AN2DL - First Homework Report
MiLi

Martim Bento, Henrique Pocinho, Salvatore Santolupo, Diego Viganó

martimrbento, pocinho, salvos, diegosdigos

11068539, 11068532, 10766335, 10739933

November 24, 2024

1 Introduction

Blood cell identification is of utmost importance in hematological studies[1]. When done correctly it can help detect several diseases, like anemia and malaria.[2–4]. This work is traditionally done by visual inspection of the cells under microscope by a professional. [5] However this approach is a time consuming manual task and requires skilled individuals, who bring their bias when judging, therefore different people may produce different results.[6] As such, recently there have been several different proposals for the automation of this process. [7–10].

In this context, the first homework for the course AN2DL asked us to develop a model to classify images of blood cells into eight classes. As such, the problem in question is of **multi-class classification**, common within the supervised learning paradigm, where, *given a set of desired outputs, the goal is to learn how to produce the correct output when presented with new inputs.*

Thus the approach to this problem was twofold. The main strategy detailed in this report is the use of **transfer learning**, *a technique where features already learned from one problem are applied to a new, similar problem.* This was done by using the feature extraction networks (**FEN's**) of already trained models. Simultaneously techniques and concepts such as **fine tuning**, **weight classes**, **data augmentation** and some others were implemented and used, aiming to increase accuracy, help with training and to reduce the risk of overfitting.

The alternative approach was the development of a custom **convolutional neural network** (CNN), which will also be mentioned.

2 Problem Analysis

The main challenge presented is to develop and train a network whose knowledge can be generalized to a test set with unknown characteristics and whose images might not be inline with the given dataset, being that they could present more noise or purposefully added alterations.

In order to do so a dataset of 13759 96x96 RGB labeled images was provided by the course faculty. An analysis of the dataset showed that out of the 13759 images, 1800 where outliers, as shown in Fig. 1. Out of the 11959 non-outlier images, 8 were repeated. Thus the repeated images and outliers were removed, and a sub-set of 11951 unique non-outlier images was obtained. From this point onward it was assumed that those 11951 images are correctly identified and labeled.

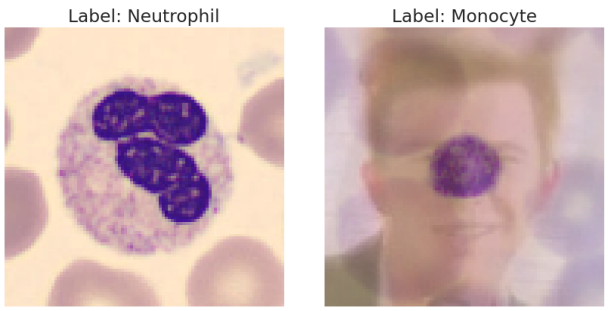

Figure 1: Examples of images taken from the dataset. On the left a cell, on the right an outlier, the known english pop-star Rick Astley[11].

For the initial conditions, the seed was fixed for all the random events, to guarantee repeatability, and the number 42 was used since it is *"the answer*

to the ultimate question of life, the universe, and everything." [12]

3 Method

The main approach consisted in transfer learning. The final model presented in Fig. 2, is composed by a FEN from EfficientNetB0, with a new classifier, which includes a global average polling (GAP) layer and a dense output layer, with softmax activation.

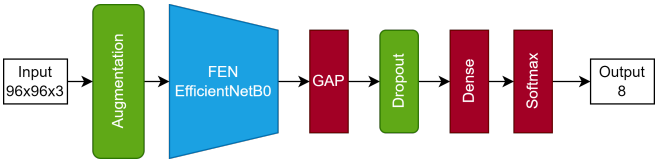

Figure 2: Final developed model. In green the layers only active during training. In blue the FEN used. In red the layers that compose the classifier.

3.1 Feature Extraction Network

Initially, MobileNetV3Small was chosen due to its small size. Thus, the first new classifier consisted in a dense output layer with a softmax activation function. The architecture was iterated on, by altering the classifier and by using other FENs. After further reading and testing, the aforementioned solution, in Fig. 2 was reached. A softmax activation function was chosen since the output of such a layer can be easily interpreted as a probability, thus the higher the output value, the higher the likelihood that a given image is of that given class.

The solution choice is motivated by empirical results obtained with the submission on Codabench and can be justified by the fact that the EfficientNet models outperform other state-of-the-art deep learning models in terms of Top-1 accuracy, and EfficientNetB0 has scaling of both depth and width, reduced computational complexity and uses residual blocks with skip connections to enhance overall classification accuracy [13, 14].

