Annals of Operations Research
https://doi.org/10.1007/s10479-022-04961-4

ORIGINAL RESEARCH

Periapical dental X-ray image classification using deep neural networks

Dipit Vasdev[1] · Vedika Gupta[2] · Shubham Shubham[1] · Ankit Chaudhary[1] · Nikita Jain[1] · Mehdi Salimi[3,6] · Ali Ahmadian[4,5]

Accepted: 29 August 2022

Abstract
This paper studies the problem of detection of dental diseases. Dental problems affect the vast majority of the world's population. Caries, RCT (Root Canal Treatment), Abscess, Bone Loss, and missing teeth are some of the most common dental conditions that affect people of all ages all over the world. Delayed or incorrect diagnosis may result in mistreatment, affecting not only an individual's oral health but also his or her overall health, thereby making it an important research area in medicine and engineering. We propose a pipelined Deep Neural Network (DNN) approach to detect healthy and non-healthy periapical dental X-ray images. Even a minor enhancement or improvement in existing techniques can go a long way in providing significant health benefits in the medical field. This paper has made a successful attempt to contribute a different type of pipelined approach using AlexNet in this regard. The approach is trained on a large dataset of 16,000 dental X-ray images, correctly identifying healthy and non-healthy X-ray images. We use an optimized Convolutional Neural Networks and three state-of-the-art DNN models, namely Res-Net-18, ResNet-34, and AlexNet for disease classification. In our study, the AlexNet model outperforms the other models with an accuracy of 0.852. The precision, recall and F1 scores of AlexNet also surpass the other models with a score of 0.850 across all metrics. The area under ROC curve also signifies that both the false-positive rate and false-negative rate are low. We conclude that even with a big data set and raw X-ray pictures, the AlexNet model generalizes effectively to previously unseen data and can aid in the diagnosis of a variety of dental diseases.

✉ Mehdi Salimi
msalimi@stfx.ca

✉ Ali Ahmadian
ahmadian.hosseini@unirc.it

1 Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, New Delhi, India

2 Jindal Global Business School, O.P. Jindal Global University, Sonipat, Haryana 131001, India

3 Department of Mathematics and Statistics, St. Francis Xavier University, Antigonish, NS, Canada

4 Department of Law, Economics and Human Sciences and Decisions Lab, Mediterranea University of Reggio Calabria, 89125 Reggio Calabria, Italy

5 Department of Mathematics, Near East University, Nicosia, TRNC, Mersin 10, Turkey

6 Center for Dynamics, Faculty of Mathematics, Technische Universität Dresden, Dresden, Germany

Published online: 15 September 2022

Keywords AlexNet · Convolutional Neural Network (CNN) · Dental · Periapical · ResNet · X-ray

1 Introduction

Dentistry or dental science is a branch of science that focuses on the study, diagnosis, prevention, and treatment of diseases and disorders related to oral health. Oral diseases have a high prevalence across the globe, making them an ideal subject for research. Dental clinical quality assessment is based on numerous quantitative measures, such as the number of yearly dental visits of people of all ages, including children, youngsters, and adults; new caries among children, endodontics; and extractions method proportion (Dental Quality Alliance, 2012). While these quality assessments evaluate dental clinical quality at the full-scale level, we likewise need to undertake subjective evaluations at the miniature level, for instance, to understand whether the treatment undertaken for a tooth produces the desired results. Traditionally, dental experts would assess the patients' dental condition based on the previous treatment undertaken by comparing changes between the dental images taken before and after the treatment.

Machine learning algorithms find an application in all fields of medicine-disease classification, disease prediction (Chatrati et al., 2020) and detection (Mishra et al., 2020), drug discovery, and medical decision making, thus, significantly changing the ways of practicing medicine. Such all-encompassing involvement in machine learning in medicine has paved paths for new research areas, such as computational biology, biochemistry, computer-assisted surgery, and many others (Ren et al., 2019; Jain et al. 2019; Gupta, 2019a). The success of machine learning algorithms in healthcare in recent years comes at a fortunate time when clinical records are progressively being digitalized.

One of the characterizing highlights of present-day medical services activity is that it produces enormous data identified with various medical tests conducted to rule out possible infections. Among various medical services fields, clinical images produce the most significant volume of information. Also, this data is growing exponentially because the instruments are improving at capturing information. These captured bits of information carry valuable insights about the patient's condition, the development of the sickness/anomalies, and the treatment's success. Each piece adds to the entire and is critical to assemble all into a big picture precisely as could be expected under the circumstances.

Notwithstanding, the extent of information frequently outreaches the potential outcomes of the customary investigation. Medical practitioners may not be able to analyze such a lot of information, owing to workforce limitations. In particular, this is significant because interpreting medical data is essential in fields like clinical picture examination. The other issue with the human investigation is that it is restricted and inclined to mistakes because of different elements—stress, lack of experience, and low expertise. Along these lines, deep learning is an appropriate solution to the problem. Deep learning applications can process data and extract critical insights at early stages with much more accuracy. This can help doctors to analyze data and examine test results more thoroughly. Implementing deep learning in healthcare services is an effective way to build the effectiveness of activity and precision results.

Image classification through deep learning frameworks has been dominant in the field of computer vision and imaging. The reason behind this is the ability of DNN and its variants to learn features hierarchically (Esteva et al., 2017). The current state of art gives several variants of DNN architecture, with the most popular ones being AlexNet (Han et al., 2017) and ResNet (Sonoda & Murata, 2019). Previous works (Tuzoff et al., 2019; Anantharaman et al., 2018; Yang et al., 2018 ; Prajapati et al. 2017) have been limited by dataset-specific architecture, resulting from a dataset dependency in terms of small size or region or a particular disease classification. This work transforms experts' observations into labels and provides an artificially intelligent approach to detect any common disease in a given dental scan automatically. Thus, it helps in places where there is still a scarcity of any experienced radiologist or equally competent expert.

In this paper, we collected a dataset of 16,000 periapical dental X-Ray images. The dataset comprises annotated X-Ray images under dentists' supervision divided into classes of commonly occurring diseases. The annotated dataset thus created is evaluated using the proposed modified pipelined approach for binary classification of tooth disease. Four different DNN models (AlexNet, ResNet-18, and ResNet-34, along with an optimized CNN) are deployed and compared. The performance is analyzed based on different standard performance metrics. The best performing model can provide its application in the automated preliminary examination of periapical dental X-rays, which will also help the medical community. The deployed pipelined approach gives an accuracy of 0.852 when implemented on the selected dataset. The dataset's novelty is that the images used have not been released earlier and collected by a regional dental clinic over its service course. To support the dental community in making clinical decisions, we propose a pipelined approach: periapical dental classification using DNN. for automated clinical quality evaluation.

Artificial Intelligence(AI) has been prominently used in various segments of society—ranging from analyzing wind speed behavior (Mostafaeipour et al., 2021) to prediction of marketing products(Goli et al., 2021). With the advancement of AI methods, medical practitioners are progressing more towards adapting and implementing AI methods in various medical fields. In working towards reducing computing time (Ranjbarzadeh et al., 2021), Cascade Convolutional Neural Network (C–CNN) along with Distance Wise Attention(DWA) were used to prevent overfitting. Two – route convolutional neural network(CNN) for identifying and categorizing COVID–19 infections were used for CT images(Ranjbarzadeh et al., 2021). A new machine learning strategy based on modified deep learning (DL) to determine the location of a tumor in a breast cancer patient(Jafarzadeh Ghoushchi et al., 2021) is also considered.

The motivation behind the study conducted in this paper can be derived from the evergrowing requirements of science and medical fields to build upon computational models and propose effective mechanisms to detect and efficiently handle human diseases. It is estimated that oral diseases affect close to 3.5 billion people worldwide, with caries of permanent teeth being the most common condition.[1] Therefore, the main objective of this paper is to propose a pipelined approach using the outperforming DNN model for efficient detection of dental diseases.

This paper presents a pipelined approach using AlexNet model. The reason for selection of AlexNet is that AlexNet has been known to excel at image classification tasks(Kumar et al., 2021), particularly when working with large real-world data sets such as the one used in this study. This is because it is capable of extracting both deep and baseline visual features. First, we developed a custom CNN model by tweaking and modifying its layers. We observed that the model has difficulty working on large high resolution data sets without over fitting. AlexNet solves this problem by combining model ensembles and consecutive convolution layers. Moreover, AlexNet requires fewer computations as compared to other DNNs

[1] https://www.who.int/news-room/fact-sheets/detail/oral-health.

like ResNet-18 and ResNet-34 which require intensive computations involving Graphics Processing Units (GPUs).

In this paper, we perform the following:

- We present a comprehensive workflow depicting the classification process of images in dentistry and medical sciences.
- We employed a large dataset of 16,000 dental X-ray images.
- Next, we implement most popular DNN models and compared them based on various performance measures.
- We propose a pipelined approach using AlexNet to correctly identify healthy and non-healthy X-ray images.

The best performing model i.e. AlexNet can find applications in automated dental disease classification, minimizing the possibilities of human error.

The rest of the paper is organized as follows: Sect. 2 presents the existing works on periapical dental X-ray image classification. Section 3 discusses the paper's contributions. Section 4 discusses the materials and methods employed to conduct the research. Section 5 shows the conducted experimental analysis and results' evaluation. Section 6 concludes the paper with an explanation of the obtained results and their assessment with the baseline methods with further enhancement.

2 Related work

This section presents comprehensive coverage of existing related works in the field of dental image classification. Deep learning has developed in recent years and can automatically extract image features using the original pixel information as input. To achieve accurate classification in dental X-rays, several researchers have contributed distinct image-processing algorithms. Figure 1 presents the taxonomical division of the reviewed literature in brief.

