A DEEP LEARNING MODEL FOR THE AUTOMATIC DETECTION OF MALIGNANCY IN EFFUSION CYTOLOGY

Shajahan Aboobacker[⋆] *Deepu Vijayasenan*[⋆] *Sumam David S.*[⋆] *Pooja K Suresh*[†] *Saraswathy Sreeram*[†]

[⋆]Department of ECE, National Institute of Technology Karnataka, Surathkal, Karnataka, India

[†] Department of Pathology, Kasturba Medical College Mangalore, Manipal Academy of Higher Education, Manipal, Karnataka, India

ABSTRACT

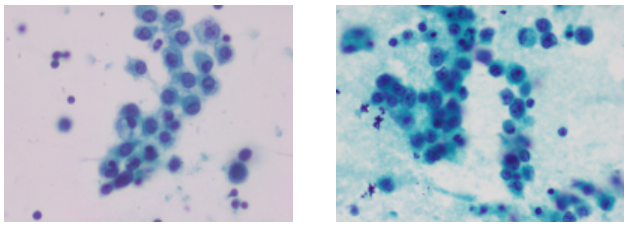
The excessive accumulation of fluid between layers of pleura covering lungs is known as pleural effusion. Pleural effusion may be due to various infections, inflammations or malignancy. The cytologists visually examine the microscopic slide to detect the malignant cells. The process is time-consuming, and interpretation of reactive cells and cells with ambiguous levels of atypia may differ between pathologists. Considerable research is happening towards the automation of fluid cytology reporting. We propose an integrated approach based on deep learning, where the network learns directly to detect the malignant cells in effusion cytology images. Architecture U-Net is used to learn the malignant and benign cells from the images and to detect the images that contain malignant cells. The model gives a precision of 0.96, recall of 0.96, and specificity of 0.97. The AUC of the ROC curve is 0.97. The model can be used as a screening tool and has a malignant cell detection rate of 0.96 with a low false alarm rate of 0.03.

***Index Terms*—** Effusion Cytology, Machine Learning, Deep Neural Network

1. INTRODUCTION

The pleural cavity is the space between outer pleura, which is connected to the wall of the chest and inner pleura, which lies over the surface of the lung. A small amount of fluid present normally between the two layers with a low concentration of protein. Excessive accumulation could be secondary to any irritation of pleura, including the infections, inflammations and malignancies. Malignant cells, mesothelial cells and inflammatory cells may constitute the cells in effusions. Primary pleural mesothelioma, lung adenocarcinoma as well as metastasis from adenocarcinoma of other sites can be seen in the pleural effusion [1]. Adenocarcinoma is one of cause for the frequently occurring malignant cells in pleural effusion [2]. Metastasis is usually seen from organs like ovary and breast in women [3]. Lymphoma can also lead to a malignant [2].

The early and efficient diagnosis increase overall survival rate of cancer patients. The cytological examination is generally conducted to diagnose disease in the pleural effusion because it is inexpensive and accurate. Sufficient amount of fluid sample extracted from the pleural cavity centrifuged and the sediment is smeared on a glass slide and stained with either Papanicolaou stain (PAP) or hematoxylin and eosin (H&E) stain. Figure 1 shows microscopic images of pleural effusion cytology stained with PAP stain.



(a) Benign Image (b) Malignant Image

Fig. 1. Effusion cytology images stained with PAP stain

The cytopathologists visually examine the slides under the microscope to find the malignant cells. Though most of the inflammatory and malignant conditions have a straightforward morphology, ambiguity arises in cases with florid mesothelial proliferation, reactive cells or malignancies with subtle atypia. This is the grey zone in which the impressions of expert cytopathologists may also differ. To mitigate this conundrum, an objective evaluation via computer-assisted algorithms could be of great utility. Often the CAD systems are used as a screening tool to eliminate benign cases. CAD systems shall have an operating point with near to zero miss rate and a small false alarm rate. The CAD systems help lighten the workload of cytologists, quicken the diagnosis procedure and eliminate the inter observer variations.

Traditionally effusion cytology is based on first manually [4, 5, 6] or automatically [1, 2, 3] identifying the nuclei and then measuring the morphometrical parameters such as nuclei area, cytoplasm area, nucleus to cytoplasm ratio and textural features to distinguish malignant cells from benign cells. Some research is done to address overlapped nuclei in the images [7, 8], but in those work, malignant cells are not identified.

Other methods depend on Machine Learning methods such as fuzzy detection algorithm [9], and artificial neural networks [10] to distinguish malignant from benign cells. All

978-1-7281-7202-6/20/$31.00 ©2020 IEEE

of these algorithms employ a two-stage process where the errors accumulate from each stage.

