

新しい深層学習手法 — 非調和学習 —

2022年3月1日

関西学院大学 工学部 知能・機械工学課程
教授 岡留 剛

技術のポイント

深層学習のミニバッチ学習を前提としたデータ拡張において、
元画像から複数の拡張を行ない、

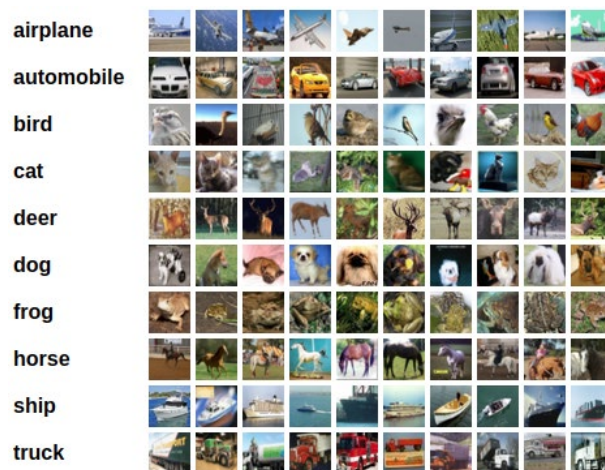
- 学習途中の分類器でそれら複数の拡張画像の誤差を見積もり、
- 誤差に応じた重みづけでサンプリングした拡張画像を1枚選択、
- ミニバッチの構成要素としてパラメータを更新。

技術の背景：深層学習

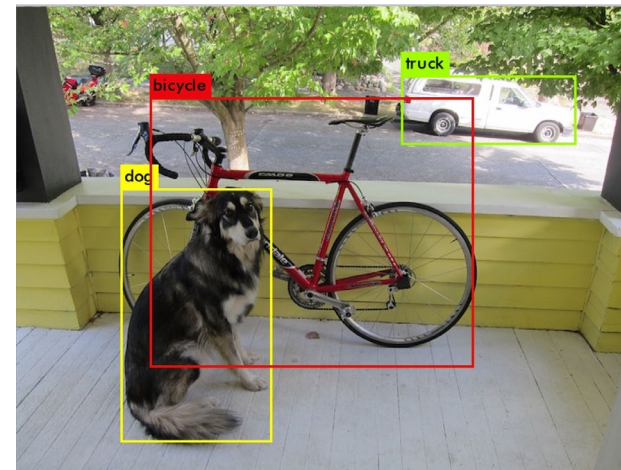
深層学習は多様なタスクで優れた性能を発揮
特に Computer Visionのタスクでその貢献は絶大

