IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.134161

HEALTHCARE APPLICATION FOR CANCER DETECTION AND ANALYSIS USING MACHINE LEARNING AND IMAGE PROCESSING

Vivek Belagali[1], Rahul[2], Yash C[3], Rahul Bhattacharya[4]

Computer Science & Engineering, Dayananda Sagar University Bengaluru, India[1-4]

I. INTRODUCTION

The cancer is most dangerous disease throughout the globe. Clinically cancer is referred as malevolent neoplasm. It is a genetic disease which is caused by unregu- lated growth of cells. As early detection of such dangerous disease could reduce the number of death [1]. The symptom of cancer is unfettered cell growth which will lead to the development of malevolent tumor, which is also harmful for the nearby tissues [2—4]. This kind of tumor can further grow and hinders the circulatory sys- tem, digestive system and nervous system and also can produce hormones that lead to hinder the proper body functionality [5—7]. The unfettered growth cell necessar- ily not harmful unless and until it does not affected the structure of DNA. If this unfettered cell growth not repaired with in the early stage this will lead the DNA to die which will cause production of unnecessary new cells. Metastatis property of cancer is more serious. The metastatis process can be defined as movement of cancer cells from one part of body to another part. This process leads to produce tumor with tissue growth [8]. Initial symptom of cancer include abnormal bleeding, form- ing new lumps, prolonged cough, change in bowl movement, unexplained weight loss, etc. Tumors can be classified into two types cancerous and non-cancerous [9]. Surgical removal of benign (non-cancerous) tumor is easy and most of the benign tumor does not grow again. The malignant tumor(cancerous) contains the larger nuclei as compare to the benign tumor. Bone cancer clinically termed as sarcomas which initiate in muscles, fibrous tissue, bone, blood vessels and the other tissues of body. Some of common types of cancer are chondrosarcoma, osteosacoma, pleo- morphic sarcoma, ewings sarcoma, fibrosarcoma. In bone cancer tumor start devo- loping in the bone, affecting the movement and growth of bones. Enchondorma is a benign tumor of bone which starts growing at the cartilage. Specially Enchondroma found in small bones of the hands. Possible area for enchondroma is the bone of upper arm, shin bone and thigh bone [10,11].

On the level of its advancement bone cancer divided into different stages.

- Stage1-It shows the non-aggressive nature of tumor, the malignancy is not spread out of bone.
- Stage 2-Same as stage1, but it shows the aggressive nature of tumor.
- Stage 3-If tumor starts growing in multiple places
- Stage 4-If cancer starts spreading in other parts of body.

This chapter is organized in five sections. 1. Introduction 2. Machine learning techniques for cancer classification. 3. Machine learning techniques for bone cancer classification. 3.1. Challenges and future prospect. 4. Conclusion. 5. References.

Machine learning techniques for cancer classification

Different types of machine learning techniques used in healthcare. The biomedical imaging modalities include ultrasound, magnetic resonance imaging (MRI), Computed Tomography (CT), histology and microscopic images. Some of the targeted organs for biomedical imaging include liver, brain, prostate, lung and exami- nation of genetic associations. The combination of Medical Imaging and Machine Learning is an excellent reference for industrial scientist, researchers of medical imag- ing, clinician and students [12—14]. Fig. 17.1. Shows the different types of medical imaging modalities used for machine learning.

MRI sequence display different type of brightness for same structures. Ultrasound is more convenient than CT and MRI. In the case of bone, ultrasound do not give the good result as the waves of ultrasound do not transmit well through bone [15,16]. Fig. 17.2A represents the axial view of CT scan while Fig. 17.2B represents the top view of CT scan. In this case CT images of pelvic bones are used. These CT scan

