Literature Review:

Past tries to approach the problem of age and gender recognition counts on the differences in the ratio between faces landmarks such as ears, mouths and eyes. Moreover, the data needed for such tasks requires special treatments when taking the images such as taking photos in predetermined positions and constrained situations. Thus, they are not appropriate for real life images and, therefore, they are far away from meeting the commercial needs.

The new advances in machine learning techniques suggested to use the so-called Convolutional Neural Networks(CNN) which proves to give a better result in such fields.

This type of Neural Networks can be developed Via the Caffe open source framework. For the papers I read, two of them used Convolutional Neural Network model based on this framework. The third paper talks about an end-to-end system which also uses convolutional neural network, but the model is hidden as the paper is describing a commercial product.

Experiments setup:

There are sources from which one can retrieve the data. Both papers revealing their network model takes Adience data. The data is then split to k-folds for cross validation. The cross validation takes place in a very interesting way called Subject exclusive cross validation. The idea of this method is to ensure that each subject’s images appear all in one-fold. Otherwise, they may appear in the testing and training folds skewing the results to be more promising than they actually are.

The convolutional network description is as follows: The first layers are convolutional that is they have predetermined and trained filters for specific features such as curves, edges and so on. After these layers, there is some ReLu activation layers and pool layers to add some non-linearity to the network. The network may have other convolutional layers after these activation layers associated with some ReLu and pool layers as well. The last layers are the connected layer. They output the class in such away it matches the features of the training data.

After the data is determined, then it is preprocessed. This step includes, face detection, facial landmark detection and alignment alongside with other features. Thereafter, the data is fed into a convolutional neural network to determined model weights. These weights are, then, used to classify input images.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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.   1. 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.   1. The classification is done by separating the tasks of classifying men’s age and women’s age. 2. 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. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 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.   1. 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. |
| daeger: deep age, gender and emotion recognition using convolutional neural networks | **10** | Afshin Debghan, Enrique G. Ortiz  Guang Shu  Syed Zain Masood | * Three training datasets are used. A 4 million Images of over 15k, identities. * 600K images labeled with real age. * Aparent age LAP dataset. | An end to end overview of the system is only given. | Real Age: 61.3  Gender: 91.0 | This paper is for the commercial product [sighthund](https://www.sighthound.com/) |