

## マスクを着用した顔の画像に対する表情認識に関する基礎検討 Basic study on facial expression recognition for images of face wearing masks

湯浅 健太<sup>†</sup>  
Kenta Yuasa

高田 秀志<sup>‡</sup>  
Hideyuki Takada

### 1 はじめに

現在、機械学習を活用した顔認識は、コンピュータでの生体認証や表情の認識など多岐にわたって一般的に用いられている。しかし、新型コロナウイルスの影響からマスクをすることによって、顔の半分以上隠れ、顔認識の利用に問題が生じている。また、マスクをすると人目で見ても表情を読み取れないことがある。特にオンライン会議のようなカメラを通した状態で見ると表情や反応が読み取れなくなる。

現在使用されている表情認識は、口や目の端点から得られる特徴点を用いて、機械学習によってモデルを生成している [1]。このような学習済みモデルは、マスクを着用していない顔画像で学習を行っているため、マスクによって顔が隠されると特徴が失われ分類ができなくなる。そこで、マスクを着用していない顔画像で学習したモデルを用いて、マスクを着用した顔画像に対して表情認識を行う方法について検討する。

本研究では、オンライン会議のようなカメラ越しで相手の顔を見る状況において、表情の認識を支援する手法として、既存のモデルでマスクを着用した顔画像の表情認識を行った場合の表情認識の評価を行う。また、表情認識の精度を高めるために、マスクで隠れた鼻や口を合成し、その画像での表情認識を行った結果についても評価を行う。

### 2 関連研究

#### 2.1 表情認識におけるモデル生成

画像認識のための学習済みモデルは、様々な手法によって生成されている。表情認識においては、顔写真から目や口といった顔パーツの輪郭の一部を特徴点とし、この特徴点同士の距離を説明変数とした機械学習によるモデルの生成手法が広く用いられている [2][3][4]。

#### 2.2 表情認識手法

機械学習による表情認識では、2.1 節で述べたようなモデルを用いて表情を区別する分類問題として捉えることができる。そのため機械学習のアルゴリズムとして CNN やディープラーニング、サポートベクターマシンと言った手法が用いられることが多い。このとき分類される表情は、怒っている顔や喜んでいる顔など判別しやすいものに限られる。

### 3 マスクを着用した顔画像に対する表情認識

本研究では、表情認識を行うための学習済みモデルとして、ブレイメン大学の Octavio Arriaga によって公開されているモデルを利用した [5]。また、Microsoft Azure によって提供されている Face App を用いて表情認識の検証も行った。本実験で表情認識を行うための顔画像は、被験者 3 名に対し、マスクをした状態で Angry, Fear, Happy, Sad, Surprise, Neutral の 6 つの表情をするように伝え、各表情をしている様子を撮影したものである。

#### 3.1 Octavio Arriaga によるモデルを用いた検証

Octavio Arriaga による学習済みモデルはデータセットとして FER2013, 学習手法として Convolutional Neural Network(CNN) 用いており、表情認識の精度は、66% となっている。このモ

デルは Macro Expression においてベンチマークとなっている [6]。また、このデータは、Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral の 7 つの感情でラベル付けされている。この学習済みモデルを使用して、マスクを着用した顔画像に対して表情認識を行った。その結果を表 1 に示す。

表 1 Octavio Arriaga のモデルを用いた検証結果

被験者	正解ラベル	表情の確率 (上位 2 つ)	
被験者 1	angry	happy: 0.252	fear: 0.230
	fear	<b>fear: 0.680</b>	angry: 0.167
	happy	neutral: 0.545	<b>happy: 0.140</b>
	neutral	fear: 0.321	<b>neutral: 0.234</b>
	sad	happy: 0.243	fear: 0.236
	surprise	fear: 0.460	angry: 0.271
被験者 2	angry	surprise: 0.679	fear: 0.119
	fear	surprise: 0.793	<b>fear: 0.111</b>
	happy	surprise: 0.347	<b>happy: 0.299</b>
	neutral	surprise: 0.708	<b>neutral: 0.218</b>
	sad	surprise: 0.696	fear: 0.250
	surprise	<b>surprise: 0.344</b>	fear: 0.291
被験者 3	angry	surprise: 0.626	happy: 0.260
	fear	surprise: 0.731	sad: 0.116
	happy	<b>happy: 0.792</b>	surprise: 0.178
	neutral	surprise: 0.555	happy: 0.284
	sad	surprise: 0.389	<b>sad: 0.197</b>
	surprise	<b>surprise: 0.851</b>	happy: 0.105

