Automated Dental Cavity Detection System Using Artificial Intelligence

Niharika Bhattacharjee

Personal Statement

It all started when I thought I had gingivitis. I came to this conclusion after consulting the most reputable source I knew, webmd.com of course.

Sitting at the doctor's office, I anxiously waited for the confirmation. Were my suspicions correct? Nope!

As my doctor explained my negative result, guilt slowly overcame me. I had just wasted three hours of my dad's time.

Driving home, I was determined to find meaning in this seemingly unfulfilling trip, and indeed I did. The visit highlighted my privilege, which ignited my "eureka" moment. A dental diagnosis requires time, a medical professional, and money. Three things millions of people globally don't have; a problem I sought to address.

With the aim of increasing dental care worldwide, I invested the following two years in creating an automated dental cavity detection system using Artificial Intelligence (AI). With a single dental photographic color image, my system can provide a cavity diagnosis and explain the diagnosis to the end-user in an understandable manner. By automating detection and explainability, which are skills dentists typically employ, I hope to assist those who can't visit the dentist due to lack of dental insurance, dentophobia, or limited dentist availability.

Through this research, I had the opportunity to work with an IBM researcher and experiment with different deep learning architectures, explainable AI algorithms, and training techniques such as curriculum learning and transfer learning. By incorporating deep learning into my project, I gained experience in an interdisciplinary field powered by linear algebra, calculus, statistics, and biology. Above all, conducting such research has motivated me to pursue a career in computational health.

To future high school researchers, curiosity and passion must come from within. For me, a personal experience prompted my exploration into the use of AI in the dentistry field. In order to conduct a research project, motivation is key. Thus, as you embark on your research journey, I encourage you to answer these important questions: Who do you want to help? How do you want to help them? Can people of all socioeconomic backgrounds benefit from your proposal? The answers you provide are all puzzle pieces. Once you put these pieces together, your own "eureka" moment will emerge.

Research Section

Abstract

Dental cavities impact over 3.9 billion people worldwide. Typically, cavity detection requires a trained dentist. However, barriers such as dentophobia, limited dentist availability, and lack of dental insurance prevent millions from receiving dental care. To address this, the student researcher created an Artificial Intelligence (AI) diagnostic system. The system uses an artificial neural network (ANN) to detect cavity presence on photographic color images and visually explains the rationale behind each diagnosis. Previous studies have only focused on cavity detection on extracted teeth showing the occlusal tooth surface. In contrast, this study's system detects cavities on images showing multiple teeth and three additional tooth surfaces (occlusal, lingual, buccal, labial). To train the system, 314 de-identified images from online sources were collected. Using transfer learning from an ImageNet1k dataset, the ResNet-27 architecture proved to be most optimal in cavity detection after earning 77.8% accuracy. Visual explanations for the system's cavity diagnoses were also generated using an algorithm called Local Interpretable Model Agnostic Explanations (LIME). After applying LIME, the system now has the ability to explain its diagnosis to an end-user in an understandable manner, which is a crucial skill employed by dentists. In the future, the student researcher aims to enhance the system's performance using semi-supervised learning techniques and additional data.

Introduction

Statement of Purpose

Dental cavities, one of the most prevalent chronic diseases in the world, impacts over 3.9 billion people worldwide (NIH, 2018). Of which 2.4 billion people suffer from untreated tooth decay, which can lead to toothaches and periapical abscesses (TIME, 2015). Typically, cavity detection requires the services of a trained dentist; however, 1 in 3 Americans don't visit a dentist on an annual basis due to cost or dentophobia (Vox, 2014). Furthermore, chances of proper cavity diagnoses are lowered in developing nations, as the dentist to patient ratio can be disproportionate (i.e., 1: 1,250,000 in Ethiopia (WHO, 2017)). Factors such as cost and limited availability of dentists often prevent people from receiving proper cavity diagnoses. To address these issues, a cavity diagnosing system using an artificial intelligence (AI) approach would be advantageous to increase access to dental care at lower costs.

Review of Literature

Limitations of Current Diagnostic Methods due to Technology and Cost

To diagnose a cavity, a dentist is often aided by x-rays or optical diagnostic devices such as the DIAGNodent™, DEXIS CariVu™, and the Canary System™. There are many limitations associated with the optical diagnostic devices mentioned above (Silvertown et al., 2017). Instead of detecting Streptococcus mutans (the cavity-causing bacteria which resides at the cavity), the DIAGNodent only detects fluorescence emitted by bacterial porphyrins, a common bacterium found everywhere in the mouth. The use of trans-illumination prevents the CariVu from detecting cavities on smooth surfaces and at around filling, while the Canary System's limiting factor is cost. At $16,000, the Canary System costs three times more than the CariVu at $5,295 (Chicago Tribune, 2019). Above all, these systems' outputs require a trained dentist to interpret.

Visual Examination

Visual examination is also commonly used for cavity diagnosis. It provides a natural way for dentists to convey their diagnosis to their patients. During the exam, a dentist can identify a patient's cavity, inform the patient about the cavity's location and development, and suggest changes to the patient's oral routine. Since the method doesn't require any equipment, it's also the most cost-efficient from an equipment perspective. However, visual examination still has one requirement that prevents millions of people from receiving a proper cavity diagnosis; the method requires a dentist. Access to dental care, in general, is difficult for many as cost, fear of pain, and lack of dental insurance pose as barriers (Gupta & Vujicic, 2019). Some of these barriers could be addressed with an AI diagnostic system that detects cavity presence and provides a diagnostic explanation at lower cost and without an initial dental visit.

