Sensors & Transducers, Vol. 249, Issue 2, February 2021, pp. 102-109

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2021
http://www.sensorsportal.com

The Automated Diagnosis Architecture and Deep Learning Algorithm for Cervical Cancer Cell Images

[1, *] Jong-Ha Lee and [2] Sangwoo Cho

[1] Dept. of the Biomedical Engineering, Keimyung University, Daegu, South Korea
[2] The Center for Advanced Technical Usability and Technologies, Keimyung University, Daegu, South Korea
[*] E-mail: Segeberg@gmail.com

Received: 22 January 2021 /Accepted: 22 February 2021 /Published: 28 February 2021

Abstract: Cervical cancer is the second most common cancer in women worldwide, with approximately 500,000 cases each year. Approximately 80 % of the reported deaths are observed in countries with poor medical conditions, such as developing countries; 230,000 people die each year from cervical cancer. When detected early, it is possible to prevent progression to invasive cancer by adequate treatment. Therefore, it is very important to detect human papillomavirus (HPV), which is known as a major cause of cervical cancer, and cervical cancer that has already progressed. The most well-known test to diagnose cervical cancer is the Pap test. Although the Pap test consumes more time to diagnose cervical cancer, this test has saved the lives of many patients through early detection of cells progressing into cervical cancer for treatment and has contributed significantly to the prevention of cervical cancer and the reduction of mortality due to cervical cancer. However, since the Pap test requires emanating the cell slide of each patient, which takes a lot of time, the pathologists performing Pap tests tend to become very tired, and it is very difficult to examine many patients. Therefore, the need for diagnostic aids to help pathologists make quick decisions is on the rise. This study aimed to address this problem by developing an automatic diagnostic aid tool for cervical cancer using Yolo V3, a deep learning algorithm. First, the RGB cell image is converted into a gray-scale image by pre-processing, and the noise in the image is removed using a 2-dimensional Gaussian smoothing filter. Next, each cell image for training was labeled by a pathologist. Finally, using the trained algorithm, the cells in the image from the Pap test are identified and marked by bounding boxes to aid the pathologist with rapid diagnosis. For training of the model, 5,631 pre-processed Pap test images were used, and the model was then tested using 563 images. In this process, the performance indicator of PASCAL Visual Object Classes was used, which was set by raising the threshold from 0 to 1.

The precision and recall values for all test images were obtained to calculate the average precision value corresponding to the recall value and to use it as an indicator to determine the performance of the algorithm. The average precision in this study was 73.34 %, and the model may be used as an auxiliary tool for pathologists performing Pap tests by improving accuracy using additional data in the future.

Keywords: Cervical, Cancer, Cell, Deep learning, AI.

1. Introduction

1.1. Background

Cervical cancer is the second most prevalent cancer among women worldwide. Approximately 500,000 women are diagnosed with cervical cancer each year, and approximately 230,000 women die each year due to cervical cancer. In particular, a majority of women who are diagnosed with cervical cancer belong to developing countries, where the incidence is 7-8 times higher than that of developed countries. In

South Korea, according to the data from Statistics Korea in 2017, a total of 3,469 cervical cancer cases were reported, and it is the fourth most common cancer among Korean women, next to breast cancer, gastric cancer, and thyroid cancer.

Compared to a previous report published by Statistics Korea, the proportion of Korean women with cervical cancer seems to have decreased, which is due to improved early diagnosis and treatment, with increased interest in cervical cancer. However, cervical cancer remains a common cancer in women, and it is fatal if left unattended. When detected early, it is possible to prevent cancer progression to an invasive tumor. In other words, early detection of human papillomavirus (HPV), the major cause of cervical cancer, and early diagnosis of cervical cancer are crucial.

1.2. Theoretical Background

1.2.1. Cervical Cancer

The uterus is divided into two parts: the body part, which occupies approximately three quarters of the uterus, and the cervical part connected to the vagina. Cervical cancer occurs in the cervix, which is the neck of the uterus. Cervical cancer occurs when normal epithelial cells in the epidermis of the cervix undergo microscopic morphological changes and form cervical intraepithelial dysplasia (or cervical intraepithelial neoplasia, CIN), a state between normal tissues and cancerous tissues; CIN then progresses to cervical epithelial cancer, with the presence of cancer cells only in the epithelium of the cervix, which is stage 0 cervical cancer. If cervical cancer is not detected at stage 0 and is not treated, it then progresses to invasive cervical cancer.

