

IDENTIFICATION AND CLASSIFICATION OF CANCER CELLS USING CAPSULE NETWORK WITH PATHOLOGICAL IMAGES

Dr. A. Pasumpon Pandian,
Professor, Department of CSE,
KGISL Institute of Technology,
Coimbatore, India.
Email: pasumponpandian32@gmail.com

Abstract: Cancer is a deadly disease that is costing the lives of many people. Over 9.6 million death is reported in 2018 due to cancer. We propose an ideal methodology to identify and classify cancer cells using pathological images with the help of capsule network. Capsule network's capability to learn patterns based on previous iterations can be exploited for this purpose. This can help in identification of cancer at early stages and work at the root cause of the disease and walk towards completely shutting down the disease. Image processing is done along with fuzzification and further, it is handled with capsule network classifier and analysed.

Keywords: Image Processing, Capsule Network, Artificial Neural Network, Cancer detection, Fuzzy system

1. INTRODUCTION

Capsule Neural Network (CapsNet) is a type of Artificial Neural Network (ANN) that is used in machine learning systems to model hierarchical relationships [1]. The biological neural organization is mimicked with a set of neurons called capsules. These capsules individually activate for various properties like position, size and hue for any type of object. A set of neurons is collectively termed as a capsule. The output of a capsule is routed to capsules of the next later thereby creating a prediction pattern over the sequence of iterations [13].

The study of cause and effects of any disease or injury is termed as pathology. For managing digital information of diseases, an image-based data environment that procures information from digital slides enabled by computer technology is used. This system is termed as digital pathology [12]. Virtual microscopy, a system of converting glass slides into digital slides, uses digital pathology for viewing, managing, sharing and analysis of data on a computer. Digital pathology has flourished into a promising avenue of diagnostic medicine with the advent of whole-slide imaging technology. It offers faster, cheaper and better diagnosis, prognosis and prediction of cancer and other significant ailments.

Medical researchers have been finding ways to cure cancer since decades, but it takes time and money for the development of new treatments. Around 40% of successful treatment involves surgery and radiation. Scientists are also working on safe ways to identify the root cause of cancer and create safe ways for shutting it down before it grows and spreads rapidly [14]. Several automated systems are also developed to calculate cancer probability and diagnose using data relevant to risk factors and symptoms.

Recent advances in Artificial Intelligence has led to development of expert medical applications and systems. Capsule networks can easily identify even rotated images and require very less training data. Hence it is a best suited technology for processing the medical images like MRI [7] or tissue biopsy images [8]. The location and size of the tumour can be detected along with the information regarding the type of tumour. The recent technologies reduce the need for complex field microscopes and time consuming methods used traditionally for determining the disease.

2. EXISTING LITERATURE

Almuntashri et al. [5] proposes an innovative technique for automatic recognition as well as classification of Gleason grading system based images of prostate cancer biopsy. The wavelet transform features and fractal analysis domains are combined together by the algorithm. Before feature extraction, biopsy images are pre-processed using effective image processing algorithm that can analyse textural complexity in terms of RGB colour channels, edge as well as segmentation data.

Zhu et al. [3] used a lung cancer survival prediction framework by integrating pathological images along with generic data. The integration of this data is beneficial as pathological images and molecular profile reveals complementary information of tumour characteristics. Model-Based Background Correction (MBBC) [16] method is used for the processing of gene expression signatures. In order to fragment individual cells from the pathological images and extract image features, a robust cell recognition and segmentation technique is used. Based on the result of cell detection, efficient geometry and texture descriptors are used for extracting an extensive set of features.

Joshi et al. [4], detects tumour blocks or lesions using a computer based procedure. They use ANN for classifying the type of tumour in MRI images to identify Astrocytoma type of tumour. Other tumour detection techniques involves image processing techniques like image segmentation, histogram equalization, morphological operations, feature extraction and image enhancement. This procedure can classify the types of tumour effectively from brain images in diverse clinical settings.

