Exploring Deep Convolutional Neural Networks as Feature Extractors for Cell Detection

Bruno C. Gregório da Silva and Ricardo J. Ferrari(✉)

Department of Computing, Federal University of São Carlos,
São Carlos, SP 13565-905, Brazil
{bruno.gregorio,rferrari}@ufscar.br
https://www.bipgroup.dc.ufscar.br

Abstract. Among different biological studies, the analysis of leukocyte recruitment is fundamental for the comprehension of immunological diseases. The task of detecting and counting cells in these studies is, however, commonly performed by visual analysis. Although many machine learning techniques have been successfully applied to cell detection, they still rely on domain knowledge, demanding high expertise to create handcrafted features capable of describing the object of interest. In this study, we explored the idea of transfer learning by using pre-trained deep convolutional neural networks (DCNN) as feature extractors for leukocytes detection. We tested several DCNN models trained on the ImageNet dataset in six different videos of mice organs from intravital video microscopy. To evaluate our extracted image features, we used the multiple template matching technique in various scenarios. Our results showed an average increase of 5.5% in the F_1-score values when compared with the traditional application of template matching using only the original image information. Code is available at: https://github.com/brunoggregorio/DCNN-feature-extraction.

Keywords: Transfer learning · Cell detection · Feature extraction · Convolutional neural network · Leukocyte recruitment · Intravital video microscopy

1 Introduction

One of the countless applications of automated image analysis involves cell detection in biological experiments. The automatic detection and counting of leukocytes in the microcirculation of living small animals, for instance, can help in the comprehension of immunological mechanisms from inflammatory processes. As a consequence, researchers can develop new drugs and therapeutic strategies to fight several diseases such as multiple sclerosis, atherosclerosis, ischemia-reperfusion injury, rheumatoid arthritis, and cancer [12]. However, this kind of analysis, which is typically done using intravital video microscopy (IVM),

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12250, pp. 91–103, 2020.
https://doi.org/10.1007/978-3-030-58802-1_7

becomes an arduous and error-prone task since it is performed by visual observation of the cellular traffic.

Different machine learning methods have been proposed in the last few years to overcome this problem. They are often designed for a particular set of images and tested on private datasets using different evaluation metrics [2]. One common approach is the use of shape information in active contours and gradient vector flow for both leukocytes detection and tracking [5,22–24,26,34,35]. Other works use adaptive template matching [1,14], image-level sets [21], and Monte Carlo [4] technique.

In our previous works, we proposed two different methods for the leukocytes detection based on the local analysis of eigenvalues obtained from Hessian matrices [28,29], and second-order momentum matrices of the phase congruency technique [10,11]. Despite promising results (most above 0.75 for F_1-score measure), these approaches were developed to enhance and detect blob-like structures in IVM images from the central nervous system (CNS) of mice. Therefore, these methods mostly fail when either the cells have distinct appearances or different image acquisition protocols.

Although the works mentioned above have presented significant results, they still rely on domain/business knowledge, demanding high expertise to create handcrafted features capable of describing the object of interest.

In the last decades, the use of artificial neural networks (ANN) or, more specifically, the convolutional neural networks (CNN) have attracted considerable attention because of its ability to learn data representations automatically while dealing with raw data.

Egmont-Petersen et al. [9], for instance, applied an ANN in IVM studies of the mesentery of mice and rats. In their work, they compared the application of an ANN using two training datasets collected from real and synthetic images of cells. Eden et al. [8] also resorted to the use of ANNs for the detection and tracking of leukocytes in IVM. The proposed cell detection approach started using a motion detection algorithm based on the image background subtraction. After this rough detection, they selected only the cells inside the vessel region and used an ANN for the classification of a sub-region as a target (cell) or a non-target, which have afterward their points analyzed by a clustering strategy. The use of these shallow ANN models, however, may not represent complex features, resulting in a low level of generalization and a weak learning of data representations.

In order to have a CNN model with a high level of generalization and without overfitting, a high number of images with labeled objects is required for training it properly. As this condition is not always satisfied, other options should be considered.

It is well-known that the first layers of deep CNNs (DCNN) trained on natural images learn more general features that can be similar to Gabor filters and color blobs [36]. This important statement suggests we can use the output of these layers as feature extractors in a process called transfer learning. Transfer learning is a popular approach in deep learning where a network developed for a specific

task may have its weights from early layers used as a feature extractor or as a starting point for training a new model and adapted in response to a new problem. This procedure can exploit the generalization of a previously well-trained architecture in another model setting.

