**Introduction**

Biometric systems are increasingly common in our everyday activities. People recognition through their own physical, physiological or behavioral traits inhibits most of the frauds often committed in security systems based on knowledge (passwords) or tokens (cards, keys, etc.). However, nowadays criminals are already developing techniques to accurately simulate the biometric characteristics of valid users, such as face, fingerprint and iris, to gain access to places or systems, process known as spoofing attack . In this context, robust countermeasure techniques must be developed and integrated into the traditional biometric applications in order

to prevent such frauds. Despite face being a promising trait due to its convenience for users, universality and acceptability, traditional face recognition systems can be easily fooled with

common printed facial photographs, which nowadays can be obtained by criminals on the worldwide network, especially due to the dissemination of social medias and networks.

Spatial image information is extremely important in tasks involving faces, such as face detection and face recognition . The different visual patterns of each facial region encode rich and discriminative information necessary to distinguish a face from other objects, and also from other faces. Regarding face spoofing detection, some works based on handcrafted features have mentioned that different spoofing cues can be extracted from different facial regions.

Recently, deep learning architectures have emerged as good alternatives for solving complex problems and have reached state-of-the-art results in many tasks due to their great power of

abstraction and robustness, working with high-level features, self-learned from the training data . Among the proposed deep learning architectures, Convolutional Neural Networks (CNN) have appeared as one of the most important classes of deep neural networks able to deal with digital images with great performances. Some CNN based state-of-the-art methods were recently proposed for face spoofing detection. However, none of them take into account the different visual aspects of each facial region and, consequently, the different local spoofing cues that could be learned by the neural networks to improve their performances. All proposed methods work on whole faces, in a holistic way, or with random and small patches,

i.e., they train the neural networks with samples extracted from random regions of the faces, all together. This can degrade the performance of the training algorithm since the backpropagation method can be distracted by the different visual information extracted from random regions of the face, instead of learning the real differences between real and fake faces in each facial region, with similar visual aspects, differing only by spoofing cues. In this context, we propose a novel CNN architecture trained in two steps for a better performance in face spoofing detection: (i) the local pre-training phase, in which each part of the model is trained on each main facial region, learning deep local features for attack detection and initializing the whole model in a great position in the search space (the network learns to detect multiple and different spoofing cues from all the facial regions); (ii) the global fine tuning phase, in which the whole model is fine-tuned based on the weights learned independently by its parts and on whole real and fake

facial images, in order to improve the model generalization.

Results obtained on two major datasets used for the evaluation of face spoofing detection techniques, Replay-Attack and CASIA FASD (Face Antispoofing Database), show that

the pre-training step on local and fixed regions of the faces improves the performance of the final model and its convergence speed. The proposed approach outperformed the state-of-theart methods while working with an efficient CNN architecture.