A CLINICAL APPLICATION OF DEEP LEARNING

MALARIA DETECTION USING CONVOLUTIONAL NEURAL NETWORKS

Scott Penco

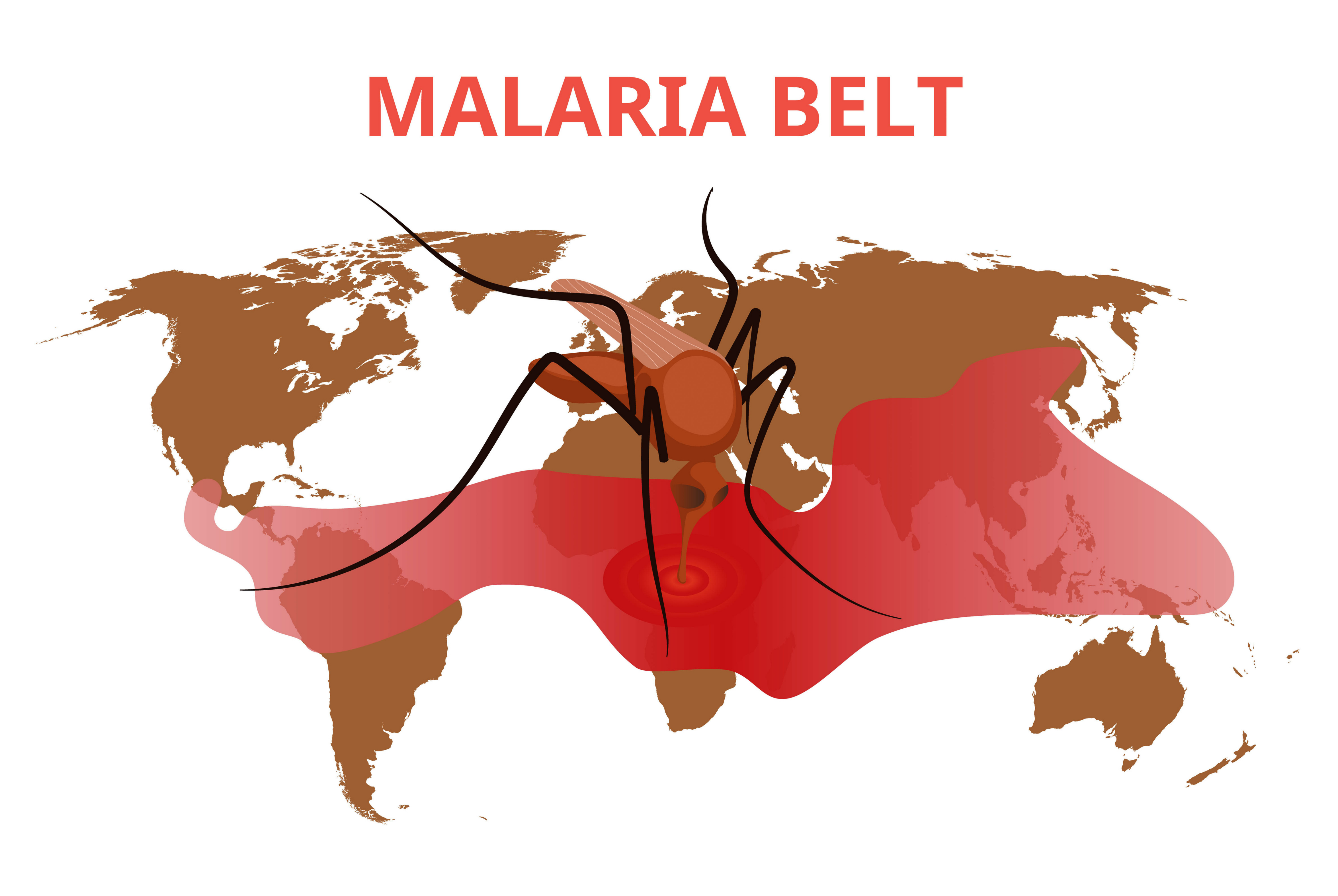
MIT-ADSP

AGENDA

- Problem Overview
- Solution Overview
- Model Comparisons
- Key Findings
- Recommendations and Potential Benefits

