

Leukocyte subtypes classification by means of image processing

Oleg Ryabchykov^{*†}, Anuradha Ramoji^{*‡}, Thomas Bocklitz^{*†}, Martin Foerster[§], Stefan Hage^{‡¶},
 Claus Kroegel[§], Michael Bauer[‡], Ute Neugebauer^{*†‡}, Juergen Popp^{*†‡}

Emails: oleg.ryabchykov@uni-jena.com, thomas.bocklitz@uni-jena.de, juergen.popp@leibniz-ipht.de

^{*} Leibniz Institute of Photonic Technology, Albert-Einstein-Str. 9, 07745 Jena, Germany

[†] Institute of Physical Chemistry and Abbe Center of Photonics, FSU Jena, Helmholtzweg 4, 07743 Jena, Germany

[‡] Center for Sepsis Control and Care (CSCC), Jena University Hospital, Erlanger Allee 101, 07747 Jena, Germany

[§] Clinic for Internal Medicine I, Department of Pneumology and Allergy/Immunology, Jena University Hospital, Erlanger Allee 101, 07747 Jena, Germany

[¶] Center for Infectious Diseases and Infection Control, Jena University Hospital, Erlanger Allee 101, 07747 Jena, Germany

Abstract—The classification of leukocyte subtypes is a routine method to diagnose many diseases, infections, and inflammations. By applying an automated cell counting procedure, it is possible to decrease analysis time and increase the number of analyzed cells per patient, thereby making the analysis more robust. Here we propose a method, which automatically differentiate between two white blood cell subtypes, which are present in blood in the highest fractions. We apply generalized pseudo-Zernike moments to transfer morphological information of the cells to features and subsequently to a classification model. The first results indicate that information from the morphology can be used to obtain efficient automatic classification, which was demonstrated for the leukocyte subtype classification of neutrophils and lymphocytes. The approach can be extended to other imaging modalities, like different types of staining, spectroscopic techniques, dark field or phase contrast microscopy.

I. INTRODUCTION

WHITE blood cells (WBCs) are also called leukocytes. These cells protect the body from infections caused by viruses and other foreign invaders like bacteria or fungi, which make WBCs an important part of the immune system. Leukocytes are produced and derived from the bone marrow and circulate through the bloodstream. A change of the number of different WBC subtypes in the blood is utilized as marker for various diseases. Therefore a blood cell count is often utilized for a routine health examination or diagnosis of specific conditions of a patient. There are five major subtypes of WBCs [1], [2]:

- neutrophils (50-70%);
- lymphocytes (25-30%);
- monocytes (3-9%);
- eosinophils (0-5%);
- basophils (0-1%).

The ranges within the brackets display the percentage of the corresponding cell subtypes in the blood, which are typical ratios for a healthy person. There are various classification approaches, which can be roughly divided into manual and automated methods of cell classification.

The manual classification is performed by a pathologist through the subjective recognition of cell subtypes on microscopic images of stained cells. This type of analysis does not require complex equipment or highly specialized chemical reagents. To simplify the identification, cells are usually stained with the Kimura stain, which colors cell nuclei in blue. Manual differentiation between varying subtypes is accomplished based on characteristics of the cell morphology, like cell size, transparency, granularity, and the shape of the cell nucleus, which are the major differences between the subtypes. Manual classification is widely used in some specific cases of diagnosis and as a “gold standard” for scientific purposes. However, variation of cell morphology within the same cell subtype is very high, and manual classification efficiency is dependent on the pathologist’s qualification and experience.

On the other side, there are various automated classification methods, based on different physical and chemical characteristics of the cells. The main advantage of the automated devices is that they efficiently analyze large number of cells in a short time. Unfortunately, their analyzing workflows include very specific combinations of chemical and physical processes. The complexity of the analysis does not allow the design of a simple portable device. Therefore, automated blood cell counting machines are usually big and expensive.

An alternative approach is an automatic image analysis of microscopic images of stained cells. In a combination with a small camera this method can become a useful tool for doctors, providing them an instant access to the information about WBCs population at bedside of a patient. There are some studies that show efficient leukocyte identification [3], [4] and segmentation [5], [6] within microscopic images. However, these studies are focused on the leukocyte count without the classification of the leukocytes into subtypes. That leads to the loss of important information about the proportions of each cell subtype. In distinction to the mentioned studies, the current manuscript describes an algorithm for the classification

of WBCs, focusing on the textural features analysis of single cell images.

The concept of the work is to extract quantitative features related to the cell morphology from the microscopic images. Subsequently, these features are used to train and evaluate a statistical model for cell subtype identification. Moreover, the same type of images as for manual classification is used, therefore, this approach allows a direct comparison to the “gold standard”. In order to use these images for an automated image analysis, standardization and preprocessing have to be carried out. However, during the pretreatment step, it is important to eliminate corrupting effects, such as uniformities in staining and lighting, but to keep the morphological information for further analysis steps.

