RESEARCH

Features-driven convolutional neural networks for cervical cell classification

Santiago Tello-Mijares[1,2]*, Jesús Bescós[2] and Francisco Flores[3]

*Correspondence:
jtello@itslerdo.edu.mx
[1]Mechatronics Postgraduate
Department, Instituto Tecnológico
Superior de Lerdo, Av.
Tecnológico, 1555, 35150 Lerdo,
México
Full list of author information is
available at the end of the article

Abstract

Background: We propose a cervical cell (CC) image classification method using a combination of classical features (size, color, etc.) with features extracted from the last layer from a convolutional neuronal network (CNN), hereinafter CNN features.

Methods: On one side, after CC segmentation into the nucleus and cytoplasm, we extract the set of classical features for every input image following well known procedures. On the other side, in the absence of the large CC images database traditionally required by CNNs to succeed, we guide the extraction of CNN features from the Herlev dataset (just 819 images) following a novel procedure: classical features are used to generate a synthetic image for every input image; these synthetic images, capturing the essence of the cells, are the ones used for training and testing the CNN.

Results: Our experiments comparing both sets of features and different classification schemes conclude that the best configuration, which achieves 100% precision and 98.61% recall in normal/abnormal CC classification (far superior to the state-of-the-art), consists of combining classical and CNN features.

Conclusion: Classifying cervical cell in Pap-smear images is generally accepted as a prerequisite to helping in the diagnosis of uterine cervical cancer. Our main contributions were the novel use of a classical-CNN features and the evaluation of several classification strategies to tackle this problem. We have thoroughly analyzed reported works on the topic and discussed our results in light of these works. This analysis reasonably validates the novelty and quality of our achievements in the absence of a fully fair framework to objectively compare our results.

Keywords: Image processing; cervical cells imaging; convolutional neural network

Background

The analysis of cervical cell (CC) images in terms of the detection of nucleus and cytoplasm deformation provides important diagnosis clues to prevent cervical cancer. CC images can be classified into seven classes[1] that differ in terms of the size and shape of the nuclei and cytoplasm: normal columnar epithelium (cyl), normal intermediate squamous (inter), normal superficial squamous (super), normal squamous non-keratinizing light dysplastic (let), abnormal squamous non-keratinizing moderate dysplastic (mod), abnormal squamous non-keratinizing severe dysplastic (svar) and abnormal carcinoma in situ (carci). However, as a minimal requirement for diagnosis is to separate normal from abnormal cells[2, 3], we have adopted the Bethesda System terminology[4], which classifies squamous intraepithelial lesion

(SIL) into two classes (see Fig. 1): normal or low grade (L-SIL), including inter and super classes; and abnormal or high grade (H-SIL), including let, mod, svar, and carci classes. Cyl cells are classified as neither normal nor abnormal[5, 6].

We use the Herlev dataset provided in [1] (from [7]) to perform tests for the selection of several classification parameters and to conduct the final evaluation. This dataset includes 819 Pap smear RGB images—0.201 mum per pixel, an average of 156x140 pixels, and 40x magnification—, with the following breakdown: L-SIL=144 (inter=70, super=74) and H-SIL=675 (let=182, mod=146, svar=197, and carci=150).

The Herlev dataset has been extensively used to test CC images classification into two classes [2, 3, 5, 6, 8, 9] (see Table 1): the best results, reported by Dounias et al.[8], applied second-order neural networks (ANNs) with a 10-fold cross validation; Tsakonas et al.[3] used genetic programming (GP); and Marinakis et al.[2] introduced particle swarm optimization (PSO) for feature selection and the one-nearest neighbor (1nn) classifier to optimize the patterns recognition. More recent works include: Nanni et al.[9], that used a support vector machine (SVM) on textures described via local binary patterns (LBP); Chankong et al.[6], which presented an automated approach to a CC classification method that uses an ANNs and a catalogue of six nucleus features and three nucleus-cytoplasm features (nine in total); and Genctav et al.[5], that described each cell using 14 different features[1] for the classification of two-class CCs using the optimal leaf ordering algorithm on the binary tree (BT) obtained by hierarchical clustering.

We here present an effective and efficient combination of cell features: a set of classical features, similar to those reported in [2, 3, 5, 6, 8, 9]; and a set of features resulting from the last layer of a CNN, say CNN features (Fig. 1). Originally proposed by LeCun et al.[10], a CNN is a neural network model with three key architectural ideas: local receptive fields, weight sharing, and sub-sampling in the spatial domain. A CNN consists of three main types of layers designed to obtain the feature maps: spatial convolution layers (Cl), subsampling pooling layers (Sl), and fully connected layers (Fl)—l is a layer index (1-6)—used to obtain the low-level, mid-level, and high-level features (Fig. 1).

