

Classification of Leukocytes Using Meta-Learning and Color Constancy Methods

Eduardo Rivas-Posada¹, Mario I. Chacon-Murguia¹, Juan A. Ramírez-Quintana¹, Carlos Arzate-Quintana²

¹ Visual Perception Laboratory, Tecnológico Nacional de México/I.T. Chihuahua, Ave. Tecnológico #2909, Chihuahua 31310, Mexico

² Biomedical Department, Universidad Autónoma de Chihuahua

ARTICLE INFO

Article history:

Received November 16, 2022

Revised December 03, 2022

Published December 21, 2022

Keywords:

AutoML;
Color Constancy;
Learning-to-learn;
Leukocytes classification;
Meta-learning;
White Blood Cells

ABSTRACT

In the human healthcare area, leukocytes are very important blood cells for the diagnosis of different pathologies, like leukemia. Recent technology and image-processing methods have contributed to the image classification of leukocytes. Especially, machine learning paradigms have been used for the classification of leukocyte images. However, reported models do not leverage the knowledge produced by the classification of leukocytes to solve similar tasks. For example, the knowledge can be reused to classify images collected with different types of microscopes and image-processing techniques. Therefore, we propose a meta-learning methodology for the classification of leukocyte images using different color constancy methods involving previous knowledge. Our methodology is trained with a specific task at the meta-level, and the knowledge produced is used to solve a different task at the base-level. For the meta-level, we implemented meta-models based on Xception, and for the base-level, we used support vector machine classifiers. Besides, we analyzed the Shades of Gray color constancy method commonly used in skin lesion diagnosis and now implemented for leukocyte images. Our methodology, at the meta-level, achieved 89.28% for precision, 95.65% for sensitivity, 91.78% for F1-score, and 94.40% for accuracy. These scores are competitive regarding the reported state-of-the-art models, especially the sensitivity which is very important for imbalanced datasets, and our meta-model outperforms previous works by +2.25%. Additionally, for the basophil images that were acquired from a chronic myeloid leukemia-positive sample, our meta-model obtained 100% for sensitivity. Moreover, we present an algorithm that generates a new conditioned output at the base-level obtaining highly competitive scores of 91.56% for sensitivity and F1 scores, 95.61% for precision, and 96.47% for accuracy. The findings indicate that our proposed meta-learning methodology can be applied to other medical image classification tasks and achieve high performances by reusing knowledge and reducing the training time for new similar tasks.

This work is licensed under a [Creative Commons Attribution-Share Alike 4.0](https://creativecommons.org/licenses/by-sa/4.0/)



Corresponding Author:

Mario I. Chacon-Murguia, Visual Perception Laboratory, Tecnológico Nacional de México/I.T. Chihuahua, Ave. Tecnológico #2909, Chihuahua 31310, Mexico.

Email: mario.cm@chihuahua.tecnm.mx

1. INTRODUCTION

The leukocytes or white blood cells like basophils, eosinophils, lymphocytes, monocytes, and neutrophils, are very important cells for human health. In the healthcare area, leukocytes have been used for the diagnosis of leukemia [1][2], in the evaluation for COVID-19 severity [3], among other pathologies [4]. Different methods [5][6] implemented to acquire medical images through a microscope, cameras, and zooms allow the

use of the machine and deep learning paradigms for the detection, segmentation, and classification of leukocytes.

Next, we present past research efforts for the detection and classification of leukocytes, or white blood cells, using traditional machine and deep learning paradigms. In the traditional machine learning area, most of the models are implemented using the Support Vector Machine (SVM) classifier in combination with other methods like; k-means [7]; k-means with morphological operations to the leukocyte's images [8]; Particle Swarm Optimization (PSO) [9]; Optimized Binary Bat algorithm (OBBA) and logistic regression [10]. In the deep learning area, most of the models are based on Convolutional Neural Networks [11], [12], [13], and a combination of CNN with other methods like; traditional feature extraction methods [14]; and SVM [15]. Besides, pretrained models have been implemented using the transfer learning paradigm, e.g., AlexNet with GoogleNet, and SVM [16]; or data augmentation techniques with Generative Adversarial Networks (GAN) with ResNet [17].

Other interesting and current methods to detect and classify leukocytes or white blood cell image based on machine and deep learning models are presented next. In [18] different machine and deep learning models like ResNet18 and DeepFlow are trained and evaluated on stain-free white blood cell images. The use of stain-free images allows a better generalization of the models because the models do not require a specific staining method for the cell images. However, in [18], it was concluded that the evaluated models do not exceed the performances of the methods based on handcraft engineering features. Another work, [19], proposes the leukocyte image classification using spatial and spectral features with an SVM. The spectral features can be used to reuse knowledge to solve similar new tasks, nonetheless, the proposed model requires a microscopic hyperspectral imaging technology. The last example is the work presented in [20]. In [20], it is proposed a model for the white blood cell leukemia image classification using a VGGNet pre-trained model as a feature extractor and a Statistically Enhanced Salp Swarm Algorithm (SESSA). The SESSA algorithm is used to remove correlated and noisy features.

However, one common situation of the previously mentioned works is that the knowledge generated by the classification of the leukocytes is not analyzed and/or used to solve similar tasks. For example, suppose that a model is trained in a particular leukocyte dataset collected with a specific type of microscope and image-processing techniques. Then, the training of the model with that dataset produces new knowledge that may be used to solve other leukocyte datasets that were acquired with different microscopes or were from different distributions. Therefore, in this work, we propose a novel methodology based on the Meta-Learning (MeL) paradigm. The MeL paradigm is part of the recent area of Automated Machine Learning (AutoML) [21] and it is inspired by cognitive and learning human processes [22], [23] of how humans can learn with little data by reusing knowledge previously learned. In the MeL paradigm, two learning processes are considered, one at the meta-level and the other at the base-level. The meta-level is considered as a level where the meta-knowledge will be produced, and this meta-knowledge can be used at the base-level to solve new similar tasks. In such a way that each level of learning has its own model, meta-model and base-model, but the base-model is leveraged by the knowledge produced by the meta-model.

The contributions of our paper are enlisted next.

- A novel methodology based on the meta-learning paradigm with Xception prior-model at the meta-level and with SVM at the base level. The base-model is leveraged by the meta-knowledge produced by classifying leukocytes from a meta-task, to solve a new base-task.
- To approximate the leukocyte images of the different tasks to a same distribution we employed the Shades of Gray (SoG) color constancy method [24][25]. SoG has been used in the medical area of skin lesion diagnosis [26][27]. However, as far as we know, this method has been not used in methodologies for leukocyte's classification. Therefore, we extended the analysis of SoG for the classification of leukocytes.
- Besides, we propose a new SoG² method based on the application of the SoG method twice with different parameters to a current image. The proposed SoG² can be used as an additional tool in the image segmentation process.
- Finally, we present an algorithm that generates a new conditioned output for the SVM used in the base-level that improves the classification of lymphocytes, monocytes, and neutrophils of the base-task.

The remainder of this paper is organized as follows. In Section 2 we formalized the MeL paradigm using 2 strategies, probabilistic and mechanistic. Then, Section 3 details the proposed MeL methodology, the datasets, and the color constancy methods employed to evaluate our work. Section 4 shows the results achieved by the proposed meta- and base-models and comparison with state-of-the-art works. Finally, in Section 5 we present the conclusions.

2. FORMALIZATION OF META-LEARNING

The formalization of the MeL paradigm is restricted to two strategies of MeL: probabilistic or mechanistic [28]. The probabilistic strategy helps to formalize the MeL paradigm by considering probability distributions. The probability distributions include previously known information. The mechanistic strategy formalizes the MeL paradigm considering the implementation of algorithms at the meta-and base-level that allow systematically to apply MeL.

