2023 2nd International Conference for Innovation in Technology (INOCON)
Bangalore, India. Mar 3-5, 2023

Bone Cancer Identification and Separation Using K-Means and KNN Classifiers

Shashikala S V
Dept. of CS&E, Adichuchanagiri University, B G Nagara,
Karnataka, India
shashisv7@gmail.com

Uma H R
Dept. of CS&E, Adichuchanagiri University, B G Nagara,
Karnataka, India
umamajupari.27@gmail.com

Sunad Kumara A N
Dept. of CS&E, Adichuchanagiri University, B G Nagara,
Karnataka, India
sunadkumar.an@gmail.com

Taranath N L*
Dept. of CSE,
Alliance College of Engineering and Design, Alliance University,
Bengaluru, Karnataka, India
taranath.nl@alliance.edu.in*

Lokesh Singh
Dept. of CSE,
Alliance College of Engineering and Design, Alliance University,
Bengaluru, Karnataka, India
lokesh.singh@alliance.edu.in

Deepti Sisodia
Dept. of CSE,
Alliance College of Engineering and Design, Alliance University,
Bengaluru, Karnataka, India
deepti.sisodia@alliance.edu.in

2023 2nd International Conference for Innovation in Technology (INOCON) | 979-8-3503-2092-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/INOCON57975.2023.10101031

***Abstract* - The unregulated cell growth will lead to the dangerous and deadliest cancer disease. In the body of humans various kinds of cancer has been detected, after going on many researches. Among all these, bone cancer is the one which will spread more widely, because of this it will leads to death. There is no anticipation for the bone cancer, therefore the bone cancer detection is more critical. Nowadays, methods like data mining and the image processing methods are utilized for the most of the studies in the process of medical image analysis. Many scientific researchers have been predicted the data, related websites and the collection of knowledge from the large databases. There are many methods used in the approaches for bone cancer detection and classification like supports vector machines, Association rule mining and fuzzy theory. In this approach of segmentation k-means will be utilized for segmenting the bone regions. In further processes for bone cancer detection the segmented image will be used by the mean intensity evaluation of the area identified. To check whether there is a presence or absence of bone cancer in the medical images that is for the classification process threshold values are used. This approach can be used for jpeg images, CT scan images. The proposed work will use the K-Nearest Neighbor (KNN) classifier as a classification technique and achieved to produce better accuracy.**

Keywords : KNN, K-Means, CT Scan, DNA

I. INTRODUCTION

Bone cancer affects almost all of the human body's organs and tissues, dividing cells into different types, spreading, and forming tumours. The bone tumour will count up and spread throughout the digestive system, nervous system, circulatory system, and it will also enter the body's hormones. As a result, the body's functioning will be impaired. Due to the actual harm to DNA, those cells are deemed cancer cells. As compared to typical cells, the cell maintains the damage and dies if the DNA is harmed. If the damaged DNA is not refurbished, it will result in the creation of obscure new cells. When bone cancer cells proliferate and scattered to other places of the body, this is referred to as metastasis. when it starts to develop into tumours that are positioned within the human body's tissues. The term "metastasis" refers to the process through which cancer cells spread from one region of the body to another, including the bones. Breast cancer, bone cancer, and prostate cancer all have a propensity to migrate to the bones. Throughout addition to the normal bone pain and weakness associated with bone cancer, this metastasized version of the disease also results in weariness and weight loss.

A bone tumour is a malignant development of tissue in bone. Because of the redundant development observed in the bone, it can be benign (non-cancerous) or malignant (cancerous). Cancer is divided into two types: "primary tumours," which form in bone cells or tissues, and "secondary tumours," which occur in other locations and spread to the body's skeleton. Primary tumours are cancers that arise in the bone, whereas secondary tumours are cancers that appear in other regions of the body. Once more, cancers and benign tumours are separated from primary tumours. Primary bone cancer is produced by the genetic proliferation of cancer cells, which begins in the bone components solely. Sarcomas are the most common kind of bone cancer. This sarcoma will invade the blood vessels of the nerves, muscles, bones, and fat tissue in the body and can spread to any part of the body. These primary tumours, which may have aberrant tissue development and produce traumatic infections or inflammation in the bone, are also known as benign tumours. Secondary bone cancer refers to metastasis bone cancer.

