

画像分類におけるTriplet Networkを用いた 半教師あり学習手法の改良*

半田光理 秦野亮 西山裕之
東京理科大学理工学研究科経営工学専攻

1 はじめに

近年、深層学習をはじめとして注目を集めている(教師あり)機械学習は、モデルを訓練するために多量のラベル付きデータが必要となる。しかし、データへのラベル付けは通常手作業で行われるため、多大な手間とコストが必要となる問題がある。そうした問題を解決するための手法のひとつとして、Triplet network(以下TN)に基づく半教師あり学習手法が存在する[1]。本手法では、部分的にラベル付けがされたデータセットから高い分類性能を持つ(TNの)モデルを学習するために、学習時の重みパラメータの調整をデータ間の類似性を表すユークリッド距離に基づく損失関数によって行う。しかし多クラス分類を行おうとしたとき、特に3クラスを超えると分類性能の低下が著しくなるという問題があった。そこで本研究では、上記の問題の克服を目標とした半教師あり学習の改良手法を提案する。具体的には、特徴空間におけるデータのばらつきを考慮したTNの新たな損失関数を与え、その効果をCifar-10データセットに基づく画像分類のタスクを用いて示す。

2 Semi-supervised learning with Triplet network

HandaらはTNを半教師あり学習法のひとつであるSelf-trainingの予測モデルとして組み込んだSemi-supervised learning with Triplet network(以下SSL with TN) [1]を提案した。本手法は、二段階から構成される(Figure 1)。第一段階では少量のラベル付きデータからTripletを作成し、サンプリングの後に予測モデルを訓練する。SSL with TNは、まず少量のラベル付きデータから識別したいデータ x と同じクラスラベルを持つ x_+ 、異なるクラスラベル x_- の3つからなるTriplet (x, x_+, x_-) を作成し、サンプリングを行った後にTriplet networkの教師あり学習を行い、ベースとなる予測モデルの訓練を行う。次に、ラベル無しデータにラベルを割り当てるために、訓練済み

の予測モデルに対してラベル無しデータを入力して特徴ベクトルを出力し、それをランダムフォレスト分類器に入力することでラベルの予測確率の取得を行う。予測ラベルが高信頼性を持つと判断できる場合、そのラベルを新たに付与した元ラベル無しデータを訓練データに追加し、予測モデルの再訓練を行う。以上のプロセスをSelf-trainingと呼び、ラベル無しデータのサイズが収束するまで繰り返し行われる。Table 1に、1554のラベル付きデータでTNを訓練した結果と60のラベル付きデータで訓練した本手法の結果、及びその他の半教師あり学習法の結果の比較を示す。この実験ではランダムフォレスト分類器に、オリジナルの画像データの特徴ベクトルに変換して入力し、訓練したものをBase lineとして使用している。本手法に基づいて得られたモデルは、3クラス分類においてBase lineより優れた分類性能を持ち、近年注目を集めている他の半教師あり学習手法に対して競争的な性能を得ていることがわかる。しかし、10クラス分類においては、良い性能は得られていない。

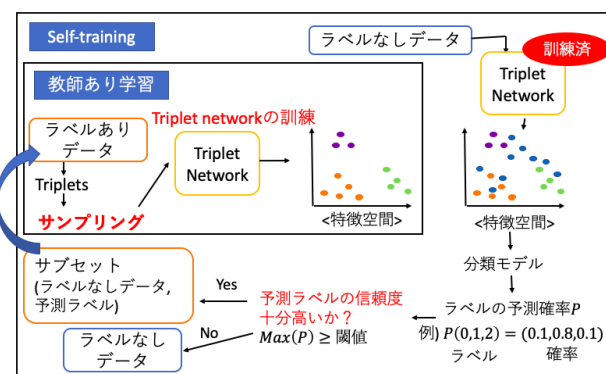


Figure 1: SSL with TNのアウトライン

Table 1: 3クラス分類の実験結果

model	precision	recall	F-measure	Test Accuracy	AUC
Base line (3 classes)	0.600	0.590	0.580	0.588	0.691
TN (1554 labeled data, 3 classes)	0.830	0.820	0.810	0.810	0.864
GAN (3 classes)	0.570	0.560	0.560	0.560	0.674
GAN with manifold learnin(3 classes)	0.750	0.750	0.740	0.745	0.809
SSL with TN (3 classes)	0.770	0.760	0.760	0.763	0.823
SSL with TN (10 classes)	0.270	0.300	0.250	0.284	0.602

*Revised semi-supervised learning with triplet network for image classification

3 提案手法

本研究の提案手法は、SSL with TNを引き継ぎ、Triplet サンプルングと損失関数の2つに着目し、改良を加える。従って、本提案手法のアウトラインはSSL with TNと同様(cf. Figure 1)であり、本研究のオリジナリティは前述の2つに帰着される。

Tripletは識別の難易度からEasy-negative、Hard-negative、Semi-hard-negativeに分類することができ、損失関数で上手く最適化するためにSemi-hard-negativeのみをサンプルングして使用する。サンプルング後のデータ空間は元のデータ空間と異なることに注意したい。TNは従来多量のラベル付きでデータで訓練されてきたが、本研究は少量のラベル付きデータを使用するため、サンプルングの影響を強く受け、各クラスのサンプル数に大きく偏りが発生する可能性がある。そこで、本研究はサンプルングの条件を緩く設定し、Easy-negativeとHard-negativeの影響を次の損失関数を定義することにより、制御する。いま、識別したい入力データのモデル出力を a 、類似データのモデル出力を p 、非類似データのモデル出力を n とすると、損失関数Lossは以下の様に定義される。

$$\text{Loss} = (1 - u) \times \text{Max}((d_{a,p} + \text{std}_{a,p} - d_{a,n} - \text{std}_{a,n} + \text{margin}), 0) + u \times \text{Max}((d_{a,p} + \text{std}_{a,p}), 0)$$