The textural information extraction from preprocessed images can be carried out by various methods [7], [8]. However, image description by means of pseudo-Zernike (PZ) moments [9] was chosen for the cell subtype identification because it was proven to be a reliable method for the recognition of shapes [10], characters [11], [12], faces [13], [14], [15], and viruses [16]. An advantage of the representation by PZ-moments is that their absolute values are independent from image rotation, which is necessary due to random orientation of the cells on a microscopic slide. The PZ-moments are derived from PZ-polynomials, which are orthogonal to each other and can be used in further statistical analysis, thus an automated classification technique can be established.

The proposed automated cell classification method is aimed to combine the simplicity of the manual classification and the advantages of automatization. The approach is based on the analysis of images, which are similar to the images used for manual “gold standard” method and are produced by common microscopy from a blood sample after non-complicated preparation. On the other side, due to automatization, extremely short classification times and objectivity, comparable with a human observer, can be achieved.

II. MATERIALS AND METHODS

A. Sample preparation

Leukocytes were isolated from the venous blood of patients admitted to the intensive care unit with informed consent according to the Ethics Committee of the Jena University Hospital (Ethic vote n 4004-02/14). Briefly, 2.7ml of blood in ethylenediaminetetraacetic acid (EDTA) was drawn freshly from an existing catheter using the BD monovettes. In case of healthy donor, blood (about 100 μ l) was collected from fingertip using lancet. Red blood cell lysis was carried out by mixing the blood with an ammonium chloride solution with a ratio of 1:5 in a 50ml falcon tube. After 5 minutes of incubation at room temperature (RT), the mixture was centrifuged for 10 minutes at 400g at RT. The WBC pellet at the bottom of the falcon tube was collected by discarding the supernatant and suspending it in a phosphate buffer solution (PBS). The WBCs were chemically fixed with 4% formaldehyde for 10 minutes, followed by washing the cells successively with PBS and 0.9% NaCl. The cells were coated on slides using cytospin and

stained with a Kimura staining solution (which stains only the cellular nucleus) and washed with distilled water. The slides were dried at RT and stored at 4 °C for maximum one hour until further use. The Kimura stained images of the WBCs (Fig. 1 *a,b*) were captured with an upright epifluorescence microscope (Axioplan 2, Carl Zeiss, Germany) equipped with an AxioCam HRc camera (Carl Zeiss, Germany). Images were acquired using Zeiss Axio Vert software (Carl Zeiss, Germany).

B. Calculations

All calculations reported in this work were carried out in Gnu R (version 3.0.2) [17] running on a Windows 7 Professional 64-bit system (Intel® Core™ i5-4570 CPU @ 3.20 GHz 2.70 GHz with 8GB RAM). In addition to the base R package, which contains the input/output, basic programming support, and arithmetic functions, some more specific algorithms were utilized from other packages. For orthogonal moment analysis the “IM” package [18] was used. A support vector machine (SVM) classification model was built with the “e1071” package [19]. Parallel computing was obtained by functions from “foreach” [20] and “doParallel” [21] package. K-means clustering from the “stats” package [17] was utilized for the background removal. The functions for principal component analysis (PCA), nonlinear least squares estimation, and the fast Fourier transform (FFT) are all contained in the base package [17]. JPEG files were loaded into the R environment via the “jpeg” package [22].

Prior to analysis, each image was converted from *sRGB* color space to *Lab* color space, one of the most common color spaces for image analysis applications. It was chosen due to the fact that, unlike additive or subtractive color models (for example *RGB* or *CMYK*), it is not optimized for image representation on a screen or for printing, but is adapted to cover the entire range of colors distinguishable by the human eye and to match the perception of these colors. In this color space, *a* and *b* components are related to chromatic color values. The *L* component of *Lab* color space closely matches the human perception of lightness, which allows to expect that in this representation cell subtypes can be identified based on their morphology. The conversion of the color space was performed by base R function “convertColor”.

Subsequently to the color space conversion, other steps, such as noise reduction, background removal and intensity normalization were performed. The details of these preprocessing steps are described in the “Results and discussion” section.

