



HAL
open science

Improving Interpretability of Leucocyte Classification with Multimodal Network

Manon Chossegros, Xavier Tannier, Daniel Stockholm

► **To cite this version:**

Manon Chossegros, Xavier Tannier, Daniel Stockholm. Improving Interpretability of Leucocyte Classification with Multimodal Network. Medical Informatics Europe (MIE 2024), Aug 2024, Athens, Greece. 10.3233/shti240602 . hal-04816434

HAL Id: hal-04816434

<https://hal.science/hal-04816434v1>

Submitted on 3 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Improving Interpretability of Leucocyte Classification with Multimodal Network

Manon CHOSSEGROS^a, Xavier TANNIER^{a,1} and Daniel STOCKHOLM^{b,c}

^a*Sorbonne Université, Inserm, Université Sorbonne Paris-Nord, Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé, LIMICS, France*

^b*Sorbonne Université, Inserm, Centre de Recherche Saint-Antoine, CRSA, Paris*

^c*PSL Research University, EPHE, Paris, France*

ORCID ID: Manon Chossegros <https://orcid.org/0000-0003-1994-8806>,

Xavier Tannier <https://orcid.org/0000-0002-2452-8868>,

Daniel Stockholm <https://orcid.org/0000-0002-5069-5256>

Abstract. White blood cell classification plays a key role in the diagnosis of hematologic diseases. Models can perform classification either from images or based on morphological features. Image-based classification generally yields higher performance, but feature-based classification is more interpretable for clinicians. In this study, we employed a Multimodal neural network to classify white blood cells, utilizing a combination of images and morphological features. We compared this approach with image-only and feature-only training. While the highest performance was achieved with image-only training, the Multimodal model provided enhanced interpretability by the computation of SHAP values, and revealed crucial morphological features for biological characterization of the cells.

Keywords. Multimodal Classification, Deep Learning, Machine Learning, White Blood Cells

1. Introduction

The classification of white blood cells (WBC) is an essential step in diagnosing scans derived from bone marrow or peripheral blood. On one hand, Deep Learning methods excel in image classification and eliminate the necessity for hand-crafted features [1], however they lack of explainability and require a substantial number of samples for effective training [2,3]. On the other hand, morphological characteristics hold importance in the final diagnosis, and therefore bring clinical value [4,5]. In this article, we proposed a Multimodal model that seeks to make the most of both approaches. This model combines very high performance with the ability to explain results, which is essential if these tools are to be accepted by experts. We compared this approach with SVM and CNN models which provided predictions based solely on features and images, respectively.

¹ Corresponding Author: Xavier Tannier; E-mail: xavier.tannier@sorbonne-universite.fr.

2. Material and Methods

This study utilized images from three different datasets: Tianjin [6], Rabin [7], and Barcelona [8]. The Barcelona dataset consisted of 17,092 images of individual normal blood cells. The Rabin-WBC dataset included 40,000 white blood cell images with 1,145 segmentation ground-truths for nuclei and cytoplasm segmentation. The Tianjin dataset comprised 6,273 images. Additionally, for this dataset, the bounding boxes of each white blood cell were provided with the images. In total, the datasets contained 1,821 basophils, 5,281 eosinophils, 6,024 lymphocytes, 14,191 neutrophils and 978 monocytes. Each dataset was split into 80/20 for train and test respectively.

2.1. Segmentation

In order to extract morphological features and characterize every part of the cell, we segmented the cytoplasm and the nucleus in every image. For the segmentation of cells in the Rabin dataset, we utilized Unet architecture [9], which was trained on ground truths provided with the dataset. The model was trained to label each pixel of the image as background, cytoplasm, or nucleus. The same architecture was employed for the Barcelona dataset. As segmentation ground truths were not provided with this dataset, they were manually annotated by our experts. Regarding the Tianjin dataset, we leveraged the bounding boxes to extract segmentation ground truths; we used SegmentAnything Model [10] to segment cell cytoplasm, and then applied the km-m nucleus segmentation method as described in [11].

