PROJECT
TITLE: Automatic dental Diagnosis using panoramic X-ray

File Name: COMS4995.2019Spring.GGOP.ac4391.tz2247.ek2993
Students: Ariel Cohen Codar ac4391, Tomer Zwi tz2247, Suhyun Kim ek2993
Columbia University

Abstract

This project introduces image segmentation and classification techniques to detect anomalies in panoramic dental X-Rays. The automation of dentists' diagnosis can reduce the time management of treatment and help to prevent misdiagnosis. The two different approaches are trained on a private dataset provided by the CHU Hospital in Paris after anonymization of the data. Classification achieved the accuracy of 98% and the image segmentation with dice loss of 0.9.

1. Introduction

Previous methods for automatic dental X-Ray diagnosis have largely focused on classification on bitewings using deepwise convolutional neural network[1][2]. Such attempts have resulted in encouraging outcome for the future of the dentistry. In the research, Detection of Tooth caries in bitewing Radiographs using deep learning, they develop computer aided (CAD) diagnosis to help dentists to found cavities. A deep fully convolutional neural network (FCNN) consisting 100+ layers, is trained to mark caries on bitewing radiographs. They found that having CAD help to improve performance and reliability and get a F1-score of 70% with CAD largely above average dentists labeling alone (50-56%) [1].

However, bitewings are prescribed only when a first consultation with a panoramic X-ray is done. Avoiding a large exposition to radiation for CAD tools is the next step in for dentistry. All the more, misdiagnosis mostly consists on ignoring potential diseases before prescribing bitewings. Research papers using panoramic X-rays focus on tooth detection rather that automatic diagnosis [3].

Our goal in this research is to provide a baseline for automatically detect potential diseases on panoramic X-Rays. Two approaches will be evaluated : classification using deepwise convolution and image segmentation using U-Net Architecture.

2. Cavity Detection

We are trying to create, for the first time, an ANN model to detect cavities in a panoramic X-rays. We use and tuned several existed ANN such as UNET and Attention UNET and try to compare between them and see which one we got a better result. To reduce high dimension and complexity of the images we decided to resize our images

2.1. Objective and Technical Challenges

The objective of this report is to show pioneer results of detecting automatic cavities on panoramic X-rays using varieties of neural network technologies such as UNET and Attention UNET and machine learning models such as classification.

The first major challenge for this project is being the first one doing this research. There were no previous research that has done before on automatic detecting oral and maxillofacial on panoramic X-Rays which was hard for us to get initial direction such as which neural network to use or machine learning models. This lead us to tried different models (such as Unet and attention Unet) for the first time working on this kind of data set . Second technical challenge is the size of our dataset which is very small. Our dataset we got from collaboration with UPEC hospital in Paris, France, and several dental clinics - we have roughly of 1000? panoramic X-rays which is not a lot for training and testing. Additionally, there is a computational challenge present in many neural network project to run a model. In this case each image has high dimensional (size of of each image) which cause to high computing. Evan with cloud computing instance (GPU and TPU) it is still run slow and sometime crushed due to lack of memory. Therefore we need to take off between features we used and the memory constraints of the computing machine. Moreover, another challenge is the imbalanced labels, because the cavity is very small part of the image we face often imbalanced labels the size of each cavity is very small compare to the size of the image which make it difficult to our Neural network to detect cavity.

Finally, the main goal of this project is to test performance of using image segmentation and classification using different neural network such as UNET and Attention Unet. A key concept is creating models that can be easily comparable results so we can choose the best one.

3. Experiment design

One strong characteristic of the dataset is that each image is a high dimension image. It includes a lot of information that some of it is not relevant to detecting cavity. Also, each image is very imbalanced and we faced a challenge to keep most of the information. The size of the cavity is very small compared to the size of the original image so we tried to extract resize the image to get the most information. We thought to use pre-trained model such as MobileNet and Vgg19 but because the size of the images we decided to approach in a different way. We approaches our problems by looking and comparing between two models - classification models and image segmentation model.

