Finding Your Celebrity Look Alike

By Chase Davis, Amanda Jacquez

Problem Definition

The goal of this project is to employ computer vision techniques with the purpose of finding out which celebrity one's face is most similar to.

Future Work

Improve performance in the future by:

- Increasing the testing set
- Increase the dataset to provide even more celebrities to use as reference
- Adapt our evaluation metrics to place weight on different features of the face to test what generates more realistic dopplegangers.

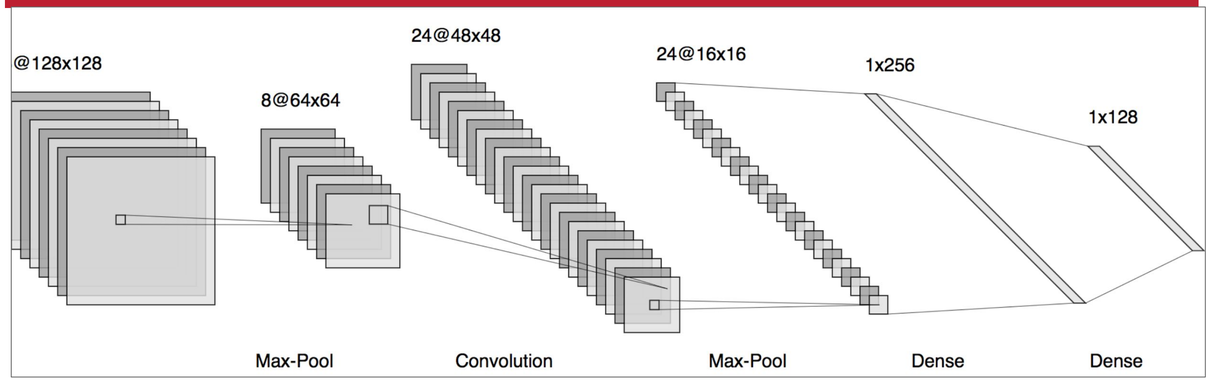
Challenges

- Making results quantitative
- Long training times made it difficult to tell when changes successfully improved training.
- The dataset was extremely large, so it was challenging when generating a smaller set to train on, how large would maintain accuracy but be feasible to test on

References

[1] "Ratings and Reviews for New Movies and TV Shows." IMDb, IMDb.com, www.imdb.com/.
[2] Serengil, Sefik, et al. "Deep Face Recognition with VGG-Face in Keras." Sefik Ilkin Serengil, 15 July 2019, sefiks.com/2018/08/06/deep-face-recognition-with-keras/.
\bibitem
[3] Tabora, Vince. "Face Detection Using OpenCV With Haar Cascade Classifiers." Medium, Becoming Human: Artificial Intelligence Magazine, 4 Feb. 2019, becominghuman.ai/face-detection-using-opencv-with-haar-cascade-classifiers-941dbb25177.
[4] Chen, C. Chen and W. H. Hsu, "Face Recognition and Retrieval Using Cross-Age Reference Coding With Cross-Age Celebrity Dataset," in IEEE Transactions on Multimedia, vol. 17, no. 6, pp. 804-815, June 2015.

Approach



- Model: We used a Convolutional Neural Network with 16 Convolutional layers with relu activation functions, 5 max pool layers and 2 dropout layers
- Layer Order: Conv, Conv, Max Pool, Conv, Conv, Max Pool, Conv, Conv, Conv, Max Pool, Conv, Conv, Conv, Max Pool, Conv, Drop Out, Conv, Drop Out, Conv, Softmax
- Scoring: Feature based absolute difference + Euclidean Difference
  - We extracted the location of each facial feature (left eye, right eye, mouth and nose) in the original image as well as the result. Copied the result and original images twice; passing first copy through grayscale filter and second copy through sobel filter (for edge detection)
  - Take the absolute difference between result image and original image of each facial feature in grayscale, sobel, and color then add these differences together.
  - Finally add Euclidean difference between color of original and color of result image.

$$|(\mathbf{S}\odot\mathbf{S})-(\mathbf{O}\odot\mathbf{O})|+\frac{1}{15}\sum_{f=0}^{4}\sum_{i=0}^{img\_width}\sum_{i=0}^{img\_height}|\mathbf{O}_{f,i,j}-\mathbf{R}_{f,i,j}|+|G(\mathbf{O})_{f,i,j}-G(\mathbf{R})_{f,i,j}|+|S(\mathbf{O})_{f,i,j}-S(\mathbf{R})_{f,i,j}|$$

Results

- Evaluated using the scoring method defined above on 50 original images
- Divided all scores my maximum score to get between 0 and 1 then multiplied by 100 to get values between 0 and 100
- We then compared median scores in each group
  - Original implementation median score: 25
  - New implementation median score: 17
  - A lower score indicates closer distance and is thus a better score
- Scoring method when being used on the same person in two seperate pictures had a small margin of error (about 5) but is still a good indicator of similar images
- We saw from results that important features tended to include observed facial features as well as complexion and hair color. Mouth seemed to be the least important feature, most likely because mouths appear differently depending on expression of emotion.

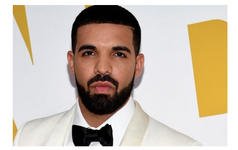
Data

We used the IMDB-WIKI dataset in our implementation. This dataset contains 524,230 images with gender and age labels attached to every picture.

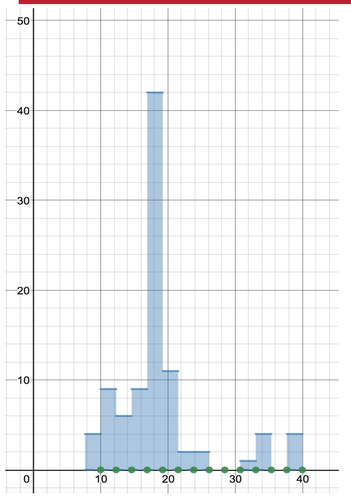
IMDB-WIKI	IMDB	Wikipedia	IMDB-WIKI used for training
524,230	461,871	62,359	260,282 images

The dataset is comprised of a combination of data from the most popular 100,000 actors as listed on the IMDB website and images from Wikipedia

When extracting these images from IMDB [1], the timestamp of which the photo was taken was removed, and images with multiple high scored face detections were removed.



Analysis



- The chart on the left represents the scaled frequency by which scores appear in the results
- A lower number implies a greater similarity, or smaller difference, between the inputted image and celebrity look alike.
- Our implementation outperforms the baseline performance

Acknowledgments

We would like to thank the CS230 teaching staff for their guidance and support throughout the quarter!