Descombes et al. [6], used thresholding for segmentation since it assists in obtaining binarized image with grey level 1 for tumour representation and grey level 0 for background representation. Intensity threshold determines the segmentation. Each pixel in the image is compared with the intensity threshold and set to black or white based on the intensity. The process is given by the equation

$$T = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} e_{i,j} * M_{i,j}}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} M_{i,j}}$$

Mobiny et al. [9] proposed a dynamic routing mechanism to improve the computational efficiency thereby speeding up the capsule network by three times. They also proposed a convolution decoder for the purpose of low reconstruction error and improved classification accuracy. The network architecture is represented as a three dimensional version. Jezer et al. [22] introduced a collective neural network and decision tree model that is used for prediction of cancer relapse. This helps patients after cancer surgery through diagnosis for planning continued treatment.

3. PROPOSED WORK

The pathological image is obtained via MRI or tissue biopsy in a digital format and fed to the image pre-processor. The pre-processed image is segmented and further enhanced. Based on the enhanced signals, cancer identification is done. Further, feature extraction is done. Based on the feature of known samples obtained from a knowledge base, fuzzification is done. Unknown samples are also transferred for fuzzification so as to enable learning in the capsule network classifier. Further, capsule network classifier is used to identify the type and stage of the cancer cells. This classified data is further sent for analysis.

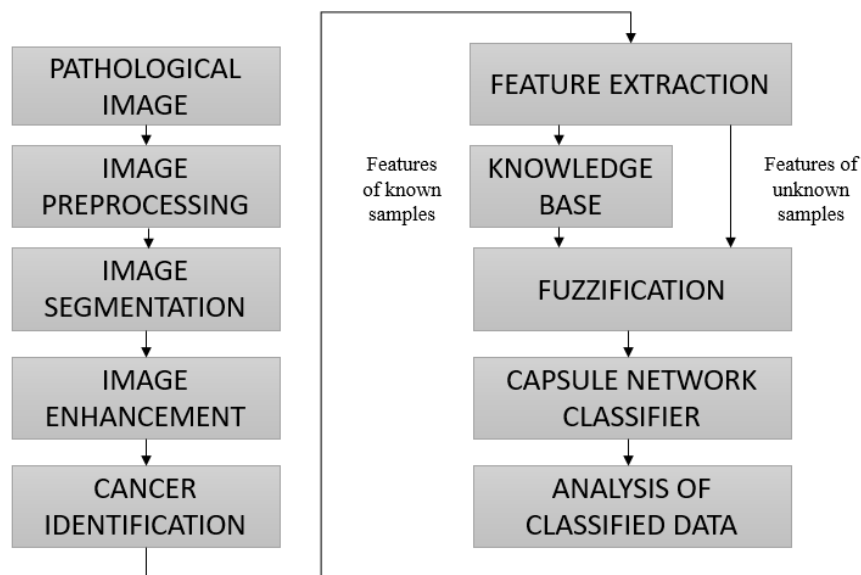


Figure 1 Block diagram of the Image Classification Unit

Image pre-processing [17] is the lowest level of abstraction in image processing. Images are enhanced, restored and brightness is adjusted at this level. Image segmentation [18] is the level in which the images are partitioned into multiple segments or sets of pixels. It is segmented based on properties such as colour, texture and intensity. This level simplifies the image into a form that can be analysed easily. It is generally used to locate objects by grouping the pixels that belong to the same object. It helps in identifying the cells available in the image. Image enhancement involves improving the quality of the image with techniques like spatial filtering, contrast enhancement, fuzzy contrast correction and density slicing. Noise removal, histogram equalization, median filtering and so on are some of the commonly used image enhancement schemes.

