An Approach for the Detection and Classification of Tumor Cells from Bone MRI Using Wavelet Transform and KNN Classifier

Conference Paper · December 2018

DOI: 10.1109/CIET.2018.8660908

3 authors:

Eftekhar Hossain
Chittagong University of Engineering & Technology

Md. Farhad Hossain
Chittagong University of Engineering & Technology

Mohammad Rahaman
Bangladesh Agricultural University

An Approach for the Detection and Classification of Tumor Cells from Bone MRI Using Wavelet Transform and KNN Classifier

[1]Eftekhar Hossain, [2]Md. Farhad Hossain and [3]Mohammad Anisur Rahaman
[1,2,3]*Dept. of Electronics and Telecommunication Engineering*
[1,2,3]*Chittagong University of Engineering and Technology*
Chittagong, Bangladesh
[1]eftekhar.13ete@gmail.com, [2]farhad.hridoy13@gmail.com and [3]anis00024@yahoo.com

***Abstract*—Bone cancer is one of the most dangerous and leading cause of early death around the globe. Therefore, early detection and classification of the bone tumor have become needed to cure the patient. This study uses a wavelet-based segmentation method for the detection of the bone tumor. The segmented bone tumor part is further processed for the classification purpose. In this study k-nearest neighbour (KNN) classifier is employed for the classification of bone tumor into benign and malignant class. A number of images are collected and gray level co-occurrence matrix (GLCM) features are extracted from these images for the creation of a learned classifier model. The obtained performance of the classification result exhibit that the KNN classifier provides 92.50% accuracy in bone tumor classification.**

***Index Terms*—Bone cancer, MRI, ADF, Feature extraction, Wavelet Transform and KNN classifier.**

I. Introduction

Bone cancer is considered as a multifarious disorder which occurs due to diverse genetic and physiological factors. It produces the uncontrolled growth of the cell making demonic bone tumors and invade to the adjacent parts of the body. Different types of bone tumor are detected in the human body. Bone cancers are also called sarcomas [1]. Most of the bone tumor is classified as benign or malignant. Benign tumors are begin in bone tissue and if the tumor cell moves to the other parts of the body through the bloodstream and lymph vessels than it is called malignant bone tumor [2]. In the year 2014, an estimation of bone tumor affected patient was provided by the American cancer society which showed that about 3020 new cases of bone tumor have been diagnosed and 1460 deaths are expected from this patient due to bone tumors. Magnetic resonance imaging (MRI) or computed tomography (CT) is used for the scanning of the bone anatomy.

Image segmentation always plays a significant part in cancer diagnosis. The actual meaning of segmentation is the split of an image into several regions and then extract the meaningful information from these regions [3]. There are a number of segmentation techniques have been applied to the MR images. For medical image analysis edge detection, thresholding-based, region-based, and clustering-based segmentation techniques are the most widely used methods. All the segmentation techniques have its own advantages and disadvantages and thus, it totally depends on the choice of the user. Besides, the segmentation and detection of bone tumor, it is necessary to classify the tumor type after detection so that it would be helpful for the specialists to provide the correct direction to the patient for early incarnation.

There have various classification methods among them four major classifiers are widely used for classification problems and these are SVM (support vector machine), KNN (K-nearest neighbour), and ANFIS (adaptive neuro-fuzzy inference system) classifier. SVM or support vector machine is a discriminative classifier which classifies by a separating hyperplane and the optimal hyperplane is found by a kernel [4]. KNN is a classifier which classifies based on feature similarity. It is a non-parametric classification technique [5]. ANFIS is a neuro-fuzzy based classifier that means it is an aggregation of the neural network and fuzzy logic in such a way that the parameters of the fuzzy systems are determined by the help neural network [6]. This paper, approaches a method for bone tumor detection using wavelet transform and also perform the classification of the benign and malignant bone tumor using k-nearest neighbor classifier.

The rest of the portion of the paper is structured in the following way: Section II provides a summarization of the previous work that was completed in the fields of bone tumor detection. Section III gives a brief description of the methodology that has been proposed for the bone tumor detection and classification. The results and discussion of the applied method have been disclosed in Section IV. Finally, Section V puts an end with conclusions and future enhancements of the work.

