

機械学習を用いた手形状分類と手話認識への応用

Hand shape classification using machine learning and application to sign language recognition

劉 冠杰

指導教員 小川 毅彦

拓殖大学大学院 機械電子システム工学専攻 小川研究室

キーワード： 機械学習、手形状認識、手話認識

1. はじめに

手話はコミュニケーションに積極的な役割を果たす言語の一形態であり、手話を利用することで、聴覚障害者と聴覚障害者の間、および聴覚障害者と健康者の間のコミュニケーションを助ける。手話では、手や指、腕を使う手指動作に加えて、顔の部位や身振りなど非手指動作が用いられるが、五十音を表す指文字を用いることで、手話単語が無い場合や一字一句正確に伝えたいときでも、相手に伝えることができる¹⁾。

本研究では、画像による手話認識を目標とし、まずは手形状の画像から指文字を認識することを目的とする。Web カメラで取り込んだ手形状の画像から、機械学習によって指文字として判別を行うシステムを構築して検討を行う。具体的には機械学習として畳み込みニューラルネットワークを用いる。

2. 機械学習

機械学習とは、コンピュータが大量のデータを学習し、分類や予測などのタスクを遂行するアルゴリズムやモデルを自動的に構築する技術である。機械学習にはさまざまな方法があるが、本研究では教師あり学習の方法で、指形状とラベルの関係を学習してモデルを作成し、分類に利用する。

2.1 畳み込みニューラルネットワーク (CNN)

畳み込みニューラルネットワーク(CNN)は入力層、畳み込み層、プーリング層、全結合層、出力層から構成される。CNNは、主に画像認識の分野で用いられる。

畳み込み層では、フィルタを用いて画像の特徴を抽出する。フィルタを画像の左上から順番に重ね、画像とフ

ィルタの値を掛け合わせる。掛け合わせた数値の総和を求め、特徴マップという新たな2次元データを取得する。プーリング層では、プーリング演算すなわち縦・横方向の空間を小さくする演算を行う。全結合層は主に最終的な特徴抽出に使用され、全結合層の最後の層は分類器として使用される²⁾。本研究では畳み込みニューラルネットワークを用いて、手話認識を行うことを目的とする。

3. 手形状分類と手話認識

手形状分類とは画像分類の一つであり、手の画像からその形状をカテゴリに分類する問題である。画像分類とは、画像情報に含まれるさまざまな特徴をもとに、オブジェクトのカテゴリを区別する問題である。手形状でコミュニケーションを行う方法として指文字がある。指文字では指の形状で50音図を表現できる¹⁾。その一部を図1に示す。

本研究では、図2に示すような、Web カメラで手の形状を取り込み、機械学習で判別し結果を出力するシステムを構築することを目的とする。本研究ではOpenCVを用いて、Web カメラからの画像を読み込み、読み込んだ画像に対して、分類処理を行う。事前に訓練済みの機械学習モデルにより分類処理を行い、手形状分類と手話認識を実現する。

