Computers in Biology and Medicine 111 (2019) 103328

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Detecting mitotic cells in HEp-2 images as anomalies via one class classifier

Krati Gupta*, Arnav Bhavsar, Anil K. Sao

School of Computing & Electrical Engineering, Indian Institute of Technology Mandi, Mandi, Himachal Pradesh, India

ARTICLE INFO

Keywords:
Auto-immune disorders
HEp-2 cells
Mitotic cells
One class classifiers
Support vector machines

ABSTRACT

We propose a novel framework for classification of mitotic v/s non-mitotic cells in a Computer Aided Diagnosis (CAD) system for Anti-Nuclear Antibodies (ANA) detection. In the proposed work, due to unique characteristics (the rare occurrence) of the mitotic cells, their identification is posed as an anomaly detection approach. This will resolve the issue of data imbalance, which can arise in the traditional binary classification paradigm for mitotic v/s non-mitotic cell image classification. Here, the characteristics of only non-mitotic/interphase cells are captured using a well-defined feature representation to characterize the non-mitotic class distribution well, and the mitotic class is posed as an anomalous class. This framework requires training data only for the majority (non-mitotic) class, to build the classification model. The feature representation of the non-mitotic class includes morphology, texture, and Convolutional Neural Network (CNN) based feature representations, coupled with Bag-of-Words (BoW) and Spatial Pyramid Pooling (SPP) based summarization techniques. For classification, in this work, we employ the One-Class Support Vector Machines (OC-SVM).

The proposed classification framework is validated on a publicly available dataset, and across various experiments, we demonstrate comparable or better performance over binary classification, attaining 0.99 (max.) F-Score in one case. The proposed framework proves to be an effective way to solve the mentioned problem statement, where there are less number of samples in one of the classes.

1. Introduction

In autoimmune disorders, the anti/auto-nuclear antibodies (ANA) combine with healthy cells and nuclear organelles and manifest specific kinds of staining patterns, representing autoimmune diseases [1–3]. For diagnosis of this disorder, Indirect ImmunoFluorescence (IIF) imaging [2,4], is still considered as a 'gold standard' test [1,3,5,6], wherein the patient blood serum is added to a specific cell substrate, consisting of Human Epithelial (HEp-2) cells and the ANA manifest as different nuclear staining patterns, which can be observed under the microscope. In practice, pathologists manually identify patterns. However, the process of manual identification can have some limitations [2,7], the assessment can suffer with intra and inter-observer variability and the task can be inefficient and labor intensive [3,8]. In this regard, computer-aided techniques involving image-based classification, aim to detect important biological characteristics of patterns, in order to assist the decision process [2].

Due to the use of cell substrate, the nuclear staining patterns can be visualized in two broad phases of the cell cycle (a series of phases taking place in a cell, leading to its cell division [9]), i.e., interphase and mitotic phase [10]. Fig. 1 depicts a few samples of mitotic and non-mitotic patterns. In this work, we focus on the task of discriminating between mitotic and interphase cells, as the identification of the mitotic cells is important, due to the following two reasons: (a) the presence of at least one mitotic pattern per whole slide image (WSI) indicates the good preparation of the sample slide [11,12] and (b) in a specimen image with mixed patterns, i.e., consisting of both mitotic and interphase patterns, the specific type of mitotic pattern can co-exist with other interphase patterns [11] and the observation of mitotic cells can help in discrimination of some similar interphase staining patterns [12].

The detection of mitotic phase patterns is a challenging task due to the following reasons: (a) As the mitosis is a transition phase in a cell cycle process, very few cells in a WSI show such behavior, hence the occurrence of mitotic patterns is typically rare. In this case, there arises a large data imbalance between mitotic and non-mitotic samples. (b) Additionally, the mitotic cells are present in a variety of shapes, hence show a large intra-class variations. The mitotic stage itself involves four different phases: prophase, metaphase, anaphase and telophase [13] (refer Fig. 1). The nuclear patterns can be visualized in any of the prominent phase of mitotic stage.

Till now, there have been several attempts to address the classification problem of interphase patterns [3,7,14–18]. However, few works

* Corresponding author. MANAS Lab, School of Computing & Electrical Engineering, Indian Institute of Technology Mandi, Mandi, Himachal Pradesh, 175005, India.

E-mail addresses: krati_gupta@students.iitmandi.ac.in (K. Gupta), arnav@iitmandi.ac.in (A. Bhavsar), anil@iitmandi.ac.in (A.K. Sao).

https://doi.org/10.1016/j.compbiomed.2019.103328
Received 12 December 2018; Received in revised form 12 June 2019; Accepted 12 June 2019

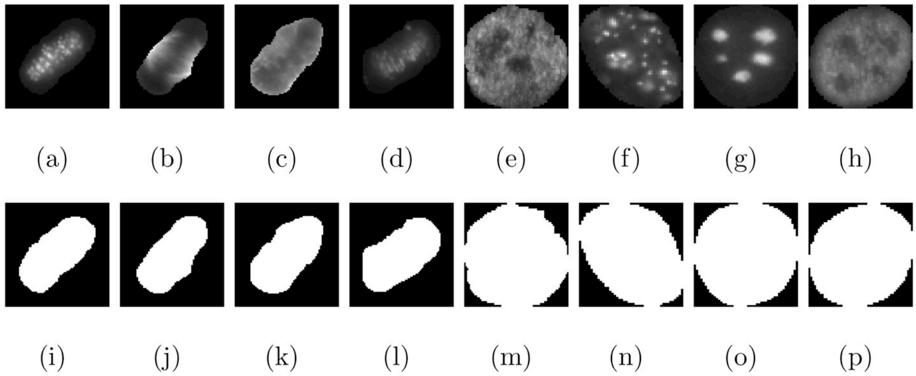


Fig. 1. Few cell images examples (a) to (d) mitotic cells & (e) to (h) non-mitotic cells, along with their DAPI masks in second row.

have attempted the mitotic cell detection for autoimmune disorders. In some of the works [11,12,19,20], authors have reported approaches for mitotic v/s non-mitotic cell classification based on morphological and textural features with the use of data augmentation techniques. However, in these techniques, generated examples from the available samples may not be able to approximate the sample distribution well, given that the actual number of mitotic cells are very less and generally lead to over and underfitting of samples. Hence, the present work focuses on building a classification framework with a novel viewpoint of considering mitotic cells as an anomaly, which in turn implies that we require only the abundantly available non-mitotic cell images for training. The important contributory aspects of the proposed framework are as follows:

(I) **Data imbalance:** The issue of data imbalance in a CAD classifier between mitotic and non-mitotic samples is crucial as there exists only 10–15% mitotic cells in a WSI [11,21]. Hence, we propose that the mitotic cell recognition problem can be treated as an anomaly detection task, where the mitotic patterns do not follow an expected pattern in a data or feature space consisting of non-mitotic cells. The One-Class Classifier (OCC) is used for anomaly detection, in which the model is trained with only the non-anomaly class (non-mitotic) samples, and then the mitotic samples are predicted as the ones which are outliers based on the model.

(II) **Majority class-specific feature representation:** The class-specific feature representation for the samples should be chosen, in order to characterize the majority class distribution well, so that the mitotic samples can be detected as an anomalous class. Here, such representation can be effectively captured using some morphology & texture-based feature representation.

Morphology based features: There is a slight but noticeable shrinkage of nuclear and cellular structures at mitotic stage [13]. This shrinkage is not observed in non-mitotic samples. Such cytological characteristics of these cells can be better consolidated using morphology (shape)-based feature representation.

Traditional texture based features: Due to some chromosomal aberrations, a prominent texture distinction can also be noted between both types of cell categories. Hence, we focus on using a filter bank based feature representation, to get a well defined texture distribution for non-mitotic samples.

CNN based features: The cells associated with both classes may vary in shape, size, and appearance and these characteristics of majority class are quite distinct from the mitotic characteristics. This behavior is captured using the multilayer Convolutional Neural Network (CNN). These learned features are demonstrated to contribute towards good discriminatory feature representation, even when using a CNN which is pre-trained on scene/object images. Here, we use the pre-trained AlexNet architecture [22].

(III) **Requirement of segmentation masks:** The Region Of Interest (ROI) is extracted using the DAPI (*4′,6-diamidino-2-phenylindole*) pattern of cell images, which is generally used to demarcate the nuclear and cellular boundaries to get binary masks (ground truth) of cells (also shown in Fig. 1(i) to 1(p)) [3]. These ground truths (GT) ROIs are used only to extract traditional morphological & filter-based features. The CNN based features are extracted without using DAPI masks, from given whole cell images, as the convolution operation is typically applied on whole images in a network. Thus, the CNN-based feature representation avoids the requirement of accurate segmentation masks in the framework, as the acquisition of DAPI mask is an extensive process [23] and it also shows carcinogenic properties [24].

Thus, in summary, the novelty of the strategy is in the notion of considering the minority mitotic class as an anomaly and applying an OCC for classification, using feature-representation involving traditional filter-banks, coupled with an effective morphology-based feature representation, as well as CNN-based features. Note that, other types of OCC may also be explored in such a scenario, where the training samples are derived only from one class. The proposed framework better inculcates all the discriminative characteristics of classes and its effectiveness is reflected in terms of classification performance. Fig. 2 depicts the proposed flow to address the problem.

The paper is organized in the following sections: the next section, Section 2 explains about the literature for general anti-nuclear antibody detection problem and its clinical aspects. Section 3 explains the proposed approach, consisting of feature representation, including preprocessing and classification schemes. Section 4 explains the experimental results and analysis, consisting of data description, evaluation protocols, results and comparisons with other classifiers and prior approaches, followed by the conclusion section. The summary and discussion constitute the last section of the manuscript.

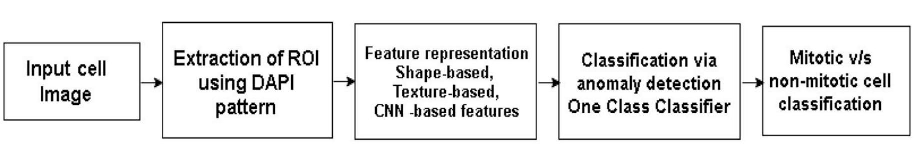


Fig. 2. Proposed flow for the mitotic cell recognition approach.

