

# Random Erasing を用いた半教師あり学習の提案

岡鼻 雄飛<sup>1</sup> 後藤 佑介<sup>1</sup>

**概要:** 近年, 多種多様な大規模データを計算機で分類するため, 機械学習に対する注目が高まっている. 一般的な機械学習の方法として, 多数のクラスラベル付き訓練事例 (以下, ラベルデータ) を用いてデータを分類する教師あり学習があり, 教師あり学習の処理性能を向上させるため, Random Erasing と呼ばれる手法を用いることが有効である. しかし, 教師あり学習では多数のラベルデータが必要であり, 人力でクラスラベル無し訓練事例 (以下, 未ラベルデータ) にラベル情報を追加するコストは非常に大きい. 本研究では, 少数のラベルデータと未ラベルデータを併用した半教師あり学習に Random Erasing を用いて高い分類精度を実現する手法を提案し, 既存手法と比較評価する.

**キーワード:** 半教師あり学習, Virtual Adversarial Training, Random Erasing

## 1. はじめに

近年, 大規模化したデータを処理するため, 計算機による機械学習が大きな注目を集めている. 一般的な機械学習の手法として, 多数のクラスラベル付き訓練事例 (以下, ラベルデータ) を用いる教師あり学習が挙げられる.

深層学習を用いた教師あり学習 [4], [5] では, 畳みこみニューラルネットワークのような複雑な構造をもつネットワークを使用することで, 既存の手法と比較して良い精度を得ることができる. 特に, 効率的に教師あり学習を行う手法である Random Erasing [14] では, 画像の分類を行うときに画像の一部を消去することで, 分類精度を向上できる.

しかし, 教師あり学習では, 多数のラベルデータが必要であり, 人力でラベル情報を追加するコストは非常に大きい. このため, ラベルデータとクラスラベル無し訓練事例 (以下, 未ラベルデータ) を併用して学習する半教師あり学習が研究されている [9], [10]. 深層学習を用いた半教師あり学習では, データの分類手法がいくつか提案されており, 高い分類精度を実現している [1], [6], [8], [11].

本研究では, 教師あり学習で用いられている手法である Random Erasing を用いた半教師あり学習を提案する. 提案手法では, 教師あり学習と半教師あり学習のアルゴリズムを組み合わせ, Random Erasing を半教師あり学習で利用できるように改良する. また, 提案手法を用いた場合

におけるデータの分類精度を評価する.

## 2. 学習方法

### 2.1 概要

はじめに, 教師あり学習, および半教師あり学習の仕組みを説明する.

### 2.2 教師あり学習

教師あり学習は, ラベルデータをもとに学習を行うことで, データを効率的に分類できる. 一方で, データの分類精度を向上させるために大量のラベルデータを用いるため, 人力でラベル情報を追加してラベルデータを作成する必要があり, コストが大きい.

### 2.3 半教師あり学習

半教師あり学習は, ラベルデータおよび未ラベルデータの両方を用いて学習を行う方法である. このため, 半教師あり学習で必要となるラベルデータの数は教師あり学習より少なくなり, ラベルデータの作成にかかるコストを削減できる. 一方で, 未ラベルデータを用いるため, 半教師あり学習の分類精度は教師あり学習に比べて低くなる.

## 3. 関連研究

半教師あり学習には多くの手法が存在する. 本章では, 深層学習を用いた半教師あり学習である Virtual Adversarial Training (VAT) [9], [10], および Neighbor Embedding [6] を説明する. また, 深層学習を用いない手法として, グラフ構造を用いる手法 [12] を説明する.

<sup>1</sup> 岡山大学大学院自然科学研究科  
Graduate School of Natural Science and Technology,  
Okayama University

### 3.1 Virtual Adversarial Training

Virtual Adversarial Training (以下, VAT) では, 正しい入力データ, および微小なノイズを追加した別の入力データとの間で Kullback Leibler (KL) divergence と呼ばれる情報量を最小化することで, 教師なし学習の誤差を定義する.

