DEEPVISION

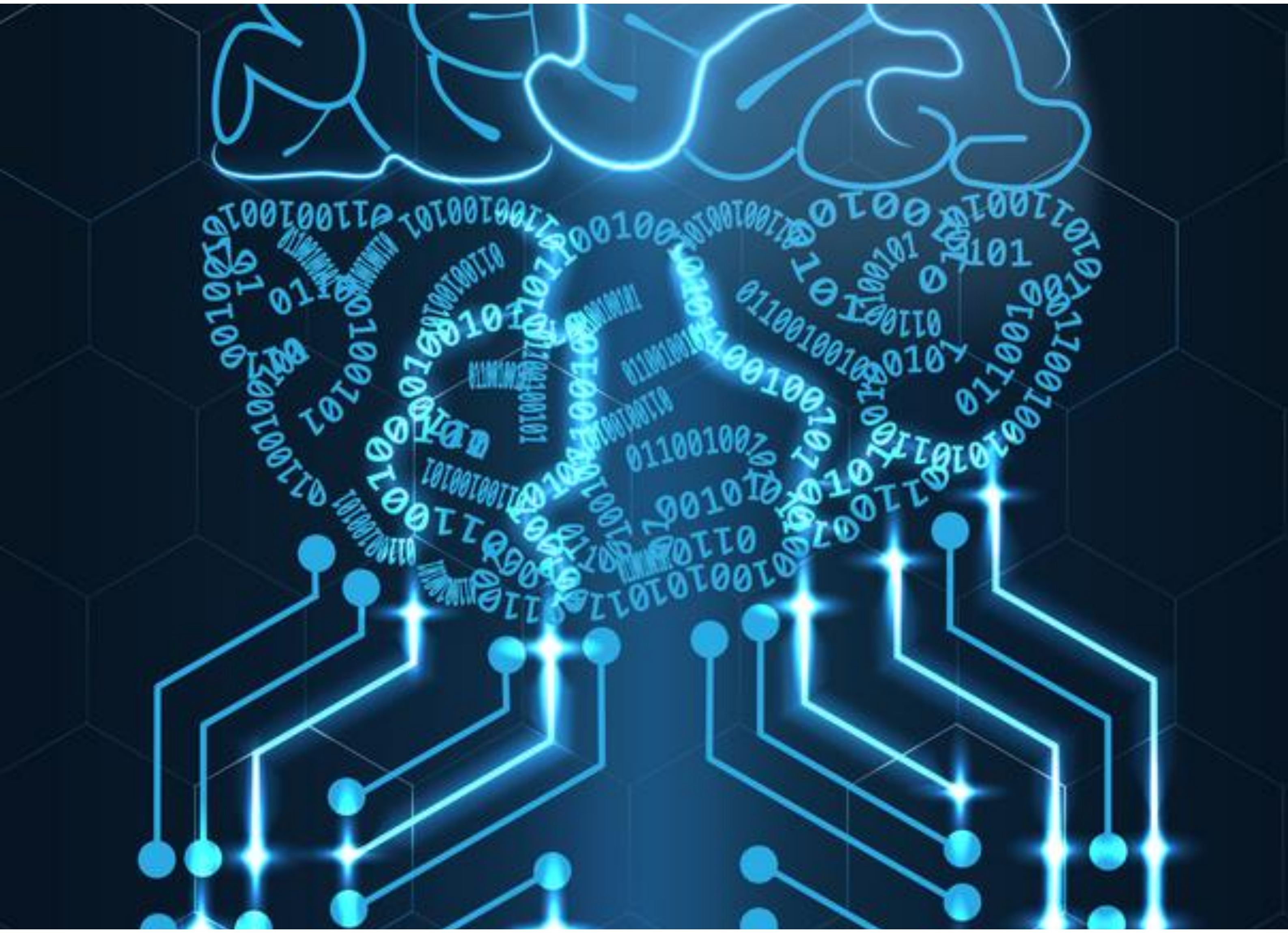
King Fahd University of Petroleum and Minerals
College of Computer Sciences and Engineering
Information and Computer Science Department
ICS 411: Senior Project (Term 172)

KFUPM KFUPM

AGE AND GENDER RECOGNITION USING DEEP LEARNING

Literature Review

Feb 11. 2018



Mustafa Al-Turki	201351850
Majed Alshaibani	201381710
Haitham Albetairi	201379790

Table of Contents

Introduction

The problem of age and gender recognition has been addressed in different works using various approaches and datasets. Approaches can be divided into those which use Convolutional Neural Network and those which do not. Non-CNN approaches use different methods for features extraction and depend on different attributes of face images to provides discriminating characteristics that help in the problem being addressed.