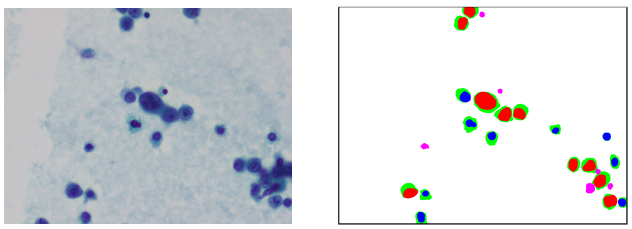
We propose an integrated approach based on deep learning, where the network learns directly to detect the malignant cells. We use U-Net architecture for the integrated segmentation and classification.

2. METHODOLOGY

The proposed algorithm is a deep learning model, where the network learns directly from the image to detect the malignant cells. The input image contains benign and malignant cells, cytoplasm and inflammatory cells. The task is to identify the presence of malignant cells. We try to classify each pixel into one of the five classes - Benign, Malignant, Cytoplasm, Inflammatory and Background. This is a semantic segmentation problem because it should understand the image in pixel level, so we used U-Net for this task.

2.1. Dataset

The work is done jointly with the Department of Pathology, KMC Mangalore, with the approval from the KMC Mangalore Institutional Ethics Committee (IEC KMC MLR 01-19/34; 16th Jan 2019). All the formalities were completed according to policies of the institution. The study involves a dataset of 212 pleural effusion cytology images of 30 patients with size 1920 $\times$ 1440. It contains benign, malignant and inflammatory cells. The images acquired by smearing the fluid on a glass slide then stained with Papanicolaou stain (PAP) or hematoxylin and eosin (H&E) and photographed with 40x optical magnification using the digital microscope camera Olympus BX53 and cellSense imaging software. We manually annotated the images to prepare the ground truth with the supervision of an experienced pathologist. Benign nuclei, malignant nuclei, cytoplasm, inflammatory cells and the background are marked with different colours. Figure 2 shows a sample image along with its annotated ground truth.



(a) (b)

Fig. 2. Sample Images (The red colour is marked for malignant, blue is for benign, green is for cytoplasm, magenta is for inflammatory and white for the background)

2.2. U-Net

U-Net is a deep neural learning model, which consists of fully convolutional layers. The data available for training is very less as we are working with medical data so we need an architecture which will work well with less amount of training data. U-Net produces excellent results for the segmentation of objects in medical images for a less amount of data [11]. It has three parts, namely a contracting (encoding) path, a bottleneck and an expanding (decoding) path as shown in the Figure 3. Contracting path contains convolutional layers and max pool layers, which reduces the spatial dimensions. The expanding path steadily retrieves the image details and dimensions through up-convolution. The bottleneck part serves as the principal connection between the encoder and decoder. There are skip connections between contracting and expanding paths to give additional information from the encoder to the decoder while up-convolution.

To improve the speed of the convolutional neural networks (CNN) we use separable convolutions. Each regular convolution is divided as depthwise and pointwise convolutions [12]. In the depthwise stage, a filter is individually convolved with each channel followed by a 1x1 convolution to merge all these channels. The separable convolution improves computation speed and decreases the size of parameters.

2.3. Experimental Setup

Even though U-Net can handle variable size input data, fixed-size patches are generated for batch training. We had 212 images with pixel size 1920 $\times$ 1440. The data was randomly partitioned as 148 training (89 benign and 59 malignant images), 21 validation (12 benign and 9 malignant images) and 43 testing (26 benign and 17 malignant images) images independent of patients. We divided the training images into patches of 512 $\times$ 512 pixels to perform batch training. By doing a horizontal and vertical shift of 128 pixels and 116 pixels, respectively, each image was cropped into 108 images. Hence we had around 16000 image patches for training. As the model is a fully convolutional network, original size was retained for the test images. The distribution of pixels among different classes in training, validation and testing is shown in the Table 1.

2.4. Training

The U-Net model was trained with the output classes using the loss function weighted cross-entropy. Since the number of background pixels was relatively large the network favoured the background compared to any other classes. In order to avoid this, a weighted cross-entropy function was used for optimization, giving less weightage for the background class. The weight of each class was determined as the inverse class frequency from the training labels. To synthesize the model

