



THE UNIVERSITY OF
CHICAGO

Health Analytics University of Chicago

Medical Image Analysis

Prepared for: Arnab Bose

March 2019

Submitted by: Ashish Mahadik

Image Data and VGG Model Setup

Total 13780 parasitized images and 13780 uninfected images were downloaded in cell_images folder

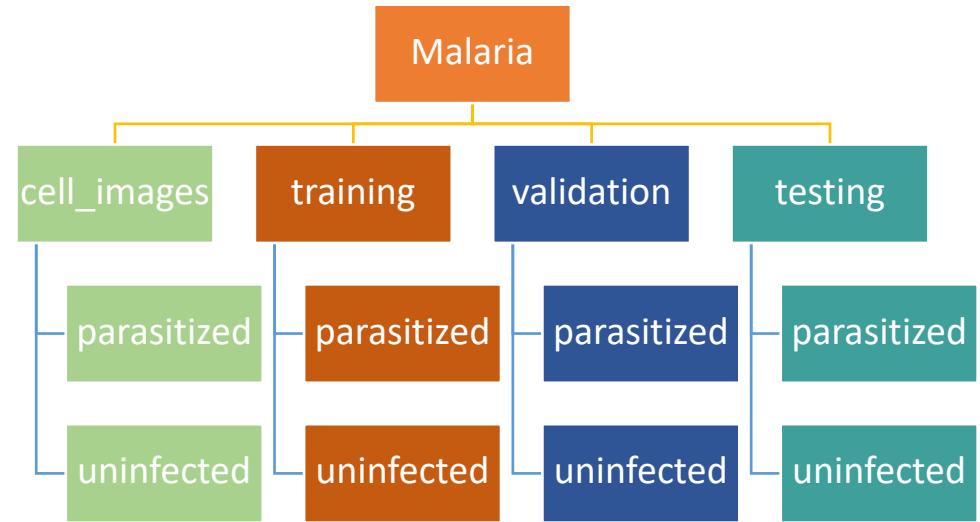
Images dataset was downloaded from cell_images.zip from NIH: <https://ceb.nlm.nih.gov/repositories/malaria-datasets/> ([Links to an external site.](#))[Links to an external site.](#)

Build_datasets.py program separated images in following structure

- 19842 images into training folder
- 2204 images into validation folder
- 5512 images into testing folder

Folder names were used as class label in classification model

Various classification model were used on top of nontrainable VGG layers



```
for layer in model_vgg19.layers:  
    layer.trainable = False  
  
x = model_vgg19.output  
x = GlobalMaxPooling2D()(x)  
x = BatchNormalization()(x)  
x= ELU()(x)  
x = Dense(256)(x,Activation = 'relu')(x)  
x= ELU(alpha=1.0)(x)  
x = Dense(2, activation='softmax')(x)
```

VGG16 model with Custom top layer - Relu Activation

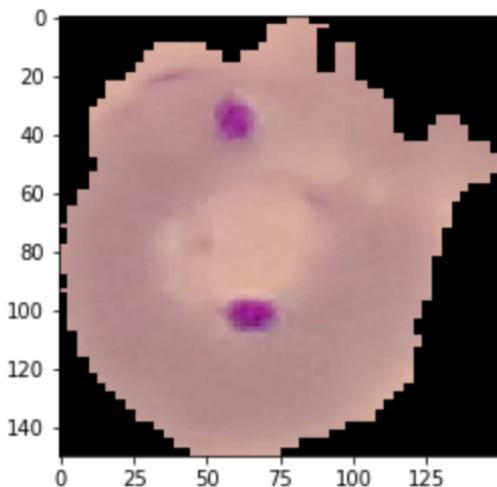
```
batch_normalization_22 (Batch Normalization)      2048  
flatten_6 (Flatten)                            (None, 8192)      0  
dense_78 (Dense)                             (None, 256)       2097408  
dropout_52 (Dropout)                          (None, 256)       0  
dense_79 (Dense)                             (None, 2)         514  
=====  
Total params: 16,814,658  
Trainable params: 2,098,946  
Non-trainable params: 14,715,712
```

Batch Normalization
Dense 256 with Relu activation
Dropout 0.5
Dense 2 with softmax activation