II. Related Work

Kishore Kumar Reddy C et al. proposed an approach for detecting the bone tumor size and stages of detected cancer. A seeded region growing algorithm was used for the segmentation of bone MR image. The algorithm performance depends on appropriate seed point selection from where the regions are started growing to nearby points on the basis of a similar criterion. After segmenting the tumor part an equation

978-1-5386-5229-9/18/$31.00 ©2018 IEEE

was used to calculate the tumor area and based on this area the stage of the bone cancer was determined [7]. Krupali D. Mistry et al. proposed an integrated method for the detection of bone cancer from MR images where the tumor extraction was done by using k-means clustering coupled with fuzzy C-means algorithm. After the segmentation, for the bone tumor identification, they use mean pixel intensity where it was used only to detect the Enchondroma type of bone cancer. The experiment was done on 31 MR images of enchondroma where the hybrid segmentation method achieved moderate accuracy and takes more processing time as compared with FCM [8]. Madhuri Avula et al. introduced an approach that detects the bone tumor from the MR images. For segmenting the image, k-means clustering algorithm was adopted and in order to detect cancer or no cancer after segmenting the tumor part from the image, they used a thresholding technique called mean pixel intensity. They perform experimentation on 400 bone MR images where they put a conclusion by showing that the mean pixel intensity values for the identification of cancer or no cancer lies in between 234 and 250. In this method, they also achieved good accuracy for bone cancer detection [9].

III. Proposed Methodology

The proposed method for the detection and classification of bone tumor is divided into two phases: training phase and testing phase. Each phase includes filtering, segmentation and feature extraction. Initially, the bone MR images are fed to the system for pre-processing such as grayscale conversion and filtering. Anisotropic diffusion filtering has been used for removing the noises from the image. Then discrete 2-D wavelet transform has been applied on the image which gives four frequency portions of the image among them three image consists high-frequency components and one image consists low frequency-component which is referred as approximate coefficients of the transformed image. Thresholding technique is applied to the approximate coefficients of the wavelet transformed image and then by calculating the highest dense pixel area the tumor part is extracted from the binary image. For the classification, four texture feature and four statistical features are extracted from the segmented image. In the training phase, this 8 features of the 180 bone MR images are used for the training of the k- nearest neighbour classifier. Different distance vectors and five nearest neighbours are used for the performance evaluation of the classifier model. In the testing phase, a new bone MR image is selected which is not been used in training. After completing all processing steps the extracted features of the new segmented bone MR image is fed to the KNN model which is created in the training phase. The KNN trained model measures the Euclidean distance from each of the new features and tries to find the nearest neighbour of the new features to classify these features into benign or malignant bone tumor feature. Total 220 bone MR images are used for training and testing of the KNN classifier. Fig. 1 shows the flow diagram of the proposed method.

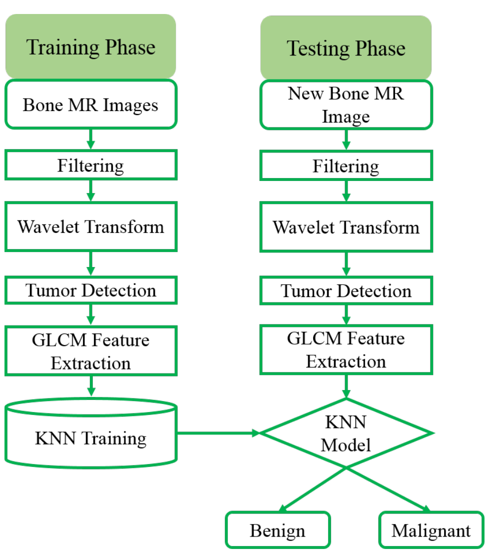


Fig. 1. Flow diagram of the proposed methodology

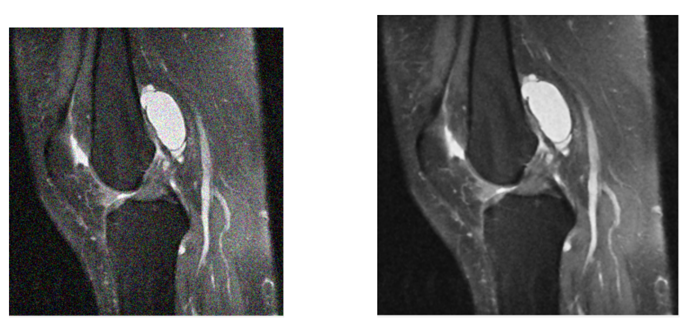
A. Bone Cancer Detection Using Wavelet Transform

