.78	0.70	0.69	0.79	0.78	0.78
3	DenseNet201	30	0.75	0.78	0.67	0.91	0.79	0.75	0.72

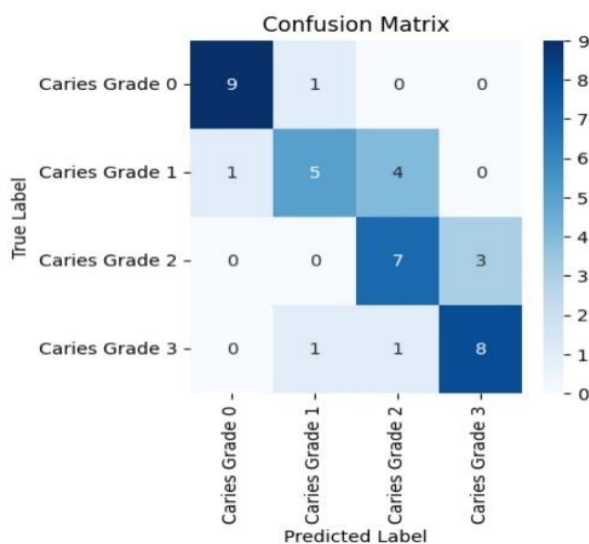


Figure 22. Confusion matrix for ResNet50V2

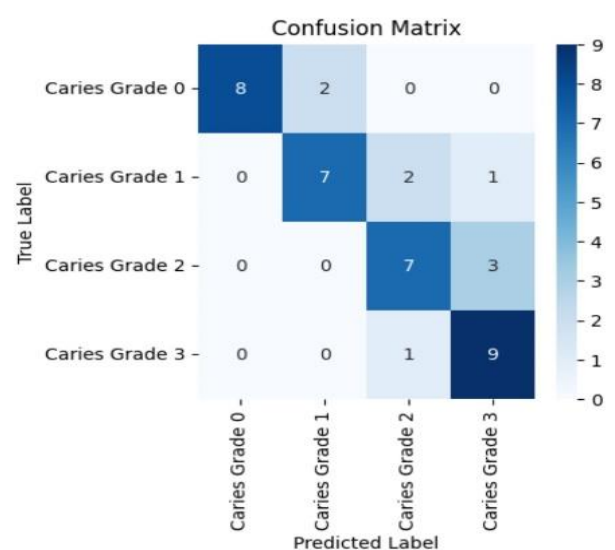


Figure 23. Confusion matrix for DenseNet169

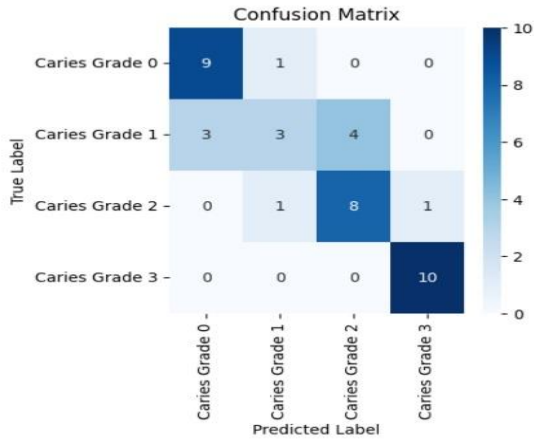


Figure 24. Confusion matrix for DenseNet201

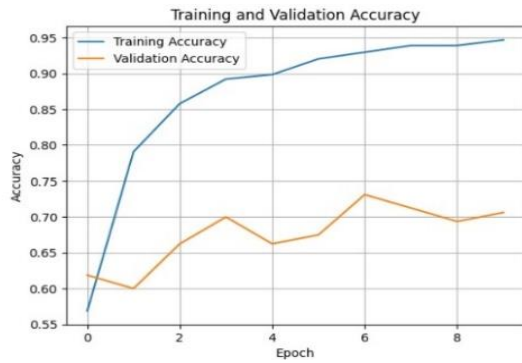


Figure 25. Training and validation accuracy for ResNet50V2

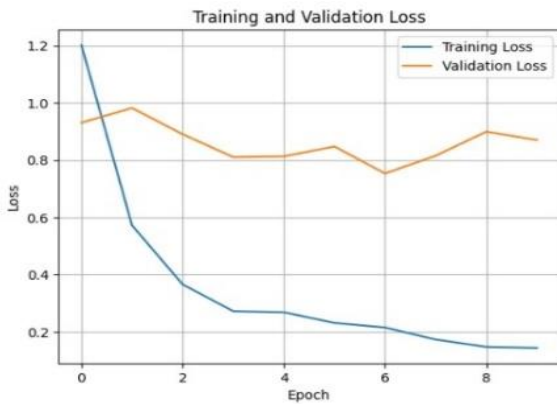


Figure 26. Training and validation loss for ResNet50V2

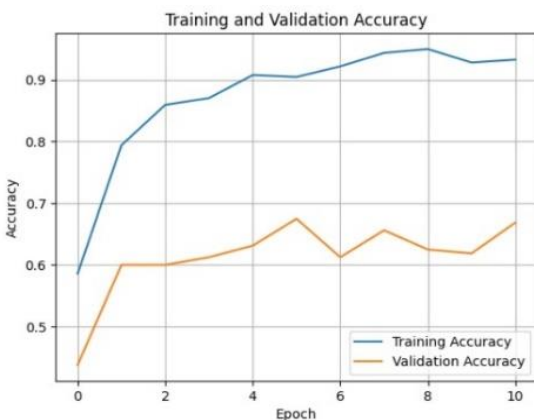


Figure 27. Training and validation accuracy for DenseNet169

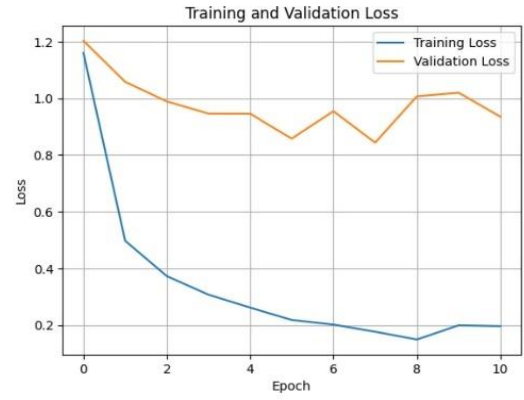


Figure 28. Training and validation loss for DenseNet169

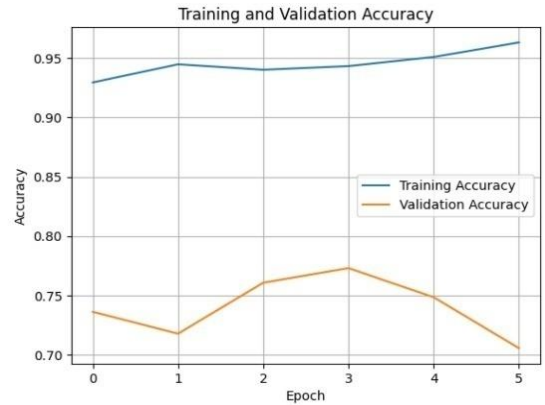


Figure 29. Training and validation accuracy for DenseNet201

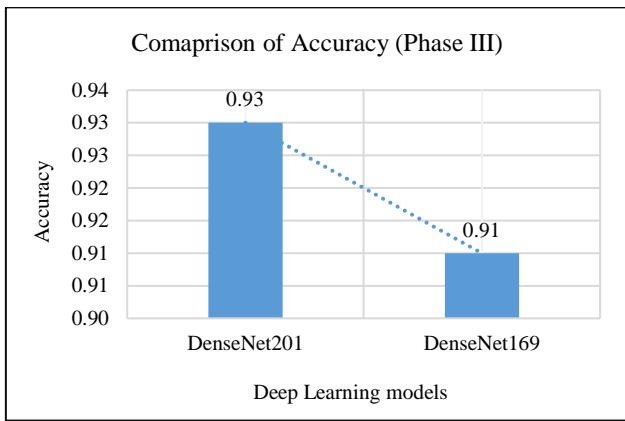


Figure 30. Training and validation loss for DenseNet201

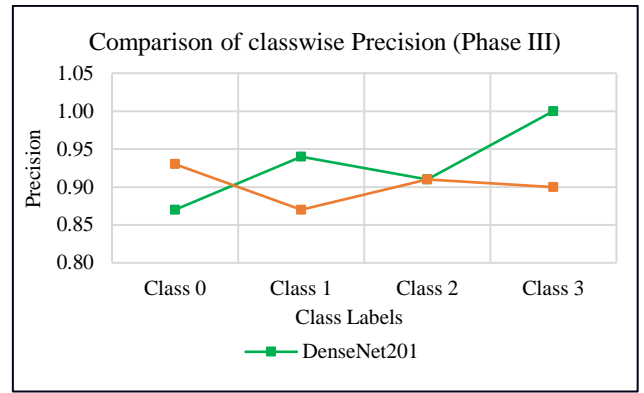
4.3 Phase III experimental results