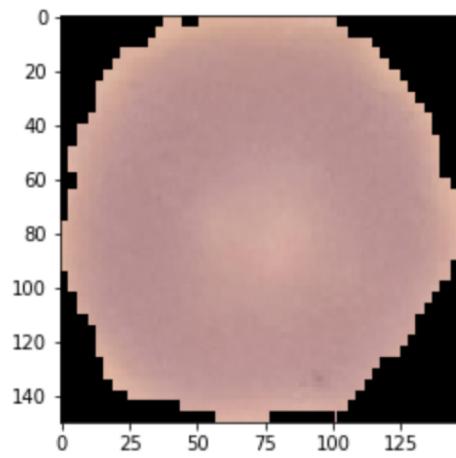
Accuracy on test images – 93.81 %

	precision	recall	f1-score	support
Parasitized	0.98	0.90	0.93	2726
Uninfected	0.91	0.98	0.94	2786
micro avg	0.94	0.94	0.94	5512
macro avg	0.94	0.94	0.94	5512
weighted avg	0.94	0.94	0.94	5512

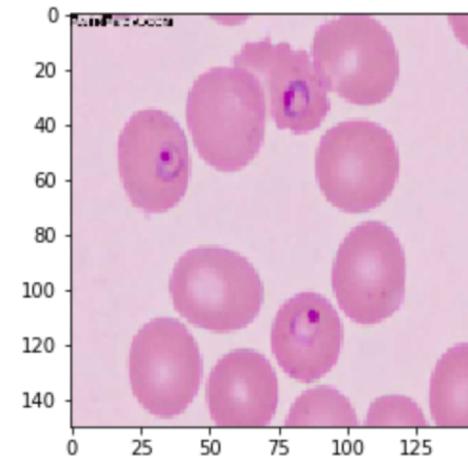
VGG16 Performance On Out Of Sample Images



