CANCER CLUMPS DETECTION USING IMAGE PROCESSING BASED ON CELL COUNTING

Jayashree P K[1], Smt. T D Shashikala[2]

PG Scholor[1], Associative Professor[2], Department of Electronics and Communication Engineering, University B.D.T College of Engineering, Davanagere-577004, Karnataka, India

---

ABSTRACT

One of the leading causes of death in people is cancer. There has been a lot of research into using image processing, classification, and techniques to identify and diagnose cancer. But the illness continues to rank among the deadliest. As a result, one of the reasons to treat cancer is not just its early detection. In the proposed method, cancer cells are identified using image processing, artificial neural network techniques, area measurement, and cell clump detection. Using the suggested method, we can automatically identify cancerous characteristics in any CT image, mammography image, or biopsy sample. There were numerous proposed algorithms, but they lacked flexibility and had uneven levels of accuracy. The system pre-processes the input images using a variety of methods, including grey scaling, binarization, inversion, and flood fill operation, before applying the proposed algorithm. If the suggested approach can successfully be used for automatically detecting cancer cells in a novel way and fine-tuned with a feedback system, it will open up new dimensions in the detection of cancer cells in the field of medical sciences.

Keywords: Image Preprocessing, Image segmentation, ResNet50.

1. INTRODUCTION

The obstreperous division of abnormal cells is referred to as cancer. Tumors can be produced by the spread of cancerous cells through the lymphatic or circulatory systems. But it's important to remember that not all tumors are malignant. Both benign (not cancerous) and malignant tumors are possible (cancerous). Over a hundred different forms of cancer have been identified, and each type has numerous Sub types with unique variations. The early stages of cancer detection are particularly difficult due to the enormous variation. In the vast majority of cases, the causes of cancer are still poorly understood. Consequently, cancer treatment becomes considerably more difficult.

Additionally, due to the disease's extreme complexity, scientists, physicians, and engineers from around the world are conducting research in the subject of cancer in an effort to better understand the disease and, in the process, discover permanent treatments for each form of cancer. Even though the process is drawn out and challenging, having greater knowledge can help doctors treat cancer patients more successfully. This inspired us to consider how cancer is discovered and to use technology to hasten the process. Researchers studying cancer might save a tone of time and improve the effectiveness of their work if they can use image processing to automatically identify cancer cells. This is because the human error component will be completely eliminated. Enormous amounts of time and also improve the effectiveness of the research, as the possibility of human error will be eliminated entirely.

2. LITERATURE SURVEY

Kumar suggested a brand-new method for detecting malignancy that makes use of clinically preferred aspects. This approach is based on the K mean cluster and segmentation premise. Cell segmentation, feature extraction, and classification for the enhancement approach are a few of the processes in this procedure. The original image was divided into parts using the image segmentation concept of image processing. During the feature extraction process, the segmented images are used to extract the features. When KNN and SVM-based classifiers are used as the last step in a classification process, the results are shown for both a wide range of photos and the images with the selected feature set. [1]

Jain proposed a novel preprocessing technique to detect lung cancer. In this study, we applied a special noise reduction method to lessen the noise difference between the input and output photos. [2]

Ramin suggested an image analysis method for locating cancer cells and counting the number of cancerous cells in the source photos. Ramin provided an image analysis method for the identification of cancer cells and quantification of the number of malignant cells in the source images. The four essential steps in this approach are preprocessing, categorization, bound regions, and cell counting. During preprocessing, noise detection for the original image was eliminated.

Following the original image's classification using the KNN method, it will group into the same cluster value. On the basis of the output from the second stage, we then count the common cells in the third step. The bound nucleus separation process uses the local thresholding technique. Regarding error ratio and standard deviation, the results are good. [3]

Using the transform Technique, Thilagavathi introduced a novel technique for counting red blood cells. Red blood cell estimation uses this approach. The five phases of this method are presented. Feature extraction, segmentation, image collection, preprocessing, and counting. Apply the basic XOR operation to two binary images after determining the lower and upper threshold values for segmentation, preprocessing with the saturation image, and segmentation. One of the most popular ways to identify cancer is often through digital mammography. Many different solutions were presented in response to the multiple categorization problems for the digital mammography image. Many features are retrieved in this employing different standard procedures and fundamental notation. The tumour in this area is calculated using the MLE technique. One of the simplest ways to diagnose breast cancer is by misreading a mammogram. By doing this, we can locate and refine the image's edges. Finally, determine the size of the tissues and distribution in a picture without segmenting it. Additionally, cancer cells now have a quick and delicate location thanks to nanotechnology. The contaminated cells in a human body can be identified and eliminated using nanotechnology. [4]

Mello suggested two techniques for finding cancer cells inside a human body. These two approaches are differentiated by their use of colour. The RGB colour format of the input image is converted to the HSL colour model in this. To the binarization, the HSL colour model is employed. The edges of the binarization image are filtered to make them more rounded. [5]

3: DATASET

Since this work is intended for all type of cancer cell detection the dataset on cancer is available in www.kaggle.com.