2.1. Probabilistic Strategy of Meta-Learning

To explain the probabilistic strategy of meta-learning, let's consider first the supervised learning in traditional machine learning models. According to [28], in supervised learning *a posteriori* probability $p(\phi_{base}|D_{base})$ is used to optimize the parameters $(\phi_{base} \in \Phi_{base}) \subset \Theta_{base}$ of a given base-model. The $D_{base}^{tr} = \{(t_1, \lambda_1), \dots, (t_K, \lambda_K)\}$ is the training base-set with K tuples of t_K inputs and λ_K labels. Thus, the optimized parameters of the base-model can be computed by

$$\phi_{base}^* = \arg \max_{\phi_{base}} \log p(\phi_{base}|D_{base}^{tr}) \quad (1)$$

By using the Bayes' theorem, (1) can be expressed as follows

$$\arg \max_{\phi_{base}} \log p(\phi_{base}|D_{base}^{tr}) = \arg \max_{\phi_{base}} \log L(D_{base}^{tr}|\phi_{base}) + \log p(\phi_{base}) \quad (2)$$

where $L(D_{base}^{tr}|\phi_{base})$ is the *likelihood* of D_{base}^{tr} given the model parameters ϕ_{base} . The optimal parameters ϕ_{base}^* are found by maximizing $L(D_{base}^{tr}|\phi_{base})$. The $\log L(D_{base}^{tr}|\phi_{base})$ function is applied to simplify the computation of the derivate to obtain the optimized parameters because the log-likelihood is easier to optimize. The *prior* probability distribution $p(\phi_{base})$ can be seen as a regularizer, that is, a regularizer of the weight parameter. Hence, (2) can be expressed using the tuples of the input data D_{base}^{tr}

$$\arg \max_{\phi_{base}} \log p(\phi_{base}|D_{base}^{tr}) = \arg \max_{\phi_{base}} \sum_k \log L(y_k|x_k, \phi_{base}) + \log p(\phi_{base}) \quad (3)$$

Most of the traditional supervised models employ a large amount of labeled D_{base}^{tr} . Nevertheless, this is not the case in the natural process of learning in humans, who use previous knowledge and little data to learn something new. In consequence, as it can be noticed in (3) the performance of the traditional supervised models depends on the amount of data of a particular task. In order to improve learning in the base-models, additional data can be incorporated in (1), by using meta-learning

$$\phi_{base}^* = \arg \max_{\phi_{base}} \log p(\phi_{base}|D_{base}^{tr}, D_{meta}^{tr}) \quad (4)$$

Each meta-set $D_{meta}^{(j)} \in D_{meta}^{tr}$ includes the tuples of input-label pairs, $D_{meta}^{(j)} = \{(t_1^{(j)}, \lambda_1^{(j)}), \dots, (t_K^{(j)}, \lambda_K^{(j)})\}$, that are used to solve a specific task. Assuming statistic independence $\phi_{base} \perp\!\!\!\perp \phi_{meta}|D_{meta}^{tr}$, where $(\phi_{meta} \in \Phi_{meta}) \subset \Theta_{meta}$ are the parameters of the meta-model, then

$$\begin{aligned} \log p(\phi_{base}|D_{base}^{tr}, D_{meta}^{tr}) &= \log \int_{\Phi_{meta}} p(\phi_{base}|D_{base}^{tr}, \phi_{meta})p(\phi_{meta}|D_{meta}^{tr}) d\phi_{meta} \\ &\approx \log p(\phi_{base}|D_{base}^{tr}, \phi_{meta}^*) + \log p(\phi_{meta}^*|D_{meta}^{tr}) \end{aligned} \quad (5)$$

Thereby,

$$\arg \max_{\phi_{base}} \log p(\phi_{base}|D_{base}^{tr}, D_{meta}^{tr}) \approx \arg \max_{\phi_{base}} \log p(\phi_{base}|D_{base}^{tr}, \phi_{meta}^*) \quad (6)$$

Therefore, the meta-learning problem under the probabilistic strategy is reduced to compute

$$\phi_{meta}^* = \arg \max_{\phi_{meta}} \log p(\phi_{meta}|D_{meta}^{tr}) \quad (7)$$

Then, (7) is used in the adaptation stage of the base-model

$$\phi_{base}^* = \arg \max_{\phi_{base}} \log p(\phi_{base}|D_{base}^{tr}, \phi_{meta}^*) \quad (8)$$

As it can be noticed, the adaptation does not depend anymore on D_{meta}^{tr} as in (4). It only depends on D_{base}^{tr} and the optimized meta-parameters ϕ_{meta}^* .

2.2. Mechanistic Strategy of Meta-Learning

To view the meta-learning problem from the mechanistic strategy, (8) is interpreted as $\phi_{base}^* = A_{base \sim \phi_{meta}^*}(D_{base}^{tr})$. Here, $A_{base \sim \phi_{meta}^*}(D_{base}^{tr})$ is the function in the meta-parameter space ϕ_{meta} that takes D_{base}^{tr} as input to obtain the optimized parameters ϕ_{base}^* in the base-domain. $A_{base \sim \phi_{meta}^*}(D_{base}^{tr})$ can be used to represent the implementation of the base-algorithm A_{base} of the base-model. The optimized meta-parameters can be obtained by $\phi_{meta}^* = A_{meta \sim \phi_{meta}^*}(D_{meta}^{tr})$, where $A_{meta \sim \phi_{meta}^*}(D_{meta}^{tr})$ is the function in the meta-parameters space ϕ_{meta} that takes D_{meta}^{tr} as input to yield the optimized parameters ϕ_{meta}^* in the meta-domain. $A_{meta \sim \phi_{meta}^*}(D_{meta}^{tr})$ can be used to represent the implementation of the meta-algorithm A_{meta} of the meta-model.

To conclude this section, we propose a general model based on MeL, Fig. 1. In Fig. 1, the diagram includes a meta-model with the meta-data that contains prior optimized and implemented models. The meta-model is in a meta-domain that can be different from the base-domain of the base-model. To establish this difference between the domains, the color of the background of the meta-model is distinct from where the base-model is trained. This means that when the MeL uses meta-data based on prior-models, a meta-model can be generated to be directly applied in a base-model and in other learning problems that can involve Out-Of-Distribution (OOD) tasks. Furthermore, it can be noticed that the diagram includes two types of lines in the contours of the boxes. The continuous lines represent well-defined elements of the model and contain known information. Instead, the dashed lines indicate the dataset, either at the base-level or at the meta-level. This is because the learning can be supervised, semi-supervised, or reinforcement, and the dataset will be labeled or unlabeled. Also, it depends on the amount of data used in MeL, e.g., for few-shot learning [29][30].

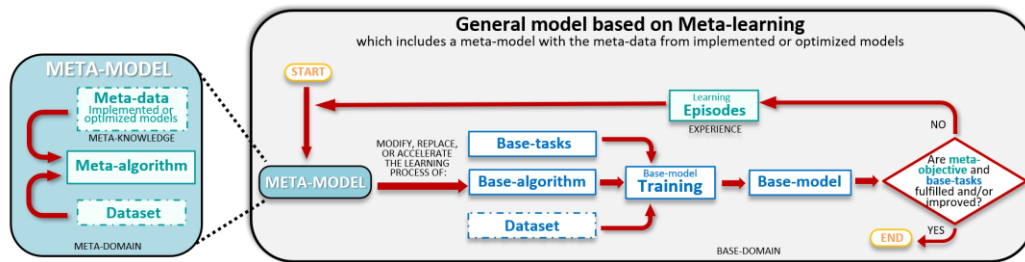


Fig. 1. Diagram of the general model based on the meta-learning paradigm

3. META-LEARNING METHOD

In this section, the datasets and the color constancy methods employed are detailed. Also, the learning processes at the meta- and base-levels of the proposed meta-learning methodology are explained.

3.1. Datasets and Color Constancy Methods

In this work, two datasets were used for the leukocytes' classification: Raabin dataset [31] and UACH (Autonomous University of Chihuahua) dataset [32]. The Raabin dataset [31] was acquired by different medical laboratories at Iran. It includes images of 5 types of leukocytes. For the neutrophils, eosinophils, monocytes, and lymphocytes the images were collected from 72 subjects from 12 to 70 years. The basophil images were acquired from a Chronic Myeloid Leukemia (CML) positive sample. The images of the Raabin dataset are used to form the meta-task $\tau_{meta} \sim D_{meta}$. On the other hand, the UACH dataset [32] was collected by the Autonomous University of Chihuahua, and it includes images of lymphocytes, monocytes, and neutrophils. The images of the UACH dataset are employed to form the base-task $\tau_{base} \sim D_{base}$. Table 1 presents the number of samples for each class per task used for the training and validation sets.

Table 1. Samples distributions for the meta- and base-task used in the MeL methodology

Classes, C	Raabin, τ_{meta}		All samples	UACH, τ_{base}	
	Training samples	Validation samples		Training	Validation
Basophil	212	89	-	-	-
Eosinophil	744	322	-	-	-
Lymphocyte	2427	1034	148	118-120	28-30
Monocyte	561	234	14	11-12	2-3
Neutrophil	6231	2660	91	72-73	18-19
Total	10175	4339	253	202-204	49-51

The collection process of the leukocyte datasets generates that some images and tasks can be considered as OOD, due to the staining method employed, the type of microscope, zoom, and camera, among other image processing operators used. Therefore, to approximate the images to a same color constancy, we implemented the Shades of Gray (SoG) method [24], [33] to all images of τ_{meta} and τ_{base} .