A. Image processing

It is a type of process which will allows for the enhancement of image features that are interested while attenuating the details which will be irrelevant for a given set of applications, and they are used for the extraction of information that are helpful regarding results of the image enhanced. Variety of images will be obtained by using the various devices which will be utilized in many issues, like in the fields of entertainment, business for example in documents maintaining, industrial, military, in civil field it will be used for the traffic purpose, security, and scientific. The humans or machines goal will be to extract useful information about

images. In order to acquire, integrate and for the interpretation of all the abundant information of visuals surrounding us that is by human perception. To get the digital images the sampled pictures that are the signals captured by two dimensional are quantized.

The agenda of image processing is classified into five assemblies they are:

- Visualizing is a technique for creating animation, diagram, or images to communicate a message.
- Image restoration and sharpening technique is used to create better image.
- Image improvement is a PC framework for perusing, seeking and getting better images from a substantial image informational collection.
- Measurement of pattern is a branch of image processing that deals with recognition of pattern and regularity in an image
- Image recognition is to identify and detect an object or a video. This method is used in many fields of security surveillance, factory automation, and toll booth monitoring.

B. Motivation

Bone cancer is a deadliest disease compared to the other cancers because once the cancer cells is seen if it is not treated at the correct time that is it should be diagnosed in the 1st stage itself otherwise it will spread from one part of the body to other, it effects the whole bones present in the body very fast and leads to death. The automated system is necessary to diagnose the bone cancer, so this kind of system is in demand since many years it not only diagnose the bone cancer accurately, also it helps to find the new cases of bone cancer. This bone cancer is being increasing these days compared to the other type of cancers. The American cancer society has estimated that there are 3020 new cases are diagnosed in the year of 2016, in that 1460 have died.

In this paper, the CT (computed tomography) scanned images will be used for diagnosing and detecting bone cancer by pre-processing like noise removal, enhancement of images, segmentation, extraction of features and classification of the cancer bone by taking the CT scan images for testing. The bone cancer images obtained from the CT scan will be first viewed by the radiologist to make final decision, then the CT scan images are viewed by the physicians to diagnose the disease there is a chances of false prediction by the doctors because if the image is not clear and if there is some abstractions in the image will result in the false prediction. But, in this paper CT scan images will be taken, the oncologists diagnose the disease along with the computer results that are obtained from segmentation and classification, so this method gives the accurate result and there will be no chance of failure or false prediction by the doctors.

C. Problem Statement

The CT scan images are given has input. Further the steps involved in detecting the bone cancer are pre-processing the CT scan images are done by first removing the noise present over the image which effects the quality of the image, after pre-processing module, the image quality will be enhanced by contrast enhancement. Detection of edges of the bone cancer is done by canny edge detection, for the bone region segmentation k-means clustering will be adapted on the pre-processed image for the foreground and background separation. The GLCM feature extraction technique is taken in ordered to extract the affected area of bone.K-Nearest Neighbors Algorithm technique is applied for classification in order to classify whether it is cancerous or not and to find the stages of cancer.

D. Scope of Bone cancer Detection system

Scope of the bone cancer detection system are as follows:

- **Medical field:** The oncologists are the radiologists uses the bone cancer images that are obtained from the computer results which will help them to predict disease accurately.
- **Mining of minerals:** This method can be utilized in the field of mining to find or to classify the rare metals and minerals.
- **Industries:** This method can be used in the manufacturing industry to determine the quality of the products and to determine the strength of the manufactured products.