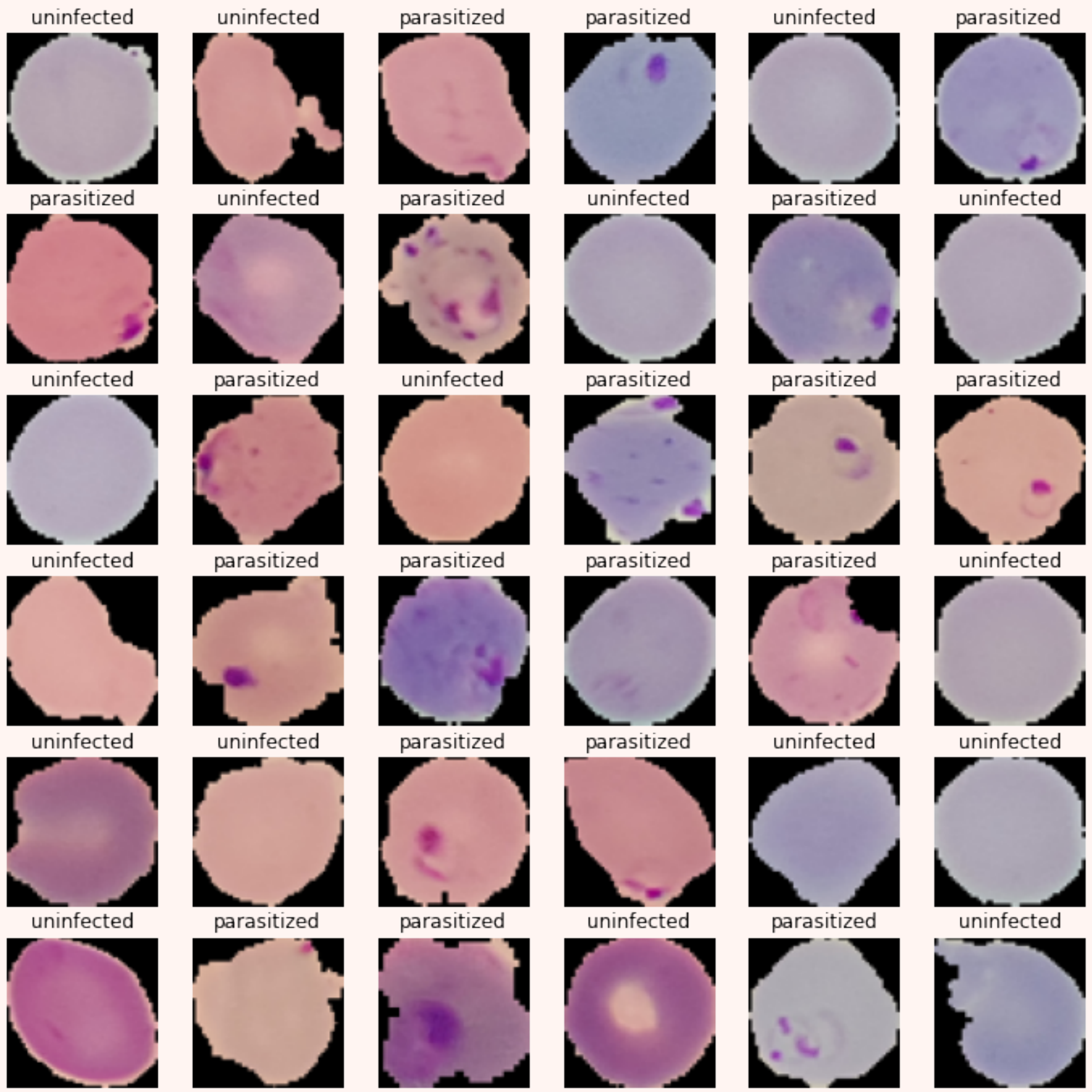
PROBLEM OVERVIEW

- Malaria is a life-threatening disease caused by parasites that are transmitted to people through the bites of infected female Anopheles mosquitoes.
- In 2020, there were an estimated 241 million cases of malaria worldwide.
- The WHO African Region carries a disproportionately high share of the global malaria burden. In 2020, the region was home to 95% of malaria cases and 96% of malaria deaths. Children under 5 accounted for about 80% of all malaria deaths in the Region.
- Where Malaria Occurs (Malaria Belt)
- How can we help solve this problem:
  - Case management
  - Surveillance

