Multimedia Tools and Applications

Dental Cavity Classification using Image Sampling and Transfer Learning

--Manuscript Draft--

Manuscript Number:	MTAP-D-23-05014
Full Title:	Dental Cavity Classification using Image Sampling and Transfer Learning
Article Type:	Track 2: Medical Applications of Multimedia
Keywords:	dental caries detection, transfer learning, deep convolutional neural networks, computer-aided diagnosis, image sampling
Corresponding Author:	Malathi G, PhD Vellore Institute of Technology Chennai Campus Chennai, Tamil Nadu INDIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Vellore Institute of Technology Chennai Campus
Corresponding Author's Secondary Institution:	
First Author:	Jason Misquitta
First Author Secondary Information:	
Order of Authors:	Jason Misquitta
	Sankalp Hegde
	Malathi G
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	This paper evaluates a dental caries detection model that utilizes transfer learning techniques and compares it with various pre-trained deep convolutional neural networks (CNNs). The model is fine-tuned on a dataset of 1500 tooth images,(obtained from Kaggle) and the extracted features from the pre-trained CNNs are used to classify the tooth images into carious and non-carious. Dental caries, also known as dental cavities, are a common oral health problem that can lead to tooth decay and loss. However, conventional methods of caries detection, such as visual inspection and radiography, have limitations in terms of accuracy, sensitivity, and radiation exposure. In this paper, a novel method of dental caries detection using image processing and transfer learning is proposed. The proposed method replaces the traditional SVM classifier with a transfer learning approach that leverages multiple pre-trained CNN models to improve the accuracy of dental caries detection. This approach allows for the exploration of various CNN architectures and enables the selection of the most suitable model for the task at hand. The method is evaluated on a test set of dental images and compared with existing methods in terms of accuracy. The results show that our method achieves superior performance of 85\% (which is among the highest for this particular dataset) in dental caries detection and can potentially be used as a reliable and efficient tool for oral health diagnosis.

Dental Cavity Classification using Image Sampling and Transfer Learning

Jason Misquitta[1], Sankalp Hegde[1], Malathi G[1]

[1]School of Computer Science and Engineering, Vellore Institute of Technology, Vandalur-Kelambakkam Road, Chennai, 600127, Tamil Nadu, India.

Contributing authors: jason.misquitta2020@vitstudent.ac.in; sankalp.hegde2020@vitstudent.ac.in; malathi.g@vit.ac.in;

Abstract

This paper evaluates a dental caries detection model that utilizes transfer learning techniques and compares it with various pre-trained deep convolutional neural networks (CNNs). The model is fine-tuned on a dataset of 1500 tooth images,(obtained from Kaggle) and the extracted features from the pre-trained CNNs are used to classify the tooth images into carious and non-carious. Dental caries, also known as dental cavities, are a common oral health problem that can lead to tooth decay and loss. However, conventional methods of caries detection, such as visual inspection and radiography, have limitations in terms of accuracy, sensitivity, and radiation exposure. In this paper, a novel method of dental caries detection using image processing and transfer learning is proposed. The proposed method replaces the traditional SVM classifier with a transfer learning approach that leverages multiple pre-trained CNN models to improve the accuracy of dental caries detection. This approach allows for the exploration of various CNN architectures and enables the selection of the most suitable model for the task at hand. The method is evaluated on a test set of dental images and compared with existing methods in terms of accuracy. The results show that our method achieves superior performance of 85% (which is among the highest for this particular dataset) in dental caries detection and can potentially be used as a reliable and efficient tool for oral health diagnosis.

Keywords: dental caries detection, transfer learning, deep convolutional neural networks, computer-aided diagnosis, image sampling

1 Introduction

Dental caries, also known as dental caries or cavities, is a common oral health problem that affects people of all ages. Dental caries are caused by the breakdown of the hard outer layer of the tooth (enamel) by bacteria that live in the mouth. These bacteria produce acids that dissolve the enamel and create holes or pits in the tooth surface. If left untreated, dental caries can progress to the deeper layers of the tooth (dentin and pulp), causing pain, infection, and tooth loss. Other preventive measures include limiting the intake of sugary and acidic foods and drinks, drinking fluoridated water, and applying dental sealants or fluoride varnish to protect the teeth.

Dental caries are a serious public health issue that can affect the quality of life and well-being of individuals and communities. Dental caries can impair the ability to eat, speak, smile, and socialize, as well as increase the risk of other oral and systemic diseases. Therefore, it is important to raise awareness about dental caries and promote preventive and curative strategies to reduce their prevalence and impact.

The motivation for this study is to address the limitations of conventional methods for dental caries detection. The aim is to develop a more accurate, sensitive, and efficient method for detecting dental caries. We propose utilizing advancements in deep learning [1], specifically convolutional neural networks (CNNs), and transfer learning techniques to improve the detection of dental caries in tooth images.

Transfer learning is an active and promising research area in machine learning that has many applications in computer vision, natural, speech recognition, language processing, recommendation systems, and more. Medical diagnosis using transfer learning is a promising technique for automated image analysis that leverages the knowledge learned from large-scale natural image datasets to improve the performance of medical image tasks.

Transfer learning is a popular machine learning method that basically shifts the knowledge learned from a source domain to a target domain, where the domains may differ in data distribution, feature space, or task specification. For example, a source domain could be a natural images dataset such as ImageNet, which contains millions of images with thousands of classes, and a target domain could be a medical image dataset such as chest X-rays, which contains thousands of images with a few classes. Transfer learning aims to use the pre-trained model on the source domain as a starting point for training or fine-tuning on the target domain, rather than training a new model from scratch. Transfer learning can overcome the challenges of data scarcity, computational cost, and time in training deep neural networks from scratch. [2]

This work presents an approach to dental caries detection using transfer learning. The basic concepts and methods of transfer learning are first introduced and how they can be applied to dental caries detection. A survey is done on existing transfer learning for dental caries detection on different types of dental images.