入力データを  $x$ , 入力データに対するラベルを  $y$ , ニューラルネットワークのパラメータを  $w$ , 微小なベクトルを  $r$  とすると,  $\Delta KL$  を

$$\Delta KL(r, x, w) = KL[p(y|x, w) || p(y|x+r, w)] \quad (1)$$

と定義する. このとき,  $r$  をランダムに作成せず,  $\Delta KL$  を最大化する  $r$  を

$$r_{v-adv} = \arg \max_r \Delta KL(r, x, w) \quad (2)$$

とする. このとき,

$$\Delta KL(r_{v-adv}, x, w) = KL[p(y|x, w) || p(y|x+r, w)] \quad (3)$$

で示される  $\Delta KL$  を最小化する. なお,  $r_{v-adv}$  は, 以下の式で求める. このとき,  $d$  はランダムに初期化した単位ベクトル,  $\beta$  および  $\epsilon$  はパラメータとする.

$$d = \nabla_r \Delta KL(r, x, w)|_{r=\beta d} \quad (4)$$

$$r_{v-adv} = \epsilon d \quad (5)$$

式 (5) で示す  $r_{v-adv}$  の値を最小化することで, 分類に適したパラメータを学習できる. また, 式 (5) において, ラベルデータではなく未ラベルデータを用いて  $r_{v-adv}$  を算出できる.

図 1 に, VAT の構成を示す. はじめに, 入力  $x$  に対してニューラルネットワークの出力となる  $network(w)$  を計算する. 次に, 式 (1)~(5) で述べた計算手順で  $r_{v-adv}$  を計算し,  $x+r_{v-adv}$  を入力として  $network(x)$  の出力を計算する. ここで, 教師あり学習における誤差 (supervised loss) は,  $x$  に対する  $network(w)$  の出力をラベル  $y$  に近づけるために用いる. 一方で, 教師なし学習の誤差 (unsupervised loss) は,  $x+r_{v-adv}$  に対する  $network(w)$  の出力を  $x$  に近づけるように最適化する. ここで, 教師あり学習の誤差は, クラスラベルが既知であるラベルデータの場合に計算できる. 一方で, 教師なし学習の誤差は, クラスラベルが未知である場合でも計算できる. このことから, VAT は, ラベルデータおよび未ラベルデータの両方を組み合わせることで半教師あり学習を行うことができる. VAT では, 評価用データセットに対して高い精度を実現しており, 教師あり学習に適用できるため, 学習手法として汎用性が高い.

MNIST に対して, VAT を用いて生成した画像を図 2 に示す. 上段は元の画像, 下段は生成した画像である. 図 2 より, VAT で生成した画像は, ノイズによる影響で全体的に不明瞭であることが分かる.

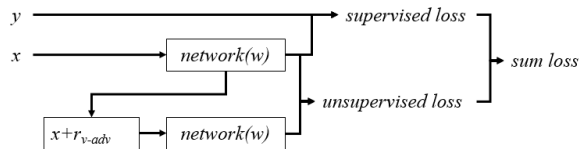


図 1 Virtual Adversarial Training (VAT) の構成



図 2 MNIST に対する VAT の適用例

### 3.2 Neighbor Embedding

Neighbor Embedding は, ユークリッド距離に基づく半教師あり学習手法である.  $x_l$  はラベルデータ,  $x_u$  は未ラベルデータ,  $z_c$  は  $x_l$  の中でクラス  $c$  に属するラベルデータとすると, 以下で最適化を行う式を定義する. なお,  $\lambda_L$  および  $\lambda_U$  は, 教師ありの誤差, および教師なしの誤差の割合をそれぞれ調整するパラメータである.