The analysis of the initial two phases of experimentation indicated that DenseNet models exhibited superior performance in the early identification of dental caries. Consequently, only DenseNet models were utilized in this stage of the study. A greater number of images, including a higher volume of augmented images in classes 2 and 3, were incorporated to enhance the resilience of class 2 and class 3 detection and classification by the models. This augmentation of data proved to be effective, as DenseNet201 achieved an overall accuracy of 93%. Moreover, DenseNet201 demonstrated class-specific precision ranging from 0.87 to 1.00, marking the most favorable outcome attained across all phases of experimentation. On the other hand, DenseNet169 exhibited precision ranging from 0.82 to 0.91, with class 3 showing a lower accuracy of 0.82. Thus, based on the

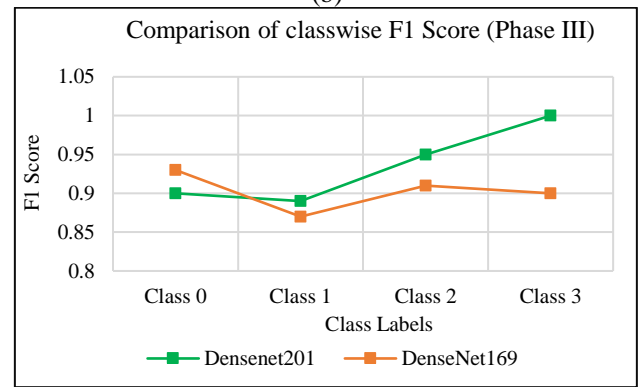
accumulated results, DenseNet201 emerged as the superior choice for early dental caries detection. The performance metrics are succinctly outlined in Table 3, while graphical representations showcasing accuracy, class-specific precision, and Comparative F1 scores can be found in Figures 31(a), 31(b), and 31(c) respectively. Additionally, confusion matrices and plots illustrating "training and validation accuracy, as well as training and validation loss," for all