An Early Experience Toward Developing Computer Aided Diagnosis for Gram-Stained Smears Images

Johanna Carvajal[†], Daniel F. Smith[†], Kun Zhao[†], Arnold Wiliem[†], Paul Finucane[◇], Peter Hobson[◇], Anthony Jennings[◇], Rodney McDougall[◇], Brian Lovell[†]

[†]The University of Queensland, Australia

[◇]Sullivan Nicolaides Pathology, Australia

Abstract

Gram stained direct smears test is clinically useful in early identification of infections. Unfortunately, this practice is considered time consuming and labour intensive. Most existing effort in this area is to perform high-magnification analysis of images taken from manually selected areas. In this paper, we address the problem of the automatic selection of candidate areas (or patches) for subsequent high-magnification analysis. Specifically, we explore and study the possibility of selecting good working areas based on low-magnification images, where bacteria are likely to be found when viewed in high-magnification images. To this end, we develop an approach to classify the areas of interest according to the textural information of the image patch. We explore and study the efficacy of traditional textural features such as Histogram of Gradients, Local Binary Patterns, and 2 Dimensional Discrete Cosine Transform. Experiments show that the best variant method is able to select working areas where it is likely to find bacteria in high-powered objective images in a wide range of images.

1. Introduction

Pathology is a key foundation in our modern healthcare system as often results from pathology tests help physicians in making accurate diagnoses. Unfortunately, some pathology tests requiring microscopy readings are considered subjective, time consuming and labour intensive. Recent effort to address these issues is by developing Computer Aided Diagnosis (CAD) systems that allow the pathologists to spend their precious time on the positive cases [1, 11, 16, 34]. In fact, the CAD systems can be used for the following scenarios: (1) rejecting easy negative cases [34]; (2) providing estimated numbers of objects of interest [4, 21] and (3) localising good areas to analyse [11, 16].

Whilst many CAD systems have been developed, in this work we address the development of CAD systems for the Gram stained direct smears test. For the sake of brevity we shall name this test as the Gram stain test. The Gram stain test offers a fast and inexpensive way for physicians to manage patients with bacterial infections [22]. Despite its benefits, the test still requires manual microscopy readings. To our knowledge, the highest state of automation for this test is only at the pre-analytical process [2]. Existing works [6, 13] primarily aim to detect the existence of bacteria, leukocytes and epithelial cells. Although progress has been made by these works, unfortunately, these works assume that the areas of interest have been pre-selected. Automatic area of interest selection is an important step to have a fully automated CAD system for this application. This is because, generally, the analysis is done by using high-powered objectives (e.g. 100x) and therefore capturing the whole slide using the high-powered objective is not practical for real scenarios due to extremely long scanning time, large image storage, and processing time. Moreover, when making a diagnosis, it is a common practice only to examine a fixed number of regions of each slide, rather than the entire slide.

The problem of automatically selecting areas of interest has been investigated in various pathology tests [1, 11]. For instance, Angulo and Flandrin [1] use a mathematical morphology approach to select areas of interest for peripheral blood smears. Unlike images taken from the Gram stain test, the peripheral blood smears images are relatively more consistent. Thus, the segmentation approach such as proposed by Angulo and Flandrin in [1] may require manual tuning for each slide; thus, making it unsuitable to address our problem. Another approach is to train an area of interest detection by constructing a training dataset consisting of positive (i.e. the locations of interest) and negative (i.e. background/other locations) samples. Some approaches proposed for histopathology image analysis follow this fashion [11]. However, these works aim at detecting much bigger structures such as suspicious lesions. Un-

like these works, our main aim is to find good areas using low-magnification images to look for the existence of bacteria when viewed at higher magnification. In other words, we are looking for areas to examine for bacteria, not bacteria itself. As bacteria size is generally extremely small. For example, an average coccus-shaped bacterium size is about 0.5-1.0 micrometer in diameter and an average bacillus size is 0.5-1.0 micrometer wide by 1.0-4.0 micrometer long [12].