2. Literature review

Anti-nuclear antibody detection is considered to be an important and challenging task for the diagnosis of autoimmune disorders. There is an extensive demand for accurate and precise strategies for classification of staining patterns, visualized on cell substrate, in order to identify the relevant disorders. The identification of the mitotic patterns enhances the difficulties in this detection process, due to their rare appearance. A detailed clinically oriented discussion of the IIF technique and ANA can be found in Refs. [1,8,25–27].

Over the last decade, many researchers have attempted methods to address the problem of HEp-2 single cell and specimen-based interphase type staining patterns classification [28–30]. The problem is explored using hand-engineered features such as shape [7,31], scale & rotational invariant descriptors [32], statistical & textural descriptors [33], and using different classifiers, such as support vector machines [6,7,14,17,18,33], decision tree [28], nearest neighbor [32], and multi-expert systems [29], and deep neural networks [15,16]. The approaches involving specimen images, are based on cell, patch and pixel-based processing [14,18,34], and use learned feature representation and classification using standard CNNs [35,36]. Such frameworks are validated over some benchmarking datasets such as ICPR2012Contest dataset [2,37] or MIVIA dataset [2,37], SNPHEp-2 [38], ICIP2013 [6], ICIP2014 [23] and I3A-Task2 specimen dataset [23]. For the mitotic cell identification problem, some works have been reported in context of cancer diagnosis [13,39,40], but not for HEp-2 images.

In this scenario, there are few works, which consider the classification problem between mitotic and non-mitotic single cells. For example, in Refs. [11,12], authors proposed a classification framework on their own dataset, consisting of relatively less number of samples, as compared to the dataset that we have used. In these, the authors used morphological and textural descriptors. In Ref. [19], the authors have addressed the issues of sample skew, using standard sample balancing methods, while the authors in Ref. [12] use hybrid techniques to address the same, wherein the authors use different data skew balancing techniques along with different classifiers, such as synthetic minority oversampling techniques and multi-expert systems. The authors of [21] have released a benchmark for the mitotic v/s non-mitotic classification problem, with another dataset consisting of relatively large number of the mitotic samples. In Ref. [10], the authors proposed an unsupervised framework to detect the mitotic patterns. We highlight that the works [10,12,19] use dataset with 70 mitotic and 1457 non-mitotic cells, which is not similar to the I3A-Task3 (also known as UQSNP_HEp-2 Task3) dataset, which we use in our study. We note that no other authors have used this publicly available I3A-Task3 dataset (https://outbox.eait.uq.edu.au/uqawilie/UQSNP_HEp2_datasets/).

Here, we propose to use a framework to identify the mitotic cells via an anomaly detection strategy. For this task, we tend to use a Support Vector Machine based One Class Classifier (OCC), typically known as OC-SVM [41], which has been extensively used in several classification tasks, such as network intrusion detection [42,43], anomaly detection using pay-load data [44], medical applications [45] etc. To the best of our knowledge, we did not find any work related to the use of anomaly detection strategy via OC-SVM to address the classification between mitotic and non-mitotic samples.

3. Proposed approach

As mentioned earlier, we propose to apply a One-Class Classifier (OCC) based anomaly detection technique, that uses the One Class Support Vector Machines (SVM). It is contrary to the approaches that use a traditional classification paradigm, wherein data from both the classes are used for training. Here, the classification models are built using only the abundant non-mitotic samples for training, and mitotic are considered as an anomaly.

The framework can be broadly divided into two steps: Feature representation & classification via anomaly detection, i.e., OCC framework. The detailed description of the proposed approach is as follows.

3.1. Feature representation scheme

Firstly, the pre-processing of given images is specifically required for contrast enhancement of low images and do not exhibit prominent discrimination from other class samples. For pre-processing, we have used the well-known γ-transform with $\gamma = 0.4$ [46] on normalized images. Here, images are normalized, by making the intensity range between 0 and 1.

For feature representation, we consider both traditional and contemporary features. Among the traditional features, the morphological features are defined in a relatively simplistic manner and hence help in more efficient processing. For texture-based features, we use two popular traditional generic filter banks, and the contemporary CNN based filters as data-dependent feature representation. The details of the features used are as follows:

(A) **Morphology-based features:** As already mentioned, due to the nuclear and cellular shrinkage during the mitosis process, the morphology-based descriptors can be used for better discrimination between mitotic and non-mitotic cell patterns. These are defined as:
- **Nuclear Area:** The nuclear area considers the total number of pixels present inside the nuclear regions. It is inferred that ideally, the nuclear area of the mitotic cells should be less than the non-mitotic ones, due to the chromosomal shrinkage inside the cells [13].
- **Nuclear Perimeter:** It defines the total number of pixels present in the periphery of the nuclear region. Note that all the morphology features are extracted from the normalized images with respect to the image sizes.

(B) **Texture-based features:** Texture-based features are motivated by the fact that there are clear textural differences between mitotic and non-mitotic cells, due to the chromosomal aberrations and deformations [13]. In the current work, we focus on using two different types of filter bank representations, that cover an appropriate range of parameters & orientations, defined as follows:
- **LM filter bank:** The LM filter bank [47] is a multi-orientation & multi-scale set of filters, consisting of a first & second derivative of Gaussian, few Laplacian of Gaussian (LoG) and 4 Gaussians, that makes a total of 84 filters. Hence, the complete filter bank yields 84-dimensional feature representation for each pixel of an image. The algorithm uses filters of size 9×9.
- **Gabor filter bank:** The Gabor filter bank [48] is a bank of 2D Gabor filters. It includes 108 filters with different frequencies of sinusoids and orientations. The final feature representation is 108-dimensional representation for each pixel in the images. Here too, the filter size is the 9×9.

Due to the specific characteristics of Gabor filter bank, it captures additional information (such as spatial frequencies), which is ignored by Gaussian-based LM filters. On the other hand, LM filters capture first and second order gradient information at various scales. Hence, both the filter banks are chosen to implement in this work.

(C) **CNN-based feature representation:** The motivation behind using the CNN is to use the learned filter responses, in which the output response of each layer are used as a data-dependent feature representation. These features are captured from the images at different levels of abstraction. Another important advantage of using such features is to avoid the requirement of DAPI patterns as the acquisition of DAPI is itself an expensive and time-consuming protocol [23].

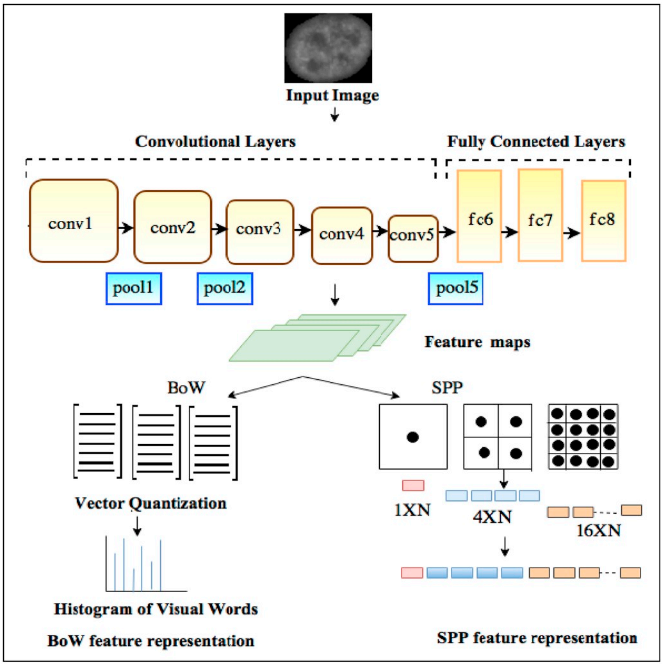


Fig. 3. Demonstration of feature representation scheme using CNN.

CNN architecture: Here, we use a CNN architecture, i.e., AlexNet [22], pre-trained on 1.3 million images in the LSVRC-2010 ImageNet training set, divided into 1000 classes. Fig. 3 shows the structure of the network, consisting of a total of 5 convolutional layers. At each convolutional layer, Rectified Linear Unit (ReLU) is used as a non-linear activation function.

Here, we analyze the responses acquired from different five convolutional layers. Note that, more contemporary CNN models could have been considered for this work. However, we have chosen a relatively simpler AlexNet architecture, as the motivation is mainly to analyze the efficacy of features at different levels of abstraction of the network. Even with the AlexNet model, we demonstrate sufficiently good quality results. Similar to the above traditional filter bank based feature representation, at each layer of CNN, the feature representation depends on the number of filters in each layer. For e.g., in the last convolution layer of AlexNet, there are 256 filters. Hence, the feature representation would be 256-dimensional for each pixel in the image.

Note that the LM and Gabor filter banks are also filter-based representation. However, the CNN filters are learned, unlike the traditional ones. The CNN filters are also more in number and arguably capture a larger variations. The features from the LM, Gabor and convolutional filter banks are represented in the form of a set of feature representations, with the variable number of filters in each case. Thus, the feature set is large, as the feature vectors are extracted from each pixel of the image. Hence, for compact representation of features, we use two well-known approaches, defined below.

Bag of Words (BoW) representation [49]**:** Here, the filter responses from all the training images are clustered and cluster centers are considered as the visual words. The BoW representation for any image is then computed as a vector of the frequency of occurrences of visual words, wherein locally computed features are binned across the visual words. Hence, the set of feature vectors of each image is mapped as a fixed and low dimensional histogram of visual words. The resultant histograms will be considered as new feature and are used to train the classification model. Here, we empirically fix the number of clusters to 64 for building the histogram representation, hence the feature representation is 64-dimensional for each image.