Our paper proposes a novel convolutional neural network (CNN) model specifically designed for dental caries detection. This model is not only compared against other well-established pre-trained CNNs using transfer learning techniques, but it also introduces a unique architecture tailored for the task. This allows for a comprehensive assessment of the performance of the proposed model in comparison to existing transfer learning approaches.

- One of the challenges in applying deep learning models to real-world scenarios is the potential discrepancy between the resolution of images used during training and the resolution of images obtained in practice. [3] To address this challenge, the paper takes a novel approach by resampling the dental image dataset into five distinct categories of resolutions: 64x64, 256x256, 512x512, 1024x1024, and the original dataset with mixed resolutions.
- The five deep learning models, including the proposed CNN model, are trained and evaluated on each of these resolution categories. This rigorous evaluation process aims to determine how well the models perform across different resolutions and whether their accuracies are affected by variations in image resolution.

2 Literature Review

The goal of [4] is to explain the different types of radiographic imaging methods that dentists can use for their practice. The article provides a summary of each radiographic method, including the necessary equipment for setup and indications for use. It describes the basic setup of commonly used radiographic techniques, highlighting their benefits and limitations, along with the situations in which they are appropriate. While advanced imaging methods may not always be necessary for treatment, they can provide additional information that justifies the higher radiation exposure for improved treatment quality.

Development and validation of a predictive model for root caries is what [5] described. The authors collected data from 208 patients and used various machine learning algorithms to develop a prediction model based on multiple patient-related factors, including age, gender, oral hygiene, and dietary habits. The study found that the developed prediction model had a higher accuracy in predicting root caries compared to traditional diagnostic methods. The article provides valuable insights into the potential of machine learning in improving dental diagnosis and treatment planning, and could pave the way for the development of more accurate and efficient diagnostic tools in dentistry.

In [6], the aim was to develop a method to detect proximal caries in panoramic X-ray images of teeth. The study involved 27 X-ray images with proximal caries, which were processed using Matlab and several morphological gradients to enhance image quality and sharpen edges. This approach allowed for more precise identification of proximal caries characteristics and their severity level. The study addressed the difficulties associated with identifying proximal caries in panoramic dental X-rays and used Multiple Morphological Gradients as the image processing technique, which yielded results that could be utilized to identify proximal caries and their severity level with high accuracy.

[7] investigates the impact of data augmentation on deep learning-based image classification. The study explores various augmentation techniques and their effectiveness in improving classification accuracy. The findings provide insights into enhancing

model performance through increased training data diversity. The focus of [8] is on dental caries as a worldwide issue for oral health and the challenges faced in India regarding its management and proposed preventive strategies and future research directions to control this widespread oral health problem in the country.

In [9], the author highlights the increasing prevalence of dental caries worldwide and the urgent need for public health measures to address this issue. The article suggests that while the causes of this increase may be debated, there are known solutions, such as water fluoridation, topical fluoride application, school oral health education programs, and proper oral hygiene practices, including brushing with fluoride toothpaste and flossing. The article emphasizes the importance of a healthy diet and regular dental visits in preventing and managing dental caries.

In this work [10], the authors aimed to evaluate the efficacy of combining the Adaptive Dragonfly (DA) algorithm and Neural Network (NN) classifier for accurately detecting caries in dental images by performing feature extraction and classification. The study highlights the potential of AI techniques in biomedical applications, as they are widely recognized for their lasting impact and effectiveness. The strategies and techniques for image classification are reviewed in [11]. Current picture classification methods, issues, and potential solutions are covered in this section. The main emphasis was on cutting-edge classification methods to increase classification accuracy. Certain important problems with categorization performance are also addressed.

A new training criterion for deep neural networks that aims to minimize the classification errors [12] was proposed. The study examines the backpropagation method and proposes a new training criterion that incorporates both cross-entropy and maximum interval minimum classification error (M3CE). The new training criterion is tested on two popular benchmark datasets, MNIST and CIFAR-10. The results show that the M3 CE criterion is effective in improving the classification performance and the proposed M3 CE-CEc approach performs well on both datasets.

In [13], the authors evaluate the performance of popular convolutional neural networks (CNNs) for real-time object recognition in video feeds. Specifically, they examine the performance of Alex Nets, GoogLeNet, and ResNet50, which are commonly used for detecting objects and categorizing object categories in photos. To test the complete potential and limitations of these networks, the authors evaluate them on three commonly used datasets: ImageNet, CIFAR10, and CIFAR100. Importantly, videos were utilised as testing datasets, not as a training dataset. They concluded that ResNet50 and GoogleLeNet are more accurate at object recognition than Alex Net. Additionally, trained CNNs behave very differently across various object categories, and some of the potential causes for this were discussed.

In [14], the authors proposed a new neural network architecture called ResNet (Residual Network) that is able to surpass human-level performance on the ImageNet classification task. The ResNet architecture is characterized by the use of residual connections, which enables very deep neural networks with hundreds of layers to be trained. The research paper also explores how deep neural networks perform under different activation functions, such as ReLU. The results show that the use of the identity mapping and the rectifier activation function improves the accuracy of the ResNet model, and that deeper networks lead to better performance. The paper provides insights into the design of deep neural networks and their performance on complex visual recognition tasks.