C. Pseudo-Zernike (PZ) Moments

As mentioned previously, PZ-moments were chosen for feature extraction from the images. These orthogonal, complex-valued moments are defined on a unit disk and are widely used for pattern recognition. The PZ-moments can describe a 2-dimensional function on the unit circle. However, the function $f(x, y)$ can represent an image if two arguments, x and y , are related to a pixel position and the function value is related to

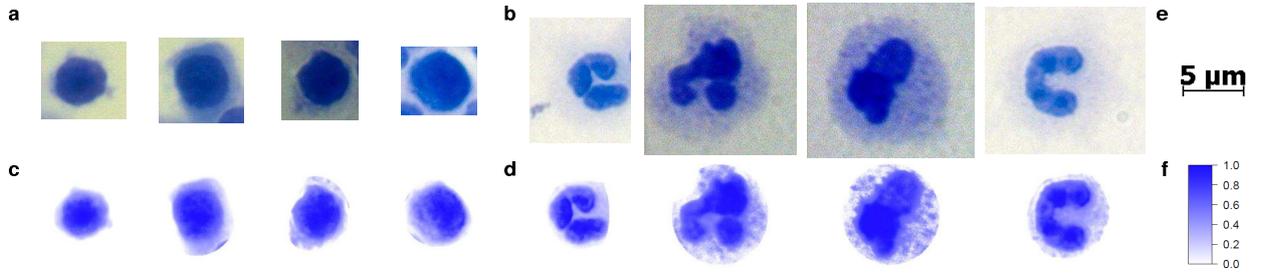


Fig. 1. Original images of two Kimura stained cell subtypes from the patients are displayed in the first row: lymphocytes (a), which are characterized by deep staining of the nuclei and a relatively small amount of cytoplasm, and neutrophils (b), which are the most common subtype that normally contain a nuclei divided into 2-5 lobes. All images are sized according to the scale (e). At the bottom preprocessed false-color equivalents of the presented images (c, d) normalized to the unit scale (f) are shown.

lightness or another color component in that pixel. The PZ-moments (A_{nl}) of an image on a unit disk are defined in radial coordinates by [23]:

$$A_{nl} = \frac{n+1}{\pi} \int_0^1 \int_0^{2\pi} [V_{nl}(r \cos \theta, r \sin \theta)]^* f(r \cos \theta, r \sin \theta) r dr d\theta.$$

In this equation $n = 0, \dots, \infty$ represents the order, the repetition is denoted by $l \leq n$, and f is the value related to the current pixel position: $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$ (polar coordinates of the pixel). V_{nl} is the orthogonal set of complex-valued PZ-polynomials, which can be written as:

$$V_{nl}(r, \theta) = R_{nl}(r) e^{jl\theta},$$

where R_{nl} represents radial polynomials with integer coefficients $D_{n,m,s}$:

$$R_{nl}(r) = \sum_{s=0}^{n-|l|} D_{n,|l|,s} r^{(n-s)},$$

$$D_{n,m,s} = \frac{(-1)^s (2n+1-s)!}{s!(n-m-s)!(n+m-s+1)!}.$$

Both the order n and the repetition l are related to the spatial frequencies of the image. However, the order n represents the spatial frequency along the unit disk's radius, while the repetition l represents the spatial frequency along the unit disk's angular coordinate. Moreover, by clarifying the idea behind order n and repetition l , the respective moments can be interpreted. Therefore, the classification model can be checked, analyzed and the morphological differences between the cell subtypes can be examined.

As it is seen from the formulas, the angular coordinate is included in the PZ-moments only within the multiplier $e^{jl\theta}$, which is related to the phase of the complex value [9], [10], [12]. Due to this fact, the absolute values of moments are independent from a rotation of the coordinate system. Thus, they are independent from the spatial alignment of the cell within the image and from the orientation on the microscopic slide. Other advantages of these particular moments are their low sensitivity to noise [10] and that the PZ-moments are orthogonal to each other.

III. RESULTS AND DISCUSSION

A. Data set

Taking into account the extremely low number of monocytes, eosinophils, and basophils in the data, only two major subtypes could be investigated in the current study. These both subtypes represent about 90% of WBCs in the blood and were included in the statistical evaluation. Thus, the training data included 28 lymphocytes and 45 neutrophils from 6 patients which were showing signs of inflammation. On the other side, the test data included 128 cells from two healthy volunteers. Unlike the training set, where some cell subtypes were sorted out, the test data included randomly selected cells without presorting or labeling according to their subtypes.

The cell subtypes included in the training data are different in sizes and cell nuclei morphology (see Fig. 1). Most notable is that the neutrophils are relatively big and have multi-lobed nuclei, while lymphocytes have almost round nuclei and are smaller. Other WBC subtypes, which were not included in the training data, are characterized by their granularity and the following properties of the cell nuclei: monocytes have kidney shaped nuclei, eosinophils have relatively small bi-lobed nuclei, and basophils have bi-lobed or tri-lobed nuclei. Although each subtype has a typical average cell size and other specific characteristics, each single cell varies from that average characteristics, which make some of its parameters dissimilar to the typical characteristics of its subtype.