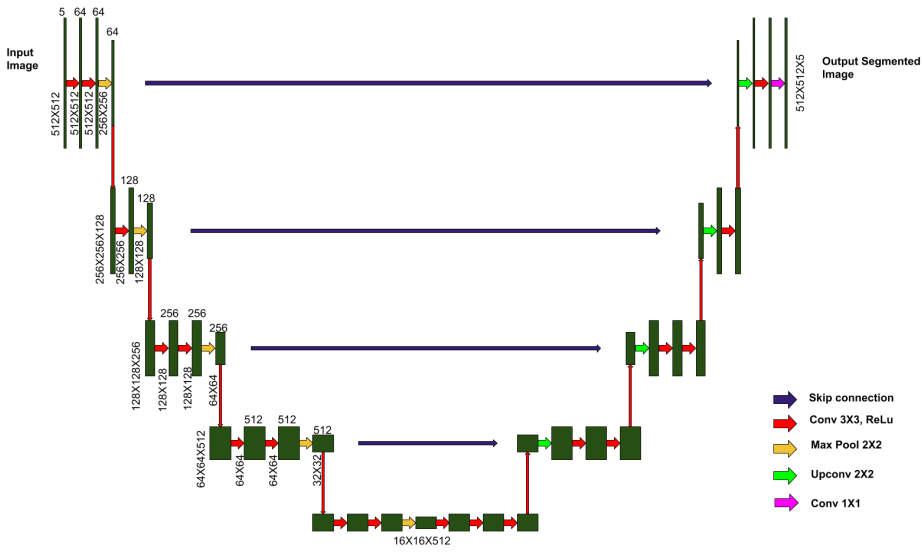


Fig. 3. U-Net Architecture

Table 1. Distribution of pixels among different classes

Class	Training	Validation	Testing
Foreground	5.9%	4.5%	4.8%
Background	94.1%	95.5%	95.2%

(a) Distribution of pixels as foreground and background

Class	Training	Validation	Testing
Benign	11.3%	20.2%	11.0%
Malignant	31.2%	18.7%	30.3%
Inflammatory	20.0%	21.8%	21.0%
Cytoplasm	37.5%	39.3%	37.7%

(b) Distribution of classes among foreground pixels

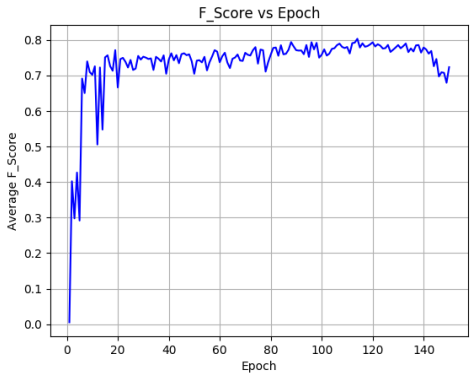


Fig. 4. Average F-Score vs Epochs

data augmentation was used. A random rotation, flipping and brightness change were performed as part of data augmentation. The model was trained for 150 epochs. The output of the model was a pixel-wise classified image. Figure 4 shows the pixel-level average f-score of benign and malignant for training phase. The f-score value reached a maximum of 0.8 at 114^{th} epoch. This model was used for testing.

3. TESTING AND POST-PROCESSING

We gave 64 images as test images to the trained model, the output of which was a five-class segmented image. Figure 5 shows a sample image, ground-truth and segmented output.

The output images often contained isolated spurious malignant patches as well as gaps in the nuclei. We performed morphological operations such as opening, closing and hole filling to reduce their effect.

The goal of the system was to filter out benign images. So each image should be either a benign or a malignant. We used the segmented result to decide whether the image is malignant or not. Hence we fixed a threshold on the percentage of malignant pixels to determine if an image is malignant. All the images that had more malignant pixels than this threshold are flagged as malignant.

4. RESULTS AND DISCUSSIONS

To study the performance of the network independent of the threshold, we considered the Receiver Operating Characteristic (ROC) curve of the system. Figure 6 shows the ROC curve

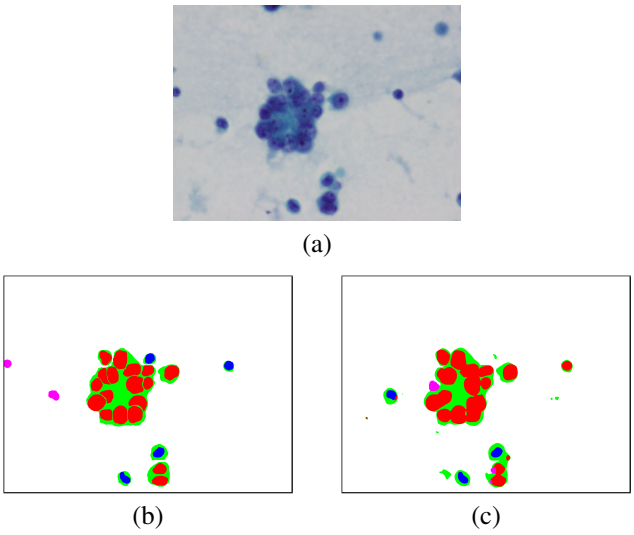


Fig. 5. Sample Images: (a) Original image (b) Ground truth (c) Predicted image

achieved by finding the sensitivity and specificity among the validation and test data for various threshold values. The Area Under the Curve (AUC) of the ROC was 0.97. The threshold to predict the malignant cases was estimated from the ROC curve. We chose the operating point, with at least 95% sensitivity that has a minimum false alarm, which was at a threshold of 0.16% malignant pixels.