● 主要なタスクとモデル

- 画像分類: VGG, ResNet ...
- 物体検出: R-CNN, YOLO
- セマンティックセグメンテーション: U-Net



引用元: <https://www.cs.toronto.edu/~kriz/cifar.html>

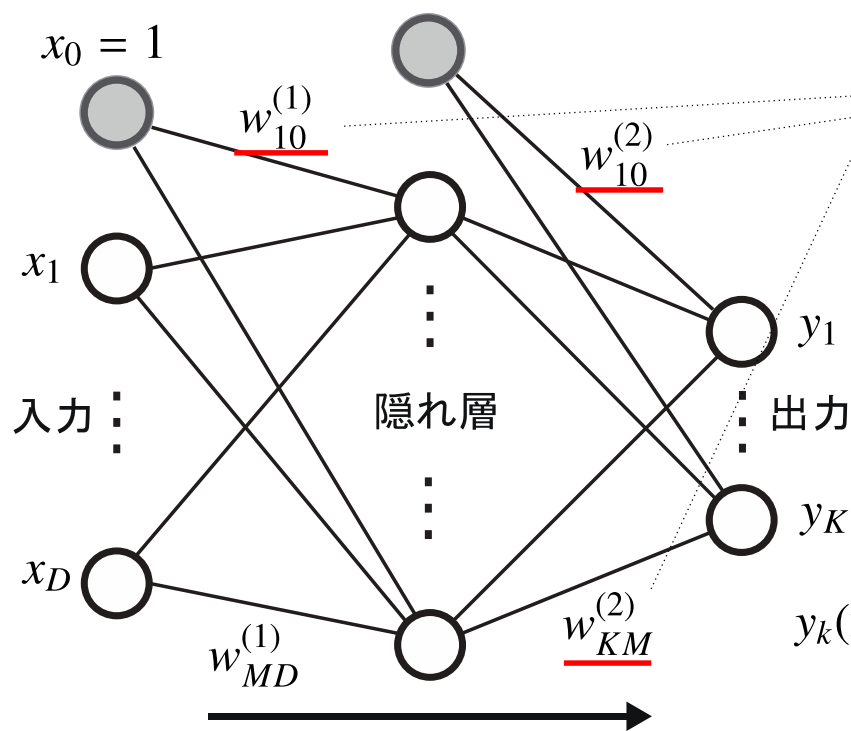


引用元: <https://pjreddie.com/darknet/yolo/>

深層学習の概要

深層学習の基礎：3層パーセプトロン

どんな計算機にもなる仕組み。



w_{ij} : 重み

重みを変えると別の計算機となる。