employed models, were presented in Figures 32-37. Table 4 shows comparison of mean confidence and standard deviation across dental caries classes. "DenseNet201" exhibits more stable and consistent confidence levels across all caries grades, while DenseNet169 shows slightly higher mean confidences but with some variability. This stability in DenseNet201 could make it a preferable choice for applications requiring reliable and consistent predictions across different classes.



(a)



(b)



(c)

Figure 31. (a) Class-wise accuracy (Phase III) (b) Comparison of class-wise precision (Phase III) (c) Comparison of class-wise F1 score (Phase III)

Table 3. Performance matrix for Phase III experimentation

Sr. No.	Model	Number of Epochs	Class 0	Class 1	Class 2	Class 3	Accuracy	Recall	F1 Score
1	DenseNet201	80	0.87	0.94	0.91	1.00	0.93	0.93	0.93
2	DenseNet169	80	0.93	0.89	1.00	0.82	0.91	0.91	0.91

Table 4. Comparison of mean confidence and standard deviation

Sr. No.	Model	Metric	Class 0	Class 1	Class 2	Class 3
1	DenseNet201	Mean Confidence	0.99	0.94	0.99	0.96
		Standard Deviation	0.02	0.12	0.03	0.09
2	DenseNet169	Mean Confidence	0.96	0.92	1.00	0.94
		Standard Deviation	0.09	0.14	0.00	0.10