Let's consider $I(x) \in \mathbb{R}^{m \times n \times 3}$ as an RGB image of $m \times n$ pixels, where $x = (x, y)$ indicates the position of each pixel. Before applying the SoG color constancy method, we resized all the images to $224 \times 224 \times 3$ and changed the pixels values $I(x) \leq 25$ for values of 255, making them white. This step was performed because some images have black regions, e.g., black pixels corresponding to the microscope contour. Then, a gamma correction is applied for each channel $ch = \{R, G, B\}$ of $I(x)$ by,

$$\tilde{I}_{ch}(x) = 255 \times \left(\frac{I_{ch}(x)}{255} \right)^{\frac{1}{\gamma}} \quad (9)$$

where γ represents the correction factor for $I_{ch}(x)$. The standard value of gamma is $\gamma = 2.2$, [25]. Next, the SoG color constancy for each channel is obtained by,

$$I_{ch}^{SoG}(x) = \frac{1}{\kappa e_{ch}} \tilde{I}_{ch}(x) \quad (10)$$

In (10), $e_{ch} \in e$ is an illuminant component that multiplied by the constant κ simulates that the pixels values of $\tilde{I}(x)$ were obtained under a certain light source. For each channel of $\tilde{I}_{ch}(x)$, κe_{ch} is computed using the Minkowski norm,

$$\left(\frac{\int (\tilde{I}_{ch}(x))^{\rho} dx}{\int dx} \right)^{\frac{1}{\rho}} = \kappa e_{ch} \quad (11)$$

We used $\rho = 6$ for the Minkowski norm in (11). Then, the constant κ for the normalization of e is computed using the Euclidean distance,

$$\kappa = \left[\sum_{ch} (norm_{ch})^2 \right]^{\frac{1}{2}} = \sqrt{(norm_R)^2 + (norm_G)^2 + (norm_B)^2} \quad (12)$$

According to [25], a perfect white light would be when $norm_{ch} = 1$ with $\kappa = \sqrt{3}$. However, in this paper we experimented with different values of κ and γ to implement the proposed MeL methodology for the classification of the leukocyte's images. For τ_{meta} , the proposed values are; $\gamma = 2.2, \kappa = 3.0$, that are the standard values used in the literature; $\gamma = 2.2, \kappa = 2.0$; and $\gamma = 1.0, \kappa = 1.0$. Besides, we propose a semi-segmentation method based on applying 2 times the SoG color constancy method, i.e., we performed SoG using the original image $I(x)$ with $\gamma_1 = 2.2, \kappa_1 = 2.0$, and then, we used the resulted image to perform SoG for the second time with $\gamma_2 = 0.2, \kappa_2 = 0.1$. We name this method as SoG². Similarly, for τ_{base} , we employed $\gamma = 2.2, \kappa = 3.0, \gamma = 2.2, \kappa = 2.0, \gamma = 1.0, \kappa = 1.0$, and for SoG² $\gamma_1 = 2.2, \kappa_1 = 2.0$ and $\gamma_2 = 0.2, \kappa_2 = 2.3$. Fig. 2 shows some images obtained with the proposed values of γ and κ using one sample image per class of the leukocyte datasets for τ_{meta} and τ_{base} . As can be noticed in Fig. 2, the images obtained with SoG with $\gamma = 2.2, \kappa = 3.0$ present a background color with shades of gray (as the name of the method). Instead, using the other proposed γ and κ values, the background is whiter, and even the color tone of some elements, like platelets and red blood cells, is diminished. Furthermore, with SoG² the nucleus of the leukocyte cells is practically segmented, especially for the images of τ_{meta} .

In the next sections, each set of images of the Raabin and UACH dataset are used for the meta-training of meta-models in τ_{meta} , and for the base-learning in τ_{base} .

3.2. Meta-Models for Meta-Learning

The first process of learning in the proposed MeL methodology consists of the meta-training of meta-models. In this work, we define a meta-model as a model trained at meta-level to produce knowledge that will be used to solve a new task. That is, to solve the base-task τ_{base} that includes little data, the base-learning leverage of the meta-knowledge produced by meta-training meta-models with τ_{meta} .

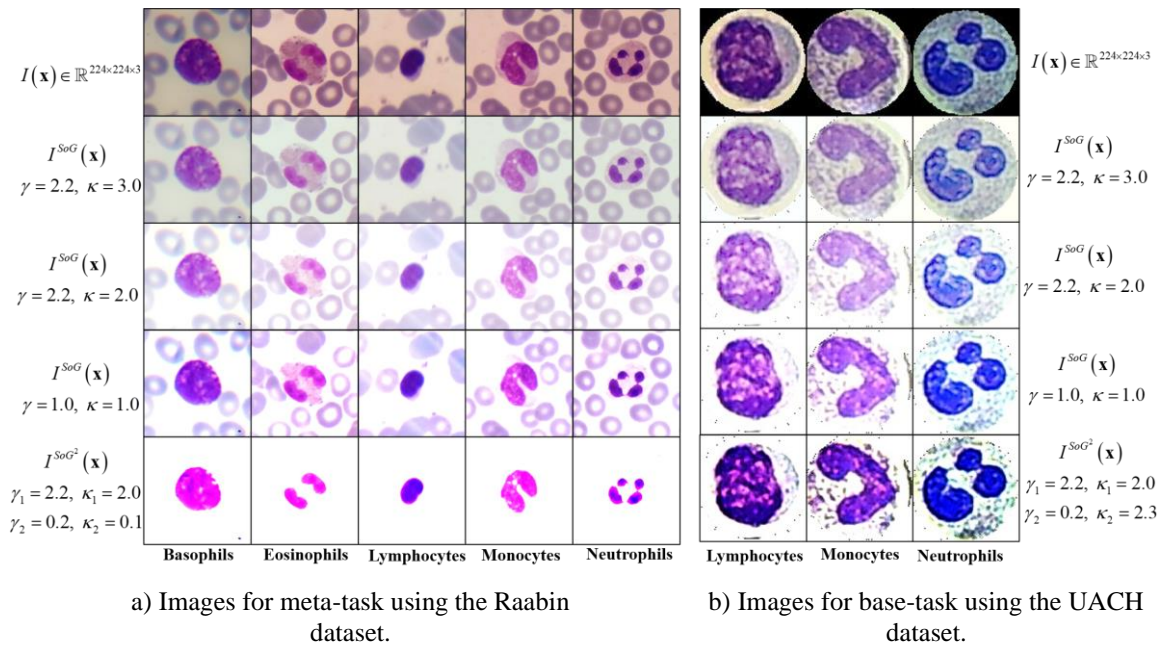


Fig. 2. Images obtained using different values of γ and κ for the Shades of Gray color constancy method

We propose 5 meta-models, all with the same architecture, but each one meta-trained with different color constancy distributions of τ_{meta} (Section 3.1). The architecture of each meta-model includes a data augmentation layer at the input, that randomly flip the input image horizontally or vertically, randomly translate it by $\pm 10\%$ and/or randomly rotate it in range of $[-72, 72]$ degrees. The output of the data augmentation layer is connected to a prior-model architecture. In this case, we employed Xception [34], [35] as prior-model, ψ_{Xce} . Xception was trained and optimized for the ImageNet dataset, and we employed the pre-trained architecture without including the last pretrained classification layer. Besides, for the meta-training, most of the parameters of ψ_{Xce} are frozen θ_{fzn} , i.e., they are not adapted in the meta-training. Only the parameters θ_{nfzn} of ψ_{Xce} that correspond from the last convolutional layer to the last layer of ψ_{Xce} are adapted. The last layer of ψ_{Xce} is connected to an average pooling layer that extracts feature vectors. Then, the feature vectors are connected to a dropout layer and next to a classification layer that includes 5 neurons, one per leukocyte class, with a softmax activation function. The softmax functions of the meta-models generates probability distributions p_j^{method} , for each color constancy method employed, $method = \{original, \gamma 2.2_ \kappa 3.0, \gamma 2.2_ \kappa 2.0, \gamma 1.0_ \kappa 1.0, SoG^2\}$. Then, the output class is $C_{method} = \underset{j}{argmax} p_j^{method}$, $j = 1, \dots, 5$. Fig. 3 shows the complete architecture of each meta-model used in the meta-training of τ_{meta} .