1) Acquisition and Filtering:

The first step of the bone tumor detection is the database collection of bone MR images. In this paper, the taken database consists of a number of benign and malignant bone tumor images. These images are collected from some orthopaedics (www.orthopaedicsone.com) and tumor (www.tumorlibrary.com) website. In the next step, the MR images are pre-processed for the detection and classification purpose. Filtering is one of them and in this paper, an anisotropic diffusion filter (ADF) are used for sharpening the edge of the bone MR image [10]. Before filtering, the bone MR image is converted to a grayscale image to reduce the computational cost. The advantage of ADF over the conventional filter is that it blurs the image by removing noises and conserve the edges. A sample bone MR image and its filtering output are shown in Fig. 2.

2) Wavelet Transform:

The Wavelet transform is a new integrated mathematical and signals processing technique which is used for complex problem-solving. It has excellent analyzing performance and



(a) (b)

Fig. 2. Filtering Output: (a) MR image of a benign bone tumor (b) Filtered bone MR image

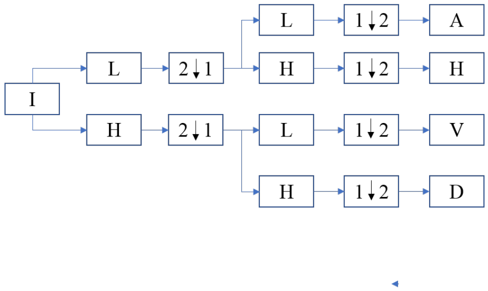
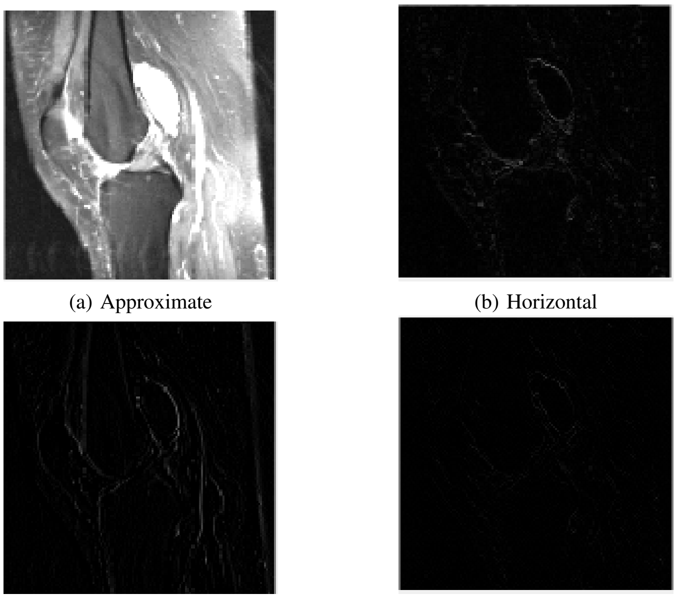


Fig. 3. 2D wavelet decomposition steps



(a) Approximate (b) Horizontal

(c) Vertical (d) Diagonal

Fig. 4. Wavelet transform output on MR image

multi-resolution characteristics, which makes it worthy for image processing. For localization and multi-resolution analysis of a signal, the wavelet transform can be applied. Moreover, it is readily applicable to all complicated images for accurate edge detection [11]. For a 2-D image signal $f(x, y)$, the wavelet decomposition algorithm can be shown in Fig. 3.

where, L-low-pass filter, H-high-pass filter, 2 $\downarrow$ 1-down sample columns, 1 $\downarrow$ 2-down samples row, I-original image, A-approximate coefficients or low-frequency components, H-horizontal detail coefficients, V-vertical detail coefficients and D-diagonal detail coefficients. The output of the wavelet transform on a bone MR image is shown in Fig. 4 where one low frequency and three high-frequency coefficients are illustrated. The high- frequency coefficients convey less information than low-frequency coefficients.

