A REPORT ON

Research and Development Project

Author:
Kumar Ashutosh
16D070043

Supervisor:
Prof. Amit Sethi

Department of Electrical Engineering
Indian Institute of Technology Bombay

May, 2018

Chapter 1

Discriminative Localization in Medical Images

1.1 Introduction

Identifying deep features in Medical Images is a challenge which is being researched upon lately. Zhou *et al* [1] develops an algorithm to visualize deep features and to get a heatmap of the activating regions. The work in [2] uses this feature for detecting pneumonia in Chest X-Rays. Our attempt it to make a generic model for easing the visualization task. Visualization is achieved without a significant change in the architecture of the CNN. This can be used to compare two models apart from the usual accuracy parameter that we consider.

1.2 Related Work

The work in [1] discussed the idea of **Class Activation Mapping**, which combined with **Global Average Pooling** produces accurate heatmaps corresponding to the desired class. The same idea is applied to various CNN architectures and the results are analyzed.
In a Convolutional Neural Network (CNN) architecture, the last dense layers are removed and replaced with a Global Average Pooling Layer which takes the average value of a layer and uses it as a feature. The feature matrices of the last layer are up-sampled and their weighted sum is the desired Class Activation Map. The weights are learned by fine-tuning the model or trained all over again (No significant difference was observed in these two cases).

1.3 Dataset

The dataset used for the work was a part of ICIAR 2018 Grand Challenge on BreAst Cancer Histology. The dataset contains 400 images each of 2048 x 1536 pixels. The dataset is divided into four classes - *benign*, *malignant*, invasive and *normal*.
For the purpose of training and validation, the data was further divided into small patches of 256 x 256 pixels. The data was cleaned for anomalies and unimportant patches were removed from training process.

1.4 Initial Training and Further Modification

The data was initially trained on *Inception V3* architecture and a validation score of 0.84 was achieved. The results are satisfactory considering the complexity of the images.
For the purpose of this research, a modification is done to the architecture and is fine-tuned. Training the data all over again also produced a similar result. The dense layer is replaced with a **Global Average Pooling** layer and finally a **Softmax Layer** is added to obtain the probability of each class.

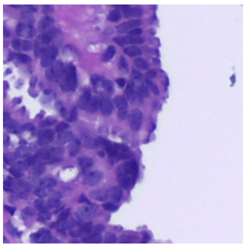
1.5 Fine Tuning

Upon Fine-tuning the modified architecture, a validation accuracy of 0.81 is observed. This compromise in the accuracy is due to the removal of the fully connected layers. This trade-off is discussed in [1] as well. Our purpose of better visualization is achieved in the configuration.

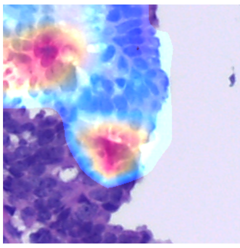
1.6 Results

Due to the removal of Fully Connected layers, the architecture is now size independent, i.e., we can use the architecture to predict an input of any size. We tested the output with both patch-wise input and whole-image input. Some interesting observations are listed along with the obtained result.

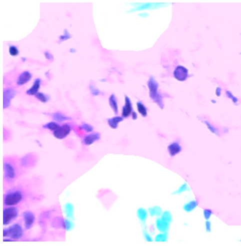
1.6.1 Patch-wise Analysis



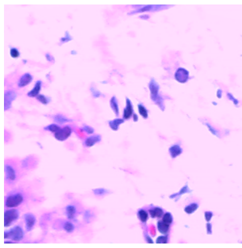
Original Class : **Benign**



CAM of : **Benign**



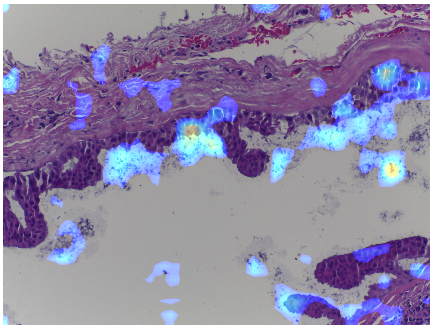
Original Class : **In-Situ**



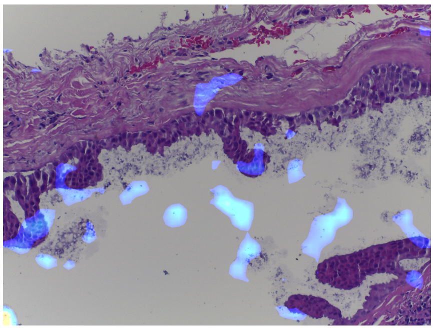
Original Class : **Benign**

The above visualizations are as per our intuition. In the first case, a CAM of Benign class on a Benign image yields a hotspot in the image. In contrast, a Benign CAM on an image from In-Situ class shows less excitation. It is important to note that the CAM values are NOT scaled with their respective class probabilities. This is done to obtain proper results worthy of comparison.

