Deep Learning Solution to Catfish Prevention

Stanford University

Kris Kaya and Stefan Swaans

{kkaya23, sswaans}@stanford.edu

Motivation

The increase in false profiles on social media has resulted in users being lured into relationships by a fictional online person, a phenomenon known as catfishing. We implemented a facial recognition algorithm that when combined with a pose recognition algorithm, provides a model for detecting false social media profiles. We compared performance using two architectures: a standard Convolutional Neural Network and a Transfer Learning Model (Inception-Resnet v2) on a classification task between 15 different subjects using accuracy as our metric

Data/Features

We used the Yale Faces Data Set, which contains 165 grayscale images of 15 individuals (14 males, one female). Each subject has 11 different images, each with a different facial expression: center-light, wearing glasses, left-light, happy, without glasses, normal, right-light, sad, sleepy, surprised, and winking. Since the data set was relatively small, we did approximately a 60-20-20 split for the train, dev, and test sets.

Our raw input features are the pixels of each grayscale image, which is 256x256 in the basic CNN and 128x128 in the transfer learning solution. We do not have any derived input features. Our features are appropriate for this task because they are standard inputs for a facial recognition model.

Figure 1. Subject 12 with a surprised expression.

Figure 2. Subject 11 winking

Hyperparameter Tuning

We systematically tuned the following hyperparameters: image size, momentum term, learning rate, and batch size. We also used learning rate decay in our second model. After testing various values we concluded that the hyperparameters below combined to create the best accuracy and lowest loss.

Hyperparameter	Value
α	1e-2.5
β	0.7
Image Size	256
Batch Size	8

Models

Two models were used:

1) Basic CNN

The first model is a 3x3 CNN -> BatchNorm -> ReLU -> 2x2 MaxPool network with Adam Optimization. We classify between 15 different people and use a sparse softmax cross entropy loss function.

2) Inception-Resnet v2

The second model that we tested used transfer learning. We used the Inception v3 architecture initially trained on the ImageNet dataset. We then trained it on our own dataset and used it to classify between the 15 different people. We again used a sparse softmax cross entropy loss function.

Figure 3. Graphical representation of the Basic CNN model.

Figure 4. Graphical representation of the Inception-Resnet v2 model.

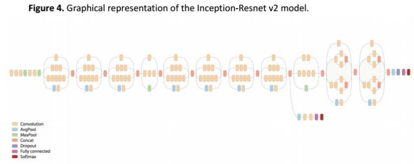


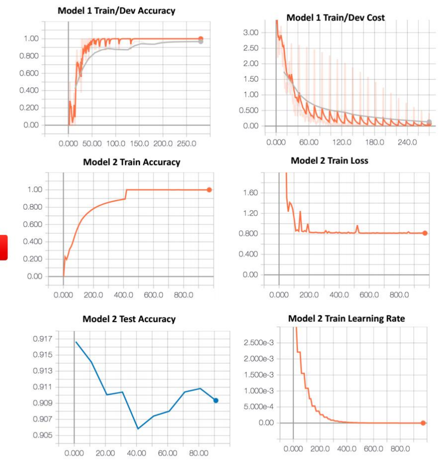
Figure 5. Softmax cross entropy

$$\mathcal{L}(y-\hat{y}) = -\sum_{i=1}^{n} y_i \log \hat{y}_i$$

Results/Discussion

We included 105 images in the training set, 30 images in the dev set, and 30 images in the test set. We found that of the two models, the transfer learning model performed better. While the Base CNN had an accuracy of 90%, Inception-Resnet v2 had a test accuracy of 91%. This was expected, as the model was pre-trained on a large dataset which we thought would result in better performance. The success of Inception-Resnet v2 suggests that since facial and pose recognition models currently exist independently, a transfer learning strategy could be useful in solving the larger problem.

Results cont.



Model	Training Accuracy	Test Accuracy
Base CNN	1	0.90
Inception-Resnet v2	0.9995	0.9108

Future

We proposed that the solution to the problem would involve both a facial recognition and a pose recognition component to detect whether or not a profile was legitimate. Because we focused on implementing the facial recognition component of our solution, the next step is to implement the pose recognition component. Once both are working, we would combine the two models to create a model that can recognize both a pose and a face in one image. We believe this model will be able to detect fake online profiles.

Contact

Kris Kaya and Stefan Swaans
Department of Computer Science
Mentor: Patrick Cho
Stanford University
{kkaya23, sswaans} @stanford.edu

References

1. P. Belhumeur, J. Hespanha, D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, July 1997, pp. 711-720.
2. Turk, M. and Pentland, A. (1991). Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3(1):71{86.
3. Phillips, P. Jonathon et al. "Overview of the face recognition grand challenge." *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* 1 (2005): 947-954 vol. 1.
4. Laurenz Wiskott , Jean-Marc Fellous , Norbert Krüger , Christopher von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, IEEE Transactions on Pattern Analysis and Machine Intelligence, v.19 n.7, p.775-779, July 1997
5. I. Naseem, R. Togneri and M. Bennamoun, "Linear Regression for Face Recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 11, pp. 2106-2112, Nov. 2010