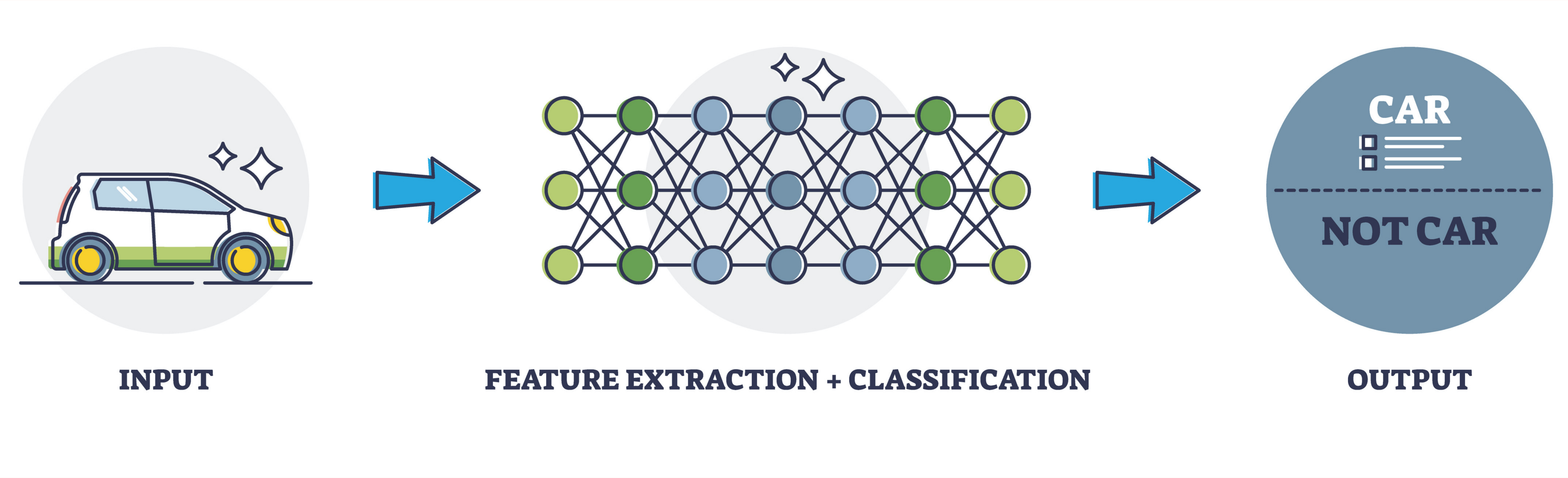


SOLUTION OVERVIEW

- Help with detection and surveillance of Malaria.
- Red Blood Cells affected with Malaria virus are morphologically different from uninfected blood cells.
- This presents a classification problem that can be solved using Deep Learning and CNNs