Bacteria are almost impossible to see using naked eyes under low-magnification objective such as 2.5x. This means that in this paper we ask a much more difficult question that states whether it is possible to extract useful information extracted from low-magnification objective images that will indicate good areas to look for the existence of bacteria.

Contributions: We list our contributions as follows. (1) In order to stimulate the interest in the community on this problem, we propose a benchmark evaluation dataset. The dataset comprises 1,293 images which were extracted from various body parts of patients with different medical conditions. It is expected that the dataset will be expanded in the future. (2) We explore various design choices such as feature choices in developing the area of interest detector. The best performing method will set the baseline for the proposed dataset.

We continue our paper as follows. Section 2 presents related works, followed by the problem definition in Section 3. The dataset is described in Section 4. Section 5 describes our proposed baseline approach. Experiments are presented in Section 6. Finally, we present the main findings and future work in Section 7.

2. Related Work

The discussion in this section is divided into three subsections: (1) the Gram stained direct smears test; (2) features used in microscopic image analysis, and (3) the problem of imbalanced training data.

Gram stained direct smears test - Existing works in this domain focus on the detection, recognition, counting, and classification of bacteria [5, 6, 14]. For instance, a contour technique is used for bacterial shape detection and identification in [5]. Hiremath *et al.* proposed an automatic classification of bacterial growth phases of bacilli cells in digital microscopic cell images [14]. To this end, geometric features were extracted and several classifiers were used, such as: k-NN, Neural Network, and Fuzzy classifiers. Finally, leukocyte and epithelial cell detection in Gram stain images was obtained via a multiple region covariance approach in [6]. The multiple region covariance approach provided an efficient way to classify cells.

Microscopic image descriptors - Computer vision texture descriptors have been used for microscopic image characterisation. Among them we can find Histogram of Gradients (HOG), Local Binary Patterns (LBP), 2-Dimensional Discrete Cosine Transform (2D-DCT), morphology, shape features, wavelets, and colour [19, 24, 35]. For instance, HOG features were used to extract anatomical information in brain images. If two anatomical structures in the brain images depict comparable appearance and are positioned in a similar area both should be considered as the same label class [29]. HOG features were also explored for renal disease detection using microscopic images of the kidneys [17].

LBP is another popular feature descriptor that has been used for microscopic images. Some of the LBP variants are examined for the classification of different types of human breast tissue images [36]. LBP is also employed to classify human embryo microscope images aiming to automatically evaluate whether the embryo is suitable for implantation in in-vitro fertilisation [37]. Texture features based on LBP, higher-order spectra, and Laws texture energy are used to describe microscopic images containing oral sub-mucous fibrosis for the oral cancer identification [18].

Besides HOG and LBP, 2D-DCT features have been also employed in microscopic images. 2D-DCT are extracted to recognise blast cells in the bone marrow of patients affected from Acute Lymphoblastic Leukaemia (ALL) [23]. The aim is to describe microscopic images with specimen blood samples and classify the image into benign or malignant. The combination of HOG and LBP features has been used to recognise signs of some lung lesions using computed tomography images [33]. HOG and 2D-DCT are fused to identify organelles from fluorescence microscopy images of HeLa cells [15]. Organelles are employed to individualise new-found genes or genes with an unknown function. A wide range of feature types are amalgamated for the detection and classification of cancer in microscopic biopsy images in [19]. The feature types used encompass texture, morphology, shape, HOG, wavelets, colour, among others.

Deep learning has become very successful for several applications within the computer vision community [28, 30]. The success is mainly due to the latest advances in central and graphics processing units and the availability of enormous amount of data [28]. However, it has been reported that deep learning techniques are powerful only when the datasets are large which, for medical images, can be infeasible or difficult to collect sufficient data [30].

Imbalanced data - An imbalanced dataset is defined as a dataset that considerably contains more data samples of one class and fewer samples of the other class [8, 20]. Imbalanced data is a common limitation in the medical image field given the difficulty to find positive (interesting,

alarm-worthy) samples, while negative (uninteresting) samples are easier to obtain. For our Gram stain problem, we aim to automate the detection of some regions that, at high-magnification, can help in the detection of bacteria. The number of interesting regions can be easily surpassed by the number of regions that are not ideal for identifying the presence of bacteria.