In this study, we explore the transfer learning approach by using different models trained on the ImageNet dataset [25] as feature extractors. The resulting feature maps are then selected and used as input for a multiple template matching (MTM) method. Our results show that features extracted from kernels trained in a different task can increase the performance of the well know template matching technique.

The rest of this paper is organized as follows: in Sect. 2, we describe our methodology and the database used in this work. Results and discussions are presented in Sect. 3, while in the last section, we make our final considerations.

2 Materials and Methods

In this section, we first describe the IVM dataset used in this work and then elaborate on the techniques applied for leukocytes detection and the metrics used to evaluate them.

2.1 IVM Dataset

To evaluate our approach, we used six videos from IVM studies with 705 frames in total that were obtained from distinct image acquisition protocols and four different animal organs: brain, spinal cord, cremaster muscle, and mesentery of mice. Figure 1 shows examples of frames from each one of the videos.

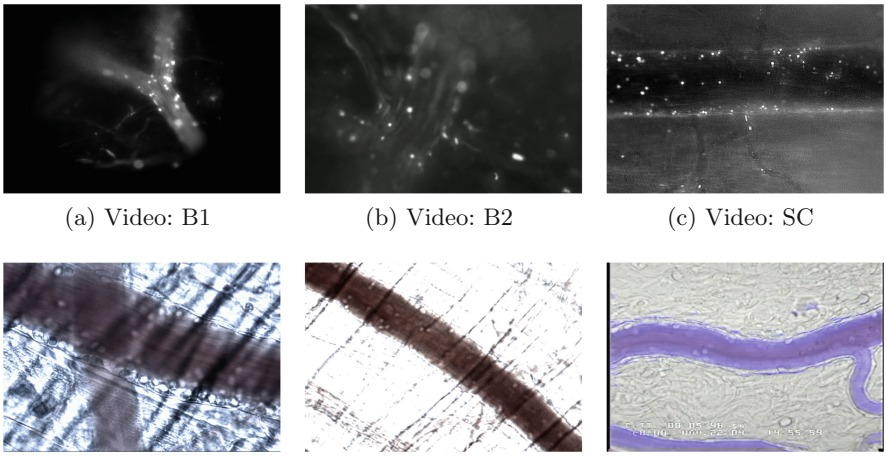
In all videos, the leukocytes were frame-by-frame manually annotated by an expert. All information necessary to describe our dataset are presented in Table 1.

Although some of these videos have a relatively small number of analyzed frames, the total number of manually annotated leukocytes is quite large (see values in Table 1), providing enough data for a proper quantitative evaluation. For more information about the experimental procedures, please refer to our previous works [28,30] and the works of our collaborators Prof. Juliana Carvalho Tavares, Ph.D.[1] [6,7] and Prof. Mônica Lopes-Ferreira, Ph.D.[2] [27].

All the images in our dataset went through the following ordered sequence of processes to 1) remove the extremely blurred images, 2) noise reduction, 3) contrast standardization, 4) video stabilization, and 5) extraction of the region of interest (ROI). The application of these preprocessing techniques is better described in our previous work [28] and is out off the scope of this paper.

[1] Department of Physiology and Biophysics, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil.

[2] Special Laboratory of Applied Toxinology (Center of Toxins Immune-Response and Cell Signaling), Butantan Institute, São Paulo, Brazil.



(a) Video: B1 (b) Video: B2 (c) Video: SC

(d) Video: C1 (e) Video: C2 (f) Video: ME

Fig. 1. Examples of video frames from different mice organs.

Table 1. Description of the IVM dataset used in this work. B1 and B2: videos from the mice brain; SC: video from the mice spinal cord; C1 and C2: videos from cremaster muscle; ME: video from the mesentery.

Category	ID	Matrix size (px)	Color	# frames analyzed	# annotated leukocytes
CNS	B1	692 × 520	No	220	5827
	B2	460 × 344	No	401	8048
	SC	692 × 520	No	21	1570
			Subtotal:	*642*	*15445*
Cremaster	C1	1392 × 1040	Yes	21	390
Muscle	C2	1392 × 1040	Yes	21	1603
			Subtotal:	*42*	*1993*
Mesentery	ME	720 × 480	Yes	21	291
			Subtotal:	*21*	*291*
				705	**17729**

2.2 Pipeline Overview

Figure 2 illustrates the pipeline of our proposed method for leukocytes detection. The main goal of our approach is to apply the concept of transfer learning using different pre-trained DCNNs and to test the model's genericity when applied for distinct targets. For that, we chose a broad list of models already trained on the ImageNet dataset [25] and used the output of their first convolutional layers in our problem, i.e., in a task entirely different from the original.