We use a CNN based on previous studies designed to process two-dimensional (2-D) images[11, 12]. The Cl and Sl layers are 2-D layers, whereas the Fl layer and output are 1-D layers. The motivation is that a CNN is advantageous for features extraction in that it is hierarchical (with multiple layers for more compactness and efficiency) and invariance-redundant (for position, size, luminance, rotation, pose-angle, noise, and distortion).

CNNs have recently been used in several applications, including hand-written digit recognition, face detection, and face recognition; a comprehensive review of CNN can be found in [13]. We propose an efficient method for extracting CC classification features from the last layer of a CNN. We demonstrate that a classification method using the combination of classical features and these CNN features is more robust and efficient than conventional state-of-the-art (SoA) methods using only classical features. Our main contributions are: the novel combination of CNN features and classical features from the cervical nucleus and cytoplasm; the generation of synthetic input images, capturing relevant cell information, to train and feed the CNN;

the generation of a representative dataset with ground-truth data; and the evaluation of several classification strategies to confront the classification problem—a tree random forest (TRF), a multilayer perceptron (MLP), and a Bayesian network (BN)—.

Implementation

Fig. 1 illustrates the general flow of the proposed cervical cell (CC) image classification method, which is a combination of classical features with CNN features.

Classical features extraction

Classical features extraction: After cell segmentation in the Lab color image, we extract the features from the segmented nucleus and cytoplasm (Fig. 1). We extract a 32 values feature-vector, including 16 values for the nucleus, 12 for the cytoplasm, and 4 more for the global cell. For the nucleus, 7 features describing shape (area, compactness, height, width, width/height, elongation, roundness), 2 features describing mean color and variance (3 Lab values each), and an additional feature for texture homogeneity (3 Lab values). For the cytoplasm, 3 features for shape (area, height, width), and the color and texture Lab features. For the global cell, 2 shape features describing cell area and compactness, and 2 more features measuring the relative area and intensity between the nucleus and the cytoplasm. All these features are then normalized for obtained the synthetic cell image.

CNN features extraction

CNNs usually operate over thousands of size-normalized input images, which requires previous adaptation of these images. Our main contribution in this phase is to replace this adaptation by the content-based generation of a synthetic cell image. For every variable-size input CC image we generate a fixed-size synthetic image capturing the semantics of the cell, by using classical features extracted from the nucleus and cytoplasm—width, height, and mean color—of the input image. We first rotate the input image to align vertically the nucleus directrix; then we fit the nucleus and the cytoplasm to respective ellipses that are then size-normalized according to maximum and minimum observations in the training set; and finally we color the ellipses with the respective mean Lab color features. These 32x32 Lab images feed the CNN (Fig. 1), which is designed to extract 16 3-dimensional vectors of high-level features.

Results

In order to classify CCs into two classes, we have performed exhaustive tests with three different classification techniques implemented in Weka (Waikato Environment for Knowledge Analysis)[14]: MLP, TRF, and BN. As TRF obtained, by far, the best results, we include here only results for this technique.

First, we compared the potential of both sets of features, testing them separately in a 2-fold setup (2 iterations, 50% training, 50% test). As shown in Fig. 2, the CNN features behave much worse than the classical ones, even with the proposed content-based adaptation of the input CC images, designed to compensate for the absence of a large database (the Herlev dataset just includes 819 images). However,

the CNN features do complement the classical features, so that their combination obtains the best results.

Secondly, we compared the proposed combination of classical and CNN features with results reported by other methods [2, 3, 5, 6, 8, 9] (Table 1). These methods also use the Herlev dataset but sometimes classify using more favorable cross-validation setups: 5-fold (5 iterations, 80% training, 20% test), 10-fold (10 iterations, 90% training, 10% test), or leave-one-out. We also include this information in order to help comparing. Table 1 shows that the proposed method outperforms results of previous works, which only consider classical features.

Conclusion

Both manual and semiautomatic methods for the analysis of Pap smear samples and timely diagnosis show three main problems that must be solved: the localization of cell nuclei (identification), the extraction of different features (extracting characteristics), and the classification as LSIL, HSIL, normal, and abnormal cells. We have presented a method that focuses on the use of CNNs to obtain a complementary set of features for classification, and, particularly, on a feature-driven technique to train and feed the CNN. Our study demonstrates that using an ensemble of classical features and CNN features from the nucleus and cytoplasm yields results far superior to the state-of-the-art when using small training datasets.