B. Workflow

To obtain a stable and efficient analytical system, an image processing workflow was developed and optimized for the specific task of leukocyte subtype classification. The data was loaded, preprocessed, and represented as a set of pseudo-Zernike moments based invariants for further analysis. The workflow is presented in more detail in Fig. 2.

Important and nontrivial steps are the image preprocessing and standardization, which have to be optimized. These procedures should reduce the variations of brightness and color tones between the images of cells within the same sample and occasional appearing variations caused by the sample preparation routine for images taken from different samples. If

the workflow presented here is applied to other imaging modalities, like holographic imaging and phase contrast microscopy, these variations are expected to be less significant. Therefore, the preprocessing procedure has to be modified individually for each microscopic imaging technique and classification task.

For the construction of the classification model based on image analysis, the measured cells were labeled according to the classification made by the pathologist. The labeled and preprocessed training data were subsequently divided into three batches for cross-validation of the model. This step of the workflow was of enormous importance for setting model parameters and estimating the model quality. Thereafter, it should not be underestimated.

Leave-batch-out-cross-validation of SVM classification was performed on the training data with different combinations of input variables. This cross-validation procedure was designed to avoid any relations between different batches of cells. Therefore, the data splitting into three batches was arranged so, that the batch reflect the measurement dates and patient's origin. Thus, the generalization performance for the prediction of an independent dataset is well estimated by the leave-one-patient-out-cross-validation. Consequently, classification models with various numbers of PZ-moments' orders and principal components were compared. The variable selection was carried out according to the highest sensitivity for cross-validation of SVM classification model. The model with highest sensitivity was chosen as an optimal one and further used for the test data prediction.

Besides high identification efficiency, the proposed algorithm has to be suitable for real-life applications. Therefore, the workflow was optimized by parallelization of each single image loading, preprocessing, and calculation of the moments. Thereby, the parallelization on hardware with a multi-core processor should decrease the calculation time for a large amount of data roughly by a factor related to the number of calculation units. We chose the number of clusters for parallel calculation as one less than the number of processor cores, which was three for the PC on which the analysis was performed. During the preliminary study stage, the amount of data was relatively small, and thus, the parallelization of calculations had a negligible effect. However, despite the insignificant improvement on a small data set, parallelization is highly important for further applications and implementation of the algorithm, especially for the case if the number of analyzed cells is on the order of thousands.

C. Preprocessing

Examples of WBC images are shown in Fig. 1 *a,b*. As it can be seen by naked eye, differences between some images, which are not related to the cell's morphology, occur. These fluctuations originate from the sample preparation procedure, which is simple and standardized. There are some systematic deviations between the cells of different patients, but also the images of cells from the same patient can differ due to the spatial alignment of the cells and non-uniform coloring of samples along microscopic slides. Moreover, parts of other

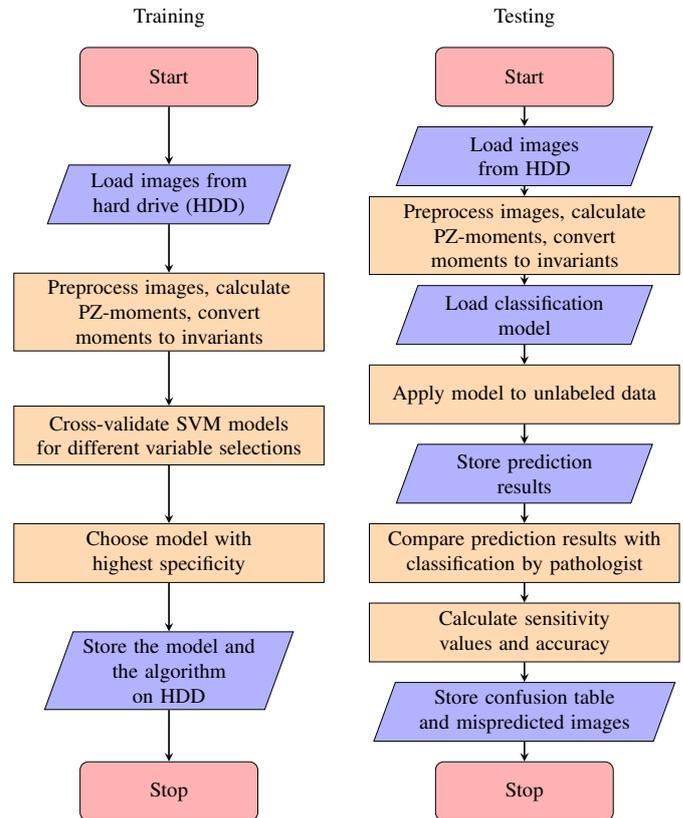


Fig. 2. Schematic workflow of the presented algorithm and the model validation.

cells are visible within some images and, additionally, other influences on the brightness, contrast, and tone are present on the microscopic images. To reduce the discussed corrupting effects, an advanced preprocessing has to be carried out before the feature extraction procedure.