There are different ways to deal with the imbalanced data problem. For example, some approaches simply adjust or optimise the relative weight parameter in the Support Vector Machine (SVM) during the training stage [38]. Another approach uses sampling techniques (under-sampling, over-sampling, and/or a combination of both) [8]. Hard Negative Mining (HNM) is an iterative technique where an initial training set is created using all positive samples and only a subset of the negative samples. The remaining set of negative samples is tested with the trained model and the resulting false positive samples replace the initial subset. This process is iterated several times until the system stabilises or until a maximum number of iteration is reached.

3. Problem Definition

Let $\boldsymbol{I}^h$ be the high-magnification tile image and $\boldsymbol{I}^l$ be the low-magnification tile image captured at the same location within a slide using different objective magnification. Generally, either 100x or 63x objectives are used to capture $\boldsymbol{I}^h$ (*i.e.,* hence we call this image a high-magnification objective image), whereas $\boldsymbol{I}^l$ is taken using a much lower objective magnification such as 5x or 2.5x. As mentioned, the time to scan a slide using the high-magnification objective is much longer than using the low-magnification objectives. This is because a higher number of tile images are required to cover the whole slide using the high-magnification objective compared to the low-magnification objective. However, not all of the slide areas are useful to scan for diagnosis purposes. Indeed, the long scanning times and large image sizes make it desirable to only scan a subset of $\boldsymbol{I}^h$ images. To this end, it is desirable to select and scan a small set of high-magnification tile images that would sufficiently cater for the pathologists/scientists to make their decisions. Generally, the tile images in this selected set may be useful to find objects of interest such as bacteria, leukocytes, and epithelial cells. In this work, the set is selected based upon the information from the low-magnification tile images.

Now we are ready to present the formal description of the problem. Let us assume that we have already scanned the whole slide using both low and high-magnification objectives. Let $\mathcal{H} = \{\boldsymbol{I}_i^h\}_{i=1}^N$ and $\mathcal{L} = \{\boldsymbol{I}_j^l\}_{j=1}^M$ be the set of tile images scanned using high-magnification and low-magnification, respectively; and $N >> M$. Let $\mathcal{R} \subset \mathcal{H}$ be the set of high-magnification objective tile images where necessary information for the pathologists/scientists is contained. In the ideal case where the full scan of high-magnification tile images are available (*i.e.,* $\mathcal{H}$ is provided), our aim is to develop a detector $\mathrm{f} : \boldsymbol{I}^h \in \mathcal{H} \mapsto \{-1, +1\}$, where $+1$ means $\boldsymbol{I}^h \in \mathcal{R}$. Basically, $\mathrm{f}(\cdot)$ is an area of interest detector. However, in a clinical scenario where a laboratory may have a high number of Gram stain tests per day, then it is almost impossible to perform a full scan of each Gram stain test using a high-magnification objective. Thus, the central problem considered in this paper is to develop $\hat{\mathrm{f}} : \boldsymbol{I}^l \in \mathcal{L} \mapsto \{-1, +1\}$, where $+1$ means the corresponding $\boldsymbol{I}^h$ taken from the same location as $\boldsymbol{I}^l$ is a member of $\mathcal{R}$.

4. Dataset

Gram stain is employed as an initial step in the identification of bacteria [3, 32]. Gram stain is a common staining technique used in microbiology to distinguish bacterial species into two groups: Gram-positive and Gram-negative [32]. Christian Gram developed this technique in 1883 [3]. Gram discovered the ability where bacteria either retain or lose the stain. The cells are stained either pink (Gram negative) or violet (Gram positive) depending on the cell wall properties.

One publicly available dataset with Gram stain images focuses on the classification of micro-organisms within the image slides [6]. The dataset used in [6] contains 150 colour images with positive and negative examples of leukocyte and epithelial cells. Another dataset presented in [5] contains 320 stained slide images not only using the Gram stain technique but also other different staining techniques. The problem addressed in [5] is the automatic detection of microbes and the extraction of bacterial clusters.

