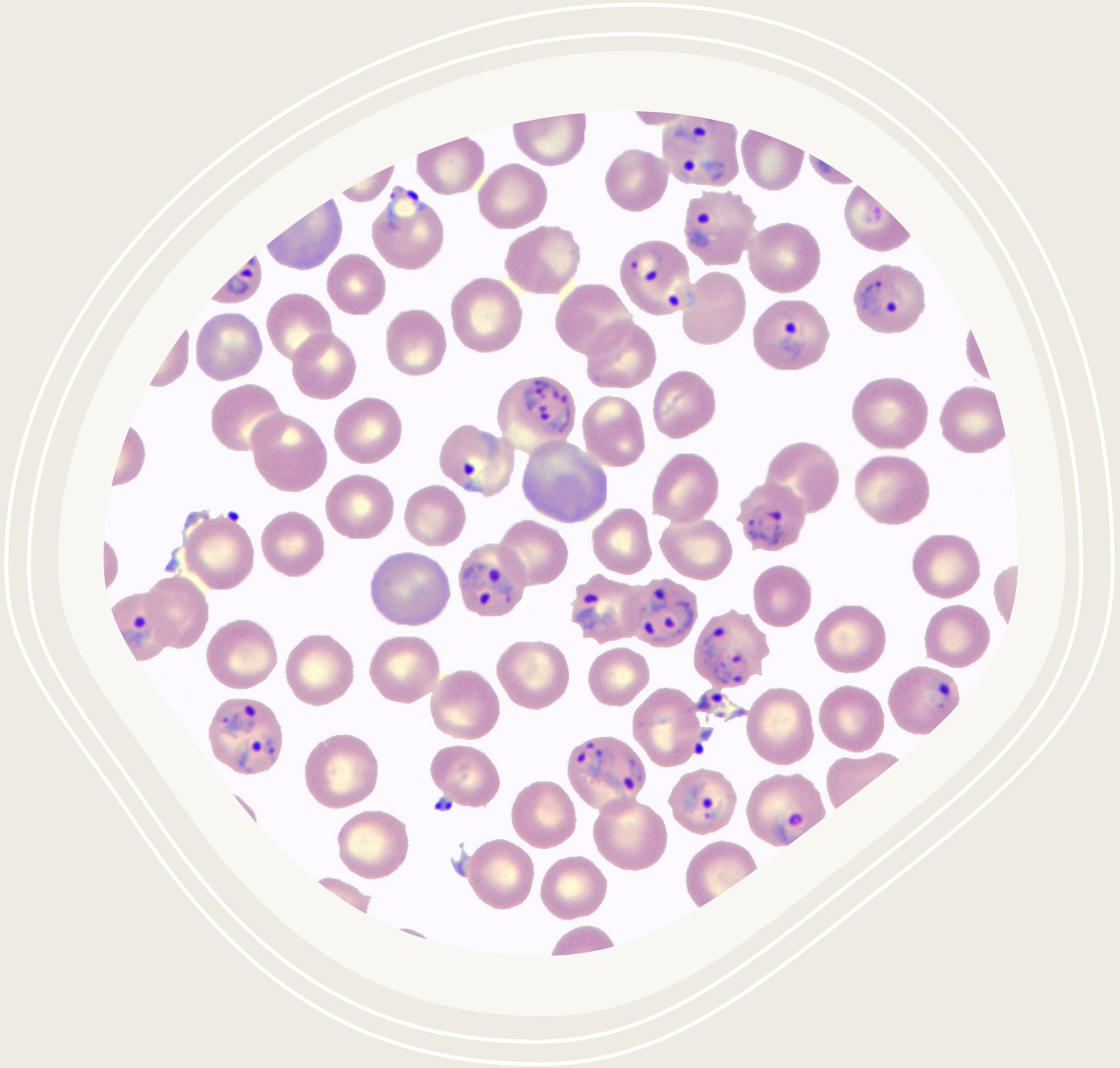


Malaria Detector CNN

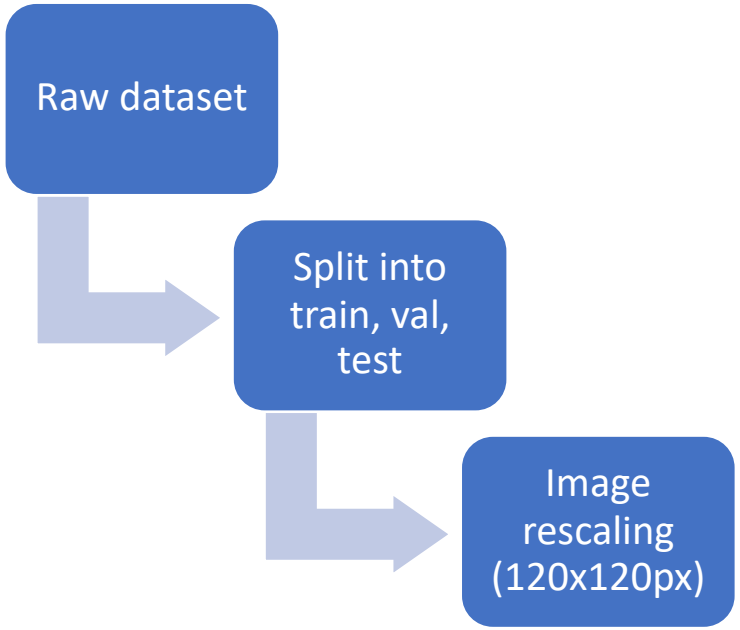
Computational Intelligence and Deep Learning