Normal cells gradually develop into cancerous cells, and such progression can take years to decades. Normal cells develop into ASCUS, L-SIL, ASCUS-H, H-SIL, and, eventually, to cervical cancer. Epithelial cells make up the outermost part of a tissue. The epithelium is located on the outermost part of the skin or the entire organ of the body, and the connective tissue is located below it. The region that separates the epithelium and the connective tissue is the basement membrane. When the cancer penetrates through this basement membrane, it is judged to have progressed to being invasive cancer. Squamous cell carcinoma (SCC), which is one of the two main types of cervical cancer, accounts for approximately 80 % of all cervical cancers. Adenocarcinoma, which is the other type, accounts for 10–20 % of all cervical cancers. Cases involving both adenocarcinoma and SCC are called mixed carcinoma, which account for 2–5 % of all cancers. SCC is then classified as 65 % of large cell non-keratinized subtype, 25 % of large cell keratinized subtype, and 5 % of small cell non-keratinized subtype. The large cell subtype shows a better prognosis than the small cell subtype.

In terms of the risk factors for cervical cancer, the incidence rate is relatively higher in women who have experienced sexual intercourse at an early age, women who have a sexual partner with HPV, and women who have multiple sexual partners, indicating that it is closely associated with sexual intercourse. It also tends to occur relatively more in smokers or in those who are less socially and economically privileged. In addition to these risk factors, it is known that sexual intercourse-induced HPV infection has a major effect on the onset of cervical cancer. Approximately 50–80 % of people are infected with HPV at least once in their lifetime. With regard to cervical cancer, HPV is divided into high-risk HPV and low-risk HPV, and there are approximately 150 kinds of viruses. The U.S. Centers for Disease Control and Prevention states that the most important cause of cervical cancer is persistent infection by HPV.

People who are engaged in sexual intercourse are constantly exposed to the risk of infection by HPV. Most of them are infected during sexual intercourse, but there is a low probability that they are infected through indirect sexual contact or by means other than sexual activity. Most HPV infections heal naturally within 6 months to 2 years. In addition, 56 % cases of cervical cancers occur in women who are not regularly screened. Therefore, early and continued participation is the most effective method to prevent cervical cancer. If abnormal cells are found in the cervix during regular screening, then appropriate treatment should be initiated.

One of the most common symptoms of cervical cancer is abnormal vaginal bleeding. Abnormal vaginal bleeding refers to vaginal bleeding after menopause or irregular bleeding before menopause but not during menstruation. Such bleeding occurs when blood vessels are distributed to newly formed cancer cells for nourishment. This bleeding is common after vaginal lavage, heavy exercise, sexual intercourse, or defecation. When cancer occurs due to secondary transmission or when cancer cells are necrotized, the secretions have strong odor, resulting in an increase in overall vaginal discharge. Cervical cancer progression and metastasis to other organs may result in kidney swelling and back pain.

Currently, regular screening is the best approach for detecting and preventing the formation of abnormal cells before developing cervical cancer. Regarding the screening method, the Pap test [1], known as the cervical cancer test, is performed, wherein cervical cells are collected using a small brush, and the pathologist directly detects abnormal cells under a microscope. Women who have had sexual intercourse or are 20 years of age or older are screened for cervical cancer, and it is recommended that they undergo the Pap test once a year. Moreover, an additional test for HPV is recommended for a more accurate examination.

When abnormal signs are detected during the examination, colposcopy and, in some cases, biopsy are performed, wherein part of the cervical tissue is removed and examined. Treatment of cervical cancer

depends on the disease progression and the patient's condition. Dysplasia and intraepithelial cancer can be cured by more than 95 % by treating the affected area with cryotherapy, laser therapy, conization, or loop electrosurgical excision procedure (LEEP). Pregnancy is also possible after treatment.

1.2.2. Cervical Cancer Screening

Cervical cancer screening is largely performed in two ways.