Artificial Intelligence (AI)

The field of AI is developing technologies that enable machines to mimic human behavior. It has been applied to several problems in medical imaging (i.e. Codella et al. (2019) for Melanoma detection, Akselrod-Ballin et al. (2019) for breast cancer prediction, Gulshan et al. (2016) for diabetic retinopathy diagnosis from fundus imaging). Machine learning (ML), a subset of AI, consists of algorithms used by machines to perform a specific task after undergoing a learning period. While classical ML algorithms rely on subject matter experts to hand craft features to extract from raw data prior to learning, deep learning (DL) consists of novel techniques that can bypass this step (Grossfeld, 2020). In imaging, DL typically uses image training/testing datasets and a multi-layer artificial neural network (ANN) architecture to learn how to complete tasks that normally require human intelligence (ex. diagnosing a cavity).

Previous Research in Use of AI in Cavity Detection

There is limited research available on diagnostic systems using AI algorithms in the dental field. Berdouses et al. (2015) used a ML algorithm called Random Forest (RF) and an object detection technique called instance segmentation to classify and highlight cavities in 103 images (Figure 1). To obtain labels for the image set, the images were provided to 2 dentists who identified and diagnosed 425 regions of interests using the International Caries Detection and Assessment System (ICDAS) rubric. The images were then fed into their RF algorithm, which correctly diagnosed 340 of the 425 regions of interest (80%).

Figure 1. The RF algorithm first isolated a region of interest (ROI) (Figure 1b). For each pixel present in the ROI, 65 features depicting the level of decay and its ICDAS class # were fed into the algorithm. The pixels in the ROI's center with the highest-class value (red marking) was used to classify the entire region. The algorithm's detection (b) and classification (c) results can be compared to the dentist's annotations as seen in Figure 1a. (Berdouses et al., 2015).

Limitations of Previous Research

The images and algorithms used in the study by Berdouses et al. (2015) did not reflect field conditions. The study's images were not reflective of what an operator of a cavity classifier would receive from future users. Each of their training images contained a single tooth, an assumption that is not realistic in practice as patients' images would consist of a mouthful of teeth. Additionally, their RF algorithm was only trained with images showing the occlusal tooth surface, and the algorithm was untrained in detecting cavities on labial, buccal, and lingual tooth surfaces. The study's ML classifier also required a data scientist to physically extract each image feature needed to make a diagnosis. The student researcher's study aims to address the limitations present in Berdouses et al.'s study.

Benefits of Using Photographic Color Images over X-Rays

While there are limited studies that have trained DL classifiers for cavity detection on photographic color images, there are many studies that have trained DL classifiers to detect cavities on X-ray images (Lee et al., 2018; Ali et al., 2016). However, since X-ray machines aren't portable, it would still require a costly visit to the dentist. A complete series of radiographic images costs $189 (How Much Do Dental X-Rays Cost Without Insurance, n.d.). For the millions of people who can't afford a dental visit, any classifier requiring X-rays is difficult to access. On the other hand, photographic images of teeth are easier to obtain by people in general and certainly in developing countries. In fact, over 3 billion people worldwide have built-in cameras in their smartphones of which 45% of the smartphone users are from emerging economies (Silver, 2019). Cheaper options are available as well, such as intraoral cameras for under $30.00. Not everyone is familiar with what their teeth look like in an X-ray image, but everyone is familiar with what their teeth look like in a photo. Thus, users can easily interpret and trust a classifier's diagnosis on photographic images, with the help of explainable AI techniques such as Local Interpretable Model-Agnostic Explanation (LIME).

Explainable AI

Local Interpretable Model-Agnostic Explanation (LIME). To develop end user trust, it's extremely important that a classifier trained to detect cavities on photographic color images can produce accurate and interpretable results. From an accuracy perspective, a false-positive diagnosis may be harmful to a patient's mental health and finances while a false negative can have significant implications on the patient's general health. Trust in AI is an emerging field. A popular explainer algorithm applied on image classifiers to develop trust is called LIME (Ribeiro et al., 2016). For a given medical image, LIME can overlay a visual explanation on top of the image, highlighting the important regions of the image that influenced the classification result (Ribeiro et al., 2016). Such explanations help the end user understand the rationale behind the model's diagnosis (Figure 2). If LIME were implemented on an accurate DL cavity detection classifier, the visual explanation would highlight which tooth surfaces heavily influenced the classifier's diagnosis. These explanations would also inform patients about the specific tooth surfaces that need more brushing. While LIME would be extremely effective, there are no trained DL classifiers publicly available that detect cavity presence on photographic images, while providing an explanation for the diagnosis using an explainable AI technique.

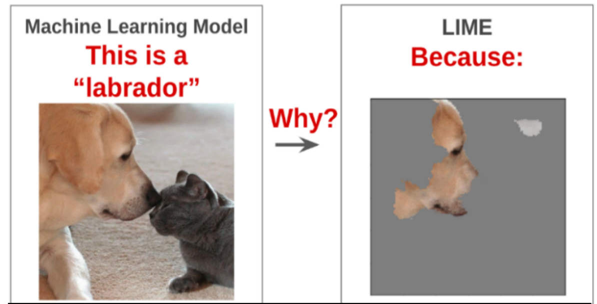

Figure 2. Visual shows a general overview of the LIME algorithm (Arteaga, 2019).