According to the chosen concept of the analysis, it was important to keep the morphological features which can be distinguished visually. The automation of the preprocessing procedure took an important part in the development of the algorithm. The original images were stored in the standard *sRGB* representation, which is designed to display images in electronic systems, such as a computer's screens. However, analysis of the color channels separately from each other can be problematic and leads to a high complexity of the classification model. Switching to a single component can be circumvented by applying a more convenient color space. As it was mentioned in "Materials and methods", the lightness *L* of *Lab* color space is closely related to the human visual perception of images. In order to keep the features used for manual classification, the *Lab* color space was used in the further analysis. Moreover, the cells used for analysis were colored by Kimura staining, which highlights the cell nucleus in blue. Due to monochromatic coloring, all variations of the chromatic values are only related to the deviations of the sample preparation process and staining. Thus, related color components (*a* and *b*) were skipped and only the lightness

L was analyzed. However, for staining procedures which stain different cell organelles or cytoplasm in different colors, normalized a and b components should be also included in the analysis.

Due to high variations between different images, even for the single L -component, the automation of preprocessing took an important part in the analysis development. Pretreatment was aimed to decrease deviations of the features extracted from images within the same cell subtype and to increase the overall identification accuracy. Consequently, the background, or non-cell area of the images, was cut off via the unsupervised k -means clustering of lightness values within each image. In order to improve the background removal, an FFT-filter was applied to the images prior to the clustering. After the background removal, the lightness distribution within each cell was standardized by means of normalization to the unit interval and equalization of the histogram.

Subsequently to the lightness standardization of the images, a 2-dimensional Gaussian function was fitted to each cell image using nonlinear least squares. Based on the coefficients of the fitted function, centers and estimated radii were determined for each cell. As the next step, background-free images of the single cells were cropped according to the estimated cells' radii. This procedure was performed, to preserve the full region of the stained nucleus with a cytoplasm area and to exclude regions of other cells, non-cell area, or unexpected artifacts which were present in some images outside of the cell area. After cropping, images were placed on frames with a determined preset size, which was chosen to fit the biggest cell expected among the analyzed cell subtypes: $13 \times 13 \mu\text{m}$, which was equivalent to 200×200 pixels. On this step the centers of the cells were also matched to the centers of the frames. Pretreated images are shown in Fig. 1 *c,d*.

D. Features extraction

As quantitative features which can be used to describe the morphology of cell images, the complex-valued pseudo-Zernike moments were chosen. However, the position of each individual cell on a slide is random and it is necessary to operate with rotationally independent features. Since the phase of the moment is related to the angular coordinate within the image plane, complex-valued moments were converted to absolute PZ-moments and then normalized to the zero-order moment. Therefore, invariants, which are not dependent on the image rotation and scale, were produced. These invariants skip all information about the phase (angular coordinate), and thus, the obtained variables are independent of the image rotation.

Unfortunately, as it is shown above in the "Materials and methods" section, the calculation of PZ-moments requires a double integration of a two-dimensional function which is a costly CPU process. Because the pre-computed images were transferred to a frame with a preset size, the algorithm for the PZ- moment calculation can be simplified. Instead of the integration, the sum of a scalar product of the image with a pre-computed complex matrix can be used. The matrices

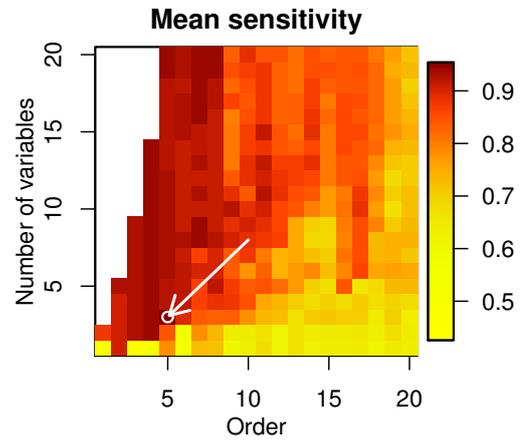


Fig. 3. Mean sensitivity of SVM leave-batch-out-cross-validation of training data. Classification models were created for a different number of selected orders of moments and for a different number of principal components (called variables in the image). The maximum value, which is related to the optimal model, is indicated with a white arrow.

TABLE I
CONFUSION TABLE FOR THE LEAVE-BATCH-OUT-CROSS-VALIDATION OF THE SVM MODEL WITH OPTIMAL VARIABLE SELECTION.

| | | Predicted | | Sensitivity |
|------|-------------|-------------|-------------|-------------|
| | | Lymphocytes | Neutrophils | |
| True | Lymphocytes | 26 | 2 | 0.893 |
| | Neutrophils | 1 | 44 | 0.978 |

related to each moment can be generated once and then stored on a hard disk drive for the further use.

