



INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

REVIEW OF CANCER DETECTION AND PREDICTION USING IMAGE PROCESSING AND MACHINE LEARNING.

Shraddha Pawar , Aishwarya Suryavanshi , Juhi Birari , Sujata Telangi , Prof. D.S.Shingate Sir

Savitribai Phule Pune university

Sandip Institute Engineering and Management , Nashik

Abstract

Today information handling is become an integral part of our day to day life as it is available in abundance. The information from the internet can be used for various purposes from Security to Healthcare management. It can decrease the load on a human being to handle the data manually. So, to highlight the importance of information handling to the society we choose medical field where images are in abundance on the internet. The analysis of images for decision making in medical perspective can be a boon to the medical field and decrease the work of the doctor and increase his productivity to a new level. The cancer has become a menace these days. Cancer is a major health problem. Computer-aided diagnosis (CAD) systems can be used to provide an insight to the cancer specialist and help him to diagnose the various stages of cancer using images and CAD. So, we thought of designing a system which will help in cancer detection and handling various stages of cancers using cancer images as the backend of our application. In our proposed system we are going to accumulate a lot of images from various cancer image databases which are available free on the internet. Then we will apply machine learning to it. Thus, our results will be divided in three categories benign, malignant and normal using machine learning algorithm CNN. Thus, our system will help and improve the decision making and productivity of a medical practitioner.

Index Terms: - Skin Disease, Median Filter, SIFT, Image Processing, Machine Learning, CNN.

INTRODUCTION

Cancer is a collective term for a large group of diseases that can spread to any body part. One important characteristic which makes cancer deadly is the rapid creation of abnormal cells that grow beyond their usual limitations and become lethal, and then attack adjoining cells and spread to other organs. The term used for this process is metastasizing. According to World Health Organization Website Cancer is the second leading cause of death globally, and is responsible for an estimated 9.6 million deaths every year. 1 in every 6 deaths in the world is caused due to cancer as per the WHO data. As with any type of cancer, the earlier it is detected, the greater is the chance for survival. The cancer has two stages benign and malignant. In benign stage the cancer is curable and in malignant stage there are high chances of patient succumbing to cancer. For cancer diagnosis the healthcare professional has to depend on his viewing and testing various symptoms which can be time consuming. If a cancer is detected in early stages it can be treated and has a high chance of survival then the patients where the cancer is detected late. So there arises a need to develop a software which will help a medical professional and patient to analyze the symptoms using image processing and machine learning techniques which will help in finding the cancer stages easily.

LITERATURE SURVEY

This topic describes the fundamentals of Image processing techniques that can be used to design a accurate Cancer detection and prediction system It helps in understanding

various ideas put forward by various technical papers published by various publishers.

- A. Fulgencio Navarro, Marcos Escudero -Vinolo and Jesus Bescos. et al. [1] authored a paper in 2019 which studies image processing as an image registration approach that outperforms top image registration techniques. Combined with the proposed lesion segmentation algorithm, this allows for the accurate extraction of features to assess the evolution of the lesion in a skin cancer scenario. Thus, it presents a case study with the lesion-size feature, paving the way for the development of automatic systems to easily evaluate skin lesion evolution. The main drawback of this paper is that this paper gives more dependance on feature extraction, Segmentation and not on prediction using machine learning.
- B. Feng Guo, Jie Yang, Yilei Chen, Bao Yao. et al. [2] authored a paper in 2018 where study shows that SIFT algorithm is invariant on rotations, translations and scaling and SIFT features have strong matching robustness for radiation transformation, perspective changes, illumination changes and noises. This paper does a comparative analysis of different results obtained by different ratio threshold and finally set 0.6 as the best value considering the balance number of matched points and matching accuracy of a given image. It is important to image recognition application. The main drawback of this paper is that this paper gives more dependance on feature extraction, Segmentation and not on prediction using machine learning.
- C. Tu Yawen and Guo Jinxu. et al. [3] authored a paper in 2018 where study identifies the possible regions in the image by collecting image classifier samples, using SIFT algorithm to extract the eigenvalues, using support vector machine (svm) classifier to classify, and using Gaussian pyramid in combination with sliding window. Then, to extract the results by the non-maximal suppression algorithm to improve the recognition rate of vehicle recognition system using SIFT. Based on this, to perform the experimental verification. The experimental process and the results analysis show that the vehicle identification accuracy rate of the above method reaches 93.4% using the above technique and combining the image processing and machine learning together.

PROPOSED SYSTEM

So, by using the above studies that are defined in the literature survey we get in to a conclusion that to get a better and more accurate results we have to use image processing and machine learning together. So, in the proposed system we will do the same.

A. GOAL AND OBJECTIVES

- To improve productivity and decision making of a Medical Practitioner.
- Maintaining an accurate database of various stages of cancer images.
- To remove noise from image using Median Filter algorithm.
- To apply SIFT for finding features and matching the image with a specific stage of skin cancer.
- To try and create an accurate skin cancer detection features dataset with three classes benign, malignant and normal.
- To apply and analyze the various cancer stages defined by the above three classes using CNN algorithm.