In addition to the meta-models presented in Fig. 3, we implemented a meta-model based on Ensemble Learning (EL). The meta-model based on ensemble learning, involving the SoG method, uses the generated probabilities distributions p_j^{method} of all the meta-trained meta-models to produce a new distribution, d_j^{EL-SoG} , for improving the leukocyte classification accuracy of τ_{meta} of the meta-models, by,

$$d_j^{EL-SoG} = \sum_{\psi} w^{method} p_j^{method} ; C^{EL-SoG} = \underset{j}{argmax} d_j \quad (13)$$

To obtain the weights w^{method} , we implemented a search space with the combinations of weights from 0.0 to 1.0 increasing 0.1 by 0.1 the value of each weight. Then, we selected the meta-model with the weights that maximized the validation accuracy.

In the next section, all the meta-trained meta-models based on ψ_{Xce} are used to solve τ_{base} .

3.3. Base-Models for Base-Learning

The second learning process in the proposed MeL methodology consists of the base-training or adaptation of the meta-knowledge generated by the meta-models into the base-models to solve a new base-task τ_{base} .

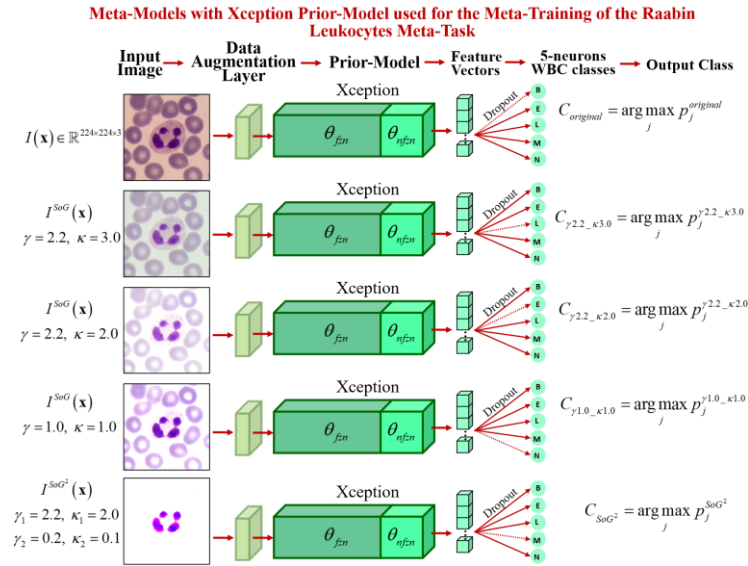


Fig. 3. Meta-models used in the meta-training for the meta-task

The base-models implemented in this paper for the classification of the images of the new base-task τ_{base} are shown in Fig. 4. These base-models use the meta-models as feature extractors and all the optimized meta-parameters of the meta-models are frozen parameters θ_{fzm}^* , i.e., these parameters are not adapted. Each base-model has a corresponding meta-model for each type of image. For instance, for the images of τ_{base} with the original format, the base-model uses the meta-model that was meta-trained using also images with the original format. Similarly, for the images obtained with the SoG methods, each base-model has its corresponding meta-model. The next step in the base-models for the classification of the images of τ_{base} is the normalization of the extracted features using L^2 . Then, we employed UMAP [36] as a dimensionality reduction method of the normalized features. Finally, Support Vector Machines (SVM) are trained with 80% of the UMAP features and validated with the remain 20%, using cross-validation with 5 folders.

Similar to the ensemble learning meta-model implemented at the meta-level, we developed a base-model based on an algorithm to improve the performance of the leukocyte classification for τ_{base} at the base-level using the base-models of Fig. 4. Given the i -th image, the proposed Algorithm 1 compute a new conditioned output O_i^{new} using all outputs O_i^{method} of the SVM_{method} , where $O_i^{method} = \{0,1,2\}$, 0 for lymphocytes, 1 for monocytes, and 2 for neutrophils.

Algorithm 1 Conditioned Output (CO) for the leukocyte classification for the UACH τ_{base}

Require: O_i^{method} for all methods, $method = \{original, \gamma2.2_kappa3.0, \gamma2.2_kappa2.0, \gamma1.0_kappa1.0, SoG^2\}$

- 1: **for** $i \leq num_samples$ **do**
- 2: compute $\alpha_i^{method} = \begin{cases} 1 & \text{iff } O_i^{method} = 0 \\ 0 & \text{otherwise} \end{cases}$
- 3: compute $\beta_i^{method} = \begin{cases} 1 & \text{iff } O_i^{method} = 1 \\ 0 & \text{otherwise} \end{cases}$
- 4: compute $\delta_i^{method} = \begin{cases} 1 & \text{iff } O_i^{method} = 2 \\ 0 & \text{otherwise} \end{cases}$
- 5: **if** $\sum_{method} \alpha_i^{method} = \sum_{method} \beta_i^{method}$ **OR** $\sum_{method} \beta_i^{method} = \sum_{method} \delta_i^{method}$ **do**
- 6: $O_i^{new} \leftarrow 1$
- 7: **continue** with the next sample i
- 8: **if** $\sum_{method} \alpha_i^{method} = \sum_{method} \delta_i^{method}$ **do**
- 9: $O_i^{new} \leftarrow 2$
- 10: **continue** with the next sample i
- 11: $O_i^{new} \leftarrow \text{mode}[O_i^{original}, O_i^{\gamma2.2_kappa3.0}, O_i^{\gamma2.2_kappa2.0}, O_i^{\gamma1.0_kappa1.0}, O_i^{SoG^2}]$

Base-Models with Meta-Trained Meta-Models for the UACH Leukocytes Classification Base-Task

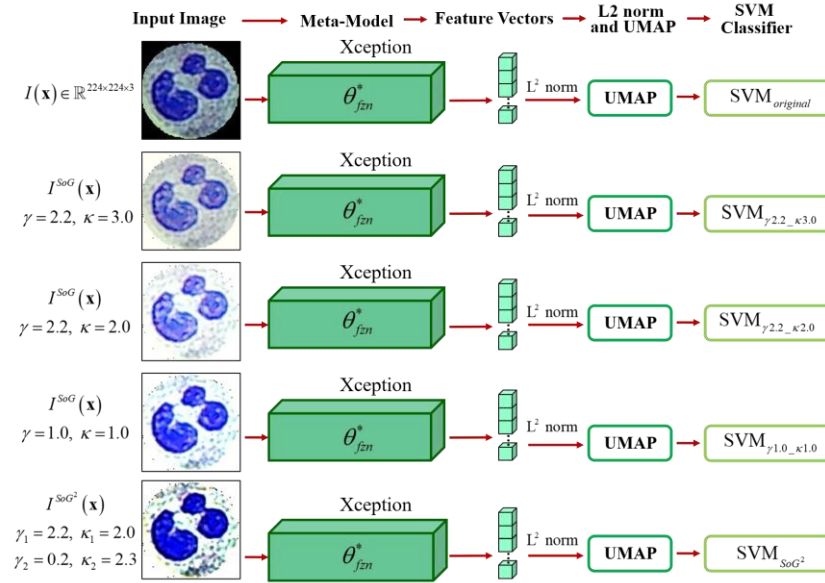


Fig. 4. Base-models used in the base-training for the base-task

As can be noticed in Algorithm 1, we applied a hierarchy criterion if it were the case of a tie between the classification outputs of the SVMs. That is, we assign class 1 to the new output O_i^{new} if 2 SVMs label the input i as 1 or monocytes, and the other 2 SVMs label the same input either as 0 or 2. It was done to give more attention to this class because only 5.53% of the total samples are monocytes in the UACH dataset. On the other hand, if it were the case that 2 SVMs classify the input image as 0, and the other 2 SVMs classify the input image as 2, then we assign 2 to the O_i^{new} because 35.97% of the samples are neutrophils.

4. RESULTS AND DISCUSSION

The results of our MeL methodology are addressed in 3 sections: results of the meta-training of the Raabin leukocytes meta-task at the meta-level; results of the base-learning adaptation for the solution of the UACH leukocytes base-task at the base-level; and results of the propose MeL methodology in comparison with state-of-the-art models.

