

Deep-Learning による手書き数字画像認識
Handwritten digit character recognition by deep-learning

阿比留大耀¹, 椎名啓司¹
Abiru Hiroaki, Keiji Shiina

Abstract: Deep-learning is an information processing technology that imitates the network of neurons in the human brain. Deep-learning makes it possible for computers to classify feature patterns contained in images, sounds, etc. by constructing multilayered neural networks.

1. はじめに

近年, 人工知能は世界中の注目を集めている. そして, その中核となる技術は, 人工知能の革命とも言われているディープラーニング [1]である. その一例として, 小売業や外食店舗内にいる消費者の購買行動のデータ分析にディープラーニングを応用している企業がある. 本論文では, ディープラーニングによる手書き数字画像認識に焦点をあて報告する.

2. 画像認識

画像認識とは, 画像データの画像内容を分析して, その特徴を抽出し, 対象物を識別するパターン認識技術の 1 つである [2].

人間は, 画像に写っているものが何であるか経験から判断することができるが, コンピューターはそのままでは画像に何が写っているかを判断することができない. しかし, 沢山の画像データから, 対象物の特徴を学習させることで, 未知の画像を与えた時に, 対象物が何であるかを確率として表現することができるようになる.

3. パターン認識

パターン認識は, 認識対象がいくつかの概念に分類できる時, 観測されたパターンをそれらの概念のうちの一つに対応させる処理である [3]. この概念をクラスあるいは類と呼ぶ. 手書き数字画像認識とは, 手書きで入力されたデータを 10 種類の数字のいずれかに対応させることである.

パターン認識を行うためには, まず, 認識対象から

何らかの特徴量を計測するための手法が必要である. その際, 一般に, 特徴量は 1 種類だけではなく, 複数の特徴量を計測し, それらを同時に用いることが多い. 文字画像認識の場合には, スキャナ等で取り込んだ画像そのものを特徴とみなすこともあるが, 文字の識別に必要な本質的な特徴のみを抽出するのが一般的である [3].

4. 機械学習

一般に, 機械学習とは与えられた情報を元に学習し, 自律的に法則やルールを見つけ出す手法やプログラムのことを指す [4].

機械学習を用いる際には, 大量の情報を元にマシンに学習させ, その学習結果は最終的にモデルとして出力される. そのようにして出来上がったモデルに分析したいデータを入力することで認識結果を最終的に得る.

さらに, 機械学習を用いることで, 大量のデータの中に共通する法則を見出すことができ, これまで人が得意としてきた作業をマシンに行わせることができるようになる [4].

5. ディープラーニング

ディープラーニング [1]とは, 人間の脳の神経細胞のネットワークを模倣した情報処理技術である [5]. ディープラーニングでは, 多層ニューラルネットワークの学習により, 画像や音声などに含まれる特徴量をコンピューター自身が発見し, 分類ルールを構築することが可能となる.

1 : 日大理工・学部・数学

Figure 1 に画像データ分類における従来手法とディープラーニングとの比較を示す。

従来では、まず特徴量を手動により設定し、その後、その特徴量を基に分類アルゴリズムにかけて分類するという手法であったが、ディープラーニングでは、手動による特徴量の設定の必要なくデータ分類が可能となったことが、機械学習技術における大きな違いである。

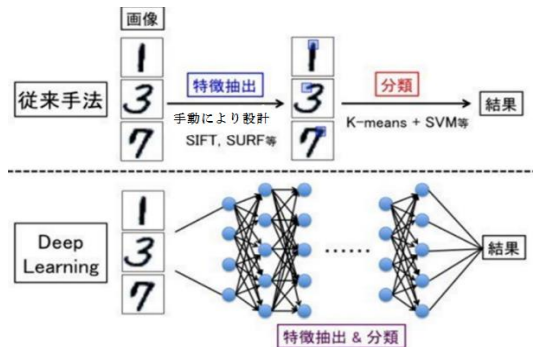


Figure 1. 画像データ分類における従来手法とディープラーニングを用いた場合との比較

一例として、今数字の手書き文字 3 を画像データとして多層ネットワークに入力したとする。Figure 2 の例に示されるように、ディープラーニングにおけるネットワーク構造では、前層における数字画像「3」の特徴抽出領域が後層において組み合わせられ、層が深くなるにつれて、その領域が広がる。

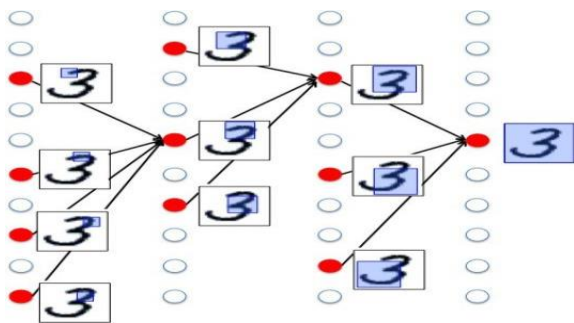


Figure 2. Deep-Learning における層ごとの特徴量抽出領域の変化

実際には、手書き数字画像データを大量に入力し学習させると、どのようなある特定の手書き数字にも共通して反応するユニットが生成される。そのような学

習を行った後、未知の手書き数字画像が入力されたとしても、ある特定の数字 (Figure 2 の例では「3」) であると認識することが可能となる。

6. Open CV

OpenCV[6]はインテルが開発したオープンソースのコンピュータビジョンライブラリであり、現在では米国のロボットの研究開発を行うベンチャー会社であるウィローガレージにて開発が続けられている。

OpenCV の主な機能として、一般的なフィルタ処理や行列演算を備え、顔認識処理にも取り組むこと等がある。

学術講演会当日では、手書き数字画像の認識実験を実際に行い、その認識結果について報告および考察する予定である。

7. 参考文献

- [1] Hinton, G. E.; Dayan, P.; Frey, B. J.; Radford, N.: "The wake-sleep algorithm for unsupervised neural networks". *Science*. 268 (5214): 1158–1161.
- [2] 大田 和樹:「事例付き様々なサービスに活用!画像認識技術とは」
<https://techacademy.jp/magazine/17068>, 2018 年 2 月.
- [3] 栗田 多喜夫:「パターン認識とニューラルネットワーク」
<https://home.hiroshima-u.ac.jp/tkurita/lecture/prnn/node1.html>, 2013 年 7 月.
- [4] JIG-SAW 株式会社:「機械学習・深層学習・AI について」
<https://clonos.jp/knowledge/detail09/>.
- [5] 木田 智士:「Deep learning で画像認識①~Deep Learning とは」
<https://lp-tech.net/articles/j4xoo>, 2016 年 11 月.
- [6] 皆川 卓也:「OpenCV で学ぶ画像認識」
<https://gihyo.jp/dev/feature/01/opencv/000>, 2008 年 5 月.