

ディープニューラルネットワークによる微表情の検出と分類

情報科学科 上村 祥之

指導教員：小林 邦和，鈴木 拓央

1 はじめに

微表情 [1] とは人が感情を隠そうとしている時に発生する表情であり，本人の意識にかかわらず発生するものである．またロボットが人間の感情を理解するてがかりになると考えられている．しかしこの微表情は表情の微小の動きであり，発生時間が短いことから検出をすることが困難である．

そこで本研究では多様な特徴抽出が行えるディープニューラルネットワークを用いて，微表情の検出を行うことを試みる．様々なディープニューラルネットワークがある中でも，画像の特徴を抽出し画像特徴ごとのわずかな違いを学習することができるネットワークである畳み込みニューラルネットワーク (CNN)[2] を用いて微表情の検出と分類を行う．

しかし CNN は多層のネットワークであるため計算に時間がかかりすぎてしまう．そこで転移学習 [3] を用いて，既存の学習済みモデルを利用し一部の層の重みを再学習することで計算時間の削減を図る．

2 提案手法

顔画像のみを入力データとするために OpenCV のカスケードを利用し，顔画像を切り出した．

さらに本研究では CNN を用い図 1 のように一部分のみを再学習する転移学習を行う．手法の流れは下記のように行う．

1. 既存の学習済みネットワークを用意する．
2. クラスを分類するための最終層である全結合層を微表情の画像データを使い再学習させる．

全結合層のみを学習させることによって本来学習する重みの数が格段に減り，計算コストを抑えることができる．これにより計算量が減り，より高速で効率的に多層なネットワークの学習を行うことができる．

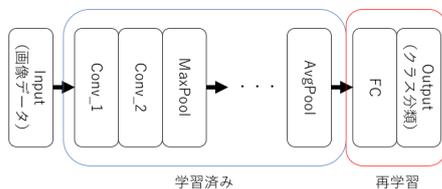


図 1: 転移学習を適用するネットワークの構造

3 計算機シミュレーション

CNN の転移学習済みのモデルは ILSVRC-2014 の分類部門で優勝した inception-v3[4] を用いる．

3.1 データセット

微表情のデータセットとして CASME II[5] を利用した．このデータセットは微表情が発生する前から微表情が発生した後までの映像データで，フレームレートは 200fps である．画像のラベルは平常時である neutral と微表情発生時である micro_expression の 2 クラスに大別され，微表情のクラス分類は happiness, disgust, surprise, repression, others の 5 クラスである．図 2 が各クラスの実際の画像である．

学習に用いる各クラスの画像の枚数が均等ではなかったことから，均等になるように微表情の検出では各クラスから 13796 枚ずつ，分類では 1605 枚ずつランダムに選択した．

3.2 結果

結果を表 1 に，誤分類された画像のクラスを表 2 に示す

表 1: シミュレーション結果

	微表情の検出	微表情の分類
精度 (%)	72.1	92.3

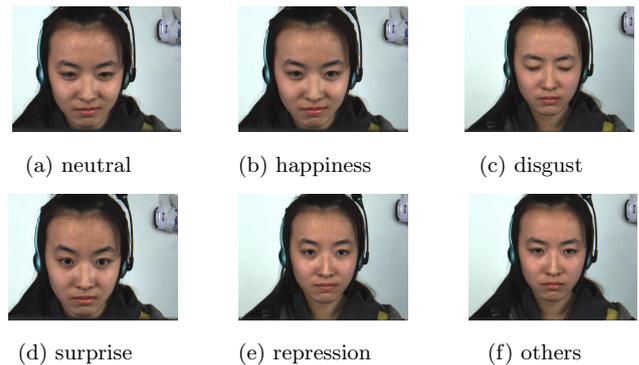


図 2: 微表情のデータセットの一部*1

表 2: シミュレーションごとの誤分類の割合

クラス	誤分類の割合 (%)
neutral	40.6
micro_expression	59.4
happiness	21.0
disgust	14.5
surprise	1.6
repression	11.3
others	51.6

3.3 考察

検出においては微表情の発生後すぐの画像と発生前の画像は違いが見えづらいということから 72.1% という精度が出たことが考えられる．最も微表情が顕著に現れたタイミングの画像を多く使うことによって精度が向上すると考えられる．

分類においてはクラス others の画像の誤分類が多かった．クラス others はどこにも分類されないものであることから，さらに細分化し新たなクラスを作るべきであると考えられる．

4 おわりに

本研究では転移学習済みの CNN を一部分再学習したネットワークを用いて微表情をどの程度検出，分類ができるかを検証することができた．

今後の課題としてデータセットを時系列データとしてみることで精度を上げることができるのではないかと考えられる．学習と実行の方法を時系列データに対応していく必要がある．

参考文献

- [1] Haggard, E. A. and Isaacs, K. S.: "Micromomentary Facial Expressions as Indicators of Ego Mechanisms in Psychotherapy", Methods of research in psychotherapy, Springer, pp.154-165(1966)
- [2] Yann LeCun, Léon Bottou, Yoshua Bengio, Patric Haffner: "Gradient-based learning applied to document recognition" Proc. of the IEEE, pp.2278-2324(1998)
- [3] 中山英樹:「深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習」, 信学技報 115(146), pp.55-59(2015)
- [4] C.Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna: "Rethinking the inception architecture for computer vision", arXiv preprint arXiv:1512.00567, (2015)
- [5] Wen-Jing Yan, Xiaobai Li, Su-Jing Wang, Guoying Zhao, Yong-Jin Liu, Yu-Hsin Chen, Xiaolan Fu: "CASME II: An Improved Spontaneous Micro-expression Database and the Baseline Evaluation", PloS One, Vol.9, No.1, p. e86041 (2014)