Pattern Recognition Letters
journal homepage: www.elsevier.com

Computer Aided Diagnosis for Anti-Nuclear Antibodies HEp-2 Images: Progress and Challenges

Peter Hobson[a], Brian C. Lovell[b], Gennaro Percannella[c], Alessia Saggese[c], Mario Vento[c,**], Arnold Wiliem[b]

[a] *Sullivan Nicolaides Pathology, Australia*
[b] *University of Queensland, Australia*
[c] *University of Salerno, Italy*

ABSTRACT

The Anti-Nuclear Antibodies (ANA) test using Human Epithelial (HEp-2) cells has been the gold standard to identify the presence of Connective Tissue Diseases (CTD) such us Systemic Lupus Erythomatus (SLE). As the ANA test is time consuming, labour intensive and subjective, there has been an on-going effort to develop image-based Computer Aided Diagnosis (CAD) systems. This paper discusses the current progress and challenges in this field and will highlight areas which require more attention.

1. Introduction

Pathology plays an integral part in our healthcare system and is critical in the diagnosis of many conditions. Recently, there has been a significant effort to develop image-based Computer Aided Diagnosis (CAD) systems for various pathology tests. Such systems will automate some processes in the test analysis (Tadrous (2010); Foggia et al. (2013); Labati et al. (2011); Wilbur (2011); Gurcan et al. (2009); Hobson et al. (2014); Wiliem et al. (2015a); Hobson et al. (2015); Samak et al. (2015); Miros et al. (2015)) and therefore will help pathologists and scientists to channel their attention on difficult cases (Tadrous (2010); Hobson et al. (2015)). For instance, a system giving top locations in a pap-smear slide that are highly likely to contain cancerous cells would avoid pathologists and scientists wasting time looking at non-diagnostic locations (Dawson).

The Anti-Nuclear Antibodies (ANA) test using Human Epithelial type 2 (HEp-2) cells in Indirect Immunofluorescence Assay (IFA) protocol has been the gold standard for detecting Connective Tissue Diseases such as Systemic Lupus Erythomatus (Meroni and Schur (2010); Wiik et al.). This is due to its high sensitivity as HEp-2 cells express a wide range of antigens. Unfortunately, as discussed by Bizzaro et al. (1998); Pham et al. (2005); Hiemann et al. (2009); Soda et al. (2009), the protocol is: (1) time consuming; (2) labour intensive; (3) subjective; (4) has low reproducibility and (5) has large inter/intra- personnel/laboratory variations.

To address these issues, there has been a steady on-going effort in the community to develop such systems (Foggia et al. (2013); Perner et al. (2002); Hiemann et al. (2009); Elbischger et al. (2009); Hsieh et al. (2009); Soda et al. (2009); Cordelli and Soda (2011); Wiliem et al. (2011); Strandmark et al. (2012); Bel Haj Ali et al. (2012); Theodorakopoulos et al. (2012); Thibault and Angulo (2012); Ghosh and Chaudhary (2012); Li and Yin (2012); Cataldo et al. (2012); Snell et al. (2012); Ersoy et al. (2012); Wiliem et al. (2013); Foggia et al. (2014); Faraki et al. (2014); Yang et al. (2014)). Some relevant benchmarking platforms have also been proposed such as ICPR2012Contest (Foggia et al. (2013, 2014)), SNPHEp-2 (Wiliem et al. (2013)), ICIP2013 (Hobson et al. (2015) [1]) and ICPR2014 (Hobson et al. (2014) [2]).

Despite their important contributions, the progress of building practical CAD systems for the ANA test is still far from complete. Many existing works primarily focus on narrow problems. For instance, works by Foggia et al. (2014); Hobson

**Corresponding author
e-mail: peter_hobson@snp.com.au (Peter Hobson), lovell@itee.uq.edu.au (Brian C. Lovell), pergen@unisa.it (Gennaro Percannella), asaggese@unisa.it (Alessia Saggese), mvento@unisa.it (Mario Vento), a.wiliem@uq.edu.au (Arnold Wiliem)

[1] The competition website and dataset available at http://nerone.diem.unisa.it/contest-icip-2013/index.shtml

[2] The competition and dataset is available at http://i3a2014.unisa.it/

et al. (2015) address the classification problem of cells images derived from ANA images. Other works aim at classifying the ANA images (Hobson et al. (2014); Wiliem et al. (2015a)).

This study aims to identify critical issues which may bring us a step closer to our final goal. We continue our discussion as follows. The discussion starts by presenting the general overview of ANA CAD systems. This allows us to have a more complete picture on how components of CAD systems are typically organized. Then, we discuss the progress and challenges made in each front in Section 3. Finally, the main findings and future directions are presented in Section 4.

2. General overview of ANA CAD systems

We first present how the ANA test is performed, and then discuss possible methods of automation that can be provided by ANA CAD systems.

2.1. Steps in ANA test

The general steps in ANA test are depicted in Fig. 1. The ANA test starts when patient serum is received. Various protocols (CDC (1996)) are employed to prepare the ANA slide. More specifically, the patient serum will be placed and incubated onto an ANA slide. The ANA slide contains HEp-2 cells. Once the ANA slide has been prepared, it is read by scientists under a fluorescence microscope. The scientists report the positivity of the slide (*i.e.* whether the test is positive or negative). The positivity is generally determined by either the fluorescence intensity of the slide or the presence of known ANA patterns. Whilst there is no further step required for negative patients, more information are extracted for positive patients.

Two pieces of information are reported for positive patients: (1) the ANA pattern and (2) the strength of the pattern. The pattern strength can be determined by serial diluting the sample until the pattern extinguishes. The dilution level before the pattern extinguishes is called the end point titre, and represents the pattern strength. In some cases the pattern strength is roughly quantized in the form of scale from one plus '+' as the weakest positive to four pluses '++++' as the strongest. In the case where there are multiple patterns appearing in the slide, individual pattern strength is reported.

Often secondary tests such as LIA, ELISA and FEIA are employed to further confirm the presence of the antibodies. This is important to ensure there is no additional pattern masked by the primary pattern. Secondary tests are also required to further analyze speckled patterns (i.e. there are numerous antibodies such as DFS-70 associated to the speckled patterns). Thus, knowing exactly which antibody is associated to a speckled pattern is important. For instance, it was recently found that the DFS-70 antibody has a relatively high prevalence in healthy individuals. Henceforth, when only DFS-70 antibody is present in a patient's serum, then the patient is unlikely to be diagnosed with a Systemic Autoimmune Rheumatic Diseases (SARD), Mahler et al. (2012).

2.2. ANA CAD system

An ideal CAD system should automate the steps depicted in Fig. 1. Patient slides can be prepared automatically to reduce the risk of error. Once the slides are prepared, they have to be digitized. The digitization process can be done via a motorized microscope equipped with a digital camera. Again, the digitization process should be done automatically as manually capturing the specimen is extremely time consuming and labour intensive; thus, defeating the whole purpose of CAD systems. Once the slide has been digitized, it is then fed into the CAD systems. The CAD systems can automate each step depicted in Fig. 1. In the next section, discussions on each issue will be elucidated.

3. Progress and challenges discussion

We organize the discussion from the simplest problem to the most challenging ones. We then end the discussion with the challenges that could be addressed in the future.

3.1. Positive-negative binary classification

The goal for this automation is to devise a classifier $f : \boldsymbol{I} \mapsto \{-1, +1\}$, where $+1$ is positive ANA and -1 is negative ANA.