3) Tumor Detection:

For further improvement, thresholding technique is applied to the low-frequency coefficients image to find the binary image of the transformed image. Thresholding is a simple technique which is used to separate the foreground and background of an image and it is calculated using (1). A black and white or

(a) (b)

Fig. 5. Tumor detection step: (a) Binary image (b) Extracted tumor part

a binary image is the output of the thresholding. Thresholding creates a binary image from grey-levels by assigning all pixels below some threshold to zero and all pixels above the threshold to one. If g(i,j) is a threshold version of p(i,j) at some global threshold T, then the output of the thresholding operation is shown in Fig. 5. p(i,j) and g(i,j) are the pixels of the input and output image respectively [12].

$$g(i,j) = \begin{cases} 1 & \text{if } p(i,j) \geqslant T \\ 0 & \text{if } p(i,j) < T \end{cases} \tag{1}$$

After thresholding, a binary image is obtained which consists only the background and the objects of the image. The output of the thresholding on an MR image is shown in Fig. 5.

Finally, the tumor part is extracted from this binary image by using object labeling algorithm which task is to assign a unique label to each of the objects in the binary image. By determining the highest dense pixel area form this objects the desired tumor portion has been separated. The algorithm steps for bone tumor extraction is followed as:

1) Let there are m number of connected components $\{cc_0, cc_l, cc_2, , cc_{m-1}\}$ in the labeled image which is constructed from the binarized image. Let the pixel densities are $\{d_0, d_1, d_2, , d_{m-1}\}$ for labels $\{cc_0, cc_1, cc_2, , cc_{m-1}\}$ respectively.
2) Find $max\,\{d_0, d_1, d_2, , d_{m-1}\}$ from this labeled connected components. Let this maximum area is d_{max}.
3) Search for the connected component $I_k, 0 \leq k \leq (m-1)$ from the set $\{cc_0, cc_1, cc_2, , cc_{m-1}\}$ which have the pixel density or area value d_{max}.
4) Remove other components whose area values are less than the d_{max} and get the segmented tumor part.

The extracted tumor output from the threshold image is shown in Fig. 5.

B. Feature Extraction

Total 8 features are extracted from the segmented bone tumor part which are:

1) Mean, $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$
2) Variance, $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$

3) Standard Deviation, $s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$

4) Entropy = $-\sum_{i=1}^{n}\sum_{j=1}^{n} p(i,j)log(p(i,j))$

5) Energy = $\sum_{i=1}^{n}\sum_{j=1}^{n}(p(i,j))^2$

6) Contrast = $\sum_{i=1}^{n}\sum_{j=1}^{n}(i,j)^2 p(i,j)$

7) Homogeneity = $\sum_{i,j=1}^{n} \frac{p_{ij}}{1+(i-j)^2}$

8) Correlation = $\sum_{i,j=1}^{n} p_{ij}\frac{(i-\mu)(j-\mu)}{\sigma^2}$

C. Classification

Classification is the process of imposing a class on a new sample on the basis of learning attained by the classifier during training. Its task is to assign associate input pattern drawn by a vector to one of the diverse predefined patterns. In this paper, the classification of bone MR image into the benign or malignant tumor is done by using the KNN classifier.

1) K-nearest neighbor Classifier: The k-nearest neighbor classifier is the most used algorithms in machine learning. It is a non-parametric technique used for classification. It does not need any previous knowledge about the structure of the data in the training set. If the new training pattern is affixed to the subsisting training set then it doesn't require retraining. Before classifying a new element vector, a comparison should be finished with the training sample using distance metrics. Its k-nearest neighbors are then considered where the class that appears most among the neighbors is given to the element to be classified. A new element is classified on the basis of the neighbors are weighted by the distance measure. The appropriate working of the scheme depends on the proper selection of the appropriate parameter such as the k which represents the number of neighbors chosen to assign the class to the new element and the distance used [13].

2) KNN Classifier Phases & Rules: KNN classifier comprise of two phases. One is training phase where the bone MR images labels with their class (benign or malignant) and another one is testing phase where the bone image are unlabeled and algorithm generates the list of k nearest data point (training data point) to label the unlabeled point and classifies their class.

KNN rules are:

- The set of stored training and testing data.
- Use euclidian distance which can given by (2) as distance parameter to measure the distance between stored records & unknown record to classify.

$$d(p_i, q_j) = \sqrt{\sum_{r=1}^{n}(p_{ir} - q_{jr})^2} \tag{2}$$

- Find k nearest neighbors & use class labels of nearest neighbors to determine the class label of unknown record by taking majority vote

D. Performance Measure

Three performance parameters accuracy, sensitivity and specificity are used to evaluate the classification performance of the KNN classifier. The parameters are measured by determining the number of true false positives and negatives.

