Data Appendix
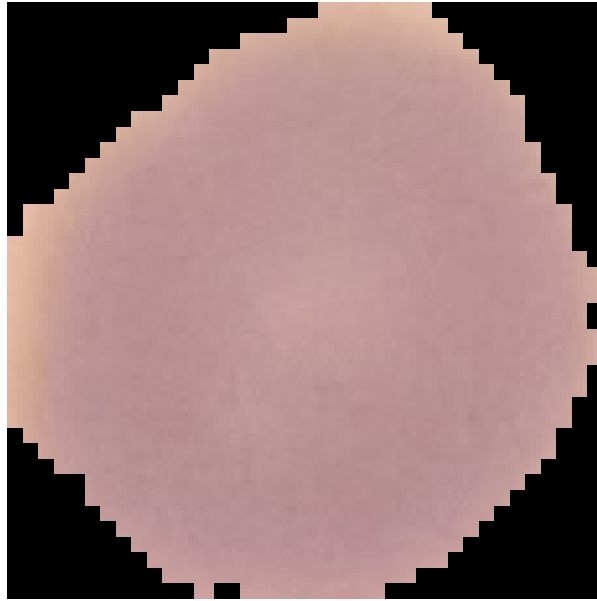**DATA download/ Main Data Set**
From the download link there are few folders when you unzip the file. The cell_images Folder is the first folder you see and inside those folders there is a cell_images, which contains all the images of the cells, parasitized which contains images only of parasitized cells and finally Uninfected which contains images of uninfected cells. The actual images are PNG files. Below are examples of images that are located in the folders one parasitized and one uninfected.

Parasitized   Uninfected
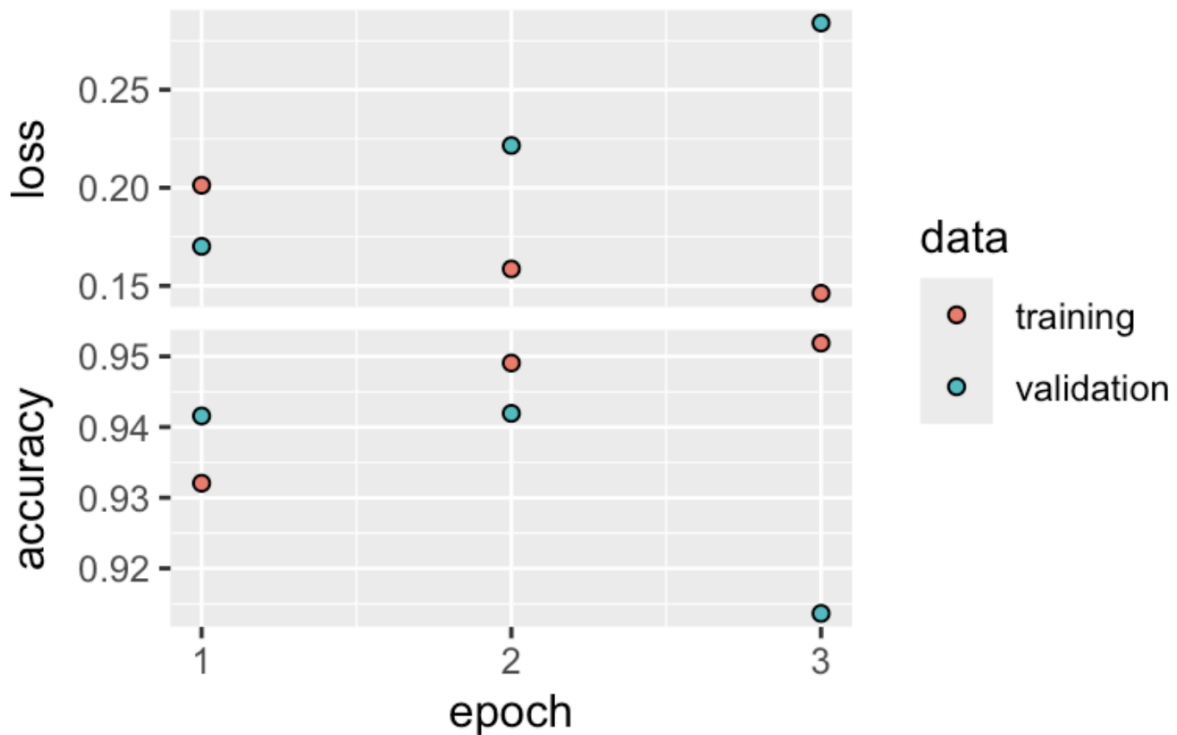
Data Dictionary

THis is the data dictionary used for the main data set. It is very basic because the image data doesn't require a lot of extra information.

Column Name	Description	Example
image	Image of the cell	
label	Label of whether the cell is infected or uninfected	Parasitized, Uninfected

Epoch vs Performance

When building the model we went through each epoch to see the performance of the model. The following graph and table describe how the performance of the model improved each epoch for both training and testing sets of data.


```
Epoch 1/3
689/689 [==============================] - 12215s 18s/step - loss: 0.2013 - accuracy: 0.9321 -
val_loss: 0.1701 - val_accuracy: 0.9416
Epoch 2/3
689/689 [==============================] - 13281s 19s/step - loss: 0.1586 - accuracy: 0.9491 -
val_loss: 0.2216 - val_accuracy: 0.9419
Epoch 3/3
689/689 [==============================] - 7800s 11s/step - loss: 0.1462 - accuracy: 0.9519 -
val_loss: 0.2840 - val_accuracy: 0.9136
```