Clinical usage - As mentioned, this is the first step in the ANA test. When a test is considered as positive, the physicians will further look into the reported information such as the positivity strength and observed positive patterns.

Progress - Due to its simplicity, this problem is one of the earliest problems attempted to be solved by the community. For instance in Soda et al. (2009), a multiple expert system is considered. More specifically, they extract a set of features from the whole image instead of from the cell content for the following reasons: (1) the background has negligible contribution to the final features as it is normally uniformly dark; (2) the cells of the same image will have similar patterns, therefore they will contribute equally; and (3) errors caused by the cell segmentation can be avoided. The features are then fed into a multiple expert system that has reject capability, Foggia et al. (1999). The reject decision is based on the classifier reliability score. More specifically, the reject decision is introduced to reduce error made by the classifier. When the system rejects an image, it means the system will let human to make the decision. We note that this approach does not follow the classifier model f described above.

A Multi-Layer Perceptrons (MLP) method is proposed in Soda and Iannello (2009). Here, a cell segmentation is used to extract the features. Early methods incorporate sophisticated classifiers and hand-crafted features. One exception is the approach described in Wiliem et al. (2011) that uses a simple average and standard deviation of the image pixel intensities.

It is also noteworthy to mention that, apart from the scientific community, there are commercial products developed for detecting positive ANA patients as reported in Bizzaro et al. (2014).

Challenges - Due to its simplicity, the works on this problem are considered relatively mature. This is evidenced by the large number of commercial products offered in this space.

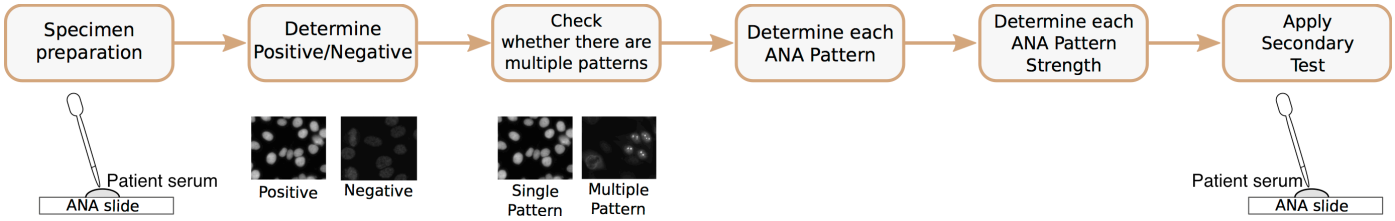

Fig. 1: Typical steps in ANA analysis.

Public benchmarking datasets - To our knowledge, there is no publicly available benchmarking dataset. Perhaps, a benchmarking dataset will provide a common platform for both laboratories and practitioners to compare different systems proposed in the literature and commercial products.

3.2. Positivity strength determination

As mentioned, there are two forms on how the positivity strength is reported: (1) a scale from one plus '+' to four pluses '++++'; or (2) the end point titre. Given a positive ANA image $\boldsymbol{I}$, the task is to develop a system that admits $\boldsymbol{I}$ and returns the strength of the pattern. The advantage of using such a system is that one could determine the pattern strength in a single dilution; thus, saving time and cost.

Clinical usage - Positivity strength information indicates the amount of auto antibody present in the patient. It is used as a guide for physicians to determine the severity of the diseases.

Progress - Most methods addressing this problem are divided into two parts: (1) positivity image feature extraction and (2) standard curve. The positivity image feature measures the pattern positivity strength. The feature is then translated into the corresponding strength in forms of either the scale or end point titre.

Fig. 2 illustrates on how most systems work. First a standard control with known strength is serially diluted using the manual method. The positivity image feature is extracted for each point. Then, the standard curve is calculated from these points. Given a positive ANA image, the positivity image feature is first extracted. Then the feature is translated into the corresponding strength via the standard curve.

Hiemann (2013) uses camera exposure time as the positivity image features. More specifically, camera exposure time is adjusted for every patient to capture a good quality picture. The key concept is that stronger positive patients generally have shorter exposure time. PolyTitre, presented in Flessland et al. (2002), uses a specialized hardware to quantify the positivity strength. Unfortunately, it requires manual operation to quantify the positivity strength. ImageTitre (Nakabayashi et al. (2001)) uses a semi-automatic approach where an image titration is applied by progressively reducing the image brightness. A human operator is still required to determine to which iterations the ANA becomes negative. To that end, recently Wiliem et al. (2011) proposed a digital titration simulation that is based on the exposure-density curve model studied in 1876 by Hurter and Driffield. The approach is combined with a positive-negative classifier which enables the whole system to run automatically without human intervention.

Challenges - Although the existing methods have been shown to be effective in determining the strength of positive ANA patients, they are still limited to patients who have a single ANA pattern. When multiple ANA patterns existed in a patient serum, then the strength of each pattern needs to be determined. To our knowledge, there is no thorough study in this area. We note that further discussion on analysis of multiple ANA patterns will be presented later in this work.

Public benchmarking datasets - To our knowledge, there is no publicly available benchmarking dataset. Releasing a dataset could help advancement in this field. To provide a complete study, the dataset should have patient sera with different end point titre and ANA patterns.

3.3. Pattern classification

Perhaps the most well-known problem in the community is on the ANA pattern classification problem. In a nutshell, the task is to classify a given image $\boldsymbol{I}$ into one of the ANA patterns.

Clinical usage - In general, a well-studied ANA pattern is associated to one or more antibodies. In other words, observing a particular pattern indicates the presence of particular set of antibodies. For instance, the nucleolar staining pattern is associated with many antibodies such as anti-Scl-70, anti-PM-Scl and anti-Th/To.

The task comprises a series of steps as follows: (1) cell segmentation; (2) mitotic-interphase cell cycle classification; (3) HEp-2 cell classification and (4) specimen image classification. We discuss each subtask in the following subsections.

3.3.1. HEp-2 cell nucleus segmentation

In this subtask the goal is to locate and segment the HEp-2 cell nucleus images from their ANA specimen image. For convenience, we shall call the HEp-2 cell nucleus image as 'HEp-2 cell image'. Fig. 3 depicts some examples of ANA specimen images. As we can see the specimen images have multiple HEp-2 cells.

Clinical usage - As mentioned, the segmentation step is an intermediate step before classification subtasks are carried on. Thus, the results from this step do not have clinical usage. However, the segmentation accuracy affects the analysis in the subsequent subtasks.

Progress - Segmenting HEp-2 cell images is of importance as most informative features are contained inside the HEp-2 cell image. This task is also considered as a challenging task. There are approaches to address this task: (1) segmenting via secondary channel, here denoted *indirect segmentation*; (2) segmenting via digital image segmentation, here called *direct segmentation*.

