Deep Metric Learning for Identification of Mitotic Patterns of HEp-2 Cell Images

Krati Gupta, Daksh Thapar, Arnav Bhavsar and Anil K. Sao
School of Computing & Electrical Engineering, Indian Institute of Technology Mandi, HP, India
{krati_gupta,d18033}@students.iitmandi.ac.in, {arnav,anil}@iitmandi.ac.in

Abstract

Automatic identification of mitotic type staining patterns in microscopy images is an important and challenging task, in computer-aided diagnosis (CAD) of autoimmune diseases. Such patterns are manifested on a HEp-2 based cell substrate and captured via Indirect immunoflourescence (IIF) based microscopy imaging technique. The present study proposes a deep metric learning methodology, in order to identify the mitotic staining patterns which are rather rare, among several other interphase patterns present in majority. Hence, the problem is framed as a mitotic v/s non-mitotic/interphase pattern classification problem. Here, the implemented network maps the input images into a latent space, in order to compare the distances between the samples, for class declaration, via a triplet-loss based framework. The framework yields good classification performance as max. 0.85 Matthews correlation coefficient in one case, with less false positive cases, when validated over a public dataset.

1. Introduction & Motivation

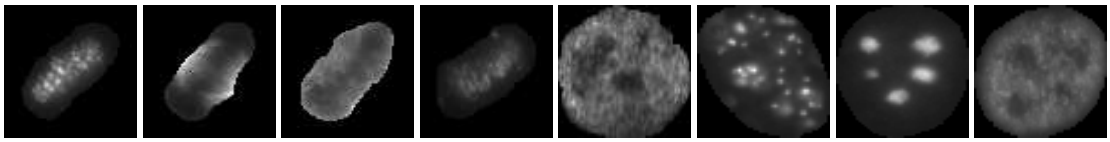
Autoimmune disorders are characterized by the presence of Anti-Nuclear Antibodies (ANA) in the patient's blood serum, wherein ANA affect healthy blood cells and tissues [8]. To confirm their presence in blood, a gold standard test [8], i.e., Indirect ImmunoFluorescence (IIF) imaging is followed, in which the patient's blood is added to a standard Human Epithelium Type-2 (HEp-2) based cell substrate (on a glass slide) and manifest in different intra-nuclear and extra-nuclear type of staining patterns [14]. Apart from these nuclear staining patterns, few distinct types of staining patterns are also visualized, known as the mitotic patterns, present in relatively rare amount. The mitotic patterns are the indicators of the mitotic phase of the cell cycle (a process for cell-division) [7], exhibiting specific clinical importance. The patterns other than the mitotic ones are collectively referred to as interphase patterns. Few types of mitotic and non-mitotic/interphase patterns are shown in Fig. 1.

The identification of each pattern type is separately required to, in turn, identify the relevant diseases. While traditionally such identification is done manually, it is a tedious and time-consuming task, motivating the need for automated machine learning based methods. In this study, we propose an approach which can contribute to such a computer-aided identification framework to triage via machine learning & pattern recognition based techniques. Such CAD systems are meant to assist the medical experts and clinicians, in their decision-making process.

As suggested in the literature, the importance of identifying mitotic over other interphase patterns lies in (a) ascertaining the presence of at least one mitotic pattern in specimen, which assures medical experts about the correct slide preparation [20] and (b) the identification of other interphase via the mitotic patterns, present in the same specimen. The above-mentioned aspects make the identification of mitotic type patterns important. Hence, the present study is framed as a task to classify mitotic v/s interphase patterns.

The important challenges in identification of mitotic patterns is their rare appearance and high intra-class variations among the different types of mitotic patterns, defined as: (a) Due to the rare appearance of mitotic patterns, these patterns always show a sample bias from other interphase class, by which the traditional classification paradigm get biased towards the majority class, as the classification model does not consider the mitotic samples much, due to the scarcity of training data from mitotic class. Hence, the data skew issue needs to be addressed in the proposed framework. (b) Moreover, the high intra-class variations among different types of mitotic patterns are effectively addressed using appropriate feature representation, that can better discriminate between mitotic and non-mitotic/interphase patterns.