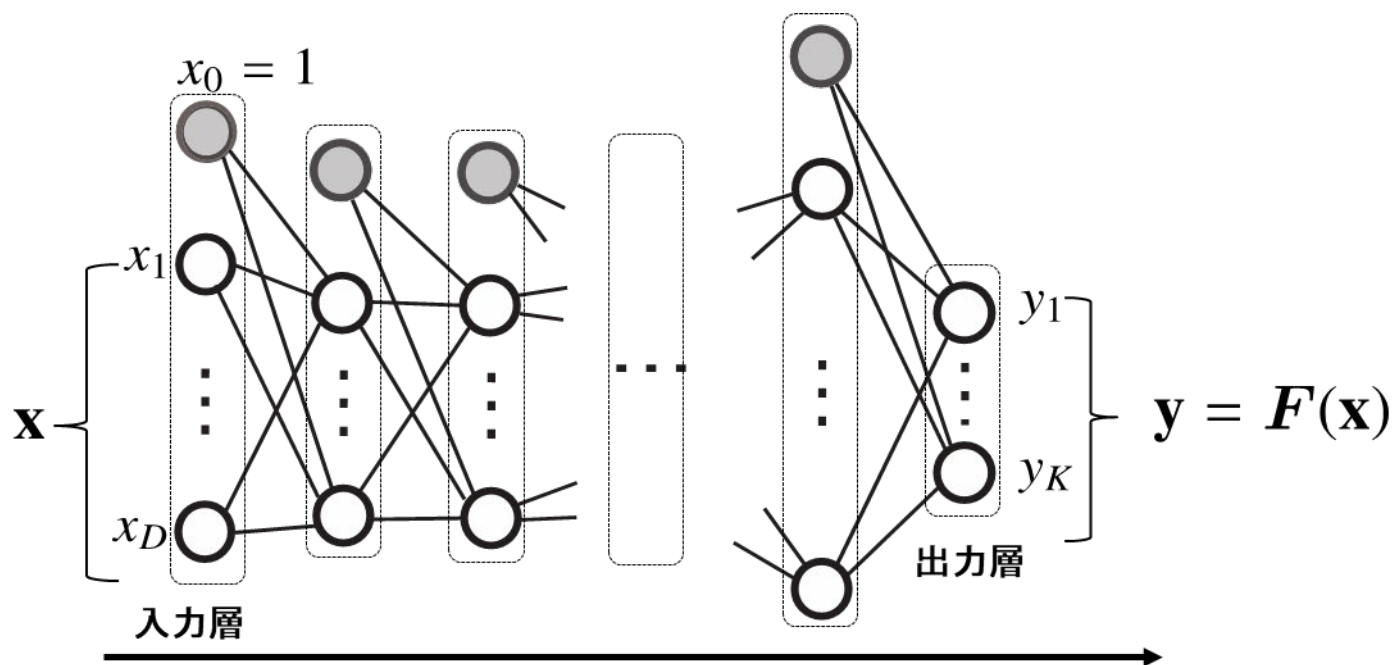
$$y_k(\mathbf{x}) = f_2 \left(\sum_{j=1}^M w_{kj}^{(2)} f_1 \left(\sum_{i=1}^D w_{ji}^{(1)} x_i + w_{j0}^{(1)} \right) + w_{k0}^{(2)} \right)$$

\mathbf{w} : すべての重みを集めたベクトル

深層学習モデル(DNN)

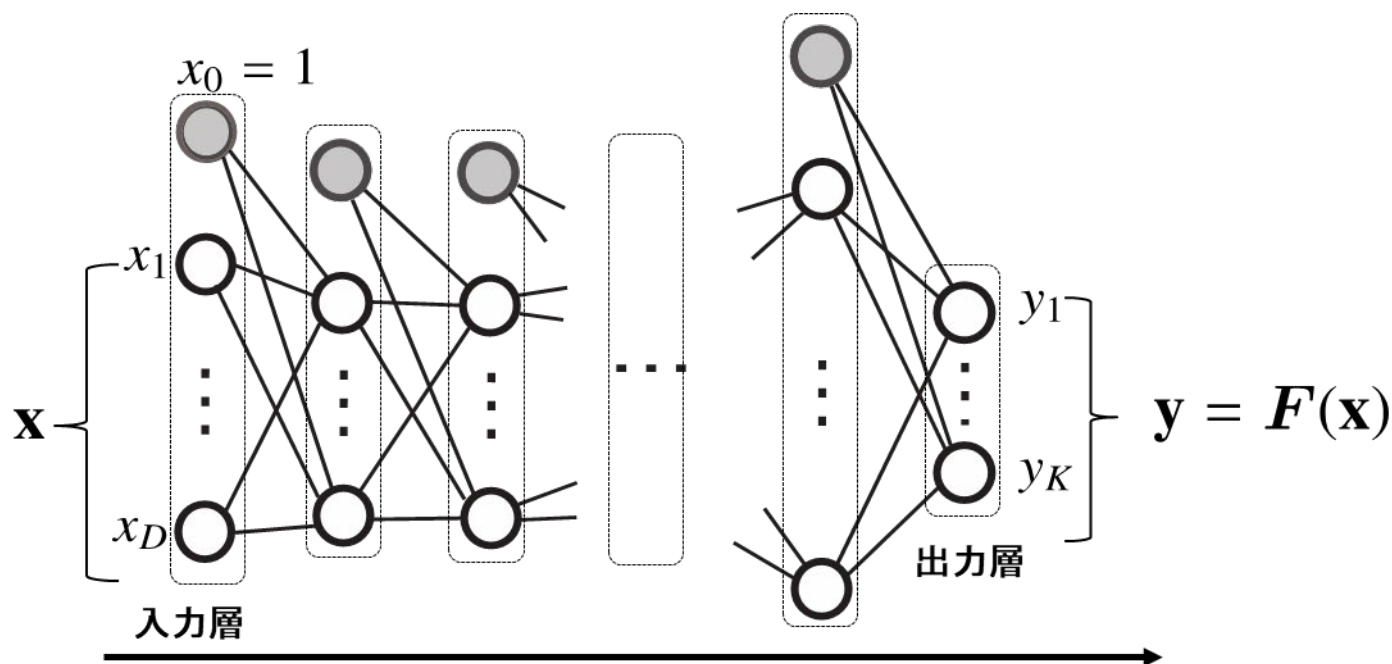
多層のパーセプトロン: より記述力の高いモデル.