Gupta et al. (2019b) proposed a CNN, which is a deep learning-based framework for classifying dysplastic tissue images. The four-way classification classified images into four classes, viz. normal tissue, mild dysplastic tissue, moderate dysplastic tissue, and severe dysplastic tissue. From the initial dataset of 672 images collected from 52 patients, 2688 images were created based on the pre-processing and data augmentation. The initial dataset consisted of images having a squamous epithelial layer of the oral cavity taken from biopsy samples. The proposed framework achieved an accuracy of 91.65% for trained data and 89.3% for test data. The framework results were also compared with manual results produced by the medical experts and thus validated.

Ren et al. (2019) proposed an efficient medical lesion image classification method. It was a super-resolution based on the deep residual networks method. Initially, a large scale super-resolution model was trained. The training procedure was based on a deep residual network. Secondly, a user-friendly interface was designed. Different scaled medical lesion images were reconstructed from the high-scale reconstruction model. Peak signal to noise ratios as well as structural similarity index values were calculated for all reconstructed images. The exploratory outcomes show that the proposed super-resolution remaking strategy accomplished better execution over different strategies thought about in this work.

In (Tuzoff et al., 2019), Tuzoff D. et al. introduced a novel arrangement dependent on convolutional neural systems (CNNs) that perform tasks for all-encompassing radiographs. The system was trained on a dataset comprising of 1352 chosen radiographs. Both tooth detection, as well as numbering studies, was performed and analyzed. The tooth detection

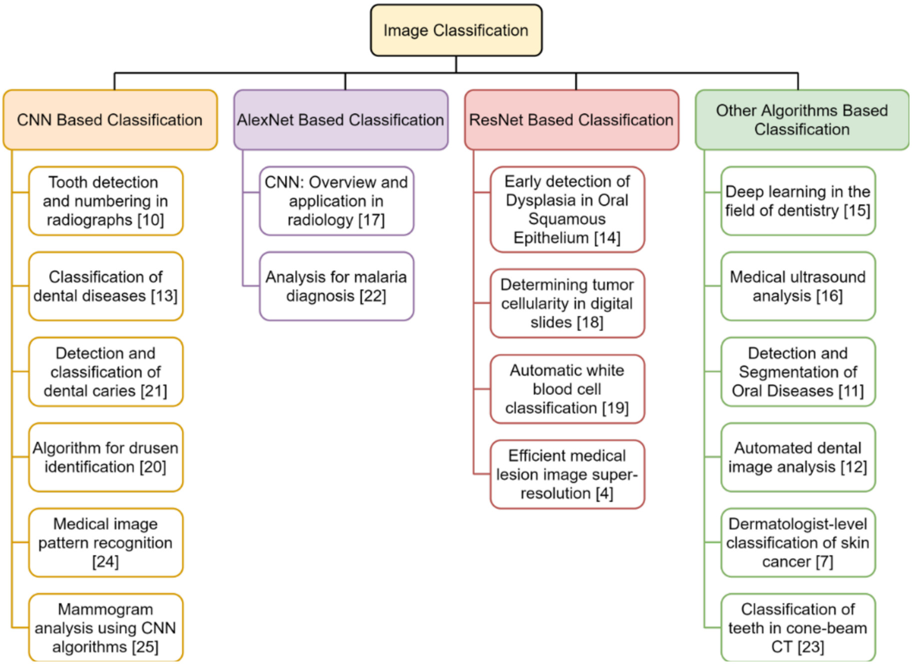


Fig. 1 Categorization of current state-of-the-art based on the technique used

module process and marked the boundaries of each tooth using Faster CNN architecture. The numbering module, however, classified the detected marked tooth using the FDI notation. It used the traditional VGG-16 CNN and the heuristic calculation to improve results as indicated by the principles for the spatial course of teeth' action. For detection, the system achieved a sensitivity of 0.9941 and a precision of 0.9945. Similarly, for numbering, a sensitivity of 0.9800 and specificity was 0.9994 were achieved. The step-by-step investigation demonstrated that the proposed method makes some errors brought about by comparable factors as those for specialists. The exhibition of the proposed automatic diagnosis helped determination arrangement was practically identical to the degree of specialists.

Artificial intelligence, commonly termed AI (Artificial Intelligence) represented by deep learning, can solve dentistry's major problems. Hwang et al. (Hwang et al., 2019) brought in an artificially intelligent view to dentistry. In this work, a comprehensive survey highlighting the recent discoveries in oral and maxillofacial radiology was presented. This survey's resources were collected from some popular databases such as PubMed, IEEE, and Scopus. The factors from 25 articles included system design, preparing information, assessment result, upsides and downsides, study item, and imaging methodology CNN was utilized as a principle organize segment.

Liu et al. (2019) concisely presented a few famous profound learning models. Afterward, they outlined and altogether talked about their applications in different explicit assignments in US picture investigation, such as grouping, recognition, and division. Finally, the open difficulties and potential patterns of things utilizing profound learning in restorative US picture examinations were discussed.

Anantharam et al. (2018) demonstrated the utilization of Mask-RCNN, the cutting edge convolutional neural system calculation for object recognition and division to the oral pathology space. Cover RCNN was initially produced for object identification and article occasion division of normal pictures. This test indicated that Mask-RCNN could likewise be utilized in a particular zone, for example, oral pathology. This paper planned for distinguishing and fragmenting mouth blisters and infections.

Yamashita et al. (2018) gave a profound view on CNN, a class of counterfeit neural systems that have gotten predominant in different computer- vision undertakings, pulling in enthusiasm over an assortment of spaces, including radiology. This audit article offered a point of view on CNN's essential ideas and the application of the same in different radiological assignments; it also talked about its difficulties and the further scope in the field of radiology. Yang et al. (2018) presented the datasets, techniques, and results coordinated to survey dental treatment attributes utilizing periapical dental X-beam pictures taken when the activities. To help dental specialists to settle on clinical choices, they proposed an instrument pipeline for robotized clinical quality assessment. A dataset built from 196 patients was used as an input. The dataset consisted of periapical dental radiography images. The images were annotated as "getting better," "getting worse," and "have no explicit change" by the assigned dental specialists. Their proposition incorporated a programmed technique with the medicinal information to edit the ROIs for clinical assessment—the apical touching locale. The presented methodology accomplished the F1 score of 0.749, equivalent to master dental specialists and radiologists' presentation.

Prajapati et al. (2017) proposed an accurate classification of diseases. A labeled dataset consisting of 251 Radio Visio-Graphy (RVG) X-ray images of three different classes was used for classification. In this paper, they investigated the exhibition of CNN to conclude the small marked dental dataset. Likewise, to improve accuracy, transfer learning was used. The paper presented experimental results for three different architectures of CNN. Akbar et al. (2018) proposed a strategy for deciding tumor cellularity in advanced slides utilizing profound learning procedures. They prepared ResNet models' progression to yield both discrete and persistent qualities and contrasted their results and scores obtained physically by a specialist pathologist. Their setups were approved on a dataset of picture patches removed from computerized slides, each containing different tumor cellularity degrees. The results obtained a precision score of 0.76. While recording tumor cellularity scores on a constant scale, ResNet showed great connections with physically distinguished scores. This indicated the potential for reproducible processing scores that are predictable with master assumptions using deep learning procedures. Habibzadeh et al. (2018) give a record of assessing white platelet differential checks through the PC-supported conclusion framework and hematology rules. We look to decide a quick, precise component for grouping and assemble data about the appropriation of white blood confirmations, which may analyze the level of any variations from the norm during the test conducted. In this work, they consider pre-handling and directed white platelets' issue into their four essential sorts, including neutrophils, eosinophils, lymphocytes, and monocytes, utilizing a continuous proposed profound learning structure.

Esteva et al. (2017) showed skin sores' arrangement utilizing a solitary CNN, prepared start to finish from pictures legitimately, utilizing just pixels and sickness names as data sources. They designed a CNN utilizing a dataset of 129,450 clinical pictures—two sets of size bigger than past datasets12—comprising 2032 distinct maladies. The CNN accomplished execution was keeping pace with every single tried master crosswise over the two errands, showing human-made brainpower fit for grouping skin malignant growth with a degree of ability practically identical to dermatologists.

Checco et al. (2006) proposed a cell neural system-based calculation for drusen recognizable proof in fundus photo. The calculation was made by various picture preparing steps: clamor decrease, histogram standardization, and a novel method of versatile division. Ali et al. (2016) proposed another strategy to recognize and distinguish dental caries utilizing X-beam pictures as a dataset and a profound neural system as a method. This system depended on a stacked meager auto-encoder and a softmax classifier. Those procedures, scanty auto-encoder, and softmax were utilized to prepare a profound neural system. This methodology was tried on a genuine dataset and has exhibited a decent presentation of discovery. Zhaohui Liang et al. (2016) proposed another and strong AI model dependent on a convolutional neural system to arrange single cells in meager blood, consequently spreading on standard magnifying instrument slides tainted or uninfected. In a ten-fold cross-approval dependent on 27,578 single cell pictures, their new 16-layer CNN model's normal precision is 97.37%.

Miki et al. (2016) explored using a profound CNN for arranging tooth types on dental cone-shaft figured tomography (CT) pictures. Areas of intrigue (ROIs), including single teeth, were removed from CT cuts. ROIs obtained from the preparation cases were utilized for preparing the DCNN. The proposed AlexNet used five convolutional layers, with three pooling layers and two association layers for diminishing the overtraining sway; they extended the data by turning the picture. The test ROIs were characterized into seven-tooth types by the prepared system. The normal order precision utilizing the expanded preparing information by picture turns and force change was 88.8%. Contrasted and the outcome without information growth, information enlargement brought about around 5% improvement in arrangement exactness. This shows the further improvement can be normal by extending the CT dataset. Lo et al. (1995) built up a few preparing techniques related to a convolution neural system for general medicinal picture design acknowledgment. A whimsical technique for utilizing turn and move invariance is likewise proposed to upgrade the neural net exhibition. Weighting coefficients of convolution pieces were shaped by the neural system through the back spread, preparing for this fake neural net. Also, radiologists' perusing technique was demonstrated to train the counterfeit neural system to perceive the predefined picture designs and those important to specialists. Their preparation systems included (a) radiologists' appraising for each presumed picture territory, (b) backpropagation of summed up dispersion, (c) coach forced capacities, (d) move and turn invariance of analysis understanding, and (e) consistency of clinical info information utilizing proper foundation decrease capacities.