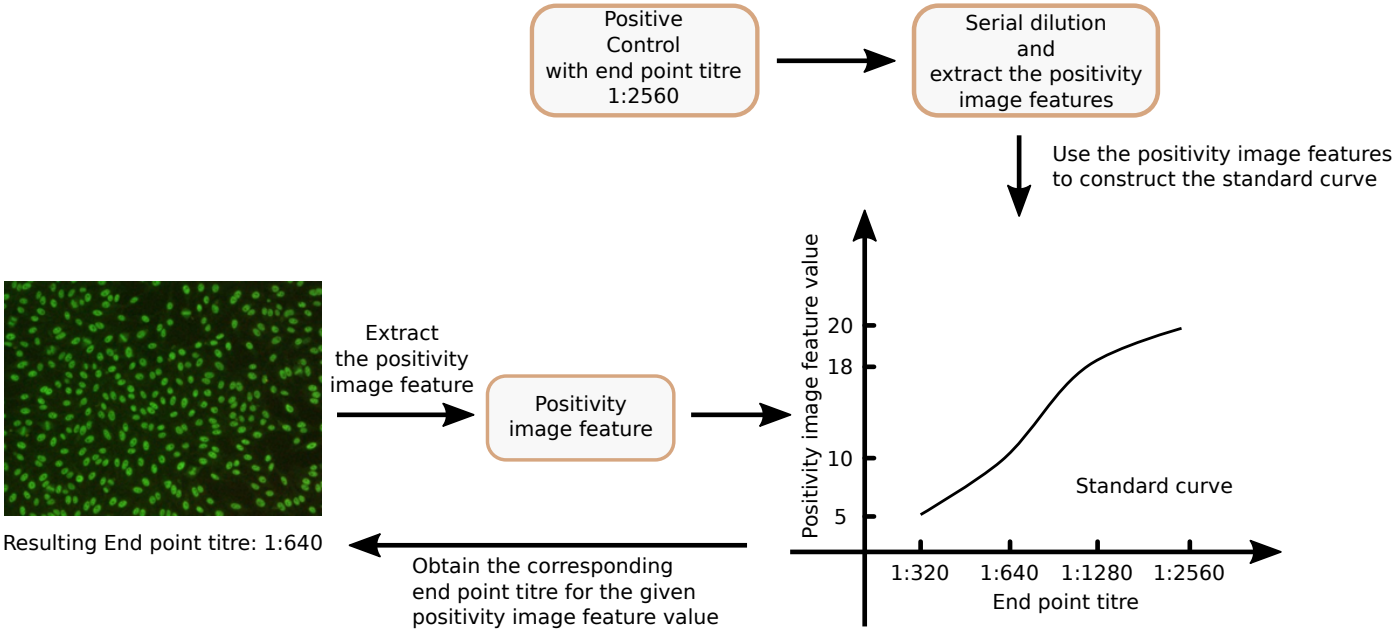

Fig. 2: An illustration on the typical steps for the positivity strength determination. First, a standard curve needs to be constructed from a positive sample with known end point titre. Then given a sample, the end point titre can be determined by first extracting the positivity image feature and use the standard curve to obtain the corresponding end point titre.

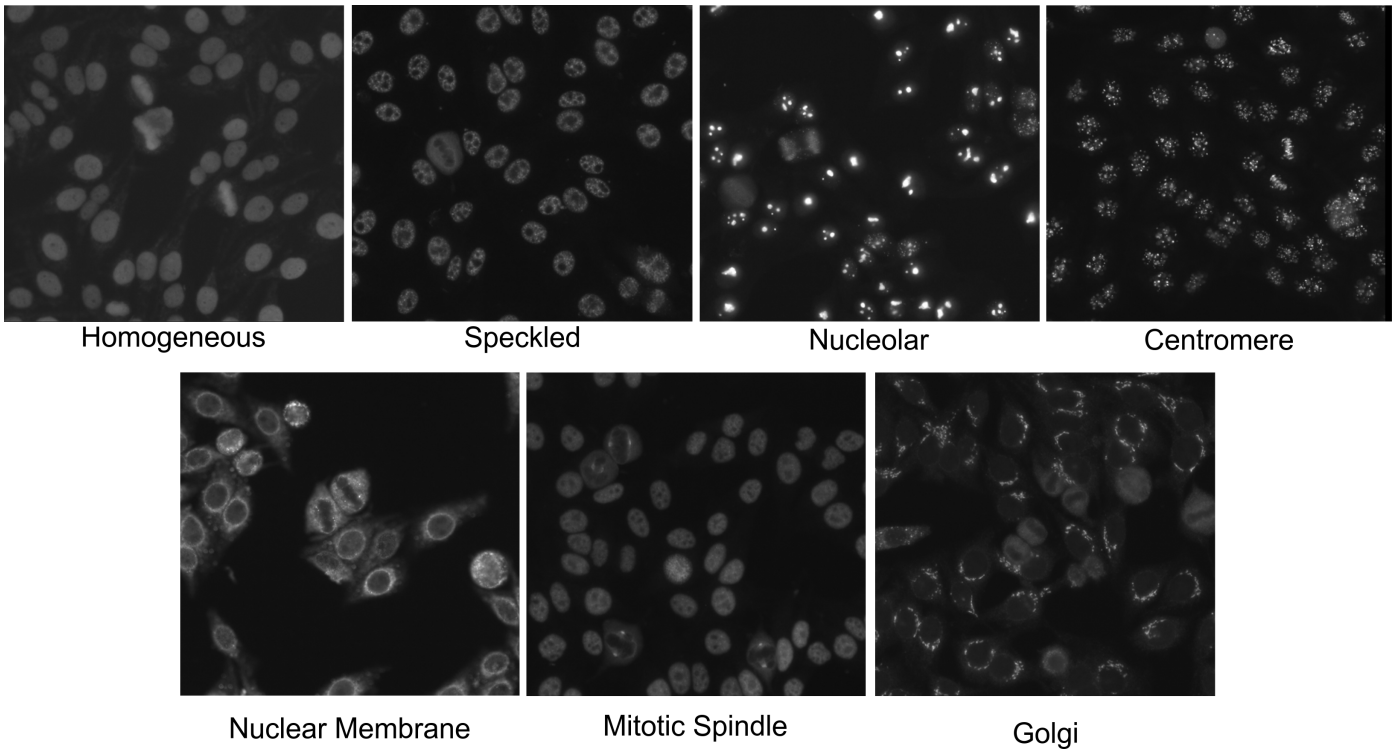

Fig. 3: Some examples of ANA specimen images with different ANA patterns: homogeneous, speckled, nucleolar, centromere, nuclear membrane, mitotic spindle and Golgi.

Early attempts in solving this problem focused on the direct segmentation approach (*i.e.* to segment the cells by developing a robust image segmentation algorithm). Early works by Huang et al. (2008a,b) describe two segmentation algorithms which are used for different group of patterns. They argue that separate approaches are required to handle variations between different ANA patterns. For instance, the cells of homogeneous patterns should be much easier to segment as they have well defined boundaries. On the other hand, centromere patterns will need different treatment as only the centromere elements are stained. We refer to Fig. 3 to see the difference between homogeneous and centromere patterns. In Perner et al. (2002)

a standard Otsu's thresholding approach is used. Percannella et al. (2012) propose an auto-learning method. The key concept is to apply an image reconstruction method for a preliminary image segmentation. Then, the result is used to train a classifier to control the dilation process.

Most early works had limited success due to variation of the HEp-2 cell appearance according to: (1) the ANA pattern; (2) the positivity strength (*i.e.* weak positives may have less defined features as they have duller images); and (3) the cell cycle (*i.e.* mitotic cells generally have different appearance to the interphase cells). To that end, recently Tonti et al. (2015) proposed a segmentation approach using an adaptive marker-controlled watershed approach. The adaptive marker selection is used to reduce the watershed algorithm to over-segmentation error and uneven illumination. Again, in this work, they group the images into two categories: (1) images with smooth texture and images with rough texture. Unfortunately, although their method showed significant improvement over the existing methods, the method is only evaluated on five ANA patterns. Thus, its efficacy on the other patterns such as mid-body and Golgi is not known. Another recent contribution on HEp-2 cells segmentation comes from Jiang et al. (2015), which propose a verification-based multithreshold probing approach. In the nutshell, the approach first generates multiple binary segmentation images by varying the image thresholding value. Then each segmented blob from a binary image is verified according to criteria such as area size, shape and position. Only segmented blobs that are passed this assessment are considered as valid segmented cell blobs. In the end, the binary images are combined together as the final segmentation result. Again, their approach is tested on four common ANA patterns such as homogeneous, speckled, nucleoar and centromere.