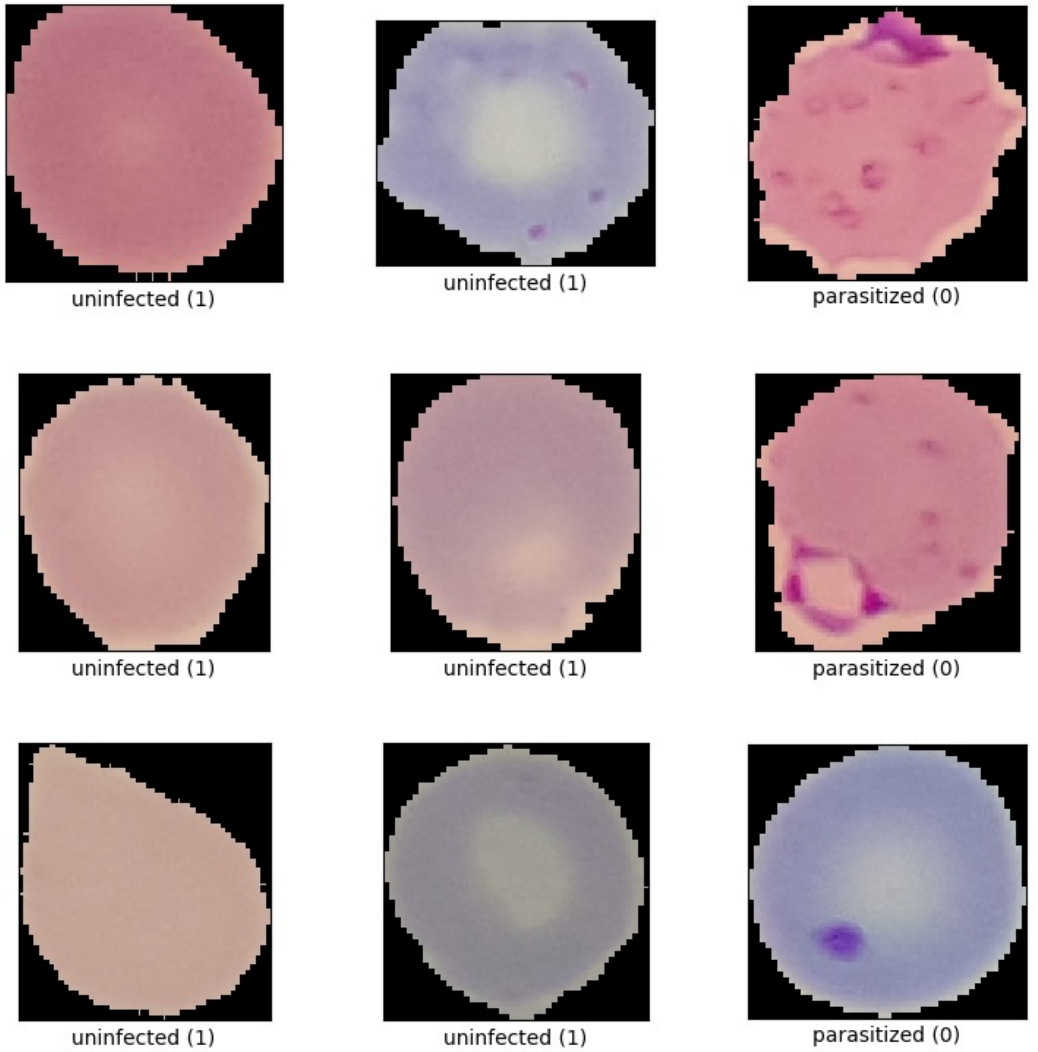
A.A. 2022-2023

Francesca Pezzuti, Pietro Tempesti

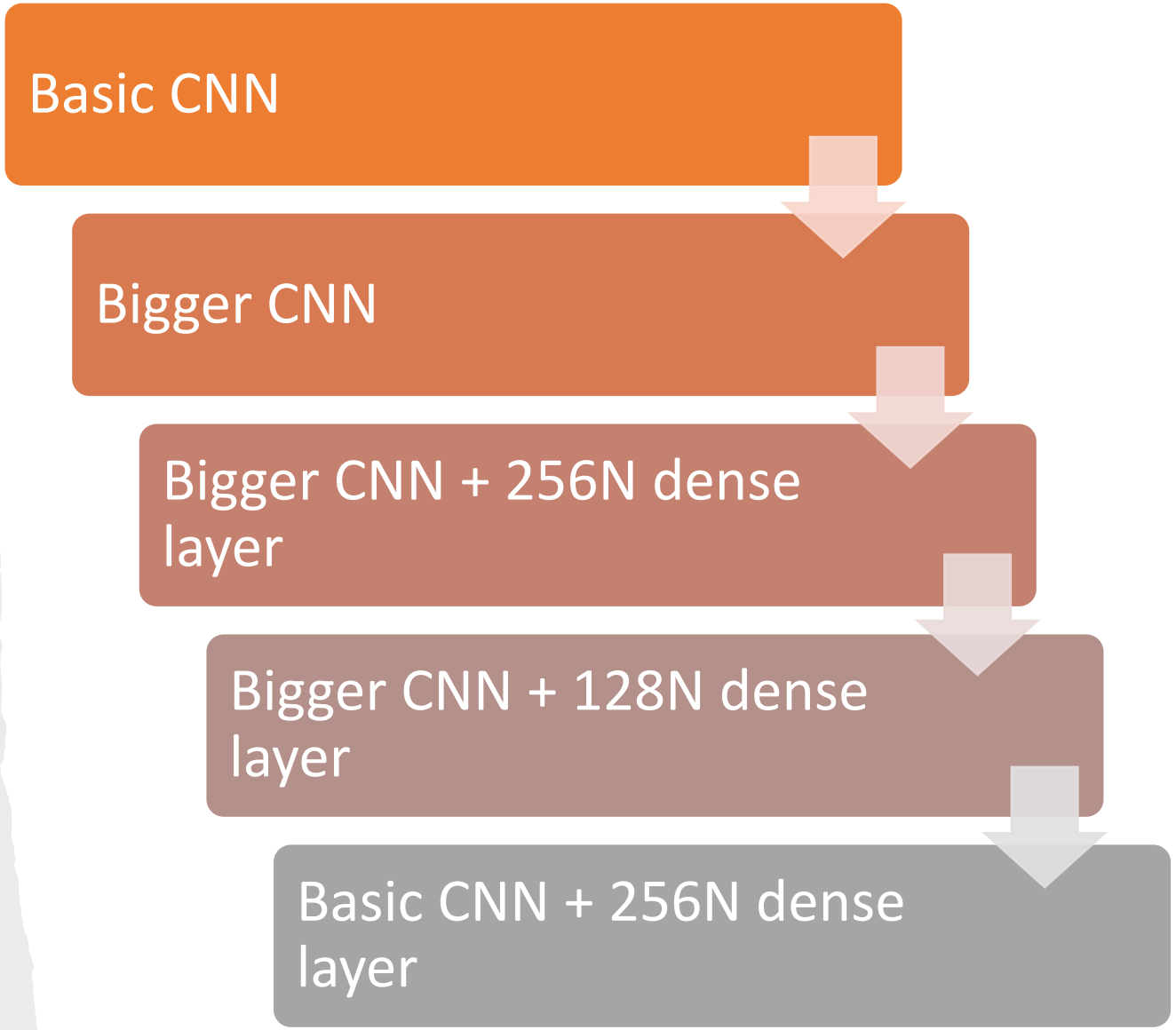
Malaria dataset

- 2 classes: uninfected & parasitized
- 27558 images, perfectly balanced
- Weight: 317.62MB





CNN from scratch



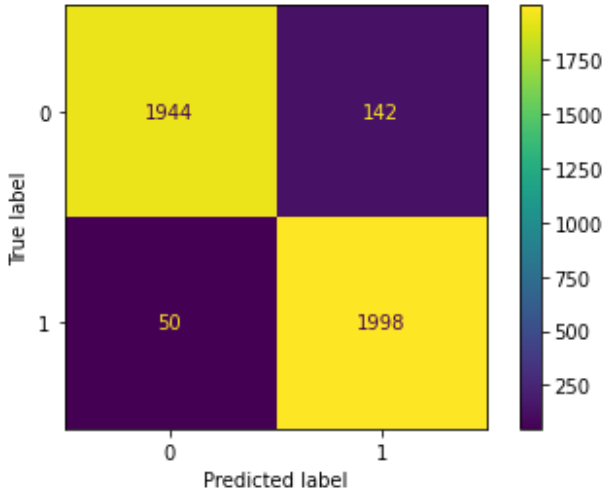
Best architecture: Bigger CNN + 256N Dense Layer

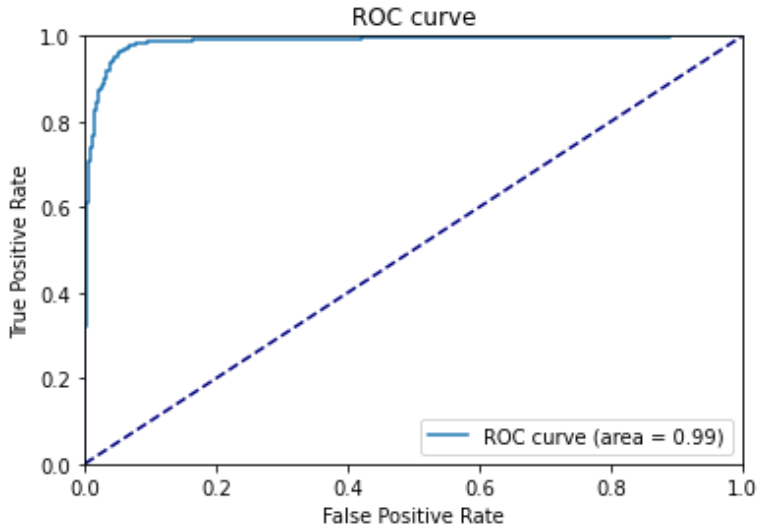
Architecture

```
Model: "256denseCNN"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_3 (InputLayer)        [(None, 120, 120, 3)]     0

 conv2d_7 (Conv2D)           (None, 118, 118, 32)      896

 max_pooling2d_7 (MaxPooling  (None, 59, 59, 32)       0
 2D)

 conv2d_8 (Conv2D)           (None, 57, 57, 64)        18496

 max_pooling2d_8 (MaxPooling  (None, 28, 28, 64)       0
 2D)

 conv2d_9 (Conv2D)           (None, 26, 26, 128)       73856

 max_pooling2d_9 (MaxPooling  (None, 13, 13, 128)      0
 2D)

 conv2d_10 (Conv2D)          (None, 11, 11, 256)       295168

 max_pooling2d_10 (MaxPoolin  (None, 5, 5, 256)        0
 g2D)

 flatten_2 (Flatten)         (None, 6400)              0

 dense_2 (Dense)             (None, 256)               1638656

 dense_3 (Dense)             (None, 1)                 257

=================================================================
Total params: 2,027,329
Trainable params: 2,027,329
Non-trainable params: 0
_________________________________________________________________
```