Lizka et al. (1995) proposed analogic CNN calculations to identify the bosom disease's major neurotic indications in x-beam mammogram pictures. The detailed calculations could recognize the smaller scale calcifications and the spicule around a given tumor portion.

In their work to reduce the computational complexity Ramin Ranjbarzadeh et al. (2021) they proposed a methodology to work only on small section of the image rather than whole image. This reduces computational complexity while also preventing overfitting. They proposed a Cascade Convolutional Neural Network (C-CNN) and a Distance Wise Attention (DWA) mechanism based on this. The model was then validated using the BRATS 2018 dataset. The results were excellent, with whole tumor, enhancing tumor, and tumor core dice scores of 0.9203, 0.9113, and 0.8726, respectively.

One of the fastest ways to diagnose patients is to use lung CT scan images to detect the condition. Finding sick tissues and segmenting them from CT slices is difficult due to comparable neighboring tissues, ambiguous boundaries, and unpredictable infections. Ramin Ranjbarzadeh et al. (2021) propose a two-route convolutional neural network (CNN) for identifying and categorizing COVID-19 infection from CT images by extracting global and local properties. To represent the input image differently, researchers used two alternative strategies: fuzzy—means clustering and local directional pattern (LDN) encoding approaches.

The results showed that the proposed framework had a precision of 96 percent, a recall of 97 percent, and a F-score of 97 percent.

One of the most frequent and dangerous tumors in women is breast cancer. Convolutional networks are a modern machine learning (ML) model that conducts picture segmentation in one learning step for the goal of detecting signals of tumor formation. Saeid Jafarzadeh Ghoushchi et al. (2021) creates a new machine learning strategy based on modified deep learning (DL) to determine the location of a tumor in a breast cancer patient. The authors use two sophisticated network-based designs (VGG) (9 layers and 2.9 million parameters) in the next study work, while the remaining networks (10 layers and 0.9 million parameters) are used in the following trials. Convolutional neural networks (CNNs) appear to be a viable choice for separating breast cancer lesions, according to the findings.

Three hybrid algorithms were presented in work by Ali Mostafaeipour et al. (2021), and their efficacy in forecasting wind speed behavior in the Abadeh region was assessed. Three hybrid algorithms: Genetic algorithm (GA), Simulated Annealing algorithm (SAA), and Shuffling Frog-Leaping algorithm (SFLA) were proposed. Meteorological data during a 10-year period from 2005 to 2015 from the Abadeh region was used, including eight inputs: maximum temperature, minimum temperature, water vapor, minimum humidity, maximum humidity, wind speed, precipitation, and sunlight hours. The algorithms' efficacy and dependability in wind speed prediction were evaluated using the metrics RMSE, MSE, MAPE, and R2. The SFLA hybrid algorithm has the lowest error and maximum reliability in forecasting wind speed behavior in this investigation, with RMSE and R2 of 0.0761 and 0.91, respectively. SAA has lower reliability in wind speed prediction than the previous hybrid algorithms.

Alireza Goli et al. (2021) proposed an integrated framework based on statistical tests, time-series neural networks, an improved multi-layer perceptron neural network (MLP), adaptive neuro-fuzzy interface system (ANFIS), and support vector regression (SVR) with novel meta-heuristic algorithms in order to obtain the prediction of dairy product demand in Iran. The innovation of this study is to improve artificial intelligence tools with the help of the newest meta-heuristic algorithms such as the grey wolf algorithm (GWO), cultural algorithm, and particle swarm optimization (PSO). The dataset consists of the monthly demand for dairy products from Pegah Golpayegan Company for a period of 60 months from 2013 to 2018. d. The Pearson correlation coefficient was used to eliminate ineffective variables: index of dairy prices, annual average gross income, and industrial production index. In the hybrid technique based on MLP, the forecast error was reduced by 1.8 times. In the ANFIS, this number was 10.63, while in the SVR approach, it was 3.2 times higher. These findings showed that the ANFIS has the highest potential for improvement using meta-heuristic techniques. It is proposed that hybrid algorithms, such as hybrid GWO-PSO, be used to improve artificial intelligence predictions in this study.

Table 1 shows a comprehensive literature survey in tabular form. Table 1 and the previous section details show that this paper presents an automated comparable classification analysis and results of tooth dental X-ray images as diseased or un-diseased. The four most sort after and hierarchically evolving deep learning architectures having a huge number of layers to prevent overfitting of data have been used. The results obtained from the application of all architectures over the complete dataset are compared.

Although previous works have worked extensively on such a topic, the literature survey reveals that there are research gaps that this study seeks to fill. The following are the research gaps in the existing literature:

- Slightly outdated CNN architectures are used, compromising performance and robustness.

Table 1 Literature survey

Existing Works	Dataset	Algorithms	Experiment result
Tissue Level Based DL Framework for Early Detection of Dysplasia (Gupta et al., 2019b)	Total of 2688 images 1882 training images 806 testing images	CNN	While training for 75 epochs, the Trained CNN achieved an accuracy of 91.65% on training data. It provided 89.3 percent accuracy on testing data
Tooth detection and numbering in panoramic radiographs using CNN (Tuzoff et al., 2019)	1352 testing panoramic radiographs 222 testing images	Faster R-CNN (Detection)	For teeth detection a sensitivity of 0.9941 and precision of 0.9945 was achieved
		VGG-16 CNN (Numbering)	For teeth numbering a sensitivity of 0.9800 and specificity is 0.9994
Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases (Anantharaman et al., 2018)	30 training images 10 testing images	Mask R-CNN	The dice coefficient score of Cold sores was 0.774 and for Cancer sores was 0.714
Automated Dental Image Analysis by Deep Learning on Small Dataset (Yang et al., 2018)	196 combinations having pre and post-treatment images	CNN AlexNet GoogleNet	The proposed network got higher accuracy in 'getting better' cases but made more wrong decisions on 'getting worse' cases. The final F1 score obtained is 0.749
Classification of Dental Diseases Using CNN and Transfer Learning (Prajapati et al., 2017)	251 X-ray images with 45 and 26 used for validation and testing, respectively	CNN	The given CNN architecture gives an accuracy of 73% on testing data
		Transfer Learning (VGG16)	The VGG16 model give an extra 15.39% compared to CNN model
		Transfer Learning (VGG16) with Fine Tune	Due to only 4 layers learning very low level features, fine tuning didn't improve the score much
Automatic white blood cell classification using pre-trained deep learning models: ResNet and Inception (Esteva et al., 2017)	Blood Cells images	ResNet-101	When the final layer was fine-tuned, it obtained an accuracy of 87.23 percent. When all layers were fine-tuned, the accuracy increased to 99.46 percent

Table 1 (continued)

Existing Works	Dataset	Algorithms	Experiment result
		ResNet-152	When the final layer was fine-tuned, it obtained an accuracy of 87.46 percent. When all layers were fine-tuned, the accuracy increased to 99.84 percent
Dermatologist-level classification of skin cancer with deep neural networks (Checco & Corinto, 2006)	129,450 clinical images	CNN/3-way	The CNN classifier had a mean accuracy of 69.5% with a standard deviation of 0.8% on 3-way classification
		CNN/9-way	The CNN classifier has a mean accuracy of 48.9% with a standard deviation of 1.9% on 9-way classification
		CNN-PA/3-way	For 3 classes, with the proposed partitioning algorithm the accuracy increases to a mean of 72.1% with 0.9% standard deviation
		CNN-PA/9-way	For 9 classes, with the proposed partitioning algorithm the accuracy increases to a mean of 55.4% with 1.7% standard deviation
Detection and Classification of Dental Caries in X-ray Images Using Deep Neural Networks (Ali et al., 2016)	The data set of images was collected from many dentists	Deep Neural Network	Predictions for teeth decay are 98 percent accurate, while predictions for normal teeth are 96.1 percent accurate. 96 percent of all tooth decay cases are correctly predicted as decayed teeth. 98 percent of all normal teeth cases are correctly classified. In total, 97 percent of predictions are correct

Table 1 (continued)

Existing Works	Dataset	Algorithms	Experiment result
CNN-Based Image Analysis for Malaria Diagnosis (Liang et al., 2016)	450 malaria images	CNN Transfer Learning	The proposed method using the new CNN model outperforms the transfer learning model. The CNN model's average classification accuracy is 97.37 percent, and the model's sensitivity, specificity, and precision are all greater than 97 percent. Whereas the transfer learning model achieved an accuracy of only 91 percent
Classification of teeth in cone-beam CT using deep convolutional neural network (Miki et al., 2016)	Original Rotation Intensity Transform Rotation + Intensity Translation	AlexNet	Samples rotated and window adjusted improved accuracies in training. In random sampling, the fill method has the best classification accuracy. Overall, increasing the number of training samples improved classification performance to 91.0 percent accuracy by rotation and intensity transformation
Artificial Convolution Neural Network for Medical Image Pattern Recognition (Lo et al., 1995)	55 chest radiographs 28 images for training 27 images for testing	Convolution	The average Az (area under the receiver operating characteristic curve) was 0.77 when CNN with delta function was used
		CNN/OA-Sym	Az was increased to 0.83 when output association was performed using CNN with a narrow Gaussian distribution

Table 1 (continued)

Existing Works	Dataset	Algorithms	Experiment result
		CNN/OA-Asym/8 Rotations	Az was increased to 0.87 by including eight different types of rotated image blocks in the input and output associations

- Because the datasets are so small, the performance of a deep CNN on such a dataset is questionable.
- The model is not trained on a larger number of epochs in some papers.
- The data is disproportionate, or some existing works have been limited to a few dental problems.