4: METHODOLOGY

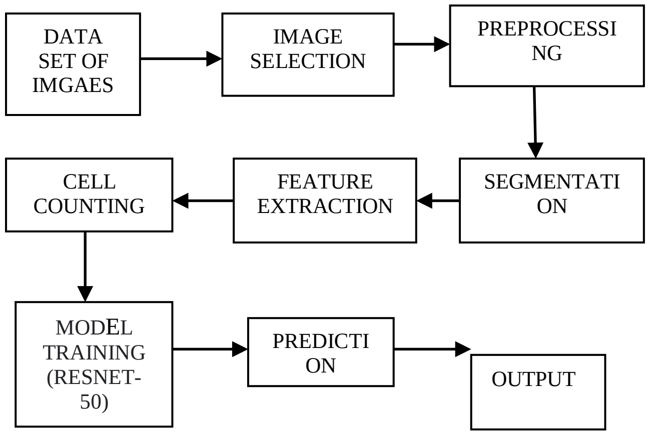

Fig: 1 Block diagram

To obtain more accurate results, the input image must go through a number of steps. These steps consist of

Image acquisition:

The cancer cell image considered hr for this work is Which is a microscopic image which is of size (90x90) Since this is too small for processing it needs resizing of the image.

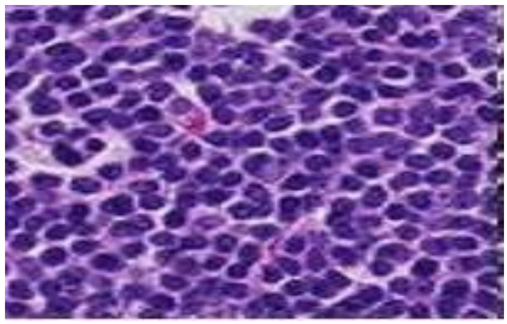

Fig 2: Input cancer cell image (size 90x 90)

RGB to Gray Conversion:

Using the RGB to Gray () function and the Luminosity Method, which states that 3-dimensional colour has three different wavelengths and their own contribution, we must take the average according to their own contribution and is given by the following equation, the original input cancer image is converted into a grayscale image to reduce the noise level of the image and for further segmentation.

$$Grayscale=0.299R+0.587g+0.114B.$$

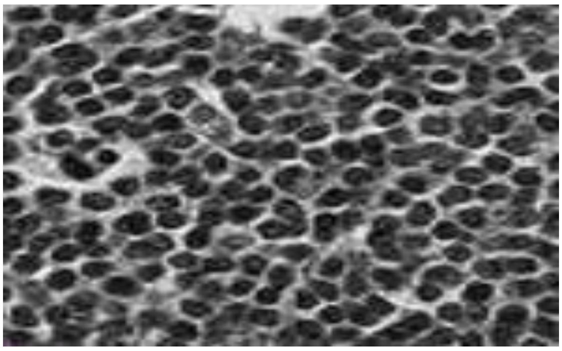

Fig 3: Input image is converted into Grayscale(size 150x150)

Binarization (Otsu's thresholding):

Otsu's thresholding method, which is based on classification, seeks the threshold that reduces intra class variance, which is determined by the weighted sum of variances for the two classes. The linear discriminant criteria used by Otsu's thresholding method presume that a picture simply consists of objects (foreground and background), with the heterogeneity and diversity of the backdrop being disregarded. To try to reduce the overlapping of the class distributions, Otsu's established the threshold.

Fig 4: Binarized image (otsu's thresholding)

Segmentation:

The technique of segmenting an image involves breaking it up into several region, to detect objects or other important details in digital image. There are many edge detector in DIP Robert edge detector, Prewitt edge detector, sobel edge detector, canny edge detector and log edge detectors here in this work canny edge detector, is used for cancer cell detection since Canny edge detector is best for visual appearance.

Canny edge detection is composed of five steps which are as follows.

