



Prediction of Disease by Analysing the Cell Image of Human

¹R. Mathumitha, ²Dr. A. Padmapriya

¹M.Phil Scholar, ²Associate Professor

^{1,2} Department of Computer Science and Engineering, Alagappa University, Karaikudi-2,
Tamilnadu, India

Abstract— *The humans are affected by different disease now-a-days. The diseases which are affected by humans are predicted in early stage or in the severity stage. The severity stage of disease cause major damage to the human. In the list of major disease, cancer is treated as the high priority which is affect human severely. The cancer can affect any part of the body. Main area of body affecting cancer such as brain, lung, blood, stomach and colon. The disease can be identified by the symptoms. The objective of the prediction is to assign patients to one of the two group either the 'benign' that is noncancerous or a 'malignant' that is cancerous. This helps to mine the data easily. In proposed work, the results of the affected cell can be easily and accurately determined by the cell image of the cancerous area of the human body. The affected and the normal cell are compared to produce the approximate identification of the diseased cell. Aim of the paper to propose a model for early detection and correct diagnosis of the disease which will help the doctor in saving the life of the patients.*

Keywords— *Data mining, cancer, proteomics, genomics, cell image, segmentation, Comparison.*

I. INTRODUCTION

The health management of hospital consist of many disease details which affects the human. Those management mine the large amount patients details. The disease reports are produced to the patients under some specific time. The severity of the disease can be varied from different levels. In that level, some disease effect can be occurred only for a week, some other disease may be cause to death. So the identification of the disease can be accurate which is more important to the patients. The identification can be done by specialist in those fields. The result might be sometime accurate or not. In this paper the disease discussed is cancer^[1].

Now-a-days the severity of the cancer disease is increased and the people affected by cancer are also increased. People trust only the health management to know the details about their body conditions. So the prediction of the disease must be accurate. In the existing system, the identification of disease by the cell value such as the cell membranes, chromosome etc^[2]. The technique used for identification of disease in the existing system produces correct results and sometime not and also it produces the high time complexity. To rectify those drawbacks, the proposed system is developed which uses the cell image comparison to find the accurate results.

The proposed system takes inputs such as normal cell images and disease affected cell images of lung, stomach, colonists, blood cells etc. Comparison technique uses both the normal and the affected cell images to produces accurate results. And also it reduces the time complexity.

II. ANALYSIS OF CANCER DISEASE

The cancer is a group of disease which are characterized by unregulated division and the spread of cells^[3]. The cancerous cells occur in solid tumors that may be appearing in the tissues in a various part of the body. By their original location they are classified into various types of cancer such as lung, colon, breast or pancreatic cancer. The analysis and the identification of disease may be categorized into two ways such as the animal and the epidermal studies.^[4]

The animal studies gives the investigator for the advantage of controlling the condition under which animal are exposed to various level to an given substance, their diet and even their genetic make-up. The analysis also includes epidemiologic studies are sometime referred to as "Natural experiments in the real world that must be evaluated for potential sources of bias or chance that may be have influenced result".

There are many types of cancer are identified are

- Bladder cancer
- Bone cancer
- Breast cancer
- Colon cancer
- Kidney cancer
- Lung and Liver cancer
- Pancreatic cancer
- Skin cancer
- Stomach cancer
- Thyroid cancer

III. EXISTING TECHNIQUES USED IN IDENTIFICATION OF CANCER DISEASE

The prediction of the cancer disease includes various techniques such as segmentation of extraction^[2] of the affected region. The disease area is segment for extraction with the help of segmentation algorithm. Those segmented area are analyzed and they are evaluated under the diagnostic indicator. These steps are processed for the analysis of the cancers. The prediction of the cancer can be varied in different types. In the view of the scan system the affected the affected are scanned and they extracted for segmentation and also for the analysis for the proper diagnosis. Another type of analyzing disease by well suited medical applications such as the measurement of proteomic and genomic^[7]. One of the most common forms of medical malpractices globally is an error in diagnosis.

IV. DRAWBACK OF EXISTING SYSTEM

The existing techniques might produce results and sometimes it produce inaccurate results also. The value produces after the extraction or the screening^[5] process of diseased area might gives the inaccurate results. The technology used for the analyzing process consists of high features which also lead to the high cost consumption. The time taken for the identification of disease also produces the high time complexity. These high technology sometime does not produces the accurate results prediction of the disease.

V. PROPOSED TECHNIQUES TO IDENTIFY THE DISEASE

The proposed techniques which used for identification of the disease rectify the early drawbacks of the existing system. The work of the proposed techniques is simple to identify the disease. The proposed technique uses the inputs as the cell images of human.