2.2. Feature Extraction

To describe the cell, morphological features such as shape, color and texture were extracted. Each feature was computed for both the cytoplasm and the nucleus, as well as for the entire cell. In total, we obtained 243 features describing every part of the cell in terms of shape, texture, and color. The resulting table was normalized by column for each dataset.

Shape Features: The cell's shape was characterized using area, perimeter, convex hull area, eccentricity, aspect ratio, roundness, circularity, solidity.

Color Features: Color features were computed from images in RGB, HSV, and CMYK spaces. Each color space was characterized by the mean, standard deviation, skewness, kurtosis, uniformity and entropy of the pixel distribution per channel. The average of these values over all channels was also calculated.

Texture Features: Texture features were computed from Gray-Level Co-Occurrence Matrix (GLCM) and Local Binary Patterns (LBP). For GLCM, we computed homogeneity, contrast, angular second moment, and energy. For LBP, we calculated energy, contrast, entropy, mean, and variance.

A Multimodal neural network was developed to classify cells using both images and morphological features. The image modality was input to a convolutional neural network based on the EfficientNet-B0 architecture. For the tabular modality, two fully connected layers were employed. The two parts of the model were eventually fused by concatenating the output layer of each branch, and the classification was performed on the concatenated vector (see **Figure 1**). This model was compared with two others: SVM for classifying cells based the tabular data only, and EfficientNet-B0 for classifying cells

based on images only. We conducted all experiments on a cluster SLURM node with 12 CPUs, 132 GB RAM, and two NVIDIA RTX 3080 GPUs with 10 GB each.

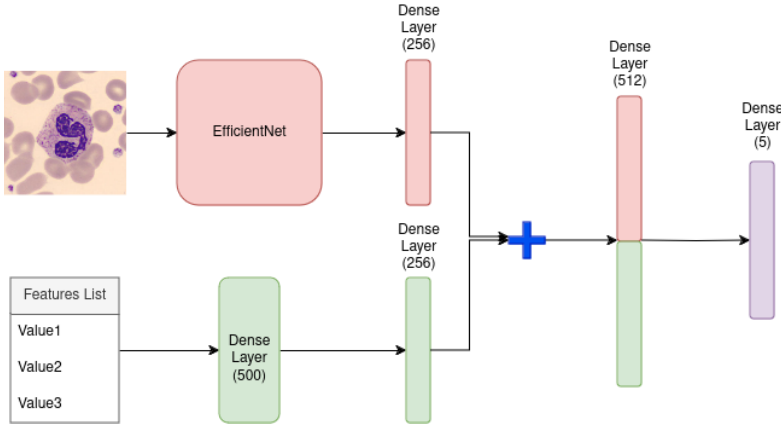


Figure 1. Architecture of the Multimodal network. The network comprises two branches, one for each modality. They are fused through concatenation before the final dense layer, which gives the class prediction. Computations were performed using Pytorch Library.

To understand the importance of each morphological feature in the Multimodal neural network, SHAP values were computed as described in [12]; by looking at the modification of the output by the model with all possible combination of features, and determining the contribution of the feature when it is added to the combination. Hence, they indicate whether the feature contributes to an increase or decrease in the output probability for each class.