DEEP LEARNING OVERVIEW

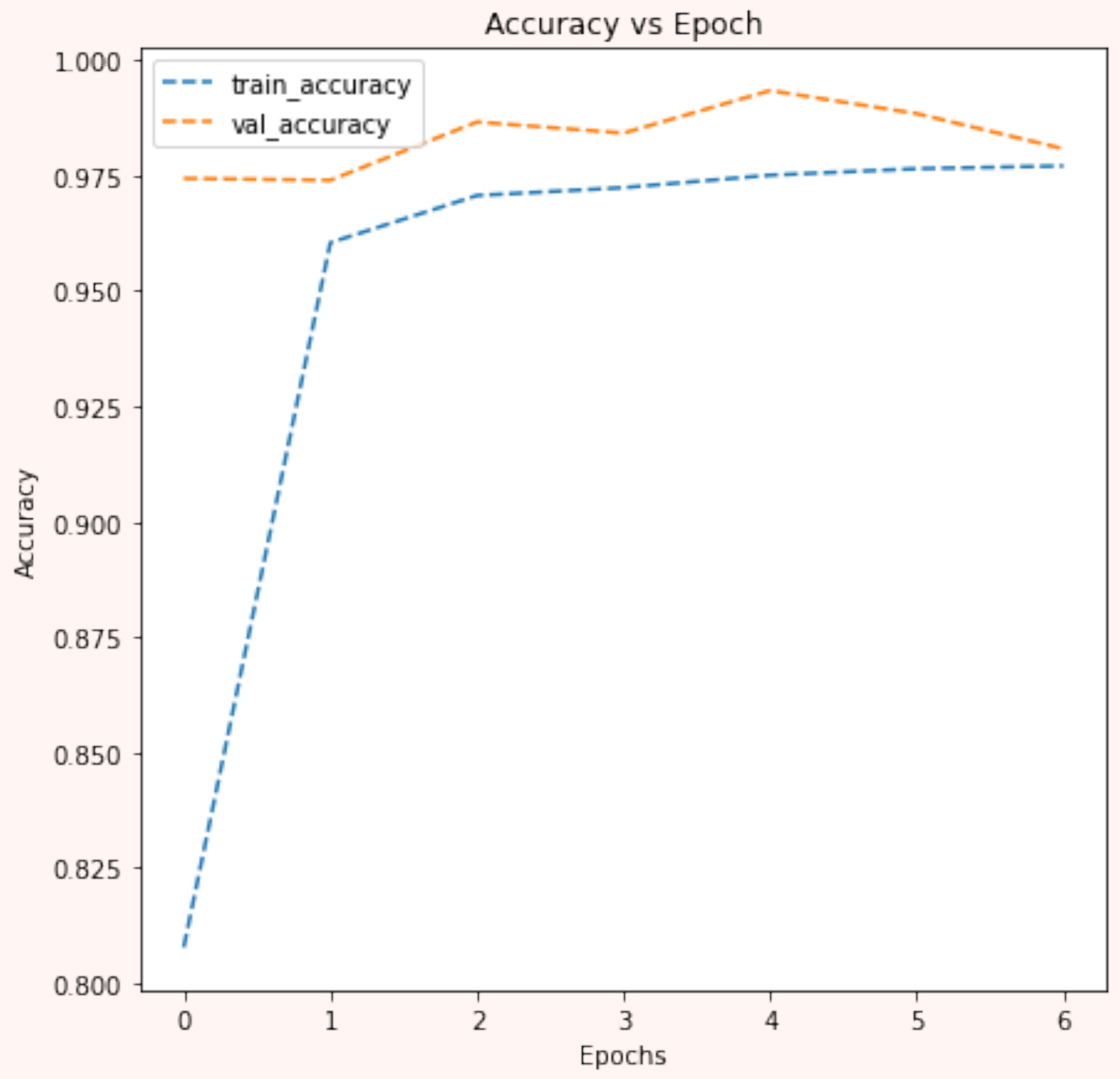


MODEL 1

Our first model uses many 2D convolutional layers, which scans the input based on the parameters. This is feature extraction.

This code passed the images through a 32 filter size, and through max-pooling downsampled the convolutional layer with the maximum value of the images, helping us determine edges in our image dataset.

