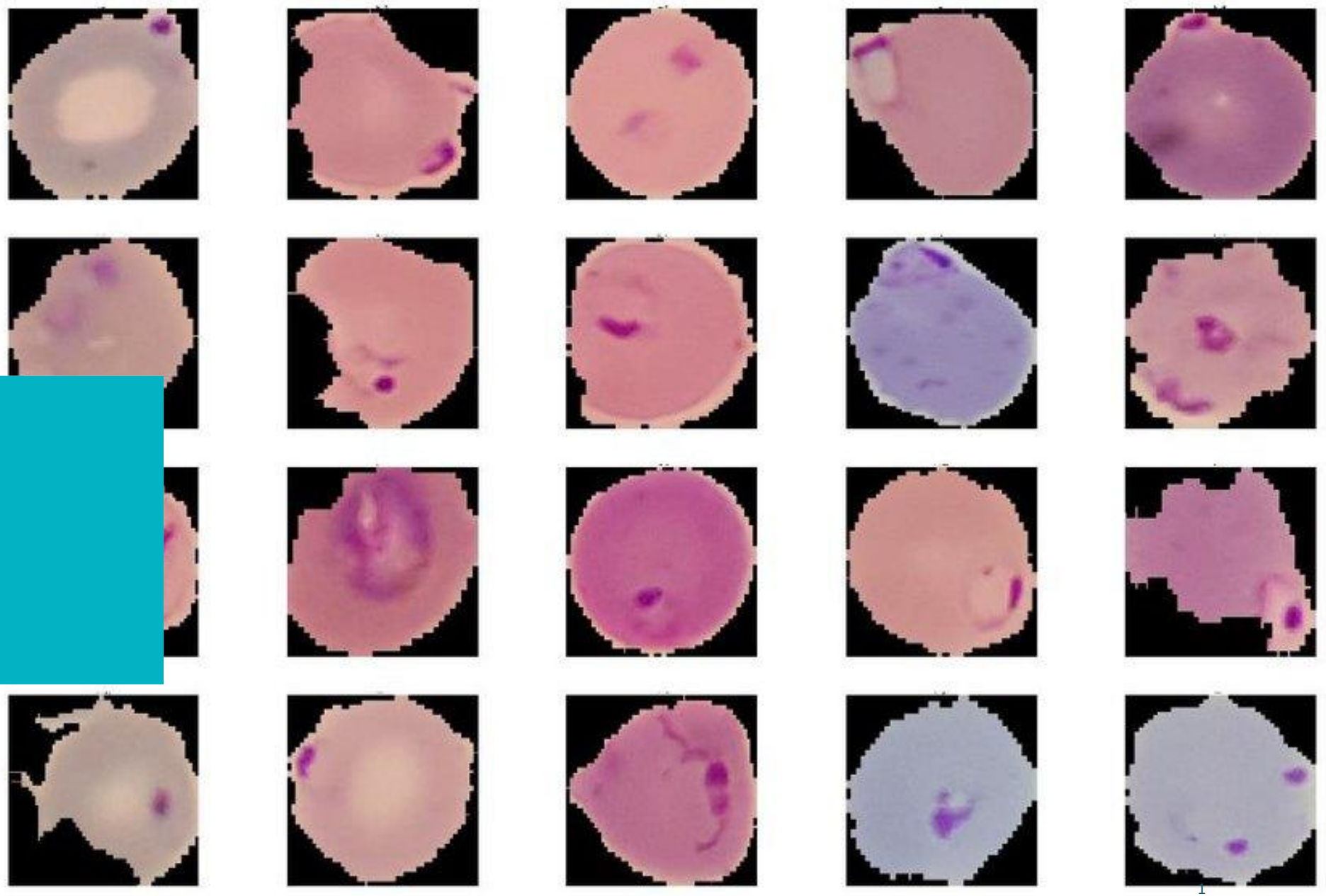