Our pipelined addresses these research gaps by working in conjunction with AlexNet, which excels at extracting deep and baseline visual features. The research gaps mentioned are addressed as follows:

- AlexNet and ResNets are modern deep CNN architectures that excel at image classification tasks.
- The dataset we used contains 16000 images that are evenly distributed across several dental condition classes. It has been collected from nearly 5000 patients over a 15-year period. The presence of such a dataset improves the model's validity, performance on real-world data, and robustness.
- We trained the model for the appropriate number of epochs (20) to avoid underfitting and overfitting.
- Our data has been evenly distributed across several dental classes, including caries, RCT, bone loss, missing teeth, and so on.

In this work, we provide a top-level categorization of dental X-ray image scans into a diseased or healthy scan, which comes to be comparable to any radiologist or dental expert's assessment.

3 Contribution outline

Integrating DNN and validating it fairly (concerning state-of-the-art works) large amounts of data to understand and analyze periapical dental diseases is the study's primary aim. This allows us to have multiple objectives:

(a) The evaluated DNN includes four well-known deep learning architectures/models, namely CNN, AlexNet, ResNet-18, and ResNet-34, for disease classification. Compared to the state-of-the-art works, the baseline architecture of AlexNet makes a difference in terms of performance (ref. Section Results). Further, employing multiple deep learning architectures/models provides a fair comparison among them.

(b) As no state-of-the-art works provide large amounts of data for validation, we obtained a dataset of 16,000 images and annotated them for dental periapical classification. This brings us a solution to quantify/test DNN appropriately and is available for research purposes (upon request).

As depicted in Fig. 12 (ref. Sect.4.3), the assessed approach works by collecting and preparing a dataset obtained from a dental clinic followed by disease detection and consequent classification through state-of-the-art baselines DNN.

4 Materials and methods

4.1 Materials

This sub-section discusses the employed datasets, the assessed deep neural network architectures, and the algorithm used.

4.1.1 Dataset collection

The dataset contains almost 16,000 unique dental X-ray images that are taken from around 5000 patients. The X-ray images were being taken for over 15 years. Some of these images may be useful, and others may not be. Some will be clear, while some will be distorted. All the images are of different sizes. Hence, pre-processing is necessary. After removing the distorted and useless images, the rest of the images were scaled to a specific size of 140 × 105 pixels by cropping the unnecessary area, resizing, or by introducing some white pixels around it. The dataset has four feature columns-: ImageID, PatientID, Gender, and Age.

4.1.2 Dataset preparation

Caries: Dental caries is a complex chronic disease that cannot be easily detected in the early stages. The occlusal, aproximal, and smooth surfaces at the gingival margins of teeth are available areas where plaque can accumulate easily and undisturbed to form dental caries.

In dental X-ray, caries can be identified by black spots on the crown of the tooth, as shown in the below dental X-Ray Images (Naam et al., 2017). Figure 2a shows a sample image of caries infected tooth from the dataset.

Root Canal Treatment (RCT): Root canal treatment is a frequently performed procedure to address pulpal and periarticular disease. It comprises several clinical steps regardless of the initial diagnosis.

RCT can be defined as an extension of caries. In dental X-Rays, RCT, similar to caries, can be identified by black spots on the tooth's crown (Almanei et al., 2017). The difference between the two is that if the black spot touches the nerve, it is identified as RCT (shown in Fig. 2b).

Abscess: Dental abscess is a frequently occurring infectious process known to health practice. The infection's fate depends on the bacteria's virulence, host resistance factors, and regional anatomy. Serious consequences arising from the spread of a dental abscess lead to significant morbidity and mortality.

In dental X-rays, an abscess can be identified by black spots around the root tip of the tooth dental X-ray Images (Chapman et al., 2013) (also shown in Fig. 2c).

Bone Loss: The black triangle between two teeth can identify bone Loss in the X-Ray images. If the triangle formed is significantly big or, in other words, is bigger than the average one, then the patient is suffering from Bone Loss. Some of the images are shown in Fig. 2d.

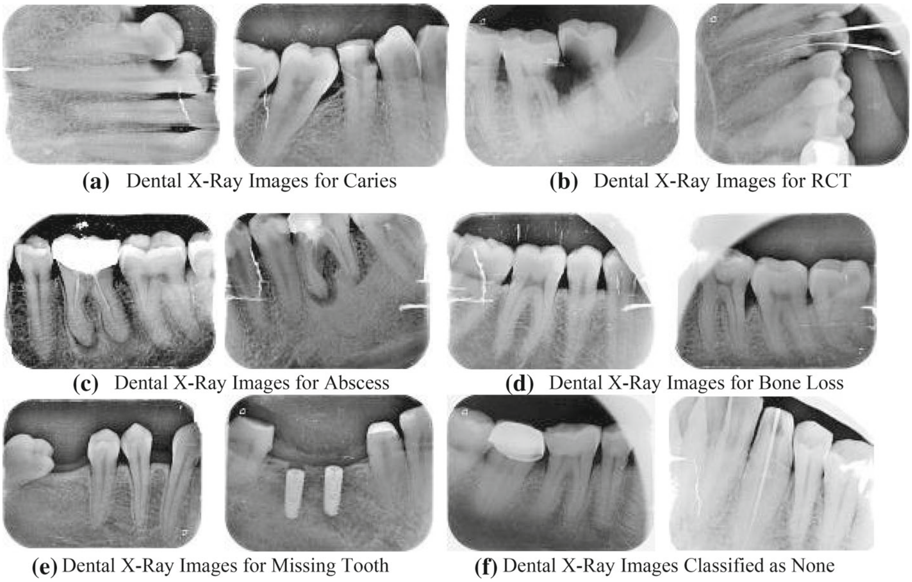


Fig. 2 Some Example of various dental conditions from the dataset

Missing Tooth: Missing tooth can be easily identified by any person as there is a big gap between two teeth. The big gap between the two teeth indicates that the tooth is missing. Some sample X-Rays of a missing tooth are shown in Fig. 2e.

None: This class is used to classify the dental X-Ray images that are not suffering from any of the above diseases. If an X-Ray image cannot be classified in any of the above categories, it is classified as "None." Some images that are classified as "None" are shown in Fig. 2f.

The dataset was first prepared from the raw data provided by the dental clinic. It contained patient information, X-ray information, and X-ray images. The information like patient age, gender, and patient ID was extracted to find patterns to associate different diseases.

After extracting the data, the data and the X-ray images were used to classify them into different diseases mentioned in the dataset description.

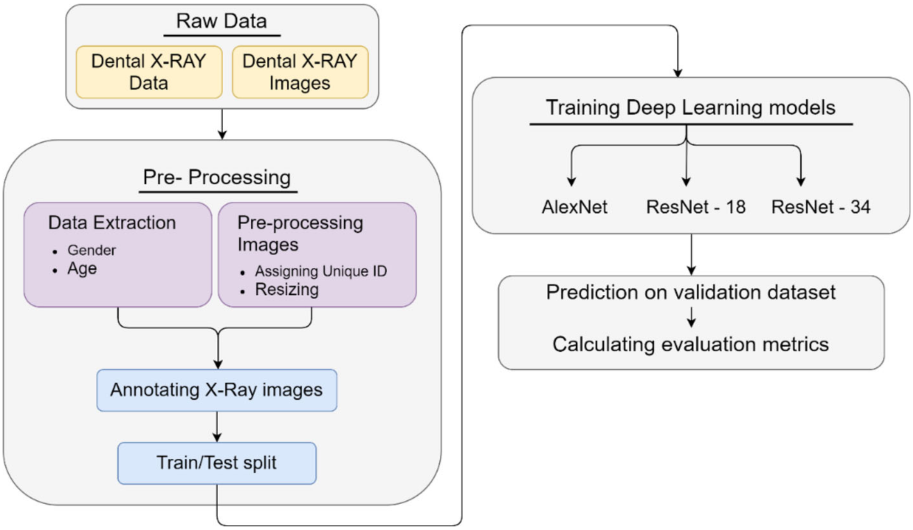


Fig. 3 Proposed modified pipelined approach for periapical disease classification

4.2 Modified approach

In the proposed work, a modified CNN model, AlexNet, ResNet-18, and ResNet-34 are the four DNNs that are used to train and test the images using the Google Colab[2] tool. The initial pre-processing of the dataset includes light addition, resining, and cropping. A Unique ID recognizes every image. Corresponding data labeling is used, which consists of attributes: Gender and Age. The resulting annotated periapical X-Ray images further trained using other baseline and optimized CNN.

The baseline models use the weights and knowledge gained from solving a specific problem and applying it to solve similar tasks. It helps leverage the weights and biases of different state-of-the-art algorithms. It hence uses it as an advantage without it being necessary to have vast amounts of data or extensive computation capabilities. In the pipelined approach, as shown in Fig. 3, the optimized CNN includes fine-tuning the model by unfreezing the specific parts of the model and re-training it on the new data with a small learning rate. Algorithm 1 explains the steps involved in implementing pre-processing and consequent DNN evaluation on the processed dataset.

[2] https://colab.research.google.com.