Methodology

General Overview

The following methodology was used to develop an AI diagnostic system that detects cavity presence on photographs and visually explains the rationale behind each diagnosis using an algorithm called LIME. With the aim of selecting an ANN architecture that was optimal for detecting cavities, the student researcher designed and trained three different ANN architectures before evaluating their performance in cavity detection using a 314-image dataset consisting of cavitated and healthy teeth images. The architectures the student researcher experimented with included a hand-designed 12-layer Convolutional Neural Network (CNN) and various extensions of pre-trained image classifiers (ResNet-18, ResNet-27). As training data was limited, transfer learning was applied to better train the networks. In addition, the capabilities of LIME were explored to provide visual explanations that could be easily interpreted by the end-users.

Training and Testing

To train the 12-layered CNN, ResNet-18, and ResNet-27 models to detects cavities on photographic color images, de-identified photographic color images exhibiting cavity presence and no cavity presence were collected by the student researcher. These images were then split into a training and testing dataset at the ratio of 80: 20, which is a standard ratio used in machine learning (Detective, 2020). By undergoing training, each ANN architecture learned how to extract the specific features or patterns (ex. shading, discoloration, and texture) associated with cavity presence. In the testing phase, each ANN architecture's performance in cavity detection was determined using the testing dataset, whose images were not seen during training.

2019 Web-Searched Dataset. To create the training and testing datasets, 314 de-identified photographic color images showing cavitated or healthy teeth. These images were taken from online sources (dental blogs, dental presentations, and journals). A dentist was then contacted to professionally diagnose the images based on the International Caries Detection and Assessment System (ICDAS) rubric (Figure 3). These diagnoses were requested so that each image had a corresponding ground truth label. During training, these labels were used to guide each ANN in learning how to detect cavities on the training dataset. As training progressed, each ANN adjusted its weights by calculating how different its predictions of the training dataset were from the labels using a loss function. The cavities present in the dataset were representative of all levels of decay. This is evident as the dentist used all ICDAS class values ranging from 0 (no cavity presence) to 6 (extensive lesion) to label the dataset. Images labeled ICDAS class 0 (no cavity presence) were represented with label - 0, and images labeled with ICDAS classes 1-6 (cavity presence) were represented with label - 1. However, all ICDAS classes were used to analyze ANN's performance. Using the labeling method above, out of the 314 images, 185 images had no cavities and 129 images had presence of a cavity. These 314 were then divided into the 251 training images (80%) and 63 testing images (20%).

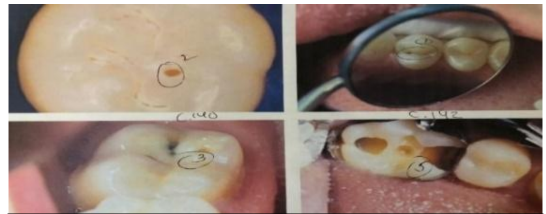

Figure 3. Examples of the dentist's annotations. (Figure by N.Bhattacharjee).

ANN Architecture 1: 12-layered Hand Designed CNN. The first ANN architecture the student researcher experimented with was a 12-layered Convolutional Neural Network (CNN), implemented on PyTorch (DL framework). This CNN's architecture was designed from scratch (Figure 4) with random weights. The training method used was simple supervised learning, in which the CNN gradually adjusted its weights using back propagation after calculating its loss. The first layers of the CNN were convolutional layers, which can extract low level features like the edges and curves of a cavity (feature extraction). The CNN's hidden layers aimed to prevent over-training, while the fully connected layers were used for classification. As the CNN was provided with limited training data, the

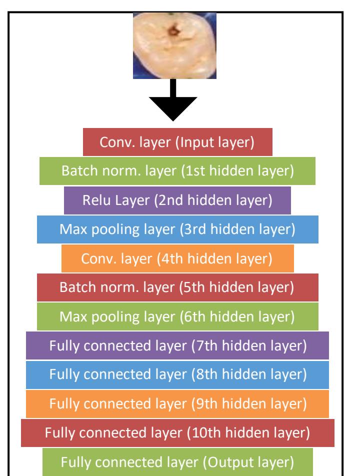

Figure 4. Framework of the initial 12-layered CNN (Figure by N.Bhattacharjee).

convolution layers were limited in their ability to learn features typically used to properly diagnose cavities. CNN architectures require a large amount of data to train properly. Collecting images was a challenge as there are no public databases present that provide cavitated and healthy teeth image datasets. The lack of data resulted in the neural network's classification accuracy to be poor. Under these circumstances, supervised learning on a CNN starting with random weights was not the right approach. Hence, the student researcher resorted to more advanced ML techniques to enhance performance.

Improving Using Transfer Learning. To improve performance while training with limited data, transfer learning was used. It is well known in the deep learning field that low-level features (ex. lines or curves) of a limited training dataset are generic features. These features are present and learnable on any large image training dataset (even a dataset of cats and dogs). When dealing with limited data, a model's performance is maximized if it previously learns general low-level features on a larger dataset independently, retains that knowledge, and then specializes in learning the higher-level features using the limited dataset. As a result, the student researcher experimented with two general purpose CNN models (ResNet-18 and ResNet-27) that were previously trained on the ImageNet1K dataset, which has over 1.2 million images. These models had already learned the low-level features needed for the detection of various shapes, including the ones needed for cavity detection. Their knowledge was preserved as the weights in the CNNs' preliminary layers were kept nearly unchanged while training on the limited dataset. The student then extended these pretrained models with additional neural network layers trained with the 314-image dental dataset. Hence, the resulting models became specialized in cavity detection since their deeper and last layers were permitted to change to learn the high-level features specific to the limited dental dataset, thus improving performance.