3.2 Training

The dataset was divided into three subsets: A training set, a validation set and a test set. The division was done as 85%, 10% and 5%, respectively. The test set acts as an holdout set, for a final local evaluation and validation, before submission. Training was done in two phases. The first only trained the classifier where all the layers of the FEN were frozen, as its purpose is to extract features, for which it is already trained. Next is the fine-tuning, where some of the FEN's *convolutional layers* were unfrozen, therefore training the feature extracting capabilities of those layers specifically for analysis of blood cell images. This should be done cautiously as it can easily lead to the model overfitting.

During both phases, early stopping was implemented, with patience of 10 and 3, respectively. Many optimizers were tested, including Adam, Nadam, AdamW and Lion. For the training of the final model, Lion was used, during both phases of training, as it provided the best results. The learning rates were also tuned to enhance the performance during training and it is also important to mention that, as shown in Fig. 2, a dropout layer was used, to help regularize the model and prevent overfitting. An attempt with the reduction of the learning rate on plateau of the validation loss was also done. However, this approach was quickly discarded as no improvement was detected. An experiment was done to test focal loss, a modified kind of cross entropy, designed to have better performance when using an imbalanced dataset, as the one used. Class weights were also tested to give more importance during training to the underrepresented classes. After comparing results, cross entropy loss with class weights performed slightly better and as such was used in the final model. To help accelerate the training and produce smaller models, mixed precision was initially used, but it generated conflicts with the augmentation layers from Keras-CV, so it was later dropped. A small price to pay for salvation.

3.3 Data Augmentation

Several augmentation pipelines with different layers were tested and the final chosen augmentation pipeline uses random flip, translation, rotation and zoom layers, from Keras, and a random augmentation and CutMix from Keras-CV. The results of the pipeline can be observed in Fig. 3. The layers' parameters were also tuned to prevent the augmentations from being too intense and making training impossible. Note that EfficientNetB0 also includes its own input preprocessing.

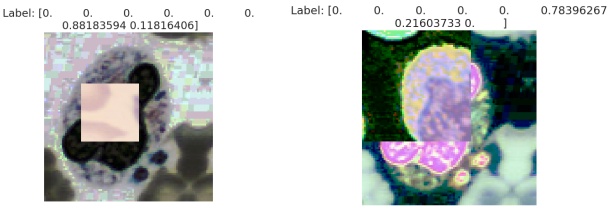

Figure 3: Example of augmented images. Note that Cut Mix also changes the labels so that there is a percentage of each cut image.

Table 1: Results obtained for several tested models.

Model	Accuracy [%]	Precision [%]	Recall [%]	Codabench Accuracy [%]
1. MobileNetV3Small	90.89	91.19	90.89	43
2. VGG16	85.79	88.12	85.79	23
3. EfficientNetV2B0	88.63	90.25	88.63	73
4. Custom CNN	80.60	82.52	80.60	66
5. EfficientNetB0	**94.65**	**94.93**	**94.65**	**81**
6. ENB0 + Weight Decay	92.14	92.43	92.14	80

4 Experiments

Several different models with varying configurations were tested, as detailed in the previous section. Some notable results are displayed, in table 1.

The first model was the first submitted attempt. It is transfer learning based, using MobileNetV3Small FEN with Adam and a very reduced augmentation pipeline. The VGG16 attempt changed the FEN and added fine-tuning.

One of the reasons for the low scores on the first few attempts was the augmentation pipeline. Some Keras-CV layers were quickly included, namely the CutMix. EfficientNetV2B0 was the first network to be tested with the more robust augmentation pipeline. It still used the Adam optimizer.

Further tweaks were made, by changing the FEN to EfficientNetB0 and the optimizer to Lion. This allowed for the highest accuracy in the Codabench set, out of all attempts. The use of weight decays, to help prevent overfitting, was also tested, during the fine-tuning phase with recourse to the AdamW optimizer, but it didn't show improvements.

A different approach was also taken, in parallel, inspired by Microcell-Net, developed by Dwivedi et al. [15].The idea was to use Microcell-Net as a base and build on top of it, creating a custom CNN. The main improvements were the addition of normalization layers, squeeze and excite mechanisms and label smoothing. However, it did not manage to overcome the current best results.

5 Results and Discussion

The highlighted experiments clarify the effects of the various features that were tested during the development of the model. The comparisons between models 1 and 2 and between models 3 and 5 of table 1 show the significant effect of the choice of the FEN. Moreover, the comparison between the different networks corroborates the earlier choice, detailed in section 3, of the EfficientNetB0 network. Models 3 and 5 also show that, although EfficientNetV2B0 is a more recent version of EfficientNetB0, it was not the best solution to this specific problem.

Comparing the first two models with the rest shows the importance of using a robust augmentation pipeline. As previously stated, models 1 and 2 used very rudimentary augmentations. The difference in accuracies obtained in Codabench is very clear, showing that adequate augmentation can allow for a much better generalization.

Finally, models 5 and 6 differ in the use of the optimizer, as well as the inclusion of weight decay. While the differences in performance are minimal, it still corroborates the additional success derived from the use of the Lion optimizer.