Spatial Pyramid Pooling (SPP) representation [50]**:** SPP is applied on the image feature maps, extracted through different layers of CNN and traditional filter bank responses. Each feature map is divided into 4 and 16 equal sub-regions and a mean pooling operator is applied on each sub-region. The extracted means from each sub-regions are concatenated to build a feature vector. The first value of the pyramid is the mean, extracted from the overall feature map. Finally, a 21 $(1 + 4+16)$ dimensional feature representation is acquired, for each image.

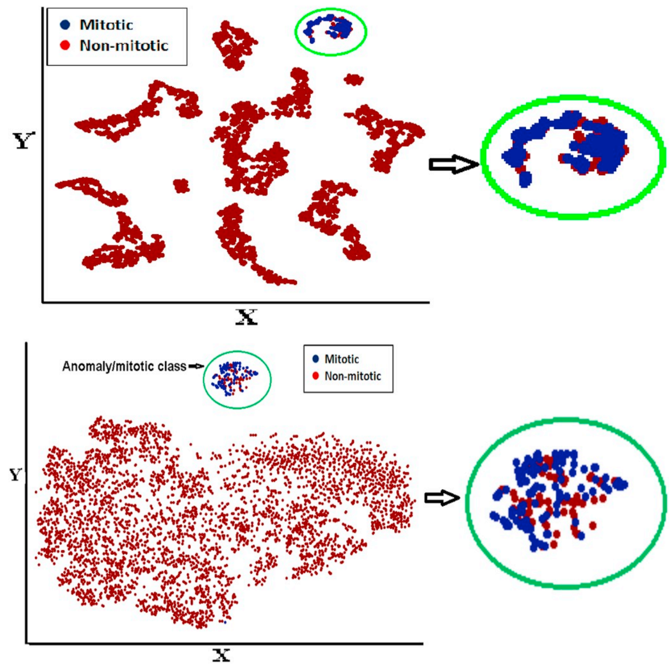


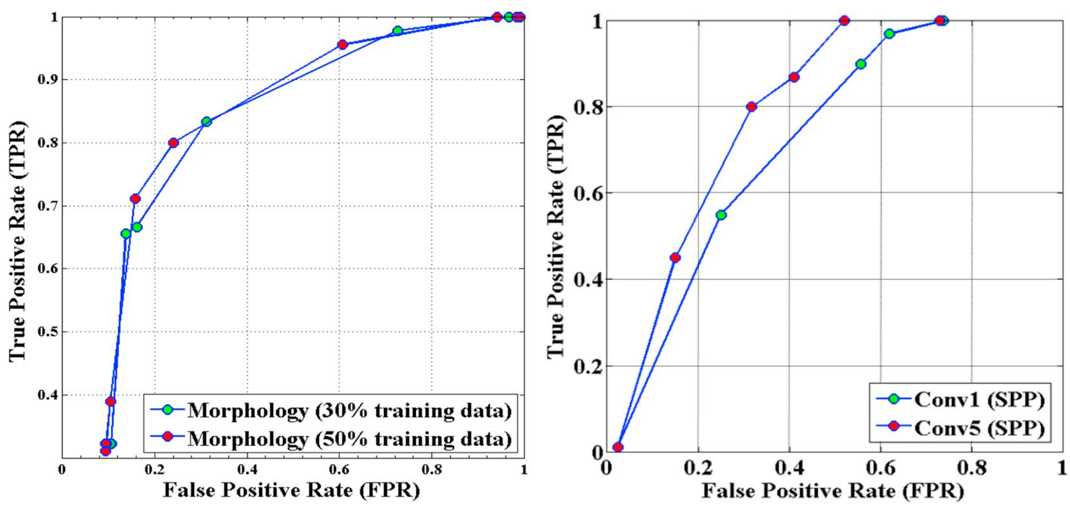
Fig. 4. 2D t-SNE plot of traditional (upper: Gabor filter bank) and CNN (lower: Conv2) feature representation. The anomaly class is represented in circle in both the plots (Better viewed in color) and the zoomed version of circles are shown separately along with the plots. Here X and Y are the two virtual feature dimensions (acquired from t-SNE technique) of high-dimensional feature representation (in reduced feature space).

In Fig. 3, the proposed flow of such feature representation scheme is demonstrated. In the figure, just as an example, only the CNN architecture is shown as a feature extractor. However, in this work, the BoW and SPP are applied on CNN as well as on traditional filter banks based responses, where we get a set of feature vectors for each image. For morphology-based features, since only two scalar-valued features are used, the feature represent is just 2-dimensional and only one such 2-dimensional feature is captured for each image. Hence, the BoW and SPP approaches need not be applied in this case.

To provide an indication of the mitotic class being considered as an anomaly, Fig. 4 shows an example t-SNE representation of filter responses, acquired from Gabor filter bank responses and Conv2 layer of CNN. We only use these two feature representations, just to demonstrate the behavior of the mitotic class distribution, being chosen as an anomaly class. The t-SNE [51] technique is used to visualize a high dimensional representation into lower virtual dimensions. Here, we notice that the mitotic samples indeed seem distinct from the majority non-mitotic samples, as they demonstrate a different behavior in feature space (also marked in a circle). Thus, the mitotic samples can be justified as anomalous examples. Note that, the t-SNE is just given to show the behavior of feature representations and the classification is done in actual feature space, not in reduced feature space.

3.2. One-class classifier (OCC): One-Class Support Vector Machines (OC-SVM)

We validate the proposed framework using One-Class Support Vector Machine (OC-SVM) [41] model as OCC, for anomaly detection.



(a) (b)

Fig. 5. Receiver Operating Curve (ROC) for (a) selected morphology/shape based features and (b) selected CNN based features.

Generally, one class models are built on majority class, i.e., non-mitotic class, in this case, and the mitotic class is treated as an anomaly. The model infers the characteristics of non-mitotic cases and from these characteristics, it can predict the test samples, which do not fit the trained model well. Hence, the non-mitotic class is considered as 'positive class' here. Such OCC's are often used for unbalanced datasets [52], unlike Binary Class Support Vector Machines (BC-SVM), wherein both class samples are used for training purpose. As OC-SVM is not as well-known as BC-SVM [53], we provide a brief description below.

In this approach, the task is to estimate a function f that takes the value as "+1" in a compact region (for majority class) which encompasses most of the data points, and "-1" elsewhere [54]. This strategy maps the input training data into kernel space and in that space, the feature representation of majority (positive) class will lie in a compact cluster, following some distribution. Our task is to estimate a maximally separated hyperplane to include most of the positive class samples on one side. In this way, the samples residing in other side can be considered as anomalous samples.

Let $x_1, x_2, \dots, x_n$ are the training examples of one class, i.e., the non-mitotic class X, where X is any subset of R^N, and φ be a kernel map $X \rightarrow F$, i.e, mapping of X into a dot product space F and the dot product can be computed using some kernel

$$k(\mathbf{x}, \mathbf{y}) = (\varphi(\mathbf{x}), \varphi(\mathbf{y})), \tag{1}$$

Then, the following objective function is defined [54]:

$$\min_{w,\rho,\varepsilon} \frac{\|w\|^2}{2} + \frac{1}{n\nu} \sum_{i=1}^{n} \varepsilon_i - \rho \tag{2}$$

$$\textit{subject to}: (w.\ \varphi(\mathbf{x}_i)) \geq \rho - \varepsilon_i \textit{ for all } i = 1..n \tag{3}$$

$$\varepsilon_i \geq 0 \textit{ for all } i = 1..n \tag{4}$$

The constructed hyperplane determines the margin between the classes. Here, w is a normal vector to the hyperplane. ε_i is slack variable, that is used to prevent over-fitting and to allow some data point to lie within the margin. As the nonzero slack variables ε_i are penalized in the objective function, the decision function,

$$f(x) = \text{sign}(w.\ \varphi(\mathbf{x}) - \rho) \tag{5}$$

will be positive for most examples of the training set. Hence, similar to the BC-SVM, here too, the primary goal is the minimization of $\|w\|$, using ν as controlling parameter.

Eq. (3) is somewhat similar to the equation used in BC-SVM [53]. However, in BC-SVM, the labels of both class samples are required in constraints, unlike OC-SVM. The parameter $\nu \in (0,1)$ sets an upper bound on the outliers and lower bound on the support vectors.

4. Experiments & results

In this section, we describe the dataset, experimental details, and results of the proposed approach.

4.1. Dataset description

To validate our approach, we have used a publicly available I3A (also known as UQSNP_HEp-2) Task-3 mitotic cell detection dataset (Weblink: https://outbox.eait.uq.edu.au/uqawilie/UQSNP_HEp2_datasets/). It comprises of 100 mitotic & 4228 non-mitotic cells. All cells are accompanied with a ground truth or mask images, which can be used to localize the cell nuclei region, required for computing the morphology-based and traditional filter-bank based features.

4.2. Evaluation metrics

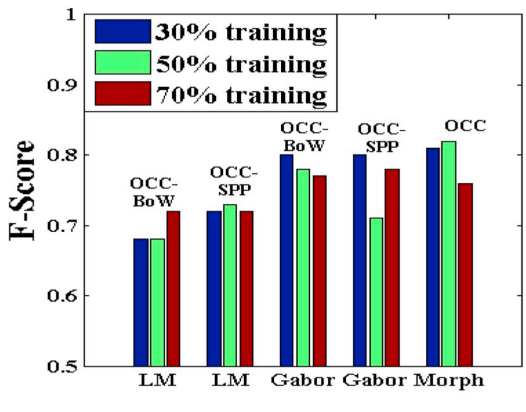
Considering the mitotic class as positive, we calculate the F-score, which is a harmonic mean of precision and recall. As the dataset is imbalanced, so it would be appropriate to give a Balanced Class Accuracy (BAC) [55], instead of using the regular classification accuracies. Both F-score and BAC are metrics for classifier evaluation, that (to some extent) handle class imbalance [55].

$$\text{F-score} = \frac{2\,.\,Precision\,.\,Recall}{(Precision + Recall)}$$

$$\text{BAC} = \frac{1}{2}.\left[\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right]$$

All the experiments are done on randomly divided data into 30%, 50% and 70% training sets, along with 60%, 40%, and 20% testing sets respectively, to ensure a statistically fair experimentation. 10% of samples are used as validation in each case. For implementation of the classifier, we use the LIBSVM toolkit [56] with Gaussian kernel. In all the experimentations, the parameter ν for OC-SVM is chosen as 0.5, acquired through validation set. The ROC curves for both types of traditional and filter bank based feature representation is shown in Fig. 5.