- True Positive (TP): Benign correctly identified as benign.
- True Negative (TN): Malignant correctly identified as malignant.
- False Positive (FP): Malignant incorrectly identified as benign.
- False Negative (FN): Benign incorrectly identified as malignant.

1) Sensitivity = $\frac{TP}{TP+FN} \times 100\%$

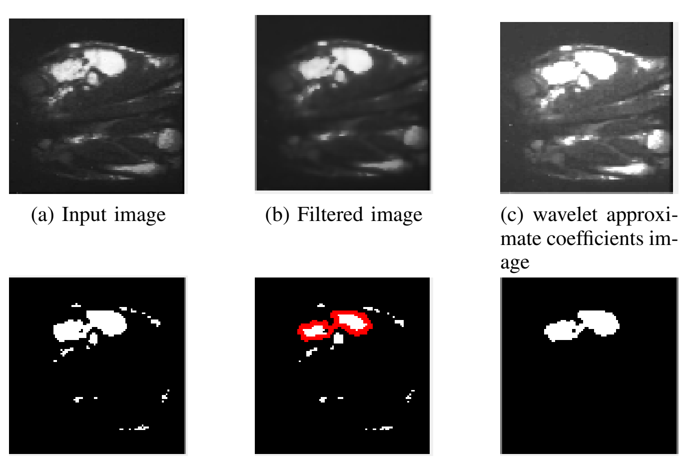
2) Specificity =$\frac{TN}{TN+FP} \times 100\%$

3) Accuracy =$\frac{TP+TN}{TP+TN+FP+FN} \times 100\%$

IV. Results & Discussion

The results of the bone tumor detection from MR images have been illustrated in Fig. 6 where the detected tumor part is indicated by a red outline. The simulation of the bone tumor detection and classification result has been done using MATLAB. The comparison of the proposed segmentation method with the existing segmentation methods has been shown in Fig. 7. Total 120 bone MR images have been used for measuring segmentation accuracy and it is done on the basis of the number of successfully segmented images over total images. The wavelet transforms based segmentation technique successfully segment 114 bone MR image and gives 95% accuracy. The Dice similarity index coefficients also have been used for the quantitative analysis of the segmentation algorithms. The dice similarity coefficient (DSC) is given by

$$DSC = 2(G \cap S)/(G + S)$$



(a) Input image (b) Filtered image (c) wavelet approximate coefficients image

(d) Binary image (e) Detected tumor (f) Tumor

Fig. 6. Steps for bone tumor detection using wavelet transform

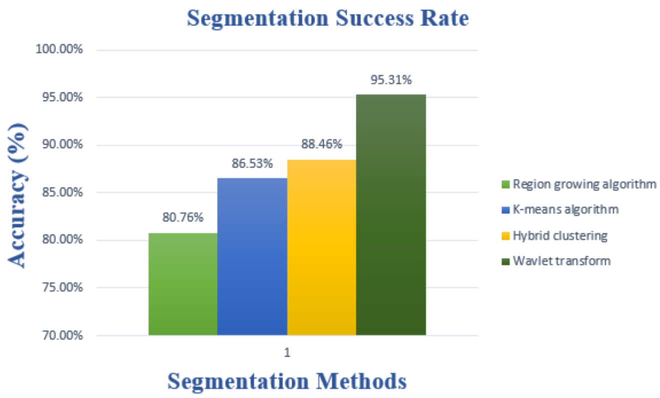
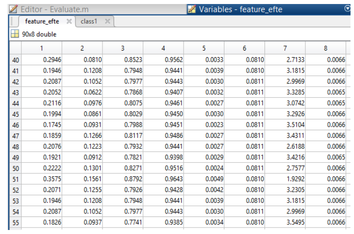


Fig. 7. Accuracy Comparison of the segmentation algorithms

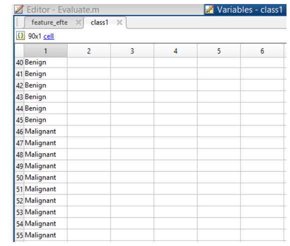
TABLE I
DICE SIMILARITY COEFFICIENT (DSC) RESULT OF BONE CANCER SEGMENTATION ALGORITHMS