Purpose of This Document

This document provides a detailed and updated literature review of some of the existing algorithms, implementations and applications related to the topic of the project: Age and Gender Recognition Using Deep Learning. First it provides description of a number of previous works in the field, then it provides a comparison and summary for each work.

Review of Work #1

Title: Age and Gender Classification using Convolutional Neutral Networks.

Authors: Gil Levi and Tal Hasser.

Age and gender classification became a major interest for computer vision scientists. The significant advances in today's social media and the explosion of data make such task a necessity. The previous work in this are did provide a solution. However, its low precision did not match the current demand and especially for commercial products. This paper tries to provide a solution to such a problem with a contemporary method using convolutional neural network. This method proves its proficiency outperforming the existing state-of-the-are methods available in its time. The method is evaluated against the well-known Adience benchmark for age and gender estimation and show a dramatic increase in the results.

Past approaches use the so-called face descriptor methods which heavily rely on face dimensions and ratios between face's landmarks. These methods do not fit into current applications as the images to classify needs to be taken with special constraints. The paper tries to correlate the recent advances in the face recognition field with the age and gender recognition using identical methodology, CNNs. The network structure is as follows:

- 96 filters in the first convolutional layer followed by (ReLu) and max pooling layers.
- 256 filters in the second convolutional layer followed by (ReLu) and max pooling layers.
- 384 filters in the last convolutional layer followed by (ReLu) and max pooling layers as well.
- Two fully connected layers.
- A third fully connected layer maps to the final classes for age or gender.

The experiment is implemented using the Caffe open source framework. Training was performed on an Amazon GPU machine. The training time was about 4 hours.

Review of Work #2

Title: Convolutional Neural Networks for Age and Gender Classification.

Author: Ari Ekmekji.

This paper summarizes a work on gender and age classification. It builds over the work of the previous paper. The main motivation for this paper are:

- To investigate the claim that deeper network architecture, reducing the number of parameters or modifying the level of the dropout would not provide better precision. These modification results in a decrease in the system precision (or at best, stay the same).
- To train the age classifier on separated genders, men and women. This results in an improve in the results.

Review of Work #3

Title: DEX: Deep EXpectation of apparent age from a single image.

Authors: Rasmus Rothe, Radu Timofte, Luc Van Gool.

This work studied the estimation of apparent age, that is the age perceived by other humans, from single face images using deep learning. The CNNs used followed the VGG-16 architecture and were pretrained on ImageNet as to benefit from the model learned to classify different objects.

A new dataset is collected due to lack of face images annotated with apparent age. A total of 524,230 face images were collected from IMDB and Wikipedia websites annotated with real ages.

The estimation problem was approached as a classification problem, where age values have been divided into 101 year labels, that is from 0 to 100.

DEX method starts by detecting and cropping the face from the image, then the result is fed to an ensemble of 20 networks.
For face detection, the Mathias et al. face detector was used, all faces were resized to ensure same location of the face in all images, finally, all images were put into 256X256 pixels size.

In age estimation step, the preprocessed image is fed to the deep CNN network which follows the VGG-16 architecture. As the problem approached as a classification problem with 101 classes, the output layer of the network has 101 nodes, each one corresponds to a probability, and the expected values is compute by summing the product of the probability and the corresponding age values.

The deep CNNs are trained on the collected dataset (IMDB-WIKI dataset), and then trained on ChaLearn LAP dataset. After that, 10 augmented versions of each images in the ChaLearn LAP dataset were used to train the networks, augmentation of each image is done using rotation, translation and rescaling of the original image. The prediction is based on the average of the networks in the ensemble, where each one is trained using different splits.

The model was the 1st place winner in the ChaLearn LAP 2015 challenge on apparent age estimation, the competition had 115 registered team.

Review of Work #4

Title: Age and Gender Recognition in the Wild with Deep Attention.
Authors: Pau Rodrıguez, Guillem Cucurull, Josep M. Gonfaus, F. Xavier Roca, Jordi Gonzalez.