The Pap test, which is most frequently performed for the diagnosis of cervical cancer, came into use in clinical practice in 1943 after George N. Papanicolaou accidentally observed malignant cells in cervical decidual cell smear in 1928, suggesting that this test could be used for the diagnosis of cervical cancer. The Pap test has been recognized as the standard early screening method for cervical cancer for the past 50 years. Using this method, abnormal findings of cells can be observed under a microscope. First, a colposcope is inserted into the vagina, then the sample is collected with a brush and applied onto a glass slide, and smearing and staining are performed on the glass slide. It is a relatively safe method, is inexpensive with low procedural and testing cost; the sample collection procedures is simple, causing less pain, and has relatively high accuracy, i.e., 75–85 %. However, the greatest limitation of the Pap test is that it may yield a false-negative result, in other words, there are cases in which a lesion is not accurately detected [2]. Despite this false-negative rate, the Pap test is the most efficient and the only screening method with proven effects of screening, and it has contributed significantly to the prevention of cervical cancer and the reduction of mortality because cervical cancer is the most essential testing method.

Liquid-based cytology (LBC) is another widely used approach after the Pap test. LBC is different from the Pap test in terms of sample preparation for testing. As in the conventional Pap test, cells are collected directly from the cervix with a small brush, but the collected cells are stored in a bottle containing a small preservative instead of directly smearing them on a microscope slide and testing them immediately.

A slide is prepared through a separate treatment when used for observation under a microscope. This method shows high stability in terms of cell damage, which is a limitation in the Pap test and demonstrates better ability to detect H-SIL. However, there is a disadvantage that the cost of the test is relatively higher than that of the Pap test.

1.3. Previous Studies

1.3.1. Machine Learning

Machine learning, defined by the renowned computer scientist Arthur Samuel as "a field of research that allows computers to learn without being explicitly programmed by humans". Machine learning algorithms are often classified into supervised learning algorithms or unsupervised learning algorithms. First, supervised learning provides correct answers to the algorithms to let them draw conclusions by learning the data. In other words, the conclusion is drawn by learning the data to find a pattern to produce a more accurate result based on the example provided and then by undergoing the feedback process through an algorithm without external intervention to correct the result. Representative supervised learning algorithms include KNN and XGBoost. Unsupervised learning refers to an algorithm that presents the correct answer by itself without giving the correct answer data to the algorithm. Representative unsupervised learning algorithms include K-means [3] and PCA [4].

Deep learning is attracting attention under the concept of machine learning. Deep learning consisting of neural networks, as shown in Fig. 1, has the accuracy that existing algorithms have not been able to achieve in various areas. This deep learning is classified as a type of supervised learning in that the correct answer data are given, but it is also classified into a new category called machine learning, with the use of neural networks due to the advent of various deep learning model structures. In particular, in this study, deep learning was used to aid in diagnosis by identifying cells from cervical cell images that have unstructured data and by distinguishing abnormal cells from normal cells.

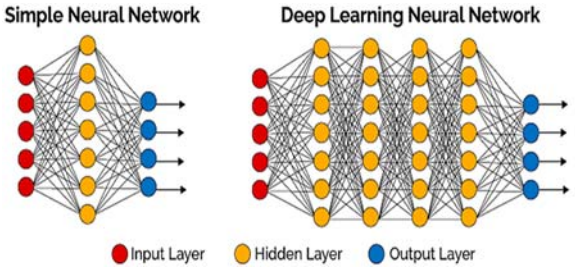

Fig. 1. Examples of differences between neural networks and deep learning networks.

1.4. Purpose

The Pap test, which is the most frequently used test to diagnose cervical cancer, detects lesions of cervical cancer, and its accuracy largely depends on the pathologist's skill, with a false-negative rate of 6–55 %. The pathologist also takes a lot of efforts to process a lot of data when examining a lesion sample. To address these problems, studies on automatized diagnosis of cervical cancer are being actively conducted. To automatically detect cervical cancer, it is necessary to divide the progression of cervical cancer by stage. There are three main ways to divide the stages: the method of dividing into normal and abnormal cells; the Bethesda System that divides progression into four stages, namely, normal cell, L-SIL, H-SIL and CIS; the World Health Organization

classification into seven stages, namely, superficial, intermediate, columea, mild dysplasia, moderate dysplasia, severe dysplasia, and carcinoma. In this study, data are classified according to "The Pap Test and Bethesda" to assist pathologists in making quick decisions. Data were collected from 987 patients from 2017 to 2018, and Yolo V3, a powerful algorithm for object detection, was used to find cells in the cervical cancer images, marking them by bounding boxes, and to distinguish abnormal cells from normal cells. This study aimed to reduce the time required for detecting lesions and to improve the efficiency of cervical cancer screening by helping the pathologist make judgments [5].