Results

	Precision	Recall	F1-Score	Support
Parasitized	0.9749	0.9319	0.9529	2086
Uninfected	0.9336	0.9756	0.9542	2048
Accuracy			0.9536	4134
Macro Avg	0.9543	0.9538	0.9535	4134
Weighted Avg	0.9545	0.9536	0.9535	4134





Hyper-Parameters Optimization

Hyper Parameters:

Activation Functions:

- Tanh, **ReLU**

Units in the dense layer:

- 128 to 256 → **160**

Learning Rate:

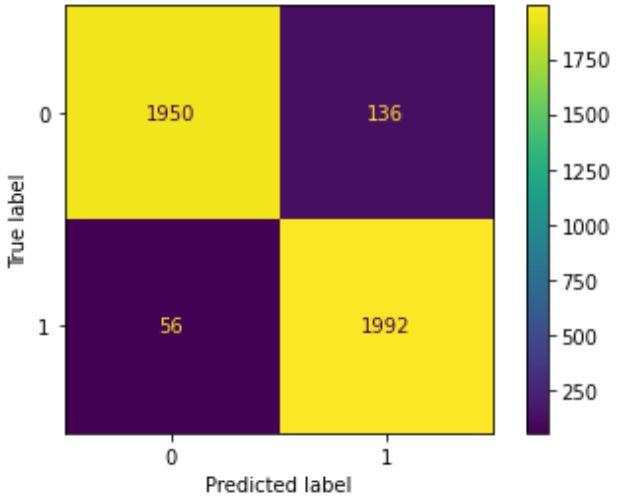
- 0.01, **0.001**, 0.0001

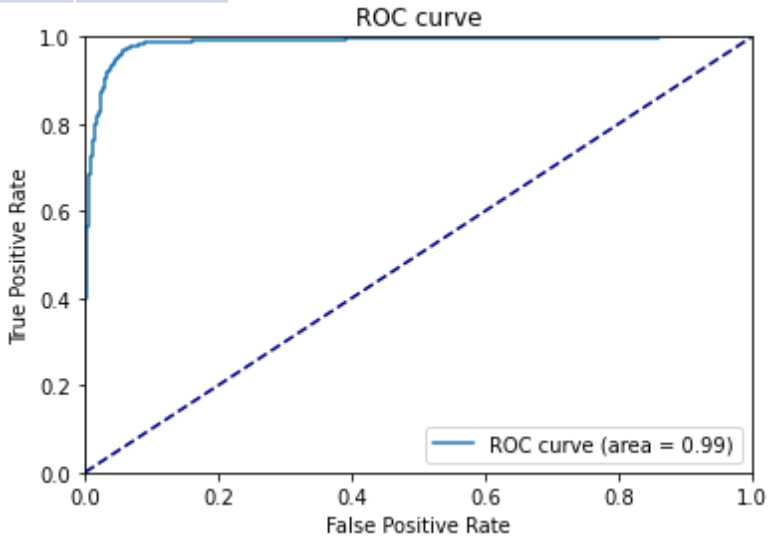
Algorithm:

HyperBand with 3 iterations

Training results:

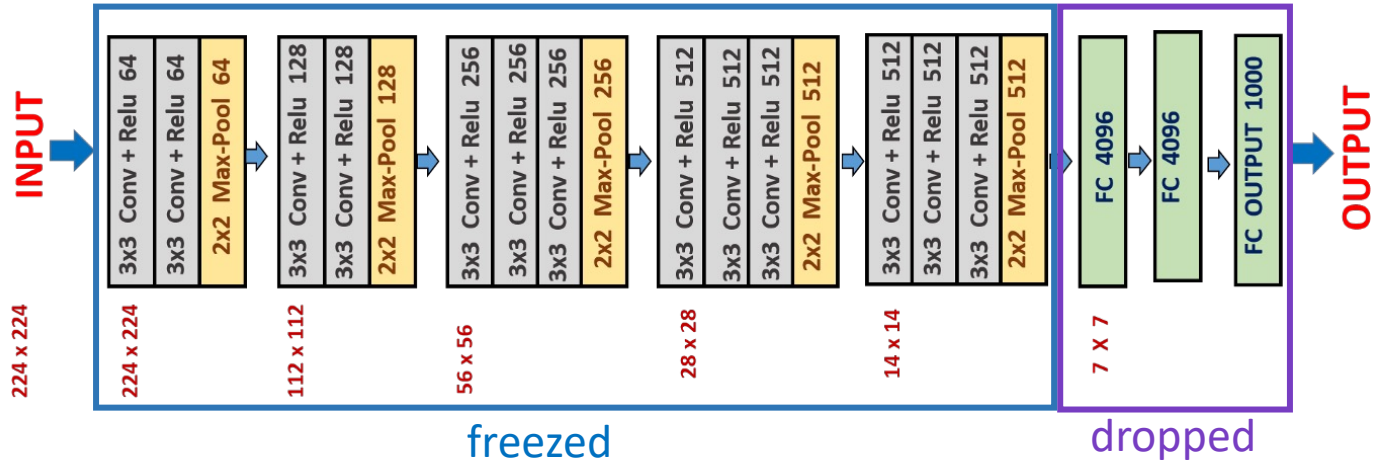
	Precision	Recall	F1-Score	Support
Parasitized	0.9721	0.9348	0.9531	2086
Uninfected	0.9361	0.9727	0.9540	2048
Accuracy			0.9536	4134
Macro Avg	0.9541	0.9537	0.9536	4134
Weighted Avg	0.9543	0.9536	0.9535	4134