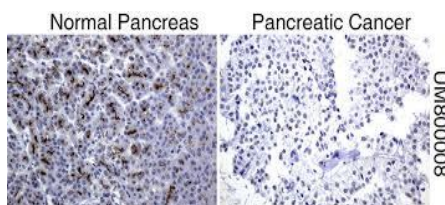


Fig 1: cell image of normal pancreas and cancerous pancreatic

Those cell image are taken from the sample tissue or the blood test. For the testing process, both the normal cell image and affected cell image are taken. The testing operation can be categorized into two process. The first process is the segmentation^[6] of the cell image. The sample cell image size is 256*256 Which is equally divided in to 16 blocks of rows and columns. Both normal and abnormal cell images are spitted into the 16 blocks of image set. Each segmented block is named and stored in a separate place.

In the testing operation, second process is defined as comparison method. The comparison process is used to find the similarity of the cell image. The similarity between the cell images defines as comparison between normal and the abnormal cell or comparison between abnormal and abnormal cell. The comparison process is considered as the main operation in the system .This is the main process to produce results.

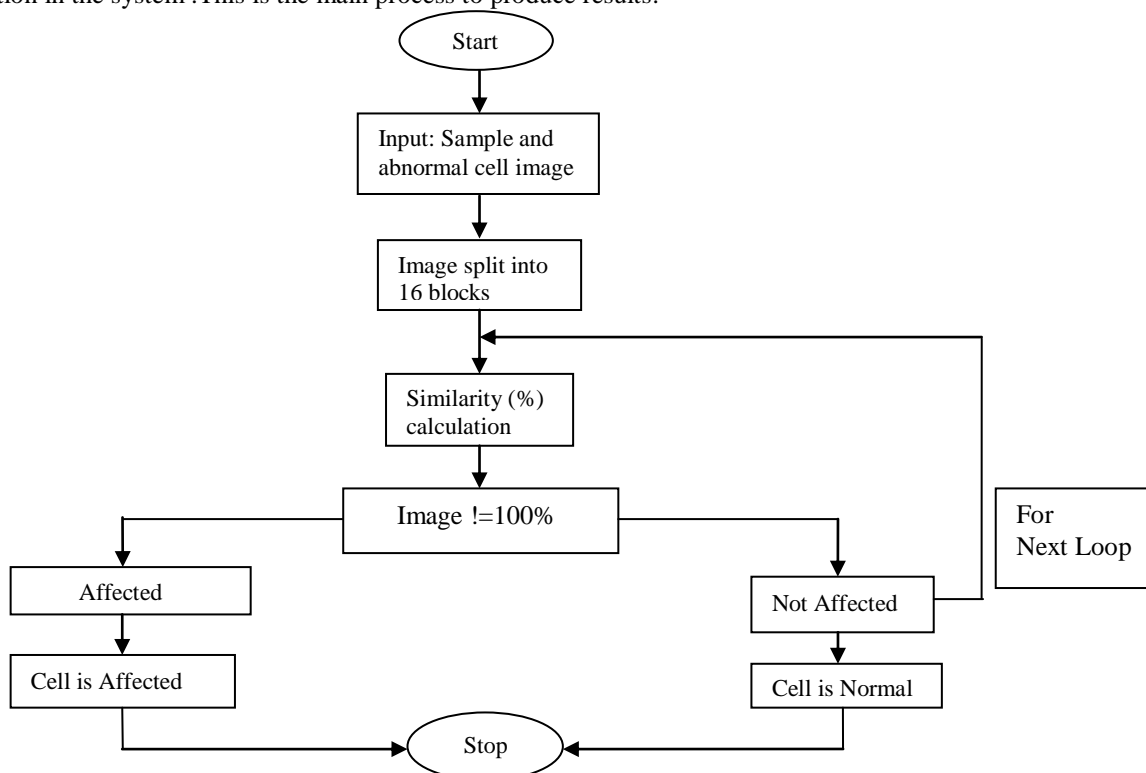


Diagram 1: Block diagram describe the step by step process in proposed work.