1. Noise reduction using Gaussian filter
2. Gradient calculation
3. Non-maximum suppression
4. Double threshold
5. Edge Tracking by Hysteresis

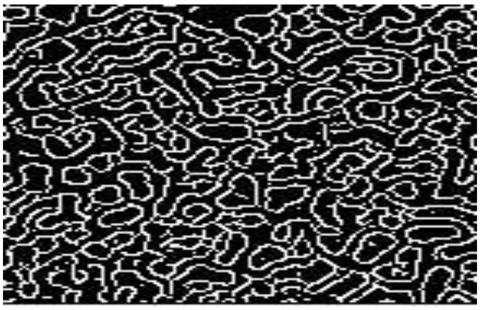

Fig 5: Canny based image

Closing:

J=imclose(I,SE)standard closing function, which uses the structuring element SE to perform morphological closing on the canny-based image. The morphological close operation is a dilation followed by an erosion, with both operations utilising the same structuring element

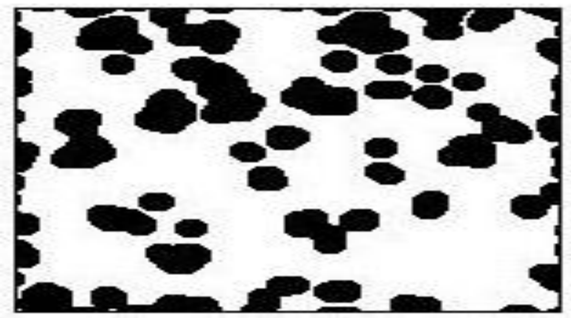

Fig 6: Closing image

Inversion: Inversion is necessary in the work carried here since the counting operation is performs only on the for ground pixels, performing inversion which make easy for cell counting just by Inverting the closing .

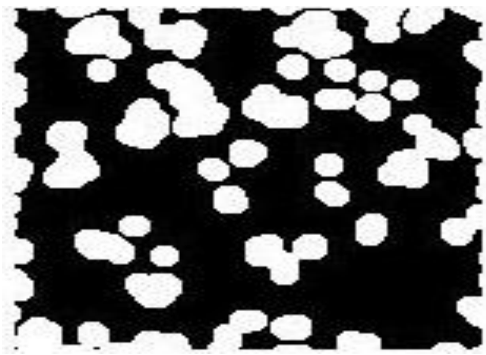

Fig 7: Inverted image

Feature extraction:

Feature extraction attempts to reduce the number of features in a dataset by generating new ones from existing ones (and then discarding the original features).This new, smaller set of features should then summarize the majority of the information in the original collection of features.

Cell Counting:

To detect the boundary lines of the original image or cells, we use a clever edge detection algorithm that finds all of the original image's edge pixels. We used the closing operation after identifying the edge pixels of the original image.

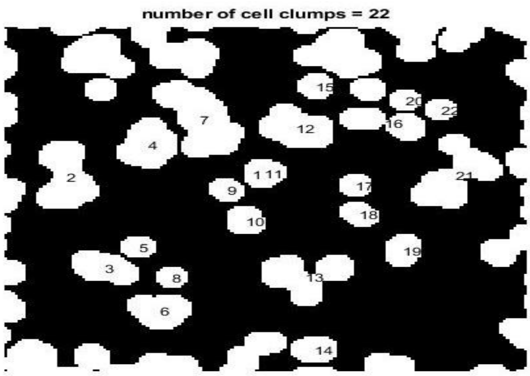

Fig: 8 Cell counted image

Model training using ResNet-50:

ResNet-50 is a convolution neural network with 50 layers. A pretrained version of the network that has been trained on over a million photos is available in the Image Net database. The pertained network can classify photos into 1000 different object categories, such as animals, a keyboard, a mouse, and a pencil.