Till now, there are very few works present, focusing on the identification of mitotic patterns. In [7, 16, 21], the authors have proposed a classification framework for mitotic v/s interphase classes, with a large sample skew and demonstrated various data skew balancing strategies to address the same. However, apart from these [7, 16, 21], other works focus on the identification of interphase patterns in a single cell and specimen images [8, 13, 14], and not of the mitotic patterns.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 1. Few cell images examples (a) to (d) mitotic cells & (e) to (h) non-mitotic cells.

Distance Metric Learning Framework
Considering the issue of the rare appearance of mitotic patterns, in this work, a distance metric learning based framework is proposed for detection of mitotic patterns, which is known to perform better in cases with few data samples [27]. The important aspects of the framework are as follows:

(A) Here, the distance metric computation is integrated with deep convolutional neural networks (D-CNN), which aims to learn useful embeddings of the data by distance comparisons of similar and dissimilar samples. The method learns a function $f(x)$ which maps input data to an embedding space ($\mathbb{R}^n$), on which a distance is computed to a measure of similarity between images, i.e., two images belonging to the same class should have smaller distance and images belonging to different classes should have larger distances. Such a learning framework can be modeled with a well-known Siamese network based framework [3, 23].

(B) The produced embeddings will be treated as a final feature representation and can be used in further aforementioned classification task of mitotic v/s non-mitotic/interphase single cell images. For classification, the embedding based features can be modeled using any traditional classifier.

(C) An important consideration of such a framework is the avoidance of data skew balancing strategies, that is generally required in such problem definition. It does not learn the distribution of a particular class and captures the class-based similarity. If any class possesses less number of samples, the triplet or pair-wise loss handles such issue. Each pair or triplet will essentially have one mitotic sample and hence addresses the issue of data imbalance.

The conventional data skew balancing strategies such as undersampling leads to the removal of potential and good samples from majority class, while the oversampling involves generation of new samples either in data or feature space from minority class samples, which do not follow the expected distribution in data or feature space. Hence, the framework used in the proposed approach is not intended to use any data skew balancing strategy and implemented with original skewed data samples only.

(D) The feature representations conventionally used in other frameworks generally focused to capture the discriminative shape and morphology-based descriptors, in order to distinguish mitotic cells from non-mitotic ones. To represent such features accurately, an exact segmentation mask is required, which is generally acquired using another chemical called DAPI (*4',6-diamidino-2-phenylindole*) [13]. On the basis of laboratory perspective, the use of a secondary chemical in entire diagnostic protocol adds on the complexity and segmentation overhead. The proposed framework involves the direct use of cell images and can also avoid the requirement of DAPI masks.

To the best of our knowledge, there is no other work, which employs such a technique for the identification of mitotic cells. However, the idea has been considered in face recognition [23] and object tracking tasks [6].

2. Related work

The detection of ANA is an important and challenging task, to identify the autoimmune disorders in affected human beings. There are some existing papers, demonstrating the clinical aspects of the problem statement [12, 18], including the importance of identifying mitotic patterns in whole slide or specimen images [7, 8, 20].

Till now, there are many works, explaining the problem of identifying interphase type patterns in HEp-2 single cell and specimen images, using traditional hand-crafted as well as deep neural networks (DNN) based feature representations and classifiers [4, 10, 11, 17, 19, 22, 26], validated over publicly available datasets. However, very few authors demonstrated the works, specific to the identification of mitotic type patterns in single cell images. For example, in [7, 16], authors proposed a classification framework on their own skewed dataset, which is different from the public dataset that we use in our study. In already existing works, the authors have addressed the issue of data skew between mitotic and interphase classes, using standard skew balancing methodologies, including morphological and textural feature descriptors [16, 20, 21, 24].

The deep metric learning framework focuses on learning a non-linear projection function, which transforms an input image from an input space to an embedding space, where images from the same class will be grouped together and

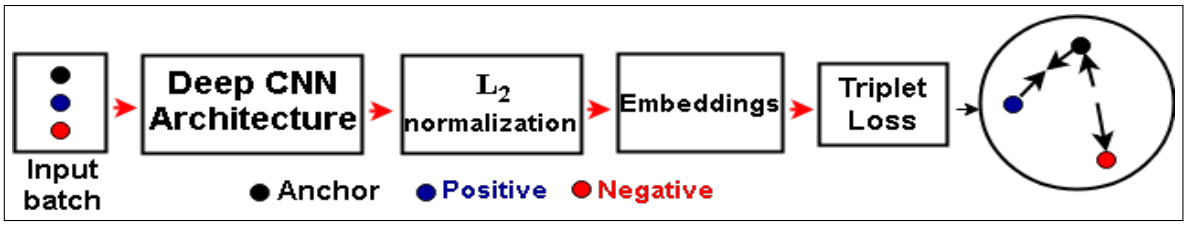


Figure 2. The flow of the model: Input batch layer is connected to Deep CNN architecture.

