

# DSAA PROJECT

## REPORT

### Team SyNc

- Ishita Vohra 20171054
- Samhita Kanaparthi 2018121005
- Vani Sancheti 20171179

### Methods Considered:

We have considered various methods including Text Extraction, calculated correlation using dot product for images stored as flattened linear vectors and linear regression using the given dataset for prediction.

Text extraction methods are not used here because there are images having cropped noise which would contribute hugely towards the deviation between two similar images.

When implemented the correlation using flattened images the accuracy found was 50% for the test data which was very less compared then the one currently implemented.

### Method Used:

Correlation is a method which depicts the probability that a linear relationship exists between the two images.

The coefficient for correlation is a measure for the correlation value.

The formula used for calculating the coefficient is:

$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$

where  $x_0$  and  $y_0$  are intensity values of the corresponding images, and  $\bar{x}$  and  $\bar{y}$  are mean intensity values of the corresponding image.

It condenses the comparison of two 2-D images down to a single scalar value,  $R$ . It is completely invariant to linear transformations of  $x$  and  $y$ . So,  $r$  is insensitive (within some limits) to uniform variations in brightness or contrast across an image.

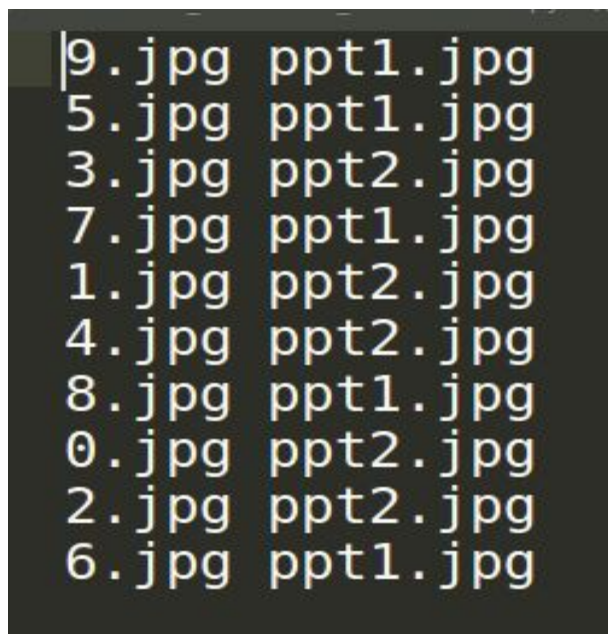
#### **Libraries Used:**

- Numpy
- cv2 from OpenCV
- Matplotlib.image

Used **Python3** for implementing the algorithm.

**Run using Command:** `python3 20171054_20171179_2018121005.py <Path to Slides directory> <Path to Frames directory>`

#### **Output for Test Sample:**



#### **Accuracy:**

1. The algorithm we have implemented shows **100%** accuracy for the sample test case.
2. It also performs pretty well with **91%** accuracy when checked with the provided dataset.
3. As we could see even though the core algorithm we choose was naive, The modifications we implemented gave us a great advantage on the result.