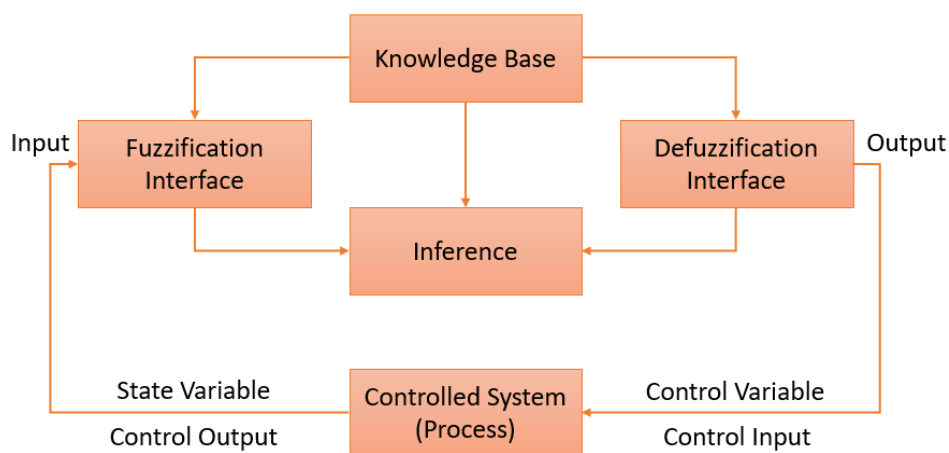


Figure 2 Fuzzy Logic Controller [22]

Based on this set of processed images, cancer identification is done. If there is any cancerous tumour identified in the image, it is further processed for the study of cancerous cells and tissues. Otherwise, the processing is stopped at this level. Feature extraction [19] is used to analyse the segments of the image. Feature is used to describe an entire picture, a collection of objects or a single object in the picture. This level also helps in elimination of false detection of cancerous modules that have been carried over from the previous stage. Binarization, normalization, thresholding, masking approach of feature extraction is used to enhance only the cancerous cells. The images are labelled with classes like: (i) normal tissue, (ii) in situ carcinoma, (iii) benign lesion, and (iv) invasive carcinoma. Grey Level Co-occurrence Matrix (GLCM) is used to differentiate abnormal brain tumours from normal ones.

The enhanced cancerous cells are compared to the knowledge database and converted into a fuzzy value by means of fuzzification [20]. Suitable member functions like age, score, and so on are used in the fuzzification process and calculations are done using triangular and trapezoidal functions. The knowledge base contains a database of the sample cancerous cells as a rule base. It also encompasses diagnosis rules and procedure to identify the cancerous cells. At the fuzzification stage, type-2 fuzzy logic is used to attain an appropriate diagnosis. Fuzzification is done by assigning the image with one or more membership values that is inclusive of the most important properties. Local fuzzification, histogram based grey-level fuzzification and feature fuzzification are the common fuzzification techniques [15].

Finally, the images are transferred to the capsule network classifier. This method is more advanced and advantageous compared to the traditional convolution neural networks (CNN) [21]. A layer of reshaped and compressed primary capsules from the previous convolutional layer is used along with another layer of cancer capsules that represent the type of image. Routing-by-agreement or max-pooling routing algorithm is used for enabling better learning of the neural network. The classification of the features are done and a parametric embedding technique, t-SNE is applied for dimensionality reduction.

The final classified data with all the required information is sent for analysis by the physician and researchers to assist in identification of the stage and root cause of the disease and provide the required treatment.