A new model [15], called the TransferGAN, that learns to generate data in the target domain by transferring the learned knowledge from the source domain was introduced. This approach allows for the effective transfer of knowledge, even when the source and target domains are vastly different. Overall, the proposed method demonstrates the potential of learning to transfer as a promising direction for transfer learning research.

In [16], the authors propose a method to convert grayscale images into binary images using the Otsu thresholding algorithm which automatically computes an optimal threshold value to separate the foreground and background pixels of an image. The paper provides a detailed description of the Otsu algorithm and its implementation for image binarization. The proposed method is tested on a variety of images and the results demonstrate the effectiveness of the Otsu thresholding algorithm in producing accurate and reliable binary images.

The authors [17], provide an overview of various image denoising algorithms and introduces a new one to point out that image denoising is an important preprocessing step in many applications and is particularly challenging due to the tradeoff between preserving image details and removing noise. The paper describes several approaches, including wavelet-based methods, total variation regularization, and nonlinear filtering techniques. According to the paper, a recently proposed algorithm named Non-Local Means (NLM) achieves outstanding results in both peak signal-to-noise ratio and visual quality. This algorithm is founded on the concept of patch-based denoising.

3 Design

This methodology that was used in this work is potrayed in Fig 1.

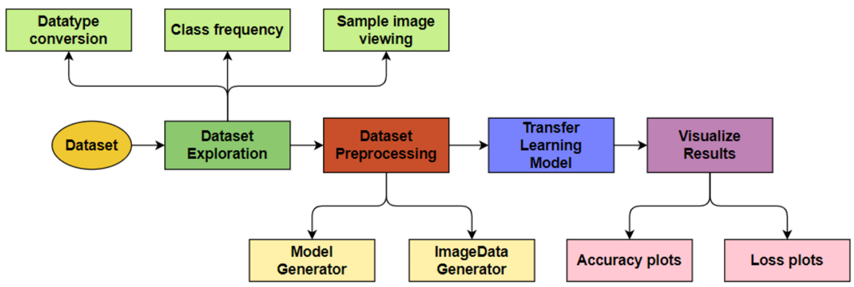


Fig. 1: The overall methodology

Dataset : The dataset employed for this study comprises the teeth decay dataset by Sáng Dinh, which is accessible on the Kaggle data repository. This dataset encompasses over 1400 images of teeth classified as carious or non-carious. The dataset is shown in Fig 2. The primary objective of this dataset is to present a comprehensive set of teeth images that can aid the advancement of computer vision methodologies for detecting dental caries. Each image within the dataset is assigned a label indicating whether it depicts "caries" or "no-caries." These labels serve as reference points for training and testing classification models.

In the training dataset, there are 315 images of "no-caries" teeth and 945 images of "caries" teeth. It's a proper 3:1 ratio split between them.

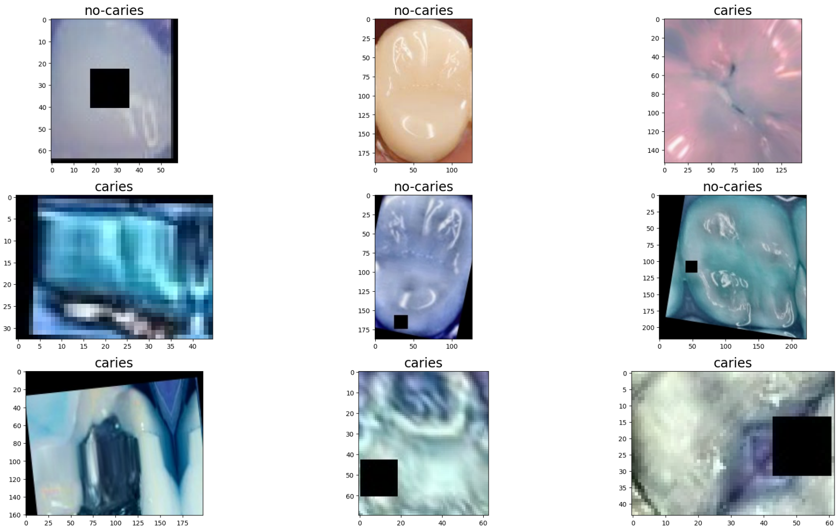


Fig. 2: Dataset

A further detailed view of the operations utilized in the image augmentation phase is presented in Fig 3. Within the scope of this project, various image processing techniques were employed to augment the dataset using the ImageDataGenerator function in Python. Initially, the rescale parameter was established at 1.0/255. to standardize the pixel values of the input images. Furthermore, a rotation range of 30 degrees was utilized to allow for image rotation within the prescribed range. Additionally, both the width-shift-range and height-shift-range were set at 0.1 to enable horizontal and vertical shifting, respectively.

Furthermore, a shear range of 0.1 was deployed to introduce shear transformations on the images. A zoom-range of 0.1 was also implemented to enable image zooming in or out. Additionally, the horizontal-flip and vertical-flip parameters were set to True to facilitate the flipping of the images. To account for any gaps that might arise during image transformations, the fill-mode was specified as 'nearest'. Finally, the brightness-range of (0.5, 1.5) was adopted to adjust the brightness of the images. The test-generator also utilized the rescale parameter of 1./255. to normalize the pixel values of the images during testing.