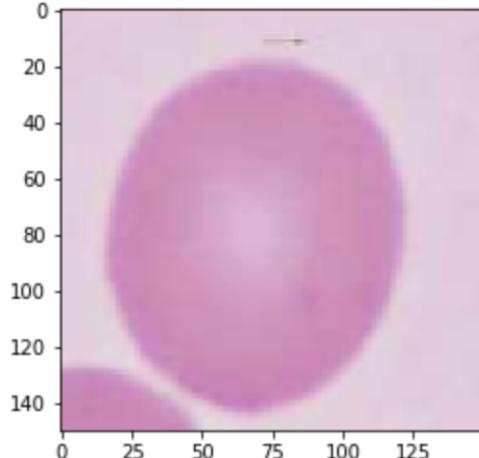
Parasitized



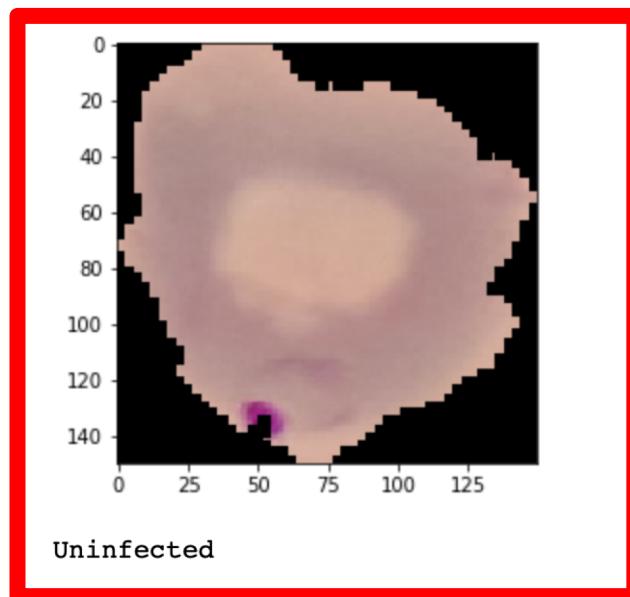
Uninfected



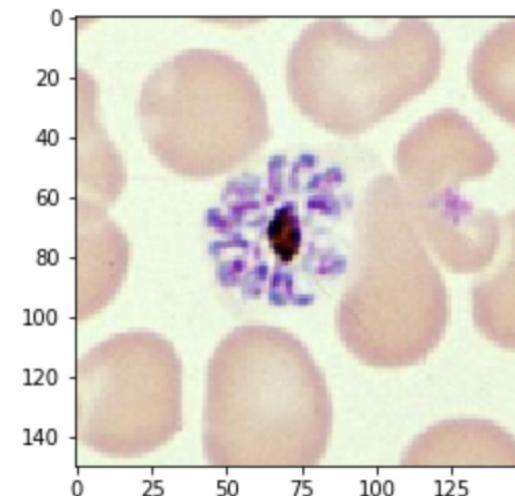
Parasitized



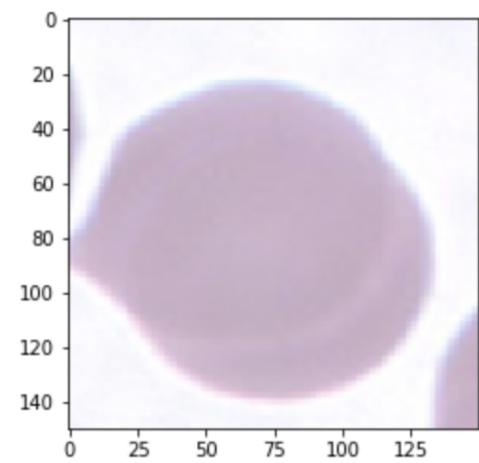
Uninfected



Uninfected



Parasitized



Uninfected

VGG19 model with Custom top layer – ELU activation

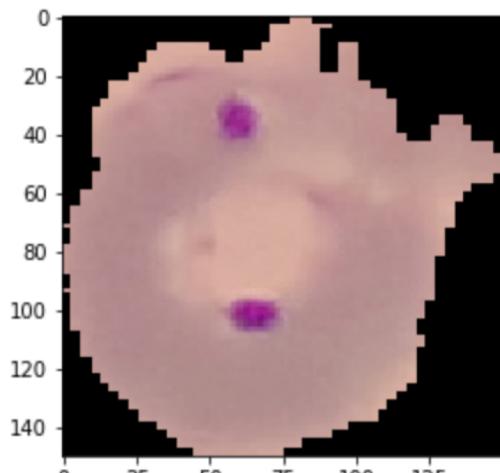
batch_normalization_24 (Batch Normalization)	(None, 4, 4, 512)	2048
flatten_8 (Flatten)	(None, 8192)	0
dense_82 (Dense)	(None, 256)	2097408
elu_2 (ELU)	(None, 256)	0
dropout_54 (Dropout)	(None, 256)	0
dense_83 (Dense)	(None, 2)	514
<hr/>		
Total params: 22,124,354		

Accuracy on test images – 92.85 %

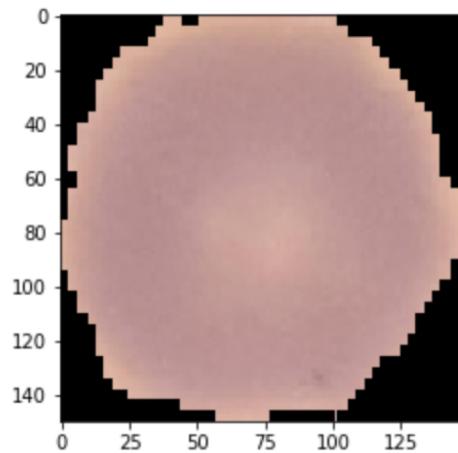
		precision	recall	f1-score	support
Batch Normalization					
Dense 256 with ELU activation	Parasitized	0.95	0.90	0.93	2726
Dropout 0.5	Uninfected	0.91	0.96	0.93	2786
Dense 2 with softmax activation	micro avg	0.93	0.93	0.93	5512
	macro avg	0.93	0.93	0.93	5512
	weighted avg	0.93	0.93	0.93	5512

ELU – Exponential Linear Unit

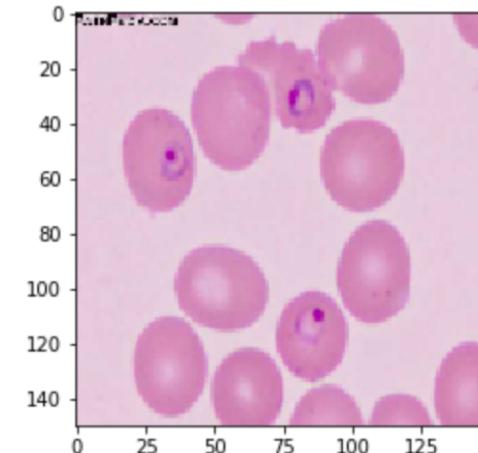
VGG19 Performance using Out Of Sample Images