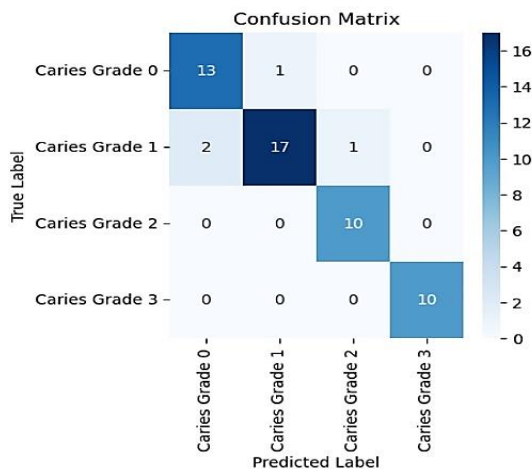


Figure 32. Confusion matrix for DenseNet201

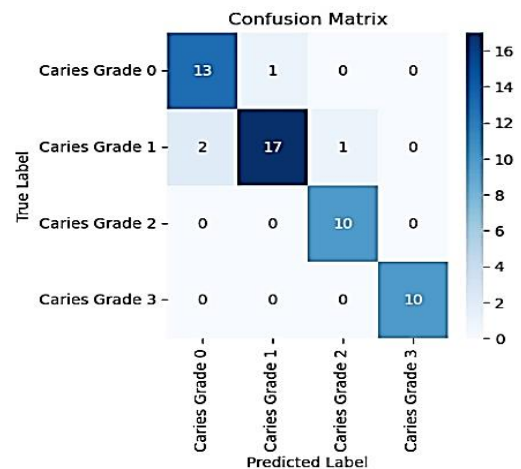


Figure 33. Confusion matrix for DenseNet169



Figure 34. Training and validation accuracy for DenseNet169

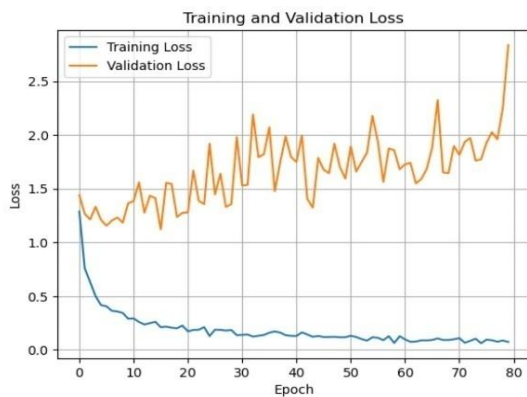


Figure 35. Training and validation loss for DenseNet169



Figure 36. Training and validation accuracy for DenseNet201

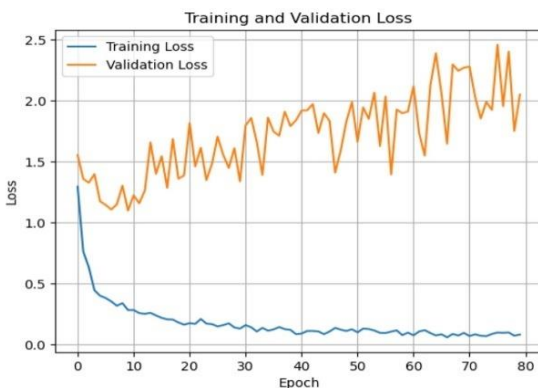


Figure 37. Training and validation loss for DenseNet201

Potential advantages of the modified DenseNet201 model

1. Efficient parameter utilization: DenseNet’s dense connectivity ensures efficient parameter usage, reducing redundant calculations and improving computational efficiency. This compactness contributes to faster training and inference times, which is beneficial for applications requiring real-time predictions.

2. Improved generalization: Utilizing pre-trained weights, along with robust “regularization techniques” such as “Dropout and Batch Normalization”, enhances the model’s generalization capabilities, even when data is limited. These techniques collectively improve the model’s resilience against overfitting, resulting in more reliable performance on unseen data.

3. Transfer learning and domain adaptation: The model’s architecture supports transfer learning, allowing adaptation to specific domains or datasets while leveraging the robust features learned from DenseNet201. Fine-tuning through dense layers ensures the capture of domain-specific features, thus boosting accuracy in niche applications.

4. Scalability and adaptability: The modular design of the architecture allows for easy modifications, such as adding more dense layers or adjusting dropout rates, to suit different datasets and problem complexities. This adaptability makes the model a versatile choice for various image classification tasks, ranging from medical imaging to natural scene analysis.

5. Resource efficiency: Global Average Pooling and Batch Normalization contribute to resource efficiency by minimizing memory usage and computational overhead, making the model suitable for deployment on devices with limited resources.

6. Computational efficiency: The model required 21min and 24.8 sec for training which is advantageous for clinical implementations where quick model updates or real-time training may be necessary.

Evaluation of the proposed model for optimum results

Table 5 shows the comparison of Densenet201 model variants.