For image segmentation; we first take our image and doing pre processing - because the image is very large we crop the image to take the most important data (looking on panoramic X-rays which is around the teeth is not important for detecting cavities). First we crop the original image to small size which includes the most important data of the panoramic X-rays and mask which the cavity part is white pixels while all other image is black pixel. After we tried to run two main models (UNET and Attention UNET). Then post processing - comparing between the mask and the original and color in green the mask on the original image).

For classification the goal is getting an image and return either there is cavity or not. We crop the image and the mask classifying between positive and negative image and mask. Then we run through the negative image and label 0 and positive image label 1. After labeling we run fully convolutional networks model t on the mask/original images.

We will compares between the models and check which model will lead to a better result.

4. Implementation

In this section, we evaluate the proposed model with extensive qualitative and quantitative experiments.

4.0 Data Collection

(1) Collecting unbiased dataset of X-Ray panoramic

As we said above, we collected a private dataset of hundreds of panoramic X-Ray with the collaboration with the dentist Phoebe Kamioner and the agreement of the CHU Hospital in Paris. We were able to write a script used in Paris to anonymized the dataset by removing the name of the patient on each x-ray.

(2) Merging the ground truth of dental experts

All the X-rays have been manually labels and review by two dentists using the software labelbox. The objective was to identify problematic regions on the X-Ray and propose, if necessary, a complementary exam such as a local X-Ray. For the purpose of this study, we will consider the problem as cavities even if a semantic segmentation could precised the different specific diagnosis.

4.1. Data preparation

(1) Image enhancement techniques (ex. CLAHE)

X-Ray images are grayscale images with a very poor contrast enhancement and a high variance of gray level. To prevent the model to learn bad featured, we used filters to contrast the images and normalized the grayscale distribution. We found that, keeping the visibility of the tooth and the gradient of gray distribution around them is not easy. Simple histogram equalization was not good and reduced the visibility of the teeth. We decided to use a CLAHE filter (Contrast Limited Adaptive Histogram Equalization) that enhance the local contrasts on images.

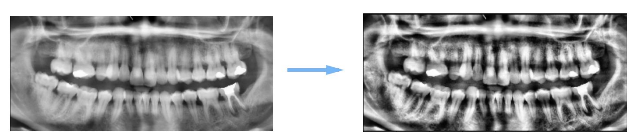


Figure 1. example of Clahe filtering

(2) Data augmentation (Flip, Rotation, Elastic deformation, Black patches and Contrast adjustments

(3) Resizing techniques (Resize, Grid, Crop)

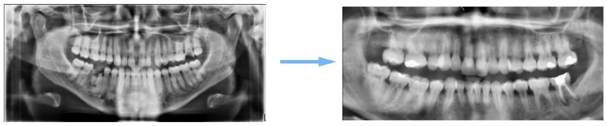


Figure 2. example of center cropped image

4.2. Deep Learning Network

We approached in two different ways: image classification and image segmentation. We have two different models for each method detailed below.

(1) Classification based : Transfert learning and multi-paths techniques

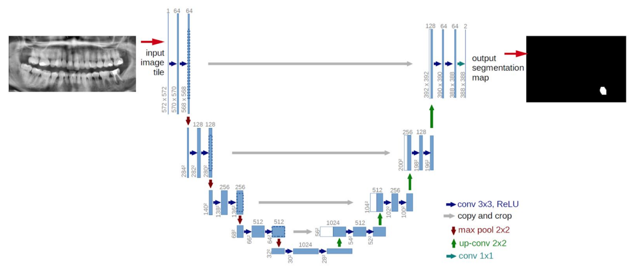
We have implemented a deep neural network using convolutional neural network for the classification task. The following figure describes the DNN of our model.

