|  |  |
| --- | --- |
| Text  Description automatically generated with medium confidence | **Machine Learning II**  **Final Project Proposal:**  **Mojahid Osman and Mina Hanna** |

**What we are trying to solve:**

Fake images, news and videos are all over the internet and social media which is creating a lot of negative impact to individual and organizations. In our final project we want to take a small action and use Deep Neural Network to help identifying fake images.

**Dataset:**

The main dataset we will use in this project is the PS Battle Dataset, (<https://paperswithcode.com/dataset/ps-battles>) which has around 102K images where all images and their edited versions are labeled along with other data like author, width, and height. The dataset has a total of 11,143 original image and 90,887 manipulated images.

A picture containing arthropod, invertebrate, lobster

Description automatically generated

[Original image]

A picture containing person, outdoor, sport, group

Description automatically generated A person swimming underwater

Description automatically generated with low confidence

[Photoshopped images]

Also, there is another dataset that is mainly focused on facial edits (<https://paperswithcode.com/dataset/faceforensics-1>) that also might be potentially used in our training but we still need to do more research to see if this will be beneficial or it can be used to build different models and we use an ensemble to capture the best results.

**Deep Neural Network and Reference Material:**

The plan is to use CNN for such detection with potential use pre-trained available neural networks as part of the experiments we plan to make.

We plan to have more understanding of the impact of photoshop manipulation on images to be able to decide the CNN parameters like kernel size, padding and pooling. We will also look at the classic pre-trained networks like AlexNet or VGG-19 to design our CNN following some of their lesson learned.

We plan to use TensorFlow as a framework for our work.

**Metrics that we will use:**

In our classification model we initially plan to utilize confusion matrix and AUC for model assessment we also plan to source other photoshopped images to test the model.

**Potential enhancements:**

Depending on the progress, we might enhance the deliveries of the project with the following:

1. Use a GUI package to build an application that can be used by broader audience
2. Expand detection to detect famous filters as another manipulation avenue

**Timelines:**

|  |  |
| --- | --- |
| Week #1 | All EDA analysis and research for the initial design of the CNN (already started) |
| Week #2 | Start building models and analyze different performance |
| Week #3 | Work on final fine tuning and enhancement if time permits |
| Week #4 | Finalize papers and presentations |