ANN Architecture 2 and 3: ResNet 18 and ResNet 27. The two pre-trained ANN architectures used were a ResNet-18 and ResNet-27. They were variants of the Residual Network (ResNet) architecture. The ResNet-18 model had 18 convolutional layers with residual blocks/skipping blocks (Figure 5) and 1 fully connected layer. The ResNet-27 had 27 convolutional layers with residual blocks and 1 fully connected layer. Both Res-Net variants used

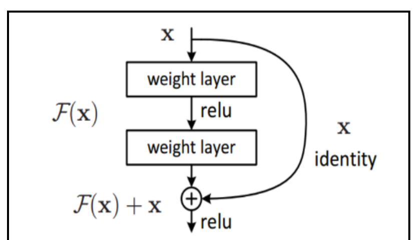

Figure 5. Architecture of a residual block (Sahoo, 2018).

residual blocks as part of their design to solve the vanishing gradient (VG) problem that arises for ANNs with many layers (Wang, 2019).

Local Interpretable Model-Agnostic Explanations (LIME). Aside from experimenting with transfer learning, the capabilities of LIME in providing explanations for a model's cavity diagnosis were also explored in a PyTorch environment. LIME promotes trust by presenting an explanation for a given model's prediction. It produces the explanation by approximating the behavior of the complex model with very simple linear models (i.e. hyperplanes). LIME was applied on a ResNet-18 model trained on the dental data. This was done since PyTorch offered built-in support for ResNet-18 architectures. The explanations generated by LIME successfully highlighted the image regions that drove the ResNet-18 to its diagnoses. Specifically, when the ResNet-18 predicted cavity presence for a given image, LIME's explanation highlighted the tooth surfaces that heavily influenced the model's diagnosis. By applying LIME, the study's AI diagnostic system now has the ability to explain its cavity diagnosis to future end-users in an understandable manner.

Equipment Used. For ResNet-27 experiments, the student researcher used Caffe, which is a deep learning framework written in C++. Experiments conducted on the ResNet-18 architecture and the hand-coded 12-layered CNN architecture were done on PyTorch, a Python-based deep learning framework. The student researcher implemented this framework in Jupyter Notebook, a web-based interactive computational environment.

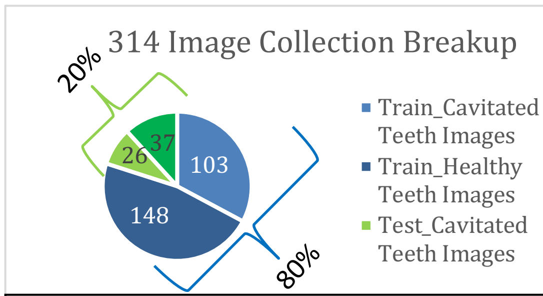

Figure 7. Breakup of experimental data used in training and evaluation (Figure by N.Bhattacharjee).

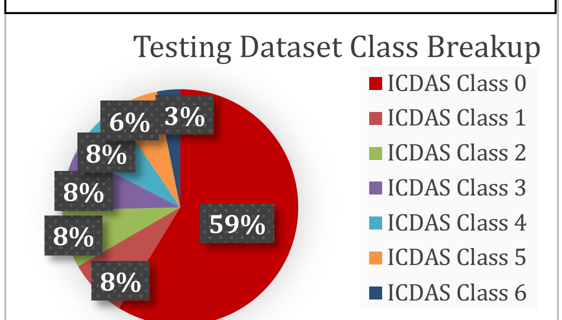

Figure 8. The test dataset representing all levels of decay. (Figure by N.Bhattacharjee).

Results

2019 Web-Searched Dataset. The 314-image collection acquired from public sources was split into a training and testing dataset in the ratio of 80:20. The training and testing dataset consisted of 251 and 63 images respectively (Figure 7). The datasets consisted of all 7 classes of decay as per the ICDAS rubric in the proportions shown in Figure 8. Since the student researcher's main objective was to diagnose images based on cavity presence only (not into the specific 7 ICDAS classes), the ICDAS 1-6 images were combined into one group (cavity presence), and ICDAS class

0 images were in another (no cavity presence). However, analyses were done using all 7 ICDAS classes.

Artificial Neural Network (ANN) Models. The performances of the hand-designed CNN, ResNet-18, and ResNet-27 models in cavity detection were compared. Table 1 by N.Bhattacharjee shows each model's architecture in terms of their layers and whether the architecture had a pre-trained ImageNet1K base model available.

Table 1: Summary of Each ANN Model's

	Initial CNN	ResNet-18	ResNet-27
Number (#) of Layers	12	18	27
# of Convolutional Layers	3	17	26
# of Fully Connected layers	5	1	1
ImageNet1K base model available for transfer learning	No	Yes	Yes

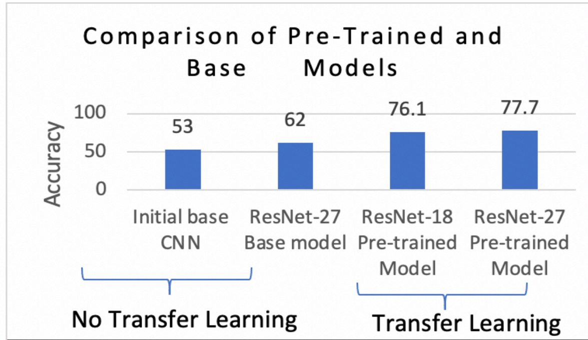

Figure 9. Overall performance of the 3 architectures while transfer learning experiments were being conducted (Figure by N. Bhattacharjee).