- 特定の関数 F を計算する.
 - ベクトル \mathbf{x} を入力層が受けとり,
 - ベクトル $\mathbf{y} = F(\mathbf{x})$ を出力層が出力.
 - 中間の層は, 前の層からリンクをつたわってきた情報をもとに単純な計算をおこない, その結果をリンクをつたって次の層にわたす.



学習：深層学習モデルの訓練概要

- あたえられたデータ $\mathcal{D} = \{(\mathbf{x}_1, \mathbf{t}_1), \dots, (\mathbf{x}_N, \mathbf{t}_N)\}$ に対し、重み \mathbf{w} を決定することが学習
 - ラベル: \mathbf{t}_n は、入力 \mathbf{x}_n に対する目的変数値
 - ✓ 回帰のとき: K 次元実ベクトル
 - ✓ K 個の2クラス分類: 成分を0または1とする K 次元ベクトル
- データに対する誤差の最小化により \mathbf{w} を決定.

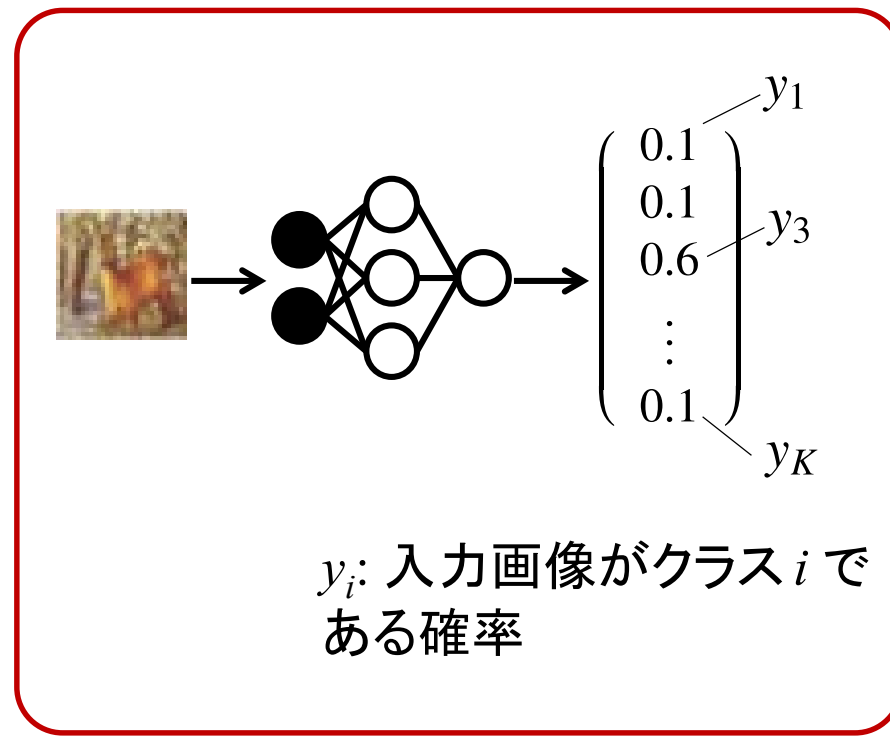
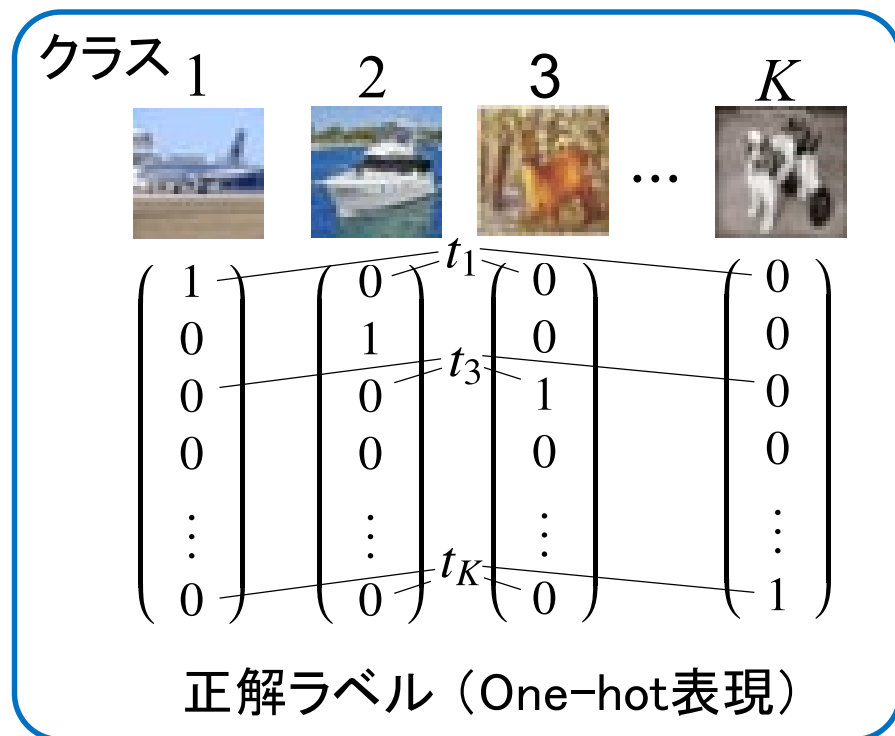