3. Results and Discussion

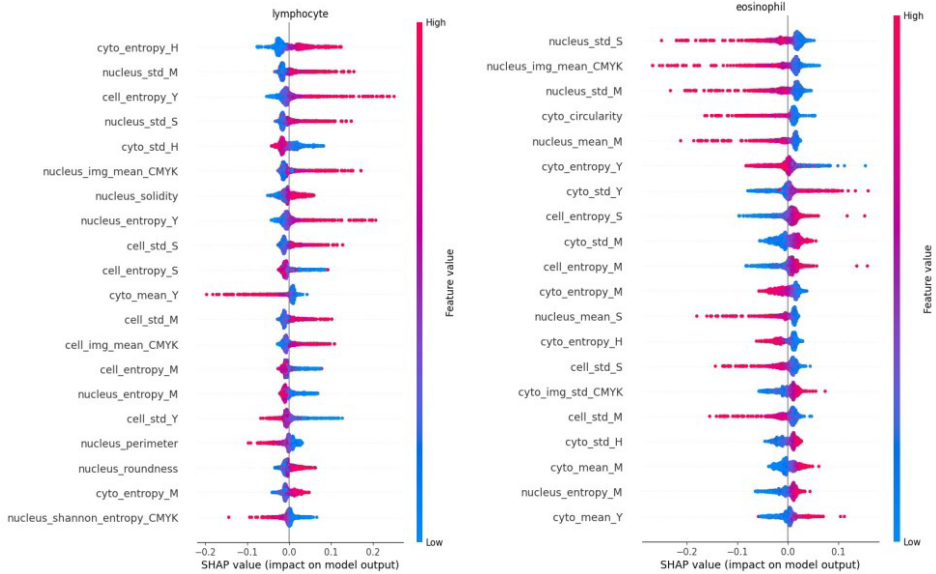
Table 1. Classification results for every different model (Multimodal Network, SVM and EfficientNet-B0) on each dataset. Accuracy is defined as the overall accuracy; it is the proportion of all rightly classified images.

Source	Training	Acc	Recall					Precision				
			Bas	Eos	Lym	Mon	Neu	Bas	Eos	Lym	Mon	Neu
Rabin	Multimodal	98.49	100	98.98	98.63	84.51	99.66	100	92.38	98.29	95.24	99.66
	Features Only	98.16	100	95.41	97.77	92.25	99.19	98.11	94.44	98.27	91.61	99.25
	Images Only	99.39	100	99.0	98.52	96.51	99.93	100	100	99.30	95.16	99.77
Barcelona	Multimodal	98.41	98.64	99.49	99.53	99.13	97.37	97.75	99.83	94.64	93.47	100
	Features Only	98.45	97.73	99.32	97.65	98.7	97.86	99.08	98.65	98.11	95.8	99.17
	Images Only	99.81	100	100	99.5	100	99.75	99.54	99.82	99.54	99.61	100
Tianjin	Multimodal	98.18	95.52	100	95.45	97.56	99.31	100	99.1	98.75	90.91	99.72
	Features Only	97.86	95.52	99.54	97.88	94.63	98.47	91.43	96.89	97.88	95.57	99.44
	Images Only	98.76	95.58	99.59	98.59	98.59	96.61	100	99.18	97.95	95.74	99.71

For each dataset, the reported accuracy in **Table 1** aligns with the state-of-the-art standards [6,8,7]. Notably, the highest accuracy scores are achieved when the model is trained with images only, while the lowest accuracy scores are observed for models trained with features only. However, the average accuracies being very close, it is challenging to conclude an improvement from feature-only training to image-only training. Reasonably, we can infer that the main advantage of utilizing features lies in the enhanced interpretability and clinical relevance of the morphological cell description.

3.1. SHAP Values

We investigated the importance of each morphological feature in the Multimodal network by computing SHAP values. In **Figure 2**, we showed the main features involved in lymphocyte and eosinophil classification in the Tianjin dataset. Each point represents a cell. Positive SHAP values reveal that the feature increases the likelihood of the class prediction, and reciprocally for negative SHAP value. The color of the dot indicates whether it is a high or a low value of the feature that impacts the probability of the class.



a) SHAP values for lymphocyte predictions: high roundness and low perimeter of the nucleus lead to higher chance of the cell being a lymphocyte.

b) SHAP values for eosinophil predictions. Magenta and Yellow components of the image considerably impact the probability of the cell being an eosinophil.