The ROC curves are demonstrated for two selected features for traditional (Fig. 5(a)) as well as CNN responses (Fig. 5(b)), on the validation set. The curves are plotted for varying Gaussian (Radial Basis



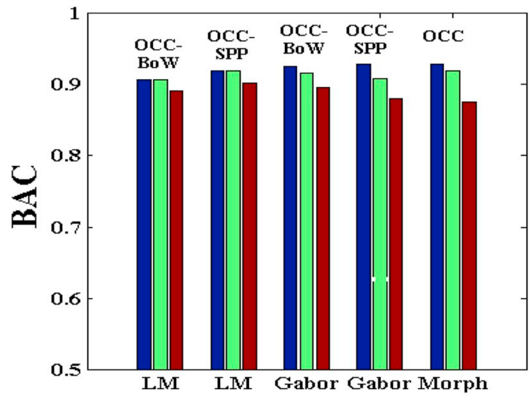


Fig. 6. Experiments of OCC with traditional features using BoW & SPP approaches. Morph represents morphological features. The color convention is same for both figures.

Function (RBF)) kernel parameters at a fixed value of ν. The selected gamma parameter for RBF kernel is 0.001 for LM, Gabor and CNN-based features and 0.5 for morphology features. The curves are demonstrated only for selected data distribution, just to show the behavior of ROC curve across such types of features.

4.3. One-class classification

Here, we discuss results with traditional feature representation, consisting of morphology and filter-bank based texture feature representation, and then using the CNN-based feature representations.

4.3.1. Traditional feature representations

For filter-banks (LM and Gabor) based features, we apply both BoW and SPP approaches to analyze the performance. In Fig. 6, we note that the performances of both BoW & SPP approaches are similar. The simple but effective morphology-based features are performing well and yield the F-score 0.81 with 30% training data. The Gabor bank outperforms the LM bank. This could be due to the high orientation sensitivity, and their aility to capture local spatial frequencies of the Gabor filters, that typically represent textural characteristics well. Finally, we observe that the relatively more simplistic morphology features are quite effective in capturing the discriminative behavior between both the classes.

We note that the False Positive (FP) cases are high in some cases, which is the reason for low F-score. Here, we infer that since the model is trained only on one class and we have a high number of negative samples, the chances of rising in FP is high. In a few cases, as the training data increases, the performance considerably decreases. This might happen because of the over-fitting of training samples, that negatively impacts the performance of the model on new data.

4.3.2. CNN-based learned features

For CNN-based learned features too (refer Fig. 7), we note that the BoW representation outperforms the SPP, with F-score of more than 0.90 in almost all cases. This feature representation extracted through each convolutional layer yields the True Positive (TP) rate 100% and very low FP rate. In all cases, the approach shows a good performance, which is also better than using traditional features. In Fig. 7, the trend of classification performance of CNN-based features are almost similar to that of traditional features, i.e., the decrement of performance with increment in training samples, that could imply over-fitting of the model.

We conclude that CNN features prove to be an effective feature representation, without the requirement of DAPI masks, even when considered in an anomaly detection scenario (i.e., using features of only one class for training).

Decision-level fusion (DLF): We also propose to use a decision level fusion for CNN-based learned features to reduce the FP cases. The decision-based fusion is done only for the CNN-based features. This is because CNN involves multi-layered features which are defined at different levels of abstraction, and hence each layer may learn different characteristics.

In the decision-based fusion of the samples, we calculate the predicted labels for all test samples from all convolutional layers. A decision level fusion strategy is applied on predicted responses on all test samples, using a majority voting rule, i.e., if a test sample is predicted as positive samples for more than half of the used layers, then it will be predicted as positive in final prediction. For the present case, the TP is reasonably high even without the decision fusion. However, the decision fusion can help in reducing the FP cases. We believe that any sample which is predicted as FP using features from some convolutional layer response (and classified via a traditional classifier), can be inversely interpreted based on the features of other layers. Hence, we propose to use a decision-based fusion for predicted test labels from all convolutional layers of the used network. We note that there is a noticeable improvement in the performance of fusion and BoW yields a good F-score value of 0.99 in one case, which is quite useful.

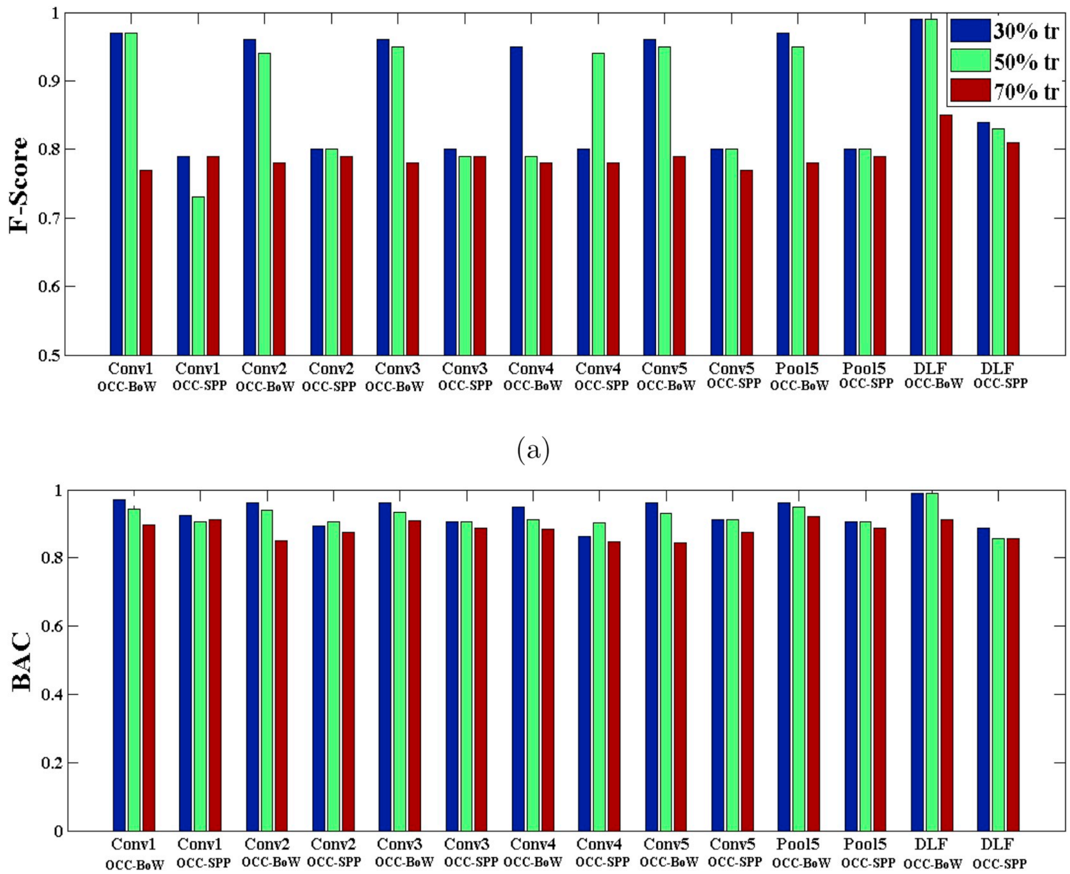
In Fig. 8, we show the average performance of CNN features, across all layers. As noted, the performance of decision-based fusion is always higher than the CNN features from individual layers. This shows the importance of such a fusion strategy.

4.4. Comparison with data-augmentation techniques

Here, we compare the proposed anomaly detection approach with those using explicit data-balancing techniques. Here, we use two techniques, i.e., undersampling (US) of majority and oversampling (OS) of minority class samples, implemented using a binary class classifier, i.e., SVM (BC-SVM).

The US is implemented by removing the randomly selected samples from the majority class, i.e., non-mitotic (in all random sets) at the data space. On the other hand, OS is done in the data space by doing the rotation of images in a range of angle between 0° to 150°, with a constant difference of 10°. The oversampling of minority class is done 40 times.

In the experimental results shown in Figs. 9 and 10, some selected features are used in case of the US and OS approaches, from a wide range of used features employed in the work. We observe that the F-score, as well as BAC in both the approaches, are good. In a few cases



(a)

(b)

Fig. 7. Experiments of OCC (One-Class Classifier) with CNN-based learned features using BoW & SPP approaches. The color bar convention is same for both figures. tr (in legend) represents the training data and DLF refers to Decision-level fusion.

with CNN features, the variations in the results are high for both the cases of US and OS. While comparing the results with the proposed approach with OCC, the augmentation methods are better for traditional features, due to the use of only one class samples in training. However, for the CNN case, the results with data augmentation are similar or lesser than the OCC approach. This comparison clearly shows the efficacy of OCC, when using a set of richer features, such as CNN features.