誤差: 深層学習モデルの訓練

クラス確率を出力する深層学習では、誤差関数として、通常、交差エントロピー誤差関数が用いられる。

交差エントロピー誤差損失: $l = - \sum_{k=1}^K t_k \ln y_k$

$(y_1, \dots, y_k)^T$: DNNの出力
 $(t_1, \dots, t_k)^T$: 正解ラベル



ミニバッチ学習：深層学習モデルの訓練

通常，深層学習 (DNN) では，

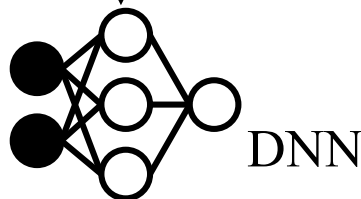
- 全データを，ミニバッチと呼ばれる単位にデータを分割し，
- 1つのミニバッチに対し，損失を小さくするようにパラメータの更新を行ない，
- それをすべてのミニバッチについて行なう。(さらに繰り返す)



<https://www.cs.toronto.edu/~kriz/cifar.html>

利点

- 過学習の抑制
- 膨大なデータに対して学習が可能

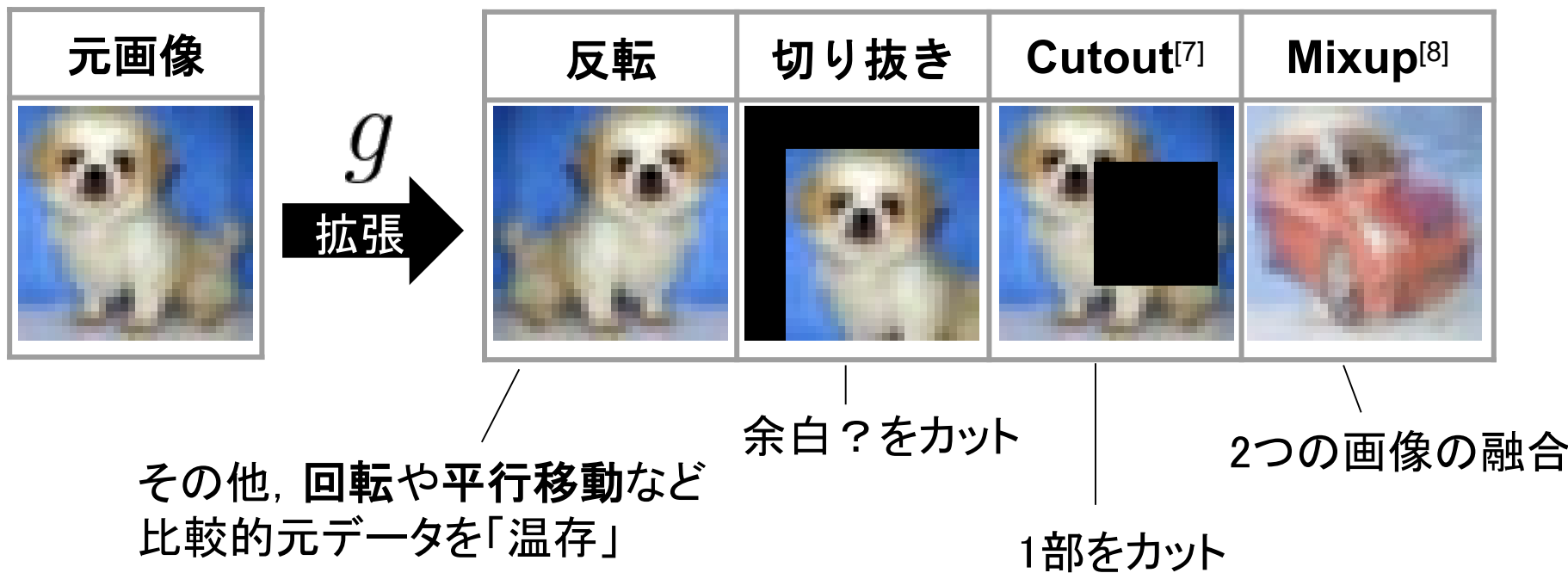


データ拡張：深層学習モデルの訓練

一般に、データを水増しして、データの多様性を促進させることで汎化性能(新しいデータに対する分類性能)が向上する。

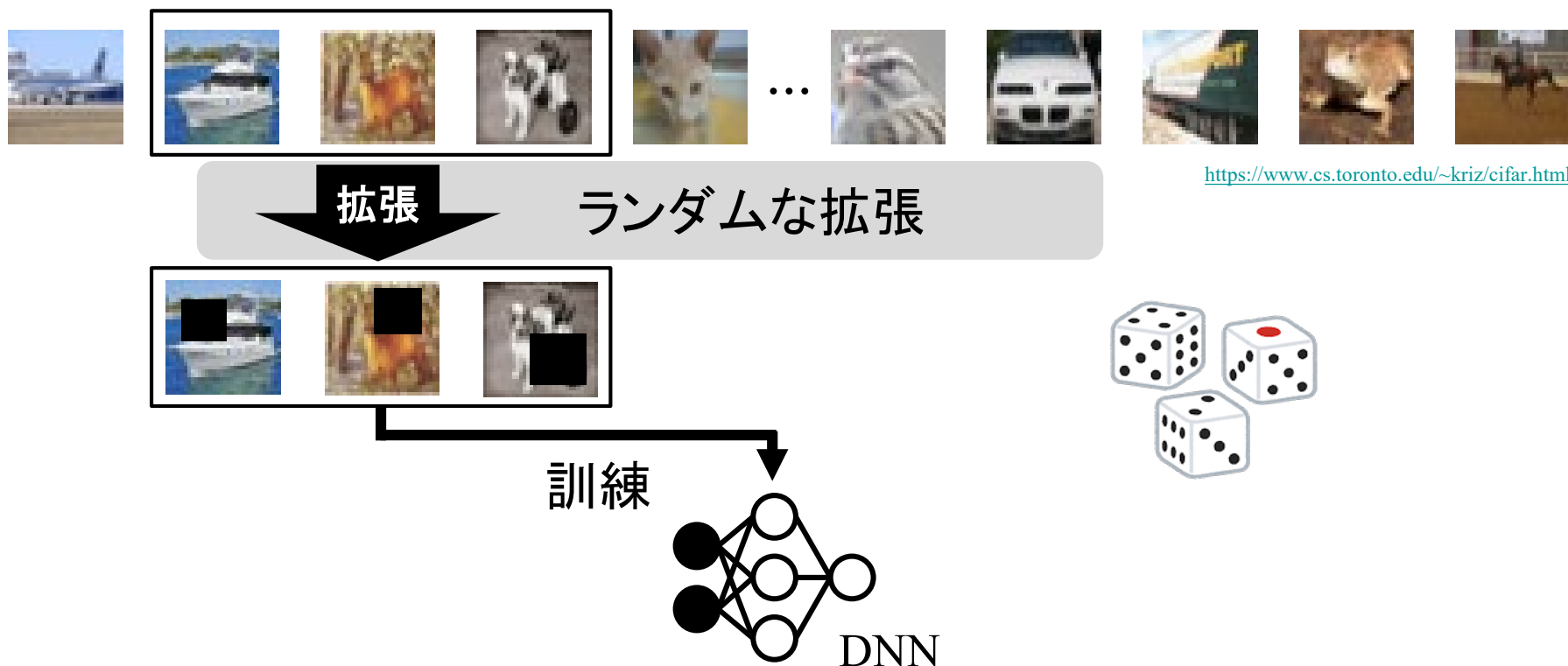
代表的なデータ拡張

<https://www.cs.toronto.edu/~kriz/cifar.html>



ミニバッチ学習におけるデータ拡張

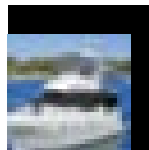
- どのデータ拡張手法でも
- それぞれの元データを「ランダム」に拡張
 - ランダムな拡張: 例えば、Cutout手法においては、元画像から切り取る部分の大きさと位置をランダムに選択



新技術の特徴・従来技術との比較

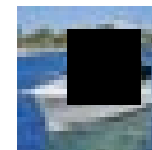
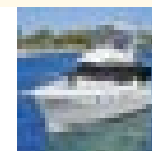
従来技術