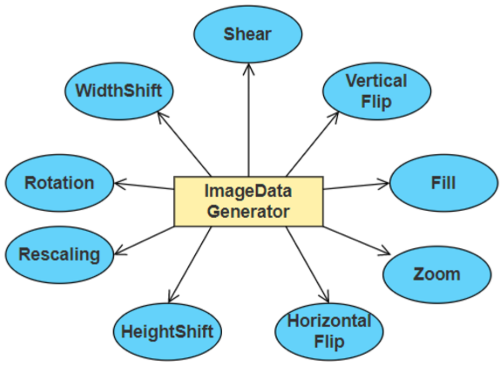


Fig. 3: Detailed view of image augmentation

4 Methodology

4.1 Conventional Deep Learning Approach

In this work, a type of artificial neural network called a convolutional neural network (CNN) was used for image classification as shown in Fig 4. The CNN consisted of several layers that extracted and transformed features from the input images. The CNN used for this project consisted of four convolutional layers. The first layer in the network took the input image and applied a set of filters to it. These filters are learned to detect certain features in the image, such as edges or corners. Next, the layer's output was fed into an activation function that added non-linearity to the model, enabling it to comprehend more intricate connections between the features. The next layer in the network downsampled the output of the previous layer, which reduced the computational complexity of the model [18] and provided a degree of invariance to small translations in the input image. This process was repeated several times with increasing numbers of filters, allowing the model to learn increasingly complex features.

After the convolutional layers, the output was flattened into a one-dimensional array and fed into a dense layer. The dense layer consisted of several nodes that performed calculations on the input data. The outputs of this layer were then passed through another activation function to introduce non-linearity into the model. To prevent overfitting, a couple of dropout layers were included in the model. These layers randomly dropped out some of the nodes during training, which helped to prevent the model from memorizing the training data and instead learned to generalize to new examples. The final layer in the CNN outputted the predicted class of the input image. The model was trained using a loss function that measured the difference between the predicted and actual class labels. The optimizer was used to minimize this loss function during training, and the accuracy of the model was evaluated using a metric that measured the percentage of correctly classified examples. Finally, the model was trained using a batch size of 50. The model was trained for 20 epochs.

```
________________________________________________________________
Layer (type)                Output Shape              Param #
================================================================
conv2d (Conv2D)             (None, 222, 222, 32)      896

activation (Activation)     (None, 222, 222, 32)      0

max_pooling2d (MaxPooling2D  (None, 111, 111, 32)     0
)

conv2d_1 (Conv2D)           (None, 109, 109, 64)      18496

activation_1 (Activation)   (None, 109, 109, 64)      0

max_pooling2d_1 (MaxPooling  (None, 54, 54, 64)       0
2D)

conv2d_2 (Conv2D)           (None, 52, 52, 128)       73856

activation_2 (Activation)   (None, 52, 52, 128)       0

max_pooling2d_2 (MaxPooling  (None, 26, 26, 128)      0
2D)

conv2d_3 (Conv2D)           (None, 24, 24, 64)        73792

activation_3 (Activation)   (None, 24, 24, 64)        0

max_pooling2d_3 (MaxPooling  (None, 12, 12, 64)       0
2D)

flatten (Flatten)           (None, 9216)              0

dense (Dense)               (None, 256)               2359552

activation_4 (Activation)   (None, 256)               0

dropout (Dropout)           (None, 256)               0

dense_1 (Dense)             (None, 64)                16448

activation_5 (Activation)   (None, 64)                0

dropout_1 (Dropout)         (None, 64)                0

dense_2 (Dense)             (None, 2)                 130

activation_6 (Activation)   (None, 2)                 0

================================================================
```

Fig. 4: Architecture of Proposed CNN model

4.2 Transfer Learning Approaches

In this work four different transfer learning models were implemented, each using a different pre-trained model. The pre-trained models that were used are:

- VGG16
- ResNet50
- IncpetionV3
- MobileNetV2

ImageNet dataset was used to train each pre-trained model. Batch size of 10 was used to train transfer learning models and compiled using the Adam optimizer and binary cross-entropy loss.

- **VGG16**
  VGG16, which had previously been trained on the ImageNet dataset, was utilized for Dental Caries Detection. The VGG16 model was customized by incorporating additional layers on top of the pre-existing, pre-trained layers to adapt it for this specific task. In order to prevent changes to the pre-trained layers, they were "frozen" during training. A global average pooling layer was applied to the output of the last convolutional layer to decrease the dimensionality of the feature maps. [15] To facilitate the training process, fully connected layers were added, with each layer followed by a batch normalization layer and a dropout layer to avoid overfitting. The model was optimized using the Adam optimizer followed by binary cross-entropy loss function and was trained for 20 epochs, which was determined to be the most effective.

- **ResNet50**
  For the ResNet50 transfer learning model, the base model is loaded from the ImageNet dataset. The weights of the pre-trained layers in the ResNet50 model will remain unchanged during training, which is referred to as "frozen." This is because these layers have already learned useful features from the ImageNet dataset, and the goal is to fine-tune the last few layers of the network for the new task of teeth classification. Some fully connected layers are added on top of the ResNet50 model. These layers are responsible for learning the mapping from the extracted features to the final output of the model, which is a probability score of the input image belonging to the positive class (teeth) or negative class (not teeth). The first fully connected layer has 512 units and uses the ReLU activation function. Batch normalization and dropout regularization are applied to prevent overfitting. This process is repeated for two more fully connected layers with 256 and 128 units, respectively. Finally, a sigmoid activation function is applied to the last layer to output a probability score between 0 and 1. This model achieved its highest validation accuracy when trained for 25 epochs.

