

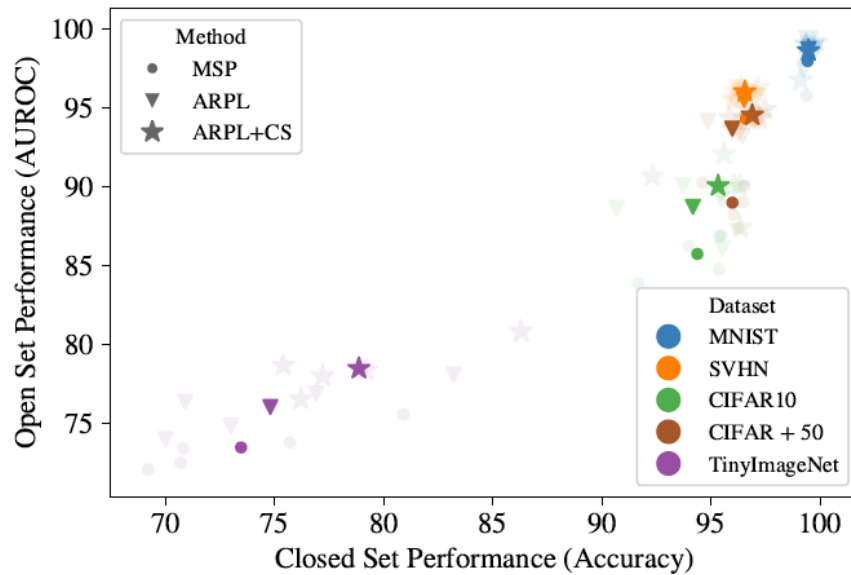
OPEN-SET RECOGNITION: A GOOD CLOSED-SET CLASSIFIER IS ALL YOU NEED?

- 概要
 - Open-set recognition (OSR)
 - テスト時で訓練時に遭遇していないクラスを観測した時に、それを無知と区別するタスク
 - The standard baseline for OSR - maximum softmax probability (MSP) baseline
 - 訓練時
 - クラスのクロスエントロピー損失
 - テスト時
 - 入力が既存のクラスに属しているか、属していないかをソフトマックス関数による確率の形で出力する
 - 一言
 - 閉集合と開集合のパフォーマンスは密接に関連しており、これは様々なデータセット、目的関数、モデル構造に対して言えることがこの論文からわかった
 - 閉集合のパフォーマンスを向上させることが重要
 - 強力なデータ拡張
 - 学習率のスケーリング
 - label smoothing (ラベル正則化)
 - ラベル平滑化では、CNNなどの出力層(クラス信頼度のsoftmax確率)を、交差エントロピー損失関数(負の対数損失)を用いて学習す

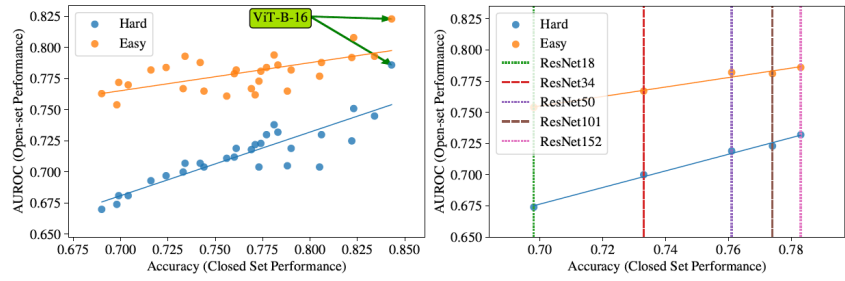
る際に、まず負のエントロピーを正則化項として加える。これにより、 $\{0,1\}$ だけで構成される「ハードな」正解のラベルに対して、パラメータ ϵ の割合だけハードラベルから差し引いた確率値を、全クラスに分配して小さなノイズ値として追加する。すると、元のハードなsoftmax確率の変わりに、「滑らかな」**ソフトラベル**（ラベル確率ベクトルの教師ラベル）に代替させて学習できる。そのおかげで、ハードラベルへの過剰適合を防ぐことができるという仕組みである