This work addresses the problem of age and gender recognition of faces in images in the wild, the problem is due to considerable differences in resolution, deformation and occlusion, and that CNN is highly sensitive of such variations. The proposed solution is influenced by recent advances in attention mechanism. The used feedforward attention mechanism can extract the most useful and discriminative parts of a given image which yields to better age and gender classification. Benchmarking is done using Adience, Image of Groups and MORPH II datasets.

Review of Work #5

Titles: Age and Gender Recognition Using Informative Features of Various Types.
Authors: Ehsan Fazl-Ersi, M. Esmaeel Mousa-Pasandi, Robert Laganière.

This study done in 2014 suggests integrating multiple descriptors to identify age and gender, as opposed to previous methods which focused on one only. They showed that their method results in superior accuracy as compared to those previous methods.

The Method

They relied on three descriptors, Local Binary Pattern (LBP), Scalar Invariant Feature Transform (SIFT), and Color Histograms (CH). Each one of these descriptors analyzes a different aspect of the images, shape, texture, and color respectively. They used SVMs with three RBF kernels, one for each descriptor, for their model. Since, our team will build a CNN instead, that part is not useful to our research. They select the most impactful and non-overlapping features using Ullman's feature selection method, which focuses on feature response vector elimination. The previous work that they will compare their results to relied on Principle Component Analysis (PCA) for feature extraction and LBPs for their descriptor.

The Dataset

They opted not to use a controlled dataset and instead trained and tested their model on Gallagher's natural environment images since their aim was to help the development of real world applications. The dataset contains about 28,000 labelled faces from Flicker. They split the set using a fivefold method proposed by Dago-Casas et al, aiming to have an average number of males/females and age group in each fold. One fold is used for testing and four for training. They also reduced the number of frequently appearing age groups to make all age group appearances even.

The Results

The accuracy of the resultant model was about 5% higher (for gender recognition) yielding a result of almost 92% when using all three descriptors as opposed to models from previous work. However, age recognition still proved to be an issue with accuracy of only 63%.

Review of Work #6

Title: Comparison of Recent Machine Learning Techniques for Gender Recognition from Facial Images.

Authors: Jospeh Lemley, Sami Abdul-Wahed, Dipayan Banik, Razvan Andonie.

This study is about different strategies for tackling the problem of gender recognition. They experimented with various methodologies for feature extraction and classification. And ended with a comparison of the results. They focused on the accuracy of the models as the main measure of performance.

The Method

They used two classification methods, SVMs and CNNs, and three feature extraction methods, Principle Component Analysis (PCA), Histogram of Gradients (HOG), and Dual Tree Complex Wavelet Transform (DTCWT). They used the scikit-learn SVM library for the implementation, and tested the SVMs with unfiltered pixels and all three features extraction techniques. For the CNN, they designed three hidden layers connected to a softmax layer as the output. All three layers are rectified linear convolutional layers with:

- 4x4 kernel for the first and second, and a 3x3 for the third.
- 2x2 pool shape and stride for all.
- 128, 256, 512 output channels respectively.
- Randomized initial weights varying by 0.5

The Dataset

They trained and tested their models on two different datasets, one containing images with optimal scenarios (FERET), while the other has images with different levels of lightning, angles, etc. (Adience). They had two different experiment result sets, employing the strategies mentioned above, one on each dataset. For each experiment, they used 70% for training, repeating each experiment at least ten times.

The Results

Comparing all these results, CNNs proved to be the optimal solution, yielding the highest accuracy on both sets out of all the tested ways of classification. The average accuracy on the FERET set for the CNN was 96% while the closest was DTCWT on an SVM (90%). While for the Adience set the CNNs outperformed the competition by about 20%. A comment made on why that might be the case is that CNNs shine when datasets grow larger. And since today's public image datasets are huge, our team decided to go with them.

Comparison and Summary

STUDY	**LENGTH (PAGES)**	**AUTHERS**	- **DATASET DESCRIPTION**	**METHODOLOGY**	**RESULTS**	**COMMENTS**
AGE AND GENDER CLASSIFICATOIN USING CONVOLUTION NEURAL NETWORKS	9	Gil Levi Tal Hassner	- Adience face dataset, for testing and training. - 26.580 photos of 2.284 unique subjects collected from Flicker.	**1-** Network Architecture: The network architecture is relatively shallow to prevent over-fitting the data. **2-** Training and Testing: Dataset is divided into 5 subject exclusive folds.	Gender: 0.859 Age: exact: 50.7 1-off: 84.7	This method outperforms all previous methods to its date, 2015.

