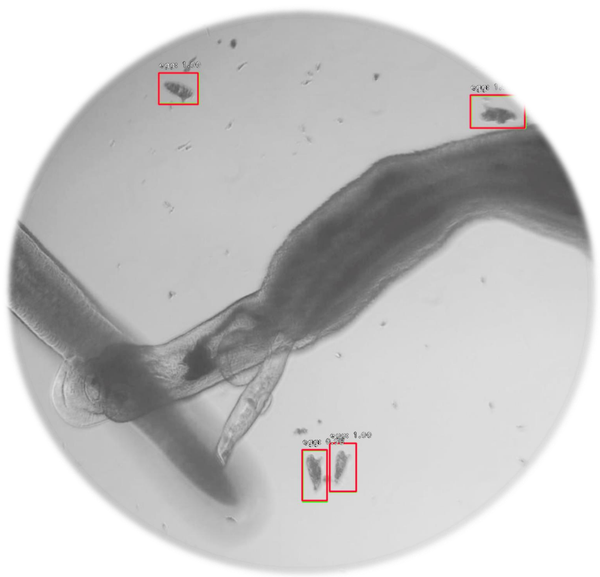
An AI hunter for automating Schistosoma egg counting



Gilda Padalino[1], Giampaolo Pagliuca[2], Josephine Forde-Thomas[1], Karl F. Hoffmann[1]

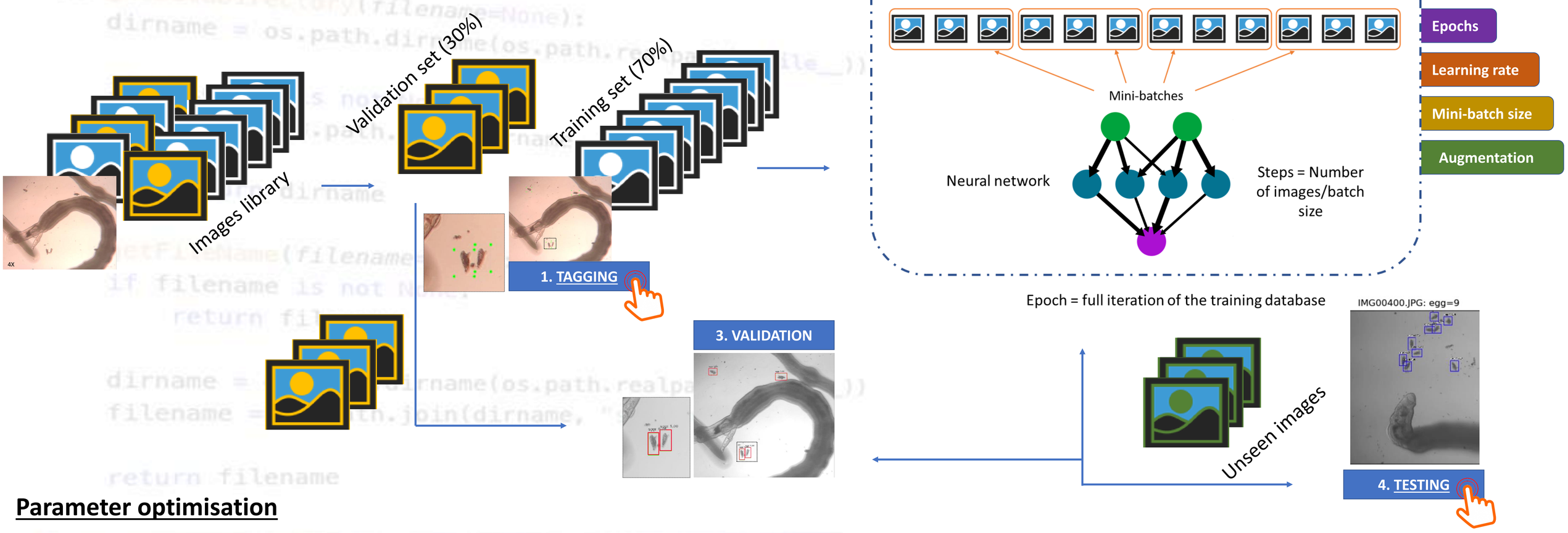
[1] Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, UK
[2] Research Scientist - Open-source contributor