There are no datasets that can be used for the task of selecting candidate areas for subsequent high-magnification analysis. We have created a novel dataset. This dataset consists of eight Gram stain images. We used a Zeiss Imager Z2 microscope and the images are captured using a PixeLINK PL-B623CF camera. Figure 2 shows the slide images at 2.5x objective.

To create the dataset, 200 areas or patches were randomly selected per image. Each patch size is 51×38 pixels to match the size of the image when viewed with the 63x objective. In total, we have collected 1,600 patches from the eight different slide images. Three subjects independently labelled each patch. For the labelling procedure, each subject is shown three images: *(i)* an image captured at low resolution where the potential candidate area (patch) is highlighted, *(ii)* the patch itself, and *(iii)* the high-magnification image corresponding to that patch (See Figure 1 for an example). Based on that observation, each subject assigned a label number that can be one, two, three, or four. A label *four* indicates that the specific patch is considered a best candidate area for high-magnification, a label *three* is an area that pathologists may look if no bacteria are seen in ar-

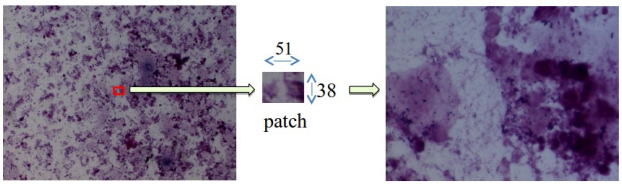


Figure 1: A patch captured using a low-magnification objective is used in conjunction with its high-magnification image to label the dataset. Best viewed in colour.

eas with labels *four*, label *two* is a dense and dark area where it is difficult to make a diagnosis, and label *one* is usually a background area or an area with some kind of unwanted artefacts such as dirt, dust, oil, etc. The results of the labelling process were analysed and we found that all subjects agreed for 1,293 images. We decided to discard the remaining 307 images in order to create an unbiased dataset. These images often contained out-of-focus or incorrectly stained areas where the subjects had different opinions on their usefulness. This dataset is available for download from the following URL: *http://www.itee.uq.edu.au/sas/datasets/*.

5. Proposed Baseline Approach

Using the randomly selected patches from low-magnification images, we extract the following descriptors: HOG, LBP, and 2D-DCT. Each patch size is $r \times c$. To obtain HOG descriptors, the horizontal and vertical gradients are calculated in small cells. The gradient of each patch contains two values, the magnitude and direction of gradient. HOG computes a histogram of oriented gradients per cell, normalises the result, and returns a descriptor which is the concatenation of the bins per cell [7]. The parameters for HOG are the cell size (pixels per cell) and the number of orientations (o). LBP is a popular texture descriptor which is locally computed by comparing each pixel with its neighbouring pixels [25]. To this end, two parameters are used, the number of points n and the radius r of the circle. DCT is a finite representation of a sequence of data points in terms of a sum of cosine functions oscillating at different frequencies [31]. 2D-DCT is a multi-dimension variant of DCT where DCTs are calculated along each dimension. In total, $r \times c$ coefficients are generated using 2D-DCT. Given that the energy coefficients are in descending order, only the first 2D-DCT coefficients are used which correspond to the top-left-most region ($q \times w$) of the resulting $r \times c$ matrix [27]. The first coefficients represent low-frequency information.

To deal with imbalanced data, we employ the hard negative mining (HNM) approach. We construct an initial training set using all positive samples and a subset of the negative samples. The model is trained using SVM with this initial training set. The model is tested using the remaining negative samples. The resulting false positives (hard negative samples) replace the negative samples in the original subset, always maintaining the same ratio of negative and positive samples. This process is repeated until the number of false positives is considerably reduced or until we reach a maximum number of iterations.

6. Experiments

We use the new dataset described in Section 4. Instead of using the four label categories of the original dataset, we combine the images labelled as *one* and *two* as negative samples, as those are samples representing areas that a pathologist will never select for high-magnification. As positive samples we use all samples with labelled as *four*. Samples labelled as *three* are not considered.