Layer (type)	Output Shape	Param #
ImageInput (InputLayer)	(None, 224, 224, 3)	0
Conv1_1 (Conv2D)	(None, 224, 224, 64)	1792
Conv1_2 (Conv2D)	(None, 224, 224, 64)	36928
pool1 (MaxPooling2D)	(None, 112, 112, 64)	0
Conv2_1 (SeparableConv2D)	(None, 112, 112, 128)	8896
Conv2_2 (SeparableConv2D)	(None, 112, 112, 128)	17664
pool2 (MaxPooling2D)	(None, 56, 56, 128)	0
Conv3_1 (SeparableConv2D)	(None, 56, 56, 256)	34176
bn1 (BatchNormalization)	(None, 56, 56, 256)	1024
Conv3_2 (SeparableConv2D)	(None, 56, 56, 256)	68096
bn2 (BatchNormalization)	(None, 56, 56, 256)	1024
Conv3_3 (SeparableConv2D)	(None, 56, 56, 256)	68096
pool3 (MaxPooling2D)	(None, 28, 28, 256)	0
Conv4_1 (SeparableConv2D)	(None, 28, 28, 512)	133888
bn3 (BatchNormalization)	(None, 28, 28, 512)	2048
Conv4_2 (SeparableConv2D)	(None, 28, 28, 512)	267264
bn4 (BatchNormalization)	(None, 28, 28, 512)	2048
Conv4_3 (SeparableConv2D)	(None, 28, 28, 512)	267264
pool4 (MaxPooling2D)	(None, 14, 14, 512)	0
flatten (Flatten)	(None, 100352)	0
fc1 (Dense)	(None, 1024)	102761472
dropout1 (Dropout)	(None, 1024)	0
fc2 (Dense)	(None, 512)	524800
dropout2 (Dropout)	(None, 512)	0
fc3 (Dense)	(None, 2)	1026

Total params: 104,197,506
Trainable params: 104,194,434
Non-trainable params: 3,072

(2) Image segmentation: U-Net [2] & Attention U-Net [3]

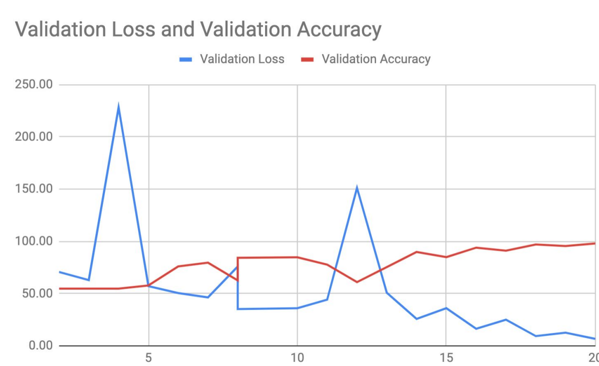
We have implemented a unet model with attention gate and without attention gate. All models are trained using the Adam optimiser [9], batch-normalisation, deep-supervision [10], and standard data-augmentation techniques (affine transformations, axial flips, random crops).



5. Results

5.1. Project Results And Discussion

We have run the classification experiment on 1792 cavity-positive images and 2167 cavity-negative images with 20 epochs. Validation was done on 527 cavity-positive images and 1982 cavity-negative images. and the validation loss and accuracy are presented in the graph below. Validation accuracy started from approximately 50% initially and by the end of the training, we have achieved the accuracy close to 100%. The Validation loss started with about 70 and it decreased to 20 by the end of training. We identified the crossing point where the validation accuracy and validation loss met is the epoch of 13, which has the validation loss of 75 and the accuracy of 89%.



For classification, we see that the accuracy is 98% at the end of the 20th epoch while the loss is 7. We assume that the model started overfitting to our dataset starting from

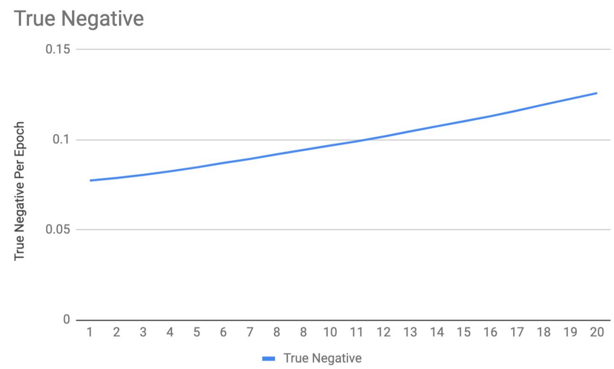
13th epoch when the validation accuracy curve and the validation loss curve meets.

For image segmentation, we have implemented a state-of-the-art U-Net model. We used multiple metrics including true negative, true positive and dice loss.