4.1. Results of the Meta-Learning of the Raabin Leukocytes Meta-Task

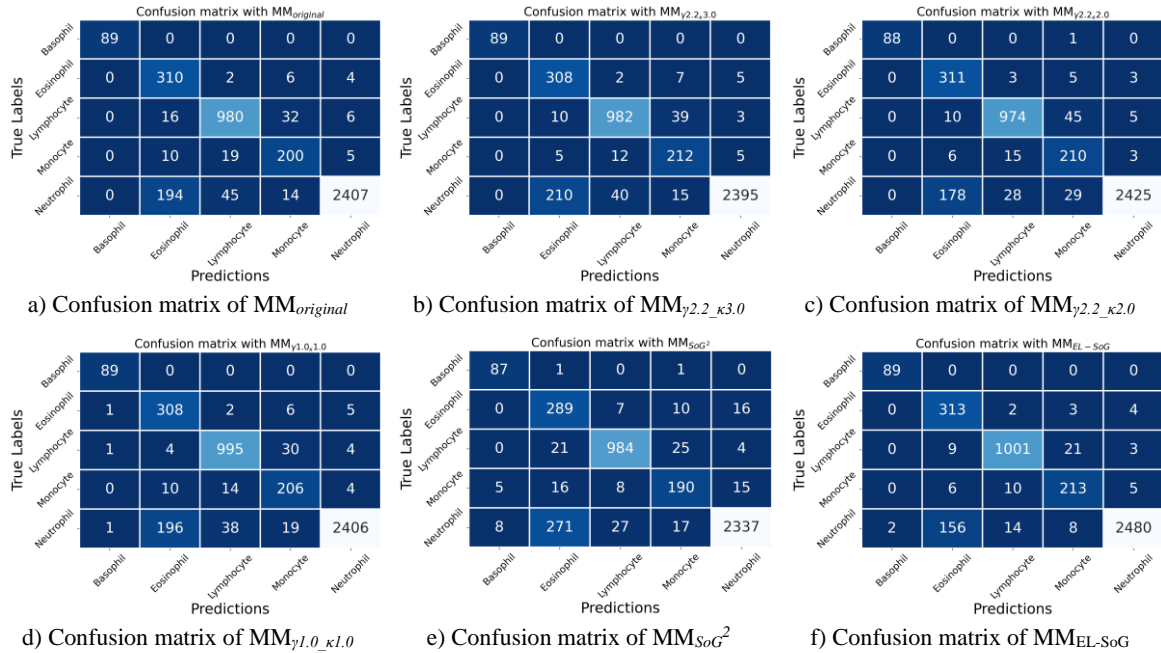
The meta-models MM_{method} based on Xception prior-model for each color constancy method (Section 3.2) were meta-trained for 40 epochs with Adam optimizer. The first 20 epochs with an initial learning rate of 0.001 that was decreasing by 0.1 every 5 epochs. After the first 20 epochs, the best meta-model is meta-trained for 20 epochs more with a learning rate of 1×10^{-7} .

On the other hand, the meta-models based on ensemble learning MM_{EL-SoG} were obtained by weighting the probability distributions of MM_{method} . The main meta-model based on EL was obtained by weighting all the probability distributions $d^{EL-SoG} = 0.2p^{original} + 0.0p^{\gamma 2.2, \kappa 3.0} + 0.7p^{\gamma 2.2, \kappa 2.0} + 0.4p^{\gamma 1.0, \kappa 1.0} + 0.9p^{SoG^2}$. To observe if the SoG^2 method helped to improve the leukocyte classification, we also generated other 3 versions of meta-models based on EL. In the version 1, v1, we removed the MM_{SoG^2} from the ensemble obtaining $d_{v1}^{EL-SoG} = 0.8p^{original} + 0.7p^{\gamma 2.2, \kappa 3.0} + 0.3p^{\gamma 2.2, \kappa 2.0} + 1.0p^{\gamma 1.0, \kappa 1.0}$. In the v2 we removed the $MM_{original}$ with the original images, and we obtained $d_{v2}^{EL-SoG} = 0.0p^{\gamma 2.2, \kappa 3.0} + 0.7p^{\gamma 2.2, \kappa 2.0} + 0.2p^{\gamma 1.0, \kappa 1.0} + 0.7p^{SoG^2}$. We noticed in the v2 that the $MM_{\gamma 2.2, \kappa 3.0}$ that uses the traditional values of γ and κ was weighted with 0.0, so, we removed it to produce v3 of EL, obtaining $d_{v3}^{EL-SoG} = 0.9p^{\gamma 2.2, \kappa 2.0} + 0.7p^{SoG^2}$. The performances achieved with the validation set by all the meta-models are presented in Table 2. Besides, the confusion matrices of the meta-models MM_{method} and of the MM_{EL-SoG} are presented in Fig. 5.

The best performances in Table 2 were obtained by MM_{EL-SoG} with a precision of 0.8928, sensitivity of 0.9565, F1-score of 0.9178, and accuracy of 0.9440. However, it is important to notice that the EL with two models: MM_{SoG^2} in combination with $MM_{\gamma 2.2, \kappa 2.0}$, achieved better sensitivity, F1-score, and accuracy than the $MM_{EL-SoG-v1}$ that is composed by 4 meta-models without including the MM_{SoG^2} . In consequence, our SoG^2 improves the leukocyte classification for the τ_{meta} .

Table 2. Performances achieved by the meta-models used in the MeL methodology

Meta-models based on the color constancy method	Precision	Sensitivity	F1-score	Accuracy
MM _{original}	0.8619	0.9340	0.8881	0.9186
MM _{γ2.2_κ3.0}	0.8594	0.9425	0.8901	0.9186
MM _{γ2.2_κ2.0}	0.8581	0.9411	0.8896	0.9237
MM _{γ1.0_κ1.0}	0.8589	0.9407	0.8904	0.9228
MM _{SoG²}	0.8159	0.9034	0.8459	0.8958
MM _{EL-SoG-v1}	0.8814	0.9477	0.9068	0.9343
MM _{EL-SoG-v3}	0.8797	0.9530	0.9085	0.9380
MM _{EL-SoG-v2}	0.8869	0.9550	0.9138	0.9417
MM_{EL-SoG}	0.8928	0.9565	0.9178	0.9440

**Fig. 5.** Confusion matrices obtained with the meta-models

To conclude the results for the meta-learning of τ_{meta} , Table 3 presents the precision, sensitivity, and F1-score for each class, obtained with the MM_{EL-SoG}. It can be seen in Table 3 that our MM_{EL-SoG} classify perfectly the basophils, and the most difficult class for our meta-model are the eosinophils.

Table 3. Performances achieved by the meta-models MM_{EL-SoG} for class of leukocyte

Class of leukocyte	Precision	Sensitivity	F1-score
Basophil	0.9780	1.0000	0.9889
Eosinophil	0.6467	0.9720	0.7767
Lymphocyte	0.9747	0.9681	0.9714
Monocyte	0.8694	0.9103	0.8894
Neutrophil	0.9952	0.9323	0.9627

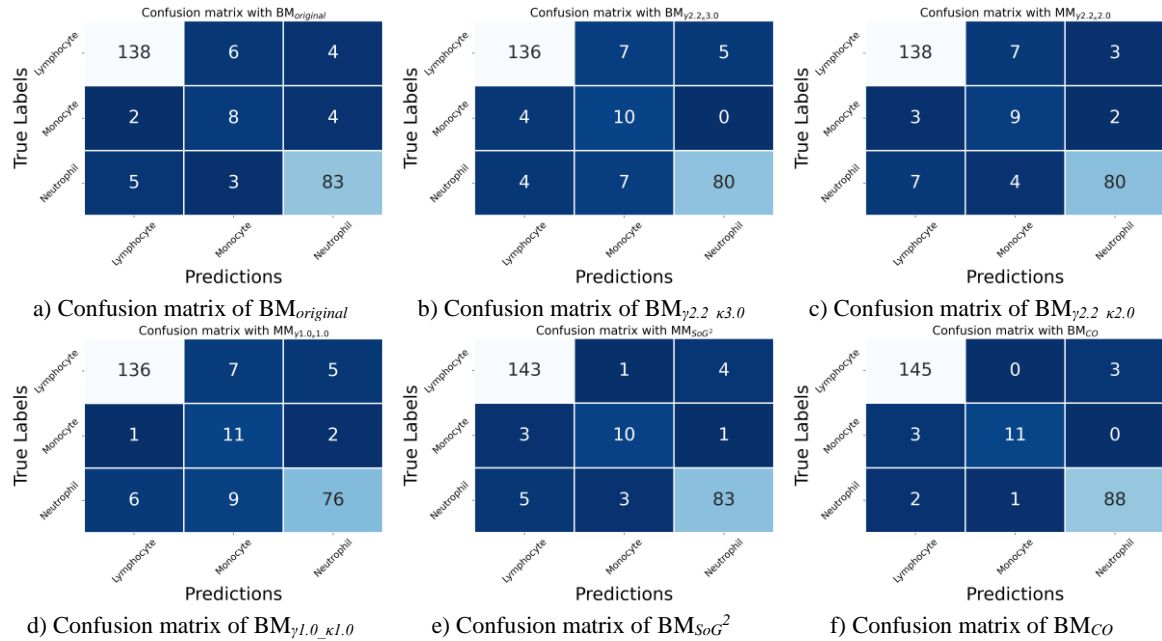
4.2. Results of the Base-Learning of the UACH Leukocytes Base-Task

The average results of the cross-validation with 5 folders for the leukocyte classification base-task τ_{base} with the UACH dataset using the base-models BM_{method} and the base-model based on the proposed algorithm BM_{CO} are shown in Table 4. Besides, the confusion matrices for all base-models are presented in Fig. 6.

