Gender and adult/non-adult classification results in large datasets based on local features and convolutional neuronal networks

Jesús B. Alonso, Modesto Castrillón-Santana, Maria De Marsico, David Freire-Obregón, Javier Lorenzo-Navarro, Michele Nappi, Enrique Ramn-Balmaseda, Daniel Riccio, Carlos Travieso-González

University of Las Palmas de Gran Canaria, Sapienza University of Rome, University of Salerno, University of Naples Federico II

Biometrics is the science devoted to automatically recognize people based on any physical or behavioral human characteristic, instead to something possessed, memorized, attached or injected. Frequently used human biometric traits are face, iris, gait, DNA, fingerprints, etc.

These characteristics are assumed to be strong, as they present the distinctiveness and permanence to clearly differentiate any two individuals. However, there is additional information that may be extracted from those mentioned biometric traits, e.g. gender, race, apparent age and mood from a face. Certainly, this extra information is not unique. In fact, these attributes lack to be distinctive enough to distinguish between two individuals, but they serve to describe people in meaningful non-overlapping categories. These attributes are known as soft biometrics, and may be certainly fused for recognition, combined with classical biometric traits, help to reduce the search space, or provide user profiling.

Given a facial image of an individual, we have studied two of these attributes, gender and age, evaluating a collection of representation alternatives. For both problems, we have considered a wide and challenging experimental, trying to avoid images captured under restricted conditions. In both cases, we have considered a bi-class problem, reducing the age estimation to an adult/non-adult classification to assist an automatic Child Abuse Material detector.

For gender classification, after reviewing the literature related to large datasets, we have studied MORPH, Labeled Faces in the Wild (LFW) and The Images of Groups (GROUPS). For adult/non-adult classification, after observing the lack in large datasets, we have combined FG-Net, Adience, GenderChildren, GROUPS and Boys2Men.

In both classification problems, different facial areas have been used to extract local descriptors: 1) head and shoulders (HS), 2) face (F), 3) periocular (P) and 4) mouth (M). In total, 15 different variants have been evaluated covering a large collection of grid setups. Additionally, a Convolutional Neuronal Network (CNN) architecture comprising three convolutional layers and two fully connected layers was evaluated for the HS pattern.

The final results indicate that for both problems hand crafted features reported slightly worse accuracies than CNNs based solutions. However, the com-

bination based on score level fusion of local descriptors and CNNs improves remarkably the classification rate.