



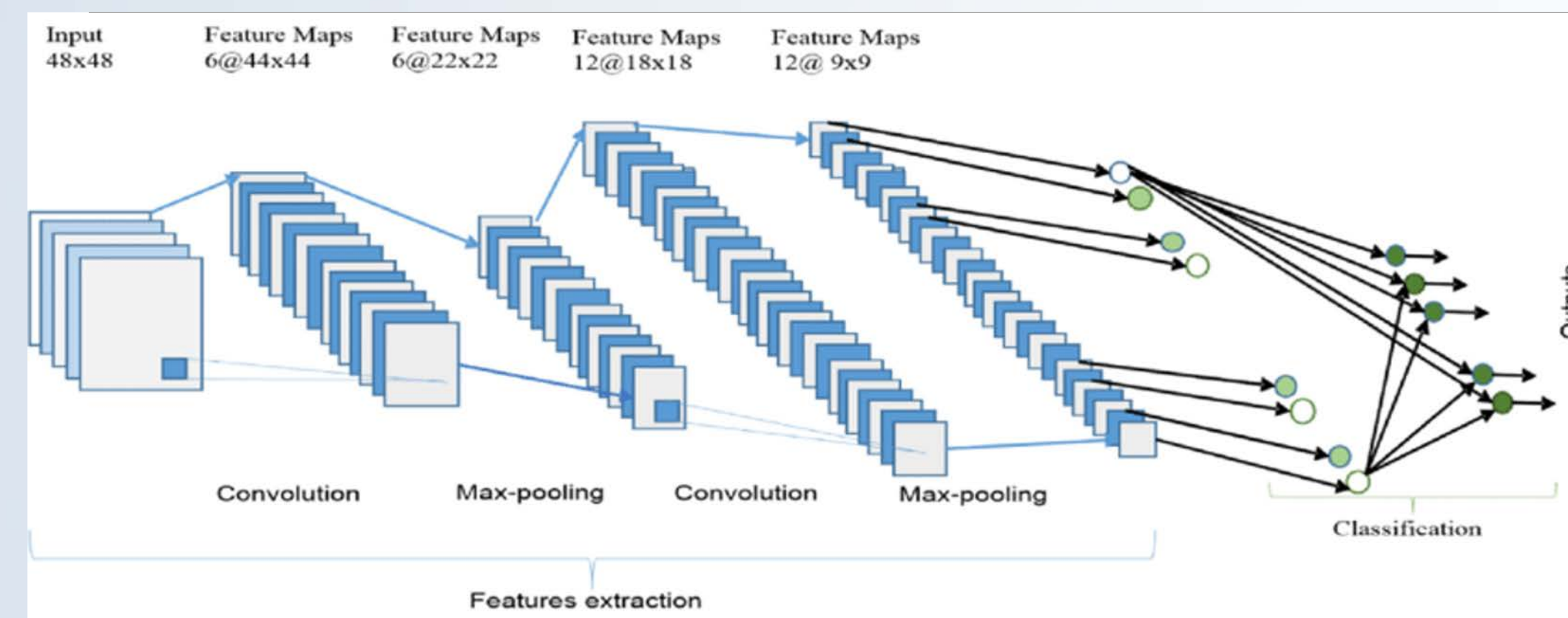
BACKGROUND

Americans are increasingly bicycling to commute, to exercise, or just for fun. The increase in bicycle riders are also means the increase in bicycle related accidents. Saddleye is a new startup company whose goal is to improve bicycle safety by using machine learning and sensors to detect hazards and notify the rider. The model built during this project is called Trail detection. Trail detection will will inform riders where the trail is and if the rider is in the trail. The model is expected to work on various trails, such as asphalt, concrete, dirt, gravel, wet, dry, or off-road.

PURPOSE AND HYPOTHESIS

Convolutional neural networks (CNNs) are a deep learning technique aimed primarily at image recognition. We decided that since we are analyzing images, this would be the best model to choose.

To analyze whether or not the subject was on or off the trail, we used a two-dimensional CNN. We had two convolution layers and two pooling layers with a dense layer. We had 100 epochs of data with early stopping to stop training if the model didn't improve.



