World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:16, No:10, 2022

Enhanced Multi-Scale Feature Extraction Using a DCNN by Proposing Dynamic Soft Margin SoftMax for Face Emotion Detection

Authors : Armin Nabaei, M. Omair Ahmad, M. N. S. Swamy

Abstract : Many facial expression and emotion recognition methods in the traditional approaches of using LDA, PCA, and EBGM have been proposed. In recent years deep learning models have provided a unique platform addressing by automatically extracting the features for the detection of facial expression and emotions. However, deep networks require large training datasets to extract automatic features effectively. In this work, we propose an efficient emotion detection algorithm using face images when only small datasets are available for training. We design a deep network whose feature extraction capability is enhanced by utilizing several parallel modules between the input and output of the network, each focusing on the extraction of different types of coarse features with fined grained details to break the symmetry of produced information. In fact, we leverage long range dependencies, which is one of the main drawback of CNNs. We develop this work by introducing a Dynamic Soft-Margin SoftMax.The conventional SoftMax suffers from reaching to gold labels very soon, which take the model to over-fitting. Because it's not able to determine adequately discriminant feature vectors for some variant class labels. We reduced the risk of over-fitting by using a dynamic shape of input tensor instead of static in SoftMax layer with specifying a desired Soft- Margin. In fact, it acts as a controller to how hard the model should work to push dissimilar embedding vectors apart. For the proposed Categorical Loss, by the objective of compacting the same class labels and separating different class labels in the normalized log domain.We select penalty for those predictions with high divergence from ground-truth labels.So, we shorten correct feature vectors and enlarge false prediction tensors, it means we assign more weights for those classes with conjunction to each other (namely, "hard labels to learn"). By doing this work, we constrain the model to generate more discriminate feature vectors for variant class labels. Finally, for the proposed optimizer, our focus is on solving weak convergence of Adam optimizer for a non-convex problem. Our noteworthy optimizer is working by an alternative updating gradient procedure with an exponential weighted moving average function for faster convergence and exploiting a weight decay method to help drastically reducing the learning rate near optima to reach the dominant local minimum. We demonstrate the superiority of our proposed work by surpassing the first rank of three widely used Facial Expression Recognition datasets with 93.30% on FER-2013, and 16% improvement compare to the first rank after 10 years, reaching to 90.73% on RAF-DB, and 100% k-fold average accuracy for CK+ dataset, and shown to provide a top performance to that provided by other networks, which require much larger training datasets.

Keywords : computer vision, facial expression recognition, machine learning, algorithms, depp learning, neural networks
Conference Title : ICANN 2022 : International Conference on Artificial Neural Networks
Conference Location : Los Angeles, United States
Conference Dates : October 27-28, 2022