

深層学習に基づくオンライン授業における疲労検出システム

劉 昊 齋藤 友彦

湘南工科大学 工学部 情報工学科

1. はじめに

近年、COVID-19の影響により、ZoomやMeetなどを使ったオンライン授業が急速に普及した。本研究では、オンライン授業に映った受講者の映像(図1)から受講者の疲労度を検出するシステムの開発を目的とする。



図1 オンライン授業に映った受講者の映像

これまで授業の様子を写した映像から受講者の疲労度や集中度を測定する試みが多くなされている[1,2]。しかしながら、これらの多くは計測用のカメラを別に設置している。本研究では、COVID-19以降普及した、一般的なオンライン授業やゼミを想定し、図1のような映像のみから疲労度を推定するシステムの構築を目指す。

本研究で提案するシステムの概要は次の通りである。オンライン授業の映像からハフ変換による直線検出を使い、受講生一人一人が映る枠を求める。そして、各受講生が映る枠に対して顔検出[3]、顔のランドマーク検出[4]を行う。また、本研究では疲労度を表す指標として、目を閉じる[5]、欠伸、という二つの行為に注目し、これらの検出を行う。

本研究では、実験を通じて次の二つを検証する。一つは、受講生一人一人が映る枠を求めることの顔検出精度への影響である。もう一つは提案する欠伸の検出精度である。

2. 疲労度検出システム

2.1 直線検出を用いた枠の推定と顔検出

オンライン授業の映像に対して、ハフ変換による直線検出を行い、それぞれの受講生が映る映像と映像の境界線を求める。得られた境界線の交点を求め、それぞれの受講生が映る映像の枠

Fatigue detection system for online classes based on deep learning.

Liu Hao, Saito Tomohiko

Shonan Institute of Technology, Faculty of Engineering,
Department of Information Science.

を決定する。得られた枠のそれぞれに対してYolov8[3]を適用し、顔を検出する。なお、本システムでは、一般的なオンライン授業を想定し、各枠に映る受講生は一人と想定する。顔検出においても、各枠で検出され最もスコアの高い顔を受講生の顔とする。図2に直線検出なしとあり、それぞれにおける顔検出結果の例を示す。

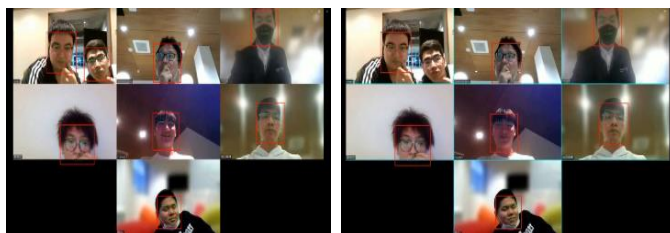


図2 顔検出(左は直線検出なし、右はあり)

2.2 ランドマーク検出

2.1で検出された顔に対して、PFLD[4]を用いて顔のランドマークを検出する。PFLDでは次の損失関数を用いる。

$$L = \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N \left(\sum_{c=1}^C \omega_n^c \sum_{k=1}^K (1 - \cos \theta_n^k) \right) \|d_n^m\|_2^2 \quad (1)$$

M は訓練画像の数、 N はランドマーク数、 $\|d_n^m\|_2$ は m 枚目の画像、 n 番目のランドマークにおける推論値とターゲットとのL2ノルムである。

$\theta_1, \theta_2, \theta_3 (K=3)$ は実際のヨー、ピッチ、ロールと推定値との偏差角である。また、重み関数 ω_n^c はクラス C に属するサンプルの割合に応じて調整される。なお、本研究では検出精度を向上させるため、1枚の訓練画像に対して、左右を反転させた画像、及び -30° から 30° までランダムに回転させた9枚の画像を加えた全てを訓練画像とする。

2.3 疲労検出

2.3.1 目を閉じる

文献[5]と同様の手法を用いて、目を閉じている(寝ている)ことの検出を行う。図3に2.2で得られたランドマーク(目の部分に着目)の例を示す。このとき、以下の定義式を用いてEAR(Eye Aspect Ratio)を計算する。

$$EAR = \frac{\|P_{61} - P_{67}\| + \|P_{63} - P_{65}\|}{2\|P_{60} - P_{64}\|} \quad (2)$$