ALGORITHM 1: Periapical pre-processing and classification using Deep Neural Networks

```
This algorithm takes the extracted images
if (videos)
  images = ExtractFrames(videos)
end if.
for each image
  NewImage = Preprocessing (image)
  function Resize(NewImage, MinSize, MaxSize,  Padding)
          resized_imge = cv.resize (img, (140,105))
     return resized_imge, Window, ScalingFactor, Padding
  end function
end For
function Preprocessing(image){
       OpenCV.ImageRead(image)
   OpenCV.convert(image)
return ProcessedImage
function Resize(Image, MinSize, MaxSize, Padding){
     if Padding is true
        Image resized to MaxSize x MaxSize
     else
   Image resized to MinSize x MaxSize
     endif
returns Image
function Sequential
x_train,x_val,y_train,y_val = SplitData(x,y,test_size= 0.3)
for each model ϵ F do   // F is the evaluated Neural Network model
        model = fit(x_train,y_train, epochs= 20)
     z = predict(x_test)
           y_pred = append(z)
end for
model.compile (loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
model.confusionmatrix(y_pred,y_test)
accuracy = compare(y_pred,y_val)
return y_pred,accuracy
end function
```

The four deep neural networks are trained on the processed data batches created. Based on the results (ref. Sect. 5), the best-trained model for every deep neural network is used to classify diseased and un-diseased on the periapical dental images.

4.3 Methods used and mathematical analysis

This section discusses in detail the mathematical analysis of the methods used in the proposed pipelined approach. The various DNNs used are—CNN, ResNet -18, ResNet-34, and AlexNet.

CNN: The CNN model takes an image of size 105 $\times$ 140 pixels in Red–Green–Blue (RGB) format. The modified CNN contains four different layers; the first two layers are convolutional. A max-pooling layer then follows these two layers. Finally, the output is flattened and dropped so that it can be fed to a dense layer.

The basic layers of the four Deep Learning models were designed based on the fol\ lowing:

Convolutional Operation: It is done on the image of size a*b, with a kernel size of p, stride size r, and padding q, produces an output of $size = \frac{(a-p+2q)}{r+1} \times \frac{b-p+2q}{r+1}$.

In neural systems, piece size shows neurons' particular field, authorizing neighborhood availability of neurons to the past. K, which is the resultant output derived from the convolutional operation of matrix X (U*V) and matrix Y(M*N) is:

$$\mathrm{K(i, j)} = \sum_{v=0}^{u=1} \sum_{v=0}^{V=1} X(u, v) \times Y(i-u, j-u) \tag{1}$$

Here, O < i < U + M-1 and 0 < I < 0 + N-1.

For computing K (0, 0), Y is first turned by 180° about the middle component, and the inside is slide so as it lies on the top of U (0, 0). After this, each element of Y is multiplied by every element of X. To calculate K (0, 0), all the products are summed together. The operation that doesn't consider flipping of grid Y is known as cross-connection activity. This is expressed as:

$$\mathrm{K(i, j)} = \sum_{v=0}^{u=1} \sum_{v=0}^{V=1} X(u, v) \times Y(u-1, v-1) \tag{2}$$

Maximum Pooling Operation: Max-Pooling is an aggregation operation that extracts the max value in a region on the image. Doing this also helps in controlling over-fitting. Dropout Regularization: "Dropout" refers to randomly dropping the neuron, which can be both hidden and visible, in a neural network. This technique was developed to tackle the over-fitting problem in a neural network.

Non-Linearity Layers: After convoluting layers, some non-linear operations are applied with the "activation functions." ReLU, as shown in Eq. 3, is a rectifier function that is half-wave and helps prevent overfitting. The variations of ReLU such as leaky ReLU, Noisy ReLU, ELU.

$$f(u) = (0, u) \tag{3}$$

$$f(u) = \{u, u > 00.01u, \text{otherwise} \tag{4}$$

$$f(u) = \{u, u \geq 0a\left(e^{u} - 1\right), \text{ otherwise} \tag{5}$$

Here Eq. (3) is of Simple ReLU, Eq. (4) is of Leaky ReLU, and Eq. (5) is of Exponential LU(ELU), a is a hyper-parameter in Eq. (5). Softmax function gives an output value between 0 and 1, is usually used for two or more classes. It is defined as:

$$f(u_i) = \frac{e^{u_i}}{\sum_{j=0}^{k} e^{u_j}} \tag{6}$$

Normalization: ReLUs has the desirable property that they do not require input normalization to prevent them from saturating. Denoting $c_{x,y}^{k}$ as the activity of neuron calculated for kernel k at position (x, y). After applying the above Nonlinearity of the ReLU, the normalized output

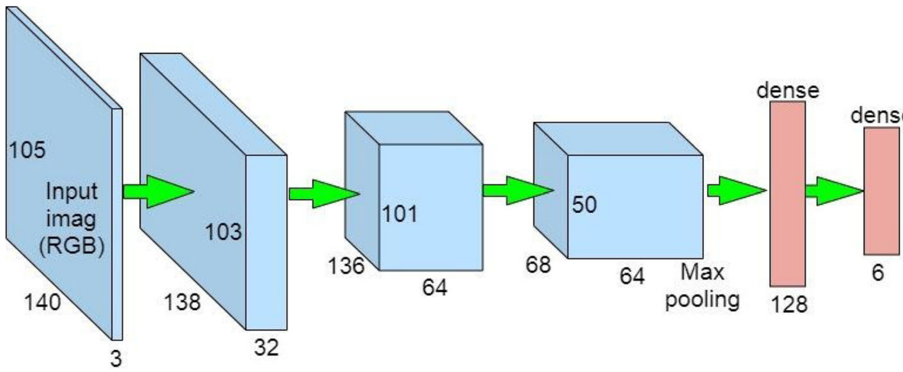


Fig. 4 Architecture of the optimized CNN Model designed

is given by $o_{x,y}^{k}$ and Eq. 7.

$$o_{x,y}^{k} = \frac{c_{x,y}^{k}}{k + \propto \sum_{j=\max\left(0,k-\frac{n}{2}\right)}^{N-1,k+\frac{n}{2}} c_{x,y}^{k}} \tag{7}$$

Figures 4, 5a, b, and 6 respectively show the specific architectures of the four models used- optimized CNN, ResNet-18, ResNet-34, and AlexNet.

The architecture of both the ResNet models as depicted in Fig. 5, have been described as follows:-

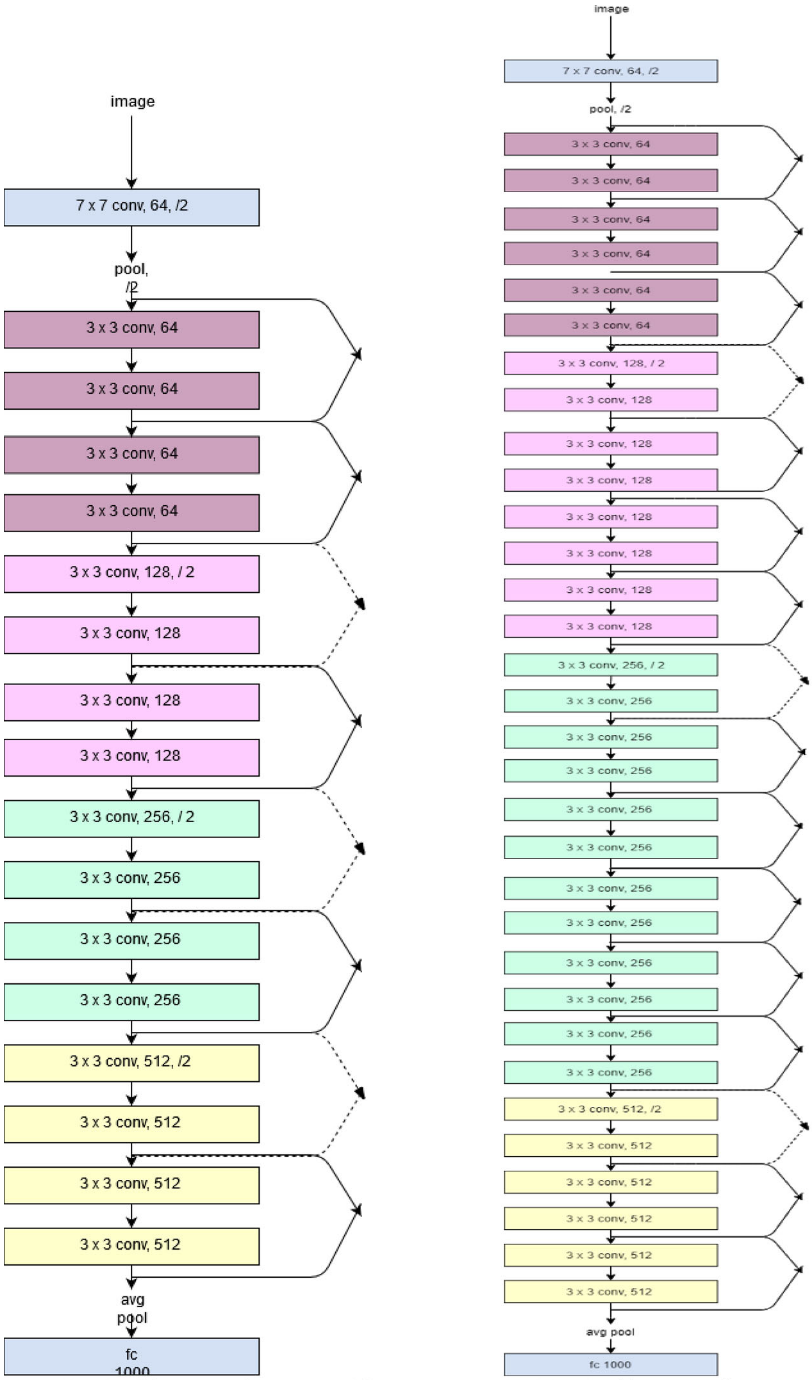
ResNet-18 consists of a total of 18 layers. Each module has four convolutional layers (excluding 1 × 1 the convolutional layer). It is similar to the ResNet-34 but only with fewer layers (He et al., 2016).

