Deep Learning – Final Project.

Idan Kelman, Doron Chapnitsky:

**A picture containing reptile, turtle

Description automatically generated**

**Info About Us:**

Idan kelman 207190828, [kelmanidan@gmail.com](mailto:kelmanidan@gmail.com)

Doron Chapnitsky 316231190, [chapnitskydoron@gmail.com](mailto:chapnitskydoron@gmail.com)

**Introduction**:

An approach to hand gesture classification via convolutional neural network.

**The problem:**

Since the last four decades, almost all forms of human gestures have been studied and used either as a natural or intuitive method to interact with computational devices. To compliment, all input–output technologies have been supportive of gesture-oriented interactions.

A picture containing cosmetic

Description automatically generated

The use of gestures acts as a much more attractive yet effective alternative of complex interface devices.

Gestures are as natural as computers completely integrated into our life. We were curious if we would be able to classify between different hand gestures, and the applications are enormous:

If we would be able to correctly classify between the basic hand gestures that we are training our model on, then the applications can be directly reduced to other problems such as :

Translating the sign language, replacing a keyboard, replacing other forms of interaction between a human and a machine and so on .

**Predecessors to this work:**

The problem that we are trying to face is pretty common , therefore there had been a lot of research on this topic .

Even though there had been a lot of research , they often solve a small sub-set of hand gestures , for example : one might detect finger indices , the other might detect swipe hand movements , and so on : as an example : we found the following paper by Google , called On-device Real-time Hand Gesture Recognition [https://arxiv.org/pdf/2111.00038.pdf]. in their work ,they suggest a small range of emogi type of symbols as follows :

A picture containing text

Description automatically generated

The main differences between our work and Google’s is in the steps of the classification:

Google suggest a classification of the key points of the hand as the basis of the hand skeleton tracker, which will be later fed onto a Neural Network to predict the hand gesture.

Diagram

Description automatically generated

Our method is a bit different; we are using a CNN network to classify directly from the input image. another key difference between the two works, is that our model is predicting more types of hand gestures, and different kinds of them.

**Our Dataset:**

We used a custom dataset that find online, that is containing 10 different classes of hand gestures, 20,000 different greyscale images that were shot by 10 different people .

* We normalized our data in order for it to fit into the CNN, by resizing the images to 128x128.
* Since we want to be able to recognize the hand gestures from a video, we had to make sure that we have all of the angles of the hands, therefor we also applied some horizontal and vertical random rotations to the images
* On some of the images, we also made it harder for our CNN to recognize what’s the hand gesture inside the image, by merging two images into one, containing a distraction.
* In order to correctly use the dataset, we then finally organized the batches in custom classes for our model.

**Methods:**

The main method

There are various different methods that might be good in order to try and classify