



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 6, June 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.512

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Highly Accurate Penta Cancer Detection System Using Deep Learning

M.Pawadhareeni, M.Subhiksha, S.Tamilselvi, A.Velu

U.G. Student, Department of Electronics and Communication Engineering, Meenakshi Sundararajan Engineering College, Chennai, Tamil Nadu, India

U.G. Student, Department of Electronics and Communication Engineering, Meenakshi Sundararajan Engineering College, Chennai, Tamil Nadu, India

U.G. Student, Department of Electronics and Communication Engineering, Meenakshi Sundararajan Engineering College, Chennai, Tamil Nadu, India

Assistant Professor, Department of Electronics and Communication Engineering, Meenakshi Sundararajan Engineering College, Chennai, Tamil Nadu, India

ABSTRACT: The project is about prediction of five types of cancer namely Skin cancer, Lung cancer, Brain tumor, Cervical cancer and Thyroid cancer using deep learning algorithms. The objective of this project is to classify between the different types of cancer and detect its presence without the need of a number of consultations from different doctors. There are four main steps data sets collection, data augmentation, training the model and web application development. Here the data sets are collected and augmentation is done to improve the number of data sets. These data sets are given as input to training model. The model is trained using various deep learning algorithms like VGG16 and Resnet50. Then the model is tested using different data sets. Finally a web application is developed using React JS where an image is given as input and the type of cancer is displayed as output. We have carried out the analysis using different cancer and non cancer images. We have been instrumental in achieving good accuracy in output as compared to other existing systems.

KEYWORDS: Cancer Prediction, Web application, Datasets, Efficiency, Convolutional Neural Network

I. INTRODUCTION

Cancer is the leading cause of deaths worldwide. According to the American cancer society, 96,480 deaths are expected due to skin cancer, 142,670 from lung cancer, 42,260 from breast cancer, 31,620 from prostate cancer, and 17,760 deaths from brain cancer in 2019 (American Cancer Society, new cancer release report 2019). Both researchers and doctors are facing the challenges of fighting cancer. Early detection of cancer is the top priority for saving the lives of many. Typically, visual examination and manual techniques are used for these types of a cancer diagnosis. This manual interpretation of medical images demands high time consumption and is highly prone to mistakes and results in the risk of losing people's life. The existing system, a highly automatic and objective model named Robust Collaborative Clustering methods and Radiomic Features for Cancer detection uses matrix tri-factorization enhanced by adaptive sparsity regularization for simultaneous feature dimensionality reduction and denoising. But it is focused only on single type of cancer and the accuracy is comparatively low. The proposed system uses VGG16 algorithm for classification of images that produces very high accuracy. The novelty of the system is its ability to identify five types of cancer using single application.

II. METHODOLOGY

In this project, we are going to determine presence of lung cancer, skin cancer, brain tumor, cervical cancer or thyroid cancer. By this project, we can able to determine the presence of cancer in the CT images or x-rays. So, initially the first step will be dataset collection where we will be collecting dataset such as CT images or x-rays which are used by the laborites to analysis the presence of cancer from various resources through internet.



After that, we will be splitting those datasets into different categories that is we will be splitting the dataset into training and testing dataset. In training datasets, we will be using the dataset for training the module whereas testing dataset is used to evaluate the model when it is been completely ready. So training dataset first undergoes the process called dataset augmentation, where the dataset is multiplied into many datasets then it will undergo the process called preprocessing, which is to make all sizes into single size. We train that datasets by extracting the features using VGG16 algorithms.

It undergoes a process called optimization which will optimize the model and loss minimization which will reduce the noises generated during training. In the last it will be undergoing a process is called model seriation which will be evaluated after generating model using the testing dataset and predict the presence of cancer. A web application using a JavaScript framework reactJS will be developed where an input x-ray image will be given to predict which type of cancer. Thus, this method provides an effective and cheap method to determine the presence of lung cancer, skin cancer or brain tumor than the methodologies used nowadays.