全体で 18 枚のうち、4 枚の顔画像が正しく認識され、精度は約 22% となった。この結果から、マスクを着用することによって表情認識の精度が著しく低下することがわかった。

#### 3.2 Microsoft Azure によるモデルを用いた検証

Microsoft によって提供されている Microsoft Azure の画像内の顔を分析する AI サービスである Face App を用いて、3.1 節で用いたものと同じ画像を用いて表情認識を行った。

表情認識の結果、このモデルを用いた検証では 16 枚の顔画像において表情認識の結果が得られなかった。これは、表情認識の前に行う、顔の検出がマスクによってできなかったことが原因であると考えられる。表情認識を行うことができた 2 つの顔データは同じ被験者のもので、どちらも正しく認識することができていた。

### 4 マスクで隠された顔パーツの合成

3 節で述べた検証結果から、マスクを着用した顔画像に対して表情認識精度が著しく低下することがわかった。これを解決するために、本研究では、マスクを着用した顔画像に対する表情認識精度の向上を図るため、マスクで隠れた顔パーツの合成を行う。本手法における顔パーツの合成は、図 1 のように鼻から顎までの顔半分を切り出したものをマスクを着用した顔画像に合成する。パーツの合成は Adobe Photoshop を用いて手動で行った。マスクで隠れていない目や眉毛と言った特徴で表情認識に差が出ないようにするため、全ての写真に対して同じ顔パーツを用いて合成する。また、合成する顔パーツは、photoAC が提供する AI 人物素材 [7] から切り取って利用している。

<sup>†</sup> 立命館大学大学院情報理工学研究所

<sup>‡</sup> 立命館大学情報理工学部

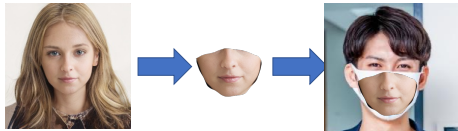


図1 顔パーツの切り取りと合成

## 5 評価実験

3節で行った検証と同じモデルを用いて、マスクを着用した顔画像に対して顔のパーツを合成した写真の表情認識を行った。

### 5.1 パーツを合成した顔画像に対する表情認識

#### 5.1.1 Octavio Arriaga による学習済みモデルを用いた検証

Octavio Arriaga による学習済みモデルを用いた表情認識結果を表2に示す。

表2 Octavio Arriaga のモデルを用いた認識結果

被験者	正解ラベル	表情の確率 (上位2つ)	
被験者 1	angry	neutral: 0.296	sad: 0.2538
	fear	neutral: 0.457	sad: 0.261
	happy	<b>happy: 0.464</b>	neutral: 0.451
	neutral	happy: 0.331	sad: 0.203
	sad	neutral: 0.345	happy: 0.247
	surprise	sad: 0.401	neutral: 0.185
被験者 2	angry	neutral: 0.489	sad: 0.245
	fear	neutral: 0.399	sad: 0.237
	happy	sad: 0.445	neutral: 0.259
	neutral	<b>neutral: 0.654</b>	sad: 0.213
	sad	neutral: 0.440	<b>sad: 0.359</b>
	surprise	neutral: 0.433	happy: 0.324
被験者 3	angry	neutral: 0.506	sad: 0.261
	fear	sad: 0.525	neutral: 0.247
	happy	sad: 0.450	neutral: 0.317
	neutral	sad: 0.550	<b>neutral: 0.245</b>
	sad	<b>sad: 0.368</b>	angry: 0.311
	surprise	neutral: 0.368	angry: 0.198

認識の結果、正しく認識できていたものは、3枚だけであった。正しく認識できていたものは、全て異なる被験者で正しい表情も happy, neutral, sad と全て異なっていた。

