User-Guided Global Explanations for Deep Image Recognition: A User Study

Mandana Hamidi-Haines[1], Zhongang Qi[2], Alan Paul Fern[1], Li Fuxin[1], and Prasad Tadepalli[1]

[1]Oregon State University
[2]Tencent

June 25, 2021

Abstract

We study a user-guided approach for producing global explanations of deep networks for image recognition. The global explanations are produced with respect to a test data set and give the overall frequency of different "recognition reasons" across the data. Each reason corresponds to a small number of the most significant human-recognizable visual concepts used by the network. The key challenge is that the visual concepts cannot be predetermined and those concepts will often not correspond to existing vocabulary or have labelled data sets. We address this issue via an interactive-naming interface, which allows users to freely cluster significant image regions in the data into visually similar concepts. Our main contribution is a user study on two visual recognition tasks. The results show that the participants were able to produce a small number of visual concepts sufficient for explanation and that there was significant agreement among the concepts, and hence global explanations, produced by different participants.

Hosted file

interative_naming_appliedailetter_2021.pdf available at https://authorea.com/users/422024/articles/527826-user-guided-global-explanations-for-deep-image-recognition-a-user-study