

Self-Training With Noisy Student Improves ImageNet Classification

備考

著者

Qizhe Xie, Minh-Thang Luong, Eduard Hovy, Quoc V. Le

掲載

Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10687-10698, 2020.

Abstract

本論文では、ImageNetにおいて88.4%の上位1位精度を達成する自己学習手法を提案する。これは、35億枚の弱ラベル付きInstagram画像を必要とする最先端モデルよりも2.0%高い精度となる。また、ロバストネステストセットでは、ImageNet-Aのtop-1精度を61.0%から83.7%に、ImageNet-Cの平均破損誤差を45.7から28.3に、ImageNet-Pの平均反転率を27.8から12.2へと改善することができました。

この結果を得るために、まずラベル付きImageNet画像に対してEfficientNetモデルを学習し、それを教師として3億枚の未ラベル画像に対して擬似ラベルを生成します。次に、ラベル付き画像と擬似ラベル付き画像の組み合わせに対して、より大きなEfficientNetを生徒モデルとして学習させます。このプロセスを、生徒を教師として戻すことで繰り返す。擬似ラベルの生成中、教師は擬似ラベルができるだけ正確になるように、ノイズを与えません。しかし、生徒の学習時には、ドロップアウト、確率的深度、RandAugmentによるデータ増大などのノイズを生徒に注入し、生徒が教師よりもうまく汎化できるようにする。