

View Reviews

Paper ID

32

Paper Title

Real Time Hand Gesture Recognition in Industry

Track Name

Main Track

Reviewer #1

Not Submitted

Reviewer #2

Not Submitted

Reviewer #4

Not Submitted

Reviewer #5

Questions

1. Overall evaluation

strong reject

2. Relevance to the conference

relevant

4. Review

Authors proposed a system for real time hand gesture recognition and motivate the application within the context of human-robot interaction and Industry 4.0. Overall, the paper is an easy read, however, it lists many details which are known to the experts in the field (the target audience of the ISPA conference and you are not writing an educational text) while omitting several major issues, which is not acceptable.

First, for an industrial application having only positive classes in your experiment is not acceptable. The proposed dataset contains only ten selected gestures and is missing the category of not-a-gesture. The experiments and the evaluation of the proposed algorithm must clearly show that the proposed method recognizes selected gestures and that it does not misinterpret normal hand movement as a gesture; the latter is not clearly demonstrated.

Next, the number of training samples is small w.r.t the number of learning parameters. This is especially suspect considering reported accuracy range from 26% to 83% (see Section IV.). This means that you cannot be certain in the generalization capabilities of the trained model which is not acceptable.

Some other issues:

Remember that you are writing a technical paper for the expert audience, and not an educational material for

beginners. An example is Table I.: What is the purpose of Table I. with regard to the topic of the paper? It is not a kernel used in your network, it not a 3D kernel, and it is a well used approximation to the Laplacian.

Sketches (or selected frames from video sequences) of gestures would improve the presentation.

Your recording device is a simple webcam, while the work [6] uses a depth camera which makes it clearly inapplicable to your setup. This makes the whole introductory paragraph of Section IV. redundant, and the explanation in the second paragraph makes no sense at all.

Intermixing implementation details, code details, and interface reduces the clarity of the presentation. ISPA is a technical conference on image analysis so the presentation should focus on algorithmic parts of the analysis pipeline.

Reviewer #6

Questions

1. Overall evaluation

weak accept

2. Relevance to the conference

relevant

4. Review

The authors developed a neural network that classifies gestures from short videos. They provide a decent introduction with the literature overview. They explain their dataset and data processing procedure. The authors explained their architecture in detail and provide experiment results. Their work and contributions are sufficient.

However, I have some remarks on the way the paper was written. The paper lacks references and an introduction into what are "cobots". While explaining the dataset authors don't mention if the dataset is now publicly available and how did they split the data into train/test subsets. The model is well explained. The training procedure lacks clarity. What are "packs", are they the same as batches or? Why do you train on each pack three times "in a row". Why don't you use epochs, what is the difference? Experimental results are good, but it is not clear what is the test set. The paper lacks references on the OpenCV and Tkinter libraries. Figures and text is written in English, but on figure 4 and 5 there are some french words. It is confusing and should be in English. In section Experiments results there is a sentence:

" On the live test, we often got a precision of +90%." This sentence is of no use since it does not provide a precision number. It sais "often got a precision of +90%". What does often mean, what +90% means?