

データ拡張による指文字画像の分類精度の向上

1. はじめに

我が国における聴覚障害者の総数は約34万人とされている。その中には手話を習得していない人も含まれており、コミュニケーションの困難さを解決するために手話や指文字を認識するシステムが提案されている。先行研究として、安価なUSBカメラのみを用いて指文字を認識するシステム[1]が提案されているが、未学習の被験者ではあまり良い結果が得られていない。また、データ取得の段階で肌色抽出がうまくいかなかった場合にも認識率が低下してしまうことが考えられる。そこで、本研究では学習データにデータ拡張を施すことで、指文字認識の分類精度向上を目指す。

2. データ拡張による指文字画像の分類精度向上

2.1 提案手法の概要

本研究ではUSBカメラに向かって指文字を表現し、画像から手領域のみを切り出し、CNNを用いて入力された指文字の分類を行う。なお、先行研究[1]では撮影した画像をそのまま学習データとしていたが、本研究では画像にデータ拡張を行い、データ数を増加させることで分類精度を向上させる。

ディープラーニングを用いた画像分類には大量の画像データが必要なため、既存の学習データを加工してデータ数を増加させるデータ拡張が幅広い分野で用いられている[2]。主に、学習データに対して左右や上下の反転、変形、明るさの変更などの操作が行われる。また、Cutoutのような、画像に大きくマスクを入れて欠損させることでモデルの頑健性を向上させる手法なども存在する。

2.2 指文字画像のデータ拡張

図1のような画像を学習データやテストデータとして扱う際、肌色領域の欠落によって正しく学習や分類を行うことができない可能性がある。この問題点を解決するため、以下の2つのデータ拡張を提案する。

1つ目のデータ拡張として、肌色マスクを付加する。入力画像に対してランダムな位置に正方形を配置し、その範囲内の背景以外の平均画素値を算出して、その色のマスクをかける(図2)。撮影環境が暗い場合など、肌色抽出がうまくいわずに手の一部が欠けてしまっているような学習データに対して、その欠けを補う効果が見込まれる。

2つ目のデータ拡張として、黒色マスクを付加する。入力画像に対してランダムな位置に黒い正方形のマスクをかける(図3)。これにより、テストデータに肌色の欠落が生じた場合でも正しく分類できる効果が見込まれる。



図1 拡張前



図2 肌色マスク



図3 黒色マスク

表1 評価実験の結果

データ拡張		分類精度
肌色マスク	なし	70.2%
	あり	73.7%
黒色マスク	なし	64.3%
	あり	69.0%

3. 評価実験

本実験では学習データとして、明るい環境と暗い環境それぞれで撮影した指文字画像を用意し、これらの画像に対してデータ拡張を行う。すべての画像に従来のデータ拡張である回転、明るさの調整を行う。次に、生成された画像に対して元画像が暗い環境で撮影したものには肌色マスク、明るい環境で撮影したものには黒色マスクを付加する。これらの生成された画像データと元画像データをCNNに学習させる。テストデータには元画像データと異なる明るさのものをそれぞれ用意し、分類を行う。実験で得られた結果を表1に示す。実験の結果、肌色マスクや黒色マスク付加のない従来のデータ拡張のみの場合と比較し、どちらの手法でも分類精度の向上が見られた。

4. おわりに

本研究では指文字認識の分類精度向上を目指し、指文字画像に肌色マスクや黒色マスクを付加するデータ拡張を行う方法を提案した。実験の結果どちらの手法でも分類精度の向上が見られた。今後の課題として、動きのある指文字の分類などが挙げられる。

【参考文献】

- [1] 小林蒼空：MediaPipe Hands と CNN を用いた手軽な指文字認識システムの提案，福岡工業大学情報工学部情報工学科卒業論文，2024
- [2] 小林賢一，辻順平，能登正人：ディープラーニングを用いた画像処理による農作物病害診断への Data Augmentation の応用，情報処理学会第79回全国大会講演論文集，2017