4.5. Comparison with CNN baseline classifier

As we use CNN based features with SVM classifier, we also show its comparison with the baseline CNN. For baseline CNN, we use the same architecture (from which the features are extracted) with Fully Connected (FC) layers used as the classifier. In this case, no fine-tuning has been done for baseline classification as the feature extraction is also done the same way using the pre-trained architecture. Importantly, the baseline classifier is implemented using both the classes for training. In Fig. 11, we show the results for baseline CNN with both US and OS techniques. We observe that the baseline classifier with US or OS is also showing similar performance to that of the OCC approach. For the comparison, we have again mentioned the best results acquired from OCC (in the same figure), which are acquired for the case of DLF approach using BoW. We note that the best results of OCC, are actually better than the CNN baseline approach. Hence, one can infer that the OCC is also a good alternative to address this problem statement, that

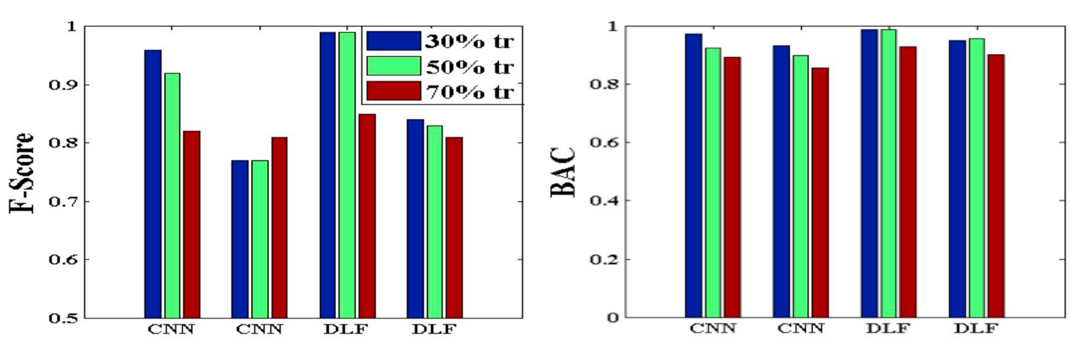


Fig. 8. Average performance of CNN-based features using BoW & SPP approaches and decision-level fusion (DLF). The color bar convention is same for both figures.

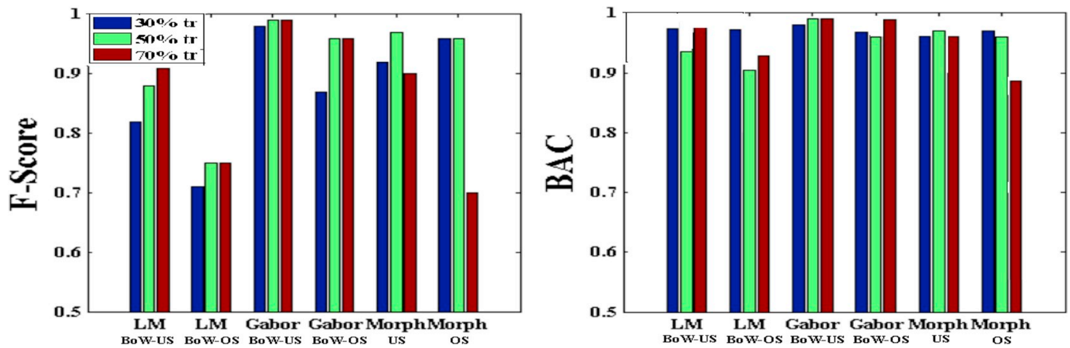


Fig. 9. Experimental results for data-augmentation techniques with selected traditional features. The color bar convention is same for both figures.

too avoiding the requirement of any data augmentation technique.

4.6. Comparison with BC-SVM (without data-augmentation)

In Section 4.4, the proposed approach is compared with BC-SVM, using data augmentation techniques, while here, we compare OC-SVM with results of BC-SVM, without using any data augmentation. The comparison is worth exploring, to ensure the effectiveness of the proposed OC-SVM. All the experiments are done only with 30% training data (Fig. 12). It is observed that the performance of OCC is less than that of BCC, when using traditional features. While, in case of CNN, the OCC outperforms BCC for BoW and it is similar to BCC for the SPP representation. We believe that in BoW, the less number of samples for the minority class is not enough for good quality clustering, which determines the final feature representation. Interestingly, for the BoW with traditional filters, the BCC is performing better. This could be because, with less number of traditional filters, the feature space is of a smaller dimension, than that with the CNN filters. Hence, the clustering would be arguably easier and better for the traditional features. On the other hand, SPP does not involve clustering, and the pooling process possibly helps in mitigating the intra-class variations. Finally, considering the overall comparison of CNN features, for most cases, the results of OCC are clearly better or very close to that of the BCC.

For traditional features, though the OCC performs less than BCC, the results are encouraging, considering the fact that OCC requires samples of only one class (an important practical advantage). In any case, our best results depict a very high F-score (0.99), and some other cases in our experimentation also yield scores close to the highest. This indicates that on this dataset (with the present training-testing protocol) the task is effectively solved.

4.7. Comparisons with other frameworks

The proposed framework is compared with two popular existing methods (defined below), which also address the issue of data imbalance.

Random Forest-based Ensemble Classifier: The first methodology is based on random forest which can be considered as ensemble classifier [57]. The single binary classifier built with datasets that suffer from imbalanced sample distribution shows bias towards the majority class. Hence, an ensemble of classifiers is known to increase the classification performance of single classifiers by integrating many of them. The random forests are built on several decision trees and each tree is built using the bagging of samples, via uniform random sampling, selected from the data (with replacement). Each decision tree in the forest considers random but balanced samples from both classes. The final prediction is made by aggregating (majority voting rule) the predictions from all classifiers. It is also beneficial to ensemble a larger variety of features in such frameworks.

SMOTE oversampling technique: Another popular method, that we use for comparison is that of SMOTE (Synthetic Minority Oversampling Technique) [58], where samples are generated along the line joining few of the k-nearest neighbors from the minority class in feature space (where k is empirically chosen). This technique importantly leads to a better feature representation & distribution than conventional oversampling techniques which typically augment data by geometric or noisy variations of the input data.

Here, the performance of both the methods are given using only one experimental setting, i.e., with 30% data for training from both the classes, and we report an average result over 10 random trials. In the classification performance of random forest classifier shown in Fig. 13, the F-score acquired through the integration of all features is 1, which is slightly better than the F-score acquired only through the decision-level fusion of convolution features (0.99), extracted through all convolutional layers of the network (refer Fig. 7). Similarly, in the classification performance of SMOTE technique, the F-Score acquired through Conv5 feature as well as the integration of all morphology, LM, Gabor, and convolutional features are the highest, i.e., 0.99.

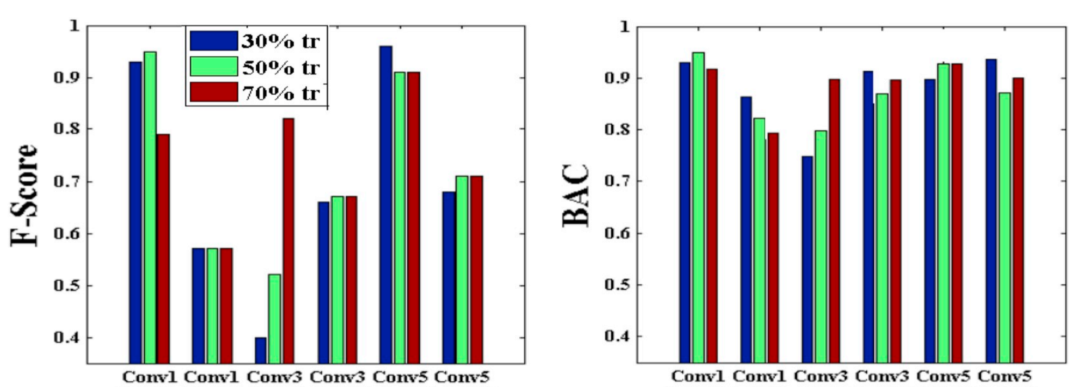


Fig. 10. Experimental results for data-augmentation techniques with selected CNN features. The color bar convention is same for both figures.

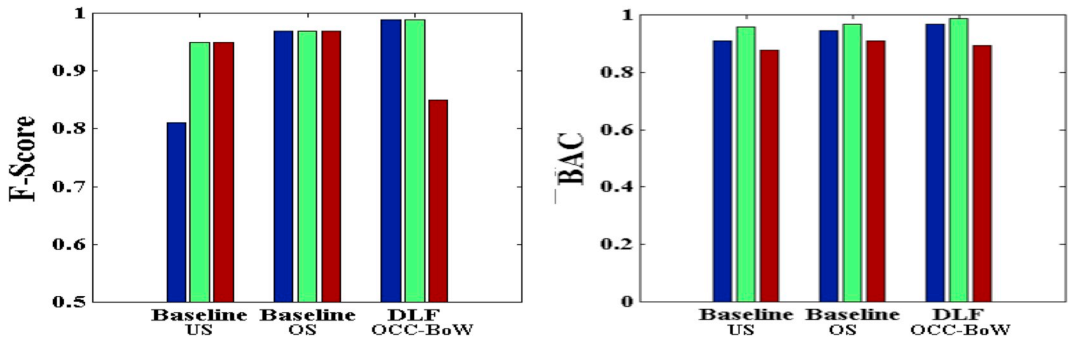


Fig. 11. Experimental baseline CNN results of mitotic v/s non-mitotic cell classification. Here features are extracted from the pool5 layer of CNN architecture. US & OS stand for Undersampling and Oversampling techniques, respectively. The color bar conventions are similar as previous figures, i.e., blue bar (I bar) is for 30% training data, green bar (II bar) is for 40% training data and red bar (III bar) is for 50% training data (Better visualized in color).

The last bar of the figure shows the performance of proposed approach at same feature representation, integrating all features (morphology, LM, Gabor & convolutional features), for comparative analysis. Here, the random forest classifier yields F-score 1, while the SMOTE and proposed method yield 0.99 and 0.98 respectively. We importantly note that the proposed approach based on OCC uses only majority class in training, while the ensemble classifier and SMOTE technique uses both class samples in training. In spite of this, the difference in the performance is marginal. Thus, the proposed framework of anomaly detection can be considered as an equally good alternative which does not require the data from the minority class for training.

4.8. Fine tuning of network

Till now, the experimentation is done using a pre-trained architecture, which is trained on a different domain of images. Here, we fine-tune the network using the current dataset, to analyze the effect of pre-trained and fine-tuned network with our dataset. As noted in Fig. 7, there was some scope for improvement on performance, hence we fine-tune the last two convolution layers Conv4 and Conv5 and analyze the performance. In Fig. 14, the experimentations are done with two different experimental setting and results are quoted with OCC, except baseline CNN classifier which uses both classes for training. We note that the F-score for fine-tuned Conv4 and Conv5 are similar or less in few cases, to the previously used features with the pre-trained network. We infer that the OCC with the pre-trained network, using a traditional classifier proves to be effective, thus avoiding the fine-tuning overhead.

