

In this paper, we aim to learn discriminative representation for facial action unit (AU) detection from large amount of videos without manual annotations. Inspired