The other approach to segment HEp-2 cells is via a secondary channel (*i.e.* indirect segmentation). In this setting, the segmentation is done indirectly from the ANA image. For instance, as described in Hobson et al. (2015), it is used a secondary fluorescence counter stain called 4',6-diamidino-2-phenylindole (DAPI). As DAPI is directed against cell DNA, the cell nucleus will always be well delineated. Thus, a high quality segmentation can be determined. However, adding DAPI counter stain may require additional work and cost. To that end, the work discussed in Krause et al. uses Propidium Iodided (PI) to get the cell shape. Generally, indirect segmentation methods obtain high quality segmentation results. However, the shortcoming of indirect segmentation is that it requires extra cost for the second channel.

Despite its shortcomings, indirect segmentation enables us to significantly advance the whole field in developing CAD systems for ANA test. One example is that the work in Hobson et al. (2015) is used to produce an extremely large benchmarking dataset for addressing classification problems. We note that the datasets produced by the indirect segmentation method can also be used to further study the direct segmentation methods.

Challenges - Despite its high quality results, the indirect segmentation methods could be more expensive and slower than direct segmentation methods. Therefore, although the HEp-2 cell segmentation problem can be considered as a solved problem due to the advent of indirect segmentation approaches, these approaches can only be used under specific circumstances where it is possible to acquire images of the secondary channel. To that end, it is still preferably to develop a direct segmentation approach that obtains segmentation quality which is equal to those produced by the indirect segmentation approaches. Such an approach will also be more economical.

Public benchmarking datasets - Both the ICIP2013 (Hobson et al. (2015)) and the ICPR2014 (Hobson et al. (2014)) datasets were produced using an indirect segmentation approach, thus they are perfect datasets to advance the direct segmentation approaches. This is because both datasets provide images of the secondary channel which can be used as the truth segmentation labels. However, data cleansing is required to remove adjacent HEp-2 cells which may create a single connected region. This task can be carried out by first applying an automatic approach that separates single cell and connected cells based on the cell morphology features. Then, a manual process is required to completely eliminate the errors produced from the automatic process.

3.3.2. Mitotic-interphase cell cycle classification

In general, the cell cycle is divided into two phases: (1) interphase and (2) mitosis phases. Thus, given a cell image $\boldsymbol{I}^{[\text{cell}]}$, the goal of this subtask is to develop a function $f : \boldsymbol{I}^{[\text{cell}]} \mapsto \{mitosis, interphase\}$ that determines the cell cycle phase of the cell image.

Clinical usage - When reading ANA slides, often scientists consider the staining of cells in various phases of the cell cycle for the following reasons. First, cells in different phases cycle express different antigens or antigens in different concentration. This means, some antibodies may be more observable in a particular cell cycle. As such, the scientists need to see both interphase and mitotic cells before reporting the test result. Since the number of mitotic cells may be significantly lower than the interphase cells, then often the scientists need to observe a minimum number of mitotic cells before proceeding with the analysis.

Progress - We note that there are a only few works focusing on this problem. The following is the discussion of each work in chronological order. In early attempts, Foggia et al. (2010) use various hand-crafted features extracted from the image such as morphological, intensity histogram, grey-level co-occurence matrix and circular Linear Binary Pattern (LBP). Five classifiers are tested to perform the classification: (1) k Nearest Neighbor; (2) Multiple Layer Perceptron (MLP); (3) Support Vector Machine (SVM); (4) Naïve Bayes; and (5) Adaboost. From the evaluation, Adaboost is shown to be superior to the other classifiers.

As a consequence of the fact that the mitotic phase lasts less than one tenth of the whole cell lifecycle, we note that mitotic cells normally appear once or twice every 10 interphase cells in a specimen image. The seminal study by Foggia et al. (2010) is performed using a dataset that maintains the balance between the number of mitotic and interphase cells. Thus, in order to overcome this limitation and to account for the skewed distribution of the mitotic and interphase cells that normally occurs in the daily practice Percannella et al. (2011) and Iannello et al.

(2014) propose an approach based on multi-objective optimization. Technically, two classifiers are utilized: (1) a classifier trained on the original skewed dataset and (2) a classifier that employs a technique to address the class imbalance problem. A threshold based on the reliability value of the former classifier is used to choose which classifier is selected for the final output.

All the above cited papers on mitotic cell recognition adopt a public dataset including six common ANA staining patterns which are related to mitotic cells that can be grouped into two categories: (1) the cells whose chromatin is stained and (2) the cells whose chromatin is not stained. However, when other HEp-2 interphase staining patterns are considered, the set of categories of mitotic cells has to be extended. Two examples illustrating such situation are in fig. 4 where they are shown the mitotic cells associated to the mitotic spindle pattern and to the mixed patterns between homogeneous and speckled. Whilst, the mitotic spindle pattern has only the tiny mitotic spindle dots stained, when homogeneous and speckled patterns co-exist, both chromatin and the cell cytoplasm are stained.

To address the above issues, recently, Miros et al. (2015) described an approach that utilizes the DAPI channel to detect mitotic cells. The intuition is that as the DNA content of cells undergoing mitosis phase is doubled, then its fluorescent intensity will be much higher than the fluorescent intensity of interphase cells. The shape, intensity and texture features extracted from the DAPI image are used as the descriptors. From their experiments, the proposed approach is able to achieve good performance in an extremely difficult scenario. By using the DAPI channel, they are able to address the problem without making the restrictive assumption on the group of mitotic cells. In other words, the approach is to develop a structure that will all work for ANA patterns. The work also proposes an evaluation metric adapted from the pedestrian detection. Here, the performance is evaluated by using the false positive rate (FPR) and false negative rate (FNR). This evaluation metric is shown to have a relationship to the clinical applications. For instance, if each slide must have at least 10% mitotic cells, therefore to identify at a minimum of one mitotic in the population, one could allow for an FNR of 0.9. In addition, the work proposes a large scale benchmarking dataset comprising more than 23,000 interphase cells and 1,092 mitotic cells.

Challenges - Although it seems that the problem can be solved by reducing its complexity using the DAPI channel as discussed in Miros et al. (2015), the use of a secondary channel may not be practical or economical for laboratories. Therefore, it is still preferable to develop an approach that works on the primary channel. We note that, the results from works employing a secondary channel allow us to construct large datasets which in turn will enable us to further study the classification problem in the primary channel.

Public benchmarking datasets - There are two known publicly available datasets, respectively by Foggia et al. (2010) and by Miros et al. (2015). The former is the first dataset ever made publicly available to the scientific community and provides cell images extracted from the primary FITC channel and has been manually annotated by medical doctors (including the segmentation masks); the dataset in Miros et al. (2015) is very recent and offers both cell images of both the FITC and DAPI channels; annotation was obtained automatically and then verified by medical doctors.