E. Objectives

The main objectives for the proposed method of the detection of bone cancer are:

- To Provide patient care that is appropriate, and effective for the treatment of cancer.
- To provide consistent approach towards detection of cancer.
- To define the treatment procedure for the patient who will undergo radiotherapy.
- To detect whether cancer has spread or not.
- To detect the bone cancer in early stages.
- To provide better accurate result by classifying the bone cancer by using the K-nearest neighbor algorithm

II. Review Of Literature :

Sinthia P and K. Sujatha [1], The K-implies computation and edge locating methodology will be used in this method to describe an unique approach to dealing with bone ailments. To recognise the edge, this approach used Sobel edge identification. Only the outskirt pixels are identified by the Sobel edge finder. The tumour area is identified using the K-Means clustering method. The most difficult step in K-Means clustering computation is determining the number of clusteres.

Mokhled S. Al-Tarawneh [2],To denoise the picture, this system used the Gabor filter. The Gabor filter produces the best results. Two segmentation techniques are used to divide the picture. Both calculations—marker-controlled watershed segmentation and threshold approach—are used. In

comparison to the thresholding technique, the marker-controlled segmentation strategy produces superior results.

FatmaTaher and NaoufelWerghi [3],Segmentation of two techniques was used in this procedure. To segment the picture, fluffy c-implies clustering and Hopfield neural are used. The PC Aided Diagnosis framework was built to differentiate malignant development in its early stages. This study tests two segmentation techniques on 1000 sample photos. When compared to fluffy clustering, HNN has a superior order.

Anita chaudhary [4],In order to reduce noise, this system uses a Gabor filter. Watershed segmentation is completed using two segmentation strategies: thresholding and marker-controlled watershed segmentation. To identify the tumour, features are removed. This study divides three features: zone, boundary, and roundness.

MaduriAvula [5], This approach specifies a way for detecting bone disease in MR images by exploiting Mean pixel power. The data The MR image is denoised, and the K-Means clustering calculation is used to segregate the tumour portion. The aggregate of pixel intensity is decided to the extracated tumour region to identify the power of mean pixel from the separated tumour portion and that with all out number of pixels. To discriminate between benign and malignant development, the intensity of the mean pixel is calculated. When the mean pixel intensity esteem exceeds the edge esteem, it is deemed cancerous.

NooshinHadavi and Md.JanNordin [6], This method makes use of a Gabor filter to eliminate noise from images. To segment the image, area developing calculations are used. To distinguish cancerous development, certain aspects are deleted from the split image and attached to the new computation cell automata.

Kishor Kumar Reddy [7], It outlines a technique for measuring the size of tumours and malignant bone growth systems using local development calculations. By applying locale developing computation, this process divided the region of interest. The number of pixels in the extricated tumour component determines the tumour measurement. The sickness arrangement is recognised based on the absolute pixel esteem. The picture is used to determine the seed point, which is difficult to pick accurately.

III. METHODOLOGY

High level design will give out an overview of the system flow. However, this will give more information for understanding the logic to the users. The system high level design deals with the connection between the different modules and its functionality. Here it will provide the knowledge regarding the system design and architecture. Following are the issues that will be seen in this part which are the basic components for designing.

A. Design considerations

The design considerations of "Bone cancer detection using k-means segmentation and KNN classification" are:

- The algorithm can accept the image file of various extensions such as .dicom ,.jpeg, .png, .bmp, .tiff, .gif type of images either color or gray.
- Resizing of the image or the image enhancement can be done by doubling the original image, and also it can be done by decreasing or increasing the blurred image, which will increase the quality of the image. Pre-processing technique is applied.
- Edge detection is done in order to observe all the boundaries of the bones to detect the bone cancer in which part it is affected by using the technique canny edge detector.
- K-means segmentation technique will be used for the segmentation of bone cancer for the foreground and backward separation of the bone.
- Feature extraction will be applied for the extraction of the texture of bone in order to see how much the cancer is spread.
- KNN classification is applied in order to check whether it is cancerous or non-cancerous, and if it is cancerous it will determine at which stage the bone cancer is present.