ResNet-34 consists of one convolutional layer and a pooling layer followed by four similar layers. The layers perform 3 × 3 convolution with a feature map of dimension [64, 128, 256, 512] respectively, bypassing the input every two convolutional layers (He et al., 2016). Also, the dimensions of width and height remain the same during the entire layer. The dotted line represents a change in the dimension of the input volume. Note that this reduction between layers is achieved by increasing the stride at each layer's first convolution instead of a pooling operation.

The description of Fig. 6, is given as follows:-

AlexNet: The input of AlexNet is always an image of size 250 × 250 pixels and in RGB format. If the image is not in this shape, it is converted to that format. AlexNet contains eight different layers; the first five layers are convolutional. Max-pooling layers then follow them. Finally, the last three layers are fully connected (Krizhevsky et al., 2012). It uses the non-saturating ReLU activation function, which showed improved results and better training performance over tanh and sigmoid. The weight update rule of AlexNet using stochastic gradient descent is given by Eq. 8.

$$v_{i+1} = 0.9v_i - 0.0005 \cdot \varepsilon \cdot w_i b - \varepsilon \cdot \frac{\partial L}{\partial w}|_{w_i D_i} \tag{8}$$



(a) Architecture of ResNet – 18 **(b)** Architecture of ResNet-34

Fig. 5 Architectures of two ResNet modes namely **a** ResNet-18 and **b** RestNet-34

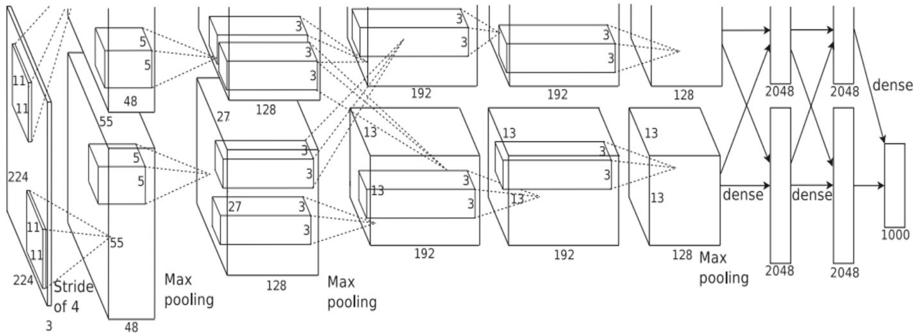


Fig. 6 Architecture of AlexNet

Equation 9 shows the Local Response Normalization that is used in the AlexNet model to aid generalization.

$$b_{x,y}^{i} = a_{x,y}^{i} / \left(k + \alpha \sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} \left(a_{x,y}^{j} \right)^{2} \right)^{\beta} \tag{9}$$

5 Experimental results

After applying a DNN on fairly large data collection, the data goes through a pre-processing phase. This phase performs resizing, cropping, and augmenting, finally producing a unique dataset of 16,000 dental X-ray images as per Algorithm 1 (ref. Section 4). Further, these images are given as input to four deep neural network models, as described in Sect. 4.1. The following sub-section presents the experimental results based on binary classification into diseased and un-diseased, i.e., an unhealthy tooth with any of the mentioned disease and healthy tooth, respectively.

5.1 Results

Initially, a six-way classification is performed by all four models for finer disease classification, followed by a four-way partition or classification scheme for all four models to achieve a coarse grain analysis of the dataset. The study indicated that the results were fairly observable, but were constrained by skewed and coarse annotation. Later, binary classification is obtained on the entire dataset to get certain finer insights. The progress in accuracy demonstrates the validity effectiveness of the partition diseased and un-diseased so formed. Though the dentists label the dataset by a run-through scan, they are not necessarily confirmed as per specific treatment making the metric inconclusive and proving that the models CNN, AlexNet, and ResNet are learning after every epoch. Table 2 represents the complete dataset results for two classes for all models used—the metrics used for comparison—precision, recall, F1-score, an accuracy score. We further present the intermediate results of all the DNNs. The performance metrics considered for this study are as follows:

Table 2 Result of different models on the complete dataset for dental disease classification

	Accuracy	Precision	Recall	F1 Score
Modified CNN	0.612	0.780	0.617	0.549
AlexNet	0.852	0.850	0.850	0.850
ResNet-18	0.815	0.815	0.815	0.815
ResNet-34	0.493	0.247	0.500	0.330

1. *Accuracy:* It is the proportion of correctly classified points (prediction) to total predictions. It has a value between 0 and 1. It is the most intuitive metric and can be useful when both test and training accuracies are compared to check if the model has overfitted or underfitted the dataset. Its formula is given in Eq. 10.

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})} \tag{10}$$

2. *Precision:* It is the proportion of correctly predicted positive points to total positive points. High precision thus means that the false positive rate is low which is beneficial for a model. Its formula is given in Eq. 11.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{11}$$

3. *Recall:* It is the proportion of correctly predicted positive observation to the all observations in actual class. It provides and indication of missed positive predictions. Its formula is given in Eq. 12.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{12}$$

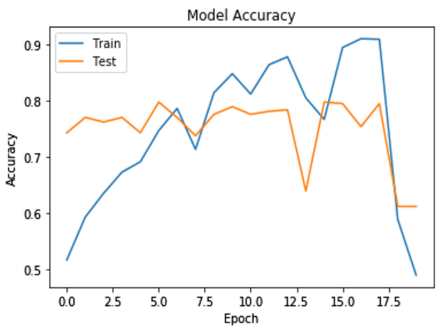
4. *F1 Score:* It is a weighted average of precision and recall. It helps combine precision and recall and provides a better measure of model performance. It's formula is given in Eq. 13.

$$\text{F1 - score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

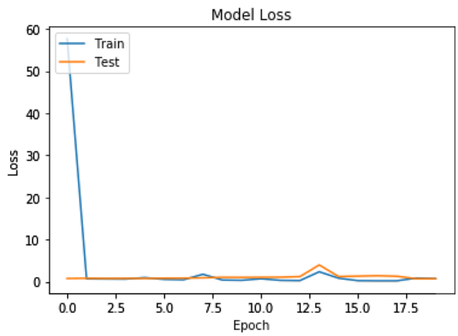
5. *Confusion Matrix:* It is a binary classification matrix of dimension 2×2, with real values on one axis and predicted values on the other.

Where,
True Positive (TP) is an outcome in which the model forecasts the positive class properly.
True Negative (TN) is an outcome in which the model predicts the negative class properly.
False Positive (FP) is an outcome in which the model forecasts the positive class inaccurately.
False Negative (FN) is a result in which the model forecasts the negative class inaccurately.

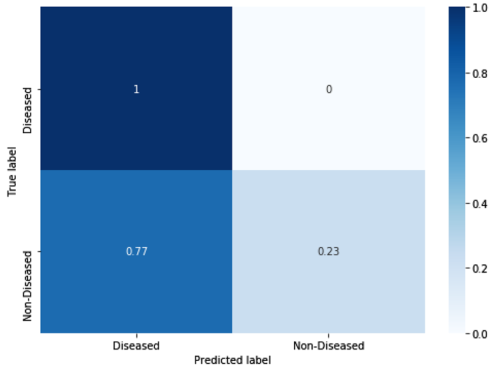
(i) *CNN with added Layers:* Fig. 7a, b and c show that training and testing accuracy vs. epochs, training loss with testing loss vs. epochs and confusion matrix. The accuracy score obtained using this redefined CNN architecture is 0.612, as shown in Fig. 8. Using optimized CNN, the precision-recall score evaluates to 0.78 and 0.617, respectively, for both classes. This score accounts for the best amongst all the tweaked CNN, i.e., freezing and un-freezing layers.



(a) Model Accuracy vs Epochs



(b) Model Loss vs Epochs



(c) Confusion Matrix

Fig. 7 Performance of the CNN model with added layers shown **a** Accuracy vs Epochs and **b** Loss vs Epochs **c** Confusion Matrix

(ii) *AlexNet:* Fig. 9a, b and c show the training and testing accuracy vs. epochs, training loss with testing loss vs. epochs and confusion matrix. The accuracy obtained using this model architecture is 0.852, as shown in Fig. 10. A larger area under the curve can be obtained in an ideal scenario with high recall and precision values. Since the area under the curve is large, indicating high precision (accurate results) and high recall (majority positive results). This signifies that both the false-positive rate and false-negative rate are low.

(iii) *ResNet-18:* Fig. 11a, b and c demonstrate training and testing accuracy vs. epochs, training loss with testing loss vs. epochs and confusion matrix. The accuracy score obtained using this is 0.815 (also shown in Fig. 12).

(iv) *ResNet-34:* Fig. 13a, b and c demonstrate training and testing accuracy vs. epochs, training loss with testing loss vs. epochs and confusion matrix. The accuracy obtained using this model is 0.612, as shown in Fig. 14.

The Table 3 depicts comparative analysis between out approach and existing approaches. However, a complete and fair comparison involving the results of existing approaches the proposed approach is not directly possible since to the best of our understanding, there has been no work on a large dataset as used in our study (16,000 images).