3.3.3. HEp-2 cell classification

Recently, the HEp-2 cell classification problem appears to be attracting considerable attention from the community. The problem can be defined as follows. Each HEp-2 cell is represented by the three-tuple $\{\boldsymbol{I}, \boldsymbol{M}, \delta\}$ which consists of: the primary channel, the FITC image channel which carries the pattern information $\boldsymbol{I}$; (2) a binary cell mask image $\boldsymbol{M}$ that could be extracted manually or from the secondary channel and (3) the fluorescence intensity $\delta = \{\text{weak}, \text{strong}\}$. The goal is to construct a classifier that classifies the cell image into one of the known classes.

Due to its popularity, this paper does not provide a comprehensive discussion of this subject. Full discussion of the HEp-2 cell classification can be found in recent review papers published by Foggia et al. (2013, 2014); Hobson et al. (2015).

Challenges - Despite significant progress, the HEp-2 cell classification problem is still far from being solved. This is because current studies only focus on a small set of ANA patterns. Although the benchmarking platform described in Hobson et al. (2015) offers two less common patterns such as Golgi and nuclear membrane, there are other ANA patterns that are important requiring further study. Two examples are: (1) the classification problem of different speckled patterns and (2) the classification problem when multiple patterns co-exist. We discuss and explore these challenges in sections 3.4.1 and 3.4.2, respectively.

Public benchmarking datasets - As expected, due to its popularity, the HEp-2 cell classification problem has at least three publicly available datasets: (1) ICPR2012 (Foggia et al. (2013)), ICIP2013 (Hobson et al. (2015)) and SNPHEp-2 (Wiliem et al. (2013)).

The ICPR2012 dataset comprises 1,457 cells extracted from 28 specimen images. The dataset provides five ANA patterns: homogeneous, coarse speckled, fine speckled, nucleolar, centromere and one Anti-Neutrophil Cytoplasmic Antibody (ANCA) pattern: cytoplasmic. Each image was acquired by a fluorescence microscope with the following hardware: 40-fold magnification objective lens, 50W mercury vapour lamp and a colored microscope CCD camera. The cell image masks were hand labeled.

The SNPHep-2 dataset contains five ANA patterns: homogeneous, coarse speckled, fine speckled, nucleolar and centromere. There are 1,884 cells extracted from 40 specimen images. Each image was acquired by a fluorescence microscope with the following hardware: 20-fold magnification objective lens, LED illumination source and a monochrome high dynamic cooled microscope camera. Note that the SNPHEp-2 dataset is a smaller part of the ICIP2013 dataset.

Perhaps ICIP2013 is arguably one of the largest benchmarking dataset for HEp-2 cell classification. It comprises 68,429 cell images extracted from 419 unique positive patient sera. The dataset provides six patterns: homogeneous, speckled, nucleolar, centromere, Golgi and nuclear membrane. The first four patterns are common ANA patterns and the last two are less

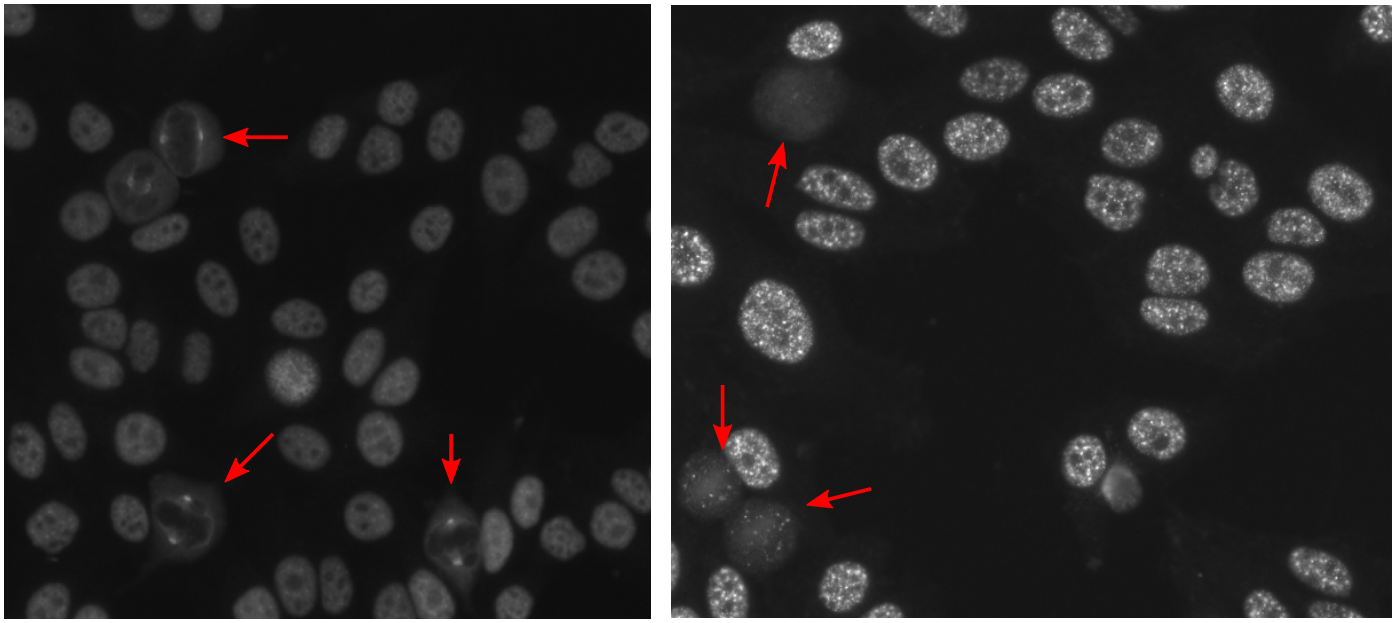

Fig. 4: Example images for mitotic spindle and mixture of homogeneous and speckled patterns. The red arrows point at the mitotic cells. Note that in both patterns mitotic cells are not within the categorization made by the existing cell cycle classification approaches. The mitotic cells of mitotic spindle pattern normally only have their mitotic spindle parts stained. Whilst, for mixed pattern, both chromatin and the nucleus are stained.

common patterns. The dataset was acquired using the same hardware as the SNPHEp-2 dataset.

3.3.4. ANA specimen image classification

Whilst there are many works addressing the problem of HEp-2 cell image classification, study of the ANA specimen image classification problem has started. This is because, it is difficult to address ANA specimen image classification problem without a good understanding of the HEp-2 cell image classification problem.

An ANA specimen image, or ANA image in short, generally contains a population of HEp-2 cells (refer to fig. 3). The general approach to address the ANA image classification problem is to classify each associated HEp-2 cell into one of the known classes. Then, the ANA image is classified into the class that frequently occurs in the population (Foggia et al. (2013)). We shall name this approach as the *MaxPopulation* approach. Unfortunately, although it seems to work, the MaxPopulation approach can not be used to classify several ANA patterns. A good counter example is the group of patterns related to HEp-2 cell cycle such as mitotic spindle Hobson et al. (2014).

One problem with the MaxPopulation approach is that it does not have a good model for the associated HEp-2 cell class population. It is shown in Hobson et al. (2014); Wiliem et al. (2015b); Manivannan et al. (2015) that modelling this population could be the key to obtain discriminative ANA image descriptors. More specifically, Hobson et al. (2014) proposed the CellBank approach that uses the HEp-2 cell image classification results to obtain the ANA image descriptor. First the approach counts the number of HEp-2 cell images classified as each known ANA class. Then the histogram of the counting is used as the descriptor. The CellBank approach has shown to be more effective than the MaxPopulation approach in classifying the mitotic spindle pattern. Another possible way is to model the associated HEp-2 cell population as a bag of words. In this setting, Wiliem et al. (2015b) propose the Bag of Cells (BoC) approach by considering each HEp-2 cell image classification result as a visual word. We note that apart from these baselines, more approaches have been proposed by the community (Manivannan et al. (2015, 2014a,b); Gragnaniello et al. (2014); Gupta et al. (2014); Cascio et al. (2014); Gao et al. (2014); Ensafi et al. (2014); Theodorakopoulos et al. (2014); Nanni et al. (2014); Codrescu (2014)). Due to the first benchmarking initiative hosted at the International Conference on Pattern Recognition 2014 (ICPR), we expect to see an explosion of works addressing this problem.