B. Architectural diagram

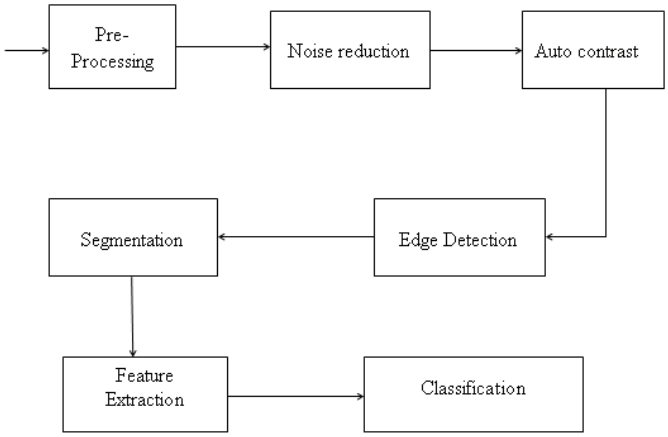

Fig. 1. Architectural diagram of bone cancer

Figure 1 illustrates the Architectural diagram of bone cancer detection. The architecture system diagram will describes the overall representation of entire bone cancer stages detection system. This model illustrates characteristics in each module and its sub-module associated with the main system. It explains the flow of the overall system working, the levels in which the data flow, the modules and the processes which should be an approach to determine the bone cancer. Some technical terms that would be understood by the systems are used by the high level design. CT scan image are accepted in this case. If there is noise present over the original image then by the use median filtering Median value will be calculated to the each channel and it is combined for the noise removal, if there is no noise present then the noise is added it may be the noise with salt and pepper, to remove this type of noises median filter method is used and Gaussian filter will be used for the removal of Gaussian noises. Gray scale images is taken

as input, then apply Gaussian filtering which sharpens the image or the image get enhanced, that is the image gets auto enhanced. Thus, the edge detected image is obtained. Gaussian noise, Poisson noise and Exponential noises etc are the various types of noises. Then edge is detected with the canny edge detector which is for extraction of the boundaries on the bone CT scan image. To segment bone cancer image k-means clustering segmentation is considered to differentiate actual or the normal bone with the cancerous bone. The bone texture will be extracted to examine the correct prediction of the bone cancer structure by using GLCM feature extraction technique. KNN classifier will be used in the classification of cancerous or non-cancerous image and to determine whether it is a normal and a cancerous bone that is to detect at which stage the cancer is present that is whether it is in the stag1, stage 2 and stage 3, this KNN classifier gives the accurate result with good accuracy.