A. Algorithm used in the prediction of the disease:

- Step 1: Start
- Step 2: Get the input image such as normal and the abnormal cell image.
- Step 3: Split the given sample image into 16 blocks.
- Step 4: Split the given affected image into 16 blocks.
- Step 5: For each block
- Step 6: Find the similarity percentage between the sample cell image and the affected cell image.
- Step 7: If similarity not equal to 100% then the cell is affected go to step 10
- Step 8: else Next block(Loop continues)
- Step 9: Print “The cell is normal”
- Step 10: Print “ The cell is affected”
- Step 11: Stop

The function of the algorithm defines as first the program starts then it gets the two input value as a image such as normal cell image and the abnormal cell image. Then these cell images are splited into 16 blocks. Some condition is applied for those blocks such as for loop and if condition. The for loop used to check every block in both image which is splitted into 16 blocks. Then the if condition is applied to the block to check the similarity between the splitted blocks of the both images by pixel wise. If the similarity percentage is 100% then it checks the another block and so on. Finally it gives the result as “The cell is Normal”. But if the similarity is not equal to 100% then the cell is defined as the affected cell. In sometime the percentage are defined as point variation so the cell is partially affected. The result may be divided into three cases.

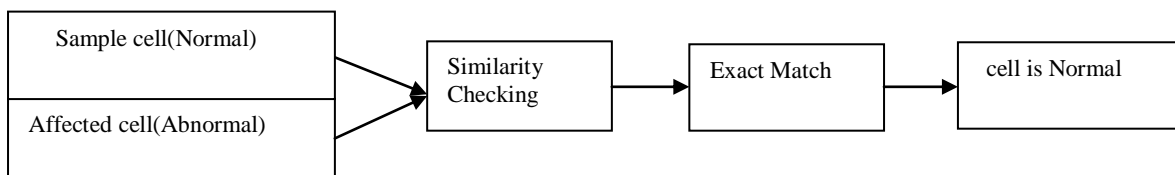


Diagram 2: comparison operation of first case.

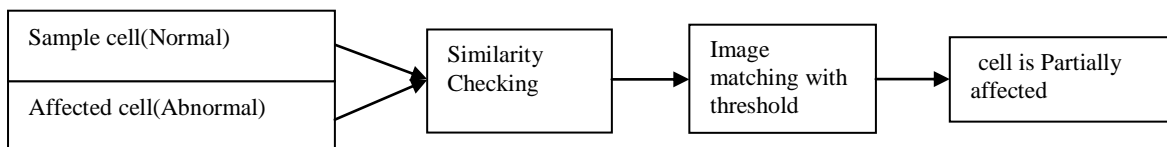


Diagram 3: Comparison operation of second case.

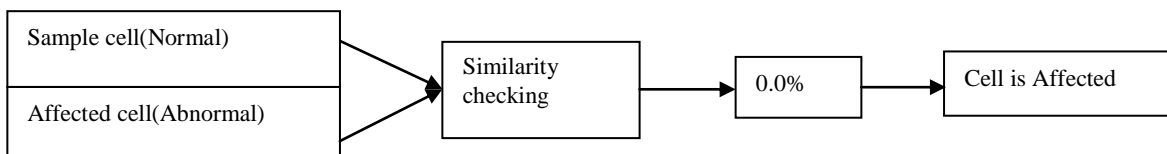


Diagram 4: Comaprison operation of third case.

The three case can be defined as

Case 1: If similarity equals to 100%, the cell image is defined as normal.

Case 2: If similarity does not equal to 100% but they produce result in point variation. The cell is defined as partially affected and produces result as affected.

Case 3: If similarity does not equal to 100% and the result is 0.0%. Then the result produced is cell affected.

By the two process in the system, the result can be produced accurately. Comparison checks the images by pixel wise to find the similarity of the image. This process helps to find the accurate results, which are more helpful for the patients.

VI. ADVANTAGE OF PROPOSED SYSTEM

The proposed system consist of more advantage such as

A. Software used in System:

The proposed work was done by the software such as core java consists of simple coding and they are object oriented. The main usage of java coding is platform and device independent. It is said to be reliable language. It also prohibits the file from introducing virus and deleting or modifying the files. Maintaining different version of an application is very easy in java. Java is capable of dynamically linking in class library, methods and objects.

B. Reduced time complexity:

With the comparison of existing system, the time complexity is reduced in the proposed system. The time taken for producing results of the system is low. Within the fraction of second the results might be displayed. The worst case of the system is said as if the similarity is not found in the first block it is declared as affected but sometime the affected area in

the cell image is appeared in the lost block. So the time taken is slightly increased in such case. So these case are defined as worst case in the proposed system.

C. Memory space management:

The space managed by both the existing and the proposed system are same. There is no difference between the memory occupation.

D. Accuracy of data:

The result produced by the system are accurate of 80% when compared to the existing system. The System produces the result is not only whether only affected or not but they are also says how many percentage they are affected.