Sl.no	Wavelet transform	Region growing	K-means clustering	Hybrid clustering
1	0.9883	0.8823	0.9123	0.9366
2	0.9528	0.8153	0.9021	0.9489
3	0.9727	0.8733	0.9213	0.9425
4	0.9620	0.8526	0.9183	0.9566
5	0.9451	0.8519	0.9367	0.9412
6	0.9782	0.8823	0.9257	0.9411
7	0.9658	0.8344	0.9519	0.9578
8	0.9336	0.8621	0.9536	0.9423
9	0.9561	0.8153	0.9357	0.9555
10	0.9588	0.9021	0.9293	0.9469


Editor - Evaluate.m | Variables - feature_efte
feature_efte | class1
90x8 double

	1	2	3	4	5	6	7	8
40	0.2946	0.0810	0.8523	0.9562	0.0033	0.0810	2.7133	0.0066
41	0.1946	0.1208	0.7948	0.9441	0.0039	0.0810	3.1815	0.0066
42	0.2087	0.1052	0.7977	0.9443	0.0030	0.0811	2.9969	0.0066
43	0.2052	0.0622	0.7868	0.9407	0.0032	0.0811	3.3285	0.0065
44	0.2116	0.0976	0.8075	0.9461	0.0027	0.0811	3.0742	0.0065
45	0.1994	0.0861	0.8029	0.9450	0.0030	0.0811	3.2926	0.0066
46	0.1745	0.0931	0.7988	0.9451	0.0023	0.0811	3.5104	0.0066
47	0.1859	0.1266	0.8117	0.9486	0.0027	0.0811	3.4311	0.0066
48	0.2076	0.1223	0.7932	0.9441	0.0027	0.0811	2.6188	0.0066
49	0.1921	0.0912	0.7821	0.9398	0.0029	0.0811	3.4216	0.0065
50	0.2222	0.1301	0.8271	0.9516	0.0024	0.0811	2.7577	0.0066
51	0.3575	0.1561	0.8792	0.9643	0.0049	0.0810	1.9292	0.0066
52	0.2071	0.1255	0.7926	0.9428	0.0042	0.0810	3.2305	0.0066
53	0.1946	0.1208	0.7948	0.9441	0.0039	0.0810	3.1815	0.0066
54	0.2087	0.1052	0.7977	0.9443	0.0030	0.0811	2.9969	0.0066
55	0.1826	0.0937	0.7741	0.9385	0.0034	0.0810	3.5495	0.0066

Fig. 8. Extracted features from the bone images


Editor - Evaluate.m | Variables - class1
feature_efte | class1
90x1 cell

	1	2	3	4	5	6
40	Benign					
41	Benign					
42	Benign					
43	Benign					
44	Benign					
45	Benign					
46	Malignant					
47	Malignant					
48	Malignant					
49	Malignant					
50	Malignant					
51	Malignant					
52	Malignant					
53	Malignant					
54	Malignant					
55	Malignant					

Fig. 9. Class labels of the features

TABLE II
PERFORMANCE ANALYSIS BONE CANCER CLASSIFICATION USING KNN CLASSIFIER

Parameters \ Distances	Euclidean	City block	Hamming
True Positive (TP)	18	16	15
True Negative (TN)	19	17	16
False Positive (FP)	1	3	4
False Negative (FN)	2	4	5
Accuracy (%)	92.50%	82.50%	77.50%
Sensitivity (%)	90.00%	80.00%	75.00%
Specificity (%)	95.00%	85.00%	80.00%

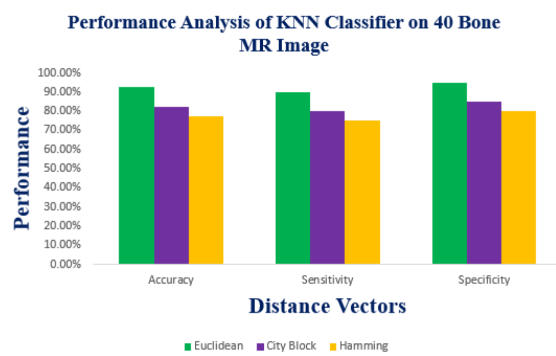


Fig. 10. Performance comparison of KNN classifier

where G is the ground truth image and S is the segmented image. The evaluated DSC value for the different algorithms on only 10 MR segmented images have been shown in Table. I and it can be said that the wavelet transform based segmentation technique provides 96.123% mean DSC value which is higher than all the other existing segmentation techniques.