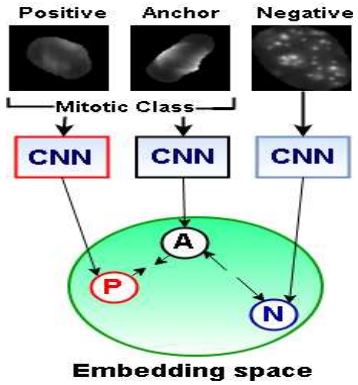


Figure 3. A triplet Siamese matching network.

images from different classes will be grouped and stretched apart [29]. Such type of frameworks are well implemented for face recognition [23], object detection [6] & person re-identification [28], but to the best of our knowledge, no works have used the similar kind of strategy for mitotic type identification problem. In general medical image analysis context, recently few works have used such metric learning based models for various tasks, e.g., multi-modal image registration [5] and multi-labeled radiographs [1].

3. Proposed approach

The proposed framework of classification between mitotic v/s interphase images involves the Siamese CNN & has following important components (Fig. 2):

3.1. Triplet Loss function

In the feature space, this loss improves the margin between the distance of images from the same class, and distance of those from different classes [15, 23]. The triplet loss minimizes the distance between an anchor image x_i^a and a positive image x_i^p (both from mitotic class), both belong to the same class and maximizes the distance between the anchor image x_i^a & a negative x_i^n of different classes. In addition, the generated embeddings are assumed to be on the d-dimensional hypersphere, i.e., $|| f(x) ||_2 = 1$. Then, the loss that has to be minimized, as:

$$L = max[|| f(x_i^a) - f(x_i^p) ||_2^2 - || f(x_i^a) - f(x_i^n) ||_2^2 + \alpha, 0]$$

$$\forall \ (f(x_i^a), f(x_i^p, f(x_i^n)) \in T)$$

where α represents the margin between positive and negative pairs images, and T is the set of possible triplets, from the training set. A triplet Siamese matching network is demonstrated in Fig. 3.

3.2. Triplet Selection: Hard Sample Mining

A high number of triplets will be generating from the data samples and passing all triplets through the network will lead to slower convergence. Hence, to select the active triplets, in order to assure the fast convergence, we apply hard sample mining, in which for any given anchor image x_i^a, the x_i^p and x_i^n are chosen for which,

$$|| f(x_i^a) - f(x_i^n) ||_2^2 - || f(x_i^a) - f(x_i^p) ||_2^2 < \alpha \quad (1)$$

$$|| f(x_i^a) - f(x_i^p) ||_2^2 < || f(x_i^a) - f(x_i^n) ||_2^2 \quad (2)$$

Here, Eq. 1 ensures that the triplets selected are violating the margin but there might be some cases, where the negative samples could be closer to the positive ones, leading to very high loss. To prevent it, only those triplets are selected for which Eq. 2 satisfies.

Hence, we apply an online triplet mining technique to select the hard positive/negative samples from mini batches and it should be ensured that a minimum number of samples of any one class should be present in each mini batch.

3.3. CNN based embeddings

Here, a network architecture has been designed, including 11 convolutional layers. The details of the architecture are provided in Table 1. The Rectified Linear Unit (ReLU) activation function is used in each convolutional layer and the convolutional layers are followed by Global Average Pooling (GAP) layer, which is importantly required in case of fewer data samples. If the flattening is directly applied after convolutional layers, then the parameters are very large in number, which is effectively decreased by GAP layer. After GAP layer, four dense layers are used. Once, the model is trained and validated using only training & validation dataset, the embeddings are extracted for all the training images, and can be used in any classifier for further