C. Flow chart for Pre-processing module

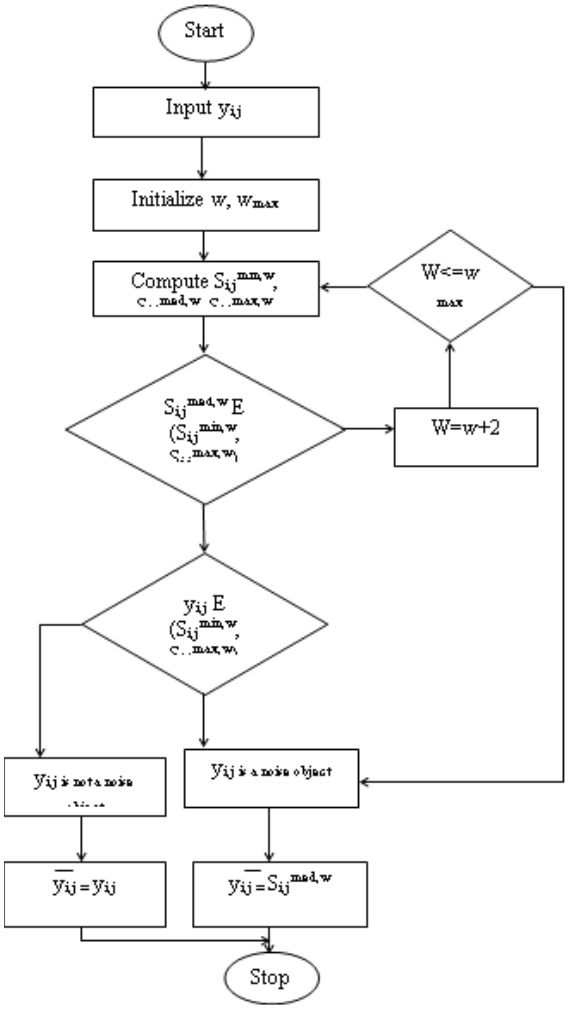

Fig. 2. Flowchart for pre-processing module

Figure 2 illustrates the flowchart for pre-processing module in bone cancer detection. First the CT scan is given as input, resizing of image is taken by doubling method. If the given input image has noise, then the noise will be removed by using Median filtering.Median values are calculated for each channels of the image and it is combined. This median filtering technique process, usually used for removal noises in images or in a signal. If there is no noise in the image the noise is added else the next further steps can be computed, if noise is present then it should be removed by using filtering approaches to the particular types of noises, in this method if there is an this type of noises is seen, then the median filtering noise removal approach is utilized for this type of noise otherwise if there is Gaussian noise exists in the bone image, then the Gaussian filter approach will be used for removing the noise. Median filter which are less sensitive than compared to the extreme values, this is a better approach to remove the noises because it will not affect the sharpness of the image and also it preserves the edges where first it will read the original CT scan image, next it will add the noises to the image, then by using the average filter for the pixel values it will filter the noise image, then it uses the median filtering method to filter the image by calculating the median value by splitting it into a 3x3 sized blocks of images because to get rid of noise efficiently in all the parts of the image where it will removes the noise affected pixels and obtain the clear noise removed image.

This noise in the image is produced if the image not captured properly in the case of CT scan image if the scanning of the patient is not taken properly due to some movements there will be a noise in the image or if the scanner is not proper also there will be a unwanted distortions in the original image which cannot be used for the further processes because it does not produce the proper result so it is important to remove the noise if there is a noise else not necessary, where in the original process if the CT scan of the patient is not proper again he or she as to go for scanning which will be long procedure, so with this noise removal technique it can preserve the time taken. The errors in the image selection process will change the color information of the image where the changed pixel values will not match the original object. This noise reduction technique is an important pre-processing step that is used to improve the results for the next methods, because if the image contains noise it degrades the quality by showing the different intensity values for the true pixel values for original image it will not produce the correct results for the next processes.

D. Flow chart for Edge detection

Figure 3 illustrates the flowchart of bone cancer detection using k-means clustering and KNN classification for edge detection. First the CT scan is given as input and the resizing of image is taken by doubling method. If the given input image has noise, then the noise is removed by using median filtering. After the noise removal edges will be detected.

The canny edge detector process is divided into five steps:

- First the Gaussian filtering is applied which is used to smoothened the image once the noises are removed from the image.
- Then, gradients intensity of the image is found.
- Non-maximum suppression method is applied in order to get rid of background and foreground response to edge detection.

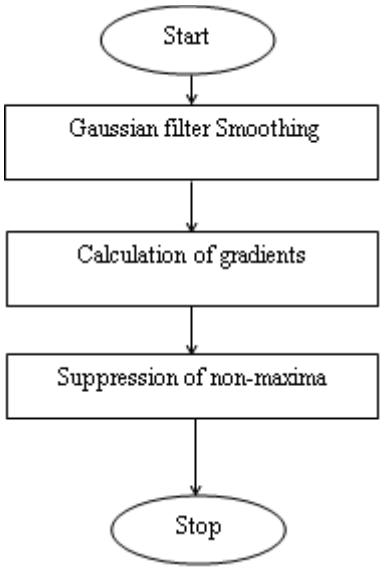

Fig. 3. Flow chart for Edge detection

- Double threshold method is applied for determining the potential edges.
- Then, finally the edges are detected by suppressing all the other edges which are weak that is not having or connected to the strong edges that is tracking the edges by using hysteresis.

IV. Implementation

A. Pseudo code for main module

The main module consisting of all the modules such as Pre-processing module, noise removal, edge detection, segmentation module, feature extraction module and classification.