From the base-models BM_{method}, the BM_{SoG²} achieved the best performances, 0.8753 ± 0.0544 , 0.8704 ± 0.0532 , 0.8704 ± 0.0380 , 0.9330 ± 0.0263 , on precision, sensitivity, F1-score, and accuracy, respectively. However, the proposed base-model BM_{CO} outperforms all the methods shown in Table 4 achieving performances above 90%, especially in accuracy with 96.47%.

Table 4. Performances achieved by the base-models used in the MeL methodology

Meta-models based on the color constancy method	Precision	Sensitivity	F1-score	Accuracy
$BM_{original}$	0.8016 ± 0.1079	0.8036 ± 0.0811	0.8036 ± 0.0921	0.9052 ± 0.0545
$BM_{\gamma 2.2 \ \kappa 3.0}$	0.7847 ± 0.0560	0.8330 ± 0.0541	0.8330 ± 0.0510	0.8933 ± 0.0317
$BM_{\gamma 2.2 \ \kappa 2.0}$	0.7800 ± 0.0407	0.8260 ± 0.0591	0.8260 ± 0.0436	0.8971 ± 0.0267
$BM_{\gamma 1.0 \ \kappa 1.0}$	0.7846 ± 0.0964	0.8514 ± 0.0893	0.8514 ± 0.0755	0.8816 ± 0.0408
BM_{SoG^2}	0.8753 ± 0.0544	0.8704 ± 0.0532	0.8704 ± 0.0380	0.9330 ± 0.0263
BM_{CO}	0.9561 ± 0.0550	0.9156 ± 0.0657	0.9156 ± 0.0565	0.9647 ± 0.0260

**Fig. 6.** Confusion matrices obtained with the base-models

4.3. Comparison of Our MeL Methodology with State-Of-The-Art Models

In this section, we compared our MeL methodology with state-of-the-art models, [Table 5](#), using the average of precision, sensitivity, F1-score, and accuracy metrics regarding the τ_{meta} and τ_{base} . For the comparison, we employed our meta-model MM_{EL-SoG} based on ensemble learning and SoG color constancy method for the Raabin dataset, and our base-model BM_{CO} with the conditioned output for the UACH dataset. In [Table 5](#), the model DRFA-Net [37] achieved the best mean scores for precision, F1-score, and accuracy. However, our MM_{EL-SoG} outperforms all models regarding the mean sensitivity metric, and its mean scores in precision, F1, and accuracy are also competitive. The sensitivity metric is crucial for imbalanced datasets because it is used to assess the true positive rates and our MeL methodology takes into consideration the classes with fewer samples. That is, our MM_{EL-SoG} achieved high sensitivity scores for all classes with an average sensitivity of 95.65%, meanwhile, other models are biased to the classes with a higher number of samples for a maximum average sensitivity score of 93.40%. To deepen this analysis of the Raabin dataset, we compared the performances by each class of leukocyte of the top-3 models of [Table 5](#) concerning the accuracy scores, Shape and Color features + SVM [38], DRFA-Net [37], and our proposed MM_{EL-SoG} , in [Table 6](#).

In [Table 6](#), our MM_{EL-SoG} outperforms all presented state-of-the-art models in terms of sensitivity for 4 of the 5 classes (except in neutrophils). For basophils and monocytes, our model achieved better performances in all metrics. The basophils and monocytes are the classes with fewer samples in the validation set of Raabin, 89 for Basophils and 234 for monocytes of 4339 samples in total. It means that the samples of basophils are 2.05% of the validation data, and monocytes are 5.39%. For imbalanced datasets, like this case, the sensitivity score is very important and our MM_{EL-SoG} takes into consideration the classes with fewer samples, excelling in the sensitivity metric of other state-of-the-art presented models. Furthermore, as it was said in [Section 3.1](#), the basophil images were acquired from a Chronic Myeloid Leukemia positive sample, giving greater importance to this particular class, and our model obtained 100% of the sensitivity.

Table 5. Comparison of our meta-model MM_{EL-SoG} for the Raabin dataset and our BM_{CO} for the UACH dataset with other state-of-the-art models using the average of Precision (P), Sensitivity (S), F1-score, and Accuracy (ACC) for all leukocyte classes

Model	Dataset	P (%)	S (%)	F1 (%)	ACC (%)
CNN-RNN [39]*	Raabin Test A	86.56	85.97	86.26	87.34
FusedCNN [40]*		87.98	85.37	86.66	90.68
NucSegNet [41]*		89.65	90.71	90.18	93.08
WBCaps [42]*		88.73	89.76	89.24	92.74
CNNDC [43]*		89.56	90.25	89.90	93.41
Shape and Color features + SVM [38]		89.79	92.76	91.04	94.65
DRFA-Net [37]		90.43	93.40	91.89	95.17
Our MM_{EL-SoG}		89.28	95.65	91.78	94.40
Handcraft features + inferences [32]	61.06	68.52	59.00	68.63	
Our BM_{CO}	UACH	0.9561±0.055	0.9156±0.065	0.9156±0.056	0.9647±0.026
		0	7	5	0

* Results reported in [37]

Table 6. Comparison of our meta-model MM_{EL-SoG} with other state-of-the-art models using Precision (P), Sensitivity (S), F1-score, and Accuracy (ACC) for each leukocyte class

Model / class	Basophil			Eosinophil			Lymphocyte			Monocyte			Neutrophil		
	P (%)	S (%)	F1 (%)	P (%)	S (%)	F1 (%)	P (%)	S (%)	F1 (%)	P (%)	S (%)	F1 (%)	P (%)	S (%)	F1 (%)
Shape-Color features + SVM [38]	96.59	95.51	96.05	72.24	91.30	80.66	97.23	95.07	96.14	84.87	86.32	85.59	98.00	95.60	96.78
DRFA-Net [37]	96.85	95.83	96.34	73.66	92.03	81.83	97.69	96.21	96.94	85.62	87.09	86.35	98.32	95.86	97.07
MM_{EL-SoG}	97.80	100	98.89	64.67	97.20	77.67	97.47	96.81	97.14	86.94	91.03	88.94	99.52	93.23	96.27

On the other hand, regarding the results in the UACH dataset in Table 5, it can be seen that leveraging the meta-knowledge acquired when solving the meta-task helped our BM_{CO} to improve significantly in comparison with the work presented in [32], which extracts features of the leukocytes manually. Besides, the precision, sensitivity, F1, and accuracy scores are in the range of the scores reported in other datasets of leukocytes, like Raabin.

As future work, our MeL methodology can be implemented in other medical areas where the data is limited and the new task can be benefited from the knowledge generated at the meta-learning, for example, for image classification of brain tumors [44], skin lesions [45], COVID-19 chest X-ray [46], among other medical images [47].

5. CONCLUSION

In this work, we proposed a meta-learning methodology for the classification of leukocytes by implementing two learning processes. The first learning process is at the meta-level and consisted in the meta-training of meta-models to solve a meta-task formed with the leukocyte's images of the Raabin dataset. The proposed meta-models are based on the Xception prior-model and are meta-trained using images with different color constancy methods. For the color constancies, we implemented the Shades of Gray method that has been used with images for skin diagnosis. However, we proposed different parameters for the SoG method for leukocyte image classification. Furthermore, we propose SoG^2 , that is a method based on the implementation of SoG twice and with different parameters, that can be used as a tool in the image segmentation processes. The performances obtained for the meta-learning are competitive regarding the state-of-the-art models, especially for the sensitivity metric, which is very important for imbalanced datasets, achieving 95.65%, +2.25% better than the best sensitivity score of the previous reported works. In addition, for the basophil images that were acquired from a chronic myeloid leukemia positive sample, our MM_{EL-SoG} obtained 100% for sensitivity.

On the other hand, the second learning process of our meta-learning methodology is at the base-level. In this case the base-models leveraged the knowledge produced by solving the meta-task to solve a new base-task, formed with leukocyte images of the UACH dataset. The base-models are adapted using feature vectors produced by the meta-models. Then, the feature vectors are normalized with L^2 and their dimension are reduced with UMAP. The outputs of the UMAP are classified with SVM. It is important to notice that we reused the knowledge produced at the meta-level to automatically extract features of a new dataset of leukocytes. The performances achieved at the base-level are competitive with other datasets evaluated in the state-of-the-art

with values around 91% for sensitivity and F1 scores, 95% for precision, and 96% for accuracy for the base-task.