4. RESULT

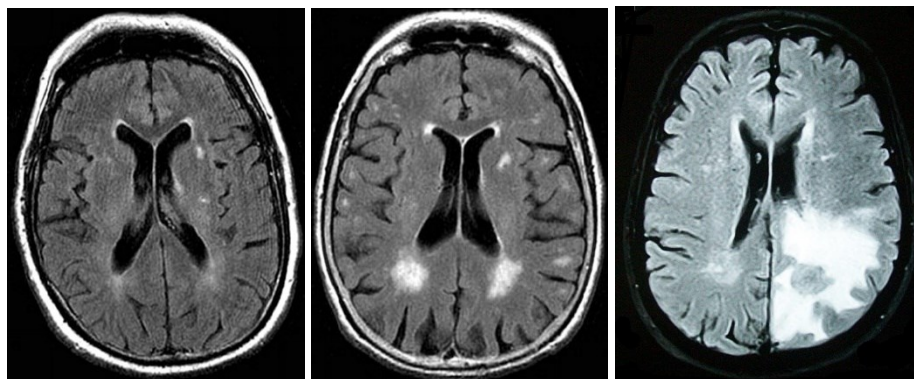


Figure 3 MRI Images of human brain with cancerous cells

Figure 3 represents the MRI images of human brain in which the white areas represent the cancerous region. The intensity of affected area varies from person to person based on which the stage of cancer can be interpreted.

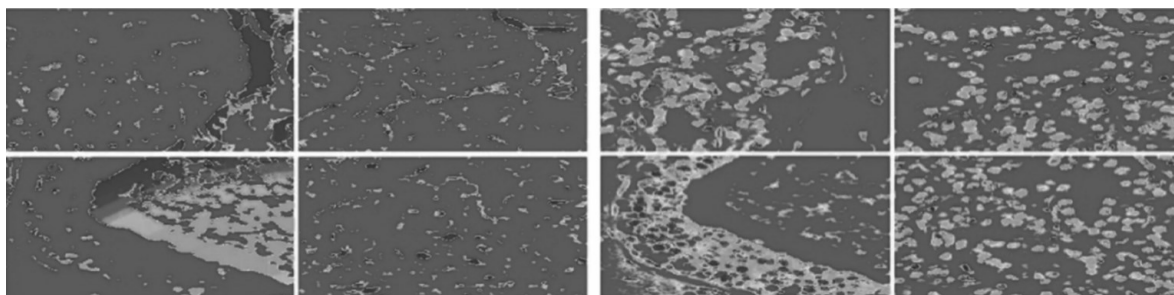


Figure 4 A histological image that helps in identification of nuclei, cytoplasm and other layers of the cell

Figure 4 represents a histological image that contains the information of the cancerous cell nuclei, cytoplasm and other layers. Analysis of this data helps in identification of the root cause of the tumour and thereby helps in guiding towards providing proper treatment based on the information gathered. More features can be included such as metabolic, genetic and anatomical attributes. With the increase in the data in knowledge base and the number of iterations, the accuracy of the system will be improved.

5. CONCLUSION AND FUTURE SCOPE

The proposed cancer cell identification and classification scheme from pathological images using capsule network aids the physicians with an expert approach to treat the problem. The risk factors along with probable symptoms are stored in the knowledge base and a capsule network classifier uses this information for the processing of the pathological images obtained in digital format. Future work involves exploring the unsupervised learning approach of capsule network as well as creating a user friendly interface for the access of data that has been processed and provided by the capsule network.

References

- [1] Sabour, Sara, Nicholas Frosst, and Geoffrey E. Hinton. "Dynamic routing between capsules." In *Advances in neural information processing systems*, pp. 3856-3866. 2017.
- [2] Almutashri, Ali, Sos Aghaian, Ian Thompson, Danny Rabah, Osman Zin Al-Abdin, and Marlo Nicolas. "Gleason grade-based automatic classification of prostate cancer pathological images." In 2011 IEEE International Conference on Systems, Man, and Cybernetics, pp. 2696-2701. IEEE, 2011.
- [3] Zhu, Xinliang, Jiawen Yao, Xin Luo, Guanghua Xiao, Yang Xie, Adi Gazdar, and Junzhou Huang. "Lung cancer survival prediction from pathological images and genetic data—an integration study." In 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pp. 1173-1176. IEEE, 2016.
- [4] Joshi, Dipali M., N. K. Rana, and VMi Misra. "Classification of brain cancer using artificial neural network." In 2010 2nd International Conference on Electronic Computer Technology, pp. 112-116. IEEE, 2010.
- [5] Lau, Phooi Yee, Frank CT Voon, and Shinji Ozawa. "The detection and visualization of brain tumors on T2-weighted MRI images using multiparameter feature blocks." In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, pp. 5104-5107. IEEE, 2006.
- [6] Descombes, Xavier, Frithjof Kruggel, Gert Wollny, and Hermann Josef Gertz. "An object-based approach for detecting small brain lesions: application to Virchow-Robin spaces." *IEEE transactions on medical imaging* 23, no. 2 (2004): 246-255.
- [7] Afshar, Parnian, Arash Mohammadi, and Konstantinos N. Plataniotis. "Brain tumor type classification via capsule networks." In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 3129-3133. IEEE, 2018.