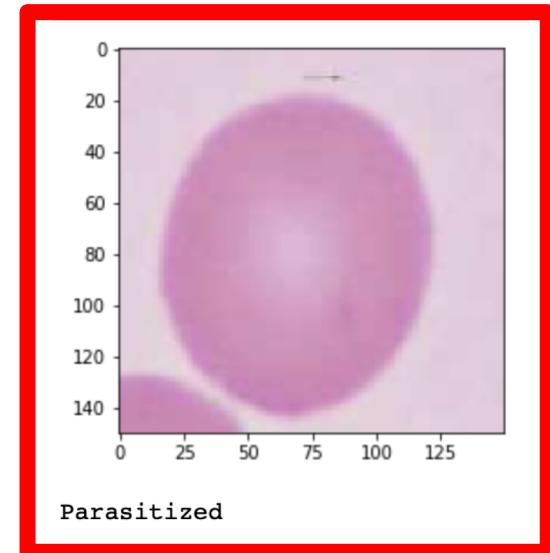
Parasitized



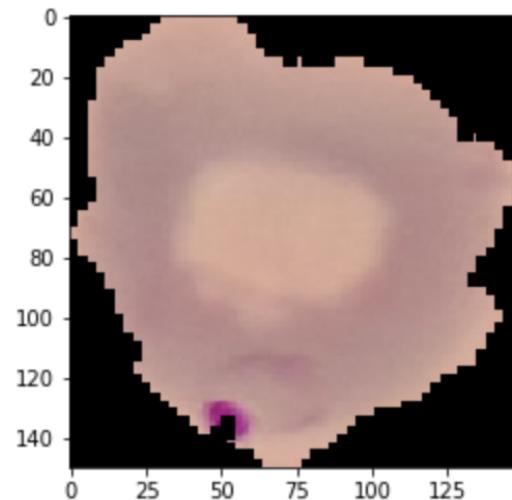
Uninfected



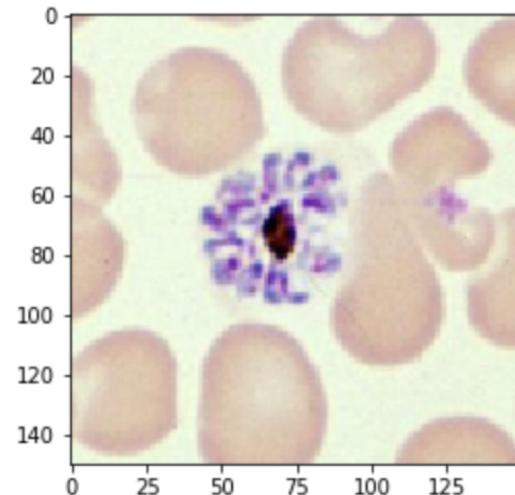
Parasitized



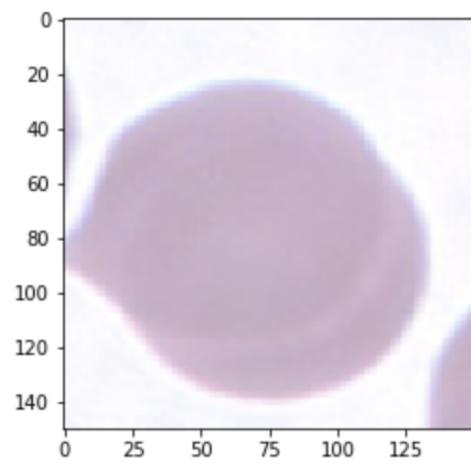
Parasitized



Parasitized



Parasitized



Uninfected

Key Learnings While Tuning Model Parameters

- Model performance improved after changing learning rate from 0.1 to 0.01
- Model performance improved by changing image size from **64* 64 to 128*128 to 150*150 to 224*224**
- Training slows down by changing from **64* 64 to 128*128 to 150*150 to 224*224**
- I settled for image size of 150 *150 for balance between accuracy and speed
- Changing model from VGG16 to VGG19 did not improve performance significantly
- Changing activation from Relu a ELU (exponential linear activation) did not improve performance significantly but helped in detecting dots towards edge.

Appendix

- References:
 - <https://www.pyimagesearch.com/2018/12/03/deep-learning-and-medical-image-analysis-with-keras/>
 - <https://ceb.nlm.nih.gov/repositories/malaria-datasets/> (Links to an external site.)Links to an external site.