ランダムに変換データを選択

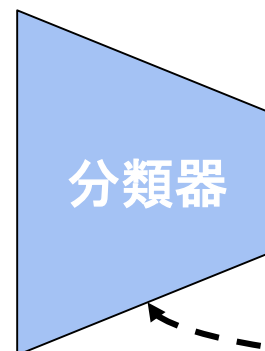


提案技術

分類器にとって
不確実性の高い
変換データを選択



分類器の性能向上

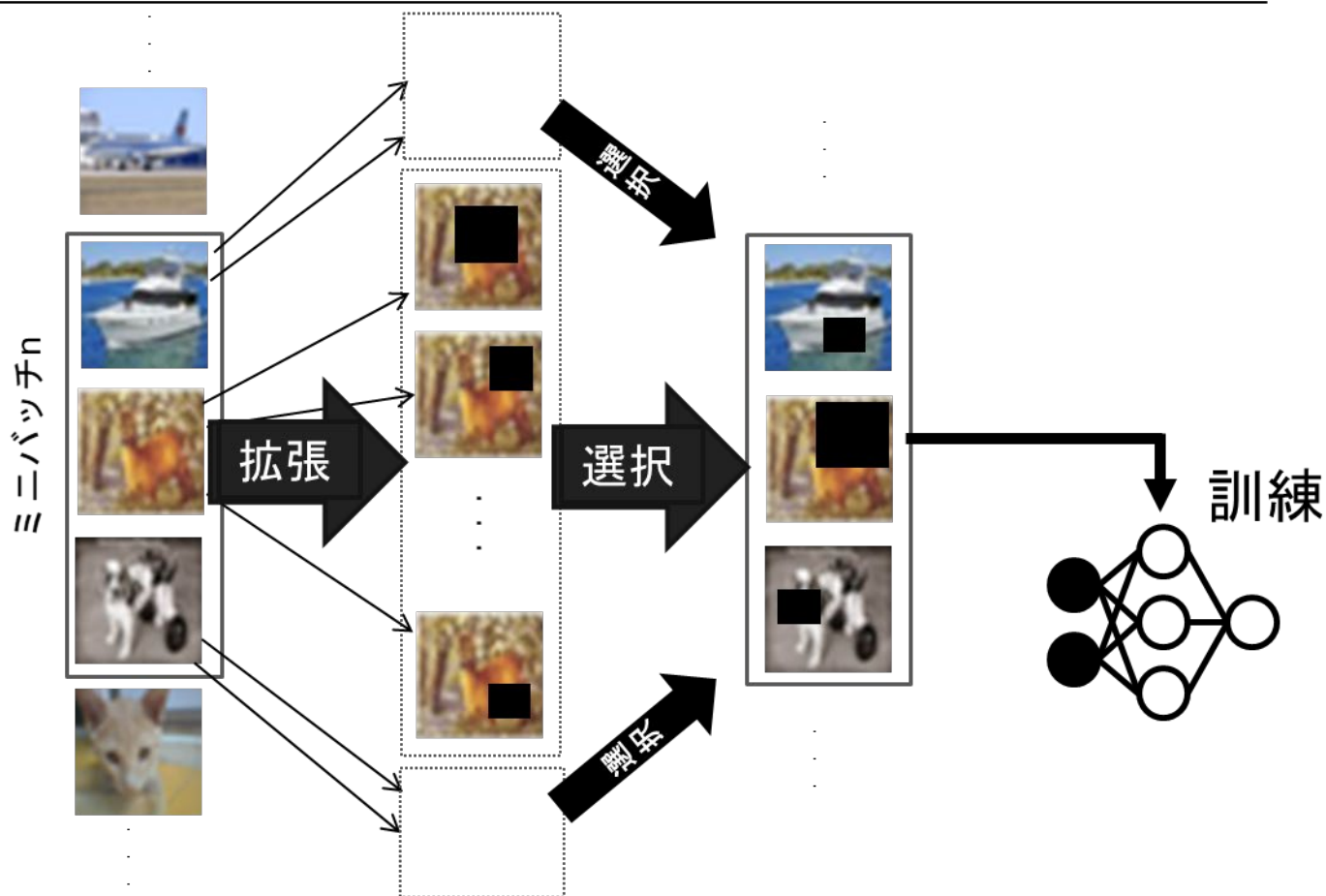


不確実性大

提案技術: 基本手順

ミニバッチ中の元画像それぞれにつき,

1. 元画像1枚をM枚に拡張(拡張手法は問わない),
2. 拡張されたM枚のうち1枚を選び, 元画像に代えてミニバッチへ.

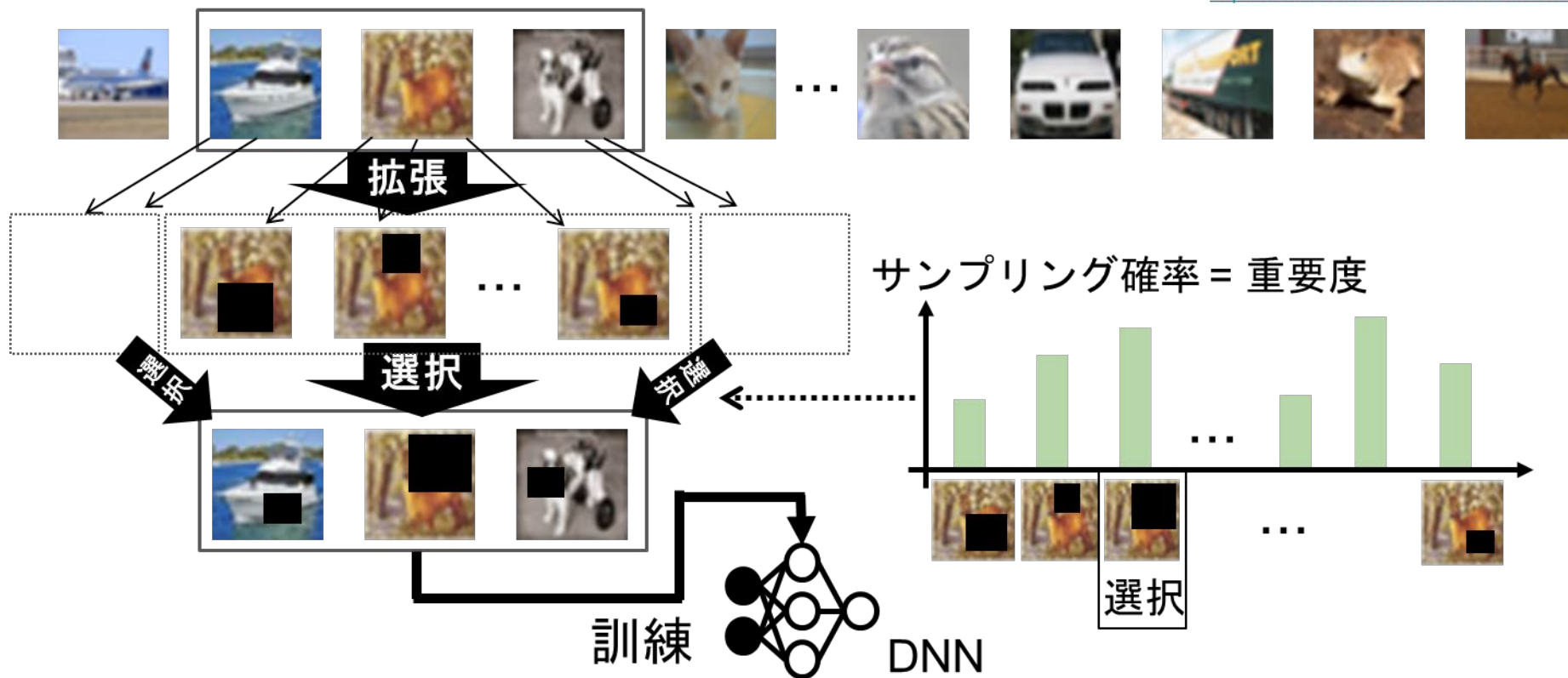


提案技術: 概要

ミニバッチ中の元画像それぞれにつき,

1. 元画像1枚をM枚に**拡張**(拡張手法は問わない),
2. 拡張されたM枚のそれぞれの**重要度**を計算,
3. 重要度を加味したサンプリングで1枚を選択(**重要度サンプリング**),
4. 元画像にかえてサンプリングした1枚をミニバッチへ.