Background

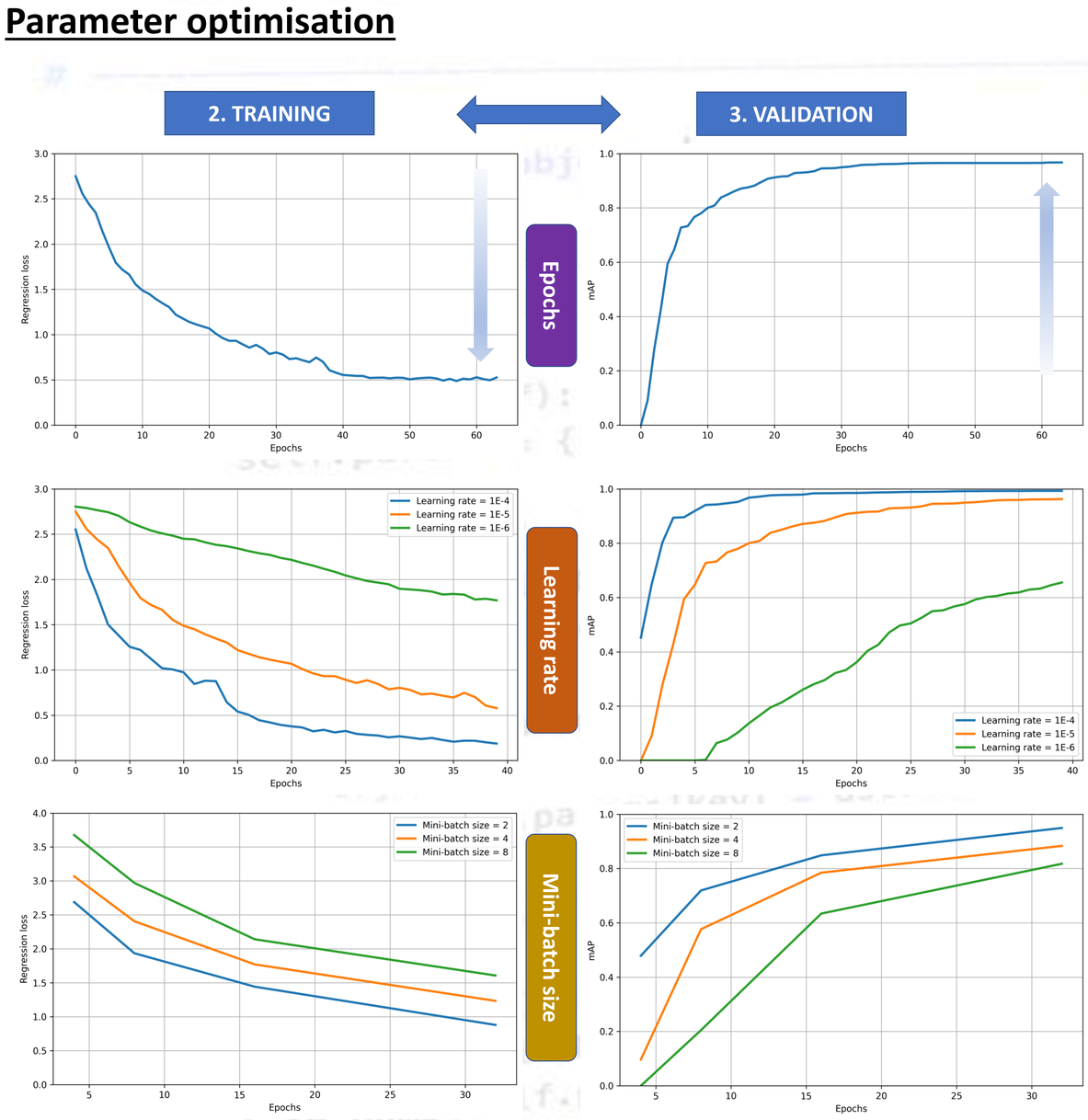
Egg counting has been for years a time-consuming process, requiring technical expertise of appropriately trained individuals using (predominantly) microscopic methods [1, 2]. The development of a system to automate parasitic helminth egg counts would represent a substantial improvement to the current process. Towards this direction, the machine learning technology and the state-of-the-art object detection models have a great potential.