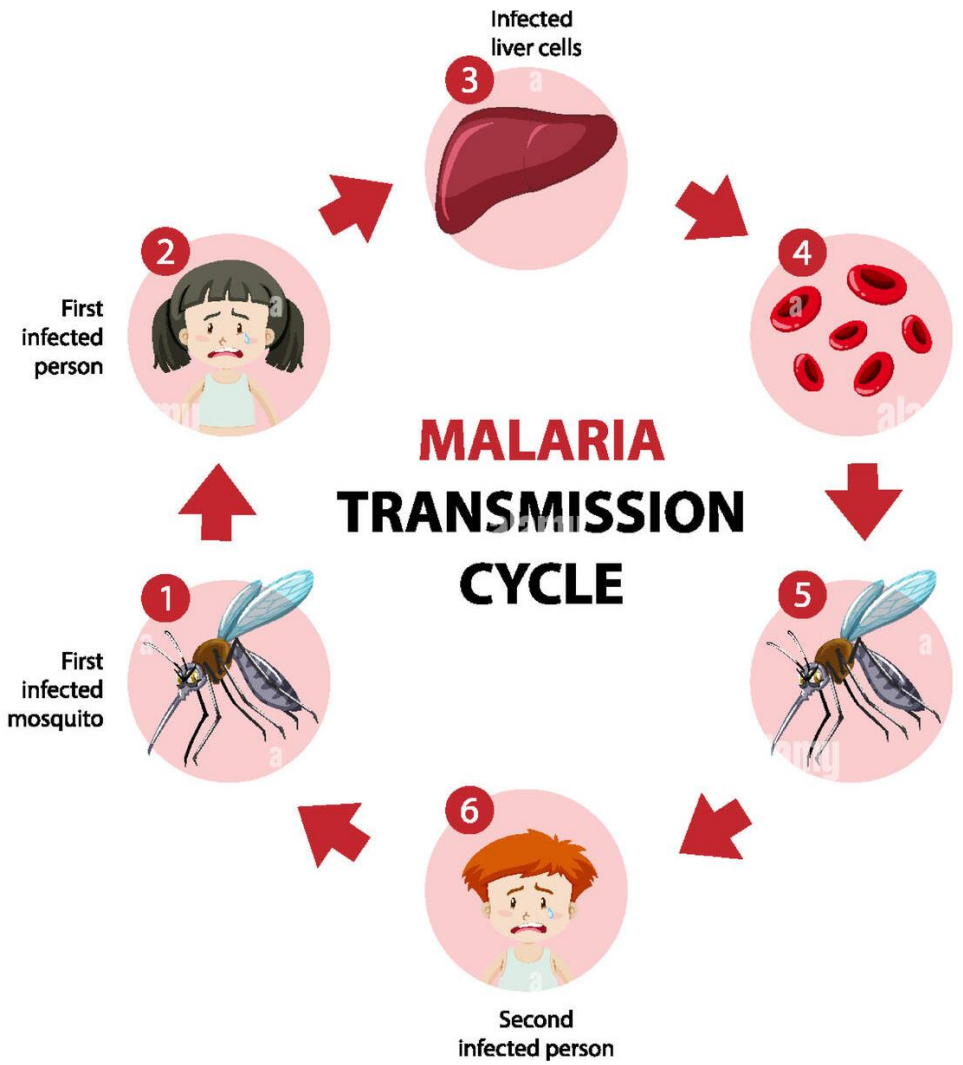
MALARIA DETECTION

