

移転メトリック学習に基づいたマルチタスク学習モデルの開発

Development of Multitask Learning Model Based on Transfer Metric learning

○学 岳 思濛 (神戸大) 准 小澤 誠一 (神戸大)

Simeng Yue, Kobe University

Seiichi OZAWA, Kobe University

Multi-task learning is a learning paradigm which allows a system to learn related multiple tasks in parallel or sequentially and to improve the generalization performance for a learning task by transferring the knowledge to other related tasks. In order to transfer knowledge effectively, it is necessary to estimate the relatedness between tasks properly. In this paper, we propose a multi-task learning algorithm for pattern recognition where the task relatedness is estimated based on the metric learning. Experimental results demonstrate the effectiveness of our method.

Key Words: multi-task learnig, pattern recognition, distance metric learning, incremental learning

1. はじめに

マルチタスク学習 [?] とは、関連のある複数のタスク (問題) を同時または逐次的にシステムに学習させ、各タスクを個別に学習するよりも、少ない訓練データで高い性能を得ようとする学習方式である。現実の環境では十分な訓練データが与えられない場合はよくある。そのため、過去の学習で得た知識を移転して、新しいタスクを高速に学習することのできる仕組みを導入することが望ましい。また、知識移転を効果的に行うためには、新しく学習するタスクに、過去のどの知識を移転するのかを適切に決める必要がある。このため、タスク間の関連度を正確に見積もることが、マルチタスク学習の研究においては重要な課題となっている。

本研究では、Zhang ら [?] が提案している Transfer Metric Learning に基づき、タスク間関連度をタスク共分散行列で表す方法を取り入れ、Resource Allocating Network with Long Term Memory (RAN-LTM)[?] のマルチタスク学習への拡張を行う。

2. 逐次マルチタスク学習

逐次マルチタスク学習は、学習すべきタスクが動的に変動するオンライン学習に分類される。ここでいうタスクとは、一つ一つのパターン認識の問題を指す。また、本研究では、各タスクの訓練データは一つずつ、または少数のかたまりでシステムに提示されると仮定する。また、本研究で用いる訓練データには、その訓練データがどのタスクに属するかを示すタスクラベルは含まれないとする。したがって、学習すべきタスクが変動したことを、システムは自動的に検知する必要がある。新しいタスクが検知される

2.1 タスク変動検知

識別器が一つ目のタスクの学習をある程度完了したと仮定する。ある程度とは、与えられる入力値に対して、なんらかの識別結果を返す程度である。このとき、ある入力に対して出た識別結果と訓練データとの誤差が非常に大きいとき、識別器はタスクの変動検知を行うことが出来る。しかし、実際には識別面付近のデータに対しては、タスクが変動してなくても誤差が大きくなることもありうる。したがって、誤差の大きな一つの訓練データからだけではタスク変動検知は行わず、タスク変動が起こったと識別器が予測した時点から、いくつかの訓練データにわたって誤差が大きくなったときにタスクの変動を検知する。

2.2 RAN-LTM

RAN とは Generalized Radial Basis Function を基本構造にもつ Radial 基底関数を用いた 3 層ネットワークで構成される。RAN は未学習領域の訓練データが提示されたとき逐次的に新しい隠れユニットを追加していくことで近似能力を高めることができる。Resource Allocating Network with Long-Term Memory (RAN-LTM) とは、長期記憶 (Long-Term Memory (LTM)) と呼ばれる記憶モジュールから構成される。このモデルでは RAN の学習によって獲得された入出力関係から記憶アイテムを抽出し、LTM に長期的に格納される。新しい訓練データの学習において、干渉の抑制に効果的な記憶アイテムのみが想起される。干渉とは、新しく提示された訓練データに対して結合加重が修正された際、過去に獲得された入出力関係が破壊されてしまうことをいう。RAN-LTM では記憶アイテムを訓練データと同時に学習することで、効率よく干渉を抑えることが出来る。