2. Materials and Methods

2.1. Materials

2.1.1. Experimental Environment

This study was conducted using 64 GB RAM, Intel i5-9600KF, and NVIDIA GeForce RTX 2080TI. The model was configured using Tensorflow version 1.14, the Python-based machine learning library of Google, and it was produced using Python packages such as Numpy and Open-cv.

2.1.2. Data and Labeling

The data of 987 patients were collected during 2017-2018. In 5,631 pieces of data, 2,827 normal cells and 2,804 abnormal cells were classified by a pathologist to construct ground truth. Fig. 2 shows the sample of images.

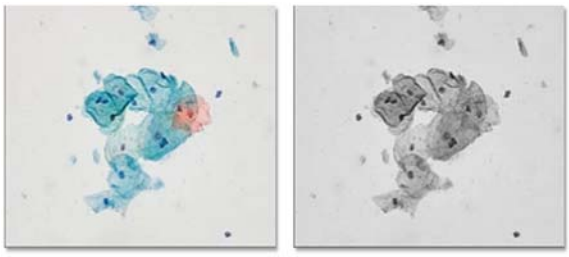

Fig. 2. Raw image (left) and pre-processed image (right).

2.2. Methods

2.2.1. Data Pre-processing Process

The cell images are displayed in color, as Papanicolaou staining has been applied to all the cells. The cytoplasm contains alkaline substances and has a strong affinity to dyes; such characteristics allow Papanicolaou staining to stain cells in various colors. After staining, the structure of the cell appears more clearly. Since the cytoplasm appears bright, providing a clear view of the nuclear chromatin, it is easy to distinguish abnormal cells from the normal ones. Papanicolaou staining is applied only to facilitate the identification of morphological features of cervical cells. The color difference caused by staining does not act as a factor that distinguishes negative cells from the positive cells when analyzing the cervical cancer images; therefore, the RGB image is converted to gray-scale image and used for training to prevent the model from recognizing the colors as a factor for training as shown in Fig. 3. In addition, 2-DGSF (2-Dimensional Gaussian Smoothing Filter) was used to remove noise in the cell image [6].

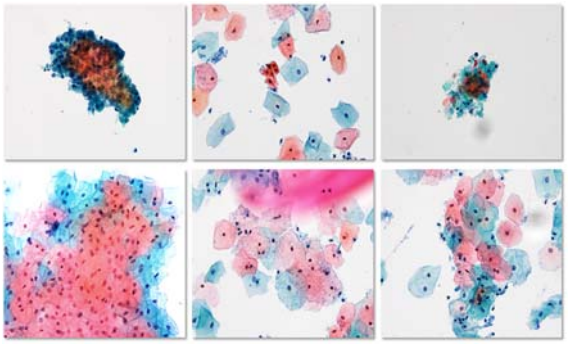

Fig. 3. Data used for training.

2.2.2. Design of Deep Learning-based Object Recognition Algorithm Yolo V3 Model

Yolo V3 [7], an object recognition algorithm based on deep learning, has proven performance in various fields. In particular, the algorithm is light enough to run smoothly on low-performance devices while maintaining a high level of accuracy. This study therefore used Yolo V3 to build a model that distinguishes abnormal cells from the normal ones among the cells in the cervical cancer images by marking the cells using bounding boxes. First, Yolo V3 predicts the position of the bounding boxes by generating anchor boxes, which are also bounding boxes with predefined shapes, by K-means clustering. While the previous Yolo 9000 algorithm predicted the center point of the grid, the Yolo V2 algorithm calculates the distance from the upper left offset and predicts the ratio of the width and height of the anchor box to be adjusted by involution.