Roula Krayem





INTRODUCTION

- What is Malaria?
- Malaria caused 627000 deaths in 2020
- Diagnoses methods including Blood Smear test
- Microscopic images of Red Blood Cells (RBCs)

https://www.alamy.com/malaria-transmission-cycle-information-infographic-image444548497.html?imageid=D0BAED1C-4F8E-43D2-A932-04A6E64DD7E9&p=540308&pn=1&searchId=cb74d113acdde3fc17256d93a4d8b6e3&searchtype=0

PROBLEM DEFINITION

Millions of Malaria cases and handers of thousands deaths annually

Providers study each RBCs sample individually under the microscope

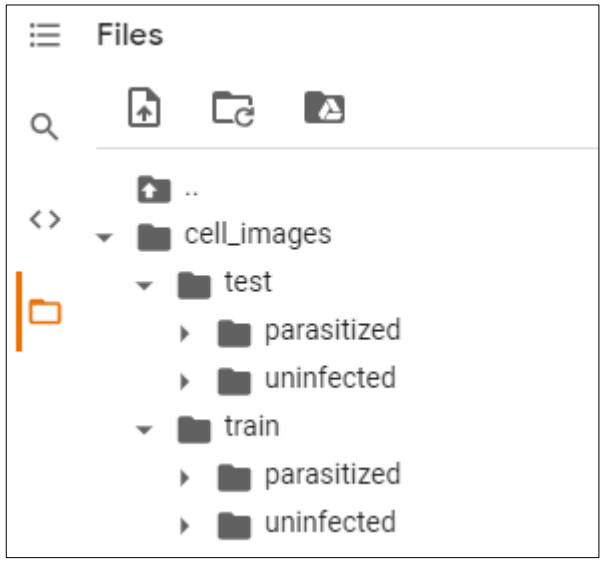
Slow process that could delay treatment

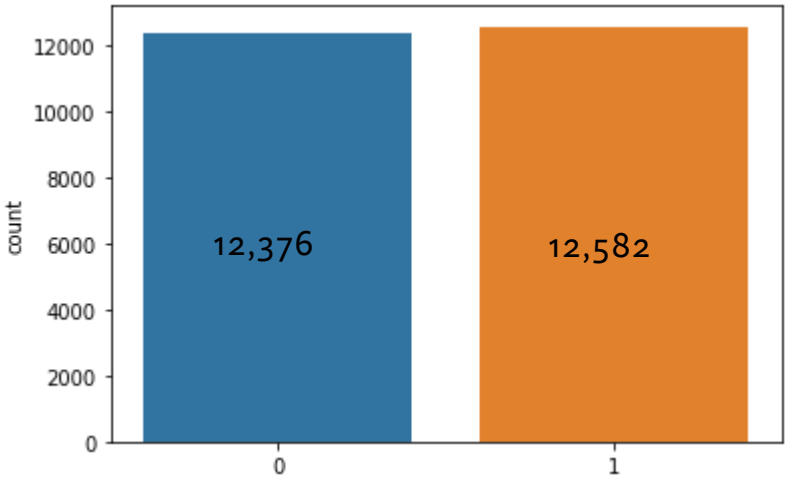
SOLUTION

Build a Neural Network model to expedite and facilitate Malaria detection

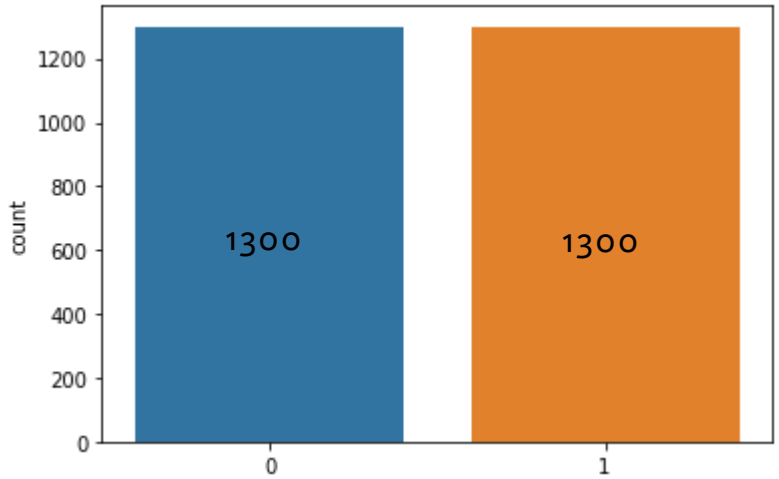
STEP 1	STEP 2	STEP 3
Perform data Exploration and note key takeaways	Build and test multiple Neural Network models to reach the best performing Model	Select the best performing model and provide recommendations