Each patch with size 51×38 pixels is converted into the gray scale colour. We use the leave-one-out cross validation approach (LOOCV). The LOOCV protocol takes out all the patches from one Gram stain slide image for testing and uses all patches from the seven remaining Gram stain slide images for training. This is executed for every testing image in a rotating basis, and the overall accuracy is obtained by averaging the accuracy of all validations.

We report the results using the Receiver Operating Characteristic curve (ROC curve). The ROC curve shows the performance of a binary classifier [10, 26]. Given a classifier and a testing sample, there are four possible outcomes. If the sample is positive and it is classified as positive, it is counted as a true positive, but if it is classified as negative, it is counted as a false negative. If the sample is negative and it is classified as negative, it is counted as a true negative, but if it is classified as positive, it is counted as a false positive. In a ROC curve the true positive rate is plotted as a function of the false positive rate for different thresholds [10]. The closer the curve follows the upper left corner, the higher the overall accuracy of the test [26].

For the calculation of HOG features we fix the cell size to 17×19 pixels and vary the number of orientations $o = \{2, 4, 8, 16, 32, 64\}$. For LBP features we vary both the radius $r = \{2, 3, 4\}$ and the number of points $n = \{2, 4, 6, 8, 10\}$. For 2D-DCT features we take the top-left-most region with $q = w$ with $q = \{1, 2, \cdots, 10\}$. Figure 3 depicts in *(a)*, a positive patch sample, and in *(b), (c)*, and *(d)* the visualisation of the three types of extracted features under consideration. It can be seen how HOG features in Figure 3*(b)* divide each patch into several cells and each cell is represented by a star which shows strength of the edge orientations in the histogram. The LBP features in Figure 3*(c)* depict the texture patterns. The 2D-DCT features are visualised using a heat-map in Figure 3*(d)*. Look-

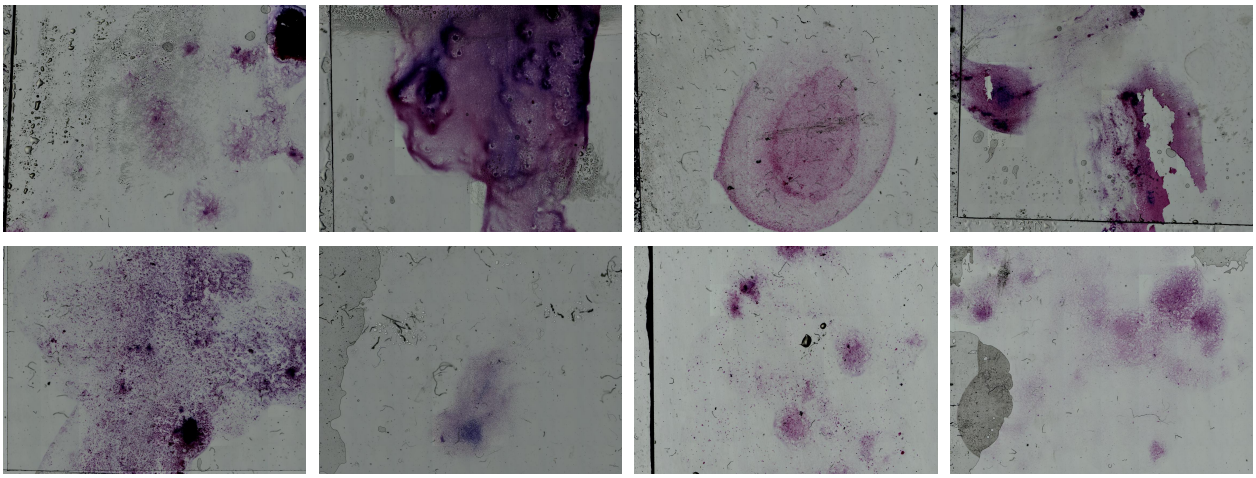
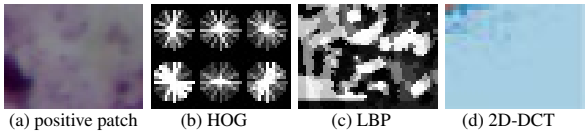


Figure 2: Novel dataset containing 8 Gram-stain slide images. Images are captured using a 2.5x objective (low-magnification).