In addition, the backbone network of Yolo V3 helped the model become much lighter. As the VGG model, the backbone network of the previous Yolo 9000, was very complex, the Darknet-19 architecture [8], which uses a relatively small number of parameters and has good performance, was applied.

A layer consisting of 3×3 Convolution and 1×1 Convolution was continuously configured, and the stride is set to 2 to lower the resolution of the feature map when performing convolution instead of MaxPooling. Then, the residual value is obtained by skip connection, and after passing through the average

pooling and fully connected layer in the last layer, the result is obtained through softmax. In Yolo V3, Darknet-53 with much more layers by applying the skip connection proposed by ResNet [9] to Darknet-19 as shown in Fig. 4.

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Fig. 4. Darknet 53 architecture, the backbone network of Yolo V3.

These characteristics enable the configuration of a lightweight model that can be mounted on a single board computer with low computing capacity, which is suitable for application as a diagnostic tool in areas with relatively poor medical infrastructure. Therefore, this study aimed to use Yolo V3 to distinguish abnormal cells from normal cells in cervical cancer cell images. Fig. 5 shows the proposed system.

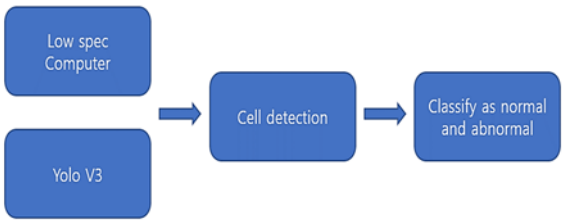

Fig. 5. Proposed assistance system for cervical cancer diagnosis.

2.2.3. Training

In the training process, among 5,631 pieces of data, 2,827 normal cells and 2,804 abnormal cells were identified and labeled by a pathologist.

The training parameters consisted of 600 epochs in a batch size of four each, and the weights learned from the COCO data set, which is the public data set of Microsoft from the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), were applied by fine-tuning. The COCO data set is composed of 80 categories, each of which is labeled. It consists of approximately 300,000 images and 1.5 million objects, and information of each image can be obtained as a JSON file. Fine-tuning is a type of transfer-leaning among weight initialization techniques that takes the weight of an existing trained model for retraining the data to be trained. Fine-tuning greatly reduces the epochs required for training and improves the performance of the model.

For the learning rate applied to learning, the first stage was set up to 200 epochs, and the value 1e-4 was applied for training. The second stage was set up to 600 epochs, and the value 1e-6 was applied for loss convergence. As shown in Fig. 6, training was conducted using the Tensor board of the deep learning framework Tensorflow to check the progress.

3. Results

3.1. Study Results

3.1.1. Performance Indicator

To evaluate the previously trained model, the performance indicator of the Pascal VOC competition is commonly used in object detection algorithms [10]. In this performance verification method, the average precision using recall and precision is expressed in each class, that is, the mean average precision (mAP) is measured by averaging the values for normal cells and abnormal cells. The value of recall is obtained by dividing the entire true proposition by the true items detected by the model, and this indicates how accurately the target object is identified. The value of precision is obtained by dividing all the items identified by the model in the correct answer label, which indicates how accurate the detected result is, that is, how much of the actual correct answer label is included among the detected results. The threshold is raised from 0 to 1 based on the IoU value, an indicator showing how much the correct answer label coincides with the detected bounding box, to find the average precision. Items for evaluating most object detection algorithms use this method to verify the performance of the model, and it was used in this study to assess the performance of the model. In addition, the mean accuracy was obtained by calculating the accuracy of the two classes of detected cells.

3.1.2. Model Performance Evaluation

The data used in the model evaluation were classified the 563 cervical cancer cell images; when testing cervical cancer cell images with a threshold of 0.45, it was judged by outputting the image as shown in Fig. 7.

