Project_group_01_04 17

CS512 FINAL PROJECT PROPOSAL

IMAGE SIMILARITY DETECTION FOR PROFILE VERIFICATION

- **Group Members:**

  - Abhinay Mannepalli (am2977)
  - Krishna Swaroop Pamidimukkala (kp956)
  - Yuvaneeth Pasham (yp345)
  - Faisal Junaid Bukhari Syed (fs471)

- **OVERVIEW:**

We have all heard and have been fascinated with facial recognition but not everyone thinks about how important it is to ascertain the similarity of images. Our team proposes to develop a model which detects and finds the similarities of different photos using Siamese networks. This is useful in profile verification, which can be used by many apps which deals with photos to verify the authenticity of the profile.

- **METHODOLOGY AND NOVELTY:**

Neural networks are the go-to algorithm when it comes to image classification. This is partly because they can have arbitrarily large number of trainable parameters. However, this comes at a cost of requiring a large amount of data, which is sometimes not available. So, we are trying to achieve a model, which aims to mitigate such an issue, and be able to identify the similarity with as minimum dataset as possible.

Our model will be trained on very limited instances, but they will be from the similar domain as our training example. The model that we wish to utilize is a Siamese network. Siamese networks have been extensively used for classification tasks but we propose to find out the similarity between images and assign ground reality of whether the person is an imposter vs the actual person.

As humans, we can recognize a person by his/her face by just meeting them once, and it is desirable by computers because many times data is at a minimum. This is what our team is desiring to achieve.

- **DATASETS:**

We will be working with the following dataset:

Indian Movie Face database (IMFDB) is a large unconstrained face database consisting of 34512 images of 100 Indian actors collected from more than 100 videos. (For this experiment we are using data of 16 actors, which results in 5000 training samples and 1095 testing samples.)

- **TECHNOLOGIES and FRAMEWORKS:**

    1. PyTorch
    2. Keras
    3. MatplotLib
    4. Numpy