E. Statistical model establishment and evaluation

To avoid an overfitting of the statistical model, the dimensionality of the data was reduced. A dimension reduction was obtained via a principal component analysis (PCA). The dimensionality of the retaining data set was optimized based on a leave-batch-out-cross-validation of the training data set. The parameter intervals checked for the feature extraction was 1 to 20 for orders, while repetition was chosen maximal. The score dimension of the PCA was evaluated from 1 to 20. For each parameter set the model performance was estimated based on the mean sensitivity. These values are summarized in plot Fig 3. The maximal sensitivity is marked on the plot with an arrow. This parameter set defines the optimal combination of input variables (3 principal components, based on PZ-moments up to 5th order). The model trained with these parameters was further analyzed and visualized. In table I a confusion table of training data cross-validation is given. In Fig. 4 a histogram of its probability scores, which represents SVM decision values rescaled to the unit range, is plotted.

F. Blind prediction

Model validation was performed by applying the established model to the independent data, which contained 163 microscopic images of stained WBCs. All preprocessing and feature extracting steps were performed on these unlabeled

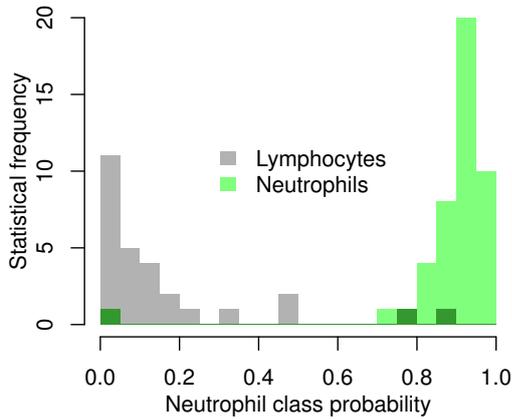


Fig. 4. Histogram for SVM posterior probabilities calculated by a leave-batch-out-cross-validation of the training data with the optimal number of variables are shown. Classification was performed between lymphocytes (gray bars) and neutrophils (green bars). The overlap of the groups is indicated within the histogram by dark green bins.

images in the same way as for the training data. In order to avoid the influence of the operator's subjectivity, a double blind prediction was carried out. Images were classified in manual mode by an experienced pathologist independently from the automated prediction. Subsequently, the statistical predictions were compared with the manual classification results. A summary of the results is visualized by a confusion table (see table II). Another representation of the classification performance is shown by means of a ROC curve in Fig 5. This curve, built for the threshold of the SVM decision values of the test data prediction, illustrates the high performance of the prediction. Moreover, the area under ROC curve (AUC) is about 0.984, which indicates an almost perfect classification. A perfect binary classification is characterized by an AUC equal to 1. Among 155 cells, which were classified as lymphocytes or neutrophils in manual mode, three images were wrongly identified by the statistical model. Such a low misclassification rate of independent test data corresponds to a high accuracy of the 2-class prediction. This accuracy was higher as 97%. Additionally, cells of the subtypes, which were not included in the training set, were present in the test data. These cells (five eosinophils and two monocytes), were predicted within the same class as neutrophils. This behavior was expected, since they feature a similar morphology as neutrophils compared to lymphocytes. Additionally, neutrophils, eosinophils, and monocytes feature a higher biological similarity and higher subjective similarity of the images. These classification results of the eosinophils and monocytes indicate that an extension of the presented model may be possible. A hierarchic layout of the classification seems optimal to incorporate eosinophils and monocytes.

IV. CONCLUSION

In this work, we presented an algorithm for a highly efficient classification between two dominant subtypes of leukocytes. The special feature of the proposed method is that by means

TABLE II
CONFUSION TABLE FOR THE PREDICTION OF THE UNLABELED TESTING DATA. CORRECT PREDICTED CELLS ARE SPECIFIED ONLY WITH THE QUANTITY OF THE IDENTIFIED CELLS. ALL INCORRECTLY PREDICTED CELLS AND CELLS, THAT RELATE TO OTHER SUBTYPES, WHICH WERE NOT INCLUDED IN THE TRAINING DATA, ARE SHOWN IN THE TABLE AS UNTREATED MICROSCOPIC (UPPER ROWS) AND PREPROCESSED (BOTTOM ROWS) IMAGES.

| | | Predicted (assorted by statistical model) | |
|--------------------------------|-------------|---|-------------|
| | | Lymphocytes | Neutrophils |
| True (assorted by pathologist) | Lymphocytes | 17 | |
| | Neutrophils | | 101 |
| | Monocytes | 0 | |
| | Eosinophils | 0 | |