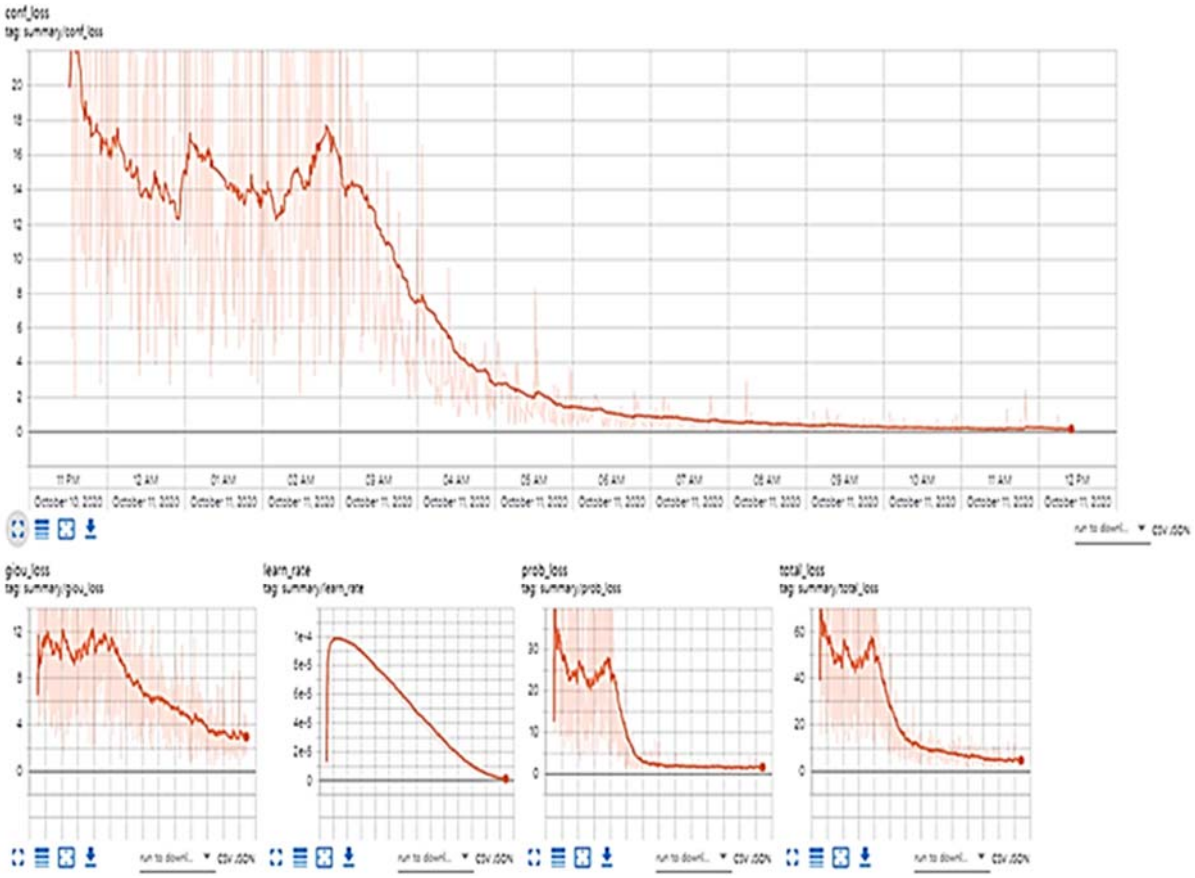

Fig. 6. Loss value of the Tensor board.