The findings of this work show that our methodology based on the meta-learning paradigm can be used for other medical areas where it is difficult to collect medical image data and the knowledge can be reused to solve similar tasks.

Acknowledgments

The authors greatly appreciate the support of Tecnológico Nacional de México / I. T. Chihuahua to perform this research under grant 10071.21-P and 14044.22-P.

REFERENCES

- [1] P. Rastogi, K. Khanna, and V. Singh, "LeuFeatx: Deep learning-based feature extractor for the diagnosis of acute leukemia from microscopic images of peripheral blood smear," *Comput. Biol. Med.*, vol. 142, no. March, p. 105236, Mar. 2022, <https://doi.org/10.1016/j.compbiomed.2022.105236>.
- [2] M. Zolfaghari and H. Sajedi, "A survey on automated detection and classification of acute leukemia and WBCs in microscopic blood cells," *Multimed. Tools Appl.*, vol. 81, no. 5, pp. 6723–6753, Jan. 2022, <https://doi.org/10.1007/s11042-022-12108-7>.
- [3] B. M. Henry *et al.*, "Red Blood Cell Distribution Width (RDW) Predicts COVID-19 Severity: A Prospective, Observational Study from the Cincinnati SARS-CoV-2 Emergency Department Cohort," *Diagnostics*, vol. 10, no. 9, p. 618, Aug. 2020, <https://doi.org/10.3390/diagnostics10090618>.
- [4] J. Kasten-Jolly and D. A. Lawrence, "Differential blood leukocyte populations based on individual variances and age," *Immunol. Res.*, vol. 70, no. 1, pp. 114–128, Feb. 2022, <https://doi.org/10.1007/s12026-021-09257-6>.
- [5] T. O'Connor, A. Anand, B. Andemariam, and B. Javidi, "Deep learning-based cell identification and disease diagnosis using spatio-temporal cellular dynamics in compact digital holographic microscopy," *Biomed. Opt. Express*, vol. 11, no. 8, pp. 4491–4508, Aug. 2020, <https://doi.org/10.1364/BOE.399020>.
- [6] T. O'Connor, J.-B. Shen, B. T. Liang, and B. Javidi, "Digital holographic deep learning of red blood cells for field-portable, rapid COVID-19 screening," *Opt. Lett.*, vol. 46, no. 10, pp. 2344–2347, May 2021, <https://doi.org/10.1364/OL.426152>.
- [7] C. G. Varma, P. Nagaraj, V. Muneeswaran, K. Muthamil Sudar, M. Mokshagni, and M. Jaswanth, "Astute Segmentation and Classification of leucocytes in blood microscopic smear images using titivated K-means clustering and robust SVM techniques," in *5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021*, pp. 818–824, 2021, <https://doi.org/10.1109/ICICCS51141.2021.9432309>.
- [8] M. Sajjad *et al.*, "Leukocytes Classification and Segmentation in Microscopic Blood Smear: A Resource-Aware Healthcare Service in Smart Cities," *IEEE Access*, vol. 5, pp. 3475–3489, 2017, <https://doi.org/10.1109/ACCESS.2016.2636218>.
- [9] N. Dong, M. die Zhai, J. fang Chang, and C. ho Wu, "A self-adaptive approach for white blood cell classification towards point-of-care testing," *Appl. Soft Comput.*, vol. 111, no. November, p. 107709, 2021, <https://doi.org/10.1016/j.asoc.2021.107709>.
- [10] D. Gupta, J. Arora, U. Agrawal, A. Khanna, and V. H. C. de Albuquerque, "Optimized Binary Bat algorithm for classification of white blood cells," *Measurement*, vol. 143, no. September, pp. 180–190, 2019, <https://doi.org/10.1016/j.measurement.2019.01.002>.
- [11] D. Yang, H. Zhao, T. Han, Q. Kang, J. Ma, and H. Lu, "Leukocyte subtypes identification using bilinear self-attention convolutional neural network," *Measurement*, vol. 173, no. March, p. 108643, 2021, <https://doi.org/10.1016/j.measurement.2020.108643>.
- [12] J. Basnet, A. Alsadoon, P. W. C. Prasad, S. Al Aloussi, and O. H. Alsadoon, "A Novel Solution of Using Deep Learning for White Blood Cells Classification: Enhanced Loss Function with Regularization and Weighted Loss (ELFRWL)," *Neural Process. Lett.*, vol. 52, no. 2, pp. 1517–1553, 2020, <https://doi.org/10.1007/s11063-020-10321-9>.
- [13] A. Khouani, M. El Habib Daho, S. A. Mahmoudi, M. A. Chikh, and B. Benzineb, "Automated recognition of white blood cells using deep learning," *Biomed. Eng. Lett.*, vol. 10, no. 3, pp. 359–367, 2020, <https://doi.org/10.1007/s13534-020-00168-3>.
- [14] R. B. Hegde, K. Prasad, H. Hebbar, and B. M. K. Singh, "Feature extraction using traditional image processing and convolutional neural network methods to classify white blood cells: a study," *Australas. Phys. Eng. Sci. Med.*, vol. 42, no. 2, pp. 627–638, 2019, <https://doi.org/10.1007/s13246-019-00742-9>.
- [15] E. Başaran, "Classification of white blood cells with SVM by selecting SqueezeNet and LIME properties by mRMR method," *Signal, Image Video Process.*, vol. 16, no. 7, pp. 1821–1829, Jan. 2022, <https://doi.org/10.1007/s11760-022-02141-2>.
- [16] A. Çınar and S. A. Tuncer, "Classification of lymphocytes, monocytes, eosinophils, and neutrophils on white blood cells using hybrid Alexnet-GoogleNet-SVM," *SN Appl. Sci.*, vol. 3, no. 4, pp. 1–11, 2021, <https://doi.org/10.1007/s42452-021-04485-9>.
- [17] L. Ma, R. Shuai, X. Ran, W. Liu, and C. Ye, "Combining DC-GAN with ResNet for blood cell image classification," *Med. Biol. Eng. Comput.*, vol. 58, no. 6, pp. 1251–1264, 2020, <https://doi.org/10.1007/s11517-020-02163-3>.
- [18] M. Lippeveld *et al.*, "Classification of Human White Blood Cells Using Machine Learning for Stain-Free Imaging

