`Breast Cancer Analysis and Evaluation using Artificial Intelligence techniques.

Team : Santhosh Kumar Padige

Under the guidance of Prof. Vahid Behzadan

Introduction:

The most common issue related to health is the breast cancer mostly in female. According to one of the surveys in 2012, because of the breast cancer 522000 deaths occur and increase the death ratio upto 14% which were in 2008.By the introduction of medical imaging it brought a huge revolution in medical field. This helps medical staff to interpret the whole case easily. Besides of medical imaging by the introduction of Artificial intelligence made a huge change in medical side. AI made the process of detection and segmentation of disease more efficient and accurately. Which overcome the process of interpretation and made it easy to interpret it more accurately.

Literature:

In literature there are many methods of Artificial intelligence which is used in medical imaging to interpret the results. Let's have to brief some work related to breast cancer.

Related to breast cancer in the last couple of years AI made a huge improvement in detection of breast cancer. Tourassi present a method of matching template which is a method of detection in machine learning and by using template matching ROI of cancer disease were generated. Alto used a method of feature based in which shape and texture features in order to diagnoses cancer. Tao used a ROI based method and classify the data into its respective classes which are benign and malignant. Zheng proposed a method of breast cancer detection in which spikes of tumor were further analyzed to get better results. Narvaez used the same method of texture which are shape and texture and to get the ROI region which can be used for annotation as well for further training and testing.

Liu proposed a method of image search based in order to detect the breast cancer. In authors used convolutional neural network in order to classify the cancer affected x-rays. In this, used a deep neural network and got the classification accuracy up to 70%. Han proposed a deep learning method and claim classification accuracy 75%.

Nuh proposed a method of deep learning for classification of breast cancer and got accuracy up to 78%.

Objective & Problem statement:

Normally the breast cancer can't be identifying in early stages and because of this the main problem regarding the breast cancer is that its death ratio is very high worldwide. According to the Who every woman out of 6 is patient of breast cancer. There is intense need of some AI method which can bring some revolution change to bring down the rate of death as much as can be.

So, the main objectives of this project are

1. To identify the cancer in early stages by Deep learning.
2. To classify the types of breast cancer.
3. To make it easy for medical stuff to interpret the results easily and quick.
4. To decrease the death rate because of breast cancer to cure them on time.

Dataset:

In this project we used Kaggle dataset. This dataset is publicly available for the research purposes. This dataset consists of two types of cancer one is Benign and other one is malignant. Therefore, we have one folder for benign breast Cancer images and other folder for the purpose of malignant breast cancer images. We have total number of 10000 breast cancer images. But originally the number of images were round about 2000 which is too small for the training of model, therefore we employ a technique of data augmentation and make the total datasets becomes four times and ease for the training of the model because deep learning needs a lot of data for good and efficient result. Some of the data are given below in the following figures for Breast cancer benign and malignant.

1. (b)

Figure: (a) shows benign breast cancer and (b) shows malignant cancer

Methodology:

In this project we are targeting medical side by deep learning, that is to classify the images weather it's belong to malignant or benign breast cancer patient. The whole pipeline will be discuss step by step

1. Input Data

The first thing is data which is our input images consist of images of benign cancer patient and malignant cancer affected patient.

2. Pre-processing

The data were not in proper use for the model to train, we did some pre-processing on these data in order to make them useful for training. Some of the preprocessing methods are