Performance of ANN Architectures

The Benefits of Using Transfer Learning. Models pre-trained on the ImageNet1K dataset produced higher accuracy compared to their base models. Pre-training allowed for better learning, which is supported by the accuracies presented in Figure 9. Without pre-training, 62% accuracy was achieved by the ResNet-27. However, with the use of pre-trained weights, accuracy increased to 77.7%.

Initial 12-layered CNN. As the CNN model wasn't pre-trained on a larger data set, the model's performance in detecting cavities was similar to random guessing (50%). As accuracy decreased with training, it was concluded that the model was inadequate in detecting cavities (Figure 10).

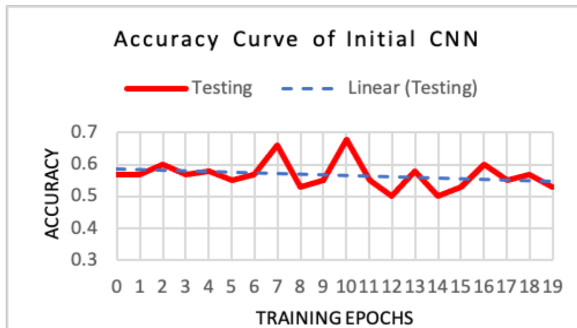

Figure 10. Although accuracy did jump at epoch 11, learning is usually indicated by a general upward accuracy trend, which was not present during testing (Figure by N.Bhattacharjee).

Pre-trained ResNet-18. Both ResNet models were pre-trained on the ImageNet1K dataset. However, the ResNet-18 diagnosed the 63-image test set with an accuracy of 76.1%. Unlike the initial CNN, the ResNet-18's accuracy values increased as more training epochs were completed (Figure 11). This indicated that pre-training in conjunction with using a ResNet architecture improved a classifier's diagnostic ability.

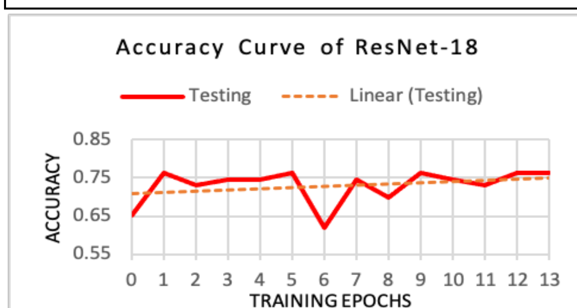

Figure 11. Ideal upward accuracy trend was exhibited. In the 6th epoch accuracy decreased to 69% but increased 10% by the 12th (Figure by N.Bhattacharjee).

Pre-trained ResNet-27. Highest accuracy values were obtained from the pre-trained ResNet-27 model on Caffe. The model diagnosed the 63-image test set with an accuracy of 77.8% on average (49/63 images) after training for 600 iterations. In the 500th iteration, accuracy increased to 79% and then plateaued 77.8%. For no cavity cases (class 0), an accuracy of 83.8% was obtained (31/37 images). For cavity cases (class 1-6), 70% accuracy was produced (18/26 images) (Figure 12). To determine which ICDAS classes were the most difficult for the model to detect, further analysis was done. For extreme decay (ICDAS 5-6), the model correctly diagnosed 5 out of 6 test cases (accuracy of 85.7%). For moderate level of decay (ICDAS class 3-4), the model correctly diagnosed 7 out of 9 cases (accuracy of 77.8%) (Figure 13). These results indicated that the ResNet-27 model's accuracy decreased as the size of the cavity decreased. Results also indicated that it was easier for the model to detect colored lesions compared to cavities with white plaque (early decay).

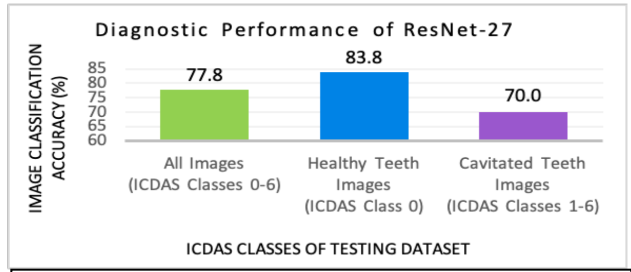

Figure 12. The ResNet-27 model's hyperparameters associated with the highest accuracy included a 64-image batch size, a learning rate of .001, and a step size of 1500. (Figure by N.Bhattacharjee)

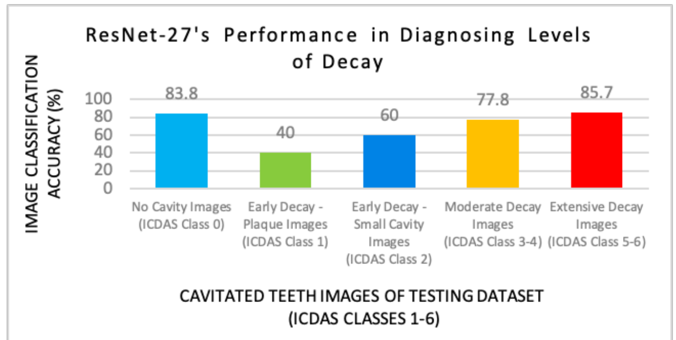

Figure 13. The model provided the most accurate diagnoses for no cavity, moderate decay, and extensive decay cases. (Figure by N.Bhattacharjee)

ResNet-27 Confusion Matrix Results. To determine the type of diagnostic errors the ResNet-27 model made, a confusion matrix was also generated (Figure 14). The model earned a sensitivity/recall score of .69, which indicated that the model can correctly identify patients with a cavity 7 out of 10 cases. The model also obtained a specificity score of .84 (model can correctly identify patients with no cavities 84 out of 100 cases). Finally, the precision score of .75 indicated that for every 100 cases the model diagnoses as having a cavity, 75 of them would actually exhibit cavity presence. By using a ResNet-27 architecture, the objectives involving training a model to detect cavities of varying levels of decay and on different surfaces were accomplished.