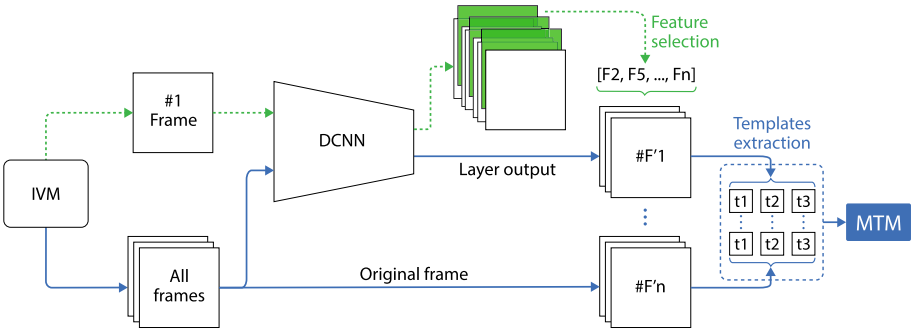


Fig. 2. Pipeline overview of the approach used in this work. (Color figure online)

Conventional DCNN models generally have a small input shape value as default. In this study, however, we decided to rescale our input images into the fixed range of 1400×1000 as the massive information contained in this kind of image is composed by small objects that are quite significant for cellular morphology characterization.

With all the images preprocessed and rescaled, we start our detection pipeline by extracting the first frame of each video and passing it forward into the DCNN model until the selected layer. In this work, each selected layer was chosen by visual inspection of its output feature images. Since our image frames present relevant information in small regions, we decided to analyze only the first convolutional layers (shallow layers) of each DCNN.

As a consequence of transfer learning, not all output images present relevant characteristics that could help in a detection process. For this reason, we performed a feature image selection capable of separating only the best set of features to be used next. To accomplish that, we extracted a small ROI previously selected and used it as a template for the template matching technique. We then get the corresponding output maps for each feature image and applied a thresholding technique on each one of then. In this case, the threshold value was set to 0.9, which results in a map of detection candidates with a high probability of being indeed cells. The accuracy of each resulting map was evaluated following the metrics described in Sect. 2.4, but to chose the best set of features, we sorted and normalized the evaluation results in order to selected only those top features whose accumulated value (or retained score) was higher than 0.1. Fig. 3 illustrates the process of feature selection.

After the first passage through our pipeline, illustrated by the green dashed arrows in Fig. 2, we have our best set of image features from the DCNN layer and can now apply our approach to all the video frames. The blue arrows in Fig. 2 show the remaining steps in our pipeline. They are similar to the previous steps, except that we now know what the best feature set is and can finally apply the MTM algorithm to identify the cell candidates. At this step, we also include the original frame image into the vector of selected features. Next in our approach,

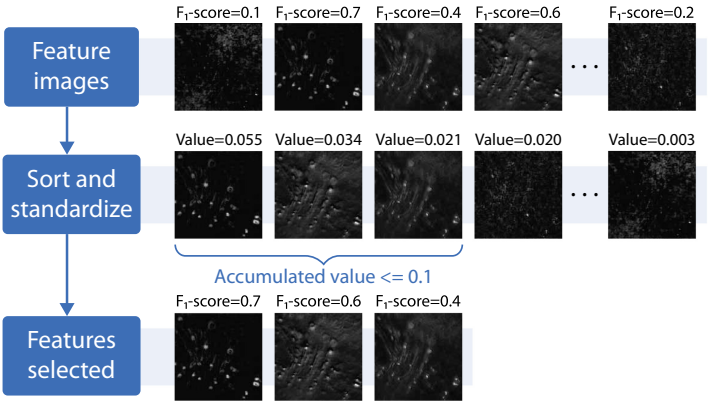


Fig. 3. Steps for feature selection.

we extracted three pre-selected ROIs from the first frames of each video to be used as template images in the entire processing.

2.3 Multiple Template Matching – MTM

As stated before, to perform our cell detection step and consequently test our proposed approach, we used the template matching (TM) technique [13,19,20] with multiple templates. The normalized cross-correlation (NCC) based TM is an algorithm of pattern recognition field that performs the detection of similar objects in an image $I(x, y)$, holding as input the image itself and a template (sub-image) $T(x, y)$ to be detected. The TM algorithm used in this work has as similarity measure the NCC coefficient, computed as:

$$\rho(x,y) = \frac{\sum_r \sum_s \left[T(r,s) - \overline{T}\right] \cdot \left[I\left(x+r, y+s\right) - \overline{I}_T\right]}{\sqrt{\sum_r \sum_s \left[T(r,s) - \overline{T}\right]^2 \cdot \sum_r \sum_s \left[I\left(x+r, y+s\right) - \overline{I}_T\right]^2}}, \tag{1}$$

where $\overline{T}$ is the average value of pixel intensities in $T(x, y)$, $\overline{I}_T$ is the average value of I in the coincident region with the current position of T, and the sums are only realized over the common coordinates of $I(x, y)$ and $T(x, y)$, delimited by the variables r and s of the summations.