Our study focuses on a pipeline approach in which data is cleaned and preprocessed prior to being used in any deep learning model. However, AlexNet has been known to excel at

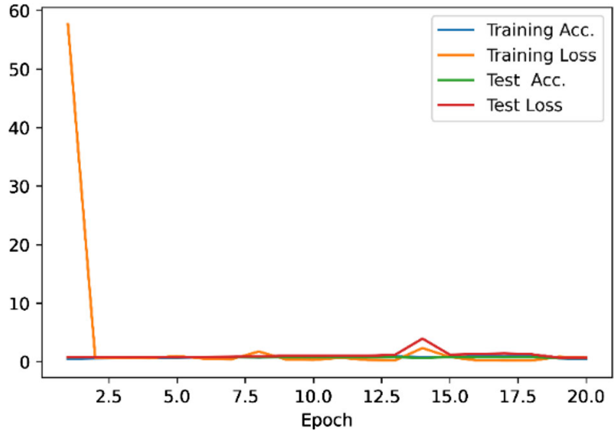
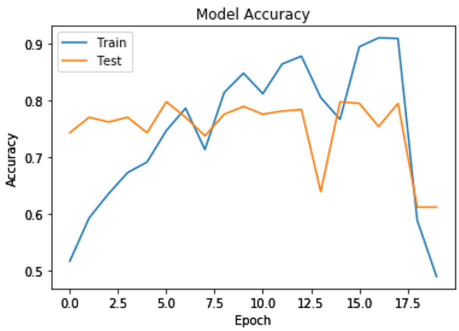
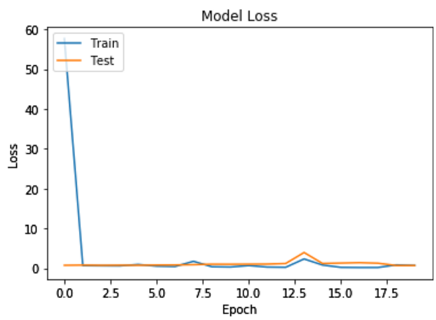


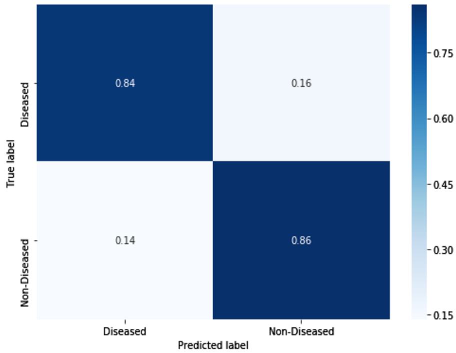
Fig. 8 Accuracy and loss in training and Test for CNN (with added layers)



(a) Model Accuracy vs Epochs



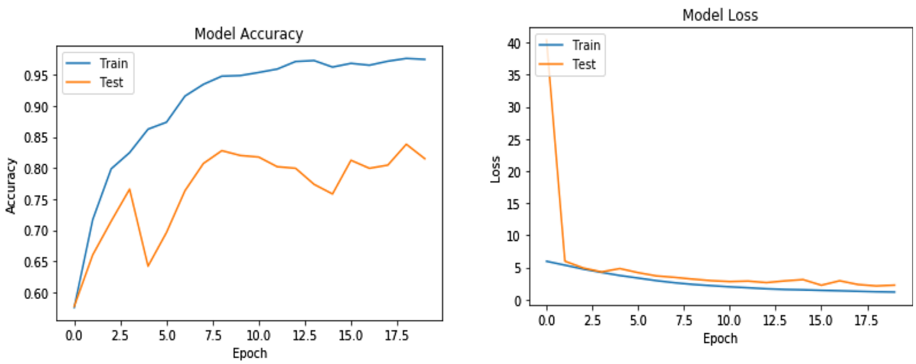
(b) Model Loss vs Epochs



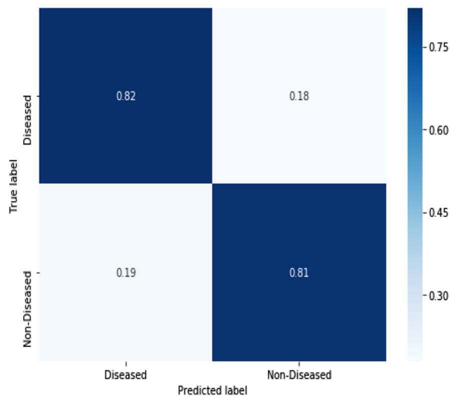
(c) Confusion Matrix

Fig. 9 Performance of the AlexNet shown as **a** Accuracy vs Epochs and **b** Loss vs Epochs **c** Confusion Matrix

Fig. 10 Accuracy and Loss in Training and Test for AlexNet



(a) Model Accuracy vs Epochs **(b)** Model Loss vs Epochs



(c) Confusion Matrix

Fig. 11 Performance of the ResNet-18 shown as **a** Accuracy vs Epochs and **b** Loss vs Epochs **c** Confusion Matrix

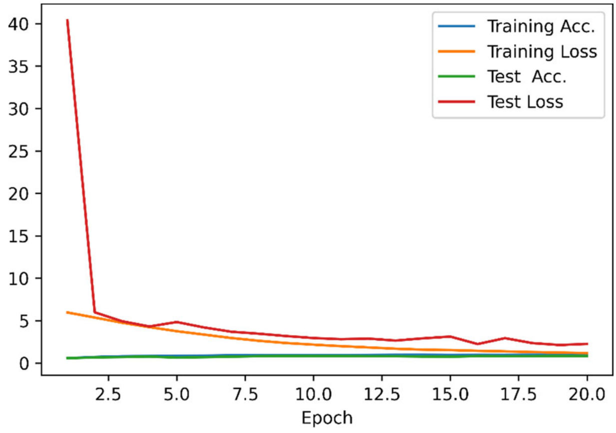


Fig. 12 Accuracy and Loss in Training and Test for ResNet-18

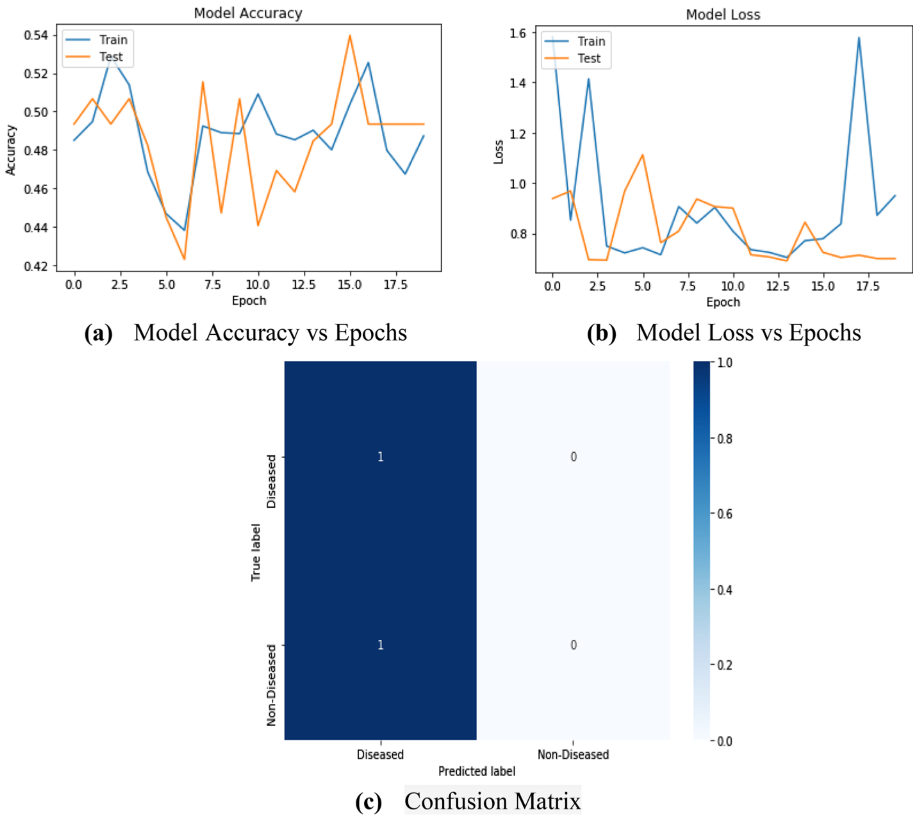


Fig. 13 Performance of the ResNet-34 shown as **a** Accuracy vs Epochs and **b** Loss vs Epochs **c** Confusion Matrix

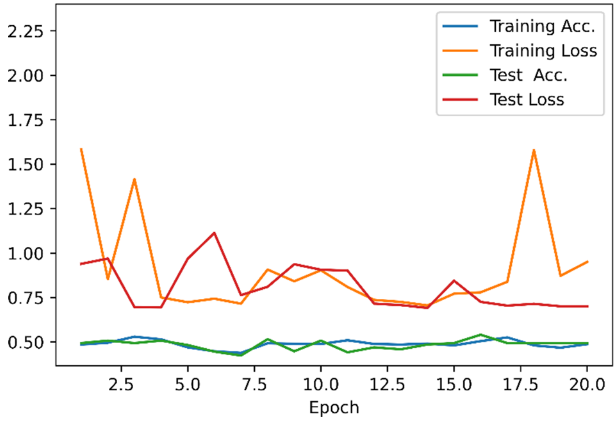


Fig. 14 Accuracy and Loss in Training and Test for ResNet-34

Table 3 Comparative Analysis of existing approaches and our pipelined approach

Parameter	Our Study	Dermatologist-level classification of skin cancer with deep neural networks (Checco & Corinto, 2006)	Automated Dental Image Analysis by Deep Learning on Small Dataset (Yang et al., 2018)	Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases (Anantharaman et al., 2018)
Accuracy	AlexNet achieves maximum accuracy of 85.2%	The maximum accuracy was obtained as 73%	The accuracy achieved was 77%	74%
Dental Dataset Size	16,000 unique dental X-ray images that are taken from around 5000 patients	–	196 combinations having pre and post-treatment images	30 training images 10 testing images
Model Used	AlexNet	CNN-PA/3-way	CNN	Mask R—CNN

image classification tasks, particularly when working with large real-world data sets such as the one used in this study. This is because it is capable of extracting both deep and baseline visual features. Even though the optimized CNN model presented has been tweaked and modifications made to the layers, the model has difficulty working on large high resolution data sets without over fitting. Alexnet solves this problem by combining model ensembles and consecutive convolution layers.