Figure 2. SHAP values obtained from the Multimodal neural network trained on the Tianjin dataset. For different color spaces (RGB, HSV and CMYK) we represented metrics by channel, `_R` for red, `_G` for green, `_M` for Magenta, `_Y` for Yellow etc.

Features such as color standard deviation or entropy of colors have a substantial impact on the final prediction. These features are linked with the diversity and the heterogeneity of pixel colors in the image, which, in turn, can be linked to the density of chromatin in the nucleus or the granulation level in the cytoplasm - crucial clinical elements for cell identification. In **Figure 2a** we observe that a high roundness or a low perimeter of the nucleus increase the likelihood of the cell being a lymphocyte, which aligns with the

characteristics of this cell. Additionally, in **Figure 2b**, magenta and yellow color features (such as *nucleus_std_M* or *cyto_entropy_Y*) were prominent in eosinophil's SHAP values, and eosinophils are characterized by their orange shade.

4. Conclusions

In this study, we compared the performances of three types of models for the classification of white blood cells across different datasets. The objective was to understand the influence of the input modality on the final prediction, whether it was morphological features or cell images. The comparison was made between an SVM, a CNN and a Multimodal neural network combining these two modalities. The results of the classification were very high for the three models. Furthermore, the inclusion of morphological features provided a valuable interpretability benefit; enabling experts to better understand the main features for prediction and link them with clinically relevant properties of the cell. The analysis of the impact of each feature was performed by computing SHAP values on the Multimodal neural network predictions.

References

- [1] Walter W, Pohlkamp C, Meggendorfer M, Nadarajah N, Kern W, Haferlach C, et al. Artificial intelligence in hematological diagnostics: Game changer or gadget? *Blood Reviews*. 2023;58:101019. doi:10.1016/j.blre.2022.101019.
- [2] Hegde RB, Prasad K, Hebbar H, Singh BMK. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. *Biocybernetics and Biomedical Engineering*. 2019;39(2):382-92. doi:10.1016/J.BBE.2019.01.005.
- [3] Castelvechi D. Can we open the black box of AI? *Nature News*. 2016;538(7623):20. doi:10.1038/538020a.
- [4] Kimura K, Tabe Y, Ai T, Takehara I, Fukuda H, Takahashi H, et al. A novel automated image analysis system using deep convolutional neural networks can assist to differentiate MDS and AA. *Scientific reports*. 2019;9(1):13385. doi:10.1038/s41598-019-49942-z.
- [5] Eckardt JN, Middeke JM, Riechert S, Schmittmann T, Sulaiman AS, Kramer M, et al. Deep learning detects acute myeloid leukemia and predicts NPM1 mutation status from bone marrow smears. *Leukemia*. 2022;36(1):111-8. doi:10.1038/s41375-021-01408-w.
- [6] Li M, et al. A deep learning model for detection of leukocytes under various interference factors. *Scientific Reports*. 2023;13(1):2160. doi:10.1038/s41598-023-29331-3.
- [7] Kouzehkanan ZM, Saghari S, Tavakoli S, Rostami P, Abaszadeh M, Mirzadeh F, et al. A large dataset of white blood cells containing cell locations and types, along with segmented nuclei and cytoplasm. *Scientific reports*. 2022;12(1):1123. doi:10.1038/s41598-021-04426-x.
- [8] Acevedo A, et al. A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. *Data in brief*. 2020;30. doi:10.1016/j.dib.2020.105474.
- [9] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18. Springer; 2015. p. 234-41.
- [10] Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, et al. Segment anything. *arXiv preprint arXiv:230402643*. 2023. doi:10.48550/arXiv.2304.02643.
- [11] Tavakoli S, Ghaffari A, Kouzehkanan ZM, Hosseini R. New segmentation and feature extraction algorithm for classification of white blood cells in peripheral smear images. *Scientific Reports*. 2021;11(1):19428. doi:10.1038/s41598-021-98599-0.
- [12] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Advances in neural information processing systems*. 2017;30. doi:10.5555/3295222.3295230.