The coefficient of correlation ρ indicates the level of similarity between the template T and the current image region. Its scale varies in the range of $[-1, 1]$ and is, therefore, normalized by the amplitudes of T and I, wherein $\rho = 1$ means the total correlation between T and the sub-region of I, $\rho = 0$ means that there is no correlation, and $\rho = -1$ means inverse correlation.

As already stated at the beginning of this section, our approach used a set o leukocyte-templates as input for the MTM algorithm. As a result, we have intensity maps with the highest coefficient values indicating the spatial locations

in the video frames with high similarity with our selected templates. In the cases where the number of templates is higher than one, a fusion step is employed by summing and normalizing all the MTM output maps. Finally, the resulting maps were thresholded by using different values. These values were defined in the range of [0.7, 0.95], with step of 0.5.

2.4 Metrics of Evaluation

For the evaluation of our approach, the resulting thresholded maps were assessed by comparing the spatial coordinates of the leukocytes' centroids that were manually identified and annotated by an expert (ground truth) with those automatically detected. In this sense, we defined the detection as true when the distance between a manually annotated centroid and an automatically detect one is less or equal than k pixels. This distance value k was estimated according to the average radius of the selected templates for each video.

Accordingly, in this study we defined the number of true positives (TPs) as the accumulated amount of leukocyte positions that were correctly detected by the algorithm, the false positives (FPs) as the accumulated amount of leukocytes automatically detected without correspondence to those manually annotated, and the false negatives (FNs) as the accumulated amount of leukocytes that the algorithm could not identify.

The measures of precision (P), recall (R), and F_1-score [15] were then used to evaluate the overall performance of our proposed approach. These measures are based on the accumulated TPs, FPs, and FNs over the sequence of video frames. They are defined as follows:

$$P = \frac{TPs}{(TPs + FPs)}, \tag{2}$$

$$R = \frac{TPs}{(TPs + FNs)}, \tag{3}$$

$$F_1\text{-}score = 2 \times \frac{P \times R}{P + R}. \tag{4}$$

3 Experimental Results and Discussions

The proposed approach for leukocytes detection was quantitatively evaluated using all the videos detailed in Sect. 2.1 for different pre-trained DCNN models. For the sake of comparison, we also applied the MTM directly to the video frames at gray level intensities, i.e., not passing through any DCNN model.

In Fig. 4, we can see the resulting F_1-score values for all the videos and models processed. For each one of them, we plotted the best values found in our experiments, considering the set of threshold values tested. Each model exhibits three different bars, colored according to the number of templates used in our experiments with the MTM algorithm.