#### 5.1.2 Microsoft Azure による学習済みモデルを用いた検証

Microsoft Azure の Face App を用いた表情認識結果を表3に示す。顔のパーツを合成することによって、全ての画像に対して表情認識を行うことができた。しかし、16枚の画像に対して neutral とする割合が最も高かった。残る2つの顔画像はいずれも surprise とする割合が最も高く、かつ認識が正しく行われていた。

#### 5.2 表情認識の検証から得られた考察

Face App を用いた検証では、提案手法によって顔を認識することができるようになり表情認識を行えるようになった。しかし、2つのモデルを使用した検証実験の結果で neutral と判定されたものは、約72%に及んだ。このように口元の表情によって表情認識の結果に大きな影響を与えることから、口元の特徴が表情を表す特徴として学習されていることが予想される。

## 6 おわりに

本研究では、マスクを着用していない顔画像を学習データとしたモデルを用いて、マスクを着用した顔画像の表情認識を行った結果について検証と、マスクによって隠れているパーツ

表3 Microsoft Azure のモデルを用いた認識結果

被験者	正解ラベル	表情の確率 (上位2つ)	
被験者 1	angry	neutral:0.975	happy:0.022
	fear	neutral: 0.855	happy: 0.135
	happy	neutral: 0.612	<b>happy: 0.387</b>
	neutral	<b>neutral: 0.956</b>	happy: 0.044
	sad	neutral: 0.731	happy: 0.267
	surprise	surprise: 0.925	neutral: 0.047
被験者 2	angry	neutral: 0.854	happy: 0.144
	fear	neutral: 0.868	happy: 0.131
	happy	neutral: 0.713	<b>happy: 0.286</b>
	neutral	<b>neutral: 0.764</b>	happy: 0.235
	sad	neutral: 0.936	happy: 0.061
	surprise	neutral: 0.687	happy: 0.312
被験者 3	angry	neutral: 0.753	happy: 0.246
	fear	neutral: 0.899	happy: 0.097
	happy	neutral: 0.645	<b>happy: 0.347</b>
	neutral	<b>neutral: 0.961</b>	happy: 0.037
	sad	neutral: 0.961	happy: 0.03
	surprise	<b>surprise: 0.501</b>	neutral: 0.476

を合成した場合の表情認識の結果について検証を行った。検証の結果、マスクを着用している場合には表情認識を正しく行うことができなかった。また、顔のパーツを合成した顔画像に対する表情認識の結果は、ほぼ neutral と判定された。

今後は、オンライン会議のような状況でマスクの着用に関わらず表情認識が行えるような学習済みモデルを構築するための手法を考案する。

## 謝辞

本研究の実験を行うにあたって顔画像を提供していただいた友人知人各位に感謝します。

## 参考文献

- [1] 小林宏, 原文雄: ニューラルネットによる人の顔の6基本表情の強さの計測に関する研究, 日本機械学会論文集 C 編, Vol. 59, No. 567, pp. 3411–3417 (1993).
- [2] Hsu, R.-L., Abdel-Mottaleb, M. and Jain, A.: Face detection in color images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 696–706 (2002).
- [3] 岡田敦志, 上村謙史, 目良和也, 黒澤義明, 竹澤寿幸: 表情・音響情報・テキスト情報からのリアルタイム感情推定システム, 人工知能学会全国大会論文集, Vol. JSAI2017, pp. 1B10S25a4–1B10S25a4 (2017).
- [4] 野宮浩揮, 宝珍輝尚: 顔特徴量の有用性推定に基づく特徴抽出による表情認識, 知能と情報, Vol. 23, No. 2, pp. 170–185 (2011).
- [5] TensorFlow.js で「感情」を認識してみよう, <https://book.mynavi.jp/manatee/detail/id=99887> (2021).
- [6] Li, S. and Deng, W.: Deep Facial Expression Recognition: A Survey, *CoRR*, Vol. abs/1804.08348, (2018).
- [7] photoAC, <https://www.photo-ac.com/main/genface> (2021).