Layer	size-in	size-out	kernel	stride
conv1	128×128×3	128×128×128	5×5	1
conv2	128×128×128	64×64×128	2×1	2
conv3	64×64×128	64×64×128	5×5	1
conv4	64×64×128	32×32×128	2×1	2
conv5	32×32×128	32×32×128	5×5	1
conv6	32×32×128	16×16×128	2×1	2
conv7	16×16×128	16×16×128	5×5	1
conv8	16×16×128	8×8×128	2×1	2
conv9	8×8×128	8×8×128	5×5	1
conv10	8×8×128	4×4×128	2×2	2
conv11	4×4×128	4×4×128	5×5	1
Global average pooling (GAP)				
Dense-512 (ReLU)				
Dense-256 (ReLU)				
Dense-128 (ReLU)				
Dense-128 (Linear)				

Table 1. The detailed structure of the network used.

classification between mitotic & non-mitotic samples.

3.4. Classification

For the classification task, we use Gaussian kernel based Support Vector Machine (SVM) classifier for prediction. It is worth noticing that the framework is not using any data augmentation & data balancing approaches, even when there is a large data imbalance between both the classes. For testing, the test samples are passed to the same network, in order to get the embeddings. Now the embeddings of both training and testing samples are used as final feature representations for SVM classifier.

4. Experiments & Results

In this section, the details of the dataset used, evaluation protocol and experimental results are presented, along with the experimental settings.

4.1. Dataset description & Evaluation protocol

To validate our approach, we have used a public I3A (also known as UQSNP_HEp-2) Task-3 mitotic cell detection dataset (https://outbox.eait.uq.edu.au/uqawilie/). It comprises of 100 mitotic cells and 4228 non-mitotic cells. Though the DAPI based segmentation masks are also provided with the original images, but the proposed approach does not require the segmentation masks.

For evaluation, considering mitotic as positive, we calculate the precision, recall and the F-score (harmonic mean of precision and recall). As the dataset is imbalanced, so the evaluation using F-score is not appropriate for the task as precision affects with the number of false positive (FP) cases, which are proportionately very high with respect to the true positive (TP) samples, as precision is defined as TP/(TP+FP). Hence, it would be appropriate to give a

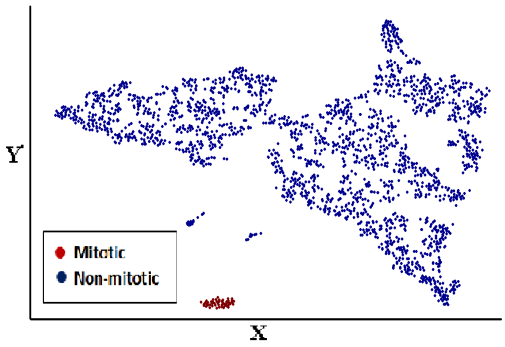


Figure 4. The t-SNE representation of embeddings of both classes (training set), in reduced virtual dimensions. Here, X & Y are two reduced virtual dimensions.

balanced class accuracy (BcA) [9], instead of using the regular accuracies. Here, one more evaluation measure is also defined, which is considered to be a good way to evaluate performance, in case of highly imbalanced classes [2], known as Matthews correlation coefficient (MCC), which considers all the four elements of a confusion matrix. Thus, both the BcA and MCC measures are defined as:

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

$$\text{MCC} = \frac{(TP.TN)-(FP.FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$

4.2. Experimental Results

The classification results are reported with three different experimental settings (we will term them as Ex-A, Ex-B and, Ex-C throughout the paper), in order to maintain the robustness of the system, i.e.,

- Ex-A: 20% samples as training data, 10% as validation data & 70% as testing data,
- Ex-B: 40% samples as training data, 10% as validation data & 50% as testing data and
- Ex-C: 60% samples as training data, 10% as validation data & 30% as testing data.

Three random sets are generated from the whole data, with the above experimental settings and an average result is reported at the end.