Activation function for output layer: sigmoid



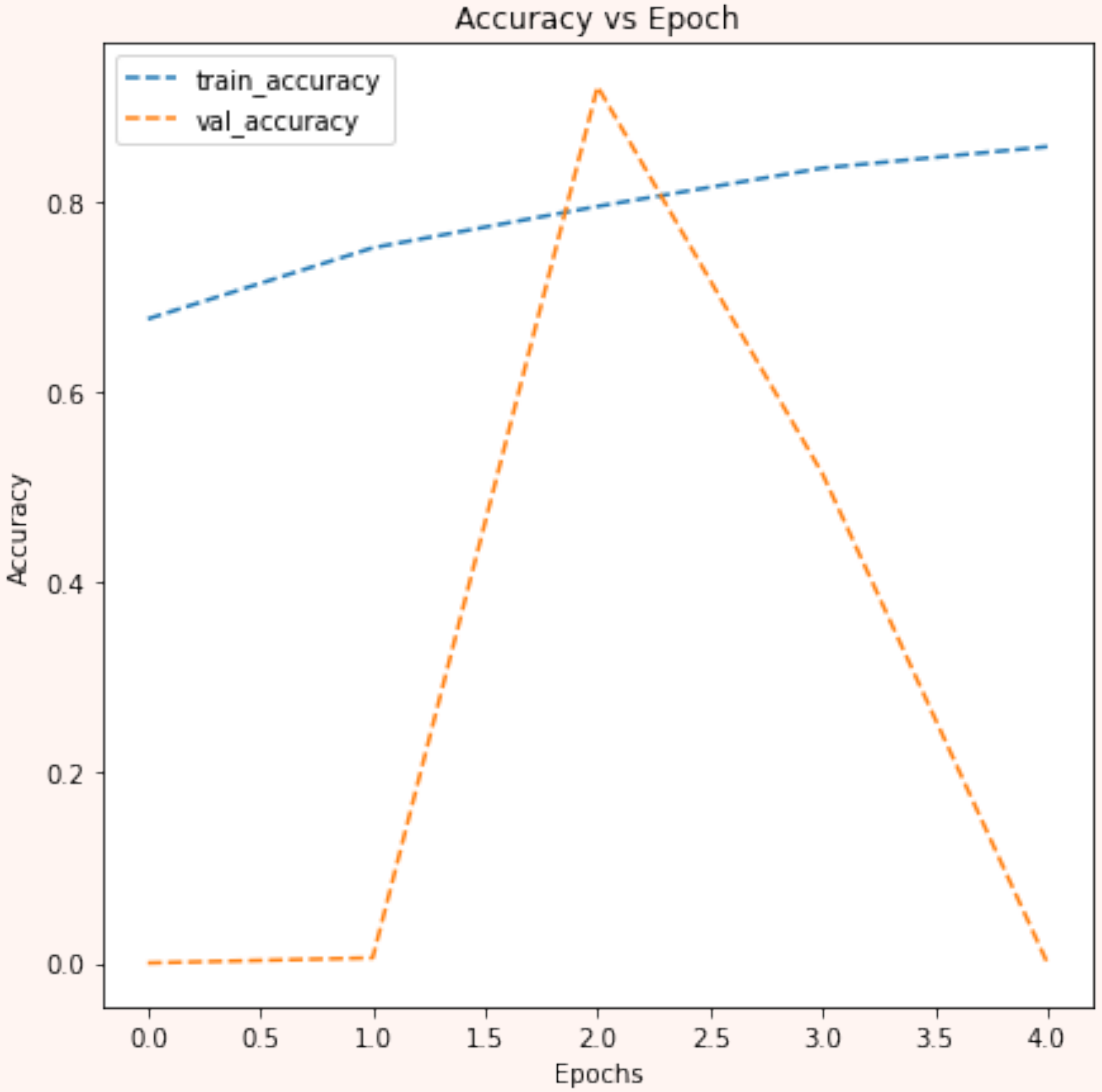
MODEL 2

Our second model introduces Batch Normalization and Leaky Relu as an activation function. This is a linear function.

Why may it not be useful for binary classification?

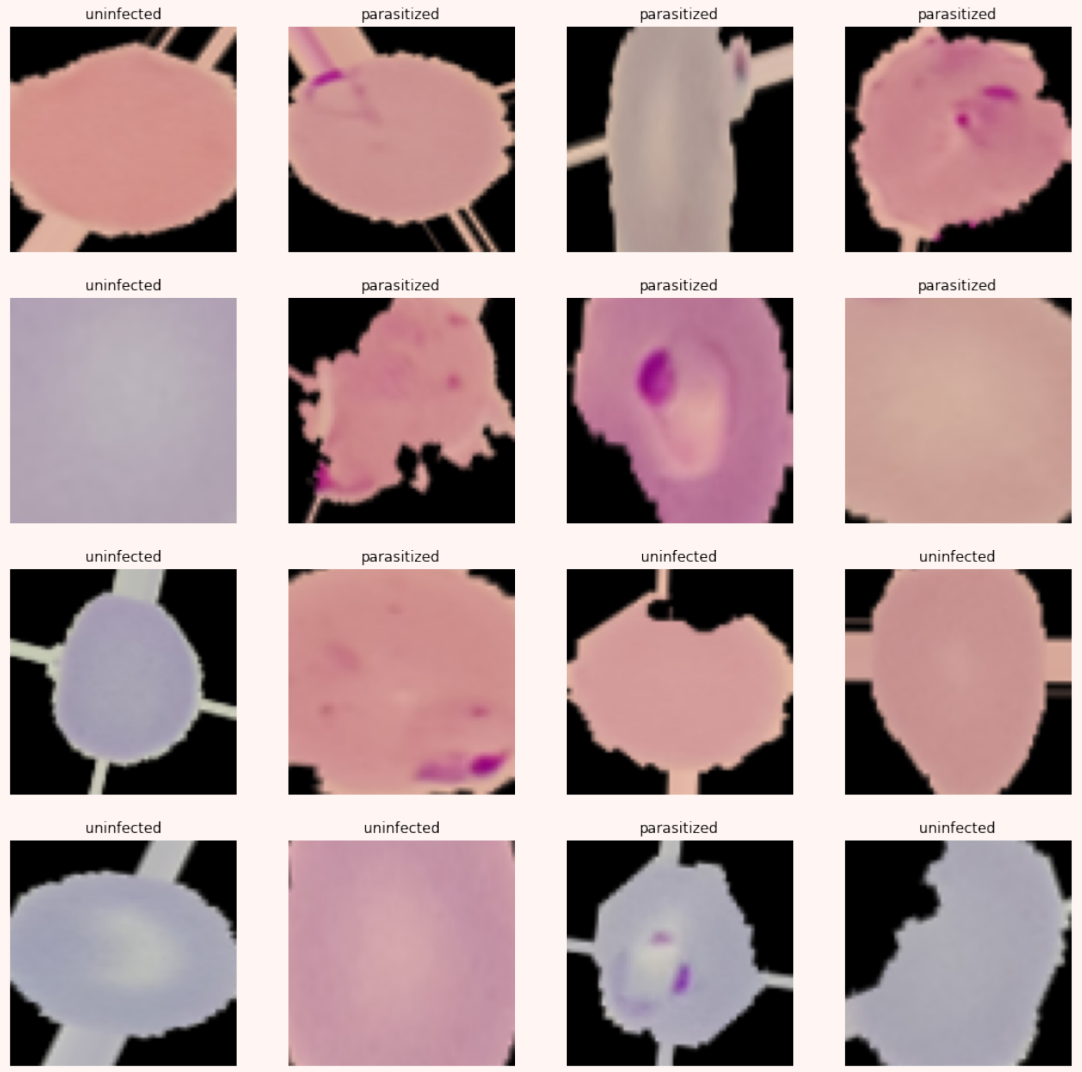
This process normalized any outlier in our data from the input layer.