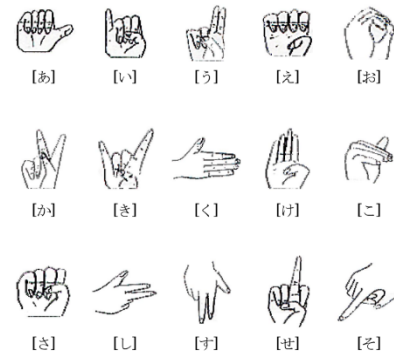


図1 指文字の一部¹⁾

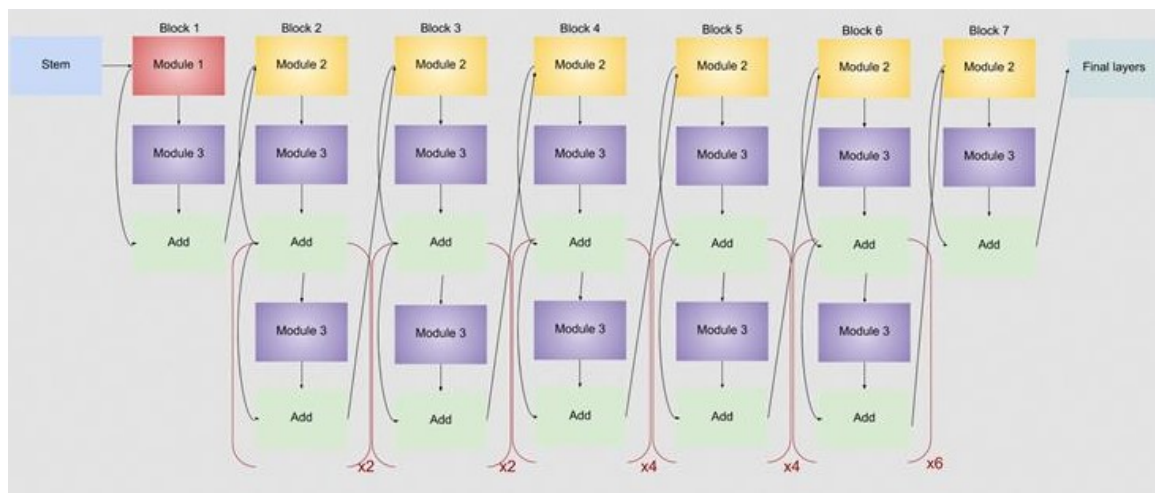


図3 用いた畳み込みニューラルネットワーク³⁾

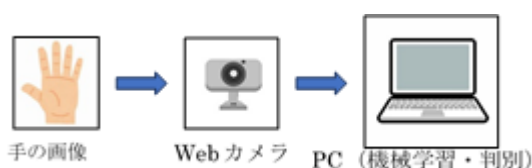


図2 システムの構成

4. 実験

指文字の学習・分類の確認のための実験を行う。最初に、学習データの作成を行う。学習データは自分で撮影した指文字画像とする。続いて CNN による学習を行う。学習データと CNN モデルを用いて学習の状況を確認する。さらに、学習済み CNN モデルを用いて指文字画像の分類を行う。用いたネットワークの構成を図3に示す³⁾。

学習データの作成を以下の通り行った。データは、画像ファイルを 46 種類の指文字とし、1 ファイル内の画像数を 200 とした。前処理を行い、画像サイズを 380×380 画素とし、0~45 までの番号のラベル付けを行った。トレーニングは 30 エポックで行い、エポック 5 のモデルを保存した。さらに学習済みのモデルを用いてテストを行った。学習結果として、30 エポックにおける誤差の推移を図4に示す。また、指文字「え」「う」の分類結果を図5に示す。

5. まとめ

今回 OpenCV を用いて、Web カメラから指文字の分類を行うシステムを構築した。指文字認識用の学習データを作成し、CNN を用いて学習・認識の実験を行った。

今後の課題として、このシステムを用いて、指文字画像から手話認識に向けて連続した指文字と手形状に基づく分類を行う方法を検討し、手話認識を実現する。

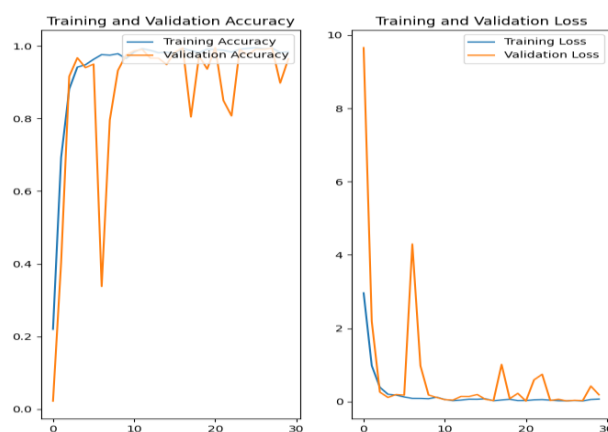


図4 学習結果

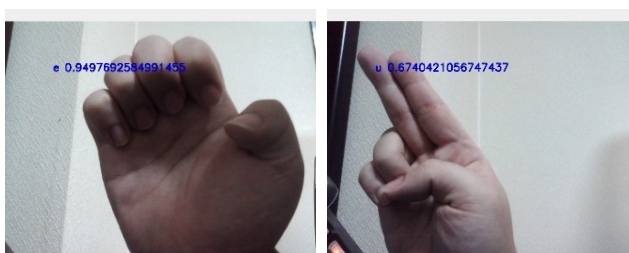


図5 指文字「え」「う」の分類結果

参考文献

- 1) 阿部 政夫 「手話-50音(指文字)」, <https://books.google.co.jp/books?id=fhctAgAAQBAJ>, 2013.
- 2) S. Raschka, V. Mirjalili, 「Python 機械学習プログラミング」, インプレス, 2020.
- 3) TensorFlow チュートリアル, 「畳み込みニューラルネットワーク」, <https://www.tensorflow.org/tutorials/images/cnn>