DOCTORAL THESIS

Advances in Long-Tailed Visual Recognition

LI, Mengke

Date of Award:
2022

Advances in Long-Tailed Visual Recognition

Abstract

Real-world data typically have a long-tailed distribution. Existing classification models that perform well on artificially balanced datasets suffer severe performance degradation on long-tailed datasets. This thesis presents three methods from different perspectives to address the issues in long-tailed visual recognition.

Key point sensitive (KPS) loss is proposed to address the issue that a deep model is prone to correctly classify the head classes while ignoring the tail classes. KPS loss assigns relatively large margins to tail classes to relieve this bias. In addition, we find that key points are more important for classification. Therefore, KPS loss simultaneously regularizes the key points strongly. Furthermore, the gradient signals of stimulus and inhibit samples for each class are re-balanced via the proposed gradient adjustment (GA) optimization strategy. This GA strategy can circumvent excessive negative signals on tail classes. KPS loss with GA significantly improves the overall classification accuracy on tail classes but sacrifices a small amount of head class accuracy.

Feature-balanced loss (FBL) is proposed to address the limitation in KPS loss and study the effect of feature norm on classification. We observe that large feature norms are helpful to obtain clear class margins and thereby propose the novel FBL, which adds an extra class-based stimulus to the logit. The class-based stimulus encourages large norms for tail classes. Moreover, we gradually increase the stimulus intensity in a curriculum learning manner. This robust training strategy helps to boost the classification performance and enables the model to be trained end-to-end. The proposed FBL incorporated with curriculum learning achieves considerable performance gains in middle and tail classes while maintaining competent performance in head classes.

Large feature norms cannot completely balance the spatial distribution of each class in feature space. This thesis further proposes the Gaussian clouded logit (GCL), which studies the effect of softmax saturation on long-tailed learning. GCL perturbs different class logits with varied amplitudes to make the loss function have different degrees of softmax saturation for each class. The tail classes are set with relatively large amplitudes to decrease softmax saturation. Therefore, samples of tail classes are more active, and their embedding space can be enlarged. To alleviate the bias in classifier, the class-based effective number (CBEN) sampling strategy is proposed. Classifier re-training with this sampling strategy can further boost the model performance. GCL with CBEN achieves superior performance compared with other state-of-the-art methods.

Comprehensive evaluation and comparison are conducted on various benchmarks. Visualization experiments and in-depth discussions of the proposed methods are provided. Experimental results demonstrate the superiority of the proposed methods.

Keywords: Long-tailed classification, Imbalanced learning, Loss modification, Logit adjustment