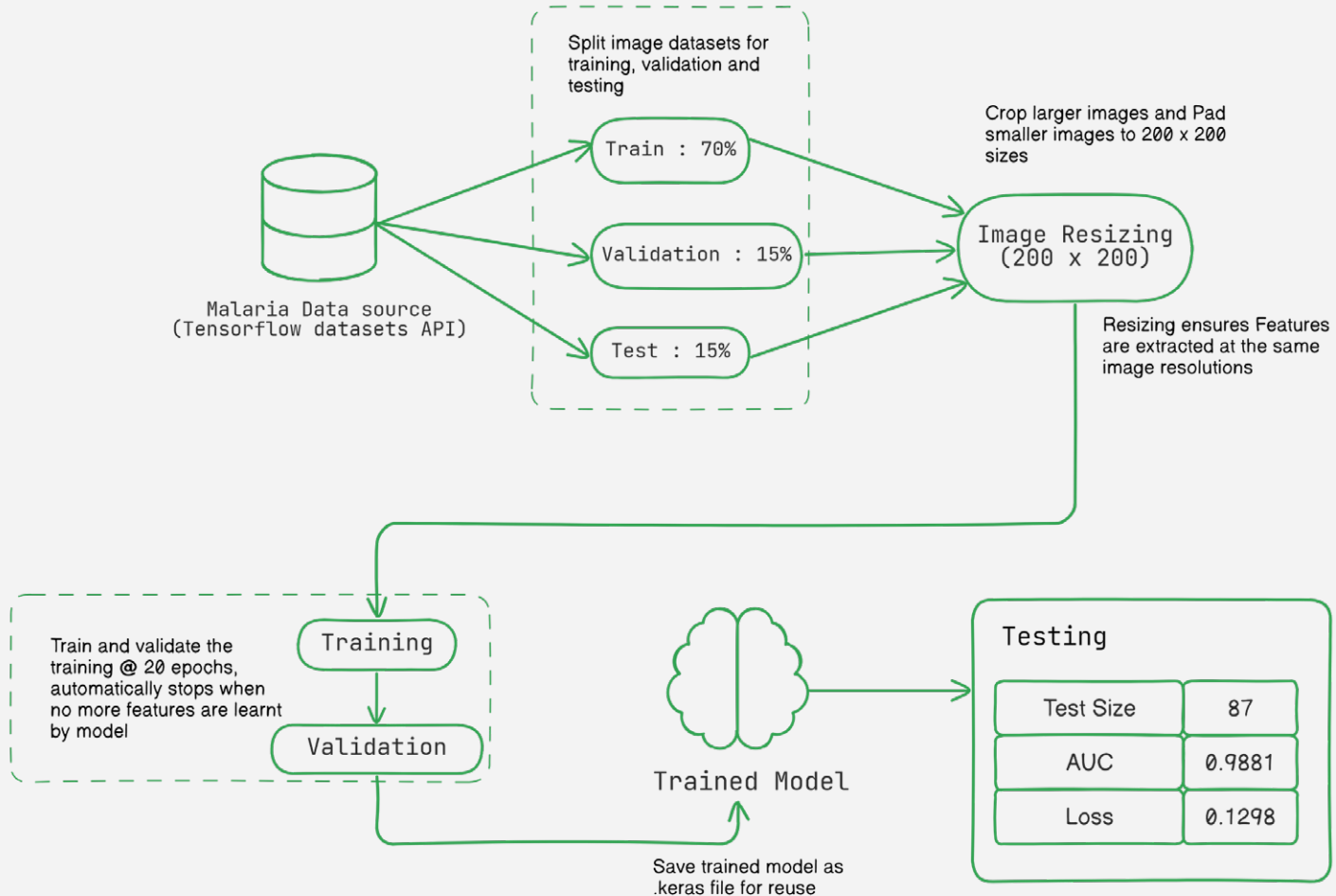# Author : Allan Wanjala Wafula

CEMA

Task:

TensorFlow is an open-source machine learning framework developed by Google for the sole purpose of building and training machine learning models. One of the datasets under the Tensorflow Image libraries is a malaria dataset which contains approximately 27,500 cell images of parasitized and uninfected cells from thin blood smear slide images of segmented cells.

Using this dataset, we would like you to create a model which is able to classify whether a blood smear is uninfected or parasitized.

Split image datasets for
training, validation and
testing

Train : 70%

Validation : 15%

Test : 15%

Malaria Data source
(Tensorflow datasets API)

Crop larger images and Pad
smaller images to 200 x 200
sizes

Image Resizing
(200 x 200)

Resizing ensures Features
are extracted at the same
image resolutions

Train and validate the
training @ 20 epochs,
automatically stops when
no more features are learnt
by model

Training

Validation

Trained Model

Save trained model as
.keras file for reuse

Testing	
Test Size	87
AUC	0.9881
Loss	0.1298

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 200, 200, 16)	448
conv2d_1 (Conv2D)	(None, 200, 200, 16)	2,320
max_pooling2d (MaxPooling2D)	(None, 100, 100, 16)	0
sequential (Sequential)	(None, 50, 50, 32)	2,160
sequential_1 (Sequential)	(None, 25, 25, 64)	7,392
sequential_2 (Sequential)	(None, 12, 12, 128)	27,072
dropout (Dropout)	(None, 12, 12, 128)	0
sequential_3 (Sequential)	(None, 6, 6, 256)	103,296
dropout_1 (Dropout)	(None, 6, 6, 256)	0
flatten (Flatten)	(None, 9216)	0
sequential_4 (Sequential)	(None, 512)	4,721,152
sequential_5 (Sequential)	(None, 128)	66,176
sequential_6 (Sequential)	(None, 64)	8,512
dense_3 (Dense)	(None, 1)	65