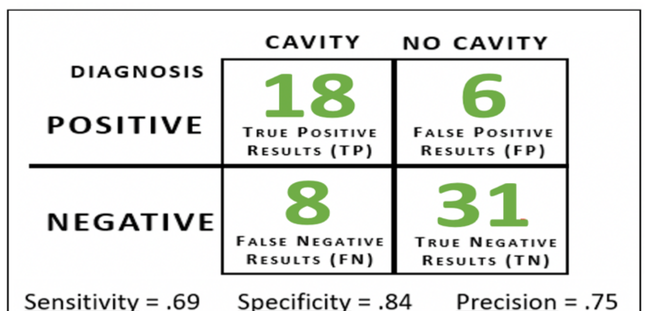

Figure 14. The red numbers represent the distribution of the 63-image test set in terms of how the model diagnosed the images (Pos. for cavity/Neg. for cavity) in comparison to the dentist's ground truth diagnoses (Cavity, No Cavity). (Figure by N.Bhattacharjee)

After comparing the performances of the 3 architectures, it was concluded the ResNet-27 model was most optimal for cavity detection.

Explaining Model Output Using LIME Algorithm. The capabilities of LIME in providing explanations for a model's cavity diagnosis were explored in PyTorch. LIME was applied on a pre-trained ResNet-18 model. The figures below represent outputs of the LIME algorithm, which are sample explanations that highlight the image regions that most influenced the ResNet-18 model's diagnoses. Figure 15 presents a sample LIME explanation that was generated after the model correctly predicted cavity presence. Figure 16 presents a sample explanation that was generated after the model correctly predicted the absence of a cavity (no cavities). After examining several explanations that LIME generated, it was concluded that LIME can effectively explain a model's cavity diagnosis and should be a key component of the cavity detection system.

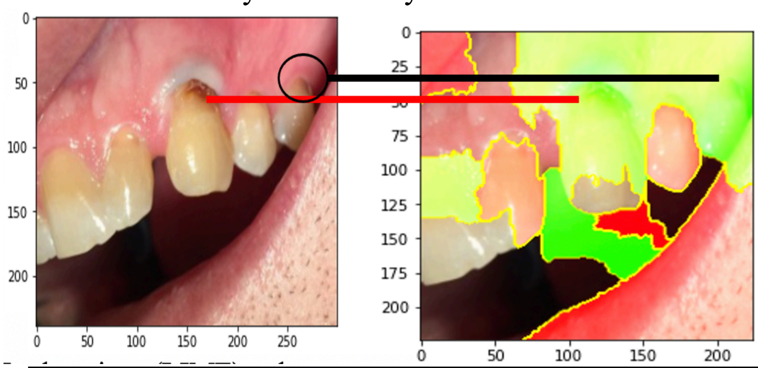

Figure 15. The ResNet model correctly classified this image as having cavity presence. The green areas highlight the image regions in which the model predicted cavity presence. The red areas highlight the image regions in which the model predicted no cavity presence. (Figure by N.Bhattacharjee)

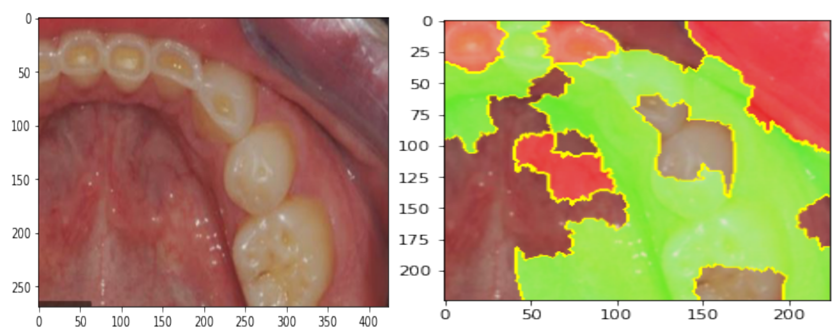

Figure 16. The ResNet model correctly classified this image as having no cavity presence. The green areas highlight the image regions in which the model predicted no cavity presence. The red areas point to regions of the image that influenced the AI model to think that there *may* be cavities (i.e. due to darker shading of gum). However, when the model was making its final diagnosis, the red area's influence was outweighed by the green regions. (Figure by N.Bhattacharjee)

Discussion

Improvements on Previous Work (Berdouses et al., 2015). The table below by N.Bhattacharjee summarizes the differences between the AI cavity detection systems developed by the student researcher and Berdouses et al, 2015.

Study	Photographic Color Images	Type of Machine Learning	Explainability	Occlusal Tooth Surface	Labial Tooth Surface	Lingual Tooth Surface	Buccal Tooth Surface	Teeth Present In Image
Student Researcher's Work	Yes	Deep Learning CNNs	Yes	Yes	Yes	Yes	Yes	Multiple
Berdouses et al., 2015	Yes	Random Forest	No	Yes	No	No	No	One

LIME's Potential. The ability to explain a diagnosis to a patient in an understandable manner is a crucial skill employed by dentists. By using LIME, the study's cavity detection system has also acquired this communication skill. To the researcher's knowledge, there are no published studies that explain their classifier's cavity diagnoses for photos. Using LIME, this study has closed the explainability gap previously present for AI cavity diagnoses.