Challenges - The challenges in the ANA image classification problem are similar to the HEp-2 cell image classification problem: (1) more comprehensive testing of the existing approaches on various ANA patterns is required; (2) the study of pattern analysis on multiple patterns is required and (3) pattern analysis of various speckled pattern is also required. Whilst it is clear for the first challenge, it is not trivial to understand the importance of the second and third challenges. For the sake of clarity, we defer the discussion of the second challenge in Section 3.4.2. One application of the third challenge is also discussed in Section 3.4.1.

Another challenge that could be worth pursuing is to redefine the ANA image classification problem into antibody detection problems. In this case, instead of classifying a given ANA image into one of the known class, the goal is to detect the existence of specific antibodies.

Public benchmarking datasets - There are two benchmarking datasets for this purpose: (1) ICPR2012 (Foggia et al. (2013)) and (2) ICPR2014 Hobson et al. (2014). Unfortunately, the ICPR2012 dataset is insufficient to perform reasonable evaluation. In the other hand, ICPR2014 offers 1,000 ANA images which has much higher number of instances.

3.4. Discussions on challenges of HEp-2 cell and ANA image classification problems

In this section, we discuss some challenges shared by both HEp-2 cell and ANA image classification problems. In addition, we also supply some discussions on possible directions worth pursuing in this domain.

3.4.1. Dense Fine Speckled (DFS) 70 detection

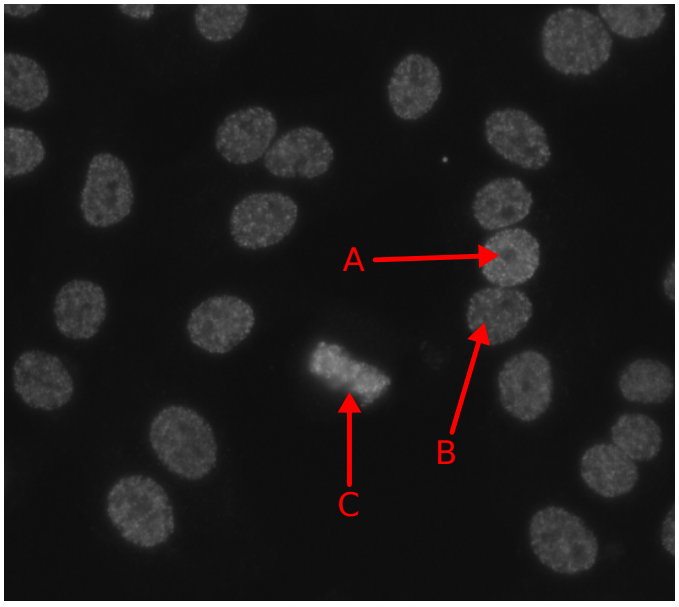

Fig. 5: Characteristics of typical DFS-70 patterns can be described as follows Mariz et al. (2011): (A) Nucleolus is not stained; (B) Grainy interphase chromatin and (C) Bright staining of the mitotic cell chromosome.

Recently, it has been shown that DFS70 auto-antibodies were detected exclusively in healthy individual (Mariz et al. (2011); Mahler et al. (2012)). In addition, the antibodies were also found in individuals with Systemic Auto-immune Rheumatic Diseases (SARD)-free within a four years follow up period (Mariz et al. (2011)). In other words, patients that only have DFS70 detected may suggest that they do not have SARD.

Clinical usage - As mentioned, detecting DFS70 pattern is useful in clinical analysis due to the recent finding suggesting that when patients only have DFS-70, then the patients may not have SARD.

When a patient has DFS70, its ANA image will exhibit a speckled pattern on the interphase HEp-2 cells and strong staining in the metaphase chromosome (Mariz et al. (2011)) as shown in Fig. 5. Therefore, the goal of this task is to detect the existence of DFS70 for any identified speckled pattern. In addition, it is also important to detect any secondary ANA pattern as one cannot exclude CTD or SARD diagnosis when additional patterns exist.

3.4.2. Mixed pattern analysis

Clinical usage - Mixed patterns have been one of the issues that have not been well studied yet in CAD system development. However, the role of mixed patterns is extremely important to the accuracy of ANA test report. For instance, if DFS70 is found and another ANA pattern co-exists in the ANA image, SARD should not be excluded. In addition, when two different patterns existed within a patient serum, then all the antibodies associated to these two patterns must be reported.

Currently, one common way to determine whether a patient has mixed patterns is to apply a manual serial dilution method on the patient serum to reduce the strength of the strongest/masking pattern. Multiple patterns are deemed to exist when another pattern is uncovered after the strength of the masking pattern drops. Unfortunately, despite its efficacy, applying serial dilution method to every patient could be extremely costly for a laboratory. Thus, one solution is to apply the serial dilution on patients having extremely high ANA positivity.

There are two main problems that have to be studied in order to develop a good CAD system: (1) the mixed pattern detection problem and (2) the mixed pattern enumeration problem.

Mixed pattern detection problem - This is a binary classification problem that will determine whether an ANA image has multiple patterns or just a single pattern.

Mixed pattern enumeration problem - Once an ANA image is determined to have multiple patterns, then the next problem is to enumerate and identify the patterns appearing in the ANA image.

3.5. Cross-hardware and assay analysis

Clinical usage - Often a CAD system is designed and tested only on a particular hardware and assay. However, this will hinder the adoption of the system into the laboratories as each one may have different needs and capabilities. So a CAD system should be able to address the hardware and assay variations.

To address this problem, there are two possible ways: (1) to develop a "calibration" or "training" phase before the system is used; (2) to design the system that can operate on images produced from different hardware and assay from the training.

Domain adaptation techniques can be applied to develop the first solution. Whilst, to develop the second solution one needs to develop a set of features and/or learning algorithm that are invariant to hardware and assay changes.

To begin studying the problem, we need to have at least one evaluation protocol. One possible protocol would be to train on the ICPR2012 dataset and test on the ICIP2013 dataset, and vice versa. Images produced for the ICPR2012 dataset are captured using different hardware and assays to the ICIP2013.

4. Conclusions

There has been a great effort to develop CAD systems for the ANA test via the Indirect Immunofluorescence protocol on HEp-2 cells. Some excellent works have advanced our understanding in developing such systems. However, despite the significant progress, we are still far from achieving our final goal.

Some challenges have been identified and discussed. For instance, there has been little work on HEp-2 pattern analysis when multiple patterns co-exist. Addressing these challenges will significantly advance the field. We also note that some challenges may be difficult to address. In this case, an indirect method could be a means to study the problem. For example, the use of a secondary counter stain enables us to generate very large benchmarking dataset for various problems such as cell segmentation, mitotic classification and HEp-2 pattern analysis problems.

Acknowledgements

This work has been partly funded by Sullivan Nicolaides Pathology, Australia and the Australian Research Council (ARC) Linkage Projects Grant LP130100230. Arnold Wiliem is funded by the Advance Queensland Early-career Research Fellowship.