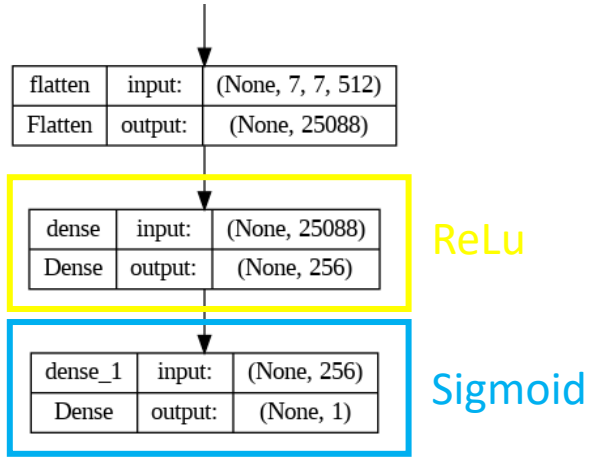


VGG16 – Feature Extraction

VGG16's architecture:

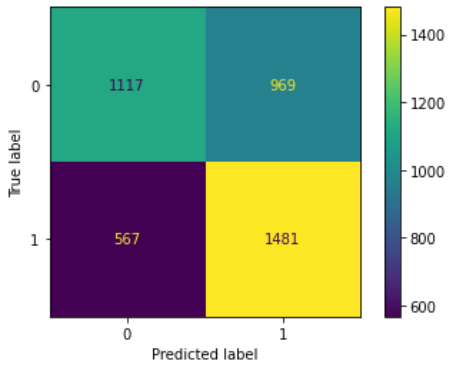


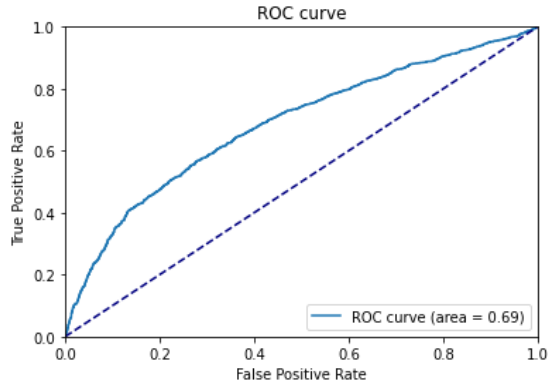
Top classifier's best architecture:



Training results:

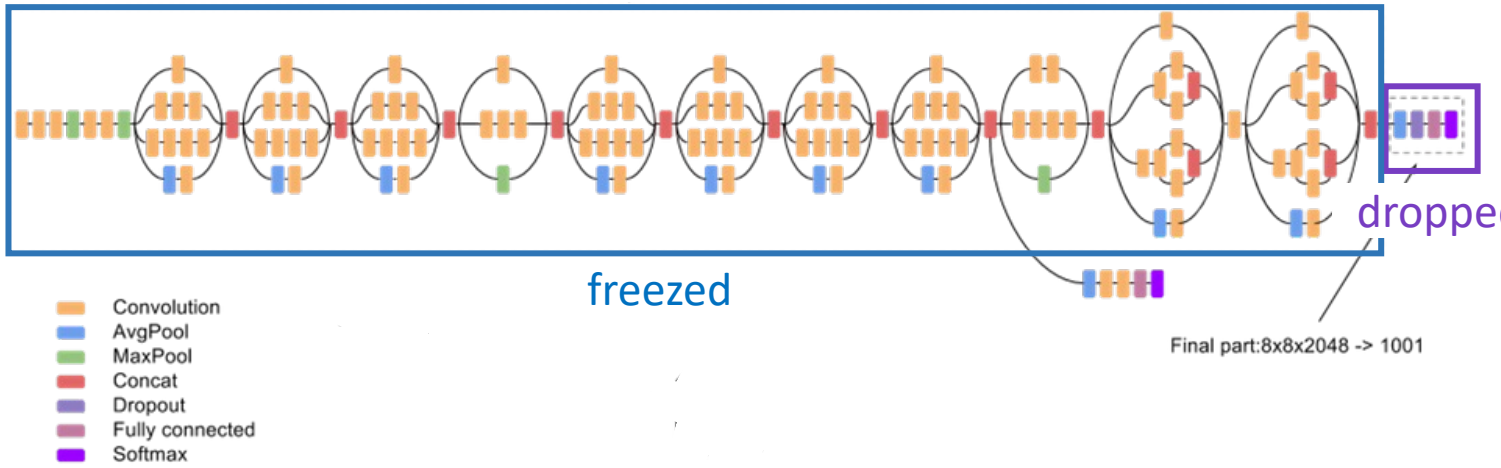
	Precision	Recall	F1-Score	Support
Parasitized	0.6633	0.5355	0.5926	2086
Uninfected	0.6045	0.7231	0.6585	2048
Accuracy			0.6284	4134
Macro Avg	0.6339	0.6293	0.6255	4134
Weighted Avg	0.6342	0.6284	0.6252	4134