For the classification of the bone tumor, total 220 MR image has been used where 180 images have been used for the training and 40 images for the testing of the KNN classifier. Eight features of each of the 180 bone MR images have been extracted and used for the training of the classifier and some samples of this features are shown in the Fig. 8 where each row represents the 8 features of one image.

The training dataset consists of 90 benign and 90 malignant bone tumor MR image. The features are sequenced column-wise as Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, and Variance. Fig. 9 shows the corresponding class labels of the features.

In this work, five nearest neighbors and different distance vectors have been used for training and testing. 40 bone MR image has been used for the performance analysis of the KNN classifier and three different distance vectors have been taken for the comparative measurement of the accuracy, sensitivity, and specificity. The classification performances and the comparative analysis plot of KNN classifier performance on bone tumor classification have been shown in Table II and in Fig. 10.

V. Conclusion and Future Work

This research work involves tumor detection from bone MRI using wavelet transform and classification of the detected tumor into benign or malignant type. The proposed approach highlighted the uses of KNN classifier for the classification of bone tumor. This method is considered a robust classification algorithm and it gives an excellent performance. The purpose of this work to develop a tool for discriminating malignant bone tumors from benign ones assisting decision making in clinical diagnosis. In future, the system can be developed to classify the not only benign or malignant tumor but also, the types of sarcomas like chondroma, Ewing sarcoma, and chondrosarcoma etc.

References

[1] V. T. Devita Jr, T. Lawrence, and S. A. Rosenberg, *Cancer: Principles & Practice of Oncology: Annual Advances in Oncology*. Lippincott Williams & Wilkins, 2012.

[2] H. D. Dorfman and B. Czerniak, "Bone cancers," *Cancer*, vol. 75, no. S1, pp. 203–210, 1995.

[3] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.

[4] L. Rong and S. Yuan, "Diagnosis of breast tumor using svm-knn classifier," in *Intelligent Systems (GCIS), 2010 Second WRI Global Congress on*, vol. 3. IEEE, 2010, pp. 95–97.

[5] C. Kalaiselvi, "Diagnosing of heart diseases using average k-nearest neighbor algorithm of data mining," in *Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on*. IEEE, 2016, pp. 3099–3103.

[6] A. Addeh, H. Demirel, and P. Zarbakhsh, "Early detection of breast cancer using optimized anfis and features selection," in *Computational Intelligence and Communication Networks (CICN), 2017 9th International Conference on*. IEEE, 2017, pp. 39–42.

[7] C. K. K. Reddy, P. Anisha, and G. Raju, "A novel approach for detecting the tumor size and bone cancer stage using region growing algorithm," in *Computational Intelligence and Communication Networks (CICN), 2015 International Conference on*. IEEE, 2015, pp. 228–233.

[8] K. D. Mistry and B. J. Talati, "Integrated approach for bone tumor detection from mri scan imagery," in *Signal and Information Processing (IConSIP), International Conference on*. IEEE, 2016, pp. 1–5.

[9] M. Avula, N. P. Lakkakula, and M. P. Raja, "Bone cancer detection from mri scan imagery using mean pixel intensity," in *Modelling Symposium (AMS), 2014 8th Asia*. IEEE, 2014, pp. 141–146.

[10] C. A. Palma, F. A. Cappabianco, J. S. Ide, and P. A. Miranda, "Anisotropic diffusion filtering operation and limitations-magnetic resonance imaging evaluation," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 3887–3892, 2014.

[11] C. Hu, L.-J. Jiang, and J. Bo, "Wavelet transform and morphology image segmentation algorism for blood cell," in *Industrial Electronics and Applications, 2009. ICIEA 2009. 4th IEEE Conference on*. IEEE, 2009, pp. 542–545.

[12] T. D. Murthy and G. Sadashivappa, "Brain tumor segmentation using thresholding, morphological operations and extraction of features of tumor," in *Advances in Electronics, Computers and Communications (ICAECC), 2014 International Conference on*. IEEE, 2014, pp. 1–6.

[13] J. S. Sánchez, R. A. Mollineda, and J. M. Sotoca, "An analysis of how training data complexity affects the nearest neighbor classifiers," *Pattern Analysis and Applications*, vol. 10, no. 3, pp. 189–201, 2007.