



HaGRID - Hand Gesture Recognition

Brody Sherwood

Problem

- **Project Goal:** Create an accurate and versatile HGR system capable of recognizing and categorizing diverse hand gestures in various environments.
- **Dataset Focus:** Utilize a subset of the extensive HaGRID dataset, incorporating annotations including bounding boxes, landmarks, and metadata.

Data

Images divided into **18** classes of gestures.



Subset of **1,000** images per class from a dataset with **552,992** images.

Contains **34,730** unique persons and at least this number of unique scenes.

The subjects are people from **18** to **65** years old.

Data

Variation in lighting, including artificial and natural light.

Taken in extreme conditions such as facing and backing to a window.

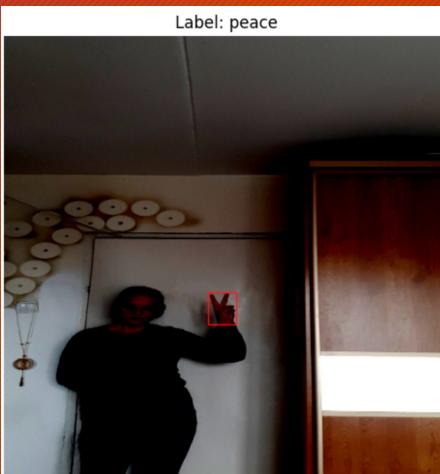
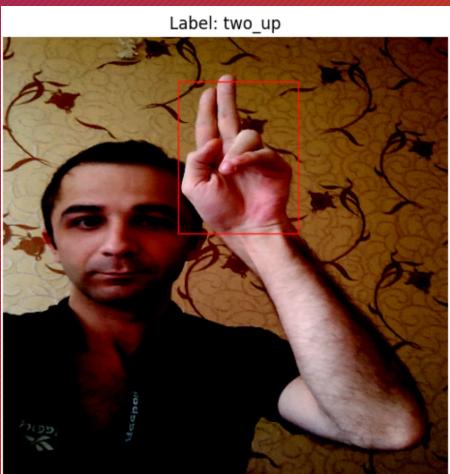
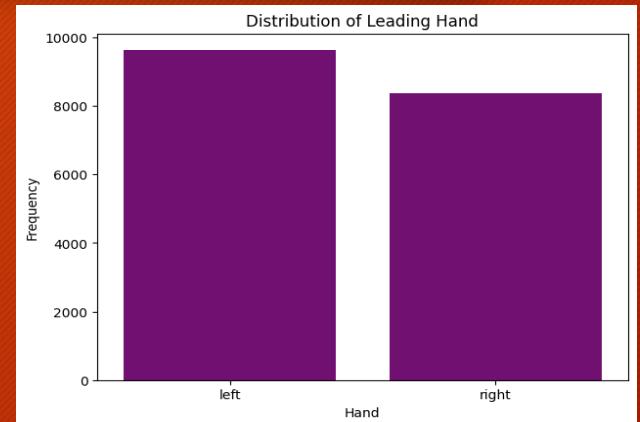
0.5 to 4 meters from the camera