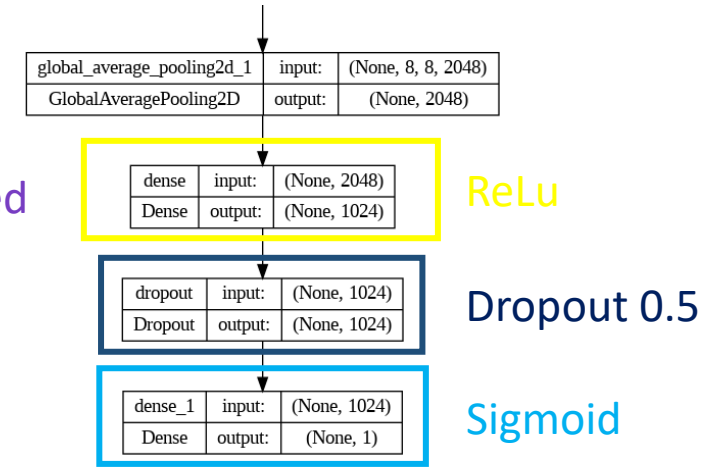


InceptionV3 – Feature Extraction

InceptionV3's architecture:

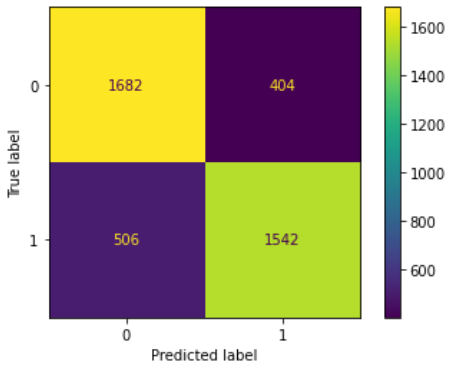


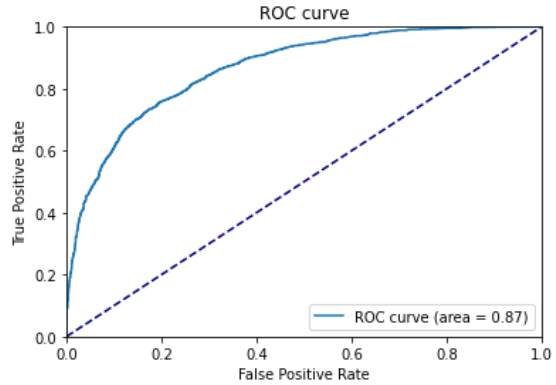
Top classifier's best architecture:



Training results:

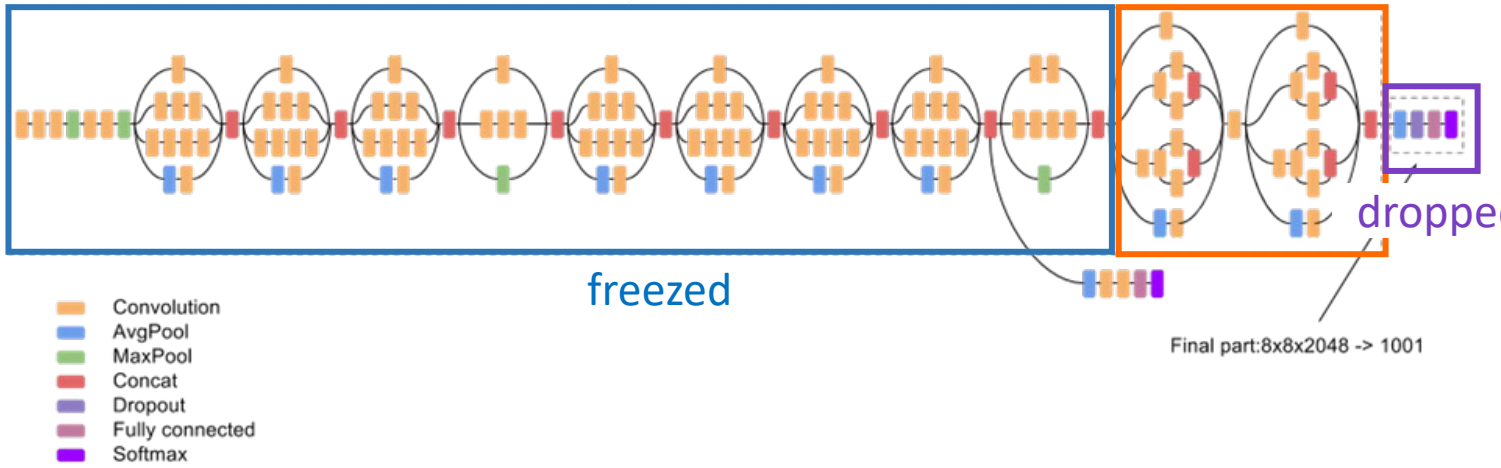
	Precision	Recall	F1-Score	Support
Parasitized	0.7687	0.8063	0.7871	2086
Uninfected	0.7924	0.7529	0.7722	2048
Accuracy			0.7799	4134
Macro Avg	0.7806	0.7796	0.7796	4134
Weighted Avg	0.6805	0.7799	0.7797	4134