(a) positive patch (b) HOG (c) LBP (d) 2D-DCT

Figure 3: (a) A positive patch sample and the extracted features: (b) HOG, (c) LBP, and (d) 2D-DCT. Best viewed in colour.

ing at the first 232 coefficients in a 17×19 grid, we can see how the energy is concentrated in the top left-most region.

HNM is used to deal with our imbalanced dataset. The initial training set is constructed using all 172 positive samples and selecting a total of $172 \times 3 = 516$ negative samples. We adopt SVM as the classifier and use the Library for Large Linear Classification (LibLinear) [9]. We train with the default type of solver and the default parameter c (cost). For comparison purposes, we also show the results with a SVM classifier without HNM (no-HNM). The ROC curves for no-HNM and HNM are shown in Figure 4 and Figure 5, respectively. For a better visualisation, we selected only the top 2 results for each feature type. 2D-DCT performance drops when the number q is increased. The best performance for 2D-DCT is given when $q = 1$, which means that the first coefficient concentrates the most part of the image energy. For both approaches, LBP produces the best trade-off performance between a low false positive rate and a high true positive rate.

HNM and no-HNM curves exhibit a similar tendency for LBP and HOG as seen in Figure 6. HNM in combination with 2D-DCT significantly improves the performance of no-HNM with 2D-DCT. This suggests that, for our task,

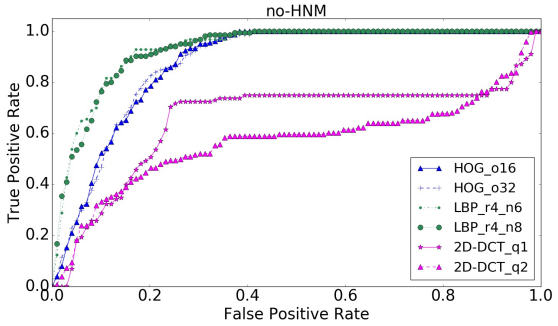


Figure 4: ROC-CURVE for no-HNM. We only show the two best results for each feature type.

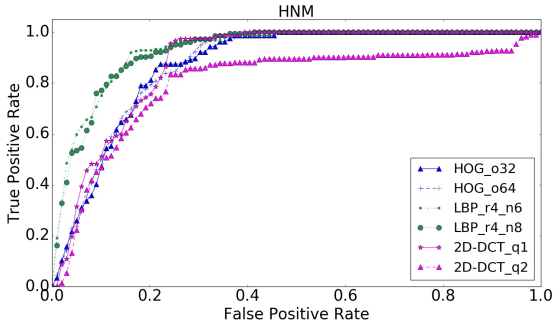


Figure 5: ROC-CURVE for HNM. As per Figure 4.

HNM is only useful under the presence of high false positives.

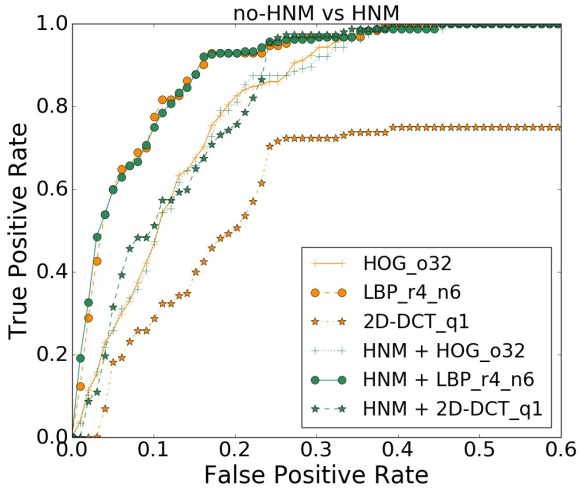


Figure 6: ROC-CURVE comparison for top results for no-HNM and HNM.