6 Conclusions

The challenge was to develop an image classification model capable of classifying different types of blood cells. In this work, several models are presented, mainly through the use of transfer learning, with varying degrees of success. The most successful model managed to correctly identify 81% of the images in the final test, on Codabench.

As many different techniques were explored, some conclusions of their impact were derived. Mainly, the need of a robust augmentation pipeline and the choice of the right FEN, in this case EfficientNetB0. From the tests performed, the Lion optimizer also showed to outperform some others.

From this work, it is clear that transfer learning is a viable strategy for the challenge in question. As such, for future work, more pre-trained models can be tested, and a more fine hyper parameter tuning can be done, while also seeking to improve the custom CNN.

As a final note, we would like to comment that the work was distributed evenly among team members and, as such, everyone participated both in the development of the models and on writing the report.

References

[1] J. Alcaraz-Quiles, Á. Molina, J. Laguna, M. Rodríguez-García, G. Gutiérrez, J. Luis Bedini, and A. Merino, "Peripheral blood morphology review and diagnostic proficiency evaluation by a new spanish eqas during the period 2011-2019," *International journal of laboratory hematology*, vol. 43, no. 1, pp. 44–51, 2021.

[2] V. Jansen, "Diagnosis of anemia—a synoptic overview and practical approach," *Transfusion and Apheresis Science*, vol. 58, no. 4, pp. 375–385, 8 2019. [Online]. Available: http://dx.doi.org/10.1016/j.transci.2019.06.012

[3] T. S. Chy and M. A. Rahaman, "Automatic sickle cell anemia detection using image processing technique," 2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE). IEEE, 11 2018, pp. 1–4. [Online]. Available: http://dx.doi.org/10.1109/ICAEEE.2018.8642984

[4] Q. Sun, S. Yang, C. Sun, and W. Yang, "An automatic method for red blood cells detection in urine sediment micrograph," 2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC). IEEE, 5 2018, pp. 241–245. [Online]. Available: http://dx.doi.org/10.1109/YAC.2018.8406379

[5] M. Campos-Medina, A. Blumer, P. Kraus-Füreder, M. Mayrhofer-Reinhartshuber, P. Kainz, and J. A. Schmid, "Ai-enhanced blood cell recognition and analysis: Advancing traditional microscopy with the web-based platform ikosa," *Journal of Molecular Pathology*, vol. 5, no. 1, pp. 28–44, 1 2024. [Online]. Available: http://dx.doi.org/10.3390/jmp5010003

[6] B. Venkatalakshmi and K. Thilagavathi, "Automatic red blood cell counting using hough transform," 2013 IEEE Conference on Information & Communication Technologies (ICT). IEEE, 4 2013, pp. 267–271. [Online]. Available: http://dx.doi.org/10.1109/CICT.2013.6558103

[7] A. Acevedo, S. Alférez, A. Merino, L. Puigví, and J. Rodellar, "Recognition of peripheral blood cell images using convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 180, p. 105020, 10 2019. [Online]. Available: http://dx.doi.org/10.1016/j.cmpb.2019.105020

[8] W. Song, P. Huang, J. Wang, Y. Shen, J. Zhang, Z. Lu, D. Li, and D. Liu, "Red blood cell classification based on attention residual feature pyramid network," *Frontiers in Medicine*, vol. 8, 12 2021. [Online]. Available: http://dx.doi.org/10.3389/fmed.2021.741407

[9] H. A. Aliyu, R. Sudirman, M. A. Abdul Razak, and M. A. Abd Wahab, "Red blood cell classification: Deep learning architecture versus support vector machine," 2018 2nd International Conference on BioSignal Analysis, Processing and Systems (ICBAPS). IEEE, 7 2018, pp. 142–147. [Online]. Available: http://dx.doi.org/10.1109/ICBAPS.2018.8527398

[10] J. Rodellar, S. Alférez, A. Acevedo, A. Molina, and A. Merino, "Image processing and machine learning in the morphological analysis of blood cells," *International Journal of Laboratory Hematology*, vol. 40, no. S1, pp. 46–53, 5 2018. [Online]. Available: http://dx.doi.org/10.1111/ijlh.12818

[11] R. Astley. Never gonna give you up. Youtube. [Online]. Available: https://www.youtube.com/watch?v=dQw4w9WgXcQ

[12] D. Adams, *The Hitchhiker's Guide to the Galaxy*. Pan Books, 1979.

[13] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. PMLR, 09–15 Jun 2019, pp. 6105–6114. [Online]. Available: https://proceedings.mlr.press/v97/tan19a.html

[14] 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. 36 188–36 197, 2018.

[15] K. Dwivedi and M. K. Dutta, "*Microcell-Net* : A deep neural network for multi-class classification of microscopic blood cell images," vol. 40, no. 7, p. e13295, Aug. 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1111/exsy.13295