Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 114 (2017) 281–287

Procedia Computer Science

www.elsevier.com/locate/procedia

Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems, CAS October 30 – November 1, 2017, Chicago, Illinois, USA

Convolutional Neural Network Based Localized Classification of Uterine Cervical Cancer Digital Histology Images.

Haidar A. Almubarak[a], R. Joe Stanley[a]*, Rodney Long[b], Sameer Antani[b], George Thoma[b], Rosemary Zuna[c], Shelliane R. Frazier[d]

[a]*Department of Electrical and Computer Engineering, Missouri University of Science and Technology, Rolla, MO USA 65401*
[b]*Lister Hill Nationa Center for Biomedical Communications for National Library of Medicine, National Institutes of Health, DHHS, Bethesda, MD, USA*
[c] *Department of Pathology for the University of Oklahoma Health Sciences Center, Oklahoma City, OK 73117, USA*
[d] *Surgical Pathology Department for the University of Missouri Hospitals and Clinics, Columbia, MO 65202, USA*

Abstract

In previous research, we introduced an automated localized, fusion-based algorithm to classify squamous epithelium into Normal, CIN1, CIN2, and CIN3 grades of cervical intraepithelial neoplasia (CIN). The approach partitioned the epithelium into 10 segments. Image processing and machine vision algorithms were used to extract features from each segment. The features were then used to classify the segment and the result was fused to classify the whole epithelium. This research extends the previous research by dividing each of the 10 segments into 3 parts and uses a convolutional neural network to classify the 3 parts. The result is then fused to classify the segments and the whole epithelium. The experimental data consists of 65 images. The proposed method accuracy is 77.25% compared to 75.75% using the previous method for the same dataset.

Peer-review under responsibility of the scientific committee of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.

Keywords: Cervical Cancer; Convolutoin Neural Networks; Data Fusion; Image classifcation; Clinical Decision Support Systems.

* Corresponding author. Tel.: +1-573-341-6896.
E-mail address: stanleyj@mst.edu

1877-0509
Peer-review under responsibility of the scientific committee of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.
10.1016/j.procs.2017.09.044

1. Introduction

Cervical cancer is the second leading cause of cancer death in women aged 20 to 39 years, in 2017 an estimate of 12,820 new cases and 4,210 is reported [1]. Screening for cervical cancer and its precursor lesions is carried out using a Papanicolaou (Pap) test. Biopsied cervical tissue histology slides are used to give a definitive evaluation; interpretation of these slides is done by an expert pathologist [2]. Pathologists seek to detect cervical intraepithelial neoplasia (CIN), which is a pre-malignant condition for cervical cancer. A cervical biopsy is classified as normal (no CIN lesion) or one of three CIN grades: CIN1 (mild dysplasia), CIN2 (moderate dysplasia), or CIN3 (severe dysplasia) by identifying the atypical cells in the epithelium by the visual inspection of histology slides [3]. Fig. 1 shows an example of different CIN grades. Delayed maturation with an increase in immature atypical cells from bottom to top of the epithelium has been observed as CIN increases in severity [4]. Computer-assisted CIN diagnosis has been studied previously in [5]–[10]; in these studies, manually handcrafted features need to be extracted using various image processing and machine learning algorithms which are time-consuming and may not be the best features to be used.

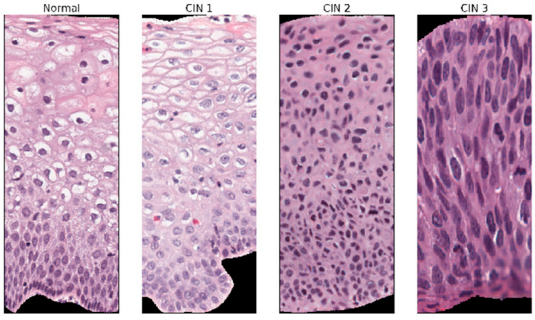

Fig. 1 CIN Grades

Convolutional neural networks (ConvNets) proved to be great in many images related domains such as image classification for very large scale datasets like ImageNet [11], face recognition [12], and breast cancer mitosis detection [13]. ConvNets does not need features to be extracted manually, instead, they use filters and convolve them with the images to extract the features. The filters are updated and tuned during the training process. In previous research, our research group used a localized fusion-based approach for CIN grade classification [6]; this localized approach divides an epithelium image into 10 segments then extract features from each segment; the futures are used to train a classifier which will classify each segment into one of the CIN grades. After classifying the segments, the whole image class is determined by voting among the 10 segments.

This research extends the localized fusion-based approach by further subdividing each segment into 3 parts: top, middle, and bottom. This division is trying to exploit the fact that the abnormality in the cells progresses from the bottom to the top of the epithelium, and analyzing the 3 parts separately then fuse the results shall improve the classification results. In this research, ConvNets will be used for feature extraction and initial classification, there will be no manually crafted features from the image or image parts. Other classification algorithms such as support vector machine, logistic regression, and random forests will be used to fuse the 3 parts and segments result to get the whole image class.

2. Methodology

The steps for processing an epithelium image for CIN classification is given as follows:

- Divide the whole image into 10 segments.
- Divide each segment into 3 parts: top, middle, bottom.
- Extract 32x32 patches (chunks) from each part.