2.3 記憶アイテム

不要な記憶データの生成をさけるために、RAN-LTM では、近似精度が十分に高まった領域から代表的な入出力が記憶アイテムとして抽出される。RAN の学習において、隠れユニットの出力が大きく、訓練データに対する出力誤差が十分に小さいとき、そのラジアル基底中心付近の領域に対する近似精度は高まっていると考えられる。そこで、活性度の高い隠れユニットのラジアル基底中心を RAN の入力としたときの出力を求め、入出力対を記憶アイテムとして長期記憶に格納する。

3. Multi-task Metric Learning

Multi-task Metric Learning [?] とは、事後確率分布の最大化によってメトリック行列のパラメータ値を推定する方法である。事後確率分布を最大化するため、尤度と事前確率を最大化すればよいことがわかっている。尤度の最大化は損失関数の最小化に置き換えて最適化される。ここでパラメータはメトリック行列であり、事前確率は Matrix Variate Normal Distribution と呼ばれる手法によって、以下の最適化問題に置き換えて求められる。

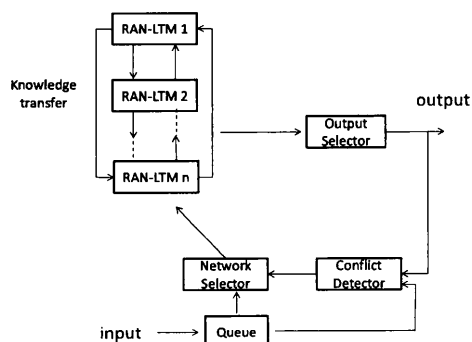


Fig. 1: マルチタスク学習システムの構造

$$\min \sum_{i=1}^m \frac{2}{n_i(n_i-1)} \sum_{j < k} g(y_{j,k}^i [1 - \|x_j^i - x_k^i\|_{A_i}^2]) + \frac{\lambda_1}{2} \sum_{i=1}^m \|A_i\|_F^2 + \frac{\lambda_2}{2} \text{tr}(U\Omega^{-1}U^T) \quad (1)$$

$$A_i \geq 0, \quad \Omega \geq 0, \quad \text{tr}(\Omega) = 1 \quad (2)$$

但し, $U = (\text{vec}(A_1), \dots, \text{vec}(A_m))$, g は損失関数であり, $y_j^i = y_{j,k}^i$ のとき $y_{j,k}^i = 1$, それ以外では $y_{j,k}^i = -1$ となる. n_i はタスク i のデータ数である. x_j^i はタスク i の j 番目のデータである. A_i は各タスクで求めたメトリック行列である. vec は, 行列をベクトルに変換する関数である. Ω は, 各タスクで求められたメトリック行列をベクトルに変換し, それらの共分散行列で与えられる. つまり, タスク間の関連度が求められる. 式 (??) の最適化問題を解くことにより, 各タスクのメトリックとタスク間の関連度が得られる.

4. Transfer Metric Learning

Multi-task Metric Learning に基づいて Transfer Metric Learning (TML) が提案されている. TML では, すべてのタスクが十分に学習してから, 新しいタスクが追加されるとする. このとき新しいタスクをターゲットタスク, 以前に学習されたタスクをソースタスクと呼ぶ. ターゲットタスクはソースタスクから求めたメトリックを利用して学習する. ターゲットタスクの学習は, 以下の最適化問題で定義される.

$$\min \frac{2}{n_m(n_m-1)} \sum_{j < k} g(y_{j,k}^m [1 - \|x_j^m - x_k^m\|_{A_m}^2]) + \frac{\lambda_1}{2} \|A_m\|_F^2 + \frac{\lambda_2}{2} \text{tr}(U\Omega^{-1}U^T) \quad (3)$$

$$A_m \geq 0, \quad \Omega \geq 0, \quad \text{tr}(\Omega) = 1 \quad (4)$$

但し, $U = (\text{vec}(A_1), \dots, \text{vec}(A_m))$ で m はタスク数である. $m-1$ 個のソースタスクを仮定しているため, ターゲットタスクは m 番目のタスクで表されることに注意する. また, ソースタスクは互いに独立であると仮定し, 重要度は等しいとする. このとき Ω は以下のように表わされる.

$$\Omega = \begin{pmatrix} \alpha I_{m-1} & w_m \\ w_m^T & w \end{pmatrix} \quad (5)$$

ここで,

$$\alpha = \frac{1-w}{m-1} \quad (6)$$

式 (??) の最適化問題を解くことで, ターゲットタスクとソースタスク間の関連度, さらにターゲットタスクのメトリック行列を求めることができる.