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.134161

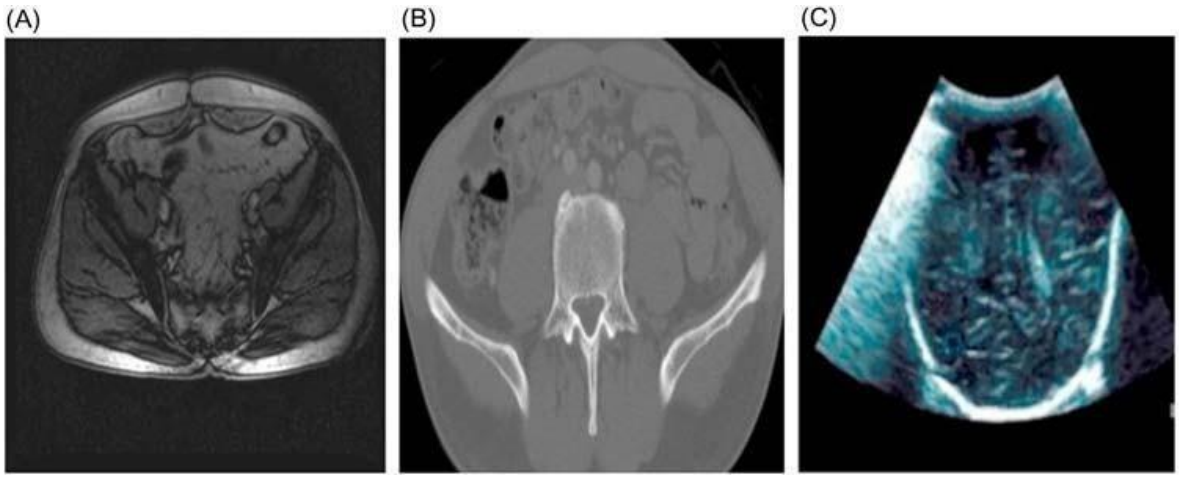

Figure 17.1 (A) An example of MRI (B) CT (C) ultra sound images.

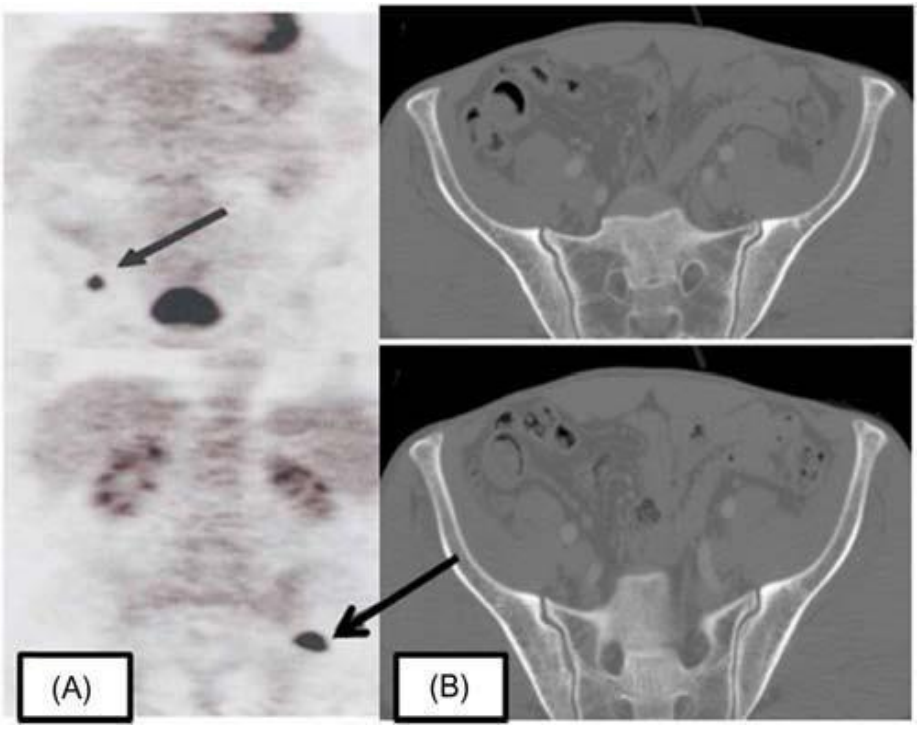

Figure 17.2 (A) Axial views of a CT scan (B) top view of CT scan.

images are used as input in Machine Learning Techniques for classifying the disease. With help of these images we can classify the abnormalities like cancer, osteoporosis, thinning of bones etc.
There are different types of classification techniques exist in machine learning.
Some of the classification techniques are as follows.