Author details

[1]Mechatronics Postgraduate Department, Instituto Tecnológico Superior de Lerdo, Av. Tecnológico, 1555, 35150 Lerdo, México. [2]Escuela Politécnica Superior, Video Processing and Understanding Lab, Universidad Autónoma de Madrid, Francisco Tomás and Valiente, 11, 28049 Madrid, Spain. [3]Postgraduate Department, Instituto Tecnológico de la Laguna, Blvd. Revolución and Av. Cuauhtemoc, 27000 Torreón, México.

References

1. Martin, E.: Pap-smear classification. Master's thesis, Technical University of Denmark, Oersted-DTU Automation (2003)
2. Y Marinakis, M.M., Dounias, G.: Particle swarm optimization for pap-smear diagnosis. Expert Systems with Applications **35**(4), 1645–1656 (2008)
3. A. Tsakonas, J.J.H.A.B.B. G. Dounias, von Keyserlingk, D.: Evolving rule-based systems in two medical domains using genetic programming. Artificial Intelligence in Medicine **32**(3), 195–216 (2004)
4. ME Plissiti, C.N., Charchanti, A.: Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering. EEE Transactions on information technology in biomedicine **15**(2), 233–241 (2011)
5. A Gençtav, S.A., Önder, S.: Unsupervised segmentation and classification of cervical cell images. Pattern recognition **45**(12), 4151–4168 (2012)
6. T Chankong, N.T.-U., Auephanwiriyakul, S.: Automatic cervical cell segmentation and classification in pap smears. Computer methods and programs in biomedicine **113**(2), 539–556 (2014)
7. Pap-smear (DTU/HERLEV) Databases and Related Studies. http://mde-lab.aegean.gr/index.php/downloads_links.html
8. G Dounias, J.J.A.T.N.A.G.P. B Bjerregaard, Panourgias, E.: Automated identification of cancerous smears using various competitive intelligent techniques. Oncology reports **15**(4), 1001–1006 (2006)
9. L Nanni, A.L., Brahnam, S.: Local binary patterns variants as texture descriptors for medical image analysis. Artificial intelligence in medicine **49**(2), 117–125 (2010)
10. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)

11. Phung, S., Bouzerdoum, A.: A pyramidal neural network for visual pattern recognition. IEEE Transactions on Neural Networks **18**(2), 329–343 (2007)
12. Vedaldi, A., Lenc, K.: Matconvnet: Convolutional neural networks for matlab. In: Proceedings of the 23rd ACM International Conference on Multimedia, pp. 689–692 (2015). ACM
13. LM Rasdi Rere, M.F., Arymurthy, A.: Metaheuristic algorithms for convolution neural network. Computational Intelligence and Neuroscience **2016** (2016)
14. Holmes, G., Donkin, A., Witten, I.H.: Weka: A machine learning workbench. In: Intelligent Information Systems, 1994. Proceedings of the 1994 Second Australian and New Zealand Conference On, pp. 357–361 (1994). IEEE

Figures

Figure 1 Overall method description. Extraction of classical features from L-SIL and H-SIL cells; generation of synthetic input images to train and test the CNN; combination of classical features and CNN features when testing with an H-SIL normal.

Figure 2 Results. Compared results using the TRF technique with classical features, with CNN features and with the aggregation of both sets of features.

Tables

Table 1 Quantitative comparison between the proposed method and those reported by other authors.

Ref	Tech	s-fold	TPR	PPV	ACC	AUC	rmse	k
PM	TRF	s=10	0.993	1	0.998	1	0.021	0.995
		s=5	0.986	1	0.997	1	0.039	0.991
		s=2	0.986	1	0.997	1	0.054	0.991
[6]	ANN	s=loo			0.992			0.97
[5]	BT	s=2						0.84
[10]	SVM	s=5				0.88		
[2]	PSO	s=10					0.12	
		s=5					0.17	
		s=2					0.22	
[9]	ANN	s=10			0.99			
[3]	GP	s=2	0.928					

ACC: accuracy; ANN: Artificial Neuronal Networks; AUC: area under the ROC; BT: Binary tree; GP: Genetic programing; k: Kappa statistic; loo: Leave one out; PM: proposed method; PPV: precision or positive predictive value; PSO: Particle swarm optimization; Ref: reference; rmse: root mean squared error; SVM: Support vector machine; Tech: technique; TPR: sensitivity, recall or true positive rate; TRF: Tree random forest.

Additional Files
Additional file 1 — Database
The entire cervical cells images database (class 1-2, H-SIL and L-SIL) and the ground-truth (as binary jpg images), both for obtained the synthetic cell images.

Additional file 2 — Interface
Classical/CNN-TRF method (as Matlab interface).

Additional file 3 — Weka Files
Weka features and associated class for test and train experiments (as ARFF Data File) and Weka TRF, MLP, BN models for final results.