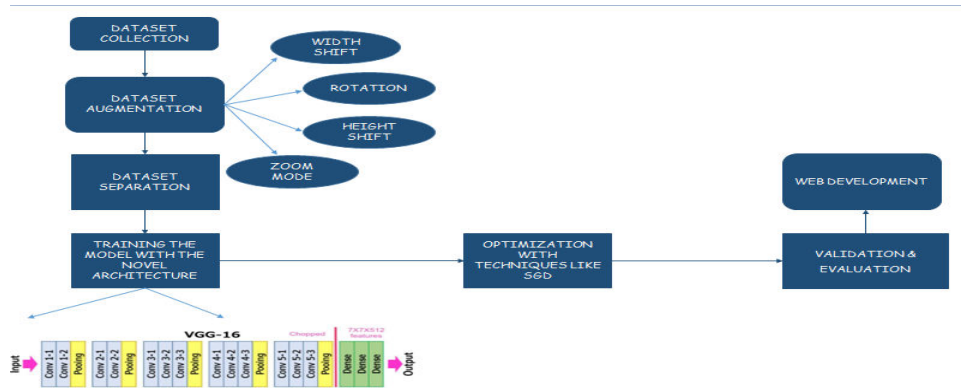


Figure 1: Block Diagram

DATA COLLECTION

A data set is a collection of data. We collected data from Kaggle website. The total data sets get divided as 75% of data for training and 25% of data for testing. The training data set is the one used to train an algorithm to understand how to apply concepts such as neural networks, to learn and produce results. It includes both input data and the expected output. The test data set is used to evaluate how well your algorithm was trained with the training data set.

DATA AUGMENTATION

Data augmentation is a technique to artificially create new training data from existing training data. This is done by applying domain-specific techniques to examples from the training data that create new and different training examples. Transforms include a range of operations from the field of image manipulation, such as shifts, flips, zooms, and much more. Image data augmentation is typically only applied to the training dataset, and not to the validation or test dataset. This is different from data preparation such as image resizing and pixel scaling; they must be performed consistently across all datasets that interact with the model.

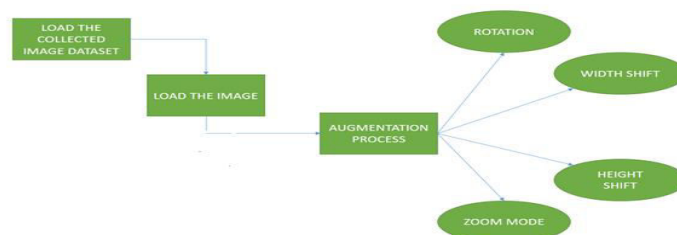


Figure 2: Augmentation Process



TRAINING THE MODEL

In this project, the VGG-16 algorithm is used for training the model. The input to conv1 layer is of fixed size 224 x 224 RGB image. The image is passed through a stack of convolutional (conv.) layers, where the filters were used with a very small receptive field: 3x3 (which is the smallest size to capture the notion of left/right, up/down, center). In one of the configurations, it also utilizes 1x1 convolution filters, which can be seen as a linear transformation of the input channels (followed by non-linearity). The convolution stride is fixed to 1 pixel; the spatial padding of conv. layer input is such that the spatial resolution is preserved after convolution, i.e. the padding is 1-pixel for 3x3 conv. layers. Spatial pooling is carried out by five max-pooling layers, which follow some of the convolutional layers (not all the conv. layers are followed by max-pooling). Max-pooling is performed over a 2x2 pixel window, with stride 2.

Three Fully-Connected (FC) layers follow a stack of convolutional layers (which has a different depth in different architectures): the first two have 4096 channels each, the third performs 1000-way ILSVRC classification and thus contains 1000 channels (one for each class). The final layer is the soft-max layer. The configuration of the fully connected layers is the same in all networks.

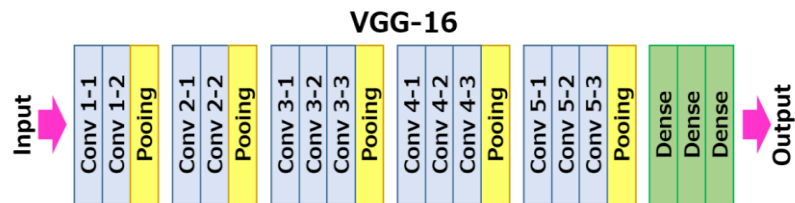


Figure 3: Architecture of VGG16

WEB APPLICATION DEVELOPMENT

React JS is a JavaScript library used in web development to build interactive elements on websites (User Interfaces), it is maintained by Facebook React makes it effortless to create interactive UIs. Due to its many advanced features like Virtual DOM, Interactive interfaces, Easy-to-use, Alternatives to other libraries, SEO friendly, Flexible Development, Quick access to developers this is been used for developing the web application. An application is developed in such a way that an input image of the cancer image will be given and an output is obtained. A database will be managed using which the prediction process happens in the backend and the result is been fetched through the web application