- **InceptionV3**
  The InceptionV3 model was pre-trained on the ImageNet dataset and its pre-trained layers were frozen to prevent them from being modified. Some fully connected layers were added on top of this model, each with a batch normalization and dropout layer to prevent overfitting of the model. The first one of these layers is with 512 units and the following two are with 256 and 128 units respectively. The last layer has a sigmoid activation function for getting a output probability for the classification of caries or no-caries. On training the model for 10 epochs, it obtained a high validation accuracy.

- **MobileNetV2**
  The MobileNetV2 was one of the models used in the research, which was pre-trained on ImageNet to extract crucial image features. To prevent the pre-trained layers of the MobileNetV2 model from being modified during the training of new layers, they were frozen. The output shape of the MobileNetV2 model was adjusted to correspond to the input shape of the additional layers. The supplementary layers included a global average pooling layer that computed a feature vector by averaging the MobileNetV2 model's output, followed by three fully connected dense layers. Each dense layer contained 512, 256, and 128 units, respectively, followed by batch normalization and dropout regularization to avoid overfitting. 20 epochs were ideal here.

5 Novelty

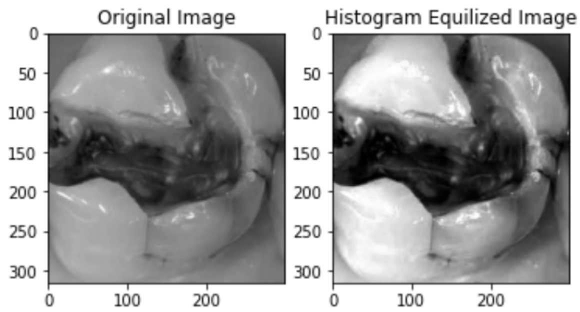
There have been many papers written on the application of deep learning models to detect dental caries. However this paper presents 2 unique novelties :

- A new CNN model is proposed and tested against the transfer learning models.
- A case may arise where a particular model is trained on a dataset that comprises of images of a certain resolution. That model is then used for detection of caries. However, the intraoral scanner that uses it, may take images of a lower resolution thus making it difficult for the model to give accurate results. Thus, the main problem here is that images of a poor resolution may be used on a model that is suited to a particular higher resolution.
  This paper counters this particular problem. The dataset has been resampled into 5 categories of resolution - 64 x 64, 256 x 256, 512 x 512, 1024 x 1024 and the original dataset (that has mixed resolutions). Each of these categories on all the 5 deep learning models. Thus, the models have been trained and evaluated on datasets of varying resolutions. This is done to judge how well the models perform. It is observed whether they are only suited to a particular resolution or not. It is also observed whether the accuracies vary with a change in sampling.

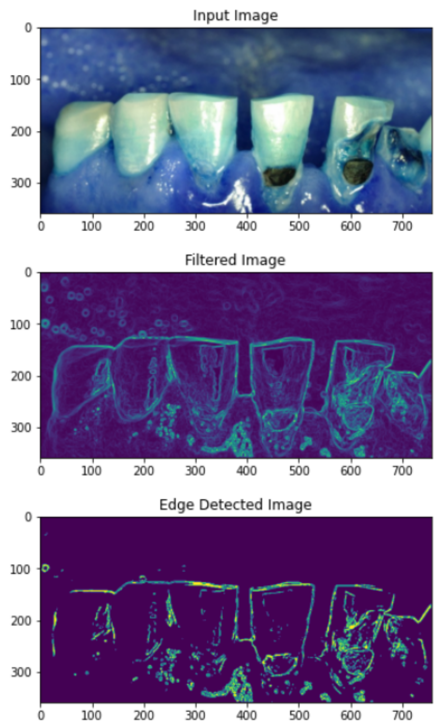
6 Implementation

A number of image processing techniques were tried to see on what basis an analysis could be done. In the end, analysis on the basis of image sampling was chosen. Some of the other processing techniques that were tried are :