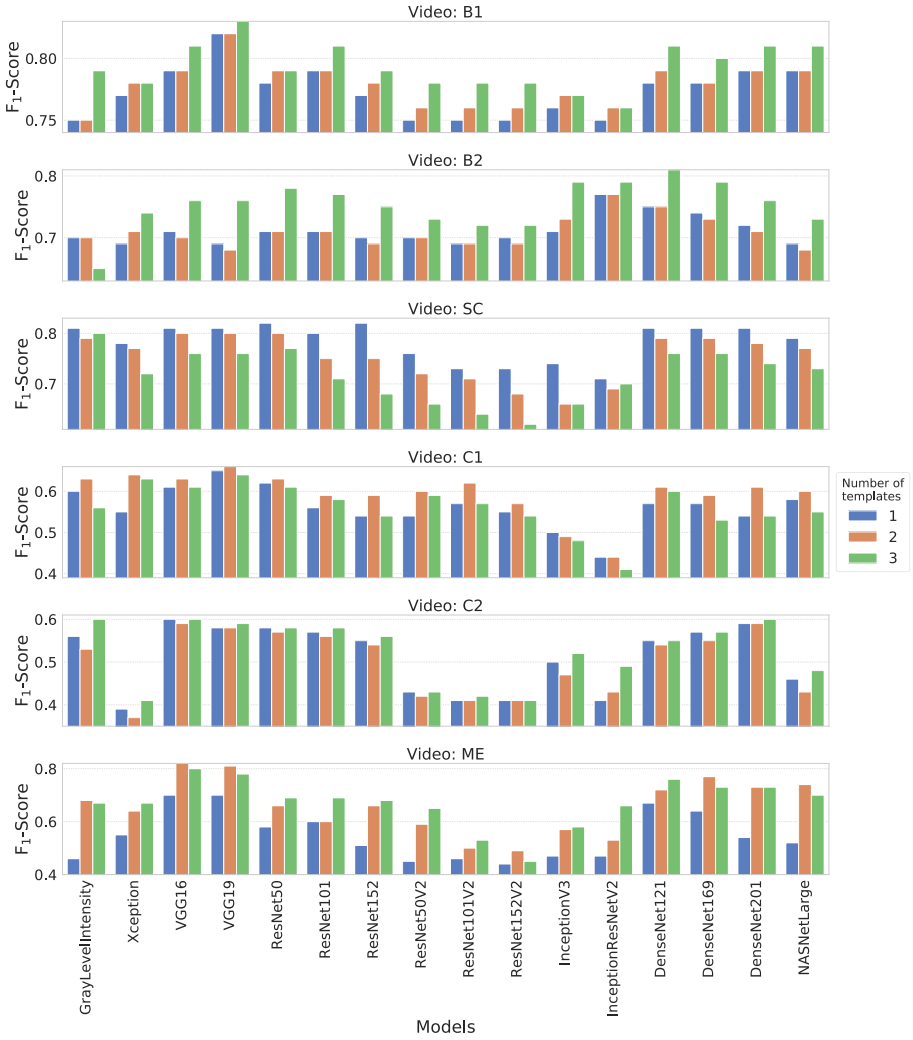


Fig. 4. Output values for each video and DCNN model tested. The colored bar indicates the number of templates used in MTM algorithm.

The plots in Fig. 4 clearly show the use of selected convolutional layers in CNNs positively contribute as generic feature extractors, even coming from models trained for a completely different task. It is also worth noticing that the use of multiple templates can help to recognize targets that are slightly different, like those present in videos B1 and B2.

Although some models did not contribute to the MTM technique in most cases, such as the ResNet50V2, ResNet101V2, ResNet152V2, InceptionV3, and InceptionResNetV2, the majority of them achieved values similar or higher than

the application of the MTM using only the gray level intensity information, which also shows the potential of this approach.

Table 2 shows the best set of values found in Fig. 4 for a better quantitative comparison of the methods. Indeed, we can observe that only in the C2 video, our strategy presented the same value as in the gray level image application, while in the remaining results, we had a considerable improvement (up to 14%). Videos C1 and C2, however, still exhibited low F_1-score values (0.66 and 0.60, respectively), which is justifiable since they have the most challenging visual aspects, with a cluttered background and cell sizes in the order of 5 pixels.

Table 2. Best F_1-score values found for each video and DCNN model.

Model	B1	B2	SC	C1	C2	ME
GrayLevelIntensity	0.79	0.70	0.81	0.63	**0.60**	0.68
Xception [3]	0.78	0.74	0.78	0.64	0.41	0.67
VGG16 [31]	0.81	0.76	0.81	0.63	**0.60**	**0.82**
VGG19 [31]	**0.83**	0.76	0.81	**0.66**	0.59	0.81
ResNet50 [16]	0.79	0.78	**0.82**	0.63	0.58	0.69
ResNet101 [16]	0.81	0.77	0.80	0.59	0.58	0.69
ResNet152 [16]	0.79	0.75	0.82	0.59	0.56	0.68
ResNet50V2 [17]	0.78	0.73	0.76	0.60	0.43	0.65
ResNet101V2 [17]	0.78	0.72	0.73	0.62	0.42	0.53
ResNet152V2 [17]	0.78	0.72	0.73	0.57	0.41	0.49
InceptionV3 [33]	0.77	0.79	0.74	0.50	0.52	0.58
InceptionResNetV2 [32]	0.76	0.79	0.71	0.44	0.49	0.66
DenseNet121 [18]	0.81	**0.81**	0.81	0.61	0.55	0.76
DenseNet169 [18]	0.80	0.79	0.81	0.59	0.57	0.77
DenseNet201 [18]	0.81	0.76	0.81	0.61	0.60	0.73
NASNetLarge [37]	0.81	0.73	0.79	0.60	0.48	0.74

Examples of output frames for each processed video are shown in Fig. 5. Each TP point found is illustrated by the green circles in the images, with its respective manual centroid annotation indicated as a cross. The blue circles represent the FP points, while the red squares are the FN ones.