DATA EXPLORATION

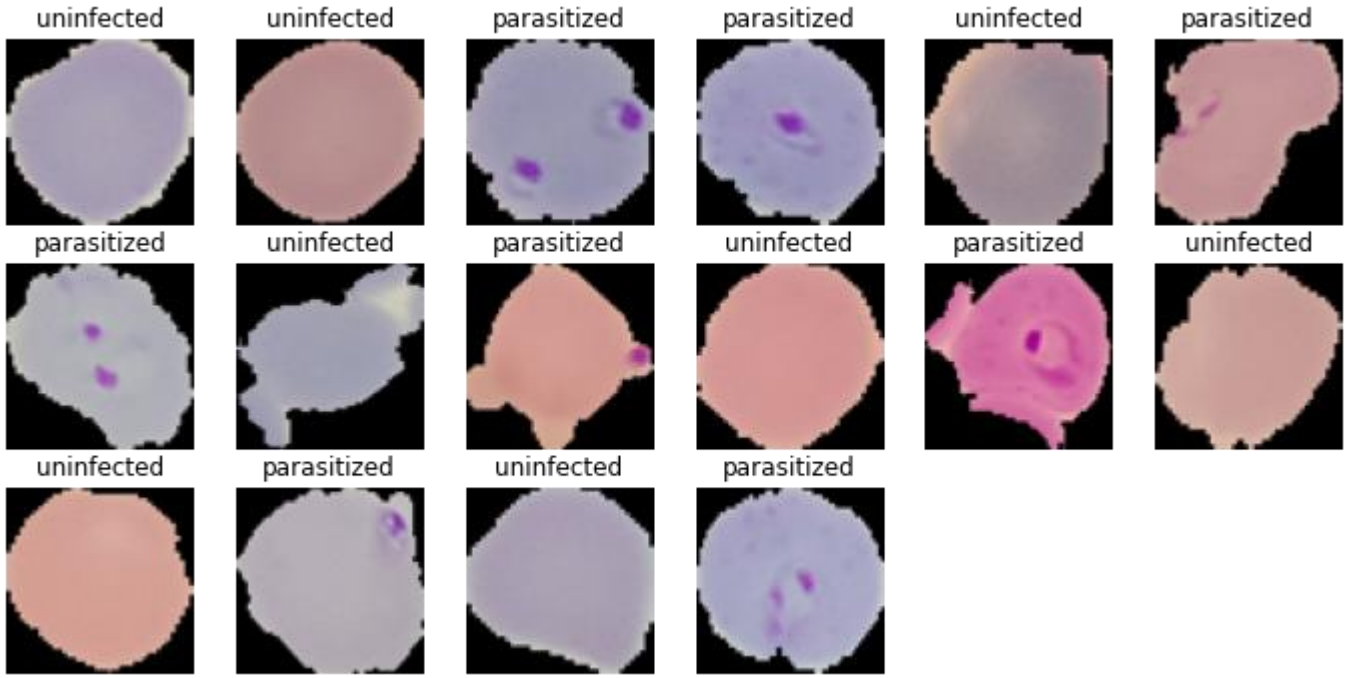




Train Data



Test Data



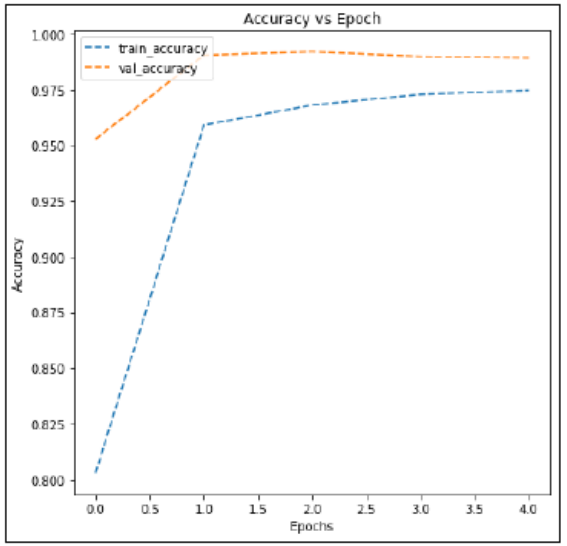
DATA PROCESSING TECHNIQUES

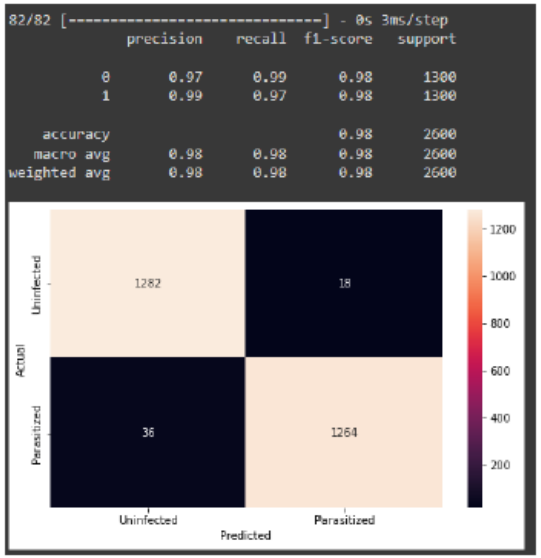
- ✓ Image resizing (64 x 64)
- ✓ Normalization
- ✓ One-Hot Encoding