7. Main Findings and Future Work

We have presented our early experience in developing a system that automatically selects candidate areas for subsequent high-magnification analysis. We have used a low-magnification objective to find potential candidate areas, trying to replicate what pathologists do when they choose areas for high-magnification viewing. We make use of existing features for texture and biomedical images. The features analysed are HOG, LBP, and 2D-DCT. Given that we have an imbalanced set of positive and negative samples we use the HNM approach to deal with this problem. Experiments have shown that the HNM approach behaves better than its no-HNM counterpart when a high number of false positive samples are generated. This is the case with 2D-DCT features. Among all the feature types analysed, LBP features exhibit the best performance.

Although our dataset is small in terms of the number of slides, we have provided 1,600 image patches with their corresponding labels. Our next step is to increase the number of slides and label them. This work can be also extended with the construction of a ranking system approach. Our dataset provides labels that rank each patch according to its importance.

Another potential area of inquiry is to merge the features and see if we can increase the system performance. Moreover, we can also experiment with other features used for medical images such as wavelets, morphological, shape, colour, among others. Undoubtedly deep learning should be also considered for our task. To overcome the lack of a large dataset, transfer learning could be an interesting alternative to explore as in [28].

8. Acknowledgements

This work has been 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

[1] J. Angulo and G. Flandrin. Automated detection of working area of peripheral blood smears using mathematical morphology. *Analytical Cellular Pathology*, 25:37–49, 2003.

[2] E. Baron, S. Mix, and W. Moradi. Clinical utility of an automated instrument for Gram staining single slides. *Journal of Clinical Microbiology*, 48(6):2014–2015, 2010.

[3] J. W. Bartholomew and T. Mittwer. The Gram stain. *Bacteriological Reviews*, 16:1, 1952.

[4] M. Bibbo, D. Kim, T. Pfeifer, H. Dytch, H. Galera-Davidson, and P. Bartels. Histometric features for grading of prostatic carcinoma. *Analytical and Quantitative Cytology and Histology*, 13:61–69, 1991.

[5] M. L. Chayadevi and G. T. Raju. *Extraction of Bacterial Clusters from Digital Microscopic Images through Statistical and Neural Network Approaches*, pages 1091–1099. 2012.

[6] M. Crossman, A. Wiliem, P. Finucane, A. Jennings, and B. C. Lovell. A multiple covariance approach for cell detection of Gram-stained smears images. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2015.

[7] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 886–893 vol. 1, 2005.

[8] R. Dubey, J. Zhou, Y. Wang, P. M. Thompson, and J. Ye.

[9] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. Liblinear: A library for large linear classification. *Machine Learning Research*, 9:1871–187, 2008.

[10] T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861 – 874, 2006.

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

[12] S. Harisha. *Microbiology*, pages 329–349. Laxmi Publications, 2006.

[13] R. G. Hauser. *Digital image analysis of Gram stained culture specimens*. PhD thesis, School of Medicine, Yale University, 2011.

[14] P. Hiremath and P. Bannigidad. Automatic identification and classification of bacilli bacterial cell growth phases. *International Journal of Computer Applications (IJCA). Special Issue on RTIPPR*, (1):48–52, 2010.

[15] X. Hong, G. Zhao, S. Zafeiriou, M. Pantic, and M. Pietikinen. Capturing correlations of local features for image representation. *Neurocomputing*, 184:99 – 106, 2016.

[16] M. Hutchinson, P. Argawal, T. Denault, B. Berger, and E. Cibas. A new look at cervical cytology. thinprep multicenter trial results. *Acta Cytologica*, 36(4):499–504, 1992.

[17] T. Kato, R. Relator, H. Ngouv, Y. Hirohashi, O. Takaki, T. Kakimoto, and K. Okada. Segmental HOG: New descriptor for glomerulus detection in kidney microscopy image. *BMC Bioinformatics*, 16(1):316, 2015.

[18] M. M. R. Krishnan, V. Venkatraghavan, U. R. Acharya, M. Pal, R. R. Paul, L. C. Min, A. K. Ray, J. Chatterjee, and C. Chakraborty. Automated oral cancer identification using histopathological images: A hybrid feature extraction paradigm. *Micron*, 43(23):352 – 364, 2012.

[19] R. Kumar, R. Srivastava, and S. Srivastava. Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features. *Medical Engineering*, 2015, 2015.