From the images in Fig. 5, we observe that FN points are often the cells very close to each other or the ones whose appearance is quite different from the rest of them. The FP points, however, mostly correspond to bright regions in the images or to the erythrocytes, which are smaller cells that appear as bright blurred points in non-consecutive frames and are not part of the manual annotations. Even so, the final results were quite promising and indicated that pre-trained DCNN models could be a good option for generic feature extraction in IVM.

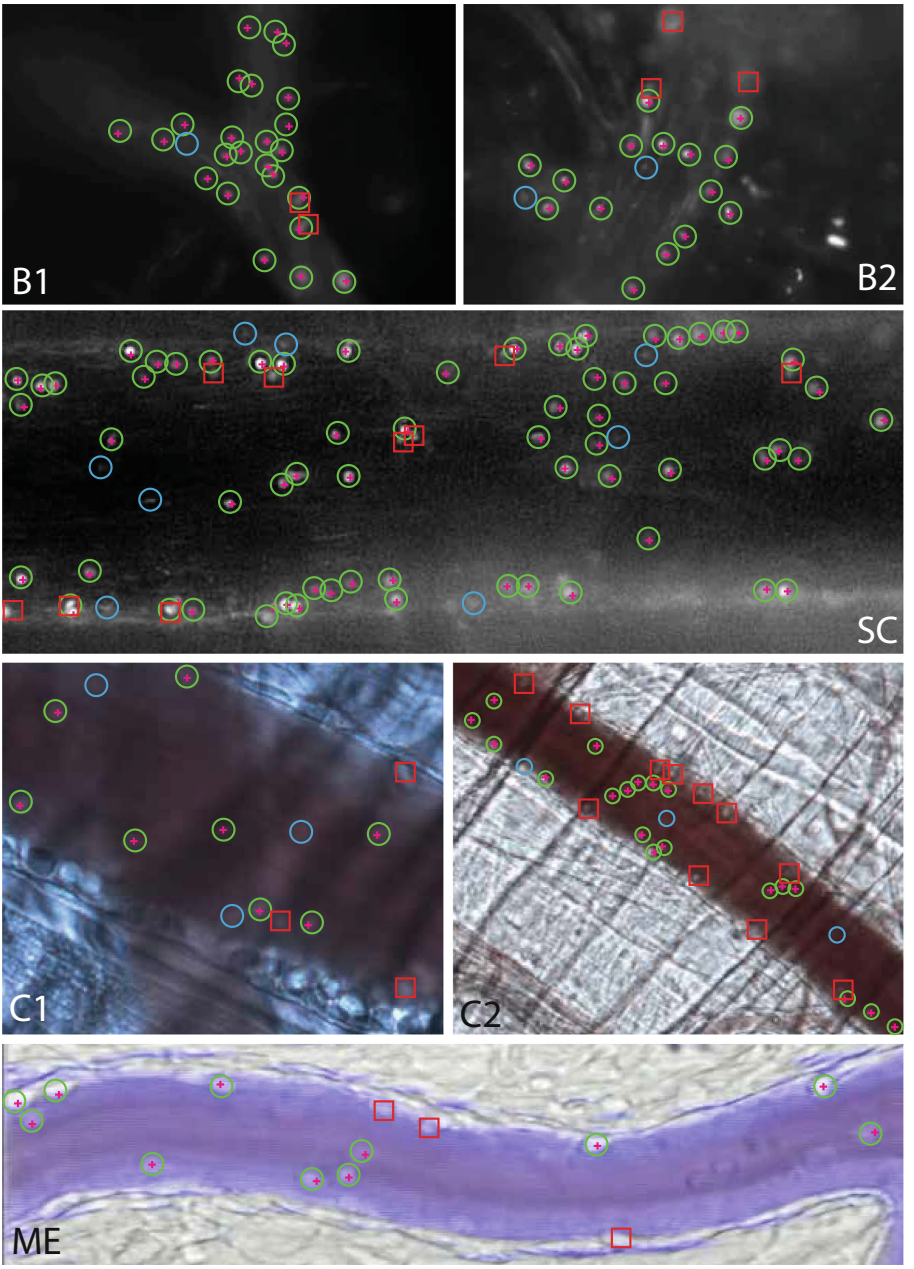


Fig. 5. Examples of MTM outputs for each video in the dataset. Green circles represent the TP points, blue circles the FP points, and red squares the FN points. (Color figure online)

4 Conclusions

Manual detection and counting of leukocytes are still fundamental tasks in IVM studies. However, visual analysis is error-prone and can generate false statistics and wrong interpretations, depending on the professional experience and expertise in the field. In this paper, we explored the transfer learning idea by using pre-trained DCNNs for image feature extraction in the application of cell detection in IVM. In order to create reliable image features, we selected the first convolutional layers of DCNNs pre-trained on ImageNet and tested them as the input for the multiple template matching algorithm.