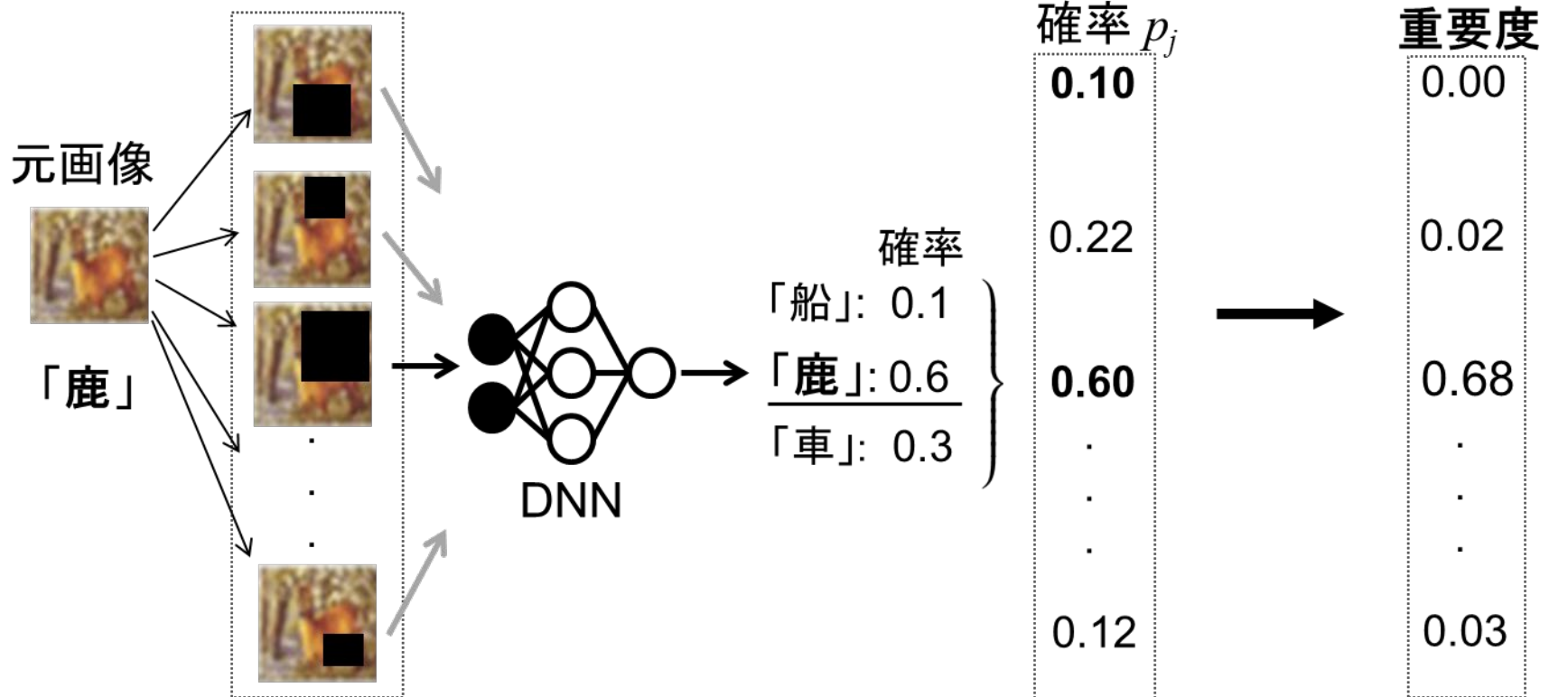
<https://www.cs.toronto.edu/~kriz/cifar.html>



提案技術詳細: 重要度

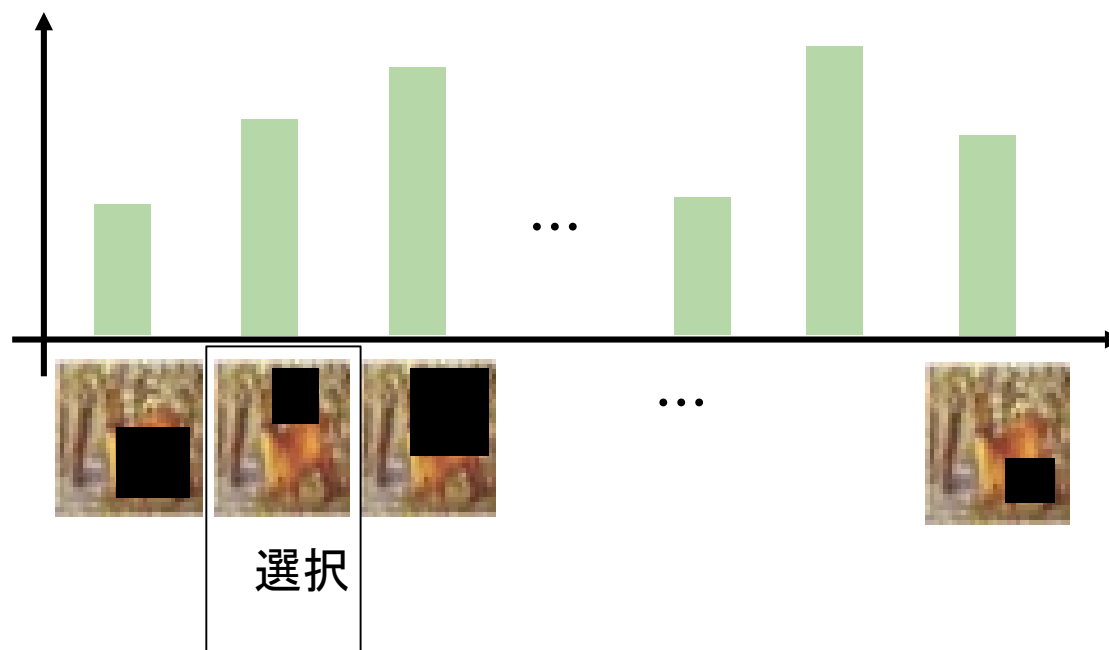
- 学習途中の分類器(DNN)に入力したときの損失 p_j を利用.
- 拡張されたM枚の画像の確率値(鹿の画像なら鹿の確率)