Decision tree algorithm

Decision tree algorithm is used for classification. In this algorithm divide and conquer method is applied for tree construction [17]. In this algorithm tree comprise of leaves and nodes, where leaves represented the class for checking the condition and nodes represented the value of attribute. The output is either true or false. The classification rules can be derived from the path from root node to the leaf node. For prediction of class of leaf rule is obtained by the traversing total number of nodes in the path [18]. Tree pruning can be obtained by removing unnecessary duplication of leaves.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.134161

Support vector machine

The support vector machine (SVM) works on learning system that uses the methodol- ogy of statistical learning for classification. SVM is defined as discriminative classifier that defined by the separating hyperplanes. Hyperplane is the optimal boundary in SVM. In labeled training dataset the output of algorithm is an optimal hyperplane. Considering the 2D space hyperplane act as line where each side represents a class. The neighboring vectors near the hyperplane are described as support vectors [19—23].

Random forests

Random Forest (RF) algorithm is one of the best algorithms for classification. RF is able for classifying large data with accuracy. It is a learning method in which number of decision trees are constructed at the time of training and outputs of the modal predicted by the individual trees. RF act as a tree predictors where every tree depends on the random vector values. The basic concept behind this is that a group of "weak learners" may come together to build a "strong learner" [8,10,24,25].

Evolutionary algorithms

Genetic Algorithm (GA) is stochastic and evolutionary method for getting optimal solutions for large and complex problem. GA involved the natural evolution as it generates the population for encoded candidate (called "chromosomes") through generations using genetic operations like crossover and mutation. At every generation, solutions are selected on the basis of fitness function, for generating offspring and creating the next generation. The initial population is generated randomly, and at every generation, candidate solution is compared and evaluated against an objective function for gaining a fitness function score. The objective function is used to measure the candidate accuracy over the training set [26,27].

Swarm intelligence

Swarm intelligence (SI) is one of the computational intelligence techniques which are used to solve complex problem. SI involves collective study of the individuals behav- ior of population interact with one another locally. Especially for biological systems nature often act as an inspiration. Simple rules are followed by agents and no central- ized control structure exists in order to predict the behavior of individual agents. The random iteration of certain degree between the agents provides an "intelligent" behavior which is then unknown to individual agents. Some of popular SI algorithms included Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC) and Ant Colony Optimization (ACO) [28—30]. Table 17.1 describes the limitation and benefits of different machine learning techniques for classification.

Machine learning techniques for bone cancer detection

Now a days Machine learning technique plays an important role in diagnosis and identifying the stages of bone cancer. Anindita Mishra et al. [8] used random forest technique for classification of bone cancer in CT images. They used different fea- tures like shade, autocorrelation, energy, variance and cluster prominence to train and test the random forest. Sinthia et al. [16] represented the approach of clustering techniques for bone abnormality classification. They used Fuzzy C-mean (FCM) and

Table 17.1 Benefits and limitations of different machine learning techniques.

Algorithm	Benefits	Limitations
Decision tree [17,18]	I. Easy to understand and effective training algorithm. II. Order of instance for training has no effect on training. III. Pruning techniques can deal the over fitting problem.	I. Mutually exclusive nature of classes. II. Dependency of decision tree on the selection of attribute. III. Error results in complex decision tree. IV. Missing attribute make it complex about the selection of branch for testing the attribute.
Support vector machine [19—23]	I. Overfitting problem is unlikely to be occur. II. Quadratic optimization problem reduced the Computational complexity. III. Easy to control the frequency of error and complexity.	I. Training is relatively slow compared to Decision Trees. II. Difficulty in determining optimal parameters of instance when the data which is used for training purpose is not linearly separable. III. Difficulty in understanding structure

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.134161

		of algorithm.
Random forest [8,10,24,25]	I. Fast and effective classification of instance. II. Robustness for irrelevant attributes. III. Can be used for both classification and regression.	I. Assumes that similar instance with similar attribute has similar classification. II. Assumes that all attribute have equal nature of relevancy. III. Complexity is high.
Genetic algorithm [26,27]	I. Simple algorithm, easy to implement. II. Can be used in feature classification and feature selection. III. Primarily used in optimization.	I. Computation or development of scoring function is non-trivial. II. Not the most efficient method to find some optima, tends to find local optima rather than global. III. Complications involved in the representation of training/output data.