1. Increasing the dropout rate from 0.2 to 0.5

The model with a dropout rate of 0.2 exhibits a high overall accuracy of 0.93, indicating strong overall performance and effective classification across all classes. The F1 scores for each class are consistently high, reflecting balanced precision and recall. Notably, the model achieves perfect recall (1.00) for class 3, suggesting it is very effective in identifying this class. Lower dropout rates tend to retain more information during training, which can be beneficial in capturing detailed features from the dataset. However increased dropout rate shows trade-off between regularization and model performance. Thus, the model with a dropout rate of $p=0.2$ appears to be the optimum choice based on overall performance metrics, including accuracy, recall, and F1 score.

2. Performance of 4BDD DenseNet201 variant

This model performed almost similar to DenseNet201 with 3BDD layers for advanced classes detection however it lagged behind in early detection of the dental caries resulting in 0.77 as class 1 accuracy. Also, overall accuracy reduced from 0.93 to 0.85. The main disadvantage of additional BDD layer is that it makes proposed model computationally less effective resulting in requirement of 28 minutes and 35.5 seconds for training. In real time applications it may affect lot in critical detection.

3. Modified DenseNet201 with learning rate scheduler

Overall accuracy is bit lower than the proposed model.

Since the lesser amount of dataset proposed model is overfitting hence learning rate scheduler was implemented but it is also showing same results of overfitting of the model as that of proposed model. so proposed model is suitable for critical detection of early detection of dental caries.

4. DenseNet201 with SGD optimizer

The obvious choice of optimizer other than Adam is SGD. But proposed model shows reduced effectiveness in class 2 (F1 score = 0.83). Even though, overall accuracy is similar to learning rate scheduler model, it does not show significant changes to the proposed model.

Table 5. Comparison of DenseNet201 model variants

Sr. No.	Model	Class 0	Class 1	Class 2	Class 3	Accuracy	Recall	F1 Score	Computational Complexity	Inference Time
1	DenseNet201 with 3BDD layer	0.87	0.94	0.91	1.00	0.93	0.93	0.93	Moderate	21min 24.8sec
2	DenseNet201 with 4 BDD layer	0.85	0.77	0.91	1.00	0.85	0.85	0.85	High	28 min 35.5sec
3	DenseNet201 with learning rate scheduler	0.93	0.94	0.91	0.83	0.91	0.91	0.91	Moderate	22min 36.4sec
4	DenseNet201 with p=0.5	0.81	0.89	1.00	0.90	0.89	0.89	0.89	Moderate	23min 24.8sec
5	DenseNet201 with SGD	0.87	0.94	0.83	1.00	0.91	0.91	0.91	Moderate	25min 25.6sec

5. CONCLUSION

This paper presents a comprehensive analysis of “deep learning architectures”, particularly residual and dense networks, applied to a novel dental dataset for the early detection of dental caries. Among the evaluated models, “DenseNet201” emerges as the optimal architecture, achieving an overall accuracy of 93% and excelling in both precision and sensitivity across all classes, especially in detecting advanced caries (class 3). The dense connections in “DenseNet” proved advantageous in handling small-scale objects like dental caries, offering superior performance compared to residual networks, which faced challenges with classifying early-stage caries as the number of training images increased.

Despite the promising results, the study is limited by the relatively small dataset size, which may hinder the model's ability to generalize to unseen data and introduce potential biases, such as class imbalance or demographic skew. Overfitting remains a concern, and the lack of validation across diverse populations limits the model's broader applicability. While computational complexity and inference time were considered, deploying the model in resource-constrained environments, such as mobile devices, remains a challenge.

Future research should focus on addressing these limitations by expanding the dataset to improve generalizability and reduce biases, exploring new architectures like transformers, and optimizing the model for mobile deployment to enable real-time, accessible dental caries detection in clinical and home settings. Furthermore, this framework could be extended to classify other dental diseases, making it a valuable tool for noninvasive, intelligent dental disease detection in preventive dentistry.

REFERENCES

[1] Çolak, H., Dülgergil, Ç.T., Dalli, M., Hamidi, M.M.

