

知能情報実験 III（データマイニング班）
指紋認証の分析と改善

185710A, 185714C, 185745C, 185752F, 185763B

提出日：2021 年 2 月 x 日

目次

1	はじめに	2
1.1	Convolutional Neural Network:CNN	2
1.2	テーマ指紋認証とは	3
2	実験方法	3
2.1	実験目的	3
2.2	データセット構築	3
2.3	モデル選定	4
2.4	パラメータ調整	4
3	実験結果	4
4	考察	5
5	意図していた実験計画との違い	5
6	まとめ	5

1 はじめに

データマイニングとはデータの中に埋め込まれている有用な知識を発掘することである。別の言い方では、データマイニングは、より良い意思決定をするために履歴データをうまく使って一般的な規則性を発見しようとする研究分野である。今回私たちのグループ3では、機会学習の基本的な考え方を実装、体験を通して学んだ。そしてその応用として、既存に存在する指紋データを使った指紋認証のプログラムの分析から、結果の改善と精度向上を目指し、その結果を可視化し、考察したことについて考察する。

1.1 Convolutional Neural Network:CNN

Convolutional Neural Network（これより CNN と呼ぶ）は畳み込みニューラルネットワークという意味であり、機械学習で画像の深層学習といえば CNN であるというほどよく使われている識別手法である。これは、ニューラルネットワークに畳み込みという操作を導入したものである。CNN について、簡単な手順を記述する。まず手順1として、画像から特徴を抽出する。フィルタを使い、入力層データの中で位置を変えながらスキャンした部分のデータと、フィルタ自身の持つデータとの差異を畳み込みの結果として畳み込みそうに書き込んだものを特徴量といい、入力層の全データをスキャンしてできた畳み込み結果の値の集まりを特徴マップという。複数のフィルタを用意することで、入力層のデータ特徴を捉えやすくしている。下の図は用意した複数フィルタのうち、一つが完全に入力層のデータの一部と同じであることを示す。

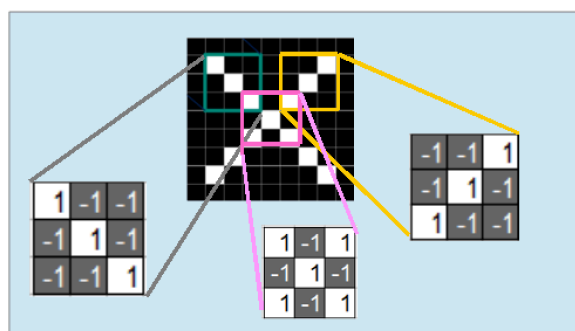


図1 CNN 解説手順1

手順2として、画像を畳み込みする。入力層のデータをフィルタのデータとピクセル毎に比較することで、畳み込み層にその類似度（特徴量）を書き込む。下記の図はフィルタを利用して特徴量を抽出し、特徴マップを作成した例である。

手順3として、画像をプーリングする。畳み込みの層の情報はプーリング層で集約する。出力に関しては、プーリング層のユニット全てと全結合し、計算結果を利用して、フィルタ、重み、バイアスを更新していく。