References

Bel Haj Ali, W., Piro, P., Giampaglia, D., Pourcher, T., Barlaud, M., 2012. Biological cell classification using bio-inspired descriptor in a boosting k-NN framework, in: International Conference on Pattern Recognition (ICPR), pp. 1–6.

Bizzaro, N., Antico, A., Platzgummer, S., Tonutti, E., Bassetti, D., Pesente, F., Tozzoli, R., Tampoia, M., Villalta, D., 2014. Automated antinuclear immunofluorescence antibody screening: A comparative study of six computer-aided diagnostic systems. Autoimmunity Reviews 13, 292–298.

Bizzaro, N., Tozzoli, R., Tonutti, E., Piazza, A., Manoni, F., Ghirardello, A., Bassetti, D., Villalta, D., Pradella, M., Rizzotti, P., 1998. Variability between methods to determine ANA, anti-dsDNA and anti-ENA autoantibodies: a collaborative study with the biomedical industry. Journal of Immunological Methods 219, 99–107.

Cascio, D., Taormina, V., Cipolla, M., Fauci, F., Vasile, S.M., Raso, G., 2014. HEp-2 cell classification with heterogeneous classes-processes based on k-nearest neighbours, in: Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), 2014 1st Workshop on, pp. 10–15. doi:10.1109/I3A.2014.17.

Cataldo, S.D., Bottino, A., Ficarra, E., Macii, E., 2012. Applying textural features to the classification of HEp-2 cell patterns in iif images, in: International Conference on Pattern Recognition (ICPR), pp. 3349–3352.

CDC, 1996. Quality assurance for the indirect immunofluorescence test for auto-antibodies to nuclear antigen (IF-ANA): Approved guideline. volume 16. NCCLS I/LA2-A. Wayne, PA.

Codrescu, C., 2014. Quadratic recurrent finite impulse response mlp for indirect immunofluorescence image recognition, in: Proceedings of the 2014 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images, IEEE Computer Society, Washington, DC, USA. pp. 49–52.

Cordelli, E., Soda, P., 2011. Color to grayscale staining pattern representation in IIF, in: International Symposium on Computer-Based Medical Systems, pp. 1–6.

Dawson, A.E., . Can we change the way we screen?: Thinprep imaging system. Cancer Cytopathology 102.

Elbischger, P., Geerts, S., Sander, K., Ziervogel-Lukas, G., Sinah, P., 2009. Algorithmic framework for HEp-2 fluorescence pattern classification to aid auto-immune diseases diagnosis, in: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 562–565.

Ensafi, S., Lu, S., Kassim, A.A., Tan, C.L., 2014. A bag of words based approach for classification of HEp-2 cell images, in: Proceedings of the 2014 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images, IEEE Computer Society, Washington, DC, USA. pp. 29–32.

Ersoy, I., Bunyak, F., Peng, J., Palaniappan, K., 2012. HEp-2 cell classification in IIF images using shareboost, in: International Conference on Pattern Recognition (ICPR), pp. 3362–3365.

Faraki, M., Harandi, M.T., Wiliem, A., Lovell, B.C., 2014. Fisher tensors for classifying human epithelial cells. Pattern Recognition 47, 2348–2359.

Flessland, K.A., Landicho, H.R., Borden, K.K., Prince, H.E., 2002. Performance characteristics of the PolyTiter immunofluorescent titration system for determination of antinuclear antibody endpoint dilution. Clinical and Diagnostic Laboratory Immunology 9, 329–332.

Foggia, P., Percannella, G., Saggese, A., Vento, M., 2014. Pattern recognition in stained HEp-2 cells: Where are we now? Pattern Recognition 47, 2305–2314.

Foggia, P., Percannella, G., Soda, P., Vento, M., 2010. Early experiences in mitotic cells recognition on HEp-2 slides, in: IEEE International Symposium on Computer-Based Medical Systems (CBMS), pp. 38–43.

Foggia, P., Percannella, G., Soda, P., Vento, M., 2013. Benchmarking HEp-2 cells classification methods. IEEE transactions on medical imaging .

Foggia, P., Sansone, C., Tortorella, F., Vento, M., 1999. Multiclassification: reject criteria for the bayesian combiner. Pattern Recognition 32, 1435 – 1447.

Gao, Z., Zhang, J., Zhou, L., Wang, L., 2014. HEp-2 image classification with convolutional neural networks, in: 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), pp. 24–28.

Ghosh, S., Chaudhary, V., 2012. Feature analysis for automatic classification of HEp-2 florescence patterns: Computer-aided diagnosis of auto-immune diseases, in: International Conference on Pattern Recognition (ICPR), pp. 174–177.

Gragnaniello, D., Sansone, C., Verdoliva, L., 2014. Biologically-inpired dense local descriptor for indirect immunofluorescence image classification, in: 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), pp. 1–5.

Gupta, K., Gupta, V., Bhavsar, A., Dileep, A.D., 2014. Class-specific hierarchical classification of HEp-2 cell images: The case of two classes, in: 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), pp. 6–9.

Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B., 2009. Histopathological image analysis: A review. Biomedical Engineering, IEEE Reviews in 2, 147–171.

Hiemann, R., 2013. Method for end-titre determination and the evaluation thereof by means of an indirect immunofluorescence assay. URL: http://www.google.com/patents/CA2702421C?cl=en. cA Patent 2,702,421.

Hiemann, R., Büttner, T., Krieger, T., Roggenbuck, D., Sack, U., Conrad, K., 2009. Challenges of automated screening and differentiation of non-organ specific autoantibodies on HEp-2 cells. Autoimmunity Reviews 9, 17–22.

Hobson, P., Lovell, B.C., Percannella, G., Vento, M., Wiliem, A., 2014. Classifying anti-nuclear antibodies hep-2 images: A benchmarking platform, in: Pattern Recognition (ICPR), 2014 22nd International Conference on, pp. 3233–3238. doi:10.1109/ICPR.2014.557.

Hobson, P., Lovell, B.C., Percannella, G., Vento, M., Wiliem, A., 2015. Benchmarking human epithelial type 2 interphase cells classification methods on a very large dataset. Artif. Intell. Med. 65, 239–250.

Hsieh, T., Huang, Y., Chung, C., Huang, Y., 2009. HEp-2 cell classification in indirect immunofluorescence images, in: Int. Conf. Information, Communications and Signal Processing, pp. 1–4.

Huang, Y., Chung, C., Hsieh, T., Jao, Y., 2008a. Outline detection for the HEp-2 cell in indirect immunfluorescence images using watershed segmentation, in: IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, pp. 423–427.

Huang, Y., Jao, Y., Hsieh, T., Chung, C., 2008b. Adaptive automatic segmentation of HEp-2 cells in indirect immunofluorescence images, in: IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, pp. 418–422.

Iannello, G., Percannella, G., Soda, P., Vento, M., 2014. Mitotic cells recognition in hep-2 images. Pattern Recognition Letters 45, 136 – 144.

Jiang, X., Percannella, G., Vento, M., 2015. A verification-based multithreshold probing approach to HEp-2 cell semgentation, in: Computer Analysis of Images and Patterns. volume 9257 of *Lecture Notes in Computer Science*, pp. 266–276.

Krause, C., Ens, K., Fechner, K., Voigt, J., Fraune, J., Rohẅader, E., Hahn, M., Danckwardt, M., Feirer, C., Barth, E., Martinetz, T., Stöcker, W., . Europattern suite technology for computer-aided immunofluorescence microscopy in autoantibody diagnostics. Cancer Cytopathology 102.