1. Resizing
2. Augmentation
3. Normalization
4. Splitting Of Dataset

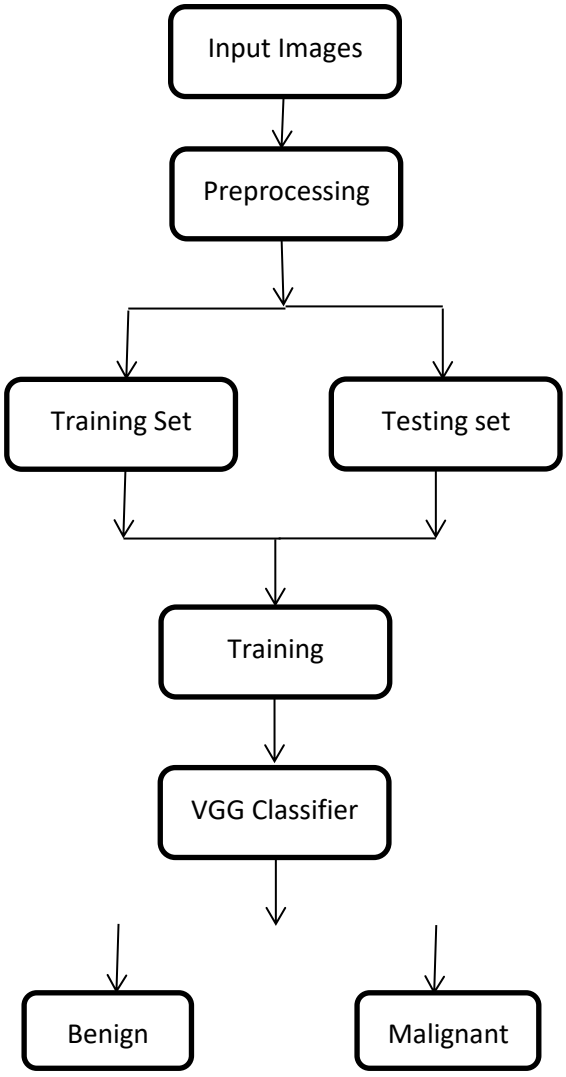
In every machine learning and deep learning splitting of dataset is very important part, in this we split data according to some percentage for training and for testing purpose. We split 80-90% for training and 10-20% data for testing.

5. Vgg Model

After splitting the data into training and testing, then we used a deep learning model which is vgg for training and prediction. Vgg is modified form of convolutional neural network having 3 convolutional layers for feature extraction.

6. Classification

The final and last step is to classify the incoming data that weather it's belonged to benign or it belongs to malignant. For classification normally used fully connected layers or regression classifier used, In this project we used fully connected layers on the bases of those features which is coming from vgg model and to classify them to its respective class. The whole proposed methodology is shown in the following figure.



Results:

As the model is trained then we used the test data in order to check our model performance. To check any model performance training and validation loss and accuracy is very important. That can be seen in the following figure.