$$L(x_l, x_u, z_1, \dots, z_c) = \lambda_L L(x_l, z_1, \dots, z_c)_L + \lambda_U L(x_u, z_1, \dots, z_c)_U \quad (6)$$

$$L(x_l, z_1, \dots, z_c)_L = -\log \frac{e^{\|Net(x_l) - Net(z_k)\|^2}}{\sum_{i=1}^c e^{\|Net(x_l) - Net(z_i)\|^2}} \quad (7)$$

$$L(x, z_1, \dots, z_c)_U = -\sum_{i=1}^c \frac{e^{\|Net(x) - Net(z_i)\|^2}}{\sum_{j=1}^c e^{\|Net(x) - Net(z_j)\|^2}} * \log \frac{e^{\|Net(x) - Net(z_i)\|^2}}{\sum_{j=1}^c e^{\|Net(x) - Net(z_j)\|^2}} \quad (8)$$

Neighbor Embedding [6] における  $L(x_l, z_1, \dots, z_c)_L$  は, ラベルありデータに対する誤差を示している. 式 (7) では, ラベルが同じデータ同士の距離を短くするとともに, ラベルが異なるデータ同士の距離を長くする. 一方で,  $L(x, z_1, \dots, z_c)_U$  は, ラベルなしデータに対する誤差を示している. この誤差を最小化することで, 識別に適した空間への写像を学習できる. Neighbor Embedding では, 識別に適した空間への写像を学習した後に,  $k$  近傍法 [6] といった識別機を用いる. 以上より, Neighbor Embedding は, MNIST において十分に高い精度を実現する.

### 3.3 グラフ構造を用いる手法

グラフ構造を用いた半教師あり学習は, グラフ構築とラベル伝播の 2 段階に分類できる. グラフ構築では, すべて

のデータを頂点とし、各データ間の距離をもとに、 $k$  近傍グラフと呼ばれるグラフを作成する。ラベル伝播では、グラフ構築の段階で作成したグラフをもとに、ラベルデータから未ラベルデータにラベルを伝播して追加する。

グラフ構造を用いた半教師あり学習において、ラベルを伝播する手法は数多く提案されている。ここでは、Local and Global Consistency [15] (以下、LGC) を説明する。LGC では、 $k$  近傍グラフに対する隣接行列を用いて、未ラベルデータにラベルを追加する。

はじめに、ラベルの伝播に用いる行列  $D$ 、および  $S$  を計算する。 $D$  は式 (9) で、 $S$  は式 (10) でそれぞれ表される行列である。

$$D_{ii} = \sum_j A_{ij} \quad (9)$$

$$S = D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \quad (10)$$

このとき、データがどのラベルに属するかを予測した結果を行列  $F$  とする。行列  $F$  の初期値  $F(0)$  は、 $F(0) = Y$  とする。このとき、ラベルを予測する計算式を以下に示す。

$$F(t+1) = \alpha S F(t) + (1-\alpha) Y \quad (11)$$

ラベルの伝播は、式 (11) が変化しなくなるまで反復した後、終了する。論文 [15] では、 $F(t)$  が最終的に式 (12) に収束することを証明している。

$$\lim_{t \rightarrow \infty} F(t) = (1-\alpha)(I - \alpha S)^{-1} Y \quad (12)$$

また、LGC 以外の手法として、Label Propagation [16], [17] が挙げられる。

グラフ構造を用いる手法に代表される深層学習を用いない半教師あり学習 [7], [12] は、深層学習を用いる手法と異なり、GPU に代表されるハードウェアを用いずに高速に学習できる。一方で、分類精度は深層学習を用いる手法に比べて低くなる。このため、近年は深層学習を用いる手法が数多く研究されている。

## 4. Random Erasing

Random Erasing は、一部の画像を消去することでデータを拡張し、画像認識の精度を向上させる手法である。画像中でランダムに消去する領域の大きさを選択し、選択した領域を規則に従い消去する。消去する領域における画素値の設定方法として、ランダム、0、画素の平均値、および画素が取りうる値の最大値の 4 種類が挙げられる。

MNIST に Random Erasing を適用した例を図 3 に示す。図 3 では、最上段が元の画像を示しており、上から順番に、消去する領域内の画素値をランダムで設定した場合、0 とした場合、画素が取りうる値の平均値 0.5 とした場合、および画素が取りうる値の最大値で設定した場合の 4 種類を示している。図 3 から、消去する領域内の値をラ