K-Means clustering algorithm to find the tumor part in bone. They concluded that FCM is better than K-Means clustering as it gives better performance in classifica- tion. Rishav Kumar et al. [21] applied the approach of Computer Aided Diagnosis (CAD) to differentiate cancerous and non-cancerous bone lesions of spine with help of Support Vector Machine (SVM). For segmentation they used Active Contour Model (ACM). Gradient Vector Flow (GVF) was used to detect moving boundary segmentation. A. Asuntha et al. [31] applied Artificial Neural Network Algorithm for detection and classification of bone cancer. They used filtering and gray scale conversion for preprocessing. For segmentation they used thresholding and superpix- el segmentation. They extracted different feature like contrast, standard deviation, mean, correlation, entropy, skewness, smoothness, variance and root mean square error. These extracted features applied to train the data in Artificial Neural Network. High time complexity was observed in that approaches. Madhuri Avula et al. [32] used the approach of mean pixel intensity thresholding for detection of bone cancer. For segmentation they used region growing algorithm. For Identification of region of interest mean pixel intensity thresholding has been used. They used threshold value for classification of cancerous and non-cancerous image. Kishore Kumar et al.

[33] used the approach of mean pixel intensity to detect bone cancer from MRI images. To detect the stages of bone cancer they used mean intensity threshold value. Ambalkar et al. [34] used K- Means clustering with thresholding approach for classification of bone cancer. For preprocessing they used RGB to gray scale conver- sion and bilateral filtering.

CHALLENGES AND FUTURE PROSPECTS

Despite the reported Machine Learning success for cancer imaging, several hurdles and limitations should be overcome way before the clinical adoption. Due to increase demand for MRI and CT, large amount of data is generating by care provider. Different standards include the DICOM and Picture Archiving and Communication System (PACS) have ensured for fast access of data and retrieval. The large amount of data is needed for the data hungry method like deep neural network. The major chal- lenge in bone cancer detection with machine learning include how we can make best use of structure of medical images and specific property of medical data in training and designing of our model. Standardized benchmarking is very important in the medical domain, especially for modalities and anatomic sites. In addition, imaging is not act as an isolated measurement of disease. Molecular signature of cancer also plays an impor- tant role in diagnosis with imaging. Potential and power of Machine Learning is increasing day by day, but some direction in terms of clinical practice is still remaining. Machine Learning techniques can be used for solving bioinformatics problem as these techniques are efficient for complex and big biological data analysis. For medical image analysis, accuracy and predictive power of Machine Learning methodologies needs fur- ther improvement. Further advancement of neural network can be done to calculate size, location and stage of bone cancer.

CONCLUSION

Machine Learning approach proved to be successful in the area of medical image diag- nosis, disease prediction, and bone cancer diagnosis and in the field of assessment of risk. In this chapter we concluded that there are many scientific challenges which need to be addressed such as noise reduction, accuracy, stages for bone cancer detection and machine

DOI: 10.17148/IJARCCE.2024.134161

learning technique plays an important role in accurate diagnosis of bone cancer. Computer based system in health care environment for diagnosis of bone can- cer at early stage proved to be very effective, especially in the country like India where mortality ratio is high and doctor to patient ratio is very less. This scenario gives rise to the medical imaging dependencies on cancer diagnosis.