Algorithm:

//Input: Bone CT scan images

//Output: Classification result

Step 1: Read the CT images

Step 2: image pre-processing steps include noise removal, contrast enhancement.

Step 3: Removing the noises in the image by using median filtering and Gaussian filtering.

Step 4: Edge detection of bone cancer is done.

Step 5: By the utilization ofK-means clustering technique and threshold value 255(tumour intensity greater than 255), bone cancer segmentation is done.

Step 6: Cancerous bone tumours segmentation is done.

Step 7: The features extracted from the output image of cancer nodule segmentation for the classification.

Step 8: The trained data is used to classifier to provide classified result.

Step 9: Testing feature such as extracted features are given for the classification.

Step 10: The classifier, classifies the result either it is a cancerous or non cancerous.

V. Results and Discussion

A. Bone Cancer Detection System Home Page

The Figure 4 shows the home page of the bone cancer detection system home page. This page include the button to load bone CT scan data, add noise, noise removal by filtering techniques, segmentation, feature extraction, classification such as classes 1, 2 3 and 4 is used for bone cancer CT scan to see whether it is a normal image represented by class1, and to see whether it is a cancerous image if it is with stage one which is represented by class 2, the stage two image is represented by class 3 and if the bone cancer is in the third stage then it is represented by class 4 then clear all and exit. The CT scan image is loaded using the load bone CT data button.

Fig. 4. Bone cancer Detection System Home Page

Here select noise button is used to select the type of noise to be added, select filter button is used to select type of filtering technique which will be useful in removing the noise, auto contrast button is used for contrasting the original image selected, edge button is used to detect the edges, select segmentation button is used to select the segmentation technique to segments the bone region, bone tumours region and cancerous tumour from the bone. Next, by clicking on feature extraction button it shows the features of cancerous tumour presented in the bone. Final button is the select classifier button, it classifies the cancer types whether it is normal or cancerous based on feature extraction, if it is cancerous then it classifies at which stage the cancer is present.

B. Input Bone CT scan image

Figure 5 illustrates the CT scan image of bone. To pre-process, initially the data must be selected. By clicking on load bone CT scan button, the CT scan image is selected. The bone cancer image data is collected from the bone dataset option in which the database is stored in the folder.

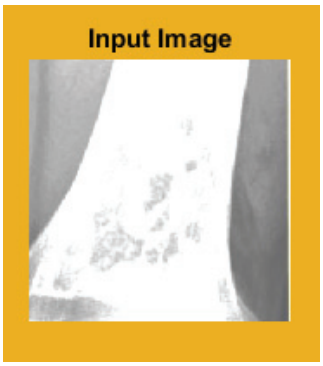

Fig. 5. Input CT scan image of bone

C. Added salt pepper noise for the bone CT scan image

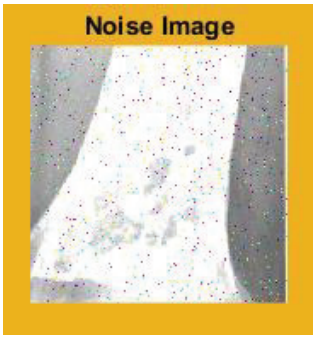

Fig. 6. Added salt pepper noise for the bone CT scan image

Figure 6 shows the salt pepper noise added image. After that the image is selected from the bone dataset. Only when if noise is not present at the bone cancer, then this type of noise will be added CT scan image, otherwise it can be processed with the further steps.

D. Median Filtered image

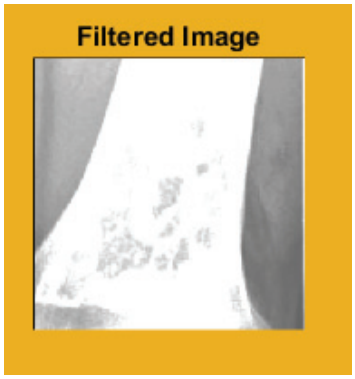

Fig. 7. Median filtered image

Figure 7 shows the filtered image, after the salt or pepper noises addition to CT scan image or if the noise exists over the original bone CT scan image then this type of noises can be efficiently get rid by utilizing median filtering technique, by calculating the median value by arranging 3x3 matrix pixels in the ascending and descending order, then the noise removed image is obtained as output by using median filtering.This noise in the image is produced if the image is not captured properly in the case of CT scan image if the scanning of the patient is not taken properly due to some movements there will be a noise in the image so that there will be a unwanted distortions in the original image which cannot be used for the further processes because it does not produce the proper result so it is important to remove the noise if there is a noise else not necessary.