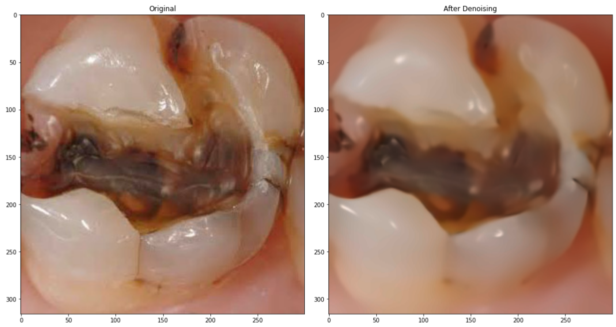
6.1 Contrast Stretching using Histogram



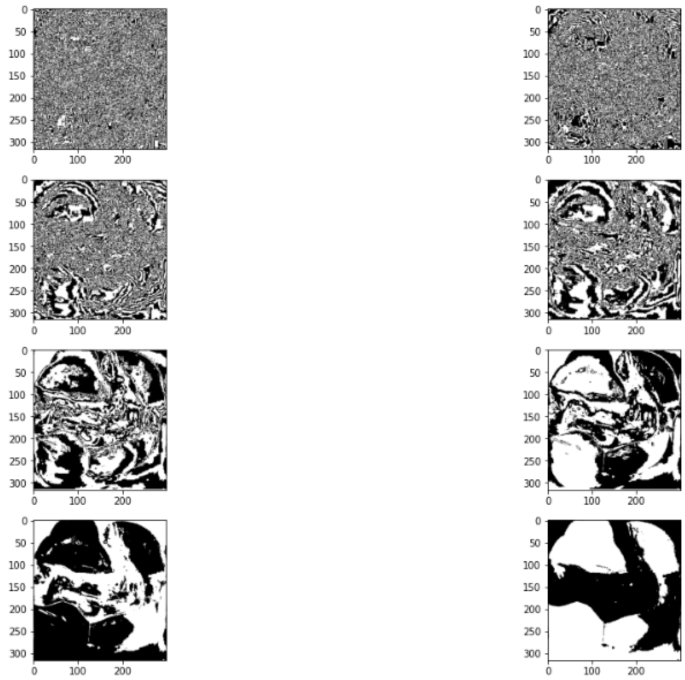
6.2 Sobel Operator for Edge Detection



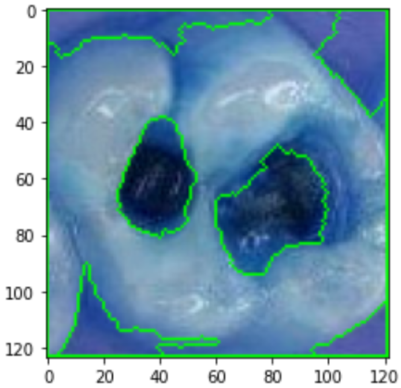
6.3 Denoising the Image



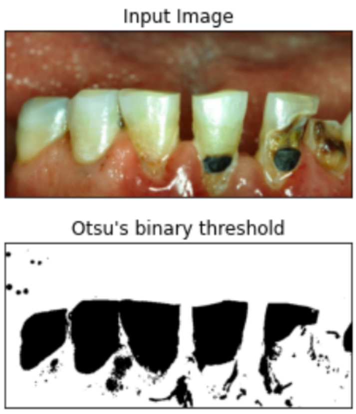
6.4 Bit Plane Slicing



6.5 Applying Watershed Segmentation

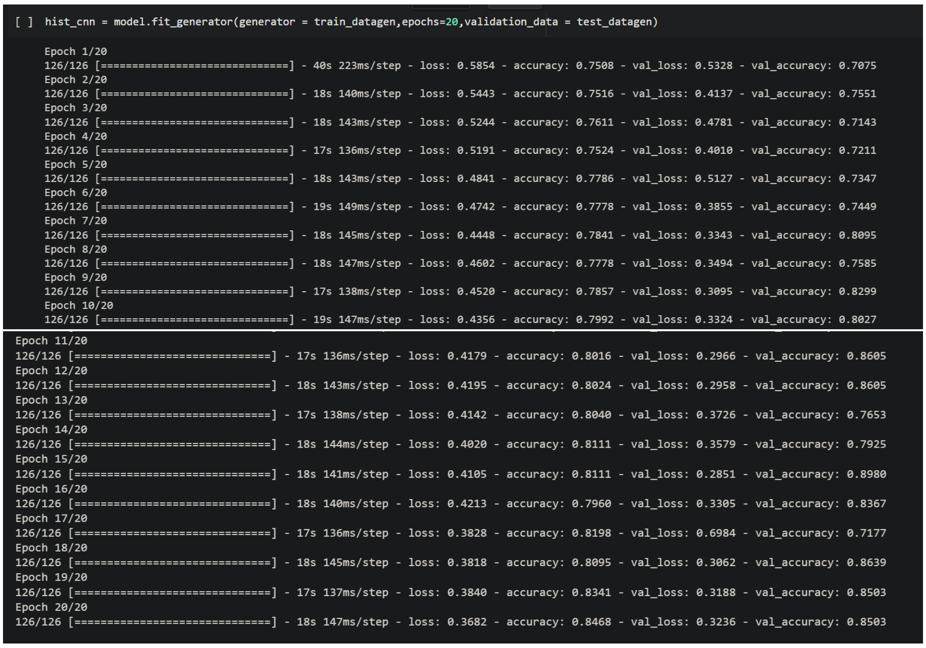


6.6 Otsu's Image segmentation (Threshold-based segmentation)



6.7 Image Sampling

Using a python code, the dataset was re-sampled into the 5 categories as mentioned in the above section. So basically there were 5 datasets containing the same images but different resolutions. Each dataset was run through the 5 deep learning models. The training accuracy, training loss, validation accuracy and validation loss was noted down for each case. Epochs of the suggested CNN model are displayed in Fig 5.



```
[ ] hist_cnn = model.fit_generator(generator = train_datagen,epochs=20,validation_data = test_datagen)

Epoch 1/20
126/126 [==============================] - 40s 223ms/step - loss: 0.5854 - accuracy: 0.7508 - val_loss: 0.5328 - val_accuracy: 0.7075
Epoch 2/20
126/126 [==============================] - 18s 140ms/step - loss: 0.5443 - accuracy: 0.7516 - val_loss: 0.4137 - val_accuracy: 0.7551
Epoch 3/20
126/126 [==============================] - 18s 143ms/step - loss: 0.5244 - accuracy: 0.7611 - val_loss: 0.4781 - val_accuracy: 0.7143
Epoch 4/20
126/126 [==============================] - 17s 136ms/step - loss: 0.5191 - accuracy: 0.7524 - val_loss: 0.4010 - val_accuracy: 0.7211
Epoch 5/20
126/126 [==============================] - 18s 143ms/step - loss: 0.4841 - accuracy: 0.7786 - val_loss: 0.5127 - val_accuracy: 0.7347
Epoch 6/20
126/126 [==============================] - 19s 149ms/step - loss: 0.4742 - accuracy: 0.7778 - val_loss: 0.3855 - val_accuracy: 0.7449
Epoch 7/20
126/126 [==============================] - 18s 145ms/step - loss: 0.4448 - accuracy: 0.7841 - val_loss: 0.3343 - val_accuracy: 0.8095
Epoch 8/20
126/126 [==============================] - 18s 147ms/step - loss: 0.4602 - accuracy: 0.7778 - val_loss: 0.3494 - val_accuracy: 0.7585
Epoch 9/20
126/126 [==============================] - 17s 138ms/step - loss: 0.4520 - accuracy: 0.7857 - val_loss: 0.3095 - val_accuracy: 0.8299
Epoch 10/20
126/126 [==============================] - 19s 147ms/step - loss: 0.4356 - accuracy: 0.7992 - val_loss: 0.3324 - val_accuracy: 0.8027
Epoch 11/20
126/126 [==============================] - 17s 136ms/step - loss: 0.4179 - accuracy: 0.8016 - val_loss: 0.2966 - val_accuracy: 0.8605
Epoch 12/20
126/126 [==============================] - 18s 143ms/step - loss: 0.4195 - accuracy: 0.8024 - val_loss: 0.2958 - val_accuracy: 0.8605
Epoch 13/20
126/126 [==============================] - 17s 138ms/step - loss: 0.4142 - accuracy: 0.8040 - val_loss: 0.3726 - val_accuracy: 0.7653
Epoch 14/20
126/126 [==============================] - 18s 144ms/step - loss: 0.4020 - accuracy: 0.8111 - val_loss: 0.3579 - val_accuracy: 0.7925
Epoch 15/20