- 開集合
 - 正則化されたソフトマックス確率の代わりにmaximum logit score (MLS)を提案した
- 評価指標
 - Semantic Shift Benchmark suite (SSB)
- テストデータが閉集合からなる
 - 訓練時に学習したクラスに関して、クラス確率を出力する
- テストデータが開集合からなる
 - クラス確率に加えて、入力既存のクラスに当てはまるかどうかのスコアを出力する
- ベースライン
 - Maximum Softmax Probability (MSP, baseline)
 - 訓練時
 - targetのラベルとソフトマックス関数の出力のクロスエントロピーで学習
 - テスト時
 - ソフトマックス関数による最大確率をスコアとする
 - ARPL
 - RPL (Reciprocal Point Learning) optimization strategyの拡張
 - 潜在空間上の「reciprocal point」との距離を測る

- 未知のクラスについては既存のクラスの「reciprocal point」との距離が遠くなる
 - ユークリッド距離とコサイン距離（コサイン類似度？）の和をスコアとした
- ARPL + CS
 - 未知のクラスの「reciprocal point」を擬似的に作成する
- 実験
 - 小規模
 - データセット
 - データセットの一部のクラスでもって学習させ、テスト時に既存のクラスと使用していないクラスを用いてOSRの精度をみる
 - MNIST, SVHN, CIFAR10 - 全て10クラス
 - これを6クラスを訓練に、4クラスが未知
 - CIFAR+N
 - CIFAR10の中から4クラスを用いて学習し、CIFAR100から10もしくは50クラスが未知
 - TinyImageNet
 - ImageNetから200クラスを抽出したもの
 - 20クラスを学習に用いて、180クラスが未知
 - テスト時に入力が既知であるか、もしくは未知であるかをバイナリで出力する
 - threshold-free area under the Receiver-Operator curve (AUROC)を評価指標として使用
 - 結果
 - 閉集合と開集合について正の相関がある

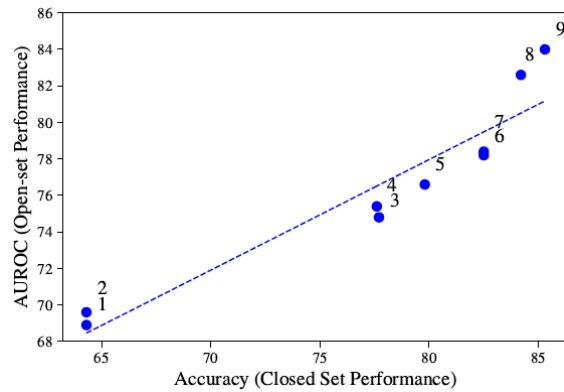


- 大規模
 - モデル
 - VGG, ResNet, EfficientNet, ViT, MLP-Mixer
 - Imagenet-21kで事前学習し、Imagenet-1kでファインチューンしたもの
 - データセット
 - Imagenet-21k-P
 - Imagenetから11kほどのクラスを選んだもの
 - その中から、Imagenet-1Kに現れない1000クラスを二つ用意する。そのうち切り分けるのが簡単なデータセットを「Easy」、難しい方を「Hard」とした
 - 結果
 - ViTだけeasy, hardについて他のモデルのような精度の挙動は示さなかった



- baselineの強化

Epochs	Scheduler	Setting					Closed-Set (Accuracy)	Open-Set (AUROC)	Combined (OSCR)
		Aug.	Logit Eval	Warmup	Label Smoothing	Ensemble			
100	Step	RandCrop	✗	✗	✗	✗	64.3	68.9	51.4
100	Step	RandCrop	✓	✗	✗	✗	64.3	69.6	50.7
200	Cosine (0)	RandCrop	✓	✗	✗	✗	77.7	74.8	64.3
200	Cosine (0)	CutOut	✓	✗	✗	✗	77.6	75.4	64.7
200	Cosine (0)	RandAug	✓	✗	✗	✗	79.8	76.6	67.3
600	Cosine (2)	RandAug	✓	✗	✗	✗	82.5	78.2	70.3
600	Cosine (2)	RandAug	✓	✓	✗	✗	82.5	78.4	70.3
600	Cosine (2)	RandAug	✓	✓	✓	✗	84.2	83.0	74.3
600	Cosine (2)	RandAug	✓	✓	✓	✓	85.3	84.0	76.1



- maximum logit score (MLS)

- 最終層でSoftmax関数に入力する前の出力のこと