[8] Iesmantas, Tomas, and Robertas Alzbutas. "Convolutional capsule network for classification of breast cancer histology images." In *International Conference Image Analysis and Recognition*, pp. 853-860. Springer, Cham, 2018.

[9] Mobiny, Aryan, and Hien Van Nguyen. "Fast capsnet for lung cancer screening." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 741-749. Springer, Cham, 2018.

[10] Sonka, Milan, Vaclav Hlavac, and Roger Boyle. "Image pre-processing." In *Image Processing, Analysis and Machine Vision*, pp. 56-111. Springer, Boston, MA, 1993.

[11] Sklansky, Jack. "Image segmentation and feature extraction." *IEEE Transactions on Systems, Man, and Cybernetics* 8, no. 4 (1978): 237-247.

[12] Al- Janabi, Shaimaa, André Huisman, and Paul J. Van Diest. "Digital pathology: current status and future perspectives." *Histopathology* 61, no. 1 (2012): 1-9.

[13] Xi, Edgar, Selina Bing, and Yang Jin. "Capsule network performance on complex data." *arXiv preprint arXiv:1712.03480* (2017).

[14] Wu, Albert W., Angela KM Lipshutz, and Peter J. Pronovost. "Effectiveness and efficiency of root cause analysis in medicine." *Jama* 299, no. 6 (2008): 685-687.

[15] Tahani, H., Keller, J. C. (1992): The fusion of information via fuzzy integration. In: Proc. NAFIPS'92, Puerto Vallarta, Mexico, 468-477

[16] Allen, Jeffrey D., Min Chen, and Yang Xie. "Model-based background correction (MBCB): R methods and GUI for Illumina Bead-array data." *Journal of cancer science & therapy* 1, no. 1 (2009): 25.

[17] Fave, Xenia, Lifei Zhang, Jinzhong Yang, Dennis Mackin, Peter Balter, Daniel Gomez, David Followill, A. Kyle Jones, and Francesco Stingo. "Impact of image preprocessing on the volume dependence and prognostic potential of radiomics features in non-small cell lung cancer." *Translational Cancer Research* 5, no. 4 (2016): 349-363.

[18] Xu, Yan, Jun-Yan Zhu, I. Eric, Chao Chang, Maode Lai, and Zhuowen Tu. "Weakly supervised histopathology cancer image segmentation and classification." *Medical image analysis* 18, no. 3 (2014): 591-604.

[19] Hira, Zena M., and Duncan F. Gillies. "A review of feature selection and feature extraction methods applied on microarray data." *Advances in bioinformatics* 2015 (2015).

[20] Saleh, Ahmed Abou Elfetouh, Sherif Ebrahim Barakat, and Ahmed Awad Ebrahim Awad. "A fuzzy decision support system for management of breast cancer." *IJACSA Editorial* (2011).

[21] Spanhol, Fabio Alexandre, Luiz S. Oliveira, Caroline Petitjean, and Laurent Heutte. "Breast cancer histopathological image classification using convolutional neural networks." In *2016 international joint conference on neural networks (IJCNN)*, pp. 2560-2567. IEEE, 2016.

[22] Farahani, Farzad Vasheghani, MH Fazel Zarandi, and Abbas Ahmadi. "Fuzzy rule based expert system for diagnosis of lung cancer." In *2015 Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS) held jointly with 2015 5th World Conference on Soft Computing (WConSC)*, pp. 1-6. IEEE, 2015.