We conducted several experiments using six different videos from IVM with one, two, and three template images. Our results showed a considerable improvement in the template matching performance, even using generic image features extracted by DCNN models trained for distinct tasks and targets. Despite the great results, in the cases where small cells are predominant, we observed some models do not produce relevant information to perform a robust detection process. However, we proved that pre-trained DCNN models could, indeed, provide generic feature images to be used in different applications.

Our future works include tests with models trained in different datasets and a detection process using different strategies for feature selection and fusion in an adaptive multiple threshold technique.

Code is available at GitHub[3].

Acknowledgments. We would like to thank our collaborators Prof. Juliana Carvalho-Tavares, Ph.D. and Prof. Mônica Lopes-Ferreira, Ph.D. for conducting the biological experiments.

Funding Statement. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (grant numbers 2013/26171-6 and 2018/08826-9).

References

1. Acton, S.T., Wethmar, K., Ley, K.: Automatic tracking of rolling leukocytes in vivo. Microvasc. Res. **63**(1), 139–148 (2002)
2. Akram, S.U., Kannala, J., Eklund, L., Heikkilä, J.: Cell tracking via proposal generation and selection. CoRR abs/1705.03386 (2017). http://arxiv.org/abs/1705.03386
3. Chollet, F.: Xception: Deep learning with depthwise separable convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1800–1807. IEEE, Honolulu (2017). https://doi.org/10.1109/CVPR.2017.195
4. Cui, J., Acton, S.T., Lin, Z.: A Monte Carlo approach to rolling leukocyte tracking in vivo. Med. Image Anal. **10**(4), 598–610 (2006)

[3] https://github.com/brunoggregorio/DCNN-feature-extraction.

5. Dong, G., Ray, N., Acton, S.T.: Intravital leukocyte detection using the gradient inverse coefficient of variation. IEEE Trans. Med. Imaging **24**(7), 910–924 (2005)
6. Dos Santos, A.C., Barsante, M.M., Arantes, R.M.E., Bernard, C., Teixeira, M.M., Carvalho-Tavares, J.: CCL2 and CCL5 mediate leukocyte adhesion in experimental autoimmune encephalomyelitis an intravital microscopy study. J. Neuroimmunol. **162**(1–2), 122–129 (2005)
7. Dos Santos, A.C., et al.: Kinin B2 receptor regulates chemokines CCL2 and CCL5 expression and modulates leukocyte recruitment and pathology in experimental autoimmune encephalomyelitis (EAE) in mice. J. Neuroinflammation **5**, 49–58 (2008)
8. Eden, E., Waisman, D., Rudzsky, M., Bitterman, H., Brod, V., Rivlin, E.: An automated method for analysis of flow characteristics of circulating particles from in vivo video microscopy. IEEE Trans. Med. Imaging **12**(8), 1011–1024 (2005)
9. Egmont-Petersen, M., Schreiner, U., Tromp, S.C., Lehmann, T.M., Slaaf, D.W., Arts, T.: Detection of leukocytes in contact with the vessel wall from in vivo microscope recordings using a neural network. IEEE Trans. Biomed. Eng. **47**(7), 941–951 (2000)
10. Elisa de Souza, K., Gregório da Silva, B.C., Carvalho-Tavares, J., Ferrari, R.J.: Automatic detection of leukocytes from intravital video microscopy using the phase congruency technique. In: Proceedigns of XI Workshop de Visão Computacional (WVC), pp. 387–391. BDBComp, São Carlos (2015)
11. Elisa de Souza, K., Gregório da Silva, B.C., Carvalho-Tavares, J., Ferrari, R.J.: Detection of leukocytes in intravital microscopy video images using the phase congruency technique. Revista de Informática Teórica e Aplicada **23**(2), 33–55 (2016)
12. Gavins, F.N.E.: Intravital microscopy: new insights into cellular interactions. Curr. Opinion Pharmacol. **12**(5), 601–607 (2012)
13. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Addison-Wesley, Boston (1992)
14. Goobic, A.P., Welser, M.E., Acton, S.T., Ley, K.: Biomedical application of target tracking in clutter. In: Conference on Signals, Systems and Computers, vol. 1, pp. 88–92. IEEE, Pacific Grove (2001)
15. Goutte, C., Gaussier, E.: A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In: Losada, D.E., Fernández-Luna, J.M. (eds.) ECIR 2005. LNCS, vol. 3408, pp. 345–359. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-31865-1_25
16. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. IEEE, Las Vegas (2016). https://doi.org/10.1109/CVPR.2016.90
17. He, K., Zhang, X., Ren, S., Sun, J.: Identity mappings in deep residual networks. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9908, pp. 630–645. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46493-0_38
18. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4700–4708. IEEE (2017)
19. Khosravi, M., Schafer, R.W.: Template matching based on a grayscale hit-or-miss transform. IEEE Trans. Image Process. **5**(6), 1060–1066 (1996)
20. Lewis, J.P.: Fast template matching. In: Vision Interface, vol. 95. pp. 120–123. Canadian Image Processing and Pattern Recognition Society, Quebec City, Canada (1995)
21. Mukherjee, D.P., Ray, N., Acton, S.T.: Level set analysis for leukocyte detection and tracking. IEEE Trans. Image Process. **13**(4), 562–572 (2004)

