

# HLCV Interim Report

Júlia Bozzino (7011915)  
Ganesh Gopalkrishna Hegde (7001558)  
Robin Kneepkens (7010559)

July 2, 2021

## Current State of Project

To quickly mention our midterm goals from the project proposal again:

- Have a versatile and robust interface with the dataset.
- Finish exploratory data analysis of the dataset.
- Implement the BirdSpotter network

We haven't exactly reached all these goals, but we have completed some other goals that in hindsight seemed to be important as well.

The interface for the dataset is as good as done. Our dataset has no native pytorch support, so we had to do quite some programming to make it compatible. At this point we can interact with it as with a native pytorch `Dataset`. We also added some support for providing transforms that are to be applied upon sample retrieval. One of the transforms we have implemented resizes the images to a fixed size without cropping. This comes in handy when we try to train our baselines that require a fixed input size. We will also implement a transform that automatically crops the sample image to its bounding box. This will also come in handy for training some comparison networks and for training the BirdSpotter.

We have also made an implementation of ResNet to be finetuned on our dataset, which will function as one of our baselines for comparison, as suggested in the feedback we got on the proposal.

Aside from that, a lot of work went into programming a light framework that is nice to interact with from the GPU cloud.

Work has also started on the BirdSpotter network, but we do not have a fully functional implementation yet.

Our core component, the spatial pyramid pooling (SPP) layer, has been implemented fully by now, as it turned out simpler than expected. We have made a custom pytorch module that implements it, so we can use it as any other module. It consists of a sequence of 2D adaptive max pooling layers from pytorch that are configured to produce the desired output.

## **Problems**

So far, no major problems have been encountered.

## **Plans**

Our near-future goals are to implement the BirdSpotter network and to ensure we can properly train our networks on the GPU cloud. Once we have that, we can start with training our networks and collecting the results for the final presentation and report.