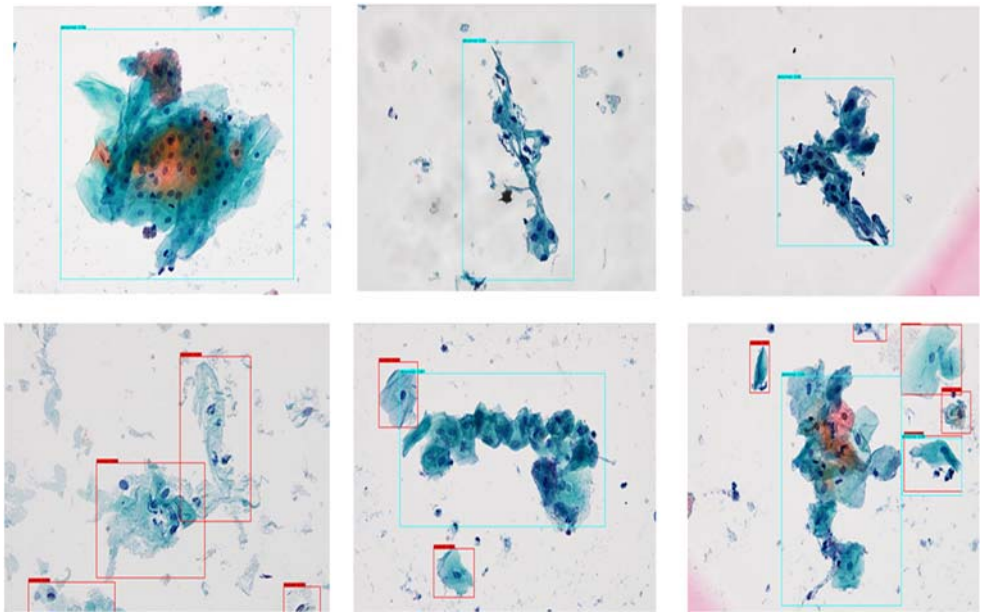

Fig. 7. Cervical cancer images classified using the model.

Fig. 8 shows our Mean average precision and accuracy result. Mean average precision was 73.34 %, which was quite high, and mean accuracy was 86.45 %. Considering the nature of the deep learning model, the larger the size of the ground truth, the higher the accuracy tends to be. It can be used as an auxiliary tool for pathologists who conduct the Pap test, by improving the accuracy using additional data later on.

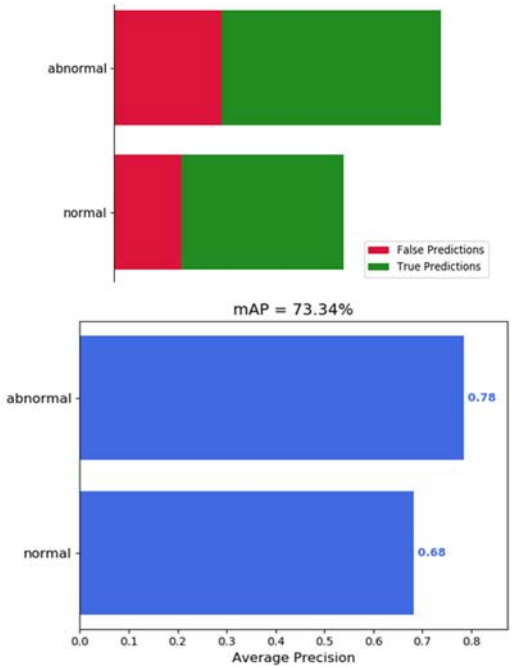

Fig. 8. Mean average precision and mean accuracy of the model.

4. Discussion

This study proposed a method to distinguish abnormal cells from normal cells in cervical cancer cell images, even with low computing power using the deep learning model Yolo V3. Despite relatively fewer data, the performance was higher than the mean average precision of other object detection algorithms using the COCO dataset which was used for transfer training. However, its current performance is not enough to assist pathologists in cervical cancer diagnosis. Nevertheless, there is ample room for improving the accuracy, as the nature of deep learning algorithms allows the accuracy to be improved by obtaining additional data. Even though various deep learning techniques such as data argumentation are not applied, its performance is competitive with that of other object detection algorithms using Microsoft COCO data set, suggesting that Yolo V3 works without problems even in pathological images.

The nature of deep learning demonstrates strong performance in unstructured data such as images; therefore, the use of a deep learning algorithm such as Yolo V3, which shows high performance even on a single board computer, as an auxiliary tool for automatic diagnosis would be useful for developing countries that have a relatively high incidence of cervical cancer. This approach will be very useful to pathologists who spend more time and therefore experience high levels of fatigue due to Pap-testing when screening for cervical cancer, especially when the number of pathologists continues to decrease.

The total number of data used for this study was 5,631, which were classified into 2,827 normal cells and 2,804 abnormal cells. Since the size of the data is significantly small, there remains room for significant improvement in the accuracy of the model by adding data.