REFERENCES

[1]. A. Binhssan, Enchondroma tumor detection, Int. J. Adv. Res. Comput. Commun. Eng. 4 (6) (2015) 1—4.
[2]. R.S. Savage, Y. Yuan, Predicting Chemo insensitivity in breast cancer with omics/digital pathology data fusion, R. Soc. Open Sci. 3 (2) (2016) 140—501.
[3]. A. Madabhushi, G. Lee, Image analysis and machine learning in digital pathology: challenges and opportunities, J. Med. Image Anal. 33 (2) (2016) 170—175.
[4]. L. Xiang, Y. Qiao, D. Nie, L. An, Q. Wang, D. Shen, Deep auto-contex convolutional neural net- works for standard-dose PET image estimation from low-dose PET/MRI, J. Neurocomput. 267 (1) (2017) 1—5.
[5]. N.E. Bylund, M. Andersson, H. Knutsson, Interactive 3D filter design for ultrasound artifact reduc- tion, IEEE International Conference on Image Processing, Genova, Italy, 2005, pp. 725—728.
[6]. S. Hameed, M.A.H. Radi, M.T. Gaata, Medical image classification approach based on texture information, J. Entropy 1 (2) (2016) 01—06.
[7]. R.V. Santosh Singh, Y. Singh, An evaluation of features extraction from lung CT images for the classification stage of malignancy, Int. Organ. Sci. Res. (IOSR) J. Comput. Eng. 9 (2) (2016) 76—79.
[8]. A. Mishra, M.V. Suhas, Classification of benign and malignant bone lesions on CT images using random forest, IEEE International Conference on Recent Trends in Electronics Information Communication Technology, 2016, Bangalore, India, pp. 1807—1810.
[9]. R. Aishwariya, M. Kalaiselvi Geetha, M. Archana, Computer aided fracture detection of X-ray images, IOSR J. Comput. Eng. 1 (1) (2008) 44—51.
[10]. G. Chu, P. Ramakrishna, H. Kim, D. Morris, J. Goldin, M. Brown, Bone tumor segmentation on bone scans using information and random forest, Int. Innov. Res. J. Eng. Technol. (IIRJET) 17 (1) (2014) 601—608.
[11]. Cem M. Deniz, S. Xiang, Segmentation of proximal femur from MR image using deep convolu- tion neural network, IEEE Trans. Magn. Reson. Med. 2 (1) (2017) 1—26.
[12]. W. Heindel, R. Gubitz, V. Vieth, M. Weckesser, O. Schober, M. Schafers, The diagnostic imaging of bone matastases, Dtscharzteblint 1 (11) (2014) 741—748.
[13]. C. Kishor Kumar Reddy, P.R. Anisha, G.V.S. Raju, A novel approach for detecting the tumor size and bone cancer stage using region growing algorithm, IEEE International Conference on Computational Intelligence and Communication Networks, Jabalpur, India, 2015, pp. 228—233.
[14]. R.S. Ananda, T. Tessamma, Automatic segmentation framework for primary tumors from brain MRIs using morphological filtering techniques, IEEE 5th International Conference on Biomedical Engineering and Informatics, Chongqing, China, 2013, pp. 238—242.
[15]. A. Chaudhary, S. Sukhraj Singh, Lung cancer detection on CT images by using image processing, IEEE International Conference on Computing Sciences, Phagwara, India, 2012, pp. 142—146.
[16]. Sinthia, K. Sujatha, A novel approach to detect the bone cancer using K-means algorithm and edge detection method, ARPN J. Eng. Appl. Sci. 11 (13) (2016) 8002—8007.
[17]. V. Podgorelec, P. Kokol, B. Stiglic, I. Rozman, Decision trees: an overview and their use in medi- cine, J. Med. Syst. 26 (5) (2002) 445—463.
[18]. Mihir N. Dalwadi, D.N. Khandhar, Kinita H. Wandra, Automatic boundary detection and genera- tion of ROI for focal liver lesion ultrasound image using texture analysis, Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) 2 (7) (2013) 2369—2373.
[19]. A. Lahsasna, R.N. Ainon, R. Zainuddin, A. Bulgiba, Design of a fuzzy-based decision support system for coronary heart disease diagnosis, J. Med. Syst. 36 (5) (2012) 3293—3306.
[20]. K. Polat, S. Gune¸s, Breast cancer diagnosis using least square support vector machine, Digit. Signal Process. 17 (4) (2007) 694—701.
[21]. R. Kumar, M.V. Suhas, Classification of benign and malignant bone lesions on CT images using support vector machine: a comparison of Kernel functions, IEEE International Conference on Recent Trends in Electronics Information Communication Technology, Bangalore, India, 2016, pp. 821—824.
[22]. H.X. Liu, R.S. Zhang, F. Luan, X.J. Yao, M.C. Liu, Z.D. Hu, et al., Diagnosing breast cancer based on support vector machines, J. Chem. Inf. Comput. Sci. 43 (3) (2003) 900—907. Jun.
[23]. M.F. Akay, Support vector machine combined with feature selection for breast cancer diagnosis, Exp. Syst. Appl. 36 (2) (2009) 3240—3247.
[24]. C. Nguyen, Y. Wang, H.N. Nguyen, Random forest classifier combined with feature selection for breast

DOI: 10.17148/IJARCCE.2024.134161

cancer diagnosis and prognostic, J. Biomed. Sci. Eng. 06 (05) (2013) 551—560.