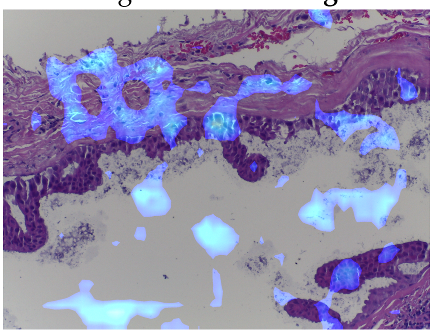
1.6.2 Whole-image Analysis



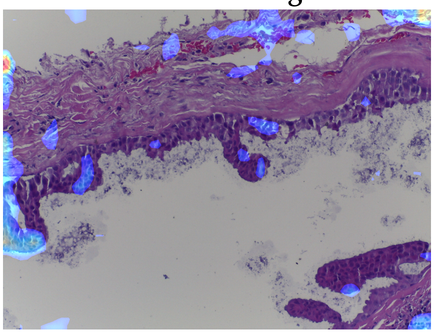
Original Class : **Benign**



CAM of : **Benign**



Original Class : **In-Situ**



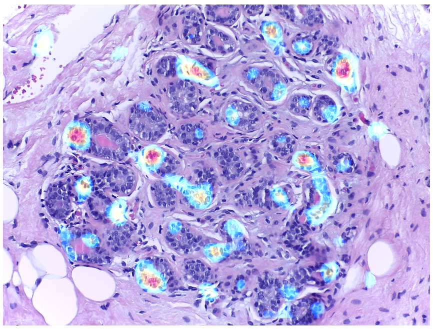
Original Class : **Benign**

This image also follows the previous trend of exciting on the correct class. In addition to this, we can observe some medium activation in case of Invasive as well. Hence we can infer that the above sample belongs to the In-Situ class with an advancement towards Invasive.

1.7 Inferences

This work produces some very interesting results based on generated Heatmaps. CAM of other classes is also informative in a sense that it gives us the information of which regions resembles the learned activation feature of an image. The major inferences that can be nicely drawn from the class activation maps are listed below.

1.7.1 Early Detection of Cancer



Normal Image with Benign CAM

The above image is classified as a Normal Image. The activation map is that of Benign. Although as per our trained model, the tissue is healthy. But this image provides an idea of vulnerable regions, thus helping in an early detection. But, it is important to note that vulnerability is also subject to the probability of this tissue being cancerous. Thus, this helps as an aid to detect the regions of possible anomalies.

1.7.2 Deep Features

In case of other datasets, it is easy to guess which part of an image influences a particular class. For example, a dataset with animals, it is easy to observe that the decision will be influenced by body and face features. This is not very trivial in Medical Images. This could help identifying deep features which may go unnoticed while manually checking for vulnerability. Also, in case of abnormal results, this heatmap would be helpful in proving reasonable justification to the resultant class.

1.8 Future Work

The same algorithm can be tested on other Medical Datasets and the performance of the model can be estimated with the help of visualization. A side by side comparison of heatmaps from various models can be used to analyze the correctness of the training. This can be used in addition to the normal accuracy comparison technique.

1.9 References

1. Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, Antonio Torralba: "Learning Deep Features for Discriminative Localization" , 2016.

2. Pranav Rajpurkar, Jeremy Irvin, Andrew NG: "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning", 2017.

Chapter 2

Annotator App

2.1 Introduction

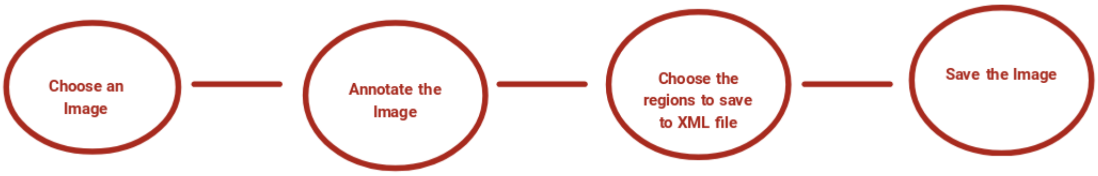
Annotation of Medical Images is a very important task which forms the basis of this whole field of Machine Learning. The task of annotation is given to specialists in their respective field. The aim of this project is to facilitate annotation and generate required XML files for producing masks. Simplifying this task of annotation would facilitate and would help in the faster annotation of datasets.

