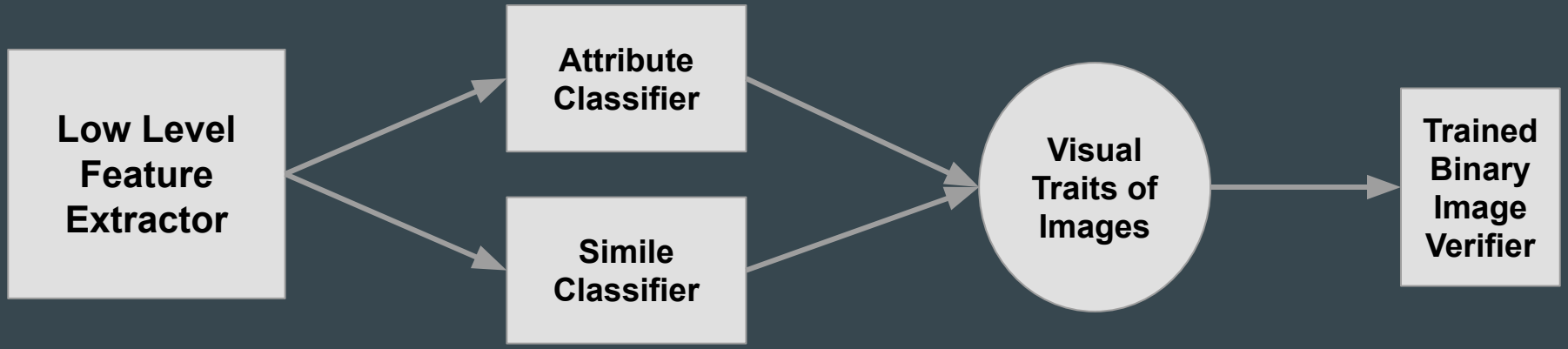
Attribute and Simile Classifiers for Face Verification

Team: Computer Visionaries
Ishaan Khare 20171153
Saraansh Tandon 20171007
Shubhankar Bhagwat 20171147

PIPELINE


Low Level Feature Extractor
Attribute Classifier
Simile Classifier
Visual Traits of Images
Trained Binary Image Verifier

Classifiers Aimed By the Paper

Attribute Classifier

1. 60-65 Binary classifiers. Cover Describable aspects of visual appearance Supervised Learning
2. Eg. Gender, Age group, skin color, etc.

Simile Classifier

1. Important aspects which cannot be described easily but are key in visual appearance.
2. Unsupervised Learning
3. Describe a person's appearance in terms of the similarity of different parts of their face to a limited set of "reference" people.

Verification Classifier

1. Given the trait vectors of two images, it decides whether they are of the same person.
2. Final Classifier of the pipeline.

PROGRESS TILL NOW

Literature Review	• Completed understanding the paper • Following the pipeline as per paper, aim to replicate the results.
Processing Dataset	• Understood the structure and split the Dataset into train and test sets for classifier preparation. • Distribution of images was not random , it was as per the research paper.
Attribute classifiers	• Trained SVM classifiers with low level features extracted from the images. Used the dataset structure as mentioned above. • Results were good accuracy wise bu there were some major issues..
Low Level Feature extraction	• Paper vaguely mentions the possible low level features, their normalization and reduction. • We filled in the gaps with our own ideas to create a low level feature set for all images.

LOW LEVEL FEATURE EXTRACTION

- Aim is to provide good meaningful and compact representations of images for training purposes.
- Extract faces from given images using inbuilt CV2 functions.
- Align all extracted faces to facilitate landmark extraction and also to bring all of them in a common frame of reference for observation.
- Extract important regions from these aligned faces, this helps pinpoint the important regions of the faces which can be used in the representation.
- This ‘information’ is interpreted and stored in the form of various local/low-level/pixel-level properties of images including :
    - Histograms of RGB,HSV versions of the image.
    - Histograms of gradient magnitude and orientations.
    - Size reduced RGB and gradient image.

IMPLEMENTATION DETAILS

- HAAR Face detection function available in opencv is used to detect faces.
- It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images.
- To achieve face alignment we constructed affine matrix to transform the image, source and destination points are derived by holding eyes as reference.
- For landmark detection we used inbuilt libraray DLIB, which is a popular face/shape landmark predictor.

ATTRIBUTE CLASSIFIER

- This is the first type of classifier. For a given attribute, this classifier learns a binary boundary denoting presence or absence of the attribute.
- The paper mentions 63 predefined attributes like ‘baby’ , ‘Bushy eyebrows’, ‘moustache’ ‘frowning’ . They also have provided annotated values corresponding to each attribute for each image.
- Also, the paper has mentioned particular pairs and image sets to be used for training so as to facilitate best possible training .
- Till now we have trained few classifiers with the low level features extracted from the images. As mentioned before , we did achieve good results accuracy-wise(which is the metric used in the paper), but there were some major issues that we need to consider.

IMPLEMENTATION DETAILS

- We have used test-train set for development as was provided on the official site of the dataset w.r.t this particular paper.
- We have used SkLearn's implementation of grid search to perform SVM classification
- The paper has mentioned that they used some version of adaboosting for feature selection to reduce the dimensionality, but the authors are not very specific on the method. So for obtaining some initial results we tried out the following two feature sets without feature selection:
  - A concatenation of all the low level features mentioned before.
  - Embeddings generated from a pre-trained neural network.
- Although some of the traits got pretty good accuracies, on observing the confusion matrices we realised that the datasets were highly imbalanced leading to very poor precision and recall.

Some Results

Features	Attribute	Accuracy
Resnet	Asian	93 %
Resnet	Attractive Woman	87.59%
Resnet	Baby	82.75 %
Resnet	Bags Under Eyes	67.85 %
Resnet	Bald	78.63 %

GOING FORWARD

- The very next step is to look at some of the feature selection methods which are fast like the filter methods. (Although the paper mostly used wrapper methods.)
- Training Simile Classifiers will be the next step. As the paper has mentioned , we must use ‘reference images’ and there special features to compare other images. The downloading of the dataset of was a difficult task due to only URLs being available but as of now the downloads are underway.
- The final classifier which is the aim of the paper, is the one which verifies whether the two images are of the same person. It uses both attribute and simile classifiers to train. This will be the final step of the pipeline
- We may try to incorporate our ideas in places where paper is not explicit in its details.