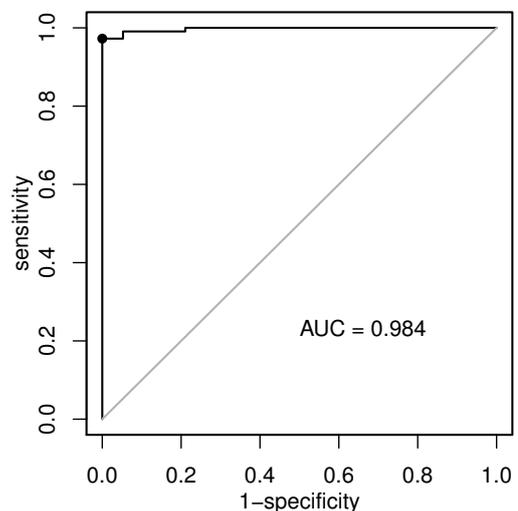


Fig. 5. The ROC curve and area under the curve (AUC) illustrate the high performance of the SVM prediction of the binary classification model between two WBC subtypes (lymphocytes and neutrophils) for independent unlabeled testing data.

of PZ-invariants the cell morphology is represented as a quantitative marker for the cell subtypes. Therefore, a combination of such common statistical methods as principal component analysis and support vector machine classification was applied to build the classification model. This approach showed a high stability against patient to patient and sample to sample variations. Moreover, an advanced image preprocessing made a further contribution to the robustness of the model. The standardization of the images decreased deviations, which occur between samples due to the sample preparation routine. Additionally, the automated framing and centering of the analyzed images of cells led to the replacement of the double numerical integration, performed for PZ-moment calculation, with a matrix product. This simplification of the calculation procedure resulted in the reduction of computation time and allowed the analysis to be performed in real-time. The classification results showed that WBCs subtypes as monocytes and eosinophils (which were not included in the model due to their low quantity in the training data) were predicted within the same class. Due to this fact, it can be assumed that the classification can be improved and extended to other cell types by a multilevel model. However, that requires a statistically significant amount of microscopic images for each leukocyte subtype in the training data set. The described approach can be applied for microscopy images taken of other staining types. Only important is that the images display the cell morphology. The method presented here may be also applied to images obtained with techniques such as fluorescence, dark field, or phase contrast microscopy.

ACKNOWLEDGMENT

Financial support of the BMBF via the integrated research and treatment center “Center for Sepsis Control and Care” (FKZ 01EO1502) and from the EU via the project “Hemo-Spec” (FP 7, CN 611682) is highly acknowledged. The authors thank Katrin Ludewig and Frank Brunkhorst for data collection and their support.