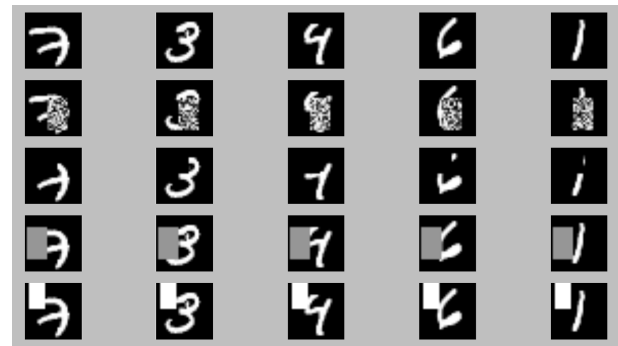


図 3 MNIST に対する Random Erasing の適用例

ンダムな値で設定した場合、一定の領域がノイズで消去される。同様に、0 で設定した場合は黒色、1 で設定した場合は白色、および平均値で設定した場合は灰色で消去される。論文 [14] では、Random Erasing において、データ拡張および正則化の両面で評価を行った。正則化に関して、Random Erasing は、Drop Out および Random Noise と比較して分類精度の向上を確認した。また、画素の設定方法として、消去する領域内の画素の値をランダムで設定する場合における画像の分類精度が一番良い。

## 5. 提案手法

### 5.1 概要

本研究では、半教師あり学習において分類精度を向上させる手法を 2 種類提案する。これらの提案手法では、誤差に対する勾配をノイズとして用いる VAT とランダムに消去する領域を選択する Random Erasing を組み合わせて、新たな半教師あり学習を定義する。

### 5.2 Virtual Adversarial Erasing

一つ目の提案手法は、深層学習を用いた半教師あり学習において、VAT を拡張させた Virtual Adversarial Erasing (以下、VAE) である。VAE では、VAT の処理手順において、画像に対してノイズを加える処理を画像の一部の領域を消去する処理に変更することで、分類精度を向上させる。

#### 5.2.1 VAE の処理手順

- VAE において、領域を消去する処理の手順を以下に示す。
- (1) 入力をもとに、クラスラベルを予測する分布 (以下、予測分布) を計算する。
  - (2) 予測分布に対する入力画像の勾配を計算する。
  - (3) 勾配の絶対値がもっとも大きい画素を中心に、画像の一部の領域を消去する。

上記の手順で領域を消去した後の画像を  $x_{erasing}$  とすると、教師なし学習における誤差は、以下の式で求められる。

$$\Delta KL(r, x, w) = KL[p(y|x, w) || p(y|x_{erasing}, w)] \quad (13)$$

VAE の概念図を図 4 に示す。VAE では、図 1 で示した  $x + r_{v-adv}$  の処理について、一部領域を消去した画像であ

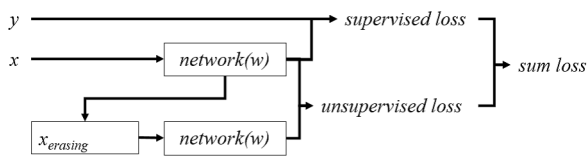


図 4 Virtual Adversarial Erasing (VAE) の概念図

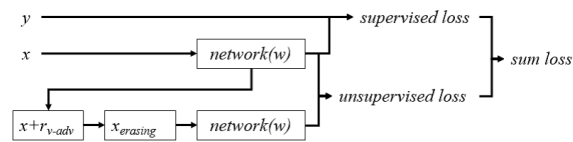


図 5 VAT + VAE の概念図



図 6 VAE による生成画像の例



図 7 VAT + VAE による生成画像の例

る  $x_{erasing}$  に置き換えている。他の処理については、VAT と同様である。

次に、既存手法である VAT と提案手法である VAE を組み合わせた手法（以下、VAT+VAE）について説明する。VAT+VAE の概念図を図 5 に示す。VAE+VAT では、図 4 と比較して、 $x_{v-adv}$  を計算した後に  $x_{erasing}$  を計算する部分を改良している。