(2013). Early childhood caries update: A review of causes, diagnoses, and treatments. *Journal of Natural Science, Biology, and Medicine*, 4(1): 29-38. <https://doi.org/10.4103/0976-9668.107257>

[2] Anil, S., Anand, P.S. (2017). Early childhood caries: Prevalence, risk factors, and prevention. *Frontiers in Pediatrics*, 5: 157. <https://doi.org/10.3389/fped.2017.00157>

[3] Meyer, F., Enax, J. (2018). Early childhood caries: Epidemiology, aetiology, and prevention. *International Journal of Dentistry*, 2018(1): 1415873. <https://doi.org/10.1155/2018/1415873>

[4] Schmoeckel, J., Gorseta, K., Splieth, C.H., Juric, H. (2020). How to intervene in the caries process: Early childhood caries—A systematic review. *Caries Research*, 54(2): 102-112. <https://doi.org/10.1159/000504335>

[5] Corrêa-Faria, P., Viana, K.A., Raggio, D.P., Hosey, M.T., Costa, L.R. (2020). Recommended procedures for the management of early childhood caries lesions—A scoping review by the children experiencing dental anxiety: Collaboration on research and education (CEDACORE). *BMC Oral Health*, 20: 1-11. <https://doi.org/10.1186/s12903-020-01067-w>

[6] Poureslami, H.R., Van Amerongen, W.E. (2009). Early childhood caries (ECC): An infectious transmissible oral disease. *The Indian Journal of Pediatrics*, 76: 191-194. <https://doi.org/10.1007/s12098-008-0216-1>

[7] Astuti, E.S.Y., Sukrama, I.D.M., Mahendra, A.N. (2019). Innate immunity signatures of early childhood caries (ECC) and severe early childhood caries (S-ECC). *Biomedical and Pharmacology Journal*, 12(3): 1129-1134. <http://doi.org/10.13005/bpj/1740>

[8] Hu, S., Sim, Y.F., Toh, J.Y., Saw, S.M., Godfrey, K.M., Chong, Y.S., Yap, F., Lee, Y.S., Shek, L.P., Tan, K.H., Chong, M.F., Hsu, C.Y.S. (2019). Infant dietary patterns and early childhood caries in a multi-ethnic Asian cohort. *Scientific Reports*, 9(1): 852. <https://doi.org/10.1038/s41598-018-37183-5>