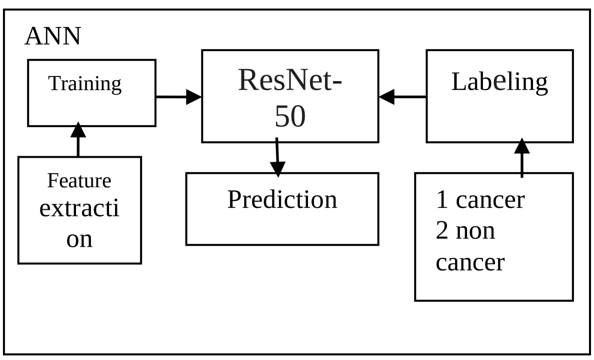

Fig 9: Working of ResNet-50

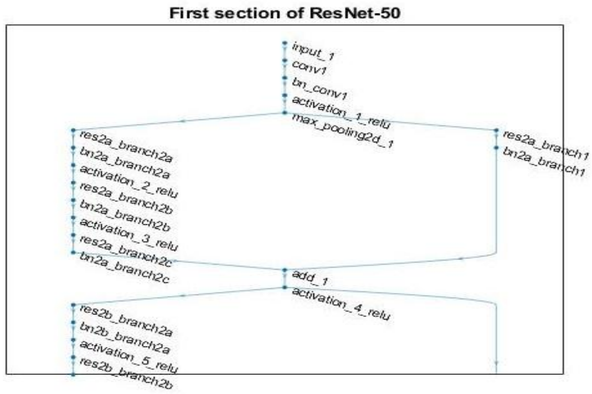

Fig 10: First section of ResNet-50 layer used

Here in this work after cell counting the image is giving as training image for ResNet-50. The 1st layer of ResNet-50 is used for the prediction, if the cells in the image is cancerous then that image is labeled as cancer and other is non cancer. The image data store contains images of various sizes, but the network requires images with dimensions of 224 by 224 by 3. Use an enhanced image data store to automatically resize the training images. Provide instructions for performing additional augmentation operations on the training images, such as randomly flipping them along the vertical axis, translating them up to 30 pixels, and scaling them up to 10% both horizontally and vertically. Data augmentation is used to keep the network from overfitting and to retain the precise characteristics of the training images.

Confusion Matrix:

The confusion matrix is used to evaluate the performance of classification models for a specific set of test data. It can only be determined once the true values of the test data are known. To assess the efficacy of a classification model, a N x N matrix known as a confusion matrix is used, where N is the total number of target classes. Here if the predicted value is logic 1 then it is cancer and if the predicted value is logic 0 then it is non cancer.

	Positive	Negative
Positive	TP	FP
Negative	FN	TN

Fig 11: Confusion matrix

5: RESULTSANDDISCUSSION

In this project work main concern is to detect the cancer clumps in a microscopic image and further classifying those images based on the cell counting. Here by selecting the one of the cancer cells images which undergo certain digital image processing techniques that help us to detect the clumps in the cancer cell image by using machine learning algorithm we are able to classify clumps and to which class the image belongs to, here mainly considering two classes cancerous and non-cancerous. This model proves to work for all the cancer cell images.

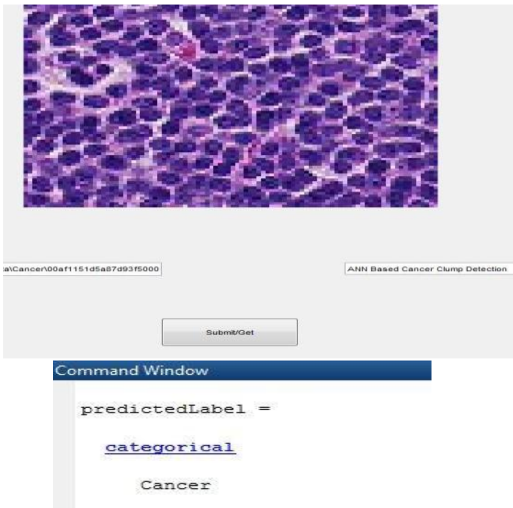

Fig 12: cancer cell prediction

Image 13: non cancer cell prediction

6: CONCLUSION

In medicine early cancer cell detection is considered as one of the challenging tasks for the doctors. The proposed work effectively detects the number of clumps at the early stage by using different image processing techniques like preprocessing, segmentation, feature extraction and training later the detected clumps in a cancer image are trained using ResNet50 algorithm to classify the cancer cells into two different classes weather it is cancer or non-cancer. In this work different cancer cell images were tested and proved successfully. This project helps us to save the patients by detecting atearly stage.

7: REFERENCES:

1. http://lymphomapictures.org/p/37/non-hodgkin-lymphoma/picture-37
   *2+ Kumar R., “Detection and Classification Using Clinically Significant and Biologically Interpretable Features” Proc of Journal of Medical Engineering,Volume 2015 (2015), Article ID 457906, 14 pages.

2. Rammin M., M., “Counting Number of Cells in Images using Genetic Algorithm,” 12th International Conference on Hybrid Intelligent Systems, Dec. 2017. pp. 185-190.

3. Thillagavathi K., “Automatic Red Blood Cell Counting in images UsingHough Transform,” Proc. of 2016 IEEE Conference on Information and Communication Technology, Apr. 2016, pp. 267-271.

4. Meello, , Marco A.,“Imaage sedments for artificial and automatic process for identifying cancer lumps Eggs,” 30th Annual International Aug. 2016. pp. 3103-3106.

5. Ammon, G.(2012, April 9). Image Segmentation for images using Digital Signal Processing. Retrieved from:

6. A history of medical imaging (2017). Retrieved from: PPT R. Boyyle and R. Vision: A First Course,On Image Processing Blackwell Scientific Publications, 2012, page. 32 - 34.

7. E. Davies e Vision: Theory, Algorithms and Practicalitiesusing Image processing and Image segmentation , Academic Press, 2011, Chap. 3.

8. Gonzalaez, R.C and Wooods, R Digital Image Processing, Addison Wesley, 2014 2, pp 414 – 428