To show the discrimination between the embeddings acquired for both the classes, via the proposed framework, a t-SNE plot (Fig. 4) has been shown with two reduced virtual dimensions for training data, however the classification has been done on actual 128-dimensional embedding representation. t-SNE [25] is a technique to visualize a high-dimensional feature representation, into two reduced virtual dimensions. In Fig. 4, it is observed that there is good discrimination between mitotic and non-mitotic/interphase samples, for training data. The approach is validated quantitatively using F-score, BcA and MCC are demonstrated for

Confusion matrix for Ex-A		
	Predicted mitotic	**Predicted interphase**
Actual mitotic (70)	69	1
Actual interphase (2960)	36	2924
Confusion matrix for Ex-B		
Actual mitotic (50)	49	1
Actual interphase (2114)	21	2093
Confusion matrix for Ex-C		
Actual mitotic (30)	29	1
Actual interphase (1268)	9	1259

Table 2. Confusion matrix acquired from the proposed framework

Experiments	Precision	Recall	F-score	BaC	MCC
Ex-A	0.66	0.99	0.79	0.98	0.80
Ex-B	0.70	0.98	0.82	0.98	0.82
Ex-C	0.76	0.96	0.85	0.97	0.85

Table 3. Results acquired from the proposed framework

the experiments in Table 3, along with the confusion matrix in Table 2. Here, we will compare and analyze all the results, based on MCC, while to show an elaborated evaluation, all the three evaluation measures are quoted here.

In all the experiments, CNN is trained using Stochastic Gradient Descent (SGD) and AdaGrad optimizer. The learning rate is .001. The decrease in the loss slows down after 200th iteration of training. The margin α is 1.5 in all experimental settings. The used CNN architecture is empirically chosen, after different parameters settings.

It is noted that the F-score is not very high because there is a high imbalance in test samples from both the classes. Hence, in spite of acquiring less false negative (FN) samples (considering interphase as a negative class), the precision is less and hence leads to less F-score (0.79, 0.82, 0.85). Note that, in practical triaging settings, the pathologists will only have to consider 36/21/9 (i.e., 1.23%/0.99%/0.71%) false positive (FP) samples, for all three cases.

While calculating BaC, which is a more appropriate measure for imbalanced class datasets, we are getting fairly good classification performance (0.98, 0.98 & 0.97). As suggested, the most appropriate evaluation protocol, i.e., MCC for different classification setting, acquired as 0.80, 0.82 & 0.85. Here also, it is observed that as the training set increases, the performance increases proportionally, but due to the high number of FP samples, the overall performance is less. Hence, for analysis of the actual number of samples, the confusion matrix is also presented in Table 2.

It is observed that the true-positive accuracy is quite high in this case. In spite of having very less (30,50 & 70 images) positive samples in training data, the framework yields good true-positive rate. This shows the effectiveness of the triplet loss based framework, where the framework is optimized to select the triplets having combinations of samples, including mitotic samples in each batch. Hence, we conclude that the framework yields good performance, even without using any data balancing or data augmentation methodologies and such type of deep metric learning based frameworks prove to be an effective and appropriate way to solve the problem.

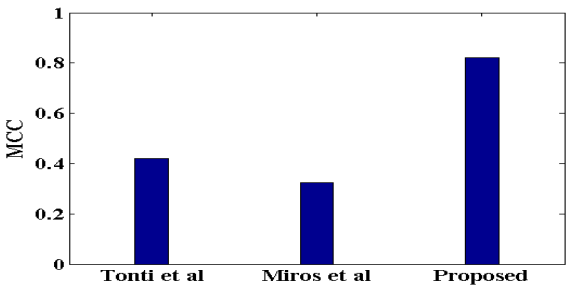


Figure 5. The comparison of proposed approach (MCC) with existing approaches.

5. Comparisons

This section presents a comparative analysis of the proposed framework with already existing approaches, CNN based baseline classifiers, and other traditional classifiers. All these comparisons are reported only on one experimental setting of Ex-B.

5.1. Comparison with prior approaches

To the best of our knowledge, no other authors have used the same dataset, that we have used in our study and the other datasets are not publicly available. Hence, the existing approaches ([20, 24]) are re-implemented using the same data, with available information of hyper-parameters. Here also, the MCC is provided for these works. It is noted that the proposed framework yields better performance among the three approaches (Fig. 5).

5.2. Comparison with cross-entropy loss

The implemented CNN model can be directly used as a feature extractor as well as the classifier, applied on raw images. For this task, the last dense layer is replaced with another dense layer, with 2 hidden nodes and softmax activation function, in order to build an end-to-end solution. To address the data imbalance between both the classes, mini-batches are created in such a way that each batch consists of the balanced number of samples, acquired from both the classes. If this mini-batch creation approach is not followed, then the classification paradigm of CNN gets biases towards the majority class.

