

Introduction

Malaria is caused by infection of the protozoan parasite *Plasmodium*. When *Plasmodium* enters the liver, it replicates rapidly and penetrates red blood cells throughout a host body's bloodstream. Prompt diagnosis of *Plasmodium* infection is essential for control and management of malaria. Although *Plasmodium* infection is easily detected with a polymerase chain reaction (PCR) test, widely available in wealthy communities, under-resourced communities have traditionally struggled to track infection. Conversely, light microscopes are used broadly across medical contexts and widely available in communities struggling with the ability to identify and track malarial infections.

Recent image recognition tools developed in the machine vision community have opened the potential for cost effective, image-based detection of *Plasmodium*. Specifically, convolutional neural networks (CNNs) have proven to be successful at complex image classification tasks and may have the ability to identify *Plasmodium* with task-specific training. Importantly, pre-trained CNNs can be further trained with transfer-learning to achieve expertise based on specific input data parameters. This means that communities with different microscopic tools, calibrated to different settings or even expressing different flaws, may be able to apply CNNs with modest transfer-learning to resolve unique challenges in their ability to detect malaria.

Here, I demonstrate that training *GoogLeNet* (Inception v1) on a transfer-learning task to identify whether microscopically-derived red blood cell images are or are not infected with *Plasmodium* is successful at detecting malarial infection with a >95% accuracy based on image properties alone.

Methods

The dataset consisted of 22,046 images of single red blood cells collected with a light microscope that were labeled by a medical pathologist as either infected with *Plasmodium* (1) or uninfected (0). Images were of a variety of sizes and shapes and contained a variety of color profiles (*Fig. 1*). In addition, an evaluation set of 2,756 labeled images was left out for model training.



Figure 1. A sample of five red blood cell images included in the training set. The cells have been determined to be infected by *Plasmodium* or not by a medical pathologist. The labels, from left to right are: Infected, Uninfected, Infected, Uninfected, Infected.

A convolutional neural network using *GoogLeNet* architecture, pre-trained on ImageNet, was adapted with transfer learning to classify whether or not microscopic images of red blood cells are infected with the parasite *Plasmodium* (Szegedy et al. 2015). The motivation to use this network architecture was a compromise between its documented classification accuracy (top-1 error: 30.22) and efficiency of training time. *GoogLeNet* requires fewer than half the number of processor operations as comparable networks, often using residual architectures, while achieving comparable performance (Fig. 2A). *GoogLeNet* contains 22 stacked convolutional layers consisting of 1x1, 3x3, and 5x5 convolutional filters (Fig. 2B). Its key innovation is approximating input sparsity through the use of “inception” modules, which reduce memory burden as compared to fully sparse architectures that may result in memory explosions. Inception modules reduce the dimensionality of sparse inputs.

Prior to applying the model, each red blood cell image was resized to a standard height and width of 250x250 pixels. On top of the standard *GoogLeNet* architecture, the output layer was fine-tuned with four fully connected layers of decreasing magnitude were added (input neurons per layer: 1024, 512, 128, 32). The Adam optimizer was used with a learning rate set to 0.001, and Cross Entropy Loss was used to assess model performance. Batch sizes of 64, 128, and 1024 samples were tested. The network trained on the largest batch size converged quickest over fewer training epochs to produce a loss under < 0.05. This network achieved the highest classification accuracy on the hidden test set, as compared to networks trained on smaller batch sizes. Indeed, this network reached peak classification accuracy of 96.8% on the hidden test set after just 3 epochs, and plateaued (with a slightly lower accuracy) with more training likely due to overfitting.

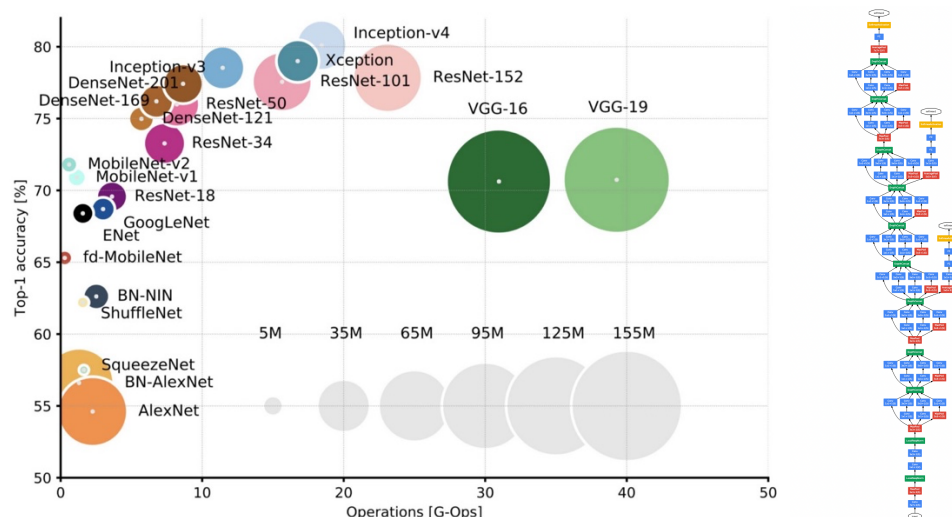


Fig. 2. (A) A comparison of common neural network architectures for image classification, with respect to peak classification accuracy as compared to number of processor operations.

Adapted from Eugenio Culurciello's: <https://culurciello.github.io/>. **(B)** The *GoogLeNet* architecture consisting of 22 convolution layers. The fully connected layers are not depicted in this schematic. Adapted from Szegedy et. al, 2016.

Concluding remarks

Here, we trained inception (Googlenet) on a transfer-learning task to identify infection of the malaria-causing parasite, *Plasmodium*, in microscopic images of red blood cells. We show that transfer-learning achieves near-human levels of expertise (96.8% test accuracy) in detecting Malaria causing infections in red blood cells.

Szegedy, C., et al. (2015) Going Deeper with Convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 7-12 June 2015, 1-9.
<https://doi.org/10.1109/CVPR.2015.7298594>