Data Selection Proposal

by Sarah Zhou

**1. Dataset**

I chose the set from [Deepfake Detection Challenge](https://www.kaggle.com/c/deepfake-detection-challenge/data). It is a large dataset of 470GB and has a pre-separated validation set of 400 videos. The test set is withheld, as it should. If needed, I could potentially scrape some videos from YouTube to augment the dataset, since it is the most common platform for videos.

**2. Methodology**

The general plan is to create a machine learning model that will train on a mixed dataset of deepfake and non deepfake videos. This should be a supervised ML project as it is essentially a classification problem on visual media.

**i. Data Preprocessing**

The dataset requires a lot of space, which will probably take a long time to load. Frames with faces and temporal features such as blinking are valuable. Neural networks require little pre-processing. Although, aligning faces, extracting and scaling the bounding boxes of eye landmark points to create new sequences of frames are pertinent[[1]](#footnote-1).

**ii. Machine learning model**

I want to estimate the probability that a piece of visual media contains deepfake.

A convoluted neural network (CNN) might be a good fit. Its multiplayer aspect allows for complicated. I can use backpropagation to calculate loss and provide feedback.

As CNNs are translation invariant, thus they can handle changes in the images, and still perform well on classification tasks. Other variants of models/network that has “position deformation invariant responses”[[2]](#footnote-2) could also be tried out. LRCN, a feature extractor on CNN, and the implementation of LSTM (long short term memory) also deem to be useful for capturing temporal patterns[[3]](#footnote-3). Overfitting is a con for this model.

**iii. Final conceptualization**

My choice for the final product is a simple landing-page web application. It will be a page where the user can provide a link to a picture/video or upload their own picture/video to detect if the content contains deepfake. The result will appear on a second page with the likelihood of deepfake indicated in percentage.

1. Deep Learning for Deepfakes Creation and Detection. <https://arxiv.org/pdf/1909.11573.pdf>. [↑](#footnote-ref-1)
2. McClure, Sean. <https://www.quora.com/What-does-space-invariant-artificial-neural-networks-SIANN-mean>. [↑](#footnote-ref-2)
3. Deep Learning for Deepfakes Creation and Detection. <https://arxiv.org/pdf/1909.11573.pdf>. [↑](#footnote-ref-3)