It is noted that the F-score and MCC acquired from this method are less than the proposed framework (Triplet loss with SVM classifier). The reason might be the involvement of less mitotic samples during training. Hence, we can conclude that mitotic samples (only 30/50/70 images in train-

Experiments	F-score	BaC	MCC
Cross-entropy loss	0.48	0.69	0.48
Nearest neighbor based classifier	0.61	0.77	0.60
Proposed	0.82	0.98	0.82

Table 4. Comparative analysis of proposed frameworks with other frameworks.

ing) are not sufficient to build a classification model, that can identify and classify the mitotic type testing images and one need to go for any other method, addressing the class sample skew. Hence, the proposed framework is a good approach for classification, even with the skewed dataset.

5.3. Nearest neighbor based baseline classifier

To show and analyze the effectiveness of proposed embeddings, the SVM based framework has been analyzed and compared with simple K-nearest neighbor method also. It is observed that KNN is performing well on produced embeddings and it is showing MCC as 0.60, which is good using this simple, yet effective classifier.

Here, analyzing the good performance of proposed framework over existing works and other baseline classifiers, it can be concluded that the proposed framework proves to be a good solution for mitotic v/s interphase classification problem, with avoidance of data augmentation & segmentation overhead.

6. Conclusion & Future aspects

Hence, in this current study, the contribution is the proposal of a framework for mitotic v/s interphase patterns classification, via deep metric learning. Importantly, the framework also avoids the requirement of segmentation masks and data balancing methods, yielding good classification performance.

The future work is focused on (a) defining a class-wise classification framework for each mitotic and interphase class with few shot learning & distance metric based framework, (b) planning to explore more sophisticated and CNN-based learned feature representation to use in the future tasks, with effective baseline classifiers and (c) demonstrating comparative analysis of proposed framework with baseline & triplet loss based approach. Hence, the overall goal is to propose a classification pipeline for all types of mitotic & interphase patterns, to be integrated with CAD- based ANA detection system.

References

[1] Mauro Annarumma and Giovanni Montana. Deep metric learning for multi-labelled radiographs. In *Proceedings of the Annual ACM Symposium on Applied Computing*, pages 34–37, 2018.

[2] Sabri Boughorbel, Fethi Jarray, and Mohammed El-Anbari. Optimal classifier for imbalanced data using Matthews correlation coefficient metric. *PLOS ONE*, 12(6):1–17, 06 2017.

[3] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. Signature verification using a "Siamese" time delay neural network. In *Advances in Neural Information Processing Systems 6*, pages 737–744. 1994.

[4] Divya BS, K. Subramaniam, and Nanjundaswamy HR. HEp-2 cell classification using artificial neural network approach. *in: Proceedings of the International Conference on Pattern Recognition (ICPR)*, pages 84–89., Dec 2016.

[5] Xi Cheng, Li Zhang, and Yefeng Zheng. Deep similarity learning for multimodal medical images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 6(3):248–252, 2018.

[6] Xingping Dong and Jianbing Shen. Triplet loss in siamese network for object tracking. In *ECCV*, 2018.

[7] P. Foggia, G. Percannella, P. Soda, and M. Vento. Early experiences in mitotic cells recognition on HEp-2 slides. In *Proc. of International Symposium on Computer-Based Medical Systems*, pages 38–43, 2010.

[8] P. Foggia, G. Percannella, P. Soda, and M. Vento. Benchmarking HEp-2 cells classification methods. *IEEE Transactions on Medical Imaging*, 32(10):1878–1889, 2013.

[9] V. García, R. A. Mollineda, and J. S. Sánchez. Index of balanced accuracy: A performance measure for skewed class distributions. *in: Proceedings of Pattern Recognition and Image Analysis*, pages 441–448., 2009.

[10] Diego Gragnaniello, Carlo Sansone, and Luisa Verdoliva. Cell image classification by a scale and rotation invariant dense local descriptor. *Pattern Recognition Letters*, 82, Part 1:72 – 78, 2016.

[11] Krati Gupta, Vibha Gupta, Anil K. Sao, Arnav Bhavsar, and A. D. Dileep. Class-specific hierarchical classification of HEp-2 cell images: The case of two classes. *2014 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A)*, 2014.