Limitations of LIME. LIME was developed in 2016, and there are extensions of LIME that achieve similar results (i.e. SHapley Additive exPlanations (SHAP)). One limitation of these explainer algorithms is that they produce explanations at the pixel level. They don't take into account semantic features (ex. texture). Another limitation is that these algorithms are applied on a trained model in a post-hoc manner. In the future, the student researcher aims to test a machine learning technique that considers explainablility even during training (ex. Retain algorithm).

Future Research

Semi-Supervised Learning Techniques. To enhance ResNet-27's performance in the future, additional data needs to be obtained. However, an increase in image quantity will also makes the data annotation process for dentists more time consuming and tedious. As a result, dentists may become hesitant in providing the labels. In the future, it's aimed that semi-supervised learning techniques can be used to decrease the reliance on dentists for labels. With this technique, only a portion of the entire training dataset is labelled by an expert. As ResNet-27 is already trained on labelled data, future images obtained won't need to be labelled. Instead, they simply will receive a pseudo label based on ResNet-27's predictions. Once all unlabeled data receives a pseudo label, the labelled data and pseudo-labeled data can be combined into one training dataset. Using this technique, ResNet-27 will have the opportunity to train on a vastly larger dataset.

Data. ResNet-27's performance can be improved with additional data. The student researcher aims to use a data augmentation technique called Generative Adversarial Networks (GANs) to generate synthetic data (Figure 17).

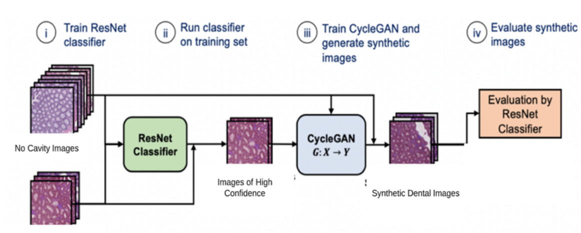

Figure 17. Methodology used to generate synthetic dental images with GANs (Image originally made by Wei et al., 2019)

Application

Figure 18. Example of an intra oral camera. (Figure by N.Bhattacharjee)

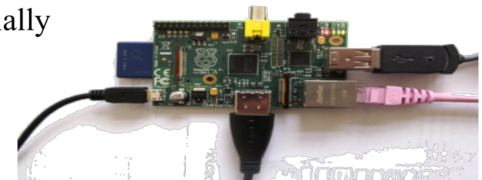

Figure 19. Example of a raspberry pi. It's essentially a mini PC with a credit card form factor (Figure by Clive Darra).

The current results on a ResNet-27 with an accuracy of 77.8% indicate that an AI diagnostic system with ANN architecture used in practical setting can be created. As suggested by a practicing dentist, images of patients' teeth can be acquired using an intra oral camera ($30) (Figure 18). Classification using DL is achievable using a computing device like a $35 Raspberry Pi (Figure 19). This affordable setup could potentially be used as first level screening for dental triage in developing nations where there is a limited dentist availability. In developed countries, the setup could potentially detect cavities at home for people who are dentaphobic or can't afford dental checkups.

Conclusion

Dental cavities, one of the most prevalent chronic disorders, impacts over 3.9 billion people worldwide. Typically, cavity detection requires the services of a trained dentist. However, barriers such as dentophobia, limited dentist availability, and lack of dental insurance prevent millions from receiving dental care. To address these issues, the student researcher created an AI diagnostic system that detects cavity presence and visually explains the rationale behind each diagnosis. The accessible system detects cavities of all levels of decay using an artificial neural network (ANN). With a single photographic color image, the system can provide a cavity diagnosis. Previously, there was a lack of studies focusing on whether AI can be used to detect early to advanced tooth decay on different surfaces (occlusal, lingual, buccal, and labial tooth surfaces) on photographic images. This study aimed to address this research gap. After collecting 314 de-identified photos from online sources, the student researcher experimented with several neural network architectures and training techniques. Using transfer learning from an ImageNet1k dataset, the ResNet-27 architecture proved to be most optimal for cavity detection after earning an accuracy of 77.8% and sensitivity score of .69. Visual explanations for the system's cavity diagnoses were generated using LIME. After applying LIME, the system now has the ability to detect cavity presence and explain its cavity diagnosis to the end-user in an understandable manner. This explainability feature was not present in previous work. By gaining two crucial skills typically employed by dentists, this study's AI diagnostic system can now provide reliable cavity diagnoses to demographics that have constantly been unaccounted for in the past.

References

Akselrod-Ballin, A., Chorev, M., Shoshan, Y., Spiro, A., Hazan, A., Melamed, R., Barkan, E., Herzel, E., Naor, S., Karavani, E., Koren, G., Goldschmidt, Y., Shalev, V., Rosen-Zvi, M., & Guindy, M. (2019). Predicting Breast Cancer by Applying Deep Learning to Linked Health Records and Mammograms. Radiology, 292(2), 331–342. https://doi.org/10.1148/radiol.2019182622

Ali, R.B., Ejbali, R., & Zaied, M. (2016). Detection and Classification of Dental Caries in X-ray Images Using Deep Neural Networks.