E. Contrast Enhancement

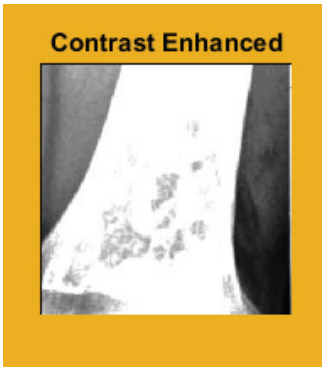

Fig. 8. Contrast Enhancement of CT scan image

Figure 8 shows contrast enhancement of the CT scan, after the removal of noise with the median filtering, then the next technique is to enhance the CT scan image by using auto contrast to increase the quality of image for the correct prediction of the cancer.

- Edge detection

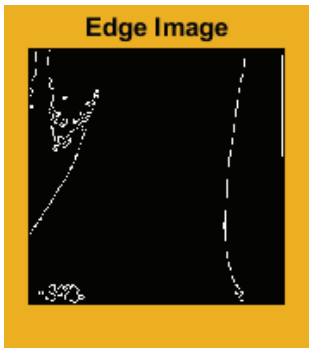

Fig. 9. Edge detection of CT scan image

Figure 9 shows the edge detection in the bones image, after contrast enhancement of the CT scan image, edges are detected for the CT scan image by using edge detector that is canny in order to extract the boundaries of the bone.

F. Segmentation of the bone cancer region

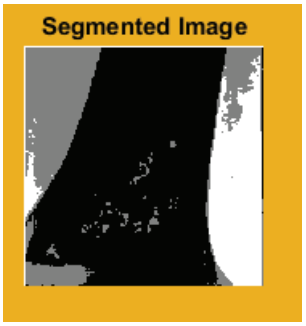

Fig. 10. Segmentation of the bone cancer image

Figure 10 shows the segmentation of the bone cancer CT scan image, after the edge of the bone are detected, then segmented of bone cancer region will be obtained by applying the K-means segmentation technique algorithm which segments the similar and dissimilar characters from the bone region in the form of clusters, in which foreground and background separation is done compare to the other segmentation techniques this k-means clustering gives the better result to differentiate the bone region between the normal and the cancerous bone image which will be useful for the feature extraction and classifier to easily predict whether it is cancerous or non-cancerous.

G. *Comparison graph of KNN and SVM classifier*

Figure 11 illustrates the comparison graph for KNN and SVM classifier in bone cancer analysis system. In the above comparison graph, X-axis will represents the SVM and KNN classifiers. Y-axis will represents the percentage of classifiers accuracy for detecting the bone cancer. In the above graph blue line indicates SVM classifier and green line indicates the KNN classifier.