- Flow Cytometry,” *Cytom. Part A*, vol. 97, no. 3, pp. 308–319, Mar. 2020, <https://doi.org/10.1002/cyto.a.23920>.
- [19] Y. Duan *et al.*, “Leukocyte classification based on spatial and spectral features of microscopic hyperspectral images,” *Opt. Laser Technol.*, vol. 112, pp. 530–538, 2019, <https://doi.org/10.1016/j.optlastec.2018.11.057>.
- [20] A. T. Sahlol, P. Kollmannsberger, and A. A. Ewees, “Efficient Classification of White Blood Cell Leukemia with Improved Swarm Optimization of Deep Features,” *Sci. Rep.*, vol. 10, no. 2536, pp. 1–11, Feb. 2020, <https://doi.org/10.1038/s41598-020-59215-9>.
- [21] X. He, K. Zhao, and X. Chu, “AutoML: A survey of the state-of-the-art,” *Knowledge-Based Syst.*, vol. 212, p. 106622, 2021, <https://doi.org/10.1016/j.knosys.2020.106622>.
- [22] A. Langdon, M. Botvinick, H. Nakahara, K. Tanaka, M. Matsumoto, and R. Kanai, “Meta-learning, social cognition and consciousness in brains and machines,” *Neural Networks*, vol. 145, pp. 80–89, Jan. 2022, <https://doi.org/10.1016/j.neunet.2021.10.004>.
- [23] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, “Meta-Learning in Neural Networks: A Survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–20, 2021, <https://doi.org/10.1109/TPAMI.2021.3079209>.
- [24] G. D. Finlayson and E. Trezzi, “Shades of gray and colour constancy,” *12th Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications*, pp. 37–41, 2004, <https://research-portal.uea.ac.uk/en/publications/shades-of-gray-and-colour-constancy>.
- [25] C. Barata, M. E. Celebi, and J. S. Marques, “Improving dermoscopy image classification using color constancy,” *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 3, pp. 1146–1152, 2014, <https://doi.org/10.1109/JBHI.2014.2336473>.
- [26] S. Milantev, V. Olyunin, N. Milanteva, I. Bykov, and I. Bessmertny, “Skin Lesion Analysis Using Ensemble of CNN with Dermoscopic Images and Metadata,” in *CEUR Workshop Proceedings*, vol. 2893, pp. 1–11, 2020, https://ceur-ws.org/Vol-2893/paper_5.pdf.
- [27] M. N. Alche, D. Acevedo, and M. Mejail, “EfficientARL: improving skin cancer diagnoses by combining lightweight attention on EfficientNet,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pp. 3354–3360, 2021, <https://doi.org/10.1109/ICCVW54120.2021.00374>.
- [28] C. Finn, “Multi-Task and Meta-Learning,” in *CS330*, Stanford, 2019, <http://cs330.stanford.edu/>.
- [29] R. Khadka *et al.*, “Meta-learning with implicit gradients in a few-shot setting for medical image segmentation,” *Comput. Biol. Med.*, vol. 143, 2022, <https://doi.org/10.1016/j.compbiomed.2022.105227>.
- [30] H. J. Ye, L. Ming, D. C. Zhan, and W. L. Chao, “Few-Shot Learning with a Strong Teacher,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 2022, <https://doi.org/10.1109/TPAMI.2022.3160362>.
- [31] Z. M. Kouzehkanan *et al.*, “A large dataset of white blood cells containing cell locations and types, along with segmented nuclei and cytoplasm,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–14, 2022, <https://doi.org/10.1038/s41598-021-04426-x>.
- [32] A. Reyes-Esparza, M. I. Chacón-Murguía, J. A. Ramírez-Quintana, and C. Arzate-Quintana, “Desarrollo de un algoritmo para reconocimiento de leucocitos en imágenes de muestras de sangre periférica,” in *Congreso Internacional en Ingeniería Electrónica, memoria ELECTRO*, pp. 1–6, 2022.
- [33] M. A. Hussain and A. S. Akbari, “Color constancy adjustment using sub-blocks of the image,” *IEEE Access*, vol. 6, pp. 46617–46629, Aug. 2018, <https://doi.org/10.1109/ACCESS.2018.2866792>.
- [34] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *30th IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1800–1807, 2017, <https://doi.org/10.1109/ACCESS.2018.2866792>.
- [35] A. Dhillon and G. K. Verma, “Convolutional neural network: a review of models, methodologies and applications to object detection,” *Prog. Artif. Intell.*, vol. 9, no. 2, pp. 85–112, 2020, <https://doi.org/10.1007/s13748-019-00203-0>.
- [36] A. Diaz-Papkovich, L. Anderson-Trocmé, and S. Gravel, “A review of UMAP in population genetics,” *J. Hum. Genet.*, vol. 66, no. 1, pp. 85–91, Oct. 2021, <https://doi.org/10.1038/s10038-020-00851-4>.
- [37] L. Jiang, C. Tang, and H. Zhou, “White blood cell classification via a discriminative region detection assisted feature aggregation network,” *Biomed. Opt. Express*, vol. 13, no. 10, pp. 5246–5260, Oct. 2022, <https://doi.org/10.1364/BOE.462905>.
- [38] S. Tavakoli, A. Ghaffari, Z. M. Kouzehkanan, and R. Hosseini, “New segmentation and feature extraction algorithm for classification of white blood cells in peripheral smear images,” *Sci. Rep.*, vol. 11, p. 19428, Sep. 2021, <https://doi.org/10.1038/s41598-021-98599-0>.
- [39] G. Liang, H. Hong, W. Xie, and L. Zheng, “Combining Convolutional Neural Network With Recursive Neural Network for Blood Cell Image Classification,” *IEEE Access*, vol. 6, pp. 36188–36197, Jul. 2018, <https://doi.org/10.1109/ACCESS.2018.2846685>.
- [40] S. Pang, A. Du, M. A. Orgun, and Z. Yu, “A novel fused convolutional neural network for biomedical image classification,” *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 107–121, 2019, <https://doi.org/10.1007/s11517-018-1819-y>.
- [41] P. P. Banik, R. Saha, and K. D. Kim, “An Automatic Nucleus Segmentation and CNN Model based Classification Method of White Blood Cell,” *Expert Syst. Appl.*, vol. 149, p. 113211, 2020, <https://doi.org/10.1016/j.eswa.2020.113211>.
- [42] Y. Y. Baydilli and Ü. Atila, “Classification of white blood cells using capsule networks,” *Comput. Med. Imaging Graph.*, vol. 80, p. 101699, 2020, <https://doi.org/10.1016/j.compmedimag.2020.101699>.
- [43] P. K. Das and S. Meher, “An efficient deep Convolutional Neural Network based detection and classification of

- Acute Lymphoblastic Leukemia,” *Expert Syst. Appl.*, vol. 183, p. 115311, 2021, <https://doi.org/10.1016/j.eswa.2021.115311>.
- [44] A. Sekhar, S. Biswas, R. Hazra, A. K. Sunaniya, A. Mukherjee, and L. Yang, “Brain Tumor Classification Using Fine-Tuned GoogLeNet Features and Machine Learning Algorithms: IoMT Enabled CAD System,” *IEEE J. Biomed. Heal. Informatics*, vol. 26, no. 3, pp. 983–991, Mar. 2022, <https://doi.org/10.1109/JBHI.2021.3100758>.
- [45] M. A. Rasel, U. H. Obaidallah, and S. A. Kareem, “Convolutional Neural Network-Based Skin Lesion Classification With Variable Nonlinear Activation Functions,” *IEEE Access*, vol. 10, pp. 83398–83414, 2022, <https://doi.org/10.1109/ACCESS.2022.3196911>.
- [46] S. A. B. P and C. S. R. Annavarapu, “Deep learning-based improved snapshot ensemble technique for COVID-19 chest X-ray classification,” *Appl. Intell.*, vol. 51, no. 5, pp. 3104–3120, May 2021, <https://doi.org/10.1007/s10489-021-02199-4>.
- [47] S. Asgari Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, and G. Hamarneh, “Deep semantic segmentation of natural and medical images: a review,” *Artif. Intell. Rev.*, vol. 54, no. 1, pp. 137–178, 2021, <https://doi.org/10.1007/s10462-020-09854-1>.

BIOGRAPHY OF AUTHORS



Eduardo Rivas-Posada, received the BE (2018) and MSc (2020) degrees in Electronic Engineering from the National Technological Institute of Mexico/Chihuahua Institute of Technology. He is currently a PhD candidate at the National Technological Institute of Mexico/Chihuahua Institute of Technology at the Visual Perception on Robotic Applications Laboratory. His research interests include meta-learning, bio-inspired algorithms, image and signal processing and computational intelligence.

d20060684@chihuahua.tecnm.mx, ORCID: 0000-0002-0135-4682



Mario I. Chacon-Murguia, received the B.Sc. (1982) and the M.Sc. (1985) degrees in Electrical Engineering from the Chihuahua Institute of Technology, Chihuahua, Mexico, and the Ph.D. (1998) in Electrical Engineering from New Mexico State University, USA. He is a Professor at the Chihuahua Institute of Technology, director of the Visual Perception Laboratory. He has published more than 218 works and published 3 books. He has directed 45 research projects for the industry and government institutions. His current research includes visual perception, image and signal processing using computational intelligence. He is an IEEE Senior member, and a member of the Mexican Research System.

mario.cm@chihuahua.tecnm.mx, ORCID: 0000-0002-5382-9424



Juan Alberto Ramirez-Quintana, received the B.Sc., M.Sc., and Ph.D. degrees in EE from the Chihuahua Institute of Technology, Chihuahua, Mexico, in 2004, 2007, and 2014, respectively. From 2008 to 2011, he was a researcher and a teaching assistant with different universities. He is currently a Research Professor at the Tecnológico Nacional de México campus I.T. Chihuahua, and the Director of Digital Signal and Artificial Intelligence Lab. His current research includes computer vision, signal processing, computational intelligence and machine learning. He is a member of the National Research System in Mexico.

juan.rq@chihuahua.tecnm.mx, ORCID 0000-0003-4445-6555



Carlos Arzate Quintana, received his bachelor's degree in 2001 as a Bacteriologist and Parasitologist Chemist at the Faculty of Chemistry of the Autonomous University of Chihuahua (UACH); completed a master's degree in immuno-biological sciences at the Faculty of Biology of the Autonomous University of Nuevo León (UANL) in 2004; obtained his PhD in Environmental Science and Technology at the Investigation Center for Advanced Materials (CIMAV) in the city of Chihuahua in 2016. He currently works as a professor and researcher at the Faculty of Medicine and Biomedical Sciences (UACH) and is the lab head of the Biomaterials laboratory. He has published several scientific articles on antimicrobial biomaterials and also works on projects for the automation of clinical diagnosis. Currently has the level 1 distinction of the national system of researchers of Mexico.

carzate@uach.mx, ORCID ID: 0000-0002-5673-5992