Exercise 2

by

Madeline Zaytsoff

ID: 40177534

Submitted in partial fulfillment of

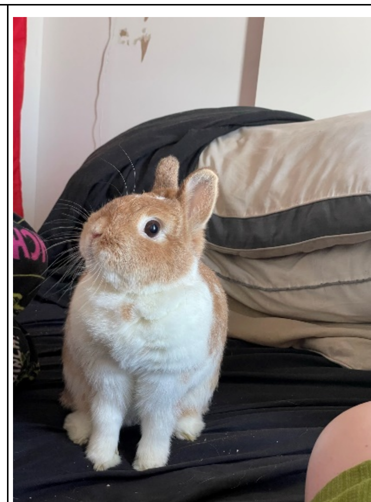
CART 451

Concordia University

Fall 2023

PART A

1. The initial data I used for this assignment consisted of a Pepsi can and a bobble-head figure. Convenience was the main factor in deciding which objects to use as both items were near me at the time of doing this exercise. I added in a third data set consisting of photos of my rabbits to act as a noise. The pop can and the bobblehead each had 51 images taken at various angles whereas the rabbit photos consisted of 10 images taken from various places. I would then manually inspect results to see if there was any misclassifications.

		
Distinct object A: Pop Can	*Distinct object B: Bobblehead*	*Random image: Murder Bun*

2. The purpose of the task was to create a minimally viable sample set of images that Edge Impulse could use to train a data set[1]. In the instance of my assignment, I am looking to have an algorithm that can determine if something is a pop can, a bobblehead, or if it's something else.

3. My first attempt at uploading data failed due to a misunderstanding with the upload parameters. I pre-formatted my images to 96x96 in photoshop before uploading and degraded the quality of my sample set. After uploading the images at full definition, I created an impulse with an image resolution of 320x320 under the assumption higher resolution would mean better outcomes. Object detection settings were first run using the MobileNetV2 architecture with a validation set size of 30%. This output gave me a precision score of 72.1%.

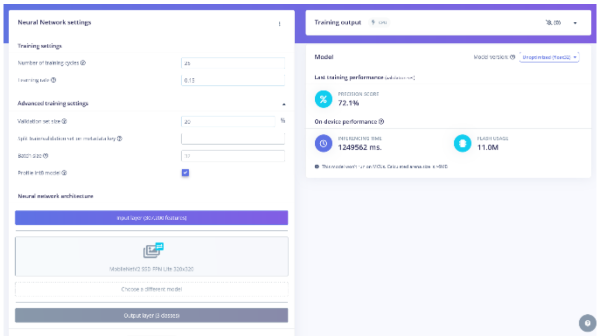
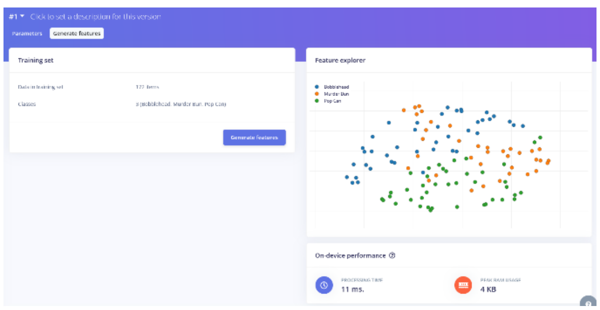
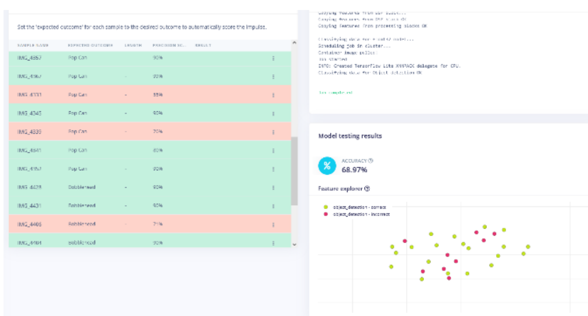
   An additional test was run using the FOMO framework, but this kept failing which appears to be due to only accepting greyscale images (mine were RGB).