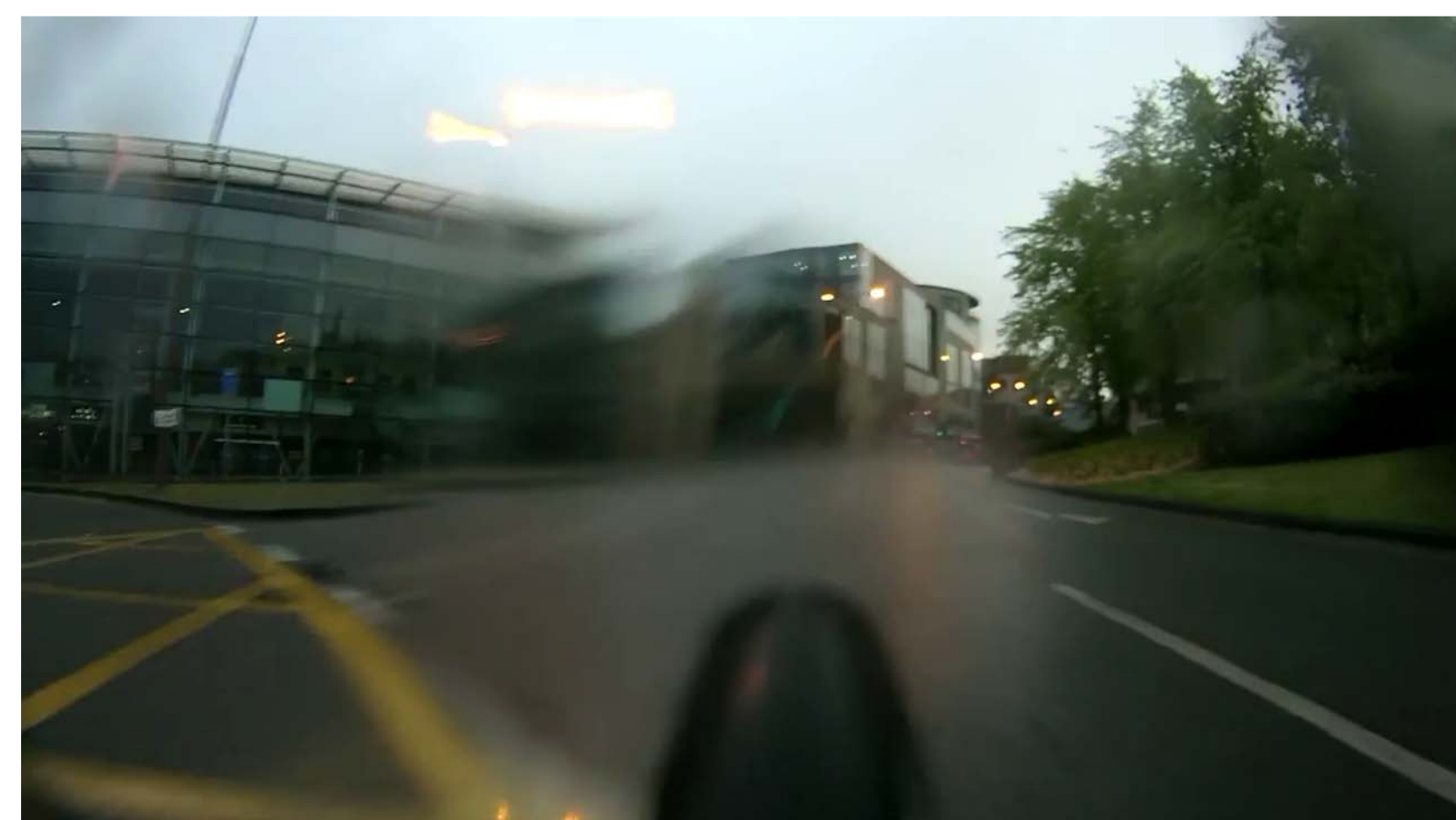