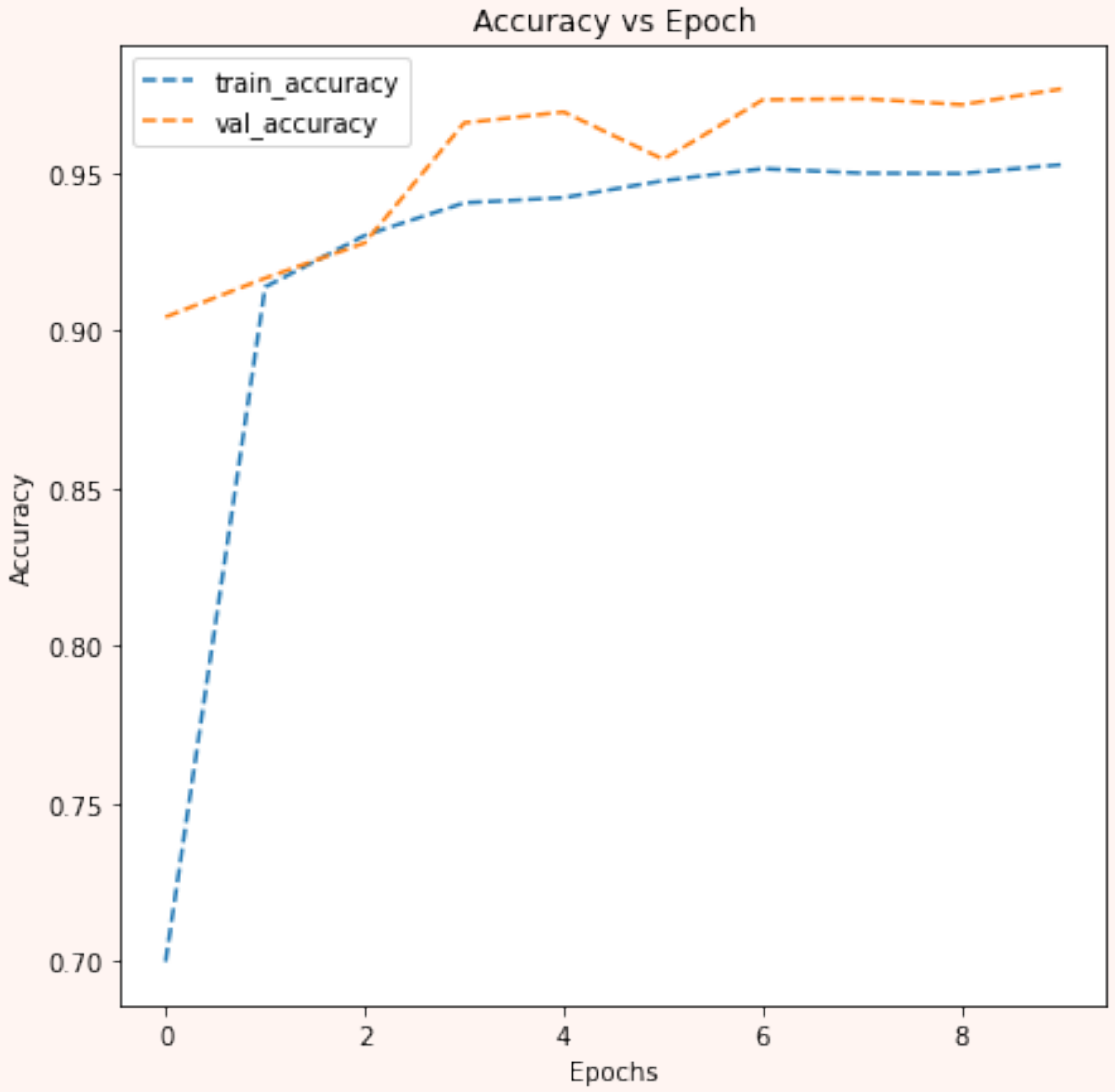
However was not very effective in determining a working classification model. Potentially Computationally expensive



MODEL 3

Model 3 introduces Data Augmentation, while the CNN and neural networks remain the same as model1. Augmentation expands and manipulates our image data set by zooming and rotating the samples.

