Ensemble Networks for Better Facial Recognition of Bearded Faces and Beyond

Edward Vendrow
evendrow@stanford.edu
Akash Singhal
akash13@stanford.edu
Amogh Dixit
addixit@stanford.edu

Introduction

- Facial recognition systems perform poorly on face with obscured features
- We propose a specialized secondary network that is trained for some particular obscurity, with a dispatcher network selecting the recognition network to use. We demonstrate this for beards.
- We achieve an accuracy of **0.9829 ± 0.0125** compared to a baseline FaceNet[1] accuracy of **0.94286 ± 0.02020** on our test set of celebrity images with and without beards.
- **With an accurate dispatcher network, our method theoretically offers strict improvements over existing face recognition systems**

Data

Examples of celebrities with and without beards:

- Needed a custom dataset because we only wanted thick beards to be fed into our specialized network and not faces with just some stubble.
- Collected 1600 images (800 with and w/o beard) with some taken from the CelebA [2] dataset and others downloaded individually from the internet for training/testing just the dispatcher network. There were upwards of 400 identities in this.
- Downloaded another ~1650 images individually of 60 different identities with and w/o beard to train/test just our specialized secondary network.
- Preprocessed the data by running face detection with MTCNN to align, crop, and resize images. Also augmented the data with random cropping, rotation, and flipping.

Results

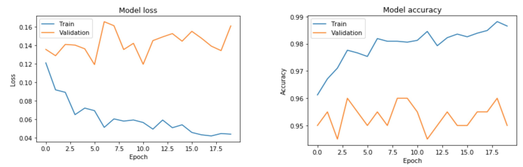
Model	Dataset	Accuracy	Validation Rate	AUC	ERR
Baseline	LFW[3]	0.9852 ± 0.0065	0.9250 ± 0.0237	0.998	0.015
Baseline	Beards	0.9429 ± 0.0202	0.1542 ± 0.0585	0.986	0.070
Specialized	Beards	0.9829 ± 0.0124	0.4657 ± 0.0948	1.000	0.013

Face Recognition: Our specialized face recognition network achieves significantly higher accuracy than the standard FaceNet model.

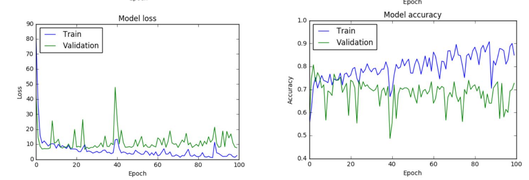
Method	Test Set Accuracy
SVM	0.53
Logistic Regression	0.70
ConvNet with Transfer Learning	0.984

Beard Recognition: Our dispatcher network recognizes the presence of a beard. Using a convolutional network architecture, we achieve high beard recognition accuracy

Deep Neural Network Training



Beard Detection using Logistic Regression

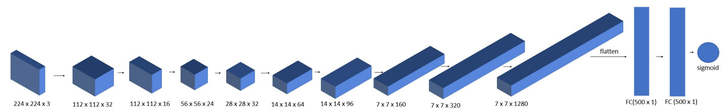


Method

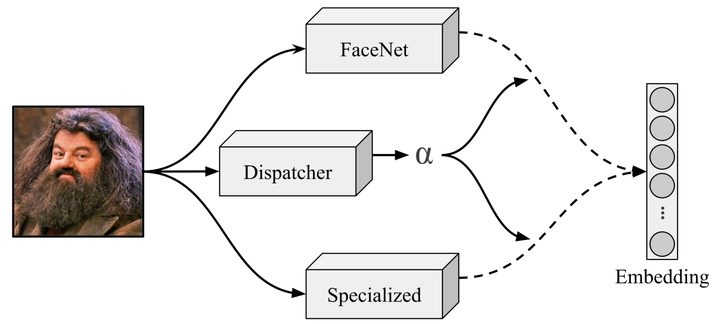
- We pass the input images through three different blocks, namely, conventional FaceNet, dispatcher, and our specialized network.
- Depending on the output of the dispatcher network (which detects the presence of a beard), we use the output of either the conventional FaceNet block or that of our specialized network.
- The specialized network is similar to FaceNet but has been trained on our additional beard/no beard (52 identities) dataset.

Architecture

Dispatcher Network Architecture:



Overall Model Architecture:



Future Work

- This architecture can be expanded to include other possible obscurities (shades, face masks, etc) by training more specialized networks with a softmax output dispatch network.
- Collect more data with more identities than the 60 identities we had for training the specialized network to get even better accuracy.
- **With an accurate dispatcher network, our method theoretically offers strict improvements over existing face recognition systems**

References

[1] Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "FaceNet: A Unified Embedding for Face Recognition and Clustering." 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015): 1-10 Crossref. Web.

[2] Liu, Ziwei, et al. "Deep Learning Face Attributes in the Wild." Proceedings of International Conference on Computer Vision (ICCV), 2015.

[3] Huang, Gary B, et. al. "Unsupervised joint alignment of complex images." International Conference on Computer Vision (ICCV), 2007.