[9] Alshunaiber, R., Alzaid, H., Meaigel, S., Aldeeri, A.,

- Adlan, A. (2019). Early childhood caries and infant's oral health; pediatricians' and family physicians' practice, knowledge and attitude in Riyadh city, Saudi Arabia. *The Saudi Dental Journal*, 31: S96-S105. <https://doi.org/10.1016/j.sdentj.2019.01.006>
- [10] Ganesh, A., Sampath, V., Sivanandam, B.P., Sangeetha, H., Ramesh, A. (2020). Risk factors for early childhood caries in toddlers: An institution-based study. *Cureus*, 12(4): 7516. <https://doi.org/10.7759/cureus.7516>
- [11] Patil, S., Kulkarni, V., Bhise, A. (2019). Algorithmic analysis for dental caries detection using an adaptive neural network architecture. *Heliyon*, 5(5): e01579. <https://doi.org/10.1016/j.heliyon.2019.e01579>
- [12] Sun, M.L., Liu, Y., Liu, G., Cui, D., Heidari, A.A., Jia, W.Y., Chen, H., Luo, Y. (2020). Application of machine learning to stomatology: A comprehensive review. *IEEE Access*, 8: 184360-184374. <https://doi.org/10.1109/ACCESS.2020.3028600>
- [13] Stratigaki, E., Jost, F.N., Kühnisch, J., Litzenburger, F., Lussi, A., Neuhaus, K.W. (2020). Clinical validation of near-infrared light transillumination for early proximal caries detection using a composite reference standard. *Journal of Dentistry*, 103: 100025. <http://doi.org/10.1016/j.jjodo.2020.100025>
- [14] Divakaran, S., Vasanth, K., Suja, D., Swedha, V. (2021). Classification of digital dental X-ray images using machine learning. In 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII), pp. 1-3. <https://doi.org/10.1109/ICBSII51839.2021.9445171>
- [15] Park, Y.H., Choi, Y.Y. (2022). Feeding practices and early childhood caries in Korean preschool children. *International Dental Journal*, 72(3): 392-398. <https://doi.org/10.1016/j.identj.2021.07.001>
- [16] Hung, M., Voss, M. W., Rosales, M. N., Li, W., Su, W., Xu, J., Bounsanga, J., Ruiz-Negrón, B., Lauren, E., Licari, F.W. (2019). Application of machine learning for diagnostic prediction of root caries. *Gerodontology*, 36(4): 395-404. <https://doi.org/10.1111/ger.12432>
- [17] Chitnis, G., Bhanushali, V., Ranade, A., Khadase, T., Pelagade, V., Chavan, J. (2020). A review of machine learning methodologies for dental disease detection. *IEEE India Council International Subsections Conference (INDISCON)*, pp. 63-65. <https://doi.org/10.1109/INDISCON50162.2020.00025>
- [18] Schwendicke, F., Golla, T., Dreher, M., Krois, J. (2019). Convolutional neural networks for dental image diagnostics: A scoping review. *Journal of Dentistry*, 91: 103226. <https://doi.org/10.1016/j.jdent.2019.103226>
- [19] Cantu, A.G., Gehrung, S., Krois, J., Chaurasia, A., Rossi, J.G., Gaudin, R., Elhennawy, K., Schwendicke, F. (2020). Detecting caries lesions of different radiographic extension on bitewings using deep learning. *Journal of Dentistry*, 100: 1034252020. <https://doi.org/10.1016/j.jdent.2020.103425>
- [20] Chen, Y.W., Stanley, K., Att, W. (2020). Artificial intelligence in dentistry: Current applications and future perspectives. *Quintessence International*, 51(3): 248-257. <https://doi.org/10.3290/j.qi.a43952>
- [21] Javid, A., Rashid, U., Khattak, A.S. (2020). Marking early lesions in labial colored dental images using a transfer learning approach. In 2020 IEEE 23rd International Multitopic Conference (INMIC), pp. 1-5. <https://doi.org/10.1109/INMIC50486.2020.9318173>
- [22] Leo, L.M., Reddy, T.K. (2021). Learning compact and discriminative hybrid neural network for dental caries classification. *Microprocessors and Microsystems*, 82: 103836. <https://doi.org/10.1016/j.micpro.2021.103836>
- [23] Myint, Z.C.K., Zaitsu, T., Oshiro, A., Ueno, M., Soe, K.K., Kawaguchi, Y. (2020). Risk indicators of dental caries and gingivitis among 10–11-year-old students in Yangon, Myanmar. *International Dental Journal*, 70(3): 167-175. <https://doi.org/10.1111/idj.12537>
- [24] Uoshima, K., Akiba, N., Nagasawa, M. (2021). Technical skill training and assessment in dental education. *Japanese Dental Science Review*, 57: 160-163. <https://doi.org/10.1016/j.jdsr.2021.08.004>
- [25] Verma, D., Puri, S., Prabhu, S., Smriti, K. (2020). Anomaly detection in panoramic dental x-rays using a hybrid deep learning and machine learning approach. In 2020 IEEE region 10 conference (TENCON), pp. 263-268. <https://doi.org/10.1109/TENCON50793.2020.9293765>
- [26] Kumar, A., Bhadauria, H.S., Singh, A. (2021). Descriptive analysis of dental X-ray images using various practical methods: A review. *PeerJ Computer Science*, 7: e620. <https://doi.org/10.7717/peerj-cs.620>
- [27] Chen, H., Zhang, K., Lyu, P., Li, H., Zhang, L., Wu, J., Lee, C.H. (2019). A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films. *Scientific Reports*, 9(1): 3840. <https://doi.org/10.1038/s41598-019-40414-y>
- [28] Tuzoff, D.V., Tuzova, L.N., Bornstein, M.M., Krasnov, A.S., Kharchenko, M.A., Nikolenko, S.I., Sveshnikov, M.M., Bednenko, G.B. (2019). Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofac Radiology*, 48(4): 20180051. <https://doi.org/10.1259/dmfr.20180051>
- [29] Reyes, L.T., Knorst, J.K., Ortiz, F.R., Ardenghi, T.M. (2021). Scope and challenges of machine learning-based diagnosis and prognosis in clinical dentistry: A literature review. *Journal of Clinical and Translational Research*, 7(4): 523. <https://doi.org/10.18053/jctres.07.202104.012>
- [30] Musri, N., Christie, B., Ichwan, S.J.A., Cahyanto, A. (2021). Deep learning convolutional neural network algorithms for the early detection and diagnosis of dental caries on periapical radiographs: A systematic review. *Imaging Science in Dentistry*, 51(3): 237. <https://doi.org/10.5624/isd.20210074>
- [31] Zhang, Y., Liao, H., Xiao, J., Jallad, N.A., Ly-Mapes, O., Luo, J. (2020). A smartphone-based system for real-time early childhood caries diagnosis. In *Medical Ultrasound, and Preterm, Perinatal and Paediatric Image Analysis: First International Workshop, ASMUS 2020, and 5th International Workshop, PIPPI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru*, pp. 233-242. https://doi.org/10.1007/978-3-030-60334-2_23
- [32] Rashid, U., Javid, A., Khan, A., Liu, L., Ahmed, A., Khalid, O., Saleem, K., Meraj, S., Iqbal, U., Nawaz, R. (2022). A hybrid mask RCNN-based tool to localize dental cavities from real-time mixed photographic images. *PeerJ Computer Science*, 8: e888. <https://peerj.com/articles/cs-888/>
- [33] Ragodos, R., Wang, T., Padilla, C., Hecht, J.T., Poletta, F.A., Orioli, I.M., Howe, B.J. (2022). Dental anomaly detection using intraoral photos via deep learning. *Scientific Reports*, 12(1): 11577.