In Fig. 14, the results of fine-tuned features using BCC is also shown, for fair comparative analysis of OCC with BCC. We note that the performance of BCC with fine-tuned features is quite less. This is an indication that when the model is trained using both class samples, it learns the characteristics of the abundant class, i.e., non-mitotic and hence yields very low classification scores for mitotic samples with less TP results. Hence, the use of OCC proves to be quite beneficial, as we

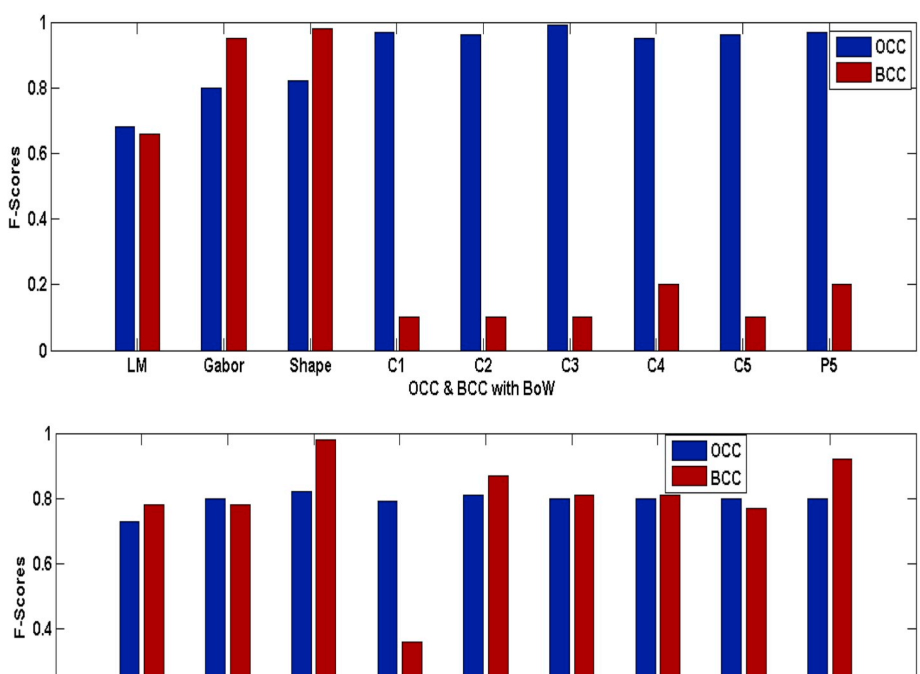


Fig. 12. Comparative analysis of BCC and OCC performance with BoW and SPP approaches on different feature representations, i.e., LM, Gabor, shape/morphology & CNN-based features (Here, C1,C2⋯C5 & P5 refer to Conv & Pool layers respectively).

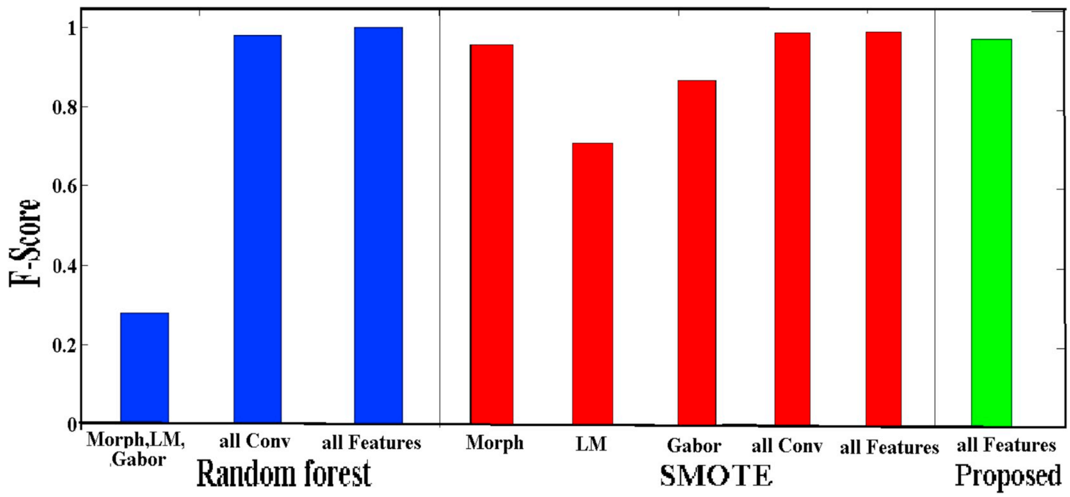


Fig. 13. Comparison of both random forest based ensemble and SMOTE technique, with proposed methodology. all Conv: Integration of Conv features, extracted from each layer. all Features: Integration of morphology, LM, Gabor & all Conv features.

test the mitotic class samples as anomalous.

Thus, concluding the results, we note that the proposed OCC based approach yields good classification performance with most feature representations and with various experimental settings of training-testing sets. The classification accuracy is equivalent or slightly less than the BCC, where we had balanced both class samples using any data-augmentation techniques. Moreover, the best classification performance is with the pre-trained CNN architecture, rather than the fine-tuned network.

4.9. Comparison with prior approaches

To the best of our knowledge, few authors have attempted the problem of mitotic cell recognition for HEp-2 cell images [12,21], but none of the works have used the dataset that we use in our current study. Moreover, none of the earlier methods have considered mitotic cell detection using anomaly detection framework. Hence, it is not possible to provide a fair comparison with prior approaches, yet to show some comparative analysis, we re-implemented some of the previous algorithms on the public dataset, that we use. The experiments are done painstakingly, using all the available hyper-parameters related information. Though the evaluation measures in prior works may be different, we evaluate the metric that we use in our algorithm. In the comparison shown in Table 1, we notice that the proposed approach clearly outperforms the earlier methods over the standardized dataset, used in this work. The reason for the performance reported in Table 1 might be the use of small size datasets used in the previous study, with different proportions of mitotic and interphase samples.

5. Ablation study

In our ablation study, the effect of different features, towards the

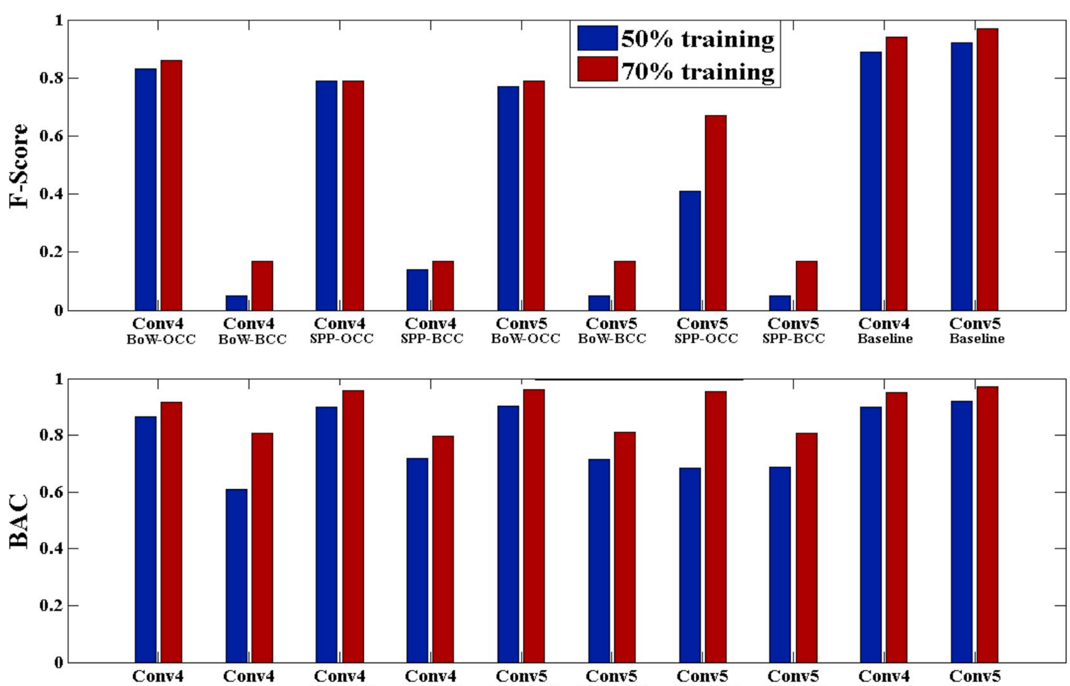


Fig. 14. Experimental results after fine-tuning of network, with fine-tuning of last two convolutional layers (results with pure data). The color convention is same for both figures.

Table 1
Comparison of proposed approach with previous ones on the current dataset.

Approach	F-score	BAC
Tonti et al. [10]	0.35	0.82
Miros et al. [21]	0.30	0.72
Proposed (traditional morphological features)	0.82	0.92
Proposed (CNN-DLF with BoW)	0.99	0.99
Proposed (CNN-DLF with SPP)	0.84	0.91

classification accuracy (F-score) is analyzed, with an experimental setting of 30% training, 10% validation, and 60% testing data. The experiments are done with OCC (Table 2). In addition, we also analyze the testing time, considering that involving more features may lead to a larger computational overhead.

- As noted, there is marginal difference amongst the testing time, acquired from different combination of features. For morphology-based features, while the testing time varies by a small amount, the F-scores acquired through both the features separately, are quite low. However, the integration of both the morphology-based features (nuclear area and nuclear perimeter) positively contributes towards a better and reasonably high performance.
- For traditional filter banks, Gabor filters perform better than LM, with a slight increase in testing time. However, the integration of both the filter yields somewhat better performance than just using Gabor filters. We observe that few testing cases, which are incorrectly classified using Gabor filters are correctly classified by LM filters and vice versa. The testing time of the integration of both the filter banks is slightly increased, with an 0.02 increment in F-Score. While combining the morphology features with both the filter banks, the testing time reduces, with an increment of 0.02 in F-Score.
- For features based on the convolutional layers-, it appears that use of only Conv5 features can also yield a high-quality performance, but can be further improved by using features from additional layers. The integration of hand-crafted features, including both morphology and filter bank features are not adding any further improvement in the classification performance. Hence, the convolutional features are themselves sufficient for good performance.