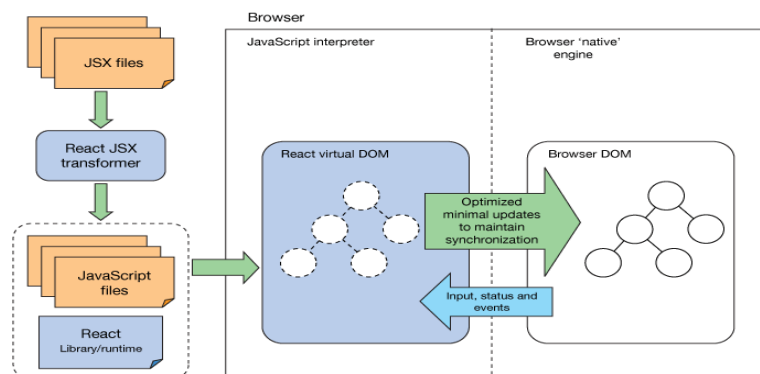


Figure 4: Web Application working



III. EXPERIMENT AND RESULT

Dataset collection involves the process of collecting different cancer dataset. These datasets are then preprocessed to form an equal aspect ratio so that it can be made ready for training with the model. The next step is data augmentation which duplicates the collected data into many images so that the prediction percentage can be increased. After this the final implementation is done where the training process takes place and the results are obtained. Then the code is run using certain commands in the command prompt. As the execution gets completed the accuracy is been plotted for the training obtained.

```
[INFO] evaluating after fine-tuning...
```

	precision	recall	f1-score	support
brain_normal	1.00	1.00	1.00	23
brain_tumour	1.00	1.00	1.00	30
cervical_cancer	0.99	1.00	0.99	65
cervical_normal	1.00	0.98	0.99	55
ignore	1.00	1.00	1.00	85
lung_cancer	1.00	1.00	1.00	87
lung_normal	1.00	1.00	1.00	79
skin_cancer	1.00	1.00	1.00	34
skin_normal	1.00	1.00	1.00	35
thyroid_cancer	1.00	1.00	1.00	25
thyroid_normal	1.00	1.00	1.00	24
accuracy			1.00	544
macro avg	1.00	1.00	1.00	544

Figure 5: Efficiency Output

Precision describes the accuracy. Recall is calculated as the ratio of no of positive samples correctly classified as positive to the total number of positive samples. The higher the recall the more positive samples are detected. F1 score is calculated as $2 * (precision * recall) / (precision + recall)$.

A web application is developed using a JavaScript framework reactJS for interactively uploading x-ray images and viewing the results. On successful login the home page is displayed to the user. The user can upload X- ray images to check for cancer. If the uploaded image is that of a cancer type the type of cancer is mentioned along with it.

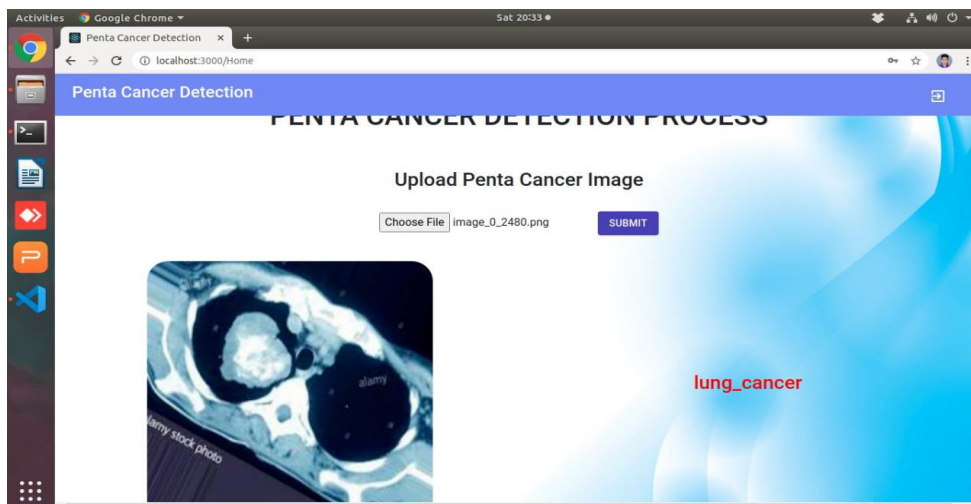


Figure 6: Lung Cancer

The above image shows that the uploaded image is that of a lung cancer



IV. CONCLUSION

The project has been successfully implemented to detect the presence of five types of cancers and determine whether the person has skin cancer, lung cancer or brain tumor and provide prior measures to avoid the disease. This also helps in providing efficient treatment in a most cheap way and eventually reduces the time required for finding the cancer. In the current state it is done manually which consumes more time and also involves human error rate.