https://doi.org/10.1038/s41598-022-15788-1

[34] Askar, H., Krois, J., Rohrer, C., Mertens, S., Elhennawy, K., Ottolenghi, L., Mazur, M., Paris, S., Schwendicke, F. (2021). Detecting white spot lesions on dental photography using deep learning: A pilot study. *Journal of Dentistry*, 107: 103615. <https://doi.org/10.1016/j.jdent.2021.103615>

[35] Li, W., Liang, Y., Zhang, X., Liu, C., He, L., Miao, L., Sun, W. (2021). A deep learning approach to automatic gingivitis screening based on classification and localization in RGB photos. *Scientific Reports*, 11(1): 16831. <https://doi.org/10.1038/s41598-021-96091-3>

[36] Kühnisch, J., Meyer, O., Hesenius, M., Hickel, R., Gruhn, V. (2022). Caries detection on intraoral images using artificial intelligence. *Journal of Dental Research*, 101(2): 158-165. <https://doi.org/10.1177/00220345211032524>

[37] Thanh, M.T.G., Van Toan, N., Ngoc, V.T.N., Tra, N.T., Giap, C.N., Nguyen, D.M. (2022). Deep learning application in dental caries detection using intraoral photos taken by smartphones. *Applied Sciences*, 12(11): 5504. <https://doi.org/10.3390/app12115504>

[38] Lee, W.F., Day, M.Y., Fang, C.Y., Nataraj, V., Wen, S.C., Chang, W.J., Teng, N.C. (2024). Establishing a novel deep learning model for detecting peri-implantitis. *Journal of Dental Sciences*, 19(2): 1165-1173. <https://doi.org/10.1016/j.jds.2023.11.017>

[39] Huang, G., Liu, Z., Maaten, L.V.D., Weinberger, K.Q. (2017). Densely connected convolutional networks. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 2261-2269. <https://doi.org/10.1109/CVPR.2017.243>

NOMENCLATURE

$I_i(x,y)$	Image $i = 1, 2, 3, 4$ in pixels
D	Dataset
$I_r(x',y')$	Rescaled image
$X^{(i)}$	Augmented input image
$X_{Base}^{(i)}$	Feature map from DenseNet201
$GAP(f)$	Global average pooling function

$X_{Pooled}^{(i)}$	Pooled feature map
$X_{Normalized1}^{(i)}$	Batch-normalized pooled feature map
$X_{Dropped1}^{(i)}$	Feature map after dropout
$X_{Dense1}^{(i)}$	Dense layer output (ReLU)
$X_{Normalized2}^{(i)}$	Batch-normalized dense layer output
$X_{Dropped2}^{(i)}$	Feature map after dropout
$X_{Dense2}^{(i)}$	Second dense layer output
$X_{Normalized3}^{(i)}$	Batch-normalized second dense output
$X_{Dropped3}^{(i)}$	Feature map after dropout
$Y_{Output}^{(i)}$	Output from softmax layer

Greek symbols

γ	Scale parameter in batch normalization
β	Shift parameter in batch normalization
μ	Mean used for normalization
σ^2	Variance used for normalization
$\phi(x_i)$	Softmax function
ϵ	Small constant added to variance

Subscripts

i	Index of the image in the dataset
x	Horizontal pixel coordinate
y	Vertical pixel coordinate
x'	Transformed horizontal pixel coordinate
y'	Transformed vertical pixel coordinate
Base	Feature map from the base model
Pooled	Feature map after global average pooling
Normalized1,2,3	First/second/third batch-normalized
Dropped1,2,3	Feature map after First/second/third dropout
Dense1,2	Output of the First/second dense layer
Output	Output from the final softmax layer