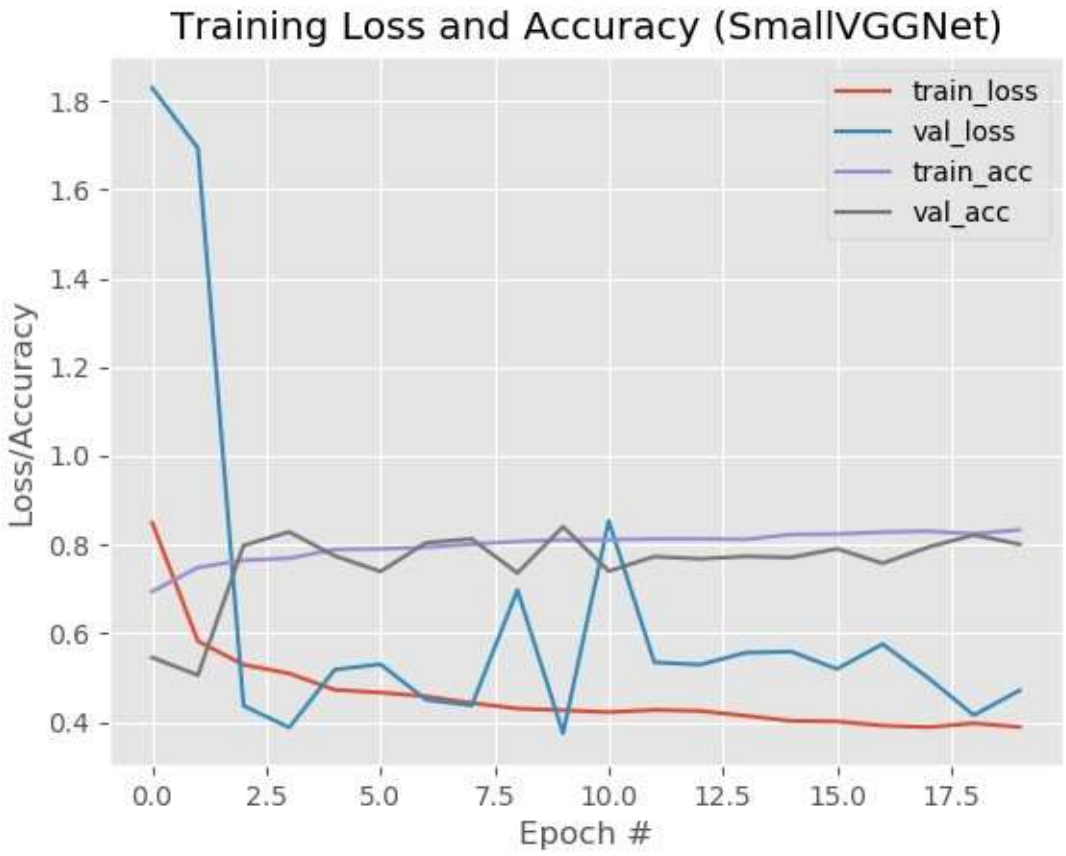


Figure: Loss and Accuracy graph

And the accuracy for both classes is 82%. After the training the model we used some UN train data in order to check our trained model prediction for images. And the results were brilliant which can be seen in the following figures.

1. (b)

Figure: (a) shows output of malignant cancer (b) shows benign cancer

Conclusion:

As form the proposed methodology it can be seen and concluded that our method of classification of breast cancer is much better than the methods which are used in literature. In literature the maximum accuracy was upto 78% and the accuracy of our proposed methodology is upto 83% which can be seen in graph. In future we are targeting to get more accuracy upto 90% by taking much more real and local data of patient and hospitals. We also targeting to make a product in which we will just give an input image to the product and that will generate the output on the screen that weather this malignant or benign.

References

[1] G. D. Tourassi, R. Vargas-Voracek, D.M. Catarious Jr and C. E. Floyd, Computer-assisted detection of mammographic masses: A template matching scheme based on mutual information. Medical Physics, Vol 30, No.8, pp: 2123–2130, 2003.

[2]. H. Alto, R. M. Rangayyan and J. L. Desautels, Content-based retrieval and analysis of mammographic masses. Journal of Electronic Imaging, Vol 14, No.2,:023016– 023016, 2005.

[3]. Y. Tao, S. Lo, M.T. Freedman and J. Xuan, A preliminary study of contentbased mammographic masses retrieval. In Medical Imaging, 65141Z. International Society for Optics and Photonics, 2007.

[4]. B. Zheng, A. Lu, L. A. Hardesty, J. H. Sumkin, C. M. Hakim, M. A. Ganott and D. A. Gur, A method to improve visual similarity of breast masses for an interactive computer-aided diagnosis environment. Medical Physics, Vol 33, No.1, pp: 111–117, 2006.

[5]. F. Narvaez, G. Diaz and E. Romero, Multi-view information fusion for automatic bi-rads description of mammographic masses. In SPIE Medical Imaging, 79630A. International Society for Optics and Photonics, 2011.

[6]. J. Liu, S. Zhang, W. Liu, X. Zhang and D. Metaxas, Scalable mammogram retrieval using anchor graph hashing. In 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), pp; 898–901, 2014.

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

[8]. F.A. Spanhol, L.S. Oliveira, C. Petitjean, C. and L. Heutte, “Breast cancer histopathological image classification using convolutional neural networks”. In International Joint Conference on Neural Networks (2016), p. 2560-2567.