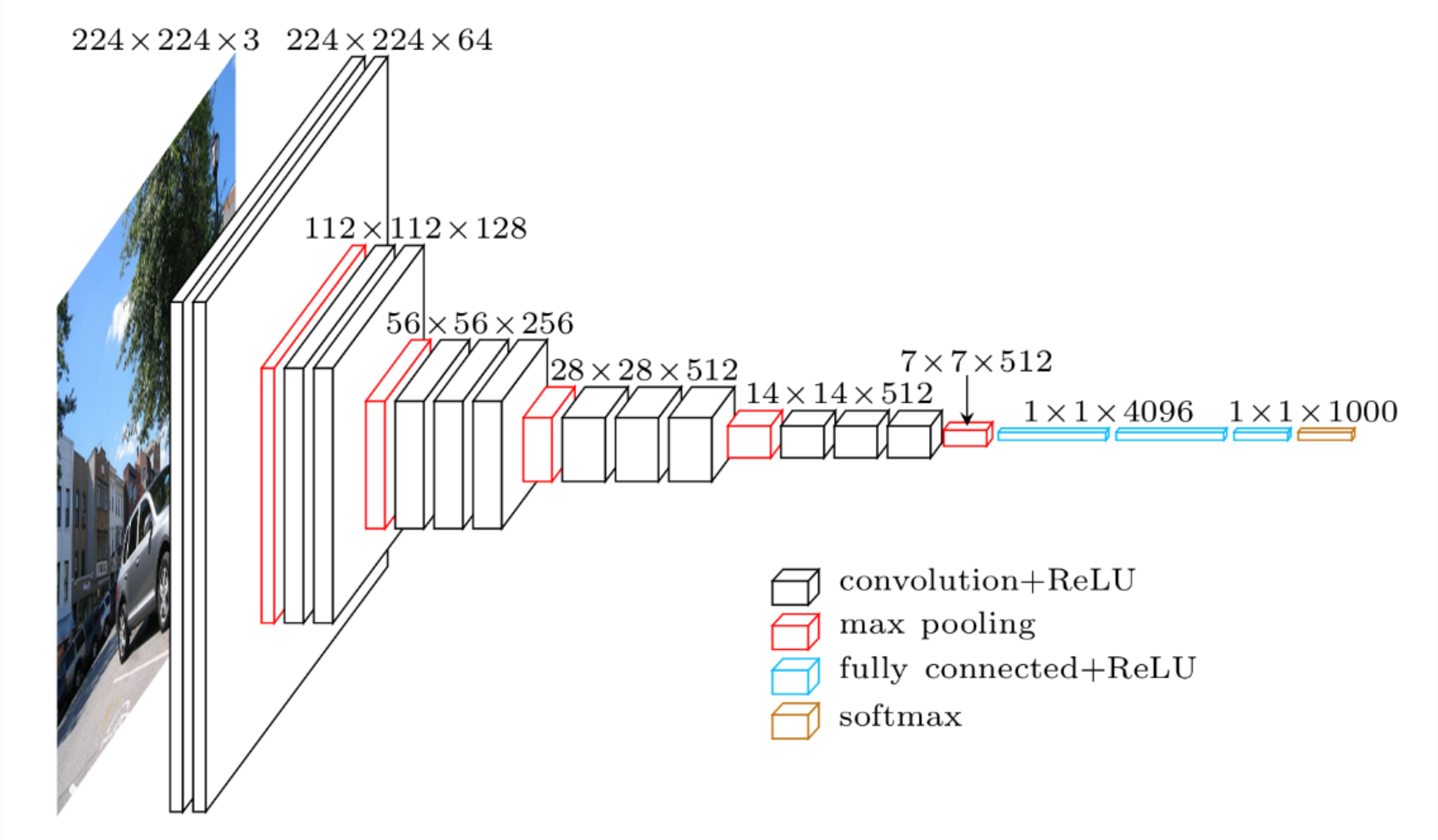


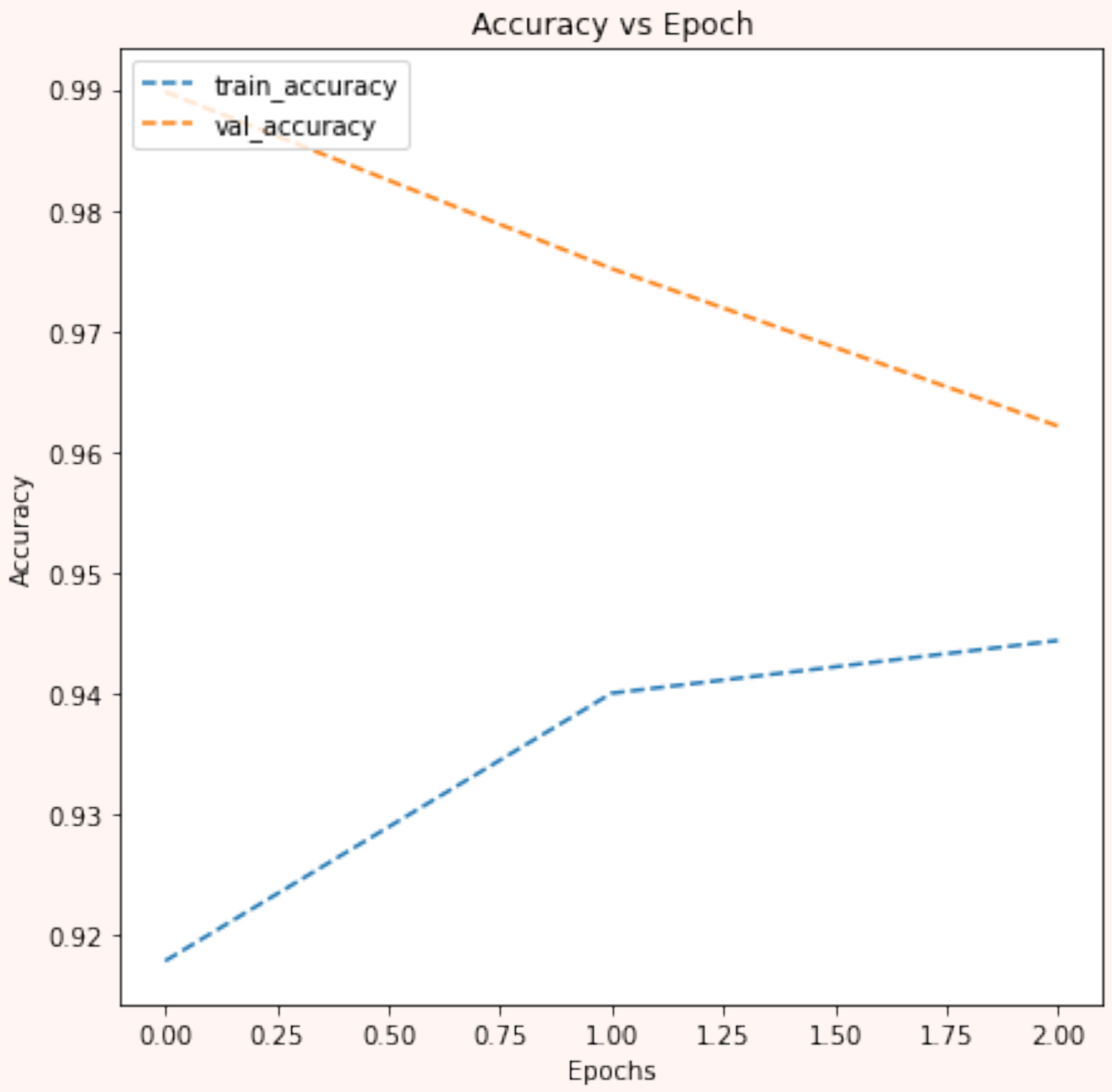


MODEL 4

Model 4 uses a pre-trained model (VGG16). This model was used to win an ImageNet competition in 2014, and is considered an excellent benchmark for image classifications. Using this model we can see how effective our built models are in this dataset versus a broad spectrum classification model.

Activation Function: Softmax, typically used in multi classification models





MODEL COMPARISONS

MODEL 1

Precision	Recall	F-1 Score
0.98	0.98	0.98

MODEL 2

Precision	Recall	F-1 Score
0.74	0.52	0.37

MODEL 3

Precision	Recall	F-1 Score
0.98	0.98	0.97

MODEL 4

Precision	Recall	F-1 Score
0.95	0.95	0.95

KEY FINDINGS

The model chosen has shown to have high accuracy on the data set given, suggesting machine learning is a viable and plausible diagnostic tool for Malaria detection

The advent of machine learning models could challenge the current zeitgeist of health care diagnosis, disease prevention, and surveillance especially on the population level.

Implementation of a new model must be compared against the current methods of diagnosis. Currently Malaria detection is determined by (1) diagnosis of a patients symptoms, (2) microscopy, or (3) Antigen detection.

RECOMMENDATIONS AND POTENTIAL BENEFITS

The next steps involved with the model would be to determine external validity, data sampling increase, test comparison and implementation into the public health domain.

The current methods of detection ,due to their need for labour and case by case nature, may be taxed more temporally for population diagnosis.

This is an important feature to note: disease tracking and surveillance is completed on a population level and not necessarily by individual patient cases.

Provide framework or foundations for other diseases that can benefit from diagnosis or tracking in Machine Learning Models