Overall, the system always yields better performance with BoW, rather than SPP, with a slight compromise in testing time. The reason could be that the BoW representation involves histograms, which are more informative than features computed by pooling over large image regions as in SPP.

6. Conclusion & future work

The proposed approach for the mitotic cell detection problem is framed as an anomaly detection problem, which incorporates effective morphology and texture-based features to represent the non-mitotic class distribution, in order to represent the mitotic class as anomalous. The framework uses an OCC, i.e., OC-SVM, which uses training data from only majority class. The algorithm yields good F-score over traditional as well as CNN-based feature representation. Moreover, we also give a comparative analysis between BoW and SPP approaches that are used to build the final feature representation scheme. To show the efficacy and adequacy of OC-SVM, we compare it with BC-SVM, which is a conventional binary classifier. Moreover, a comparative analysis is also done with v/s without using data-augmentation techniques, OC-SVM v/s CNN baseline classifier, comparison with fine-tuned network and comparative analysis with prior approaches.

Hence, the proposed framework proves to be an effective framework to solve the task of mitotic v/s non-mitotic cell classification. In future, we mean to use more sophisticated CNN architectures and feature representations, for better classification. Moreover, we also mean to address the classification among different mitotic & interphase classes, in order to build a complete pipeline, to be integrated in a CAD based system.

7. Summary & discussion

In the presented work, an automated classification strategy is shown between mitotic v/s non-mitotic/interphase patterns, with a sample skew, due to the rare occurrence of the mitotic patterns. Here, the minor class, i.e., mitotic is treated as an anomaly class and classification is framed as an anomaly detection framework. The important outcomes of the work are as follows:

- The class-specific morphological features, traditional filter bank based and CNN-based feature representation is used. The CNN based features are used to avoid the requirement of segmentation work.
- The raw set of features are compactly represented using BoW & SPP approaches. In all the experiments, the BoW is proved to be better than others.
- The two simpler & morphology-based features yield the higher or almost similar performance than other complex features, but require the exact segmentation masks. On the other hand, the Gabor outperforms LM, due to its rotational & scale-invariant properties. The CNN based features also perform well, using only one class samples in training.
- In terms of comparison with data-augmentation techniques, the used OS & US techniques perform slightly better than OCC, but require both classes in training and lead to over-fitting & under-fitting of data. Hence, here also, the OCC technique is preferred over other data augmentation techniques.
- With comparison of CNN baseline classifier, OCC using CNN based features performs better than baseline and the OCC (OC-SVM) performs less in traditional features and higher in CNN-based features than BC-SVM.

Table 2
Ablation analysis (in terms of testing time (msec/image) & F-scores) of OCC with different feature representations. Time: Testing time.

Features	BoW/SPP	Time	F-score	Features	BoW/SPP	Time	F-score
Morph (Nuclear Area)	-	1.30	0.46	Morph + LM + Gabor	BoW	2.50	0.84
Morph (Nuclear Perimeter)	-	2.05	0.66	Morph + LM + Gabor	SPP	1.70	0.86
Morph (Both)	-	2.20	0.81	LM + Gabor	BoW	2.80	0.82
LM	BoW	2.00	0.68	LM + Gabor	SPP	1.33	0.84
LM	SPP	2.00	0.72	Conv5	BoW	1.50	0.96
Gabr	BoW	2.06	0.80	Conv5	SPP	1.20	0.80
Gabor	SPP	2.08	0.80	all Conv	BoW	2.20	0.98
Morph + LM	BoW	1.60	0.82	all Conv	SPP	1.50	0.88
Morph + LM	SPP	1.20	0.82	Morph + LM + Gabor + all Conv	BoW	2.25	0.98
Morph + Gabor	BoW	1.60	0.82	Morph + LM + Gabor + all Conv	SPP	1.60	0.89
Morph + Gabor	SPP	1.60	0.82				

Hence, in all perspectives, the proposed framework proved to be an effective solution to address the problem, with a max. F-score as 0.99.

Conflicts of interest

None declared.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2019.103328.

References

[1] Y. Kumar, A. Bhatia, R. Minz, Antinuclear antibodies and their detection methods in diagnosis of connective tissue diseases: a journey revisited, Diagn. Pathol. 4 (2009) 1, https://doi.org/10.1186/1746-1596-4-1.

[2] P. Foggia, G. Percannella, P. Soda, M. Vento, Benchmarking HEp-2 cells classification methods, IEEE Trans. Med. Imaging 32 (10) (2013) 1878–1889, https://doi.org/10.1109/TMI.2013.2268163.

[3] P. Hobson, B.C. Lovell, G. Percannella, A. Saggese, M. Vento, A. Wiliem, Computer aided diagnosis for anti-nuclear antibodies HEp-2 images: progress and challenges, Pattern Recogn. Lett. 82 (Part 1) (2016) 3–11 https://doi.org/10.1016/j.patrec.2016.06.013.

[4] E. Greidinger, R. Hoffman, Antinuclear antibody testing: methods, indications, and interpretation, Lab. Med. 34 (2) (2003) 113–117, https://doi.org/10.1309/VUB9-0VTP-MEWV-3W0F.

[5] P. Luigi Meroni, N. Bizzaro, I. Cavazzana, M. Borghi, A. Tincani, Automated tests of ANA immunofluorescence as throughput autoantibody detection technology: strengths and limitations, BMC Med. 12 (2014) 38, https://doi.org/10.1186/1741-7015-12-38.

[6] P. Hobson, B.C. Lovell, G. Percannella, M. Vento, A. Wiliem, Benchmarking human epithelial type 2 interphase cells classification methods on a very large dataset, Artif. Intell. Med. 65 (3) (2015) 239–250, https://doi.org/10.1016/j.artmed.2015.08.001.

[7] K. Gupta, V. Gupta, A.K. Sao, A. Bhavsar, A.D. Dileep, Class-specific hierarchical classification of HEp-2 cell images: the case of two classes, Proceedings of 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), ICPR, 2014, pp. 6–9, , https://doi.org/10.1109/I3A.2014.13.

[8] A. Rigon, M. Infantino, M. Merone, G. Iannello, A. Tincani, I. Cavazzana, N. Caraballese, A. Radice, M. Manfredi, P. Soda, A. Afeltra, The inter-observer reading variability in anti-nuclear antibodies indirect (ANA) immunofluorescence test: a multicenter evaluation and a review of the literature, Autoimmun. Rev. 16 (2017) 1224–1229, https://doi.org/10.1016/j.autrev.2017.10.006.

[9] G. Logambal, V. Saravanan, Cancer diagnosis using automatic mitotic cell detection and segmentation in histopathological images, Proceedings of the Global Conference on Communication Technologies, GCCT), 2015, pp. 128–132, , https://doi.org/10.1109/GCCT.2015.7342638.

[10] S. Tonti, S. Di Cataldo, E. Macii, E. Ficarra, Unsupervised HEp-2 mitosis recognition in indirect immunofluorescence imaging, Proceedings of 37th IEEE Engineering in Medicine and Biology Society, EMBC), 2015, pp. 8135–8138, , https://doi.org/10.1109/EMBC.2015.7320282.

[11] P. Foggia, G. Percannella, P. Soda, M. Vento, Early experiences in mitotic cells recognition on HEp-2 slides, Proceedings of the International Symposium on Computer-Based Medical Systems, CBMS), 2010, pp. 38–43, , https://doi.org/10.1109/CBMS.2010.6042611.

[12] G. Iannello, G. Percannella, P. Soda, M. Vento, Mitotic cells recognition in HEp-2 images, Pattern Recogn. Lett. 45 (2014) 136–144, https://doi.org/10.1016/j.patrec.2014.03.011.

[13] A. Paul, D.P. Mukherjee, Mitosis detection for invasive breast cancer grading in histopathological images, IEEE Trans. Image Process. 24 (2015) 4041–4054, https://doi.org/10.1109/tip.2015.2460455.

[14] A. Wiliem, P. Hobson, B.C. Lovell, Discovering discriminative cell attributes for HEp-2 specimen image classification, Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2014, pp. 423–430, , https://doi.org/10.1109/WACV.2014.6836071.

[15] X. Jia, L. Shen, X. Zhou, S. Yu, Deep convolutional neural network based HEp-2 cell classification, Proceedings of the International Conference on Pattern Recognition, ICPR), 2016, pp. 77–80, , https://doi.org/10.1109/ICPR.2016.7899611.

[16] D. BS, K. Subramaniam, N. HR, HEp-2 cell classification using artificial neural network approach, Proceedings of the International Conference on Pattern Recognition, ICPR), 2016, pp. 84–89, , https://doi.org/10.1109/ICPR.2016.7899613.

[17] D. Gragnaniello, C. Sansone, L. Verdoliva, Cell Image Classification by a Scale and Rotation Invariant Dense Local Descriptor, Pattern Recognition Letters vol. 82, Pattern recognition Techniques for Indirect Immunofluorescence Images Analysis, 2016, pp. 72–78 Part 1 https://doi.org/10.1016/j.patrec.2016.01.007.

[18] S. Manivannan, W. Li, S. Akbar, R. Wang, J. Zhang, S.J. McKenna, An automated pattern recognition system for classifying indirect immunofluorescence images of HEp-2 cells and specimens, Pattern Recogn. 51 (2016) 12–26 https://doi.org/10.1016/j.patcog.2015.09.015.

[19] G. Percannella, P. Soda, M. Vento, Mitotic HEp-2 cells recognition under class skew, Proceedings of the International Conference on Image Analysis and Processing – ICIAP, 2011, pp. 353–362, , https://doi.org/10.1007/978-3-642-24088-137.