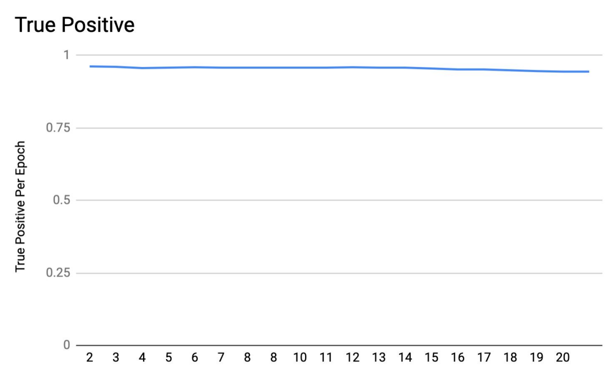
Table 1. Result of Image Segmentation with Unet

Epoch	True Negative	True Positive	Dice Loss
1	7.74E-02	9.62E-01	9.89E-01
2	7.88E-02	9.61E-01	9.89E-01
3	8.05E-02	9.56E-01	9.89E-01
4	8.25E-02	9.58E-01	9.89E-01
5	8.47E-02	9.59E-01	9.89E-01
6	8.71E-02	9.58E-01	9.89E-01
7	8.94E-02	9.58E-01	9.89E-01
8	9.19E-02	9.58E-01	9.89E-01
9	9.44E-02	9.58E-01	9.89E-01
10	9.68E-02	9.58E-01	9.89E-01
11	9.91E-02	9.59E-01	9.89E-01
12	1.02E-01	9.58E-01	9.89E-01
13	1.05E-01	9.58E-01	9.89E-01
14	1.07E-01	9.55E-01	9.89E-01
15	1.10E-01	9.52E-01	9.89E-01
16	1.13E-01	9.52E-01	9.89E-01
17	1.16E-01	9.49E-01	9.89E-01
18	1.20E-01	9.46E-01	9.89E-01
19	1.23E-01	9.44E-01	9.89E-01
20	1.26E-01	9.44E-01	9.89E-01

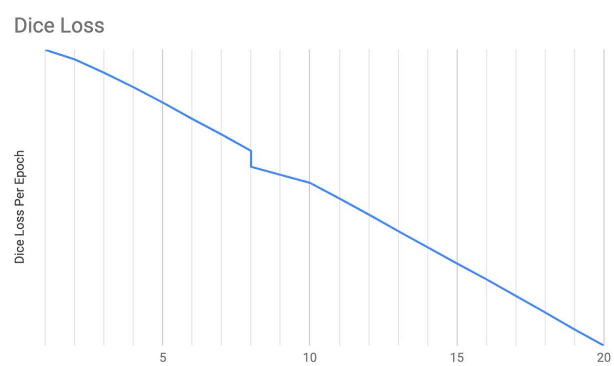
As shown in the graph below, the true negative values are increasing, which means that the negative values are actually predicted to be negative as the number of epoch grows. It can be translated that the model learns the areas of no cavities in an effective manner with increasing number of epochs.



However, true positive values did not change much, which indicates that the model has a hard time finding the areas with cavities. We believe the reason is because of the resizing we have done with image segmentation, which made the mask size even smaller. Therefore, the model does not have enough positive learning data to learn from, which makes it challenging to learn the location of cavities.



The following graph shows dice loss in a logarithmic scale. The decreasing values of dice loss shows that as the number of epoch increases, the model is able to learn how to read x-ray images.



6. Conclusion and Future Work

For the classification task, we have successfully classified images with cavities and without cavities with high accuracy. The overfitting issue can be resolved with more data. For the image segmentation task, our model learns the location without cavities as the number of epoch increases, but it is challenging to find the location with cavities. In the future, we plan to input data without resizing and also implement the U-Net model with attention gates in order to look for specific areas to detect cavities.

7. Notes

8. References

[1] Hak Gu Kim, and Yong Man Ro, “Modality-bridge Transfer Learning for Medical Image Classification”, 2017
[2] Olaf Ronneberger et al., “U-net: Convolutional networks for biomedical image segmentation,” in MICCAI, 2015.
[3] Ozan Oktay, “Attention U-Net: Learning Where to Look for the Pancreas”, 2018
[4] Md Zahangir Alom et al., “Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation,” 2018.