Moreover, when the cells identified as abnormal cells are classified into each stage according to the progression of cervical cancer, that is, when the number of classes of abnormal cells is increased, the number becomes very low, resulting in data imbalance. To this end, additional data sets should be obtained to enable the classification of cervical cancer by stage. Then, the techniques used in various deep learning should be applied to improve the accuracy and to reinforce the training against noise data. As Yolo V3 has demonstrated satisfactory performance with pathological images, it can be applied to a wide variety of diseases other than cervical cancer by constructing a model through training with other pathological images.

5. Conclusion

The Pap test, which is the most frequently performed test for the diagnosis of cervical cancer, is a method for observing abnormal findings of cells under a microscope. It is performed by inserting a colposcope into the vagina, collecting a sample with a brush, applying the sample to a glass slide, and smearing and staining the slide. It has advantages such as being inexpensive in terms of procedural and test costs, having a simple sample collection process causing little pain, and having relatively high accuracy. However, the Pap test involves a significant amount of time for diagnosis. Contrary to the recent surge in demand for testing, the number of pathologists is decreasing. Moreover, in developing countries, where early diagnosis of cervical cancer is difficult due to poor medical infrastructure, the mortality rate is 7-8 times higher than that in developed countries.

To compensate for these problems, this study proposed a diagnostic aid tool using Yolo V3, a deep learning-based object detection algorithm that can be applied to a single board computer, which is relatively easy to distribute due to low cost.

A total of 5,631 cell images were obtained to train the model, of which 2,827 normal cells and 2,804 abnormal cells were labeled by a pathologist. Moreover, as a data pre-processing process, the RGB scale was converted to gray-scale, and a 2-Dimensional Gaussian Filter was used to remove noise.

To apply fine tuning, a method of transfer learning, the weight from the Yolo V3 model trained with the COCO data set was used as the initial weight.

The results of the training showed that the mean average precision was 73.34 %, which was high, and the mean accuracy was 86.45 %.

In this study, despite the small number of data, Yolo V3 showed satisfactory performance. As Yolo V3 was proven to be applied to a single board

computer with its lightweight algorithm, it seemed possible to develop an auxiliary tool for automatic diagnosis of cervical cancer at low cost to compensate for the high fatigue experienced by pathologists and to allow utilization in developing countries with relatively poor medical infrastructure. In addition, Yolo V3 has demonstrated smooth learning of pathological images, suggesting the possibility for application in the diagnosis of various other diseases.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A1B04031182/ NRF2020R1I1A1A01072101).

References

[1]. J. H. Park, The Usefulness of Cervicovaginal Cytology as a Primary Screening Test, *The Korean Journal of Cytopathology*, Vol. 19, Issue 2, 2008, pp.107-110.

[2]. Y. T. Kim, Causes and Diagnoses of Cervical Cancer, *Korean Med Assoc*, Vol. 50, Issue 9, 2007, pp. 769-777.

[3]. T. Kanungo, D. M. Mount, N. Netanyahu, C. Piatko, R. Silverman, A. Y. Wu, An efficient k-means clustering algorithm, Analysis and implementation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, Issue 7, 2002, pp. 881-892.

[4]. W. Chen, M. J. Er, S. Wu, PCA and LDA in DCT domain, *Pattern Recognition Letters*, Vol. 26, Issue 15, 2005, pp. 2474-2482.

[5]. N. A Mat-Isa, An automated cervical pre-cancerous diagnostic system, *Artificial Intelligence in Medicine*, Vol. 42, Issue 1, 2008, pp.1-11.

[6]. B. Bosi, G. Bois, Y. Savaria, Reconfigurable pipelined 2-D convolvers for fast digital signal processing, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 7, Issue 3, pp. 299-308.

[7]. J. Redmon, A. Farhadi, Yolov3: An incremental improvement, *arXiv:1804.02767, Computer Vision and Pattern Recognition*, 2018.

[8]. J. Redmon, Darknet: Open source neural networks in c, http://pjreddie.com/darknet/, 2013–2016.

[9]. K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, USA, 2016, pp. 770–778.

[10]. M. Everingham, L.Van Gool, C. K. I Williams, J. Winn, A. Zisserman, The PASCAL Visual Object Classes (VOC) Challenge, *IJCV*, Vol. 88, 2010, pp. 303-338.