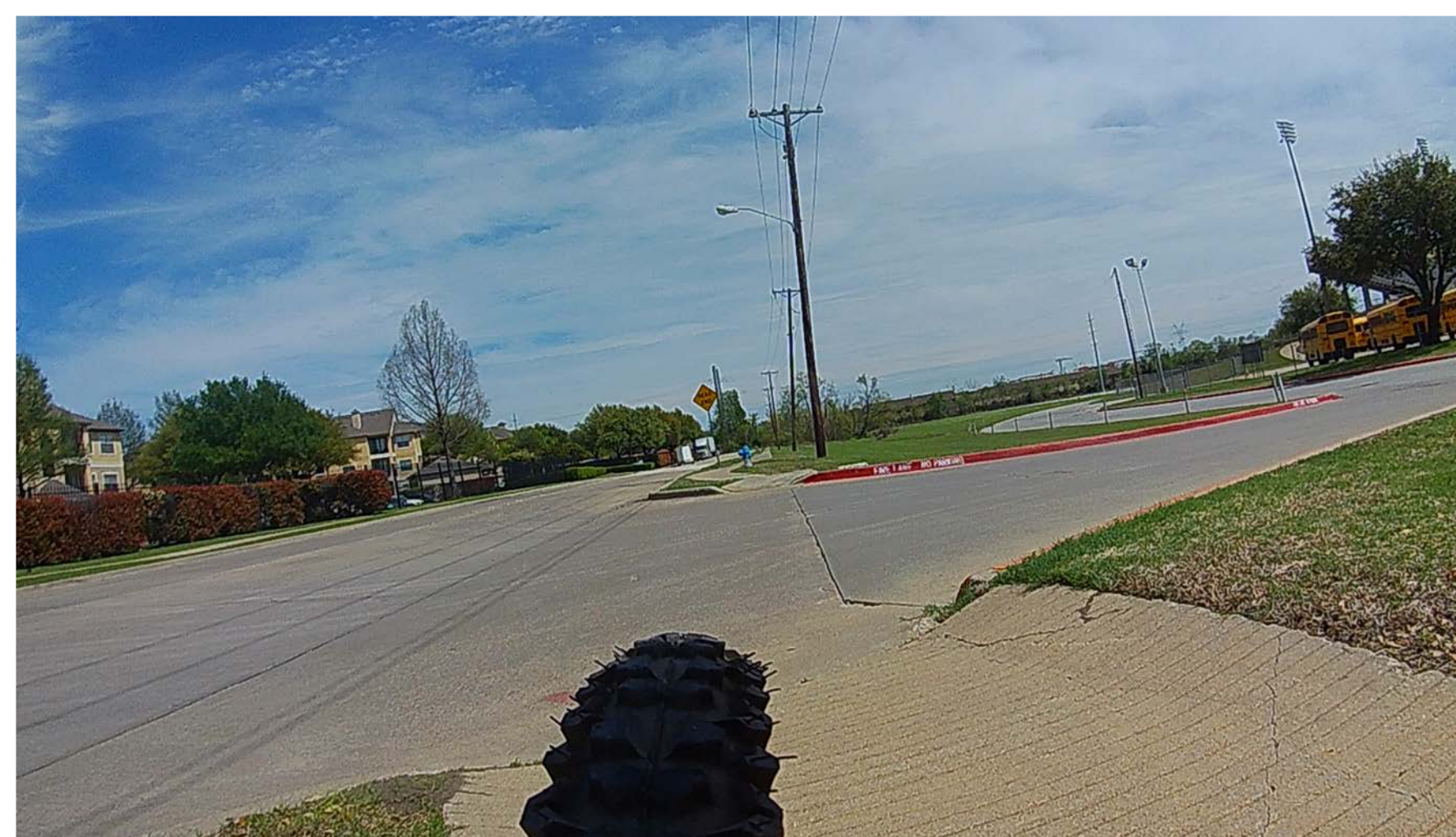
TECHNOLOGIES AND TOOLS