MODEL CONVERSATION

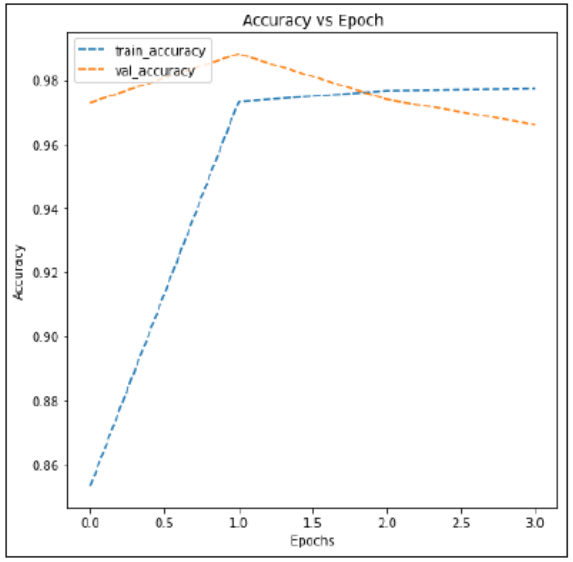
FROM SCRATCH

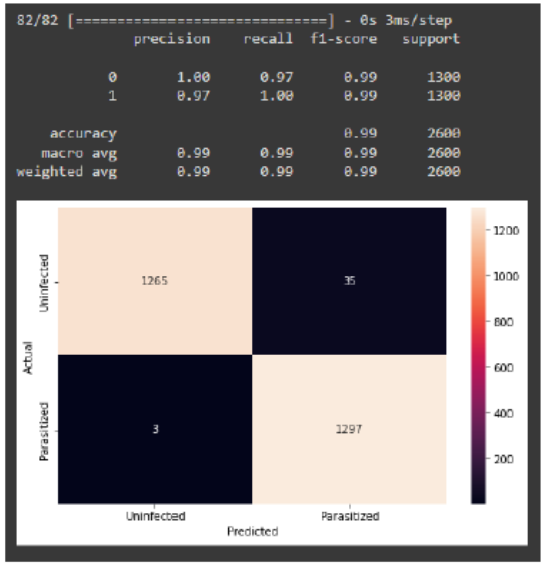
Base Model



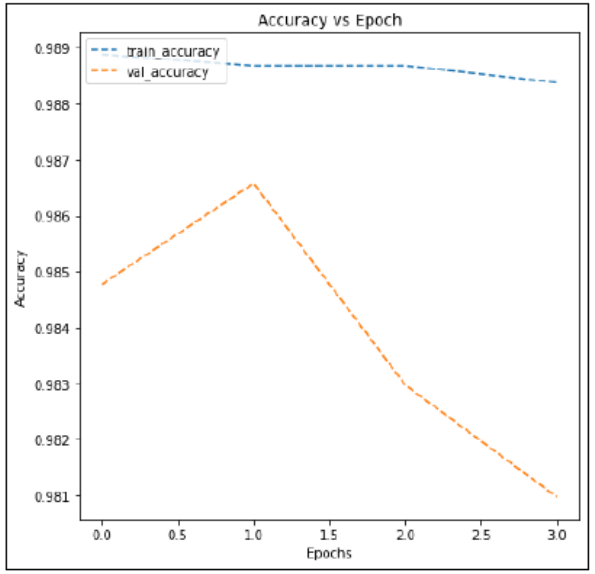


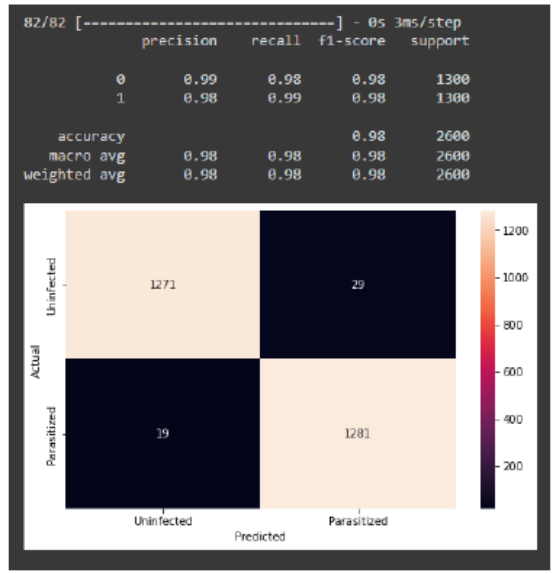
First Model

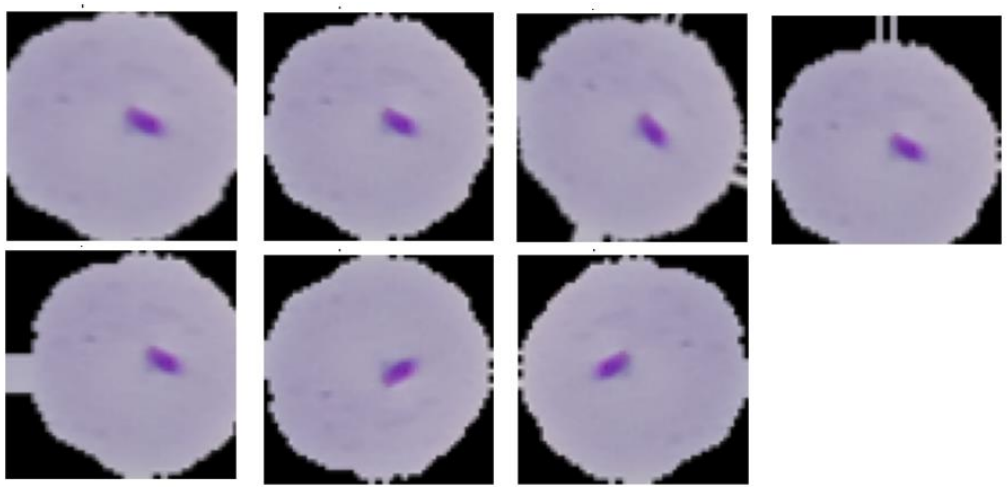




Second Model





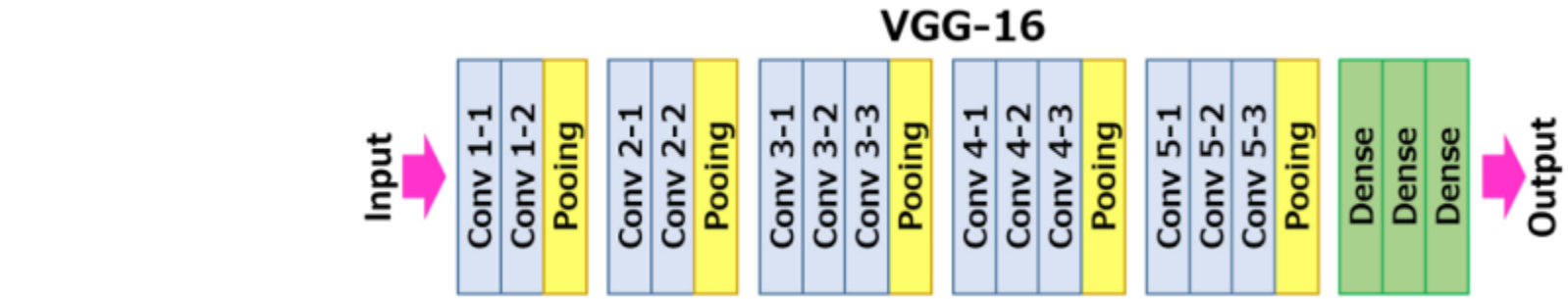


DATA AUGMENTATION

Data Image Process	Precision Unin*	Precision Para*	Recall Unin*	Recall Para*	F1 Score Unin*	F1 Score Para*	Accuracy	Misclassified Para*	Misclassified Unin*
Zoom	0.97	0.99	0.99	0.97	0.98	0.98	98%	37	17
Shear	0.98	0.99	0.99	0.98	0.98	0.98	98%	30	15
Rotation	0.98	0.98	0.98	0.98	0.98	0.98	98%	22	21
Width Shift	0.97	0.99	0.99	0.97	0.98	0.98	98%	39	11
Height Shift	0.99	0.98	0.98	0.99	0.98	0.98	98%	14	31
Vertical Flip	0.97	0.99	0.99	0.97	0.98	0.98	98%	38	15
Horizontal Flip	0.98	0.99	0.99	0.98	0.99	0.98	98%	29	10
All	0.98	0.99	0.99	0.98	0.98	0.98	98%	28	15