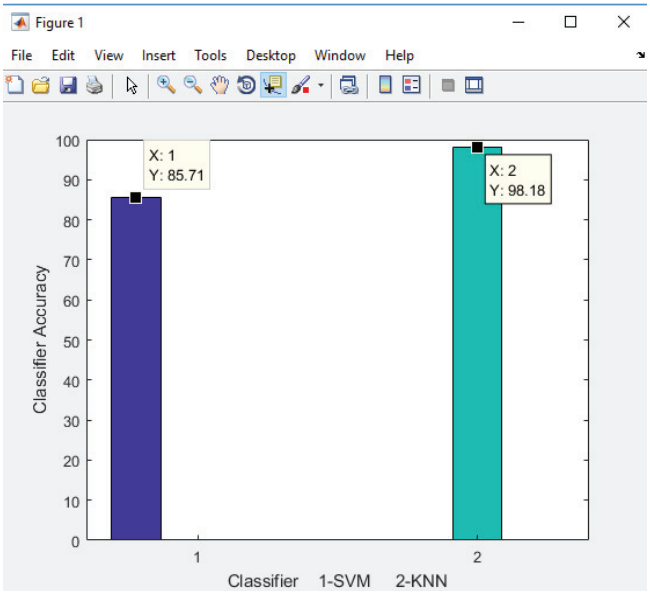

Fig. 11. Comparison graph for KNN and SVM Classifier

In the feature extraction phase numerical data will be extracted, then it will be passed as a input parameter to k-nearest neighbor classifier. Test data are termed as the extracted features. K-nearest neighbour is used to detect the bone cancer accurately whether the cancer is at the 1st stage, 2nd stage or 3rd stage else normal. This KNN classifier gives the good accuracy of 98.18% for detecting the bone cancer where the SVM fails by giving the accuracy of 85.71%, so the KNN classifier is the best method for detecting the bone cancer.

VI. Conclusion

Bone cancer in human beings is the most deadliest disease where the symptoms of this disease will not notices in the earlier stages. To prevent the death rate, it is very important for detecting the bone cancer in the early stages and needs exact determination for the people who are suffering from this disease. The bone cancer comes from other cancerous diseases like breast, bone cancer etc., which will widely spread to the bone. The proposed method will utilizes computed Tomography for the investigation of bone cancer stages growth in the CT scan images.

In this paper, KNN classification technique is used for the bone cancerous tumour detection in the images of CT scan. In pre-processing method noise removal, auto contrast and edge detection are performed, K- means segmentation technique will be used for the segmentation of the bone region. GLCM approach is taken for the feature extraction. K- Nearest Neighbor classifier will be used as a classification technique, for different kinds of images this method has been tested, and proven that KNN classifier has achieved 98.18% accuracy as shown in the snapshot.

VII. Future Enhancement

Proposed work describes the future work about using medical images for the bone cancer detection and its good performance. In future enhancement, different types of bone cancer can be classified and can take the few more parameters from the segmentation of cancerous nodule for the better classification.

References

[1] Sinthia P and K. Sujatha, "A novel approach to detect the bone cancer using K-means algorithm and edge detection method", ARPN Journal of Engineering and applied science,11(13), July 2020.

[2] Mokhled S. Al-tarawneh, "Lung cancer detection using image processing techniques", Leonardo electronic journal of practices and technologies, 20, 147-158, 2020.

[3] FatmaTaher and NaoufelWerghi."Lung cancer detection by using Artificial Neural Network and Fuzzy Clustering Methods", Americal Journal of Biomedical Engineering, 2(3), 136-142, 2019.

[4] Anita chaudhary, Sonitsukhrajsingh, "Lung cancer detection on CT images by using image processing", International conference on computing sciences, 2019.

[5] MaduriAvula, Narasimha Prasad Lakkakula, Murali Prasad raja, "Bone cancer detection from MRI scan imagery using Mean Pixel Intensity, Asia modeling symposium,2018

[6] NooshinHadavi, Md.JanNordin, Ali Shojaeipour, "Lung cancer diagnosis using CT-scan images based on cellular learning automata", IEEE, 2017.

[7] Kishor Kumar Reddy, Anisha P R, Raju G V S, "A novel approach for detecting the tumor size and bone cancer stage using region growing algorithm", International Conference on Computational Intelligence and Communication Networks, 2015.

[8] AbdulmuhssinBinhssan, "Enchondromatumor Detection", International journal of advanced research in computer and communication Engineering, 4(6), june 2015.

[9] Md. BadrulAlamMiah and Mohammad Abu Yousuf, "Detection of lung cancer from CT image using Image Processing and Neural Network", International conference on Electrical engineering and Information & communication Technology, 2016.

[10] EzhilE.Nithila, S.S.Kumar, "Automatic detection of solitary pulmonary nodules using swarm intelligence optimized neural networks on CT images", Engineering Science and Technology, an international journal, 2016.