### 5.2.2 VAE の生成画像

VAE による生成画像を図 6 に示す。図 6 では、勾配が大きい部分を消去しているため、数字の一部分が消去されている。次に、VAT+VAE による生成画像を図 7 に示す。図 7 では、勾配からのノイズに加えて画像領域を消去することで、画像の一部分が消去されるとともに、画像全体にノイズが発生している。

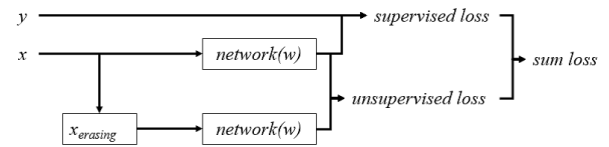


図 8 Random Erase Training (RET) の概念図



図 9 MNIST データセットの画像例

## 5.3 Random Erase Training

二つ目の提案手法は、深層学習を用いた半教師あり学習において、VAE を拡張させて Random Erasing と組み合わせた Random Erase Training（以下、RET）である。RET では、VAE を用いて消去する領域を決定する方法について、勾配を用いる代わりにランダムで行う。この場合、入力画像に対する勾配を計算する必要がなく、VAE に比べて計算速度が速くなる。

図 8 に、RET の概念図を示す。図 8 では、図 4 の場合と異なり、 $x_{erasing}$  の生成時に  $network(w)$  を計算する必要がないため、入力  $x$  から直接  $x_{erasing}$  を計算できる。

## 6. 評価

### 6.1 MNIST を用いた分類精度

MNIST [13] は、0 から 9 までの数字を手書きしたデータの集合であり、データの大きさは  $28 \times 28$  ピクセル、画像は 70,000 枚である。MNIST データセットの画像例を図 9 に示す。クラスラベルは 0 から 9 の 10 種類であり、このうち一つの数字を画像としてデータセットに格納する。評価では、50,000 枚の画像を用いて、ラベルデータは 100 枚、未ラベルデータは 49,900 枚とする。また、分類精度の評価を行うためのテストデータとして 10,000 枚を用いる。

評価に用いる手法について、提案手法では VAE, RET, および VAT+VAE の 3 種類を用いる。VAE では、消去する領域における値の設定方法について、4 章で説明した 4 種類の設定それぞれについて評価する。また、既存手法では、VAT, および Convolutional Neural Network (CNN)[4] の 2 種類を用いる。CNN は、本評価では畳みこみを用いた単純な教師あり学習として用いる。CNN ではラベルデータのみを学習に使用するため、半教師あり学習と比較して分類精度が低下する。

表 1 MNIST を用いた分類精度の評価結果

手法	精度	学習方法
VAE (ランダム)	98.80%	半教師あり
VAE (0)	98.66%	半教師あり
VAE (1)	96.83%	半教師あり
VAE (平均)	98.46%	半教師あり
RET (ランダム)	98.85%	半教師あり
VAT	98.99%	半教師あり
VAT+VAE	98.97%	半教師あり
CNN	83.70%	教師あり

MNIST を用いた場合における分類精度について、評価結果を表 1 に示す。表 1 より、半教師あり学習を用いた提案手法は、教師あり学習による既存手法と比較して、精度が大幅に向上する。また、提案手法における画素の埋め方について、VAE および RET を用いてランダムに画素を埋める場合の分類精度が他に比べて高い。提案手法でランダムに画素を埋める場合、消去する値に規則性がないため、消去する領域を一定の値で設定する他の場合に比べて分類精度は高くなる。以上より、教師あり学習だけでなく、半教師あり学習においても同様に、ランダムに画素を埋める場合の分類精度がもっとも高いことが分かる。