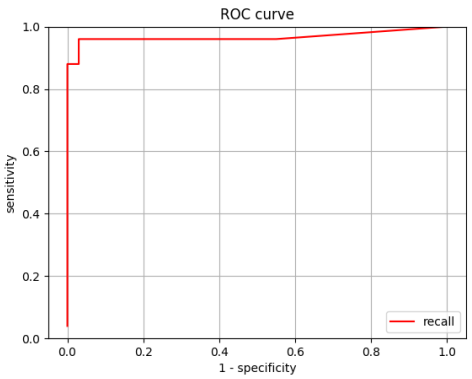


Fig. 6. ROC curve of malignant images

At the threshold value, the sensitivity (Recall) was 0.96 and False alarm (1-specificity) was 0.03. Most of the errors were due to benign cells getting classified as malignant. The results obtained were verified by the expert pathologists and they rated it as acceptable.

To test the validity of the U-Net for this application, we also implemented another semantic segmentation model Deeplab V3 plus, which is a state of art method for natural image segmentation [13, 14]. We used the same training and test data. The AUC obtained was 0.81. The sensitivity was 0.71 and false alarm rate was 0.13 for the same threshold value.

5. CONCLUSION

We used an integrated approach for detecting malignant cells in microscopic images of 40x magnification. We used a U-Net model for semantic segmentation and we fixed a threshold value for the percentage of malignant pixels to detect the malignant images. The Precision was 0.96, Recall was 0.96 and Specificity was 0.97. Unlike other systems, this model is an integrated system for effusion cytology which accepts input image and directly predicts whether the image is malignant or not. The obtained sensitivity value (0.96) and false alarm rate (0.03) satisfy the clinical requirements. U-Net showed an improvement of 20% in comparison with Deeplab V3 plus in terms of AUC of the ROC.

6. REFERENCES

[1] Khin Yadanar Win, Somsak Choomchuay, Kazuhiko Hamamoto, and Manasanan Raveesunthornkiat, "Artificial neural network based nuclei segmentation on cytology pleural effusion images," in *2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*. IEEE, 2017, pp. 245–249.

[2] Khin Yadanar Win and Somsak Choomchuay, "Automated segmentation of cell nuclei in cytology pleural fluid images using otsu thresholding," in *2017 International Conference on Digital Arts, Media and Technology (ICDAMT)*. IEEE, 2017, pp. 14–18.

[3] Khin Yadanar Win, Somsak Choomchuay, Kazuhiko Hamamoto, and Manasanan Raveesunthornkiat, "Comparative study on automated cell nuclei segmentation methods for cytology pleural effusion images," *Journal of Healthcare Engineering*, vol. 2018, 2018.

[4] Marie Moses Ambroise, Prabhavati Jothilingam, and Anita Ramdas, "Utility of nuclear morphometry in effusion cytology," *Asian Pac J Cancer Prev*, vol. 15, no. 16, pp. 6919–6922, 2014.

[5] B Arora, Savita Setia, and Bharat Rekhi, "Role of computerized morphometric analysis in diagnosis of effusion specimens," *Diagnostic Cytopathology*, vol. 34, no. 10, pp. 670–675, 2006.

[6] Zainab AJ Al-Obaidi, "Diagnostic approach of atypical cells in effusion cytology using computerized image analysis," *Al-Kindy College Medical Journal*, vol. 4, no. 1, pp. 60–62, 2007.

[7] Khin Yadanar Win, Somsak Choomchuay, and Kazuhiko Hamamoto, "K mean clustering based automated segmentation of overlapping cell nuclei in pleural effusion cytology images," in *2017 International Conference on Advanced Technologies for Communications (ATC)*. IEEE, 2017, pp. 265–269.

[8] Khin Win, Somsak Choomchuay, Kazuhiko Hamamoto, and Manasanan Raveesunthornkiat, "Detection and classification of overlapping cell nuclei in cytology effusion images using a double-strategy random forest," *Applied Sciences*, vol. 8, no. 9, pp. 1608, 2018.

[9] Lei Zhang, Qiuguang Wang, and Jiping Qi, "Research based on fuzzy algorithm of cancer cells in pleural fluid microscopic images recognition," in *2006 International Conference on Intelligent Information Hiding and Multimedia*. IEEE, 2006, pp. 211–214.

[10] Adarsh Barwad, Pranab Dey, and Shaily Susheilia, "Artificial neural network in diagnosis of metastatic carcinoma in effusion cytology," *Cytometry Part B: Clinical Cytometry*, vol. 82, no. 2, pp. 107–111, 2012.

[11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[12] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[13] Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos, "Image segmentation using deep learning: A survey," *arXiv preprint arXiv:2001.05566*, 2020.

[14] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2017.