View Reviews

Paper ID
11

Paper Title
VEGAC: Visual Saliency-based Age, Gender, and Facial Expression Classification Using Convolutional Neural Networks

REVIEWER #3

REVIEW QUESTIONS

1. Please indicate your overall rating of this paper
Reject

2. Comments visible to authors
This paper proposes a framework VEGAC to address age, gender and facial expression classification problem using visual saliency. The framework takes the off-the-shelf CNNs, e.g. VGG-16 to extract the saliency and further visual features for three classification tasks.

I have several concerns and questions.

All the components are not new. The implementation of the framework, i.e. the reweighed the cropped face images by saliency and classification module is a little clumsy. I would encourage the authors to implement two components into one end-to-end network.

The performance is not encouraging. The results on two tasks are inferior to the state-of-the-arts. For the third task, facial expression classification, the authors fail to explain what are CK+ in Table 7. Why methods in CK+ perform much better than the proposed one?

More ablation studies are needed. For instance, in order to show that saliency helps, a baseline should be without saliency reweighting.

Others:
Why reweighting ratio by saliency is set to be 70%?

REVIEWER #1

REVIEW QUESTIONS

1. Please indicate your overall rating of this paper
Borderline

2. Comments visible to authors
This is an analysis paper. The dataset bias is an important issue. However, the analysis itself is not novel and the conclusion is not impressive.

REVIEWER #2

REVIEW QUESTIONS

1. Please indicate your overall rating of this paper

Accept

2. Comments visible to authors

This paper was an interesting and well written exploration of your generalized visual saliency approach (VEGAC) to the age, gender, and facial expression classification problems. Here are a few specific comments/questions/suggestions:

WRT your data partitioning discussed in Section 3.3: In your train/validate/test process is there any reuse of data? Specifically, are you potentially reusing face images from the same person across your partitions?

You should label your axes for Table 3.

Is the face expression dataset that you use (AffectNet) acted or observed? Do you have any thoughts on your face expression results relative to the arousal-valence paradigm? For example, you note that anger and sad are easily confused. This wouldn't be surprising since they are near each other in the arousal-valence space.