また、VAE, RET, および VAT+VAE の 3 種類の提案手法と既存手法 VAT と比較した場合、提案手法の分類精度は VAT に比べてわずかに低い。MNIST において、画像の全体にノイズを加える VAT における分類の効率が他の手法と比べて高いと考えられる。しかし、提案手法と既存手法の差は小さく、ほぼ同精度で分類できる。

さらに、提案手法において、RET の分類精度は VAE (ランダム) と比較して約 0.05% 高い。RET では、入力画像に対して勾配を用いる代わりにランダムで消去する領域を決定することで、分類精度が向上する。以上より、MNIST を用いた場合、VAE と比較して RET による半教師あり学習手法の分類精度が高くなる。

## 6.2 Fashion MNIST を用いた分類精度

### 6.2.1 概要

FMNIST[3] は、服や靴といったファッションアイテムに関するデータ集合であり、データの大きさは 28×28 ピクセル、画像は 70,000 枚である。FMNIST データセットの画像例を図 10 に示す。FMNIST における分類精度の計算で用いるクラスラベルは、T-shirt/top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, および Ankle boot の 10 種類である。また、FMNIST による画像の分類は MNIST と比べて難しく、分類精度は低くなる。

### 6.2.2 教師あり学習における評価

教師あり学習に半教師あり学習の誤差を追加することで、教師あり学習に比べて分類精度が向上することが考えられる。そこで、提案手法を教師あり学習に適用した場合



図 10 Fashion MNIST (FMNIST) データセットの画像例

表 2 FMNIST を用いた分類精度の評価結果 (教師あり)

手法	精度	学習方法
VAE	93.12%	半教師あり
RET	93.24%	半教師あり
VAT	92.96%	半教師あり
CNN	92.64%	教師あり

表 3 FMNIST を用いた分類精度の評価結果 (半教師あり)

手法	精度	学習方法
VAE	84.07%	半教師あり
RET	84.35%	半教師あり
VAT	83.96%	半教師あり
CNN	82.95%	教師あり

の分類精度について、FMNIST を用いた評価結果を表 2 に示す。

表 2 より、半教師あり学習で用いる VAE, RET, および VAT の分類精度は、教師あり学習の既存手法である CNN に比べて高い。これは、半教師あり学習を行うことによって分類精度が向上したと考えられる。また、半教師あり学習における提案手法である RET の分類精度は、VAE および VAT に比べて高い。FMNIST を用いる場合、画像の一部の領域を消去する手法の方がノイズを加える手法に比べて分類の精度が高いためである。

### 6.2.3 半教師あり学習における評価

FMNIST を用いた半教師あり学習の評価結果を表 3 に示す。ラベルデータは 1,000 枚、未ラベルデータは 49,000 枚とする。

表 3 より、半教師あり学習手法である VAE, RET, および VAT は、教師あり学習手法である CNN に比べて分類精度が向上する。また、半教師あり学習手法では、RET における分類精度がもっとも高い。教師あり学習における評価と同様に、半教師あり学習においても、画像の一部の領域を消去する手法の方がノイズを加える手法に比べて分類の効率高い。このため、ノイズを加えることで、微細な色および形状の変化における認識は難しくなると考えられる。

## 7. まとめ

本研究では、教師あり学習におけるデータの分類手法である Random Erasing を半教師あり学習に拡張した手法として、勾配を用いて消去する画像の領域を決定する Virtual Adversarial Erasing (VAE), およびランダムに画像の消去を行う Random Erase Training (RET) の 2 種類を提案した。提案手法では、画像の一部領域を消去した上で半教師あり学習を行う。評価の結果、これらの提案手法はどちらも半教師あり学習として画像データを分類し、既存手法である Virtual Adversarial Training (VAT) とほぼ同様の分類精度となることが分かった。

