A2: telephoto-crow

Member: Oskar Austa

I have set up my github repo at: github.com/oskarasd123/telephoto-crow

Business understanding:

Objective:

- I am building a detection system to detect crows and other birds with large presence in cities in order to make automatic deterrent systems possible. This system should be able to detect birds and their species from images.
- The detector should be able to detect birds from low resolution (where the bird is ~30 pixels across) and low light images.

My situation:

I have a telephoto camera that I can aim automatically which I can use to gather data.
I also have a system with an rtx 3060 12GB for model training.
I can gather data in my environment and from the internet to train the model.

My goals:

My primary goal is to make a model that can reliably detect bird presence and species from a distance of 100m with my camera.
I will gather data from the internet to augment my model's feature extraction capabilities and gather my own data to further train and validate my model's capabilities.

Data understanding:

Gathering data:

I will use openly available datasets from the internet.
I will gather and annotate images gathered from my environment.

Describing data:

The data will be small images that have labeled classes. No bounding boxes will be used. There will also be images without birds.

Exploring data:

Not applicable.

Verifying data quality:

The images used for training dosen't need to be high quality since i want the model to perform well with low quality images, however the labels need to be correct. Incorrect labels can be found by using a trained model to predict and searching through badly classified images.

Project plan:

1. Find data from kaggle and make a dataloader to load the data.
    a. The simplest way to get training data is to find already packaged data from the internet. The datasets used to train the feature extractor don't all have to be about birds but I will only keep the datasets that help with real world validation data.
2. Make a model architecture and a simple training setup.
3. Gather images from my environment and label them.
    a. This might take the most time as birds aren't around all the time.

4. Make a multi dataset training setup to use datasets where labels don't match between datasets.
    a. This won't take much time as I have plenty of experience from experimenting and making models with my own data.
5. Train the feature extractor + classification head(one per dataset) on all datasets.
    a. This will take at least 7 hours per training run, more depending on how much data I gather.
6. Iteratively improve the data and training pipeline until I achieve desired accuracy.
    a. After making the training setup the areas where I can improve the most is gathering more data.

    I would guess the total time it takes is about 40 hours.

How am I supposed to keep things concise while writing 900-1900 words?