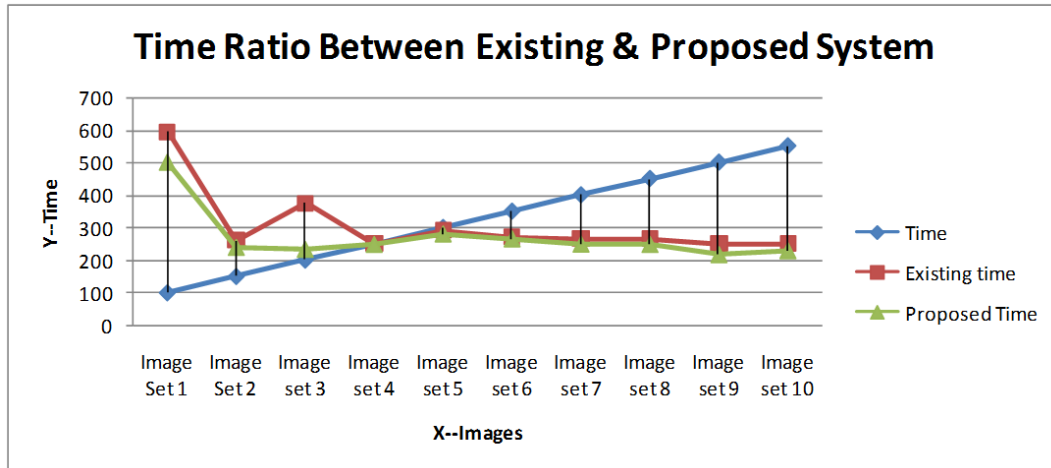


Fig 2: Graph defines the time ratio between the existing and proposed system

VII. CONCLUSIONS

The proposed work helps to predict disease not only used for cancer disease but also used in blood related or some other disease. The cancer are considered as major disease which must be predicted and give proper diagnosis treatment. The work is more useful and reduces the time complexity. The segmentation and the comparison technique of proposed system gives accurate result. These work is more useful to the health management and some cancer prediction or diagnosis center.

REFERENCES

- [1] *Cancer Research Through The Help Of Soft Computing Techniques: A Survey* Mr. Tanupriya Choudhury , Dr.Vivek Kumar , Dr. Darshika Nigam, Research Scholar (CSE), Jagannath University, Jaipur, India, Principal, Delhi College of Technology and Management, Palwal, India, Department of BioChemistry, SLS, Dr. B.R. Ambedker University, Agra, India.
- [2] *Identifying Lung Cancer Using Image Processing Techniques* , Disha Sharma, Gagandeep Jindal.
- [3] *Using Data Mining Techniques For Diagnosis and Prognosis Of Cancer Disease*, Shweta Khar Sr.Assistant Professor, Bhilai Institute of Technology, Durg-491 001, Chhattisgarh, India, Shweta.bitdurg@gmail.com
- [4] *Cancer causes and cancer research on many levels of complexity* , Sunny Y. Auyang.R.
- [5] *Oral Cancer Early Detection and Stages Using Various Methods*, G. Visalaxi , Assistant Professor, Dept. of CSE, Apollo Engineering College, Tamil Nadu, India.
- [6] *Diagnosis of Lung Cancer Prediction System Using Data Mining Classification Techniques* V.Krishnaiah, Dr.G.Narsimha, Dr.N.Subhash Chandra ,Associate Professor, Dept of CSE, CVR College of Engineering, Hyderabad, India, Assistant Professor, Dept of CSE, JNTUH College of EngineeringKondagattu, Andrapradesh, India, Professor of CSE & Principal, Holy Mary Institute of Technology and Science, Hyderabad, India.
- [7] *Application Of Data Mining In Bioinformatics*, Khalid Raza, Centre for Theoretical Physics, Jamia Millia Islamia, NewDelhi-110025,India
- [8] *”Biological data mining”*, George Tzanis, Christos berberidis, and Ioannis vlahavas Department of Informatics, Aristotle university of Thessaloniki, Greece.
- [9] *”Biological applications of multi-relational data mining”*, David page, department of biostatistics and medical informatics and department of computer sciences, University of Wisconsin Madison, WI 53706 – page@biostat.wisc.edu , Mark Craven, department of biostatistics and medical informatics and department of computer sciences, University of Wisconsin Madison, WI 53706- craven@biostat.wisc.edu
- [10] *”Application of data mining in bioinformatics”*, Khalid raza, centre of theoretical physis, jamia millia islamia, New delhi-110025, India.