A model is said to be a good fit if it can generalize and learn the features from the training data without over-fitting or under-fitting, i.e., it can generalize the features and perform well even on unseen data. Considering that the images are real X-ray images and the size of the dataset is very larger, AlexNet model is performing very well. As it can be seen from Fig. 9a and 10, the training and testing accuracy are similar throughout the training, indicating that the model hasn't over fitted to the dataset and is generalizing well on unseen data. This can be seen from the fluctuation in the graph for the validation accuracy curve.

6 Conclusion

The importance of timely and accurate dental diagnosis cannot be overstated. The sooner a dental problem is identified and treated, the more likely it is to reverse and repair the damage with minimal invasion. Even a minor enhancement or improvement in existing techniques can go a long way in providing significant health benefits in the medical field. This paper has made a successful attempt to contribute a different type of pipelined approach using AlexNet in this regard. Medical diagnoses and treatments differ from one patient to the next. While some treatment options may be beneficial to a specific patient population, others may not. However, due to the unpredictable nature of different medical treatments on different human bodies, some patients' bodies may be receptive of one kind of treatment while the other patients could be responding to different treatments.

In this paper, we performed dental X-ray image classification using four different DNNs, namely, AlexNet, ResNet-18, ResNet-34 and modified CNN on a fairly large dataset of 16,000 periapical dental X-ray images. To the best of our understanding, this dataset is the largest data considered under study for the purpose of dental disease detection. The dataset was then annotated with the most common diseases in teeth- caries, root canal treatment, abscess, bone loss, and a missing tooth. has utilized initial pre-processing followed by the state-of-the-art deep learning The approach model application. The validation accuracy was fairly observable for all the algorithms; however, AlexNet outperformed the other algorithms for X-ray classification with the accuracy of 0.852. This may be attributed to the fact that AlexNet has been known to excel at image classification tasks, particularly when working with large real-world data sets such as the one used in this study(Ref. Section 1: Introduction). We observe that AlexNet model does not overfit the dataset and generalizes effectively to new data. The idea behind dental X-ray automation is to accurately diagnose the problem by studying dental X-rays with deep learning architectures. The presented approach will not replace the manual force, that is, the clinical staff; rather, it will increase human productivity, decrease the treatment time, and efficiently treat many patients. Additionally, the pipelined approach used in this paper addresses the research gaps in existing literature as specified in the paper (Ref. Section 2: Related Work).

References

Akbar, S., Peikari, M., Salama, S., Nofech-Mozes, S., & Martel, A. L. (2018). Determining tumor cellularity in digital slides using resnet. In Medical Imaging 2018: Digital Pathology *10581*, 105810U. International Society for Optics and Photonics.

Ali, R. B., Ejbali, R., & Zaied, M. (2016). Detection and classification of dental caries in x-ray images using deep neural networks. In Int. Conf. on Software Engineering Advances (ICSEA) (p. 236).

Almanei, K., Alsulaimani, R., Alfadda, S., Albabtain, S., & Alsulaimani, R. (2017). Digitally scanned radiographs versus conventional films for determining clarity of periapical lesions and quality of root canal treatment. *The Scientific World Journal*.

Anantharaman, R., Velazquez, M., & Lee, Y. (2018). Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases. In 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE (pp. 2197–2204).

Chapman, M. N., Nadgir, R. N., Akman, A. S., Saito, N., Sekiya, K., Kaneda, T., & Sakai, O. (2013). Periapical lucency around the tooth: Radiologic evaluation and differential diagnosis. *Radiographics, 33*(1), 15–32.

Chatrati, S. P., Hossain, G., Goyal, A., Bhan, A., Bhattacharya, S., Gaurav, D., & Tiwari, S. M. (2020). Smart home health monitoring system for predicting type 2 diabetes and hypertension. *Journal of King Saud University-Computer and Information Sciences*.

Checco, P., & Corinto, F. (2006). CNN-based algorithm for drusen identification. In 2006 IEEE International Symposium on Circuits and Systems. IEEE (pp. 4).

Dental Quality Alliance (2012). Pediatric oral health quality and performance measures: Environmental scan. American Dental Association.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115.

Goli, A., Khademi-Zare, H., Tavakkoli-Moghaddam, R., Ahmad S., Sasanian, M., Malekalipour Kordestanizadeh R. (2021). An integrated approach based on artificial intelligence and novel meta-heuristic algorithms to predict demand for dairy products: A case study, *Network: Computation in Neural Systems*, *321*, 1–35 https://doi.org/10.1080/0954898X.2020.1849841

Gupta, R. K., Kaur, M., & Manhas, J. (2019b). Tissue level based deep learning framework for early detection of dysplasia in oral squamous epithelium. *Journal of Multimedia Information System, 6*(2), 81–86.

Gupta, V., Singh, V. K., Ghose, U., & Mukhija, P. (2019a). A quantitative and text-based characterization of big data research. *Journal of Intelligent and Fuzzy Systems, 36*(5), 4659–4675.

Habibzadeh, M., Jannesari, M., Rezaei, Z., Baharvand, H., & Totonchi, M. (2018). Automatic white blood cell classification using pre-trained deep learning models: ResNet and Inception. In Tenth International Conference on Machine Vision (ICMV 2017) *10696*, 1069612. International Society for Optics and Photonics.

Han, X., Zhong, Y., Cao, L., & Zhang, L. (2017). Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification. *Remote Sensing, 9*(8), 848.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770–778). 3

Hwang, J. J., Jung, Y. H., Cho, B. H., & Heo, M. S. (2019). An overview of deep learning in the field of dentistry. *Imaging Science in Dentistry, 49*(1), 1–7.

Jafarzadeh Ghoushchi, S., Ranjbarzadeh, R., Najafabadi, S. A., et al. (2021). An extended approach to the diagnosis of tumour location in breast cancer using deep learning. *Journal of Ambient Intelligence and Humanized*. https://doi.org/10.1007/s12652-021-03613-y

Jain, R., Jain, N., Aggarwal, A., & Hemanth, D. J. (2019). Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. *Cognitive Systems Research, 57*, 147–159.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems. 1097–1105.

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

Liang, Z., Powell, A., Ersoy, I., Poostchi, M., Silamut, K., Palaniappan, K., Guo, P., Hossain, M.A., Sameer, A., Maude, R.J. & Huang, J. X. (2016). CNN-based image analysis for malaria diagnosis. In 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 493–496). IEEE.

Liszka, G., Roska, T., Zarandy, A., Hegyesi, J., Kek, L., & Rekeczky, C. (1995). Mammogram analysis using CNN algorithms. In Medical Imaging 1995: Image Processing, *2434*, 461–470. International Society for Optics and Photonics.

Liu, S., Wang, Y., Yang, X., Lei, B., Liu, L., Li, S. X., .Ni, D. & Wang, T. (2019). Deep learning in medical ultrasound analysis: A review. *Engineering*.

Lo, S. C. B., Chan, H. P., Lin, J. S., Li, H., Freedman, M. T., & Mun, S. K. (1995). Artificial convolution neural network for medical image pattern recognition. *Neural Networks, 8*(7–8), 1201–1214.

Miki, Y., Muramatsu, C., Hayashi, T., Zhou, X., Hara, T., Katsumata, A., & Fujita, H. (2016). Classification of teeth in cone-beam CT using deep convolutional neural network. *Computers in Biology and Medicine, 80*, 24–29.

Mishra, S., Pandey, S. R., Hicks, D., Goyal, A., & Gaurav, D. (2020). A blood pressure and heartbeat anomaly detection and notification mobile application system. *Journal of Web Engineering*, 735–761.

Mostafaeipour, A., Goli, A., Rezaei, M., Qolipour, M., Arabnia, H.-R., Goudarzi, H., & Behnam, E. (2021). Performance of different hybrid algorithms for prediction of wind speed behavior. *Wind Engineering, 45*(2), 245–256. https://doi.org/10.1177/0309524X19882431

Naam, J., Harlan, J., Madenda, S., & Wibowo, E. P. (2017). Image processing of panoramic dental X-ray for identifying proximal caries. *TELKOMNIKA Telecommunication Computing Electronics and Control, 15*(2), 702–708.

Prajapati, S. A., Nagaraj, R., & Mitra, S. (2017). Classification of dental diseases using CNN and transfer learning. In 2017 5th International Symposium on Computational and Business Intelligence (ISCBI) (pp. 70–74). IEEE.

Ranjbarzadeh, R., Bagherian Kasgari, A., Jafarzadeh Ghoushchi, S., et al. (2021). Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images. *Science and Reports, 11*, 10930. https://doi.org/10.1038/s41598-021-90428-8

Ranjbarzadeh, R., Jafarzadeh Ghoushchin S., Bendechache, M., Amirabadi, A., Ab Rahman, M. N., Baseri Saadi, S., Aghamohammadi, A., Kooshki Forooshani, M. (2021). Lung Infection Segmentation for COVID-19 Pneumonia Based on a Cascade Convolutional Network from CT Images, *BioMed Research International*, *2021*, Article ID 5544742, 16. https://doi.org/10.1155/2021/5544742

Ren, S., Jain, D. K., Guo, K., Xu, T., & Chi, T. (2019). Towards efficient medical lesion image super-resolution based on deep residual networks. *Signal Processing: Image Communication, 75*, 1–10.

Sonoda, S., & Murata, N. (2019). Transport analysis of infinitely deep neural network. *The Journal of Machine Learning Research, 20*(1), 31–82.

Tuzoff, D. V., Tuzova, L. N., Bornstein, M. M., Krasnov, A. S., Kharchenko, M. A., Nikolenko, S. I., Sveshnikov, M. M., & Bednenko, G. B. (2019). Tooth detection and numbering in panoramic radiographs using convolutional neural networks. *Dentomaxillofacial Radiology, 48*(4), 20180051.

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging, 9*(4), 611–629.

Yang, J., Xie, Y., Liu, L., Xia, B., Cao, Z., & Guo, C. (2018). Automated dental image analysis by deep learning on small dataset. In 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC) *1*, 492–497). IEEE.