今後は、より実用的なデータセットを用いた評価を行う予定である。

## 参考文献

- [1] Rasmus, A., Valpola, H., Honkala, M., Berglund, M. and Raiko, T.: Semi-supervised Learning with Ladder Networks, Proceeding of Advances in Neural Information Processing Systems (NIPS'15), Vol. 2, pp. 3546-3554 (2015).
- [2] Goodfellow, I. J., Shlens, J. and Szegedy, C.: Explaining and Harnessing Adversarial Examples, Proceeding of International Conference on Learning Representations (ICLR 2015), arXiv (online), available from < <https://arxiv.org/pdf/1412.6572.pdf> > (accessed 2017-12-12).
- [3] Han, X., Kashif, R. and Roland, V.: Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms, arXiv (online), available from < <https://arxiv.org/pdf/1708.07747.pdf> > (accessed 2017-12-12).
- [4] He, K., Zhang, X., Ren, S. and Sun, J.: Deep Residual Learning for Image Recognition, arXiv (online), available from < <https://arxiv.org/pdf/1512.03385.pdf> > (accessed 2017-12-12).
- [5] Hoffer, E. and Aillon, N.: Deep Metric Learning using Triplet Network, Proceeding of International Conference on Learning Representations (ICLR 2015), arXiv (online), available from < <https://arxiv.org/pdf/1412.6622.pdf> > (accessed 2017-12-12).
- [6] Hoffer, E. and Ailon, N.: Semi-supervised Deep Learning by Metric Embedding, arXiv (online), available from < <https://arxiv.org/pdf/1611.01449.pdf> > (accessed 2017-12-12).
- [7] Joachims, T.: Transductive Inference for Text Classification using Support Vector Machines, Proceedings of 16th International Conference on Machine Learning, pp. 200-209 (1999).
- [8] Kingma, D., Mohamed, S., Rezende, D. J. and Welling, M.: Semi-supervised Learning with Deep Generative Models, Proceeding of Advances in Neural Information Processing Systems, pp. 3581-3589 (2014).
- [9] Miyato, T., Maeda, S., Koyama, M., Nakae, K. and Ishii, S.: Distributional Smoothing with Virtual Adversarial Examples, arXiv (online), available from < <https://arxiv.org/pdf/1507.00677.pdf> > (accessed 2017-12-12).
- [10] Miyato, T., Maeda, S., Koyama, M., Nakae, K. and Ishii, S.: Virtual Adversarial Training: a Regularization Method for Supervised and Semi-supervised Learning, arXiv (online), available from < <https://arxiv.org/pdf/1704.03976.pdf> > (accessed 2017-12-12).
- [11] Park, S., Park, J. K., Shin, S. J. and Moon, C.: Adversarial Dropout for Supervised and Semi-supervised Learning, arXiv (online), available from < <https://arxiv.org/pdf/1707.03631.pdf> > (accessed 2017-12-12).
- [12] Sousa, C., Rezende, S. and Batista, G.: Influence of Graph Construction on Semi-supervised Learning, Proceeding of European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Vol. 8190, pp. 160-175 (2013).
- [13] The MNIST Database of handwritten digits (online), available from < <http://yann.lecun.com/exdb/mnist/> > (accessed 2017-12-12).
- [14] Zhong, Z., Zheng, L., Kang, G., Li, S. and Yang, Y.: Random Erasing Data Augmentation, arXiv (online), available from < <https://arxiv.org/pdf/1708.04896.pdf> > (accessed 2017-12-12).
- [15] Zhou, D., Bousquet, O., Navin, T., Weston, J. and Scholkopf, B.: Learning with Local and Global Consistency, Proceeding of Advances in Neural Information Processing Systems, Vol. 16, pp. 321-328 (2004).
- [16] Zhu, X., and Ghahramani, Z.: Learning from Labeled and Unlabeled Data with Label Propagation, Carnegie Mellon University CALD Technical Reports, CMU-CALD-02-107, pp. 1-17 (2002).
- [17] Zhu, X. and Ghahramani, Z.: Semi-supervised Learning using Gaussian Fields and Harmonic Functions, Proceedings of 21th International Conference on Machine Learning (ICML'03), pp.912-919 (2003).