REFERENCES

- [1] A. Kratz, M. Ferraro, P. M. Sluss, and K. B. Lewandrowski, “Normal reference laboratory values,” *New England Journal of Medicine*, vol. 351, no. 15, pp. 1548–1563, 2004. doi: 10.1056/NEJMcpc049016 PMID: 15470219. [Online]. Available: <http://dx.doi.org/10.1056/NEJMcpc049016>
- [2] A. Ramoji, U. Neugebauer, T. Bocklitz, M. Foerster, M. Kiehnopf, M. Bauer, and J. Popp, “Toward a spectroscopic hemogram: Raman spectroscopic differentiation of the two most abundant leukocytes from peripheral blood,” *Analytical Chemistry*, vol. 84, no. 12, pp. 5335–5342, 2012. doi: 10.1021/ac3007363 PMID: 22721427. [Online]. Available: <http://dx.doi.org/10.1021/ac3007363>
- [3] S. Khan, A. Khan, F. S. Khattak, and A. Naseem, “An accurate and cost effective approach to blood cell count,” *International Journal of Computer Applications*, vol. 50, no. 1, 2012. doi: 10.5120/7734-0682. [Online]. Available: <http://dx.doi.org/10.5120/7734-0682>
- [4] L. Putzu and C. Di Ruberto, “White blood cells identification and counting from microscopic blood image,” *International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering*, vol. 7, no. 1, pp. 20 – 27, 2013. [Online]. Available: <http://waset.org/Publications?p=73>
- [5] M. M. G. Bhamare and D. Patil, “Automatic blood cell analysis by using digital image processing: A preliminary study,” in *International Journal of Engineering Research and Technology*, vol. 2, no. 9, ESRSA Publications, 2013. [Online]. Available: <http://www.ijert.org/view-pdf/5460/>
- [6] F. Sadeghian, Z. Seman, A. R. Ramli, B. A. Kahar, and M.-I. Saripan, “A framework for white blood cell segmentation in microscopic blood images using digital image processing,” *Biological procedures online*, vol. 11, no. 1, pp. 196–206, 2009. doi: 10.1007/s12575-009-9011-2. [Online]. Available: <http://dx.doi.org/10.1007/s12575-009-9011-2>
- [7] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *Systems, Man and Cybernetics, IEEE Transactions on*, vol. SMC-3, no. 6, pp. 610–621, 1973. doi: 10.1109/TSMC.1973.4309314. [Online]. Available: <http://dx.doi.org/10.1109/TSMC.1973.4309314>
- [8] M. Habibzadeh, A. Krzyzak, and T. Fevens, *Comparative study of feature selection for white blood cell differential counts in low resolution images*, ser. Lecture Notes in Computer Science. Springer International Publishing, 2014, vol. 8774, book section 20, pp. 216–227. ISBN 978-3-319-11655-6. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-11655-6_20
- [9] T. Xia, H. Zhu, H. Shu, P. Haigron, and L. Luo, “Image description with generalized pseudo-Zernike moments,” *Journal of the Optical Society of America A*, vol. 24, no. 1, pp. 50–59, 2007. doi: 10.1364/JOSAA.24.000050. [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-24-1-50>
- [10] S. O. Belkasim, M. Shridhar, and M. Ahmadi, “Pattern recognition with moment invariants: A comparative study and new results,” *Pattern Recognition*, vol. 24, no. 12, pp. 1117–1138, 1991. doi: 10.1016/0031-3203(91)90140-Z. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/003132039190140Z>
- [11] C. Kan and M. D. Srinath, “Invariant character recognition with Zernike and orthogonal Fourier-Mellin moments,” *Pattern Recognition*, vol. 35, no. 1, pp. 143–154, 2002. doi: 10.1016/S0031-3203(00)00179-5. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0031320300001795>
- [12] C. W. Chong, P. Raveendran, and R. Mukundan, “The scale invariants of pseudo-Zernike moments,” *Pattern Analysis and Applications*, vol. 6, no. 3, pp. 176–184, 2003. doi: 10.1007/s10044-002-0183-5. [Online]. Available: <http://dx.doi.org/10.1007/s10044-002-0183-5>
- [13] Y.-H. Pang, A. T. B. J, and D. N. C. L, “Enhanced pseudo Zernike moments in face recognition,” *IEICE Electronics Express*, vol. 2, no. 3, pp. 70–75, 2005. doi: 10.1587/elex.2.70. [Online]. Available: <http://dx.doi.org/10.1587/elex.2.70>
- [14] E. Walia, C. Singh, and N. Mittal, “Discriminative Zernike and pseudo Zernike moments for face recognition,” *Int. J. Comput. Vis. Image Process.*, vol. 2, no. 2, pp. 12–35, 2012. doi: 10.4018/ijcvip.2012040102. [Online]. Available: <http://dx.doi.org/10.4018/ijcvip.2012040102>
- [15] J. Haddadnia, M. Ahmadi, and K. Faez, “An efficient feature extraction method with pseudo-Zernike moment in rbf neural network-based human face recognition system,” *EURASIP Journal on Advances in Signal Processing*, vol. 2003, pp. 890–901, 2003. doi: 10.1155/s1110865703305128. [Online]. Available: <http://dx.doi.org/10.1155/s1110865703305128>
- [16] T. Bocklitz, E. Kämmer, S. Stöckel, D. Cialla-May, K. Weber, R. Zell, V. Deckert, and J. Popp, “Single virus detection by means of atomic force microscopy in combination with advanced image analysis,” *Journal of Structural Biology*, vol. 188, no. 1, pp. 30 – 38, 2014. doi: 10.1016/j.jsb.2014.08.008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1047847714001841>
- [17] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2015. [Online]. Available: <https://www.R-project.org/>
- [18] B. Rajwa, M. Dundar, A. Irvine, and T. Dang, *IM: Orthogonal Moment Analysis*, 2013, R package version 1.0. [Online]. Available: <https://CRAN.R-project.org/package=IM>
- [19] D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, and F. Leisch, *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*, 2015, R package version 1.6-7. [Online]. Available: <https://CRAN.R-project.org/package=e1071>
- [20] Revolution Analytics and S. Weston, *foreach: Provides Foreach Looping Construct for R*, 2015, R package version 1.4.3. [Online]. Available: <https://CRAN.R-project.org/package=foreach>

- [21] —, *doParallel: Foreach Parallel Adaptor for the 'parallel' Package*, 2015, R package version 1.0.10. [Online]. Available: <https://CRAN.R-project.org/package=doParallel>
- [22] S. Urbanek, *jpeg: Read and write JPEG images*, 2014, R package version 0.1-8. [Online]. Available: <https://CRAN.R-project.org/package=jpeg>
- [23] C. H. Teh and R. T. Chin, "On image analysis by the methods of moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 496–513, Jul 1988. doi: 10.1109/34.3913. [Online]. Available: <http://dx.doi.org/10.1109/34.3913>