

Mining the Displacement by Max-pooling in Convolutional Neural Networks

鄭, 煜辰

<https://hdl.handle.net/2324/4110518>

出版情報 : Kyushu University, 2020, 博士 (学術), 課程博士
バージョン :
権利関係 :

氏 名 : 鄭 煜辰

論 文 名 : Mining the Displacement by Max-pooling in Convolutional Neural Networks
(畳み込みニューラルネットワークにおける最大プーリングによる
変位検出)

区 分 : 甲

論 文 内 容 の 要 旨

The max-pooling operation is a common step in modern deep convolutional neural networks (CNNs), which is often introduced to obtain translation-invariant representations and downsample the feature maps of convolutional layers. However, in doing so, it loses the spatial information of the maximums. In this thesis, a novel feature is extracted from the max-pooling operation in CNNs, called displacement features. The displacement features record the location coordinates of the maximums in pooling windows of the max-pooling operation, which represents the “inter-class” and “intra-class” micro differences between different samples. Then, the class-wise trends and behaviors of the displacement features are discovered and analyzed in different ways. To verify the effectiveness of the displacement features, the displacement features are applied on two classical tasks, text recognition and offline signature verification. For text recognition tasks, the displacement features are extracted from the max-pooling layer and combined with the features resulting from max-pooling to capture the inter-class micro differences between the similar classes. The displacement features compensate for spatial information lost in the traditional max-pooling operation helps discriminate unnecessary absorptions from necessary absorptions. The extensive experiments and discussions on three text datasets, MNIST, HASY, and Chars74K-font datasets demonstrate that the proposed displacement features can improve the performance of the CNN based architectures and tackle the issues with the micro differences of max-pooling in the text recognition tasks. For offline signature verification tasks, the displacement features of the maximums in the max-pooling operation are extracted and fused with the pooling features to capture the intra-class micro differences between the genuine signatures and skilled forgeries as a feature extraction procedure. The displacement features represent the crucial differences between the genuine signatures and their corresponding skilled forgeries, which is useful for verification systems. The extensive experimental results and analysis on GPDS-150, GPDS-300, GPDS-1000, GPDS-2000, and GPDS-5000 datasets demonstrate that the proposed method can discriminate the genuine signatures and their corresponding skilled forgeries well and achieve state-of-the-art performance on these datasets.