*Para= Parasitized and Unin = Uninfected

VGG16



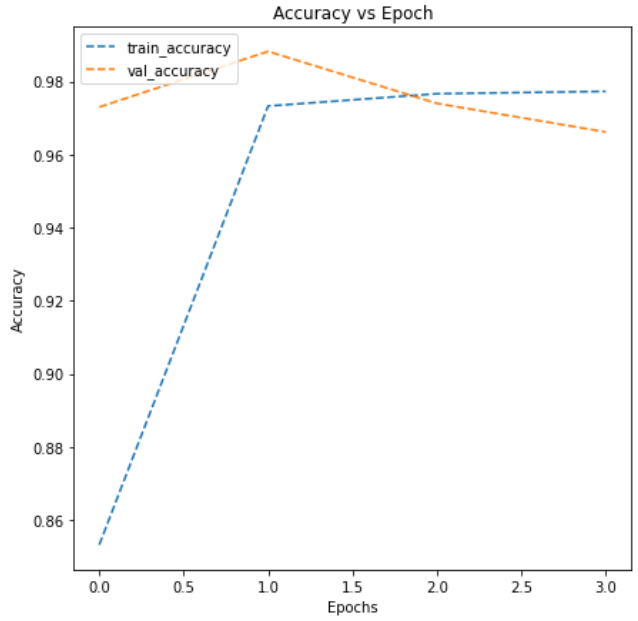
Learning rate / batch size	Precision Unin*	Precision Para*	Recall Unin*	Recall Para*	F1 Score Unin*	F1 Score Para*	Accuracy	Misclassified Para*	Misclassified Unin*
0.001/32	0.98	0.84	0.82	0.98	0.89	0.91	90%	24	238
0.00001/23	0.96	0.89	0.89	0.96	0.92	0.93	92%	50	147
0.00001/65	0.95	0.91	0.91	0.95	0.93	0.93	93%	66	123
0.0000/100	0.96	0.89	0.88	0.96	0.92	0.93	92%	48	154