[12] Daniel Hahm and Ursula Anderer. Establishment of HEp-2 cell preparation for automated analysis of ANA fluorescence pattern. *Cytometry Part A*, 69A(3):178–181, 2006.

[13] Peter Hobson, Brian C. Lovell, Gennaro Percannella, Alessia Saggese, Mario Vento, and Arnold Wiliem. Computer aided diagnosis for anti-nuclear antibodies HEp-2 images: Progress and challenges. *Pattern Recognition Letters*, 82, Part 1:3 – 11, 2016.

[14] Peter Hobson, Brian C. Lovell, Gennaro Percannella, Alessia Saggese, Mario Vento, and Arnold Wiliem. HEp-2 staining pattern recognition at cell and specimen levels: Datasets, algorithms and results. *Pattern Recognition Letters*, 82, Part 1:12 – 22, 2016.

[15] Elad Hoffer and Nir Ailon. Deep metric learning using triplet network. *CoRR*, abs/1412.6622, 2014.

[16] Giulio Iannello, Gennaro Percannella, Paolo Soda, and Mario Vento. Mitotic cells recognition in HEp-2 images. *Pattern Recognition Letters*, 45:136–144, 2014.

[17] Xi Jia, Linlin Shen, Xiande Zhou, and Shiqi Yu. Deep convolutional neural network based HEp-2 cell classification. *in: Proceedings of the International Conference on Pattern Recognition (ICPR)*, pages 77–80., Dec 2016.

[18] Yashwant Kumar, Alka Bhatia, and Ranjana Minz. Antinuclear antibodies and their detection methods in diagnosis of connective tissue diseases: A journey revisited. *Diagnostic pathology*, 4:1, 02 2009.

[19] Siyamalan Manivannan, Wenqi Li, Shazia Akbar, Ruixuan Wang, Jianguo Zhang, and Stephen J. McKenna. An automated pattern recognition system for classifying indirect immunofluorescence images of HEp-2 cells and specimens. *Pattern Recognition*, 51:12 – 26, 2016.

[20] Anastasia Miros, Arnold Wiliem, Kim Holohan, Lauren Ball, Peter Hobson, and Brian C. Lovell. A benchmarking platform for mitotic cell classification of ANA IIF HEp-2 images. In *Proc. of International Conference on Digital Image Computing: Techniques and Applications*, pages 1–6, 2015.

[21] Gennaro Percannella, Paolo Soda, and Mario Vento. Mitotic HEp-2 cells recognition under class skew. In *Proc. of ICIAP 2011*, pages 353–362. 2011.

[22] Omid Sarrafzadeh, Hossein Rabbani, Alireza Mehri Dehnavi, and Ardeshir Talebi. Analyzing features by SWLDA for the classification of HEp-2 cell images using GMM. *Pattern Recognition Letters*, 82(Part 1):44 – 55, 2016.

[23] Florian Schroff, Dmitry Kalenichenko, and James Philbin. FaceNet: A unified embedding for face recognition and clustering. *CoRR*, abs/1503.03832, 2015.

[24] Simone Tonti, Santa Di Cataldo, Enrico Macii, and Elisa Ficarra. Unsupervised HEp-2 mitosis recognition in indirect immunofluorescence imaging. In *Engineering in Medicine and Biology Society (EMBC), 37th Annual International Conference of the IEEE*, pages 8135–8138, 2015.

[25] Laurens van der Maaten and Geoffrey E. Hinton. Visualizing high-dimensional data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.

[26] A. Wiliem, P. Hobson, and B. C. Lovell. Discovering discriminative cell attributes for HEp-2 specimen image classification. In *IEEE Winter Conference on Applications of Computer Vision*, pages 423–430, 2014.

[27] Liu Yang and Rong Jin. Distance metric learning: A comprehensive survey. *Michigan State Universiy*, 2, 2006.

[28] D. Yi, Z. Lei, S. Liao, and S. Z. Li. Deep metric learning for person re-identification. In *2014 22nd International Conference on Pattern Recognition*, pages 34–39, Aug 2014.

[29] Xuefei Zhe, Shifeng Chen, and Hong Yan. Directional statistics-based deep metric learning for image classification and retrieval. *CoRR*, abs/1802.09662, 2018.