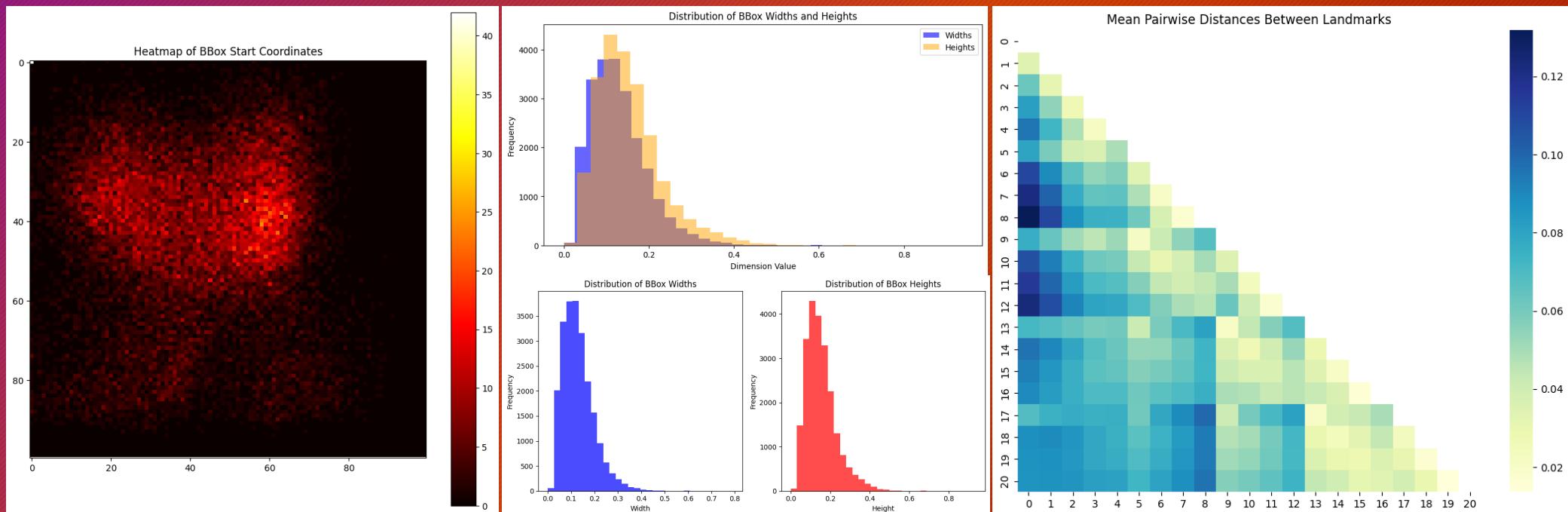
Annotations

Each image's annotations contains:

- Label (class)
- Bounding boxes of hands with gesture labels
 - 21 landmarks
 - Some with 2 bboxes (no_gesture)
- leading_hand** the image has and a **leading_conf** as its confidence.
 - Left hand has nearly 1000 more than right.



Bboxes and Landmarks



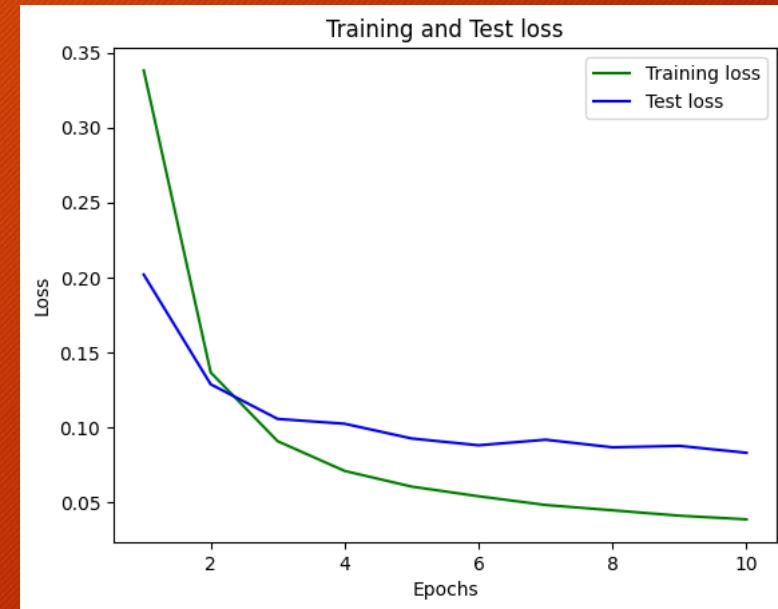
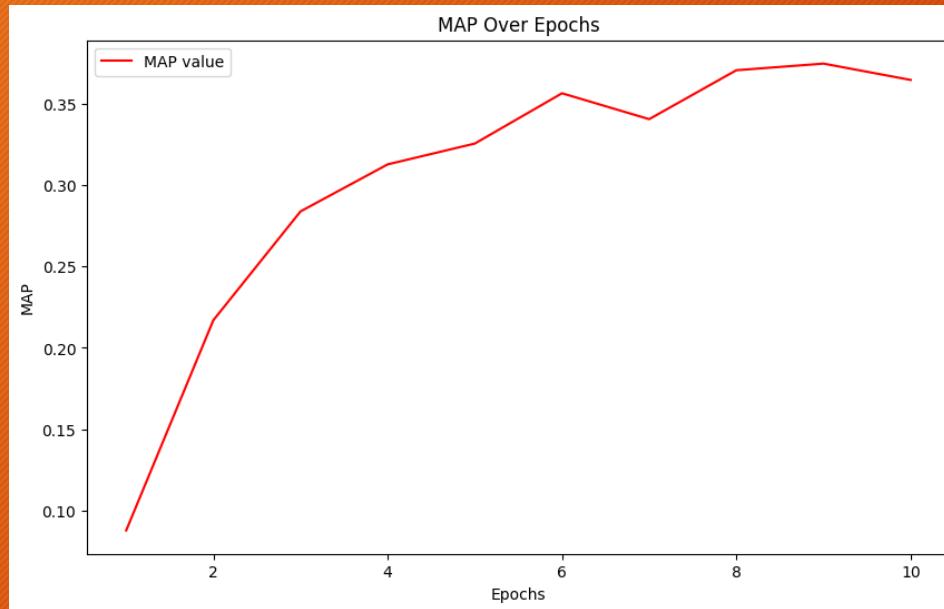
-Only 391 of 18,000 bboxes were larger than the threshold bbox area

RGB



'two_up' has the highest average **Red**, with **'call'** having both the highest average **Green** and **Blue**.

Modeling



epoch : 8 ||| training loss : 0.04083 ||| test loss : 0.08745 ||| MAP : 0.37478

Predictions



Conclusion

- The model needs refining to be optimized, just scratched the surface
- Other pretrained models should be used for comparison
- Different methodology of attack in the future
- Pytorch is hard