126/126 [==============================] - 18s 141ms/step - loss: 0.4105 - accuracy: 0.8111 - val_loss: 0.2851 - val_accuracy: 0.8980
Epoch 16/20
126/126 [==============================] - 18s 140ms/step - loss: 0.4213 - accuracy: 0.7960 - val_loss: 0.3305 - val_accuracy: 0.8367
Epoch 17/20
126/126 [==============================] - 17s 136ms/step - loss: 0.3828 - accuracy: 0.8198 - val_loss: 0.6984 - val_accuracy: 0.7177
Epoch 18/20
126/126 [==============================] - 18s 145ms/step - loss: 0.3818 - accuracy: 0.8095 - val_loss: 0.3062 - val_accuracy: 0.8639
Epoch 19/20
126/126 [==============================] - 17s 137ms/step - loss: 0.3840 - accuracy: 0.8341 - val_loss: 0.3188 - val_accuracy: 0.8503
Epoch 20/20
126/126 [==============================] - 18s 147ms/step - loss: 0.3682 - accuracy: 0.8468 - val_loss: 0.3236 - val_accuracy: 0.8503
```

Fig. 5: Epochs of Proposed CNN Model

7 Results and Discussion

Besides accuracy, loss was also plotted for each model. Both were shown over the course of the epochs. In Fig.6 and Fig.7, the plots have been displayed for the proposed CNN model. These plots are done on the original dataset. It is noticed that the

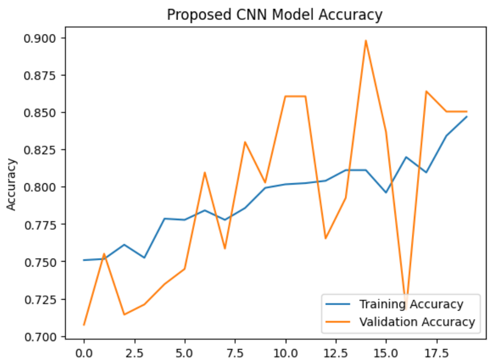


Fig. 6: Accuracy Plot for Proposed CNN Model

training accuracy and loss are much more of a gradual slope as compared to the validation accuracy and loss which is far more volatile. However, the respective graphs

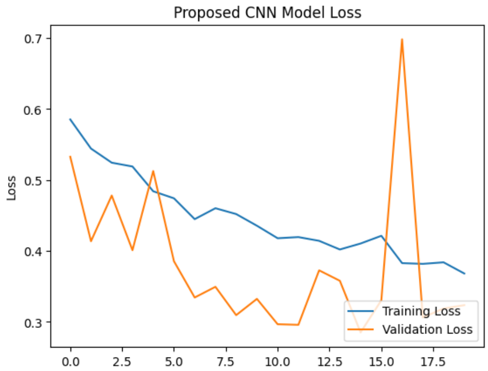


Fig. 7: Loss Plot for Proposed CNN Model

are headed in the right direction and the final training accuracy is almost equal to the final validation accuracy. These are positive signs indicating that the proposed CNN model has performed well. The evaluation metrics of each model trained on the original dataset are shown in Table 1.

Table 1: Evaluation Metrics per Model

Models	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
Proposed CNN Model	0.847	0.850	0.368	0.324
InceptionV3	0.845	0.837	0.368	0.374
MobileNetV2	0.861	0.786	0.317	0.484
VGG16	0.853	0.772	0.347	0.527
ResNet50	0.746	0.714	0.491	0.451

Table 2: Validation accuracy for each Model as per the respective Dataset

Models	Original Dataset	64x64 Dataset	256x256 Dataset	512x512 Dataset	1024x1024 Dataset
Proposed CNN Model	0.850	0.891	0.871	0.823	0.871
InceptionV3	0.837	0.837	0.833	0.850	0.827
MobileNetV2	0.786	0.803	0.816	0.861	0.857
ResNet50	0.714	0.830	0.806	0.714	0.813
VGG16	0.772	0.833	0.803	0.847	0.793

Fig. 8: Testing the model on a Random Image

8 Conclusion

A number of observations were inferred from the results in Table 2. The proposed CNN Model has performed the best in 3 out of 5 datasets. This indicates that it is expceptionally good considering the fact that the other models are pre-trained transfer learning. The 256x256 dataset has lower accuracies as compared to the other modified datasets but still considerably decent. The 4 modified datasets have higher accuracy as compared to the original dataset. This is probably because the original dataset has images of mixed resolution.