[20] J. Lee, Y. Wu, and H. Kim. Unbalanced data classification using support vector machines with active learning on scleroderma lung disease patterns. *Journal of Applied Statistics*, 42(3):676–689, 2015.

[21] J. J. Lewis, V. N. Chihota, M. van der Meulen, P. B. Fourie, K. L. Fielding, A. D. Grant, S. E. Dorman, and G. J. Churchyard. Proof-of-Concepts: Evaluation of an automated sputum smear microscopy system for tuberculosis diagnosis. *PLoS One*, 7(11), 2012.

[22] L. M. Marler, J. A. Siders, and S. Allen. *Direct Smear Atlas: A Monograph of Gram-Stained Preparations of Clinical Specimens*. Lippincott Williams and Wilkins, 2001.

[23] S. Mishra, L. Sharma, B. Majhi, and P. K. Sa. *Microscopic Image Classification Using DCT for the Detection of Acute Lymphoblastic Leukemia (ALL)*, pages 171–180. 2017.

[24] L. Nanni, M. Paci, F. L. Caetano dos Santos, H. Skottman, K. Juuti-Uusitalo, and J. Hyttinen. Texture descriptors ensembles enable image-based classification of maturation of human stem cell-derived retinal pigmented epithelium. *PLOS ONE*, 11:1–29, 02 2016.

[25] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[26] S. Park, J. Goo, and C.-H. Jo. Receiver operating characteristic (ROC) curve: Practical review for radiologists. *Korean Journal of Radiology*, 5(1):11 – 18, 2004.

[27] W. B. Pennebaker and J. L. Mitchell. *JPEG Still Image Data Compression Standard*. Kluwer Academic Publishers, 1st edition, 1992.

[28] H. Ravishankar, P. Sudhakar, R. Venkataramani, S. Thiruvenkadam, P. Annangi, N. Babu, and V. Vaidya. *Understanding the Mechanisms of Deep Transfer Learning for Medical Images*, pages 188–196. Springer International Publishing, 2016.

[29] G. Sanroma, G. Wu, Y. Gao, and D. Shen. Learning to rank atlases for multiple-atlas segmentation. *IEEE Transactions on Medical Imaging*, 33(10):1939–1953, 2014.

[30] D. Shen, G. Wu, and H.-I. Suk. Deep learning in medical image analysis. *Annual Review of Biomedical Engineering*, 19, 2016.

[31] H. K. Singh, P. K. Maurya, K. Singh, and P. Singh. Discrete cosine transform and discrete fourier transform of RGB image. In *International Journal of Computer Applications*, 2012.

[32] P. Singleton and D. Sainsbury. *Dictionary of Microbiology and Molecular Biology*, pages 329–349. John Wiley & Sons, Ltd, 2007.

[33] L. Song, X. Liu, L. Ma, C. Zhou, X. Zhao, and Y. Zhao. Using HOG-LBP features and MMP learning to recognize imaging signs of lung lesions. In *International Symposium on Computer-Based Medical Systems (CBMS)*, pages 1–4, 2012.

[34] P. J. Tadrous. 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(6):849–858, June 2010.

[35] K. D. Toennies. *Feature Detection*, pages 173–207. Springer London, London, 2017.

[36] S. Wan, H.-C. Lee, X. Huang, T. Xu, T. Xu, X. Zeng, Z. Zhang, Y. Sheikine, J. L. Connolly, J. G. Fujimoto, and C. Zhou. Integrated local binary pattern texture features for classification of breast tissue imaged by optical coherence microscopy. *Medical Image Analysis*, 38:104 – 116, 2017.

[37] L. Xu, X. Wei, Y. Yin, W. Wang, Y. Tian, and M. Zhou. *Automatic Classification of Human Embryo Microscope Images Based on LBP Feature*, pages 145–152. 2014.

[38] M. Zhan, M. M. Crane, E. V. Entchev, A. Caballero, D. A. Fernandes de Abreu, Q. Chng, and H. Lu. Automated processing of imaging data through multi-tiered classification of biological structures illustrated using caenorhabditis elegans. *PLOS Computational Biology*, 11(4):1–21, 04 2015.