

Face Super-resolution Guided by Facial Component Heatmaps

Xin Yu¹, Basura Fernando¹, Bernard Ghanem², Fatih Porikli¹, Richard Hartley¹

¹Australian National University,²King Abdullah University of Science and Technology
{xin.yu, fatih.porikli, richard.hartley}@anu.edu.au, Basuraf@gmail.com,
Bernard.Ghanem@kaust.edu.sa

Abstract. State-of-the-art face super-resolution methods leverage deep convolutional neural networks to learn a mapping between low-resolution (LR) facial patterns and their corresponding high-resolution (HR) counterparts by exploring local appearance information. However, most of these methods do not account for facial structure and suffer from degradations due to large pose variations and misalignments. In this paper, we propose a method that explicitly incorporates structural information of faces into the face super-resolution process by using a multi-task convolutional neural network (CNN). Our CNN has two branches: one for super-resolving face images and the other branch for predicting salient regions of a face coined facial component heatmaps. These heatmaps encourage the upsampling stream to generate super-resolved faces with higher-quality details. Our method not only uses low-level information (i.e., intensity similarity), but also middle-level information (i.e., face structure) to further explore spatial constraints of facial components from LR inputs images. Therefore, we are able to super-resolve very small unaligned face images (16×16 pixels) with a large upscaling factor of $8\times$, while preserving face structure. Extensive experiments demonstrate that our network achieves superior face hallucination results and outperforms the state-of-the-art.

Keywords: Face, super-resolution, hallucination, facial component localization, multi-task neural networks.

1 Introduction

Face images provide crucial clues for human observation as well as computer analysis [1,2]. However, the performance of most existing facial analysis techniques, such as face alignment [3,4] and identification [5], degrades dramatically when the resolution of a face is adversely low. Face super-resolution (FSR) [8], also known as face hallucination, provides a viable way to recover a high-resolution (HR) face image from its low-resolution (LR) counterpart and has attracted increasing interest in recent years. Modern face hallucination methods employ deep learning [9,10,7,11,6,12,13,14,15,16] and achieve state-of-the-art performance. These methods explore image intensity correspondences between LR and HR faces from large-scale face datasets. Since near-frontal faces prevail in popular