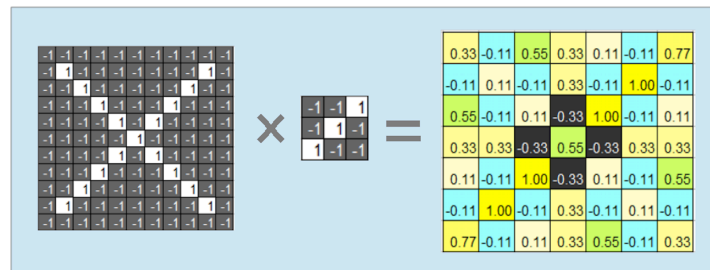


図 2 CNN 解説手順 2

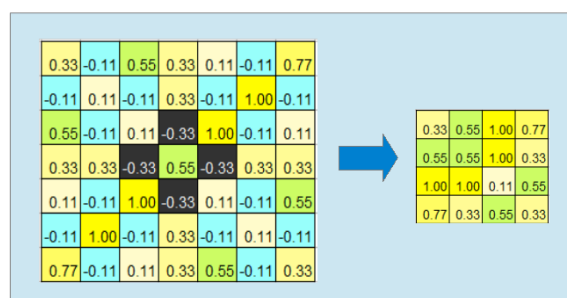


図 3 CNN 解説手順 3

1.2 テーマ指紋認証とは

本グループでは、授業の中でデータマイニングについて学び、それらの応用実験として、画像認識について分析しようと考えた。そして、データマイニングを行う識別手として CNN に目をつけ、データの分類ラベルがはっきりしており、複雑である指紋認証の分析、精度改善を行うこととなった。CNN は畳み込み処理を利用したニューラルネットワークであり、どのくらい畳み込み処理を行うのか、どのくらいニューラルネットワークを深くするのは定義されていない。

2 実験方法

2.1 実験目的

実験を通して明らかにしたいこと、確認したいこと、検証したいことを述べよう。

2.2 データセット構築

Kaggle より、指紋のデータセットを利用している。<https://www.kaggle.com/ruizgara/socofing>

2.3 モデル選定

アルゴリズムは序章で述べたとおり、CNN を利用しており、以下のコードを参考にしている。
<https://www.kaggle.com/brianzz/subjectid-finger-cnnrecognizer>
本アルゴリズムの運用において、指紋認証の正答率が十分に高い点、可読性に優れており改善手法を模索しやすい点に特に優れているため使用することにした。

2.4 パラメータ調整

CNN のパラメータ調整では、epoch 数・画像データ数・batchsize や活性化関数である LeakyReLU を変更していった。

3 実験結果

事実として得られた結果を示そう。なお、以下の点に留意すること。

- 「思う」「思われる」のような主観ではなく、客観的事実を述べること。
- 図表には適切なキャプションを付けること。
- 挿入した図表について、本文中でその読み方を述べること。その際には label, ref により相互参照すること。
- レポートにおけるグラフの作成においては、以下の点に注意する。
 - － 軸目盛および軸ラベルに関する注意事項
 - * 必ず軸ラベルを表示する
 - * 軸に単位がある場合には、ラベルに単位を付記する
 - * 軸目盛は適切な感覚で表示する
 - * 軸目盛は述べたい内容に応じて線形スケールと log スケールを使い分ける
 - * 印刷時に明瞭に読むことができるサイズで表示する
 - － 線・点・ポイントおよび凡例に関する注意事項
 - * 線・点・ポイントは、印刷時に明瞭に識別できる太さやサイズで表示する
 - * 1 つのグラフに複数にデータを表示する際には、データごとに異なる線種、線の太さ、ポイント形状などを使用する
 - * モノクロ印刷でも識別できるように線・点・ポイントを使用することが望ましい
 - * 凡例は線・点・ポイントに重ならないように注意する

4 考察

実験課題への取り組みを通し、実験の意義、実験からわかったこと、今後の展望などを述べる。失敗やつまづきがあれば、それらについての失敗分析を含めると良い。

5 意図していた実験計画との違い

グループワークとして2ヶ月程度の時間が用意されていた。ガントチャート??等、何かしら工夫して全体の計画を述べよう。これらの期間をどのように使おうとし、実際どうだったのかについて自己評価（振り返り）してみよう。大きなズレがある場合それは何故起きたのか、どうやればそのギャップを縮められそうか検討してみよう。

6 まとめ

データマイニング班の達成目標を振り返り、選んだテーマに対する機械学習の適用を通して得られた知見や学んだことをまとめよう。また今後やるべきことや後進に伝えたいこと等あれば自由に述べよう。

参考文献

- [1] CNN 解説,
<https://udemy.benesse.co.jp/data-science/ai/convolution-neural-network.html>, 2021/01/026.
- [2] 指紋データセット, <https://www.kaggle.com/ruizgara/socofing>
- [3] アルゴリズム, <https://www.kaggle.com/ruizgara/socofing>
- [4]