Machine Learning Engineer Nanodegree

Capstone Proposal

Arwa Ibrahim
November 11, 2019

Domain Background

Kinship verification is increasing attentions as it holds promise to an abundance of applications. For instance, it could be used for forensic investigations, automatic photo library management, historic lineage and genealogical studies, social-media analysis, cases of missing children and human trafficking.

The goal of kinship verification is to determine whether a pair of faces are blood relatives or not. This task has seen lots of attention, which mainly focus on parent-child pair-wise types such as: father-daughter (F-D), father-son (F-S), mother-daughter (M-D), mother-son (M-S). Also some have focused on siblings pairs such as: brother-brother (B-B), sister-sister (S-S).

Challenges preventing visual kinship recognition from transitioning from research to reality are two-fold:

1. Existing data resources for visual kinship are too small to capture true data distributions.

2. Hidden factors of visual similarities/differences between blood relatives are complex and less discriminant than the computer vision algorithms used most often for higher-level categorizations like facial recognition or object classification.

Problem Statement

The problem that is to be solved divided into two tasks:

- Kinship verification (one_to_one): intended to determine whether or not a pair of facial images are blood relatives of a particular type (e.g., parent-child). This is a classical boolean problem with system responses being either kin and non-kin (i.e., related and unrelated, respectfully). Thus, this task tackles the one-to-one view of automatic kinship recognition.

- Family classification (one_to_many): determine the family that an unseen subject belongs to. This is done by referencing faces, families modeled using facial images of all but the held-out family members, then at test-time the held out members are used to evaluate on. Types of held out family members vary by type from the youngest boy in a family tree that spans back several generations to a member whom assumes the role of being a mother, a sister, a daughter, and sits right in the middle of the family tree. Hence, this task is formulated as a one-to-many, closed-form classification problem.

Datasets and Inputs

The data is provided by Families In the Wild (FIW), the largest and most comprehensive image database for automatic kinship recognition. FIW is made-up of 11,932 family photos that span 1,000 different families, with the data distributed fairly with average of about 12 images per family, each with at least 3 and as many as 38 members.

This data supports all prior pair-wise types. In addition, it introduce grandparent-grandchild pairs (grandfather-granddaughter (GF-GD), grandfather-grandson (GF-GS), grandmother-granddaughter (GM-GD), grandmother-grandson (GM-GS)).

The FIW database contains rich labels capturing the complex hierarchical relationships inherited throughout families like family trees. Also, it is a public and noncommercial database.

Solution Statement

One of the based CNN solutions of this problem is to use pre-trained VGG-Face CNN for classification via the family labels using fully-connected layer as a final layer. Then, the CNN is fine-tuning using triplet-loss function on top.

Benchmark Model

The authors of the 1st place solution in the FIW2017 competition proposed KinNet, a fine-to-coarse deep metric learning framework for kinship verification. In the framework, the authors transferred knowledge from the large-scale-data-driven face recognition task by pre-training the network with massive data for face recognition. Then, the network was fine-tuned to find a metric space where kin-related peoples are discriminant.

Specifically, the authors adopted different variations of the residual network as the basic architecture of KinNet. Pretraining was done on subset of MS-Celeb-1M dataset, with the output dimension of fc-layer set to 41, 856. A softmax loss was added after the fc-layer to decide between the 41, 856 subjects. While fine-tuning, the fc-layer was replaced with a new fc-layer of size 1024D, which was followed by an L2-norm layer to normalize features to unit length. In the end, a soft triplet-loss was configured to force the KINs closer and NON-KINs farther apart. Since the first layer learns the most general representations (i.e., lower-layers tend to learn more basic filters), the authors froze the bottom most 7×7 convolution layer and updated other layers to adapt the model for kinship verification. To increase the number of training images, as well as balance the number of images per member, the authors proposed an augmentation strategy that makes each member of the 300 families contain the same number of images. Finally, the reported kinship verification result is derived from the fusion of different models with different network depth and cropped sizes at the input.

Evaluation Metrics

The evaluation metrics chosen is the AUC - ROC curve between the **predicted probability** and the **observed target**. AUC - ROC curve is a performance measurement for classification problem at various thresholds settings. ROC is a probability curve and AUC represents degree or measure of separability. It tells how much model is capable of distinguishing between classes. Higher the AUC, better the model is at predicting 0s (non-related) as 0s and 1s (totally related) as 1s. By analogy, Higher the AUC, better the model is at distinguishing between patients with disease and no disease.

Project Design

The research in image-based kinship verification has seen tremendous growth with an increase in the number of kinship specific databases released in the public. Research in this field started with images having frontal pose faces with limited variations. The focus of the research was on estimating handcrafted structural and textural features for verifying kin in images. Since then, several databases have been created which consist of faces with real-world variations of illumination, aging, and pose. However, one of the major challenges in image-based kinship verification research is the limited number of samples present across publicly available datasets. This poses a big challenge for developing accurate deep learning based models for this task.

The research in image-based kinship verification can be divided into two groups:

- Utilizing handcrafted features.
- Utilizing deep learning based features for classification.

The following table showcase the research progression in the area of image-based kinship verification using deep learning based features for classification.

Deep learning has been successfully applied to model representations in natural language processing, speech recognition], image segmentation, and object recognition. These algorithms learn the deep features in an abstract manner by utilizing a network of non-linear transformations in a greedy layer-wise mechanism. Among many deep learning algorithms, stacked denoising autoencoders (SDAE) and deep belief networks (DBN) are two very popular techniques. In additional to these methods there are also Filtered Contractive DBN (fcDBN).

In this project we will divide the solution into two main steps to compare their results. **The first step** is to begin with the transfer learning to create a CNN using one of the popular bottleneck features such as: VGG19, ResNet50, Inception and Xception. **The second step** is to try an autoencoder based method such as: SDAE, DBN and fcDBN.

Authors	Kinship Algorithm	Cornell Kin Database	UB KinFace Database	KinFaceW-I & KinFaceW-II Database	Other databases
Dehghan et al.	Discrimination via gated autoencoders			√	
Robinson et al.	Fine-tuning VGG Network				√
Zhang et al.	Deep Convolutional Neural Networks			√	
Li et al.	Siamese Convolutional Neural Net			√	
Li et al.	KinNet				√
Wang et al.	Denoising auto-encoder based metric learning				√
Duan et al.	Coarse to Fine Transfer	√	√	√	
Kohli et al.	Kinship verification via representation learning	√	√	√	√