*Para= Parasitized and Unin = Uninfected

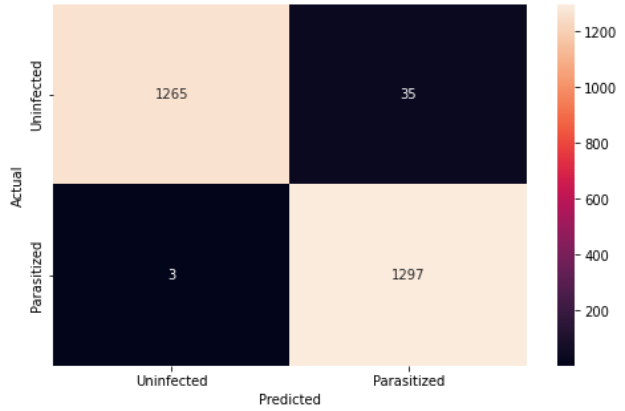
BEST MODEL

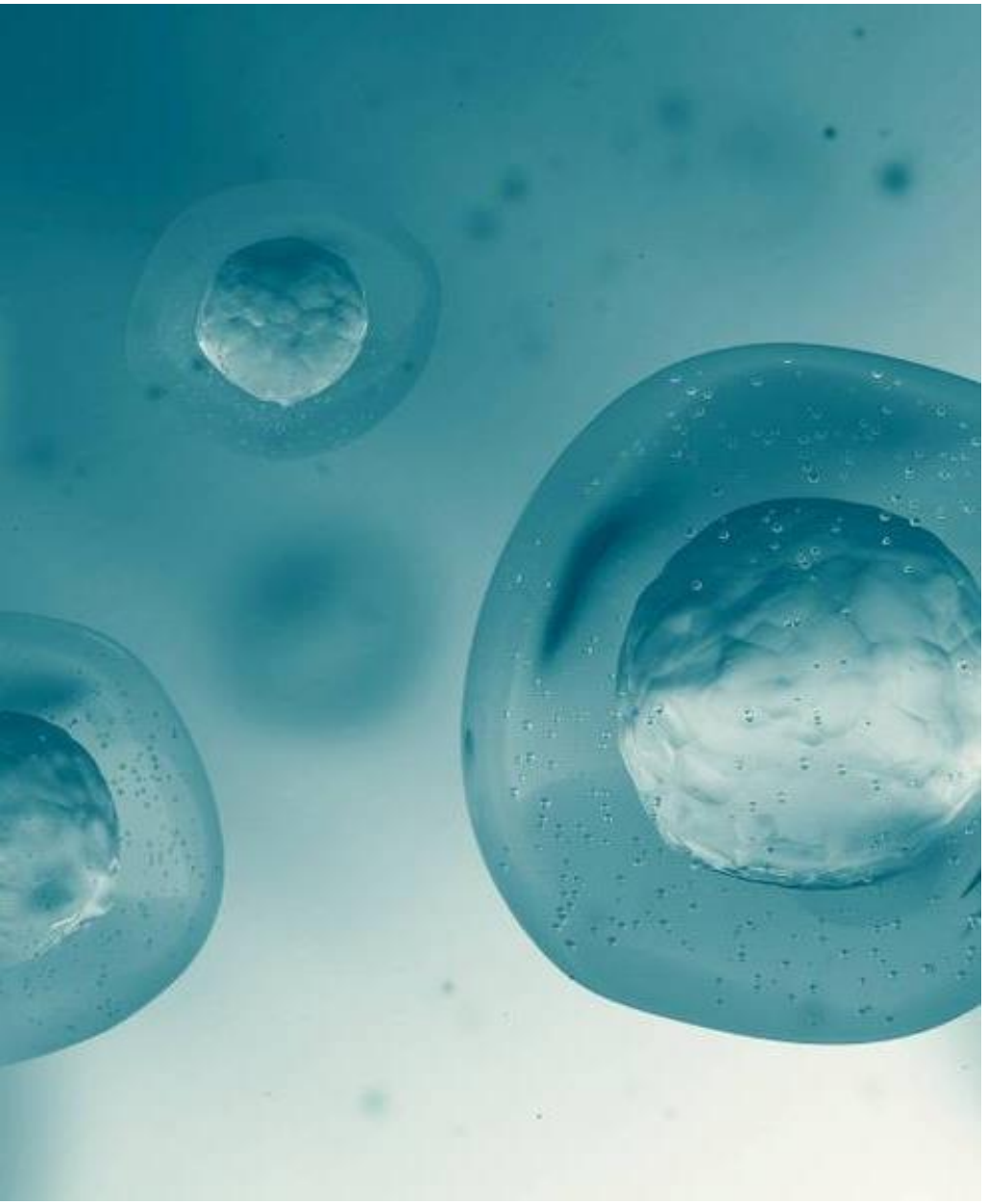
MODEL (1)

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 64, 64, 32)        416

 max_pooling2d (MaxPooling2D  (None, 32, 32, 32)       0
 )

 conv2d_1 (Conv2D)           (None, 32, 32, 32)        4128

 max_pooling2d_1 (MaxPooling  (None, 16, 16, 32)       0
 2D)

 dropout (Dropout)           (None, 16, 16, 32)        0

 conv2d_2 (Conv2D)           (None, 16, 16, 32)        4128

 max_pooling2d_2 (MaxPooling  (None, 8, 8, 32)         0
 2D)

 dropout_1 (Dropout)         (None, 8, 8, 32)          0

 conv2d_3 (Conv2D)           (None, 8, 8, 32)          4128

 max_pooling2d_3 (MaxPooling  (None, 4, 4, 32)         0
 2D)

 conv2d_4 (Conv2D)           (None, 4, 4, 32)          4128

 max_pooling2d_4 (MaxPooling  (None, 2, 2, 32)         0
 2D)

 dropout_2 (Dropout)         (None, 2, 2, 32)          0

 flatten (Flatten)           (None, 128)               0

 dense (Dense)               (None, 512)               66048

 dropout_3 (Dropout)         (None, 512)               0

 dense_1 (Dense)             (None, 2)                 1026

=================================================================
Total params: 84,002
Trainable params: 84,002
Non-trainable params: 0
_________________________________________________________________
```




```
82/82 [==============================] - 0s 3ms/step
              precision    recall  f1-score   support

           0       1.00      0.97      0.99      1300
           1       0.97      1.00      0.99      1300

    accuracy                           0.99      2600
   macro avg       0.99      0.99      0.99      2600
weighted avg       0.99      0.99      0.99      2600
```





RECOMMENDATIONS

MODEL FUTURE IMPROVEMENTS

1. Try different learning rates and batch-size when fitting the second model using the data image generator.
2. Use the Ensemble technique with multiple weak models to improve the results.

POLICYMAKERS

1. Adopting the first model to automate and accelerate Malaria detection.
2. Requiring a revision of the images classified as uninfected by a physician before making the final clinical decision.

Thank you!
Roula Krayem
Roula.Krayem@outlook.com