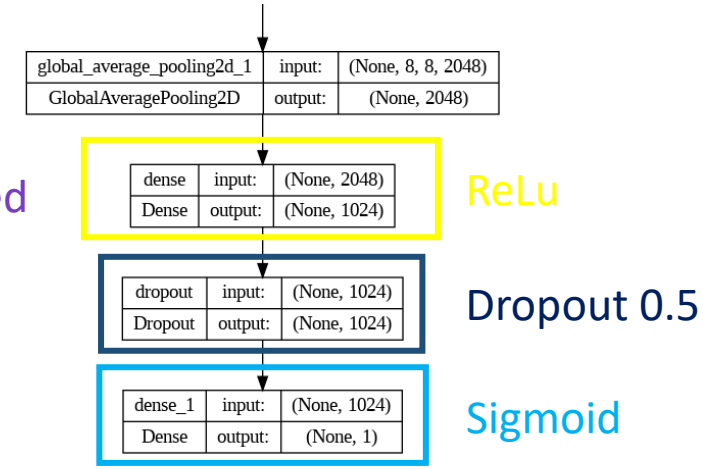


InceptionV3 – Fine tuning

InceptionV3's architecture:

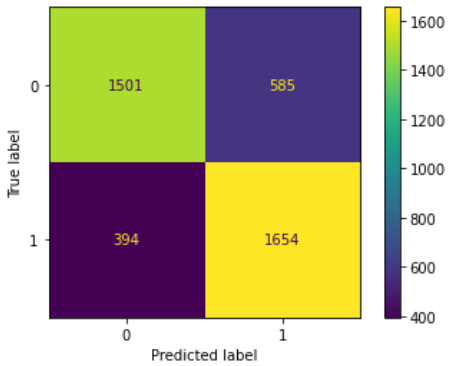


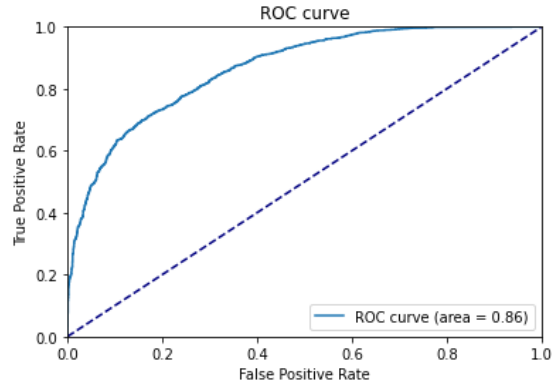
Top classifier's best architecture:



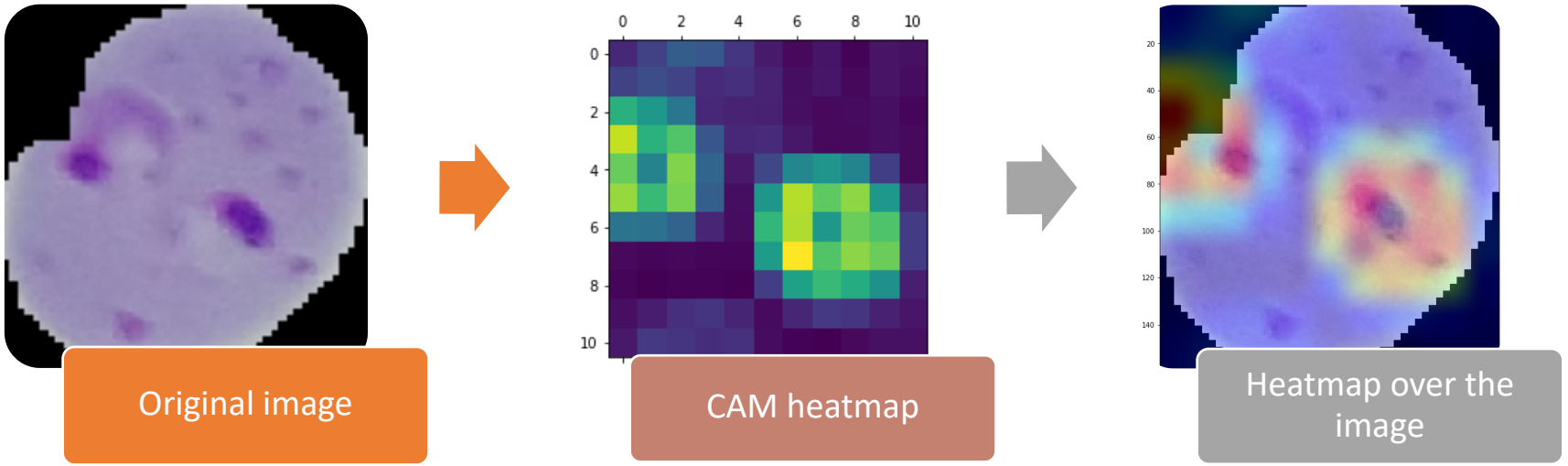
Training results:

	Precision	Recall	F1-Score	Support
Parasitized	0.7921	0.7196	0.7541	2086
Uninfected	0.7387	0.8076	0.7716	2048
Accuracy			0.7632	4134
Macro Avg	0.7654	0.7636	0.7629	4134
Weighted Avg	0.7656	0.7632	0.7628	4134

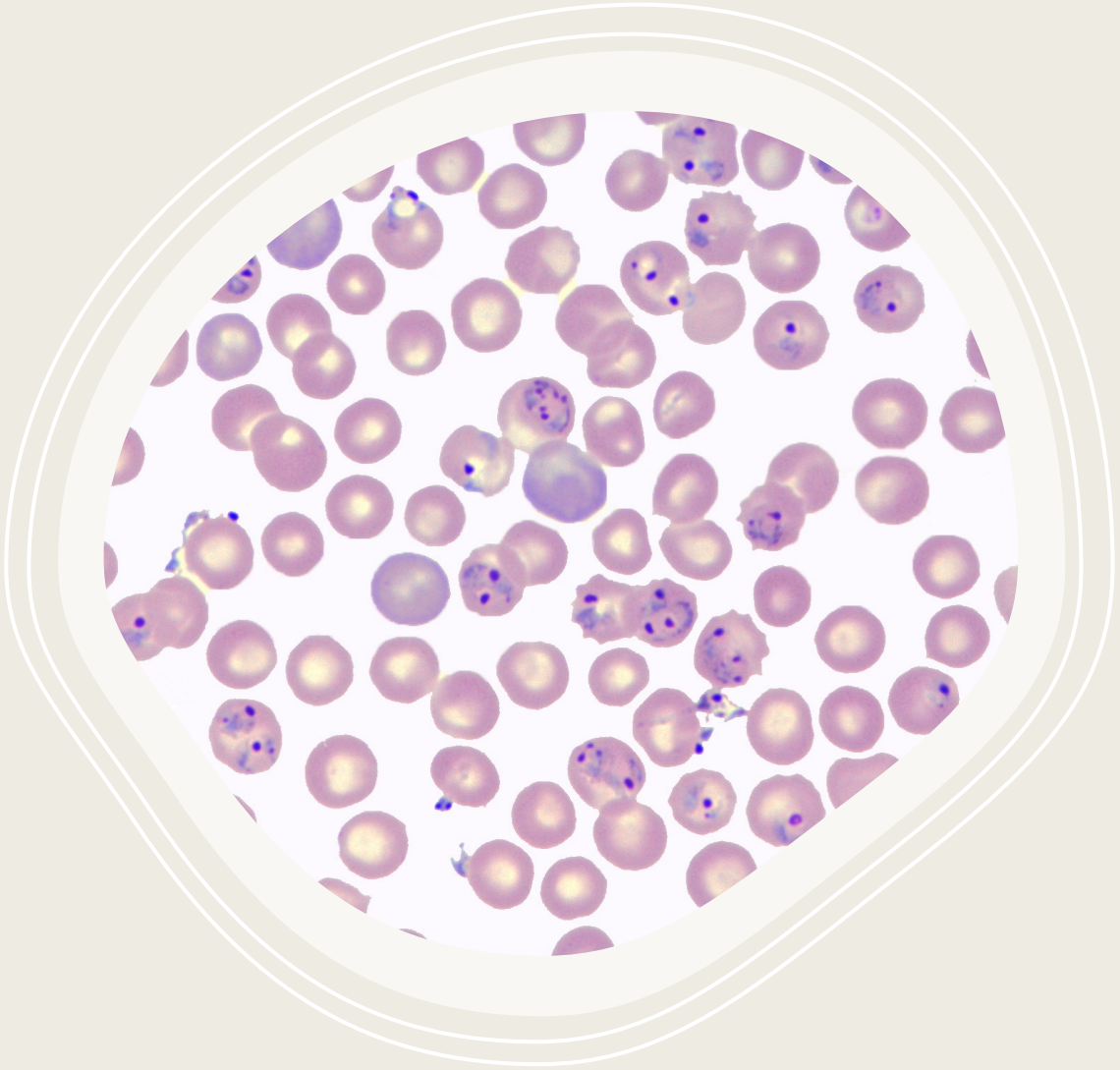




XAI – Class Activation Map visualization



- CAM visualization is a popular technique for interpreting the decisions of CNNs and can help identify which parts of an image the network is attending to for classification.
- In a medical scenario this is really important to detect possible misclassifications, or to better address some treatment directly to the infected parts of our blood cells.



Conclusions:

- **CNN from scratch**'s results are comparable to state-of-the-art approaches.
  - The optimized CNN has competitive performances despite its simple architecture.
- **XAI experiments** allowed us to learn a possible classification pattern exploitable by medical experts to better assess their diagnoses.
- **Pre-trained models**' results are not satisfactory:
  - the poor performances may be due to the fact that our dataset is composed by very simple and low-definition images, very different from the training base of these models.
- A further improvement of our study may be to leverage other pre-trained models.

Thanks for your attention!