[20] K. Gupta, A. Bhavsar, A.K. Sao, Mitotic cells detection for HEp-2 specimen images using threshold-based evaluation scheme, Proceedings of SPIE Medical Imaging, 2018, pp. 10581–10589, , https://doi.org/10.1117/12.2293524.

[21] A. Miros, A. Wiliem, K. Holohan, L. Ball, P. Hobson, B.C. Lovell, A benchmarking platform for mitotic cell classification of ANA IIF HEp-2 images, Proceedings of the International Conference on Digital Image Computing: Techniques and Applications, DICTA, 2015, pp. 1–6, , https://doi.org/10.1109/DICTA.2015.7371213.

[22] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Proceedings of Advances in Neural Information Processing Systems, 2012, pp. 1097–1105, , https://doi.org/10.1145/3065386.

[23] P. Hobson, B.C. Lovell, G. Percannella, M. Vento, A. Wiliem, Classifying anti-nuclear antibodies HEp-2 images: a benchmarking platform, Proceedings of the International Conference on Pattern Recognition, ICPR), 2014, pp. 3233–3238, , https://doi.org/10.1109/ICPR.2014.557.

[24] B. Chazotte, Labeling nuclear DNA using DAPI, Cold Spring Harb. Protoc. (1) (2011), https://doi.org/10.1101/pdb.prot5556 pdb–prot5556.

[25] D. Hahm, U. Anderer, Establishment of HEp-2 cell preparation for automated analysis of ANA fluorescence pattern, Cytometry Part A 69A (3) (2006) 178–181, https://doi.org/10.1002/cyto.a.20223.

[26] P.L. Meroni, P.H. Schur, ANA screening: an old test with new recommendations, Ann. Rheum. Dis. 69 (8) (2010) 1420–1422, https://doi.org/10.1136/ard.2009.127100.

[27] A.S. Wiik, M. Hier-Madsen, J. Forslid, P. Charles, J. Meyrowitsch, Antinuclear antibodies: a contemporary nomenclature using HEp-2 cells, J. Autoimmun. 35 (3) (2010) 276–290 special Issue - In Honour of Professor Haralampos M. Moutsopoulos https://doi.org/10.1016/j.jaut.2010.06.019.

[28] P. Perner, H. Perner, B. Muller, Mining knowledge for HEp-2 cell image classification, Artif. Intell. Med. 26 (1) (2002) 161–173 https://doi.org/10.1016/S0933-3657(02)00057-X.

[29] P. Soda, G. Iannello, Aggregation of classifiers for staining pattern recognition in antinuclear autoantibodies analysis, IEEE Trans. Inf. Technol. Biomed. 13 (3) (2009) 322–329, https://doi.org/10.1109/TITB.2008.2010855.

[30] P. Soda, G. Iannello, A multi-expert system to classify fluorescent intensity in antinuclear autoantibodies testing, IEEE Symposium on Computer-Based Medical Systems (2006) 219–224, https://doi.org/10.1109/CBMS.2006.21 CBMS'06.

[31] G.V. Ponomarev, M.D. Kazanov, Classification of ANA HEp-2 Slide Images Using Morphological Features of Stained Patterns, Pattern Recognition Letters vol. 82, Pattern recognition Techniques for Indirect Immunofluorescence Images Analysis, 2016, pp. 79–84 Part 1 https://doi.org/10.1016/j.patrec.2016.03.010.

[32] S. Ensafi, S. Lu, A.A. Kassim, C.L. Tan, Accurate HEp-2 Cell Classification Based on Sparse Coding of Superpixels, Pattern Recognition Letters vol. 82, Pattern recognition Techniques for Indirect Immunofluorescence Images Analysis, 2016, pp. 64–71 Part 1 https://doi.org/10.1016/j.patrec.2016.02.007.

[33] O. Sarrafzadeh, H. Rabbani, A.M. Dehnavi, A. Talebi, Analyzing features by SWLDA for the classification of HEp-2 cell images using GMM, Pattern Recogn. Lett. 82 (Part 1) (2016) 44–55 https://doi.org/10.1016/j.patrec.2016.03.023.

[34] K. Gupta, V. Gupta, A. Bhavsar, A.K. Sao, Class-specific hierarchical classification for HEp-2 specimen images, Proceedings of International Conference on Image Processing, ICIP, 2015, pp. 641–645, , https://doi.org/10.1109/ICIP.2015.7350877.

[35] Y. Li, L. Shen, S. Yu, HEp-2 specimen image segmentation and classification using very deep fully convolutional network, IEEE Trans. Med. Imaging (99) (2017) 1561–1572, https://doi.org/10.1109/TMI.2017.2672702.

[36] H.T.H. Phan, A. Kumar, J. Kim, D. Feng, Transfer learning of a convolution al neural network for HEp-2 cell image classification, Proceedings of the IEEE International Symposium on Biomedical Imaging, ISBI), 2016, pp. 1208–1211, , https://doi.org/10.1109/ISBI.2016.7493483.

[37] P. Foggia, G. Percannella, A. Saggese, M. Vento, Pattern recognition in stained HEp-2 cells: where are we now? Pattern Recogn. 47 (7) (2014) 2305–2314, https://doi.org/10.1016/j.patcog.2014.01.010.

[38] A. Wiliem, Y. Wong, C. Sanderson, P. Hobson, S. Chen, B. C. Lovell, Classification of Human Epithelial Type 2 Cell Indirect Immunofluoresence Images via Codebook Based Descriptors, arXiv preprint arXiv:1304.1262.

[39] O. Sertel, U.V. Catalyurek, H. Shimada, M.N. Gurcan, Computer-aided prognosis of neuroblastoma: detection of mitosis and karyorrhexis cells in digitized histological images, Proceedings IEEE Engineering in Medicine and Biology Society, 2009, pp. 1433–1436, , https://doi.org/10.1109/IEMBS.2009.5332910.

[40] V. Roullier, O. Lézoray, T.V. Thong, A. Elmoataz, Multi-resolution graph-based analysis of histopathological whole slide images: application to mitotic cell extraction and visualization, Computerized medical imaging and graphics, Off. J. Comput. Med. Imag. Soc. 35 (2011) 7–8 603–15 https://doi.org/10.1016/j.compmedimag.2011.02.005.

[41] L.M. Manevitz, M. Yousef, One-class SVMs for document classification, J. Mach. Learn. Res. 2 (2002) 139–154.

[42] M. Zhang, B. Xu, J. Gong, An anomaly detection model based on one-class SVM to detect network intrusions, Proceedings of the 11th International Conference on Mobile Ad-Hoc and Sensor Networks, MSN), 2015, pp. 102–107, , https://doi.org/10.1109/MSN.2015.40.

[43] Y. Wang, J. Wong, A. Miner, Anomaly intrusion detection using one class SVM, Proceeding of the Fifth Annual IEEE SMC Information Assurance Workshop, 2004, pp. 358–364, , https://doi.org/10.1109/IAW.2004.1437839.

[44] X.N. Nguyen, D.T. Nguyen, L.H. Vu, Pocad: a novel pay load-based one-class

classifier for anomaly detection, National Foundation for Science and Technology Development Conference on Information and Computer Science, (NICS), 2016, pp. 74–79, , https://doi.org/10.1109/NICS.2016.7725671.
[45] S. Dreiseitl, M. Osl, C. Scheibböck, M. Binder, Outlier detection with one-class SVMs: an application to melanoma prognosis, Proceedings of AMIA Symposium, 2010, pp. 172–176.
[46] R.C. Gonzalez, R.E. Woods, Digital Image Processing, Prentice Hall, 2008.
[47] T. Leung, J. Malik, Representing and recognizing the visual appearance of materials using three-dimensional textons, Int. J. Comput. Vis. 43 (1) (2001) 29–44, https://doi.org/10.1023/A:1011126920638.
[48] I. Fogel, D. Sagi, Gabor filters as texture discriminator, Biol. Cybern. 61 (2) (1989) 103–113 https://doi.org/10.1007/BF00204594.
[49] J. Yang, Y.-G. Jiang, A.G. Hauptmann, C.-W. Ngo, Evaluating bag-of-visual-words representations in scene classification, Proceedings of the International Workshop on Multimedia Information Retrieval, 2007, pp. 197–206, , https://doi.org/10.1145/1290082.1290111.
[50] S. Lazebnik, C. Schmid, J. Ponce, Beyond bags of features: spatial pyramid matching for recognizing natural scene categories, Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition vol. 2, CVPR'06, 2006, pp. 2169–2178, https://doi.org/10.1109/CVPR.2006.68.
[51] L. Van der Maaten, G. Hinton, Visualizing high-dimensional data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.
[52] B. Krawczyk, Ł. Jeleń, A. Krzyżak, T. Fevens, One-class classification decomposition for imbalanced classification of breast cancer malignancy data, Proceedings of the International Conference on Artificial Intelligence and Soft Computing, 2014, pp. 539–550 https://doi.org/10.1007/978-3-319-07173-246.
[53] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297, https://doi.org/10.1007/BF00994018.
[54] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, J. Platt, Support vector method for novelty detection, Proceedings of the International Conference on Neural Information Processing Systems, 1999, pp. 582–588.
[55] V. García, R.A. Mollineda, J.S. Sánchez, Index of balanced accuracy: a performance measure for skewed class distributions, Proceedings of Pattern Recognition and Image Analysis, 2009, pp. 441–448 https://doi.org/10.1007/978-3-642-02172-557.
[56] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (3) (2011) 27, https://doi.org/10.1145/1961189.1961199 1–27:27.
[57] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, F. Herrera, A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches, IEEE Trans. Syst. Man CybernetE. Part C (Appl. Rev.) 42 (4) (2012) 463–484, https://doi.org/10.1109/TSMCC.2011.2161285.
[58] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, SMOTE: synthetic minority over-sampling technique, J. Artif. Int. Res. 16 (1) (2002) 321–357.