22. Ray, N.: A concave cost formulation for parametric curve fitting: detection of leukocytes from intravital microscopy images. In: Proceedings of the International Conference on Image Processing, pp. 53–56. IEEE, Hong Kong (2010)
23. Ray, N., Acton, S.T.: Motion gradient vector flow: an external force for tracking rolling leukocytes with shape and size constrained active contours. IEEE Trans. Med. Imaging **23**(12), 1466–1478 (2004)
24. Ray, N., Acton, S.T., Ley, K.: Tracking leukocytes in vivo with shape and size constrained active contours. IEEE Trans. Med. Imaging **21**(10), 1222–1235 (2002)
25. Russakovsky, O., et al.: ImageNet large scale visual recognition challenge. Int. J. Comput. Vis. **115**(3), 211–252 (2015). https://doi.org/10.1007/s11263-015-0816-y
26. Sahoo, S., Ray, N., Acton, S.T.: Rolling leukocyte detection based on teardrop shape and the gradient inverse coefficient of variation. In: International Conference on Medical Information Visualisation, pp. 29–33. IEEE Computer Society, London (2006)
27. dos Santos, J.C., et al.: Stingray venom activates IL-33 producing cardiomyocytes, but not mast cell, to promote acute neutrophil-mediated injury. Sci. Rep. **7**(7912), 2045–2322 (2017). https://doi.org/10.1038/s41598-017-08395-y
28. Gregório da Silva, B.C., Carvalho-Tavares, J., Ferrari, R.J.: Detection of leukocytes in intravital video microscopy based on the analysis of Hessian matrix eigenvalues. In: 28th Conference on Graphics, Patterns and Images, pp. 345–352. IEEE, Salvador (2015). https://doi.org/10.1109/SIBGRAPI.2015.48
29. Gregório da Silva, B.C., Carvalho-Tavares, J., Ferrari, R.J.: Detecting and tracking leukocytes in intravital video microscopy using a Hessian-based spatiotemporal approach. Multidimens. Syst. Sig. Process. **30**(2), 815–839 (2018). https://doi.org/10.1007/s11045-018-0581-5
30. Gregório da Silva, B.C., Freire, P.G.L., Mello, R.F., Bernardes, D., Carvalho-Tavares, J., Ferrari, R.J.: Técnica de estabilização de movimento em microscopia intravital utilizando métodos de co-registro de imagens. In: XXIV Congresso Brasileiro de Engenharia Biomédica (CBEB), pp. 193–196. CBEB, Uberlândia, MG, Brazil (2014)
31. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognitions. In: International Conference on Learning Representations (ICLR) (2015)
32. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-first AAAI Conference on Artificial Intelligence (2017)
33. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818–2826. IEEE, Las Vegas (2016). https://doi.org/10.1109/CVPR.2016.308
34. Xu, C., Prince, J.L.: Generalized gradient vector flow external forces for active contours. Sig. Process. **71**(2), 131–139 (1998)
35. Xu, C., Prince, J.L.: Snakes, shapes, and gradient vector flow. IEEE Trans. Image Process. **7**(3), 359–369 (1998)
36. Yosinski, J., Clune, J., Bengio, Y., Lipson, H.: How transferable are features in deep neural networks? In: Advances in Neural Information Processing Systems 27, vol. 2, pp. 3320–3328. Curran Associates Inc, Montreal (2014)
37. Zoph, B., Vasudevan, V., Shlens, J., Le, Q.V.: Learning transferable architectures for scalable image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8697–8710. IEEE (2018)