Arteaga, C. (2019, October 22). Interpretable Machine Learning for Image Classification with LIME. Medium. https://towardsdatascience.com/interpretable-machine-learning-for-image-classification-with-lime-ea947e82ca13

Berdouses, E. D., Koutsouri, G. D., Tripoliti, E. E., Matsopoulos, G. K., Oulis, C. J., & Fotiadis, D. I. (2015). A computer-aided automated methodology for the detection and classification of occlusal caries from photographic color images. *Computers in Biology and Medicine*, *62*, 119–135. https://doi.org/10.1016/j.compbiomed.2015.04.016

Chicago Tribune. (2019, May 21). Canary System. Chicagotribune.Com. https://www.chicagotribune.com/business/blue-sky/chi-dentist-office-technology-canary-system-bsi-story.html

Codella, N. C. F., Nguyen, Q.-B., Pankanti, S., Gutman, D. A., Helba, B., Halpern, A. C., & Smith, J. R. (2017). Deep learning ensembles for melanoma recognition in dermoscopy images. IBM Journal of Research and Development, 61(4/5), 5:1-5:15. https://doi.org/10.1147/jrd.2017.2708299

Density of dentistry personnel (total number per 1000 population, latest available year). (2018, December 28). Retrieved from https://www.who.int/gho/health_workforce/dentistry density/en/.

D. (n.d.). *Dental Charting Symbols Tooth Chart Definition Of Tooth Chart intended for Dental Charting Symbols 24166 | Dental charting, Tooth chart, Dental assistant study*. Pinterest. https://www.pinterest.com/pin/356558495497028746/

Detective, T. D. (2020, January 31). The 80/20 Split Intuition and an Alternative Split Method. Medium. https://towardsdatascience.com/finally-why-we-use-an-80-20-split-for-training-and-test-data-plus-an-alternative-method-oh-yes-edc77e96295d

Grossfeld, B. (2020, October 12). Deep learning vs machine learning: a simple way to understand the difference. Zendesk. https://www.zendesk.com/blog/machine-learning-and-deep-learning/

Gulshan, V., Peng, L. P., & Coram, M. C. (2016, December 13). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal. JAMA Network. https://jamanetwork.com/journals/jama/fullarticle/2588763

Gupta, N. G., & Vujicic, M. V. (2019, November). Main Barriers to Getting Needed Dental Care All Relate to Affordability. http://www.ada.org/~/media/ADA/Science%20and%20Research/HPI/Files/HPIBrief_0419_1.pdf

How Much Do Dental X-Rays Cost Without Insurance? (n.d.). 1Dental. https://www.1dental.com/dental-xrays/

King, B. A., Dube, S. R., & Tynan, M. A. (2012). Current tobacco use among adults in the United States: findings from the National Adult Tobacco Survey. American journal of public health, 102(11), e93–e100. https://doi.org/10.2105/AJPH.2012.30100

Kliff, S. (2014, May 4). *One in three Americans didn't see the dentist last year*. Vox. https://www.vox.com/2014/5/4/5675878/one-in-three-americans-didnt-see-the-dentist-last-year

Lee, Jae-Hong, et al. "Detection and Diagnosis of Dental Caries Using a Deep Learning-Based Convolutional Neural Network Algorithm." Journal of Dentistry, Elsevier, 26 July 2018, www.sciencedirect.com/science/article/pii/S0300571218302252

Marcenes, W., Kassebaum, N. J., Bernabé, E., Flaxman, A., Naghavi, M., Lopez, A., & Murray, C. J. L. (2013, July). Global burden of oral conditions in 1990-2010: a systematic analysis. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4484374/.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?" Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1–10. https://doi.org/10.1145/2939672.2939778

Sahoo, S. (2018, November 29). Residual blocks — Building blocks of ResNet - Towards Data Science. Medium. https://towardsdatascience.com/residual-blocks-building-blocks-of-resnet-fd90ca15d6ec

Silver, L. (2019, December 31). Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. Retrieved from https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/.

Silvertown, J. D., Wong, B. P. Y., Abrams, S. H., Sivagurunathan, K. S., Mathews, S. M., & Amaechi, B. T. (2016). Comparison of The Canary System and DIAGNOdent for the in vitro detection of caries under opaque dental sealants. Journal of Investigative and Clinical Dentistry, 8(4), e12239. https://onlinelibrary.wiley.com/doi/abs/10.1111/jicd.12239

Statista. (2020, July 27). Global per capita tea consumption as of 2016, by country. https://www.statista.com/statistics/507950/global-per-capita-tea-consumption-by-country/

Tooth Decay (Caries or Cavities) in Children. (n.d.). Retrieved from https://www.hopkinsmedicine.org/health/conditions-and-diseases/tooth-decay-caries-or-cavities-in-children.

Toppa, S. (2015, March 5). Over 2.4 Billion People Suffer From Untreated Tooth Decay. Retrieved from https://time.com/3732925/one-third-world-tooth-decay/.

Wang, C. (2019, January 08). The Vanishing Gradient Problem. Retrieved November 12, 2020, from https://towardsdatascience.com/the-vanishing-gradient-problem-69bf08b15484

Warden, P. (2014, June 9). Deep learning on the Raspberry Pi! Pete Warden's Blog. https://petewarden.com/2014/06/09/deep-learning-on-the-raspberry-pi/

Wei, J. W., Suriawinata, A. S., Vaickus, L. V., Ren, B. R., Liu, X. L., Wei, J. W., & Hassanpour, S. H. (2019, October). Generative Image Translation for Data Augmentation in Colorectal Histopathology Images. https://arxiv.org/pdf/1910.05827.pdf