Labati, R.D., Piuri, V., Scotti, F., 2011. All-IDB: the acute lymphoblastic leukemia image database for image processing, in: 2011 18th IEEE International Conference on Image Processing (ICIP), IEEE. pp. 2045–2048.

Li, K., Yin, J., 2012. Multiclass boosting svm using different texture features in HEp-2 cell staining pattern classification, in: International Conference on Pattern Recognition (ICPR), pp. 170–173.

Mahler, M., Hanly, J.G., Fritzler, M.F., 2012. Importance of the dense fine speckled pattern of HEp-2 cells and anti-dfs70 antibodies for the diagnosis of systemic autoimmune diseases. Autoimmunity Reviews 11, 642–645.

Manivannan, S., Li, W., Akbar, S., Wang, R., Zhang, J., McKenna, S.J., 2014a. HEp-2 cell classification using multi-resolution local patterns and ensemble svms, in: 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), pp. 37–40.

Manivannan, S., Li, W., Akbar, S., Wang, R., Zhang, J., McKenna, S.J., 2014b. HEp-2 specimen classification using multi-resolution local patterns and svm, in: 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), pp. 41–44.

Manivannan, S., Li, W., Akbar, S., Wang, R., Zhang, J., McKenna, S.J., 2015. An automated pattern recognition system for classifying indirect immunfluorescence images of HEp-2 cells and specimens. Pattern Recognition .

Mariz, H., Sato, E., Barbosa, S.H., Rodrigues, S.H.Dellavance, A., Andrade, L., 2011. Pattern on the antinuclear antibody-HEp-2 test is a critical parameter for discriminating antinuclear antibody-positive healthy individuals and patients with autoimmune rheumatic diseases. Arthritis and Rheumatism 63, 191–200.

Meroni, P.L., Schur, P.H., 2010. ANA screening: an old test with new recommendations. Annals of the Rheumatic Diseases 69, 1420 –1422.

Miros, A., Wiliem, A., Holohan, K., Ball, L., Hobson, P., Lovell, B.C., 2015. A benchmarking platform for mitotic cell classification of ana iif HEp-2 images, in: IEEE Conference on Digital Image Computer: Techniques and Applications (DICTA), pp. 1–6.

Nakabayashi, T., Kumagai, T., Yamauchi, K., Sugano, M., Kuramoto, A., Fujita, K., Hidaka, H., Tozuka, M., 2001. Evaluation of the automatic fluorescent image analyzer, image titer, for quantitative analysis of antinuclear antibodies. American Journal of Clinical Pathology 115, 424–429.

Nanni, L., Paci, M., Caetano dos Santos, F.L., Hyttinen, J., 2014. Morphological and texture features for HEp-2 cells classification, in: 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), pp. 45–48.

Percannella, G., Soda, P., Vento, M., 2011. Mitotic HEp-2 cells recognition under class skew, in: Image Analysis and Processing (ICIAP), pp. 353–362.

Percannella, G., Soda, P., Vento, M., 2012. A classification-based approach to segment HEp-2 cells, in: International Symposium on Computer-based Medical Systems (CBMS), pp. 1–5.

Perner, P., Perner, H., Mller, B., 2002. Mining knowledge for HEp-2 cell image classification. Artificial Intelligence in Medicine 26, 161–173.

Pham, B., Albarede, S., Guyard, A., Burg, E., Maisonneuve, P., 2005. Impact of external quality assessment on antinuclear antibody detection performance. Lupus 14, 113–119.

Samak, A., Wiliem, A., Hobson, P., Walsh, M., Ditchmen, T., Troskie, A., Barksdale, S., Edwards, R., Jennings, A., Lovell, B.C., 2015. An optimization approach to scanning skin direct immunofluorescence specimens, in: IEEE Conference on Digital Image Computer: Techniques and Applications (DICTA), pp. 1–8.

Snell, V., Christmas, W., Kittler, J., 2012. Texture and shape in fluorescence pattern identification for auto-immune disease diagnosis, in: International Conference on Pattern Recognition (ICPR), pp. 33750–33753.

Soda, P., Iannello, G., 2009. A multi-expert system to classify fluorescent intensity in antinuclear autoantibodies testing. Computer Based Medical Systems , 219–224.

Soda, P., Iannello, G., Vento, M., 2009. A multiple expert system for classifying fluorescent intensity in antinuclear autoantibodies analysis. Pattern Analysis and Application 12, 215–226.

Strandmark, P., Ulén, J., Kahl, F., 2012. HEp-2 staining pattern classification, in: Int. Conf. Pattern Recognition, pp. 33–36.

Tadrous, P.J., 2010. Computer-assisted screening of ziehl-neelsen-stained tissue for mycobacteria. algorithm design and preliminary studies on 2,000 images. American Journal of Clinical Pathology 133, 849–858.

Theodorakopoulos, I., Kastaniotis, D., Economou, G., Fotopoulos, S., 2012. HEp-2 cells classification via fusion of morphological and textural features, in: IEEE International Conference on Bioinformatics and Bioengineering (BIBE), pp. 689–694.

Theodorakopoulos, I., Kastaniotis, D., Economou, G., Fotopoulos, S., 2014. HEp-2 cell classification using morphological features and a bundle of local gradients descriptors, in: 1st Workshop on Pattern Recognition Techniques for Indirect Immunofluorescence Images (I3A), pp. 33–36.

Thibault, G., Angulo, J., 2012. Efficient statistical/morphological cell texture characterization and classification, in: International Conference on Pattern Recognition (ICPR), pp. 2440–2443.

Tonti, S., Cataldo, S.D., Bottino, A., Ficarra, E., 2015. An automated approach to the segmentation of HEp-2 cells for the indirect immunofluorescence ana test. Computerized Medical Imaging and Graphics 40, 62–69.

Wiik, A.S., Hier-Madsen, M., Forslid, J., Charles, P., Meyrowitsch, J., . Antinuclear antibodies: A contemporary nomenclature using HEp-2 cells. Journal of Autoimmunity In Press, Corrected Proof.

Wilbur, D.C., 2011. Digital cytology: current state of the art and prospects for the future. Acta Cytologica 55, 227–238.

Wiliem, A., Hobson, P., Minchin, R., Lovell, B., 2011. An automatic image based single dilution method for end point titre quantitation of antinuclear antibodies tests using HEp-2 cells, in: Digital Image Computing: Techniques and Applications, Noosa, Australia. pp. 1–6.

Wiliem, A., Hobson, P., Minchin, R.F., Lovell, B.C., 2015a. A bag of cells approach for antinuclear antibodies HEp-2 image classification. Cytometry Part A 87, 549–557.

Wiliem, A., Hobson, P., Minchin, R.F., Lovell, B.C., 2015b. A bag of cells approach for antinuclear antibodies HEp-2 image classification. Cytometry Part A 87, 549–547.

Wiliem, A., Wong, Y., Sanderson, C., Hobson, P., Chen, S., Lovell, B.C., 2013. Classification of human epithelial type 2 cell indirect immunofluoresence images via codebook based descriptors, in: IEEE Workshop on Applications of Computer Vision (WACV), pp. 95–102.

Yang, Y., Wiliem, A., Alavi, A., Lovell, B.C., Hobson, P., 2014. Visual learning and classification of human epithelial type 2 cell images through spontaneous activity patterns. Pattern Recognition 47, 2325–2337.