InceptionV3 proved to be the best transfer learning model, followed by MobileNetV2. ResNet50 and VGG16 were usually below the average accuracy.

As seen in Fig 8, the proposed CNN model correctly classifies an input image into one of the 4 categories - healthy tooth (no decay), enamel decay, dentin decay, pulp decay. Thus, it can be used to assist a medical practitioner to classify complex cases of cavity detection.

9 Statements and Declaration

9.1 Author Contributions

J.M. and M.G. chose the dataset; J.M. re-sampled the dataset; J.M. and S.H. implemented the proposed model and compared it with transfer learning models; J.M. and S.H. prepared the original draft submission; M.G. reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

9.2 Funding

This research received no external funding.

9.3 Conflict of Interests

The authors declare no conflict of interests.

9.4 Acknowledgements

Not applicable

10 Data Availability Statement

The datasets generated during and/or analysed during the current study are available in :

- **Teeth Dataset scaled to 64x64 pixels:**
  https://www.kaggle.com/datasets/jasonfm26/sampling
  Code :https://www.kaggle.com/code/jasonfm26/evaluating-models-on-the-64x64-dataset

- **Teeth Dataset scaled to 256x256 pixels:**
  https://www.kaggle.com/datasets/jasonfm26/changed
  Code :https://www.kaggle.com/code/jasonfm26/evaluating-models-on-the-256x256-dataset

- **Teeth Dataset scaled to 512x512 pixels:**
  https://www.kaggle.com/datasets/jasonfm26/sample
  Code :https://www.kaggle.com/code/jasonfm26/evaluating-models-on-the-512x512-dataset

- **Teeth Dataset scaled to 1024x1024 pixels:**
  https://www.kaggle.com/datasets/jasonfm26/bigfile
  Code :https://www.kaggle.com/code/jasonfm26/evaluating-models-on-the-1024x1024-dataset

- **Original Teeth Dataset :**
  https://www.kaggle.com/datasets/snginh/teethdecay
  Code :https://www.kaggle.com/code/jasonfm26/evaluating-models-on-the-original-dataset
  Code :https://www.kaggle.com/code/jasonfm26/testing-proposed-cnn-model-on-random-image

NOTE: All 5 datasets were used as we performed a comparative study of transfer learning models on different datasets. The codes run on each dataset are also provided in these links. They are displayed under the “code” section.

References

[1] Lee, S., Oh, S.-i., Jo, J., Kang, S., Shin, Y., Park, J.-w.: Deep learning for early dental caries detection in bitewing radiographs. Scientific reports **11**(1), 16807 (2021)

[2] Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Transactions on knowledge and data engineering **22**(10), 1345–1359 (2009)

[3] Lee, J.-H., Kim, D.-H., Jeong, S.-N., Choi, S.-H.: Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. Journal of dentistry **77**, 106–111 (2018)

[4] Ajins, T., Anitha, C., Anupama, S., Thomas, A., Jacob, M., *et al.*: Imaging modalities for general dental practice-bird's eye view. Journal of Academy of Dental Education **2**(1), 11–14 (2016)

[5] Hung, M., Voss, M.W., Rosales, M.N., Li, W., Su, W., Xu, J., Bounsanga, J., Ruiz-Negrón, B., Lauren, E., Licari, F.W.: Application of machine learning for diagnostic prediction of root caries. Gerodontology **36**(4), 395–404 (2019)

[6] Na'am, J., Harlan, J., Madenda, S., Wibowo, E.P.: Image processing of panoramic dental x-ray for identifying proximal caries. TELKOMNIKA (Telecommunication Computing Electronics and Control) **15**(2), 702–708 (2017)

[7] Perez, L., Wang, J.: The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621 (2017)

[8] Miglani, S.: Burden of dental caries in india: Current scenario and future strategies. International journal of clinical pediatric dentistry (2020)

[9] Bagramian, R.A., Garcia-Godoy, F., Volpe, A.R., *et al.*: The global increase in dental caries. a pending public health crisis. Am J dent **22**(1), 3–8 (2009)

[10] Patil, S., Kulkarni, V., Bhise, A.: Algorithmic analysis for dental caries detection using an adaptive neural network architecture. Heliyon **5**(5) (2019)

[11] Nath, S.S., Mishra, G., Kar, J., Chakraborty, S., Dey, N.: A survey of image classification methods and techniques. In: 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), pp. 554–557 (2014). IEEE

[12] Xin, M., Wang, Y.: Research on image classification model based on deep convolution neural network. EURASIP Journal on Image and Video Processing **2019**, 1–11 (2019)

[13] Sharma, N., Jain, V., Mishra, A.: An analysis of convolutional neural networks for image classification. Procedia computer science **132**, 377–384 (2018)

[14] He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1026–1034 (2015)

[15] Wei, Y., Zhang, Y., Huang, J., Yang, Q.: Transfer learning via learning to transfer.

In: 35th International Conference on Machine Learning, ICML 2018, vol. 11, p. 8059 (2018)

[16] Yousefi, J.: Image binarization using otsu thresholding algorithm. Ontario, Canada: University of Guelph **10** (2011)

[17] Buades, A., Coll, B., Morel, J.-M.: A review of image denoising algorithms, with a new one. Multiscale modeling & simulation **4**(2), 490–530 (2005)

[18] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

Original Teeth Dataset

Click here to download Link(s) to supporting data
http://www.kaggle.com/datasets/snginh/teethdecay/code