5. 提案手法

本研究では, タスク間の関連度を求めるために TML を利用する. TML を利用する理由は 2 つあり, 1 つはタスク間の関連度を求めて, 知識移転を行うためである. もう 1 つはタスクごとに最適なメトリックを学習することで, 識別性能を向上させるためである.

図 2 に提案手法のアーキテクチャを示す. 知識移転は, ソースタスクで求められたメトリック行列をターゲットタスクの初期値とすることで実現される. これにより, タスクごとに最適化されたメトリックを求められることを期待している. そして, 各タスク学習したメトリックを RAN-LTM の RBF に適用することにより識別性能を向上することができる. 式は以下のように表す.

$$y = \exp\left(-\frac{(x-c)^T A(x-c)}{2\sigma^2}\right) \quad (7)$$

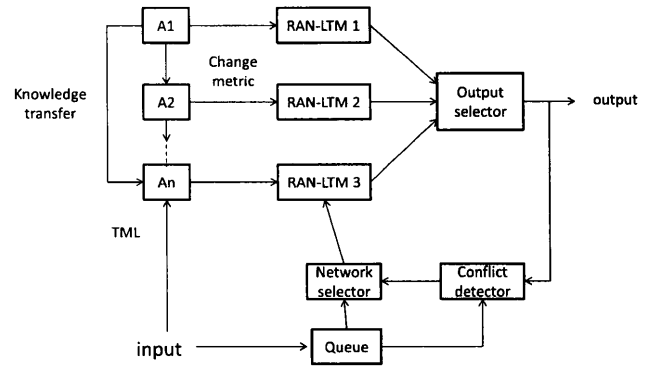


Fig. 2: 提案手法システムの構造

Table 1: タスク関連度 (Vowel データ)

	Task 1	Task 2	Task 3
Task 1	1.0	0.48	0.18
Task 2	0.48	1.0	0.22
Task 3	0.18	0.22	1.0

6. 評価実験

計算機実験によって, 提案手法の有効性を示す. 実験には UCI データセットに含まれる Ann データのほか, Vowel, Thyroid データを用いた. これらデータをマルチタスク学習に拡張するため, 複数のクラスを統合し, 異なるクラスラベルを割り当てて, 元々の認識タスク以外に 2 つのタスクを定義した.

表 ?? に, Vowel データに対して求められた関連度を示す. 関連度は 1 に近いほど, タスク間に関連があることを意味する. 表 ?? に, タスク関連度に基づいてメトリック学習を導入した提案手法 (RAN-LTM+TML) とタスク関連度を用いず, タスクごとにユークリッド距離に基づいて学習する RAN-LTM の認識率を示す. メトリック学習を導入することで, 認識性能が向上していることがわかる.

Table 2: 各データの認識率 [%]

	RAN-LTM	RAN-LTM+TML
Ann	87.0	94.1
Vowel	59.9	64.0
Thyroid	83.8	94.0

7. まとめ

本研究では, タスク間の関連度を求めて知識移転を行う学習アルゴリズムを提案した. また, メトリック学習により RAN-LTM の識別性能が改善されることを示した.

文献

- [1] S. Ozawa, A. Roy, and D. Roussinov, "A Multitask Learning Model for Online Pattern Recognition," *IEEE Trans. on Neural Networks*, vol. 20, no. 3, pp. 430-445, 2009.
- [2] Y. Zhang and D.-Y. Yeung, "Transfer Metric Learning by Learning Task Relationships," *Proc. of 16th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, pp. 1199-1208, 2010.

- [3] S. Ozawa, S. L. Toh, S. Abe, S. Pang, and N. Kasabov, "Incremental Learning of Feature Space and Classifier for Face Recognition," *Neural Networks*, vol. 18, nos. 5-6, pp. 575-584, 2005.