[25]. K.M. Cherry, S. Wang, E.B. Turkbey, R.M. Summers, Abdominal lymphadenopathy detection using random forest, Proceeding SPIE Medical Imaging, San Diego, California, 2014, pp. 3245—3268.

[26]. A. Das, M. Bhattacharya, A Study on Prognosis of Brain Tumors Using Fuzzy Logic and Genetic Algorithm Based Techniques, International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing, Shanghai, China, 2009, pp. 348—351.

[27]. G. Schaefer, T. Nakashima, Data mining of gene expression data by fuzzy and hybrid fuzzy meth- ods, IEEE Trans. Inf. Technol. Biomed. 14 (1) (2010) 23—29.

[28]. W. Zhao, C.E. Davis, Swarm intelligence based wavelet coefficient feature selection for mass spec- tral classification: an application to proteomics data, Anal. Chim. Acta 651 (1) (2009) 15—23.

[29]. P.-C. Chang, J.-J. Lin, C.-H. Liu, An attribute weight assignment and particle swarm optimization algorithm for medical database classifications, Comput. Methods Prog. Biomed. 107 (3) (2012) 382—392.

[30]. H.-L. Chen, B. Yang, G. Wang, S.-J. Wang, J. Liu, D.-Y. Liu, Support vector machine based diagnostic system for breast cancer using swarm intelligence, J. Med. Syst. 36 (4) (2011) 2505—2519.

[31]. A. Asuntha, A. Srinivasana, Bone cancer detection using artificial neural network, Int. J. Sci. Res. (IJSR) 17 (2) (2018) 56—63.

[32]. M. Avula, L. Narasimha Prasad, M.P. Raja, Bone cancer using region growing algorithm and mean pixel intensity thresholding, Int. J. Soft Comput. 13 (3) (2018) 69—80.

[33]. C. Kishor Kumar Reddy, P.R. Anisha, G.V.S. Raju, A novel approach for detecting the bone can- cer and its stage based on mean intensity and tumor size, IEEE International Conference on Researches in Applied Computer Science, Konya, Turkey, 2015, pp. 162—171.

[34]. S.S. Ambalkar, S.S. Thorat, Bone tumor detection from MRI Image using machine learning, Int. Res. J. Eng. Technol. (IRJET) 5 (5) (2018) 3561—3564. Jul.

FURTHER READING

[1]. G.W. Milligan, S.C. Soon, L.M. Sokol, The effect of cluster size, dimensionality, and the number of clus- ters on recovery of true cluster structure, IEEE Trans. Pattern Anal. Mach. Intell. 1 (2) (2004) 40—47.

[2]. K. Jalal Deen, R. Ganesan, An automated lung cancer detection from CT images based on using artificial neural network and fuzzy clustering methods, Int. J. Appl. Eng. Res. 9 (22) (2014) 17327—17343.

[3]. B.G. Patil, Cancer cells detection using digital image processing methods, Int. J. Latest Trends Eng. Technol. 3 (4) (2014) 45—49.

A. Verma, G. Khanna, A survey on digital image processing techniques for tumor detection, Indian J. Sci. Technol. 9 (4) (2016) 1—14.

[4]. S. Banik, M.R. Rangaraj, J.E. Desautels Leo, Detection of Architectural Distortion in Prior Mammograms of Interval-cancer Cases with Neural Networks, IEEE 31st Annual International Conference of the EMBS, Minneapolis, USA, 2009, pp. 6667—6670.

[5]. H. Singh, S. Verma, G.K. Marwah, The new approach for medical enhancement in texture classification and feature extraction of lung MRI images by using gabor filter with wavelet transform, Indian J. Sci. Technol. 8 (35) (2015) 1—7.