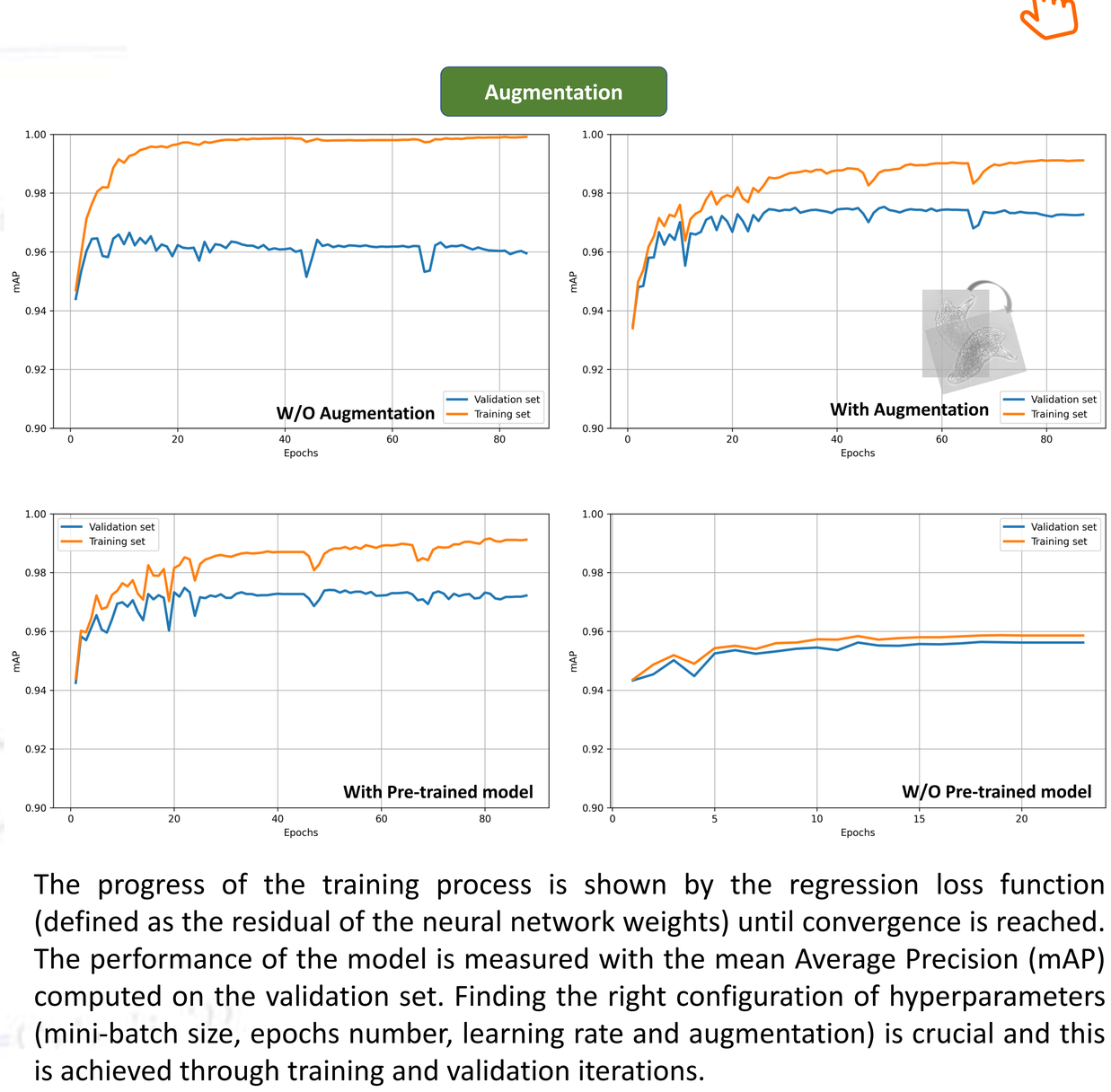
RetinaNet [3] was chosen for this egg detection 'challenge' implementing the Keras RetinaNet framework [4]. A library of 1.14k images was manually annotated (16k instances of *Schistosoma mansoni* egg) and split in a training and validation sets (70/30%). RetinaNet was trained using the first set of images and then validated with the second one. Following an iterative optimisation, the best model was tested on unseen images.