B. STATEMENT OF SCOPE

- Preprocessing of images: In this the system will remove noise from image using median filter algorithm.
- Feature Extraction of images: In this the system will extract features from image to be used for machine learning.
- Training and Testing Dataset: In this the system will create training and testing datasets to train SVM and test using SVM.
- Result prediction: In this the system will apply CNN on testing image and return the prediction for image in three classes benign, malignant and normal.

MATHEMATICAL MODELING

Mathematical model

$S = \{I, O, F, DD, NDD, \text{Success}, \text{Failure}\}$

Where,

$I = \{\text{image}\}$

$O = \{\text{Cancer Detection}\}$

$F = \{\text{browseImage}(), \text{upload}(), \text{conversion}(),$

$\text{featureExtraction}(), \text{featureMatching}(),$

$\text{generateResult}()\}$

$DD = \{\text{null}\}$

$NDD = \{\text{image}, \text{result}\}$

Success = Cancer detected successfully

Failure = Low light intensity may affect accuracy

FUTURE SCOPE

This project presents several sections on state of art techniques, analysis and comparisons on benchmark datasets for the brain tumor, breast cancer, lung cancer, skin cancer detection respectively of measure, sensitivity, specificity, accuracy, precision point of view it is feasible to implement. Different types of cancer detection and classification using machine assistance have opened up a new research area for early detection of cancer, which has shown the ability to reduce manual system impairments.

MOTIVATION OF PROJECTS

1. Machine Learning is a branch of AI that uses numerous techniques to complete tasks, improving itself after every iteration.
2. Pathologists are accurate at diagnosing cancer but have an accuracy rate of only 60% when predicting the development of cancer.
3. Machine Learning is the next step forward for us to overcome this hurdle and create a high accuracy pathology system.
4. A smart cancer detection and prediction system using a combination of image processing, features extraction, dataset generation and machine learning.

ACKNOWLEDGEMENT

We are thankful to our guide Prof. D. S. Shingate, SIEM, Nashik for the Guidance. We needed his essential guidance and suggestions.

DATA MODEL AND DESCRIPTION

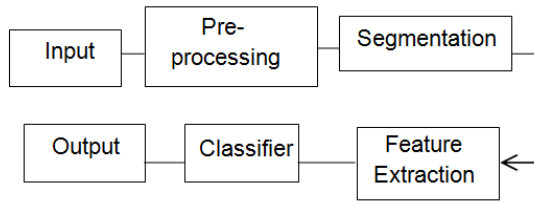
Data Description

Data objects that will be managed/manipulated by the software are described in this section. The database entities or less or data structures required to be described. Data objects are Java.

Data objects and Relationships

A relationship creates a two-way connection between two database objects. In general, relationships are named and each direction-specific connection is described using a short word or phrase. For example, in a reference relationship, one object references an object while the other object is referenced by the object.

BLOCK DIAGRAM



CONCLUSION

In this paper, we have studied various concepts that have been used by other users and thus developing a more accurate cancer detection and prediction system using image processing and machine learning together as new architecture. The image database needed will be downloaded from free databases available online. This paper mainly concentrates on two algorithms as image processing algorithm SIFT that will be used to extract features from image database that will be used as training and testing parameters for machine learning algorithm SVM. Thus, it will decrease the amount of time needed for medical practitioner and a patient to detect and understand the phase of cancer. The main drawback of this framework is that the training image quality should be good for a good end result.

REFERENCES

[1] Fulgencio Navarro, Marcos Escudero-Vinolo and Jesus Bescos., et al. "Accurate segmentation and registration of skin lesion images to evaluate lesion change" 2168-2194 (c) 2018 IEEE. Personal.

[2] Feng Guo, Jie Yang, Yilei Chen, Bao Yao., et al. " Research on Image Detection and Matching Based on SIFT Features." 2018 3rd International Conference on Control and Robotics Engineering.

[3] u Yawen and Guo Jinxu., et al. " Research on Vehicle Detection Technology Based on SIFT Feature." 2018 IEEE.

[4] M. Zhen, and J. M. R. S. Tavares. "A novel approach to segment skin lesions in dermoscopic images based on a deformable model." IEEE Journal of biomedical and health informatics,20(2), 615-623, 2016.

[5] B. Van Ginneken, et al. "Computer-aided diagnosis in chest radiography: a survey", Medical Imaging, IEEE Trans 20(12) 1228- 1241, 2001

[6] H. Cheng, et al. "Computer-aided detection and classification of microcalcifications in mammograms: a survey", Pattern recognition, 2967-2991, 2003.

AUTHORS PROFILE



Shraddha Pawar
Student, B.E. Computer Department.



Juhi Birari
Student, B.E. Computer Department



Ashiwarya Suryawanshi
Student, B.E. Computer Department



Sujata Telangi
Student, B.E. Computer Department

Prof. M.V .Korade
Assistant Professor, M.E. Computer Engineering.