So, this project reduces the time required for manual classification and eliminates the human error rate by this project. In the coming future, we review the application of the cancer determine technology in the healthcare field and it can promote for detecting various types of cancer with more accuracy. In medical field they have more chance to develop or convert this project in many ways. Thus, this project has an efficient scope in coming future where manual predicting can be converted to computerized production in a cheap way.

REFERENCES

1. Hangfan Liu, Hongming Li, Mohamad Habes, Yuemeng Li, Pamela Boimel, James Janopaul-Naylor, Ying Xiao, Edgar Ben-Josef, and Yong Fan, Senior Member, "IEEE Robust Collaborative Clustering of Subjects and Radiomic Features for Cancer Prognosis", IEEE Transaction on Biomedical Engineering [vol:2969839,2020].
2. Arnaud A. A. Setio, Francesco Ciompi, Geert Litjens, Paul Gerke, Colin Jacobs, Sarah J. van Riel, Mathilde Marie Winkler Wille, Matiullah Naqibullah, Clara I. Sanchez, Bram van Ginneken "Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks" ,IEEE Transactions on Medical Imaging.
3. Amir Mirbeik-Sabzevari, Student Member, IEEE, Negar Tavassolian, Senior Member, IEEE "Ultra-Wideband, Stable Normal and Cancer Skin Tissue Phantoms for Millimeter-Wave Skin Cancer Imaging", IEEE Transactions on Biomedical Engineering.
4. Yi Wang, Na Wang, Min Xu, Junxiong Yu, Chenchen Qin, Xiao Luo, Xin Yang, Tianfu Wang, Anhua Li, and Dong Ni* "Deeply-Supervised Networks with Threshold Loss for Cancer Detection in Automated Breast Ultrasound", IEEE Transactions on Medical Imaging.
5. Jose M. Anton-Rodriguez , Peter Julyan, Ibrahim Djoukhar, David Russell, D. Gareth Evans, Alan Jackson, and Julian C. Matthews "Comparison of a Standard Resolution PET-CT Scanner With an HRRT Brain Scanner for Imaging Small Tumors Within the Head" IEEE Transactions on radiation and plasma medical sciences, vol. 3, no. 4, July 2019.
6. Shekoofeh Azizi, Sharareh Bayat, Pingkun Yan, Amir Tahmasebi, Jin Tae Kwak, Sheng Xu, Baris Turkbey, Peter Choyke, Peter Pinto, Bradford Wood, Parvin Mousavi*, Purang Abolmaesumi* "Deep Recurrent Neural Networks for Prostate Cancer Detection: Analysis of Temporal Enhanced Ultrasound" IEEE Transactions on Medical Imaging.
7. Jun Xu,* Member, IEEE, Lei Xiang, Qingshan Liu, Senior Member, IEEE, Hannah Gilmore, Jianzhong Wu, Jinghai Tang, and Anant Madabhushi* "Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images" IEEE Transactions on Medical Imaging.
8. Koyel Mandal, Rosy Sarmah, and Dhruva Kumar Bhattacharyya "Biomarker Identification for Cancer Disease Using Biclustering Approach: An Empirical Study" IEEE/ACM Transactions on Computational Biology and Bioinformatics.
9. Sepideh Rahmatinia, Student Member, IEEE and Babak Fahimi, Fellow, IEEE University of Texas at Dallas, Department of Electrical and Computer Engineering, Richardson, TX, USA "Magneto-Thermal Modeling of Biological Tissues: A Step towards Breast Cancer Detection" IEEE Transactions on Magnetics.
10. Yanbo Wang 1, Weikang Qian 2, Bo Yuan 1,†1 Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, 200240, P.R.China 2 UM-SJTU Joint Institute, Shanghai Jiao Tong University, Shanghai, 200240, P.R.China "A graphical model of smoking-induced global instability in lung cancer†" IEEE/ACM Transactions



INNO  **SPACE**
SJIF Scientific Journal Impact Factor

**Impact Factor:
7.512**

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details