Workflow



Parameter optimisation

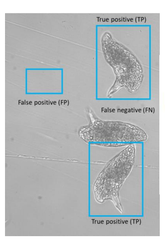




The progress of the training process is shown by the regression loss function (defined as the residual of the neural network weights) until convergence is reached. The performance of the model is measured with the mean Average Precision (mAP) computed on the validation set. Finding the right configuration of hyperparameters (mini-batch size, epochs number, learning rate and augmentation) is crucial and this is achieved through training and validation iterations.

Testing final model

Following the optimisation, the best RetinaNet inference model was tested on a set of 200 new images (containing 0-35 eggs/image). The manual and automatic count were recorded and the model was evaluated for its interpretive accuracy (based on precision and recall/sensitivity values). The protype achieved good performances with a sensitivity and precision above 88% and 91%, respectively.

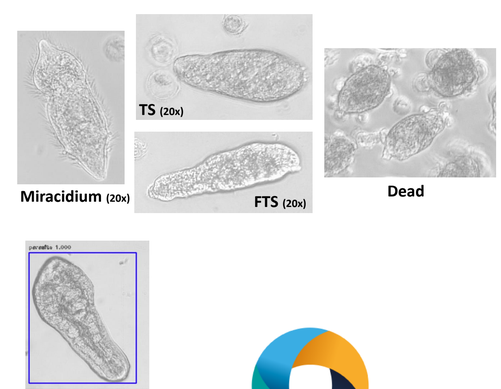


$$\text{Precision} = \frac{\#TP}{\#TP + \#FP}$$

$$\text{Recall} = \frac{\#TP}{\#TP + \#FN}$$

Next steps

This tool has a great potential for phenotypic classification and parasite enumeration also on other *S. mansoni* life cycle stages as well as other parasites (e.g. liver flukes).



References

[1] Olveda R, et al. (2010) Estimating the sensitivity and specificity of Kato-Katz stool examination technique for detection of hookworms, Ascaris lumbricoides and Trichuris trichiura infections in humans in the absence of a 'gold standard', Int J Parasitol, https://doi.org/10.1016/j.ijpara.2009.09.003 **[2]** Dalsgaard A, et al. (2014) Simple fecal flotation is a superior alternative to guadruple Kato Katz smear examination for the detection of hookworm eggs in human stool, PLoS Negl Trop Dis, https://doi.org/10.1371/journal.pntd.0003313 **[3]** Dollar P, et al. (2017) Focal Loss for Dense Object Detection, *Proc. IEEE Int. Conf. Comput. Vis.* **[4]** GitHub, "Keras implementation of RetinaNet object detection." [Online]. Available: https://github.com/fizyr/keras-retinanet . [Accessed: 21-Aug-2018].

PRIFYSGOL ABERYSTWYTH UNIVERSITY

SUPERCOMPUTING WALES
UWCHGYFRIFIADURA CYMRU