ただし、 $d_{a,p}$ は類似間距離、 $d_{a,n}$ は非類似間距離、 $\text{std}_{a,p}$ は類似データ間の特徴ベクトルの標準偏差、 $\text{std}_{a,n}$ は非類似データ間の特徴ベクトルの標準偏差、 u 及び margin はハイパーパラメータである。従来のTriplet lossはユークリッド距離を用いて類似度を計算し、類似間距離 $d_{a,p}$ を0に、非類似間距離 $d_{a,n}$ を1に近づけるように学習を行う。しかし、ユークリッド距離のみでは、各クラス内のデータのばらつきを考慮していないため、本研究は、従来のTriplet lossに $\text{std}_{a,p}$ と $\text{std}_{a,n}$ の項を追加したものを新たに用いる。

4 実験と考察

提案手法の効果を調べるために、600,000枚の画像データからなるCifar-10データセットから、少数のラベル付きデータと多数のラベル無しデータからなる実験用データセットを作成し、用いる。本研究は、SSL with TNとモデルの性能比較を行うため、10クラス分類、3クラス分類共に、同じデータセットを用いる。具体的に、ラベル付きデータは各クラス20のサンプル、ラベルなしデータは各クラス498のサンプルを用いた。3クラス分類タスクと10クラス分類タスクにおける実験結果をTable 2と3にそれぞ

れ示す。比較対象として、Base line (ランダムフォレスト)及び多量のラベル付きデータの識別結果を示す。Table 2から、本研究の提案手法は先行研究のSSL with TNよりも優れており、多量のラベルありデータで訓練されたモデルの識別精度と近い識別精度を持つことがわかる。対して、Figure 3に示される10クラス分類タスクは、比較対象であるベースライン及び多量のラベル付きデータで訓練されたモデルの識別精度が両方とも低い結果となっている。つまり、10クラス分類タスクは、3クラス分類タスクに比べて非常に分類の難しいタスクであることがわかる。従来多量のラベル付きデータで訓練されてきたTNは、少数のラベル付きデータの訓練に適応していないことがわかる。提案手法の識別精度は、多量のラベル付きデータで訓練されたTNよりも高い識別精度を達成した。これは、従来のユークリッド距離のみを用いたTriplet lossよりも、標準偏差を追加した新たな損失関数が少数のラベル付きデータを用いた学習にTNを適応させるために有効であることを示唆している。本研究の損失関数がある理由は、類似データ間の特徴ベクトルの標準偏差を小さくすることで、ユークリッド距離のみでは不十分であったデータ内のばらつきを考慮することができ、類似データ間の特徴ベクトルをベクトルの要素単位で近づけることができたからであると考えられる。

Table 2: 3クラス分類の実験結果

model	precision	recall	F-measure	Test Accuracy	AUC
Base line	0.600	0.590	0.580	0.588	0.691
TN (1554 labeled data)	0.830	0.820	0.810	0.810	0.864
SSL with TN	0.770	0.760	0.760	0.763	0.823
提案手法	0.790	0.780	0.780	0.782	0.846

Table 3: 10クラス分類の実験結果

model	precision	recall	F-measure	Test Accuracy	AUC
Base line	0.280	0.280	0.270	0.275	0.597
TN (5810 labeled data)	0.310	0.310	0.300	0.310	0.617
SSL with TN	0.310	0.300	0.300	0.301	0.612
提案手法	0.340	0.350	0.340	0.301	0.637

5 おわりに

本研究ではラベル付けのコストを削減するため、少量のラベル付きデータで高精度を達成する半教師あり学習手法を提案することを目的に、SSL with TNを改良した。データのばらつきを考慮した新たな損失関数があることが示されたため、同じくデータの分散を考慮したマハラノビス距離を適用することがさらなる発展として考えられる。

参考文献

- [1] Chien-Liang Liu Handa Hikari. “Semi-supervised learning with Triplet network”. In: *Master Thesis, NCTU*. 2019.

講演原稿の訂正箇所

発表 I D : 6R-04

発表タイトル：画像分類におけるトリプレットネットワークを用いた
半教師あり学習手法の提案

著者：半田光理，秦野 亮，西山裕之（東理大）

○ 訂正箇所

- ・参考文献に以下の研究論文を加えました。

Chien-Liang Liu Handa Hikari. "Semi-supervised learning with Triplet network". In: Master Thesis, NCTU. 2019.

- ・当初入れておりました 2 件の参考文献の情報は、上記の参考文献に内包されておりますので、割愛いたしました。

- ・第 3 節の提案手法を上記の参考文献の内容を改良させる内容として修正しました。

- ・第 5 節の実験と考察の内容も、本提案手法に基づく内容に更新しました。

- ・その他、各節においても、若干の加筆修正を行っております。

○ 変更理由

我々の手違いにて、上記の参考文献の内容を中心とした講演原稿を投稿してしまいました。そのため、このままでは著者情報に差異が生じてしまいますので、当初の講演目的の内容、そして著者情報に合わせた原稿の内容に変更させていただきたく、上記の変更を行わせていただきました。