- **重要度:** $w_j = \frac{p_j}{\sum_{n=1}^M p_n}$ (正規化確率)



提案技術詳細: 重要度サンプリング

サンプリング確率 = 重要度



実験: タスク

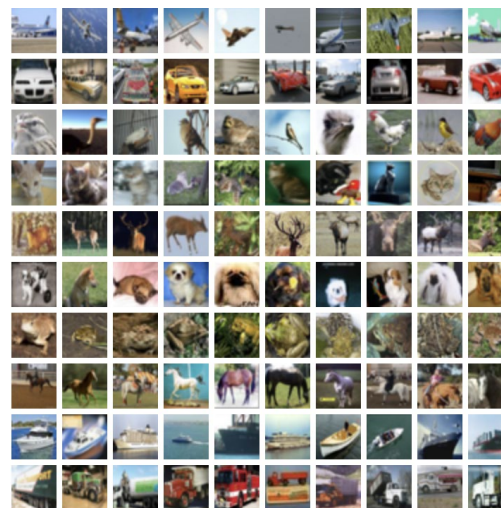
データセット

CIFAR-10

CIFAR-100

モデル: PreActResNet-18^[6]

評価指標: 誤分類率

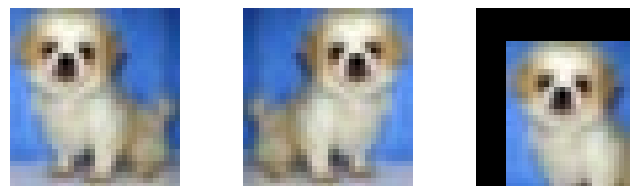


	CIFAR-10	CIFAR-100
データ数	訓練: 50,000枚 テスト: 10,000枚	
クラス数	10	100
画像サイズ	32x32x3(RGB画像)	

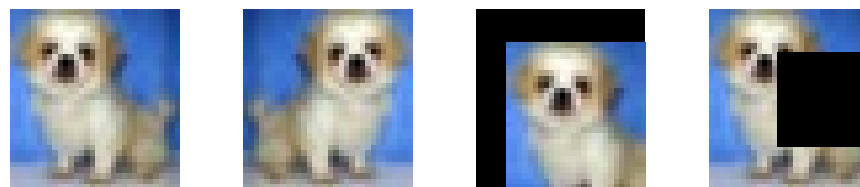
実験: 設定

データ拡張手法: 元画像 +

■ 設定A: 反転 + 切り抜き



■ 設定B: 設定A + Cutout



実験：結果（誤分類率）

- CIFAR-10

	従来法	提案法
設定A	5.37%	4.75%
設定B	3.86%	3.81%

- CIFAR-100

	従来法	提案法
設定A	24.3%	23.4%
設定B	23.6%	22.6%

まとめ

深層学習のミニバッチ学習を前提としたデータ拡張において、元画像から複数の拡張を行い、損失に応じた重みづけでサンプリングした拡張画像を1枚選択し、DNNを訓練する手法。

画像認識では、従来法と比べて高い精度がでる。

データ増強を行なって、深層学習モデルを訓練しているほとんどのタスクに適用可能。

ただし、学習時間は既存手法の約1.8倍

深層学習の応用

- 自動翻訳
- 医療における画像診断や健康診断
- 自動運転
- 画像からの人物同定

実用化に向けた課題

- 本技術の有用性について、一般的な画像認識の学習では検証済であるが、個別の案件での検証が必要。
- 本技術の適用可能分野として画像認識以外では、例えば、画像中の物体検知・人物同定分野や、音声認識や音楽生成の分野などがありうる。
- 学習ドメインごとに、適切なデータ拡張のしかたや、拡張の個数など、細かなチューニングが必要となる。

企業への期待

- 比較的データが少ない特定のドメインで、機械学習による分類器や回帰モデルを構築する際のデータ拡張
- 転移学習の枠組みで、分類器や回帰モデルを構築する際の転移先ドメインの手持ちのデータ拡張などに利用できると思われる。

本技術は、現行の形で、比較的簡単に、導入することができるので、試験的に実験してもらいたい。

その際に、技術アドバイス等、本学との連携が必要であれば検討したい。

本技術に関する知的財産権

- 発明の名称 : 情報劣化処理型水増しを用いた学習方法及び学習装置
- 出願番号 : 特願2019-209179
- 出願人 : 学校法人関西学院
- 発明者 : 岡留 剛, 井手 敦也

お問い合わせ先

関西学院大学

研究推進社会連携機構 知財産学連携センター

TEL : 079-565-9052

FAX : 079-565-7910

e-mail : industry-academia@kwansei.ac.jp