STUDY	**LENGTH (PAGES)**	**AUTHERS**	**DATASET DESCRIPTION**	**METHODOLOGY**	**RESULTS**	**COMMENTS**
CONVOLUTIONAL NEURAL NETWORKS FOR AGE AND GENDER CLASSIFICATION	7	Ari Ekmekji	- Adience face dataset, for testing and training. - 26.580 photos of 2.284 unique subjects collected from ***Flicker***. - Images used are front facing. Total of 20,000. - Images are originally of size 768x768, preprocessed to 256x256.	1- Network Architecture: The network architecture is relatively shallow to prevent over-fitting the data. 2- Training and Testing: Dataset is divided into 6 subject exclusive folds, each of these folds is, then, divided into male and female each is further divided into 8 age groups. 3- The classification is done by separating the tasks of classifying men’s age and women’s age. 4- The approach used is to first classify data on gender and then classify on age for each gender separately. This shows better results.	Age: 0-2: 0.27 4-6: 0.76 8-13: 0.76 15-20: 0.92 25-32: 0.78 38-43: 0.87 48-53: 0.79 60+: 0.76 all: 0.79 Gender: exact: 54.5 1-off: 84.1	The work of this paper is based on the work of the next paper in this table.

STUDY	LENGTH (PAGES)	AUTHERS	DATASET DESCRIPTION	METHODOLOGY	RESULTS	COMMENTS
DEX: DEEP EXPECTATION OF APPARENT AGE FROM A SINGLE IMAGE	6	Rasmus Rothe, Radu Timofte, Luc Van Gool	- IMDB-WIKI Dataset - 524,230 images collected from **IMDB and Wikipedia websites.** - LAP dataset - 4699 face images labeled by averaging 10 opinion of independent users of two online applications.	1- Network Architecture: VGG-16 Architecture, a deep network with 16 layers. An ensemble of 20 networks is used in training and testing. 2- The problem is approached as a classification problem with 101 classes. 3- The age is estimated using the expected value of the final softmax layer. 4- The results are evaluated using Mean Absolute Error MAE and ε-error.	ε-error: 0.264975	This work won the 1st place in ChaLearn LAP 2015 Challenge.

STUDY	LENGTH (PAGES)	AUTHERS	DATASET DESCRIPTION	RESULTS
AGE AND GENDER RECOGNITION IN THE WILD WITH DEEP ATTENTION	34	Pau Rodrıguez, Guillem Cucurull, Josep M. Gonfaus, F. Xavier Roca, Jordi Gonzalez.	- Adience dataset. - Image of Group dataset. - MORPH II dataset.	Age: On Adience: 61.8±2.1 On Image of Group: 60.0 On MORPH II: MAE = 2.56 Gender: On Adience: 93.0±1.8 On Image of Group: 86.9

STUDY	**LENGTH (PAGES)**	**AUTHERS**	**DATASET DESCRIPTION**	**RESULTS**
AGE AND GENDER RECOGNITION USING INFORMATIVE FEATURES OF VARIOUS TYPES	5	Ehsan Fazl-Ersi, M. Esmaeel Mousa-Pasandi, Robert Laganière	- Gallagher’s natural environment images - contains about 28,000 labelled faces from Flicker.	Age: 63% Gender: 92%

Team Contribution

ID	Name	Tasks
201351850	Mustafa Al-Turki	Resources collection, Work #3, Work #4, Document Preparing.
201381710	Majed Alshaibani	Work #1, Work #2
201379790	Haitham Albetairi	Work #5, Work #6

References

[1] Age and Gender Classification using Convolutional Neural Networks,

http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.722.9654&rep=rep1&type=pdf

[2] Convolutional Neural Networks for Age and Gender Classification,

http://cs231n.stanford.edu/reports/2016/pdfs/003_Report.pdf

[3] DEX: Deep EXpectation of Apparent Age from a Single Image,

https://www.vision.ee.ethz.ch/publications/papers/proceedings/eth_biwi_01229.pdf

[4] Age and Gender Recognition in the Wild with Deep Attention,

https://www.researchgate.net/publication/

318084198_Age_and_Gender_Recognition_in_the_Wild_with_Deep_Attention

[5] Age and Gender Recognition Using Informative Features of Various Types,

http://www.site.uottawa.ca/~laganier/publications/esmaeelICIP2014.pdf

[6] Comparison of Recent Machine Learning Techniques for Gender Recognition from Facial

Images, http://ceur-ws.org/Vol-1584/paper21.pdf