The softwares used throughout the project are listed below

- Python
- TensorFlow
- Keras
- Jetson Nano
- Linux

The hardware used throughout the project are listed below

- Bicycle
- Bean - Jetson Nano

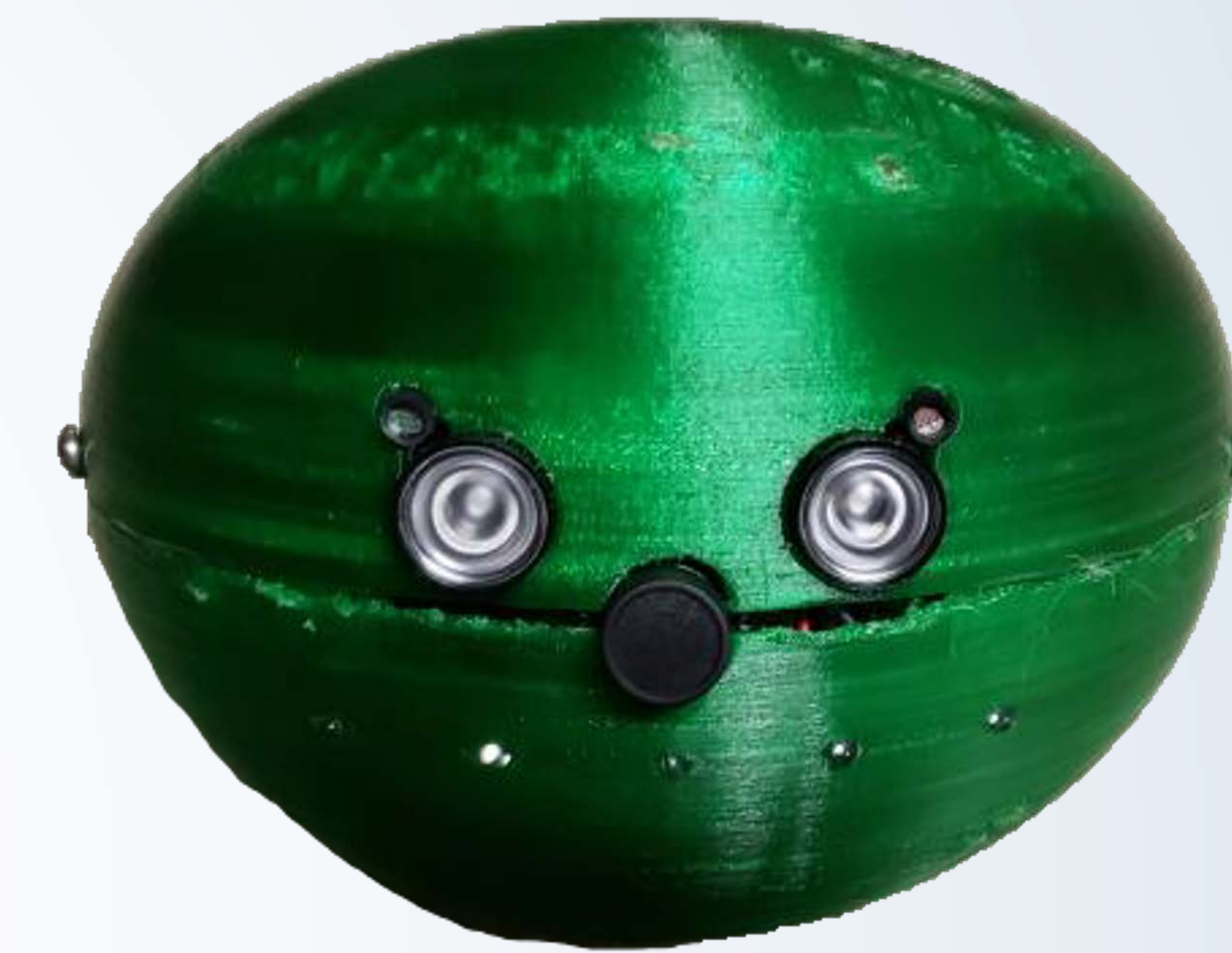


DATASET

Our dataset consists of videos from the rear POV of a bike being ridden. These are videos from various locations, roads, and environments. Variety in riding conditions were limited, as we were not able to get many samples with rain or any samples with snow. The sources of these videos are:

- Videos gathered from riding a bike ourselves
- Videos gathered online

In total, we have 500,000+ samples to train our model on.



CONCLUSIONS



The current design rationale fits the needs of a recreational bicyclist that wants a digital read out of the path they are riding and a monitor of their daily activity for health reasons. The unit small, lightweight, and has efficient battery usage. Data generated from the unit is used to build a model that does the trail identification.

- Current issues include, data generation, classification, and standardization. As two groups we have decided to share data, we are taking steps to make sure we collect and label data in a standardized way. These include sharing all documentation with the other team as well as code. Although we operate on separate Githubs, we share a Discord and Slack for communication, as well as a Google Drive for storage. Another issue that we are facing is building the model, as we have to make sure enough data is collected before we start building the model.

BIBLIOGRAPHY

1. https://drive.google.com/open?id=1uhYdE4NqTACrGMazhxo1qB_ptOQh9YyY
2. https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/customization/custom_training_walkthrough.ipynb