4. *Accuracy* – 68.97%
   Precision – 72.1%

---

[1] Please note, I was not in class on the day this exercised was held and therefore only have a second-hand account of what we were doing in class. This question has been answered to the best of my ability without a first-hand account.

5.

	*Screencap of the object detection output. After running the MobileNetV2 architecture, I received a precision score of 72.1%.*
	*The data sorting between the three image sets shows overlap, but also the beginning of some clustering. I suspect the overlap is largely caused by the background noise that showed up in my photos.*
	*Setting the image threshold to 0.05 allows me to achieve an accuracy score of 68.97%. The images that didn't score well were poor quality or obstructed.*

6. In addition to having a larger data set to work with, I could get my model to perform better with the quality of photos I'm using. The lighting isn't great and the objects aren't nearly as in focus as they could be.

PART B

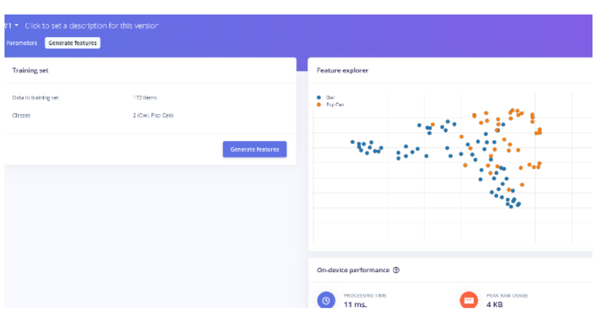
1. For this model, I used my pre-existing images of the pop can and compared it to an image library of wooden owls created by Sabine for class. The owl library had 71 images inside of it and it was chosen due to the clear quality of all the photos. All owl images were clear of any other objects (save for one) and had clean lighting.

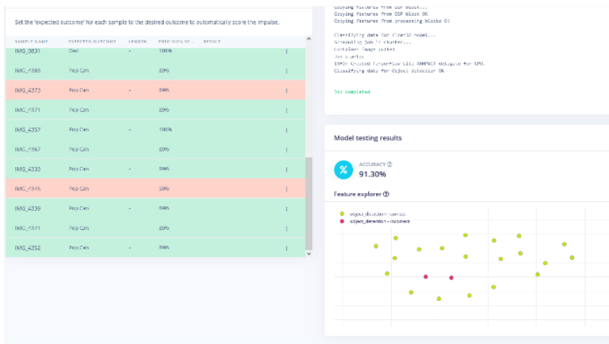
		
Distinct object: Pop Can		*Sourced object: Wooden Owl*

2. The purpose of the task was to contrast the result of a well-documented data set against and poorly curated one. My pop can photos featured lots of noise in the background compared to the noiselessness of the owl photos.

3. *Accuracy* – 91.03%
Precision – 79.4%

4.

	*Clustering is beginning to happen when object detection is checking between the wooden owl and the pop can. However, it's clear that the owl images are much more confidently identified by the system.*

	*With an accuracy score of 91.3%, all the owl photos in my test set are correctly identified. The pop can predictions also seemed to improve a bit but are still lower in result quality.*

5. As was mentioned earlier, the quality of the documentation is what created a better result compared to the first model. Improvements could still be made with future collections such as having a background color different than the owl.

PART C

What?

A portable camera that looks and initially operates like any other. When a photo is taken, what is processed as the output is an AI's interpretation of the scene.

Why?

On the research end of things, this device would allow for more natural sample collection in the real world while getting feedback from the pre-existing model. Subjects that should classify but don't are then documented for later processing.

On the "for fun" end of things, I like the idea of having a device whose entire purpose is to provide an alternative perspective simulation. Maybe there's something you're not noticing in your own photos?

Where?

The design would only require a hobby microcontroller, so it could be fashioned for mobile usage. Existing designs from consumer cameras could be incorporated for greater control over how the images are output.

With whom?

Researchers testing their existing models and weirdo art students such as me would probably get the most out of this prototype due to its output inconsistency. This prototype would be ideal for either testing an existing model on the go or just seeing what weird things you can make.

Storyboard

The user starts by taking a snapshot of a scene or object which is interpreted by a trained data set. Before the image is shown on the display, a trained data set tells an ai generation tool to create an image based on what it thinks it sees.