2.2 Related Software

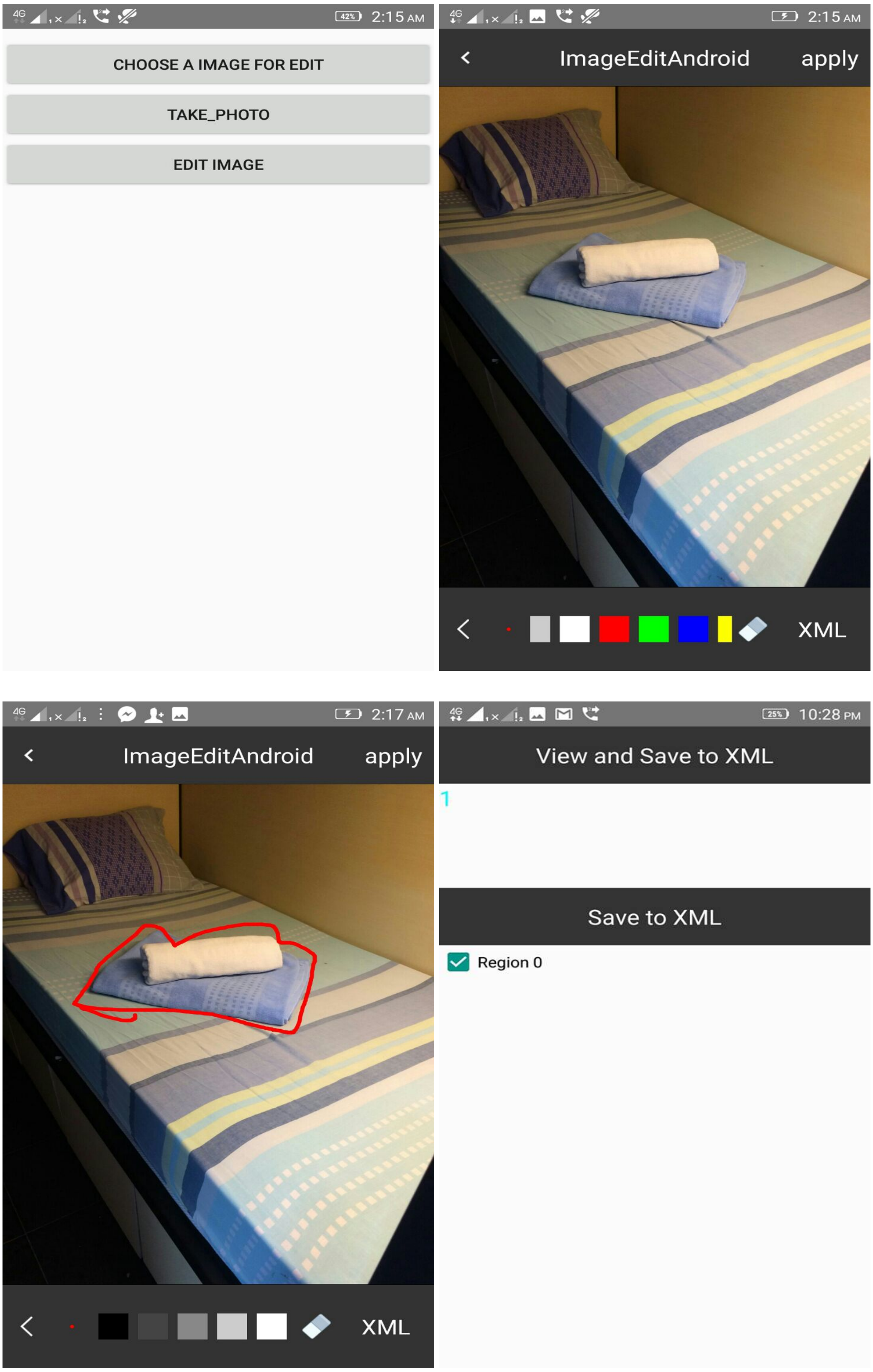
Aperio ImageScope is the currently used application for annotation task. Although the application offers a wide range of functionalities and is easy to use, this software is available only for Windows Operating Systems. With the diminishing demand for Windows OS and huge popularity of Android Applications, our attempt is to create an Android Application which annotates the data with a very simple and easy to use interface.

2.3 Block Diagram of Functionality

This sections represents the functioning of the App with the help of Block Diagrams.



2.4 Application Layout



2.5 Future Work

The following features is going to be added to the App to make user experience better:

- An option to load multiple images simultaneously so as to make the annotation task faster.
- Feature to zoom the image while Annotating.

2.6 References

- Image Editor Android App. https://github.com/siwangqishiq/ImageEditor-Android.