ただし、 $P_{60}, P_{61}, P_{63}, P_{64}, P_{65}, P_{67}$ は図3における該当番号の座標位置を示す。また、式(2)は左目のEARを表す。左及び右目のEARの平均値を求め、その値が2.5秒以上0.28未満である場合

画面に「TIRED」と出力する。



図3 目のランドマーク

2.3.2 欠伸

2.2 で得られたランドマークを使い、欠伸の検出を行う。図4に2.2で得られたランドマーク（口の部分に着目）の例を示す。本研究では次式で定義される MAR (Mouth Aspect Ratio) を提案する。

$$MAR = \frac{\|P_{89} - P_{95}\| + \|P_{91} - P_{93}\|}{2\|P_{88} - P_{92}\|} \quad (3)$$

ただし、 P_{88} , P_{89} , P_{91} , P_{92} , P_{93} , P_{95} は図4における該当番号の座標位置を示す。MARが4秒以上0.1以上である場合、欠伸をしたと判定し、画面に「TIRED」と出力する。

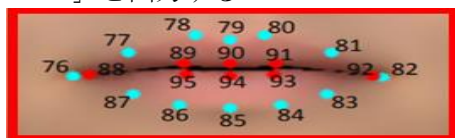


図4 口のランドマーク

3. 実験

3.1 顔検出に関する実験

2.1 で述べた受講生一人一人が映る枠を求めることで、顔検出にどのような影響があるかを実験により検証する。図1のように7名で実施したゼミの映像（約50分）の内、人の顔のみが映った時間（画面共有等をしていない時間）の画像をランダムに100枚抽出し、全ての受講生の顔検出が成功している割合を算出した。なお、本実験で使用した映像は実際のゼミの映像であり、誰が受講生（ゼミに参加している学生）かはあらかじめ分かっている。特に図1の一番左上の枠に2名映っているが、一人は受講生、もう一人は見学者である。この映像において見学者の顔を検出してしまった場合、検出失敗と判断する。実験の結果、成功率は直線検出をしない場合は19%、する場合は73%となった。

3.2 欠伸の検出に関する実験

2.3.2 で述べた欠伸の検出の精度について実験により検証する。図5のように一人のみが映った映像（約5分）の中で、欠伸の演技を50回行い、それらが検出されたかを検証した。実験の結果50回全ての欠伸を検出した。

4. 考察

3.1.の実験の結果、直線検出をし、受講生一

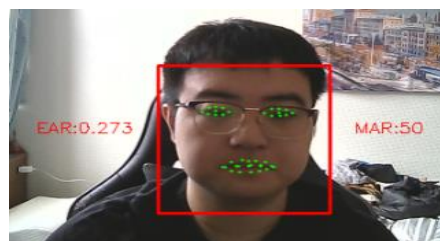


図5 欠伸の実験結果

人一人が映る枠を定めることで、顔検出の成功確率が上がることが確認された。これは、本実験で使用した映像の大部分にゼミには関係のない見学者が映っており、それを受講生の顔として検出していたからである。今回使用した映像のように、授業やゼミに関係のない人が映る場合、本システムで提案した受講生一人一人の枠を定め、その枠内で一つの顔検出を適用するという方法は特に有効である。また、3.2の実験の結果、欠伸の演技の全てを検出した。従って、本研究で提案した欠伸の検出法が有効であることが確かめられた。しかしながら、今回の実験では、単純に口を大きく開ける欠伸の演技のみをしたので、今後は様々なタイプの欠伸を調べ、それらの検出法を検証する必要がある。

5. まとめ

本研究では、オンライン授業の映像から受講生の疲労度を検出するシステムの提案、実装を行った。本稿では、疲労を表す指標として、目を閉じる、欠伸のみに着目した。今後の課題として、疲労について詳細に調べ、より精密な疲労度推定法の提案が挙げられる。

謝辞

本研究は JSPS 科研費 18K11585, 22K02811 の助成を受けたものです。

参考文献

- [1] 藤白智也, 浅利恭美, 宮田真宏, 山田徹志, 大森隆司, “授業中の画像からの姿勢推定による集中度の評価,” 人工知能学会第34回全国大会論文集, (2020).
- [2] 府馬央昂, 鷹野孝典, “受講生画像からの表情・姿勢推定に基づく学習状態判定機能を備えた遠隔講義システムの開発,” JSiSE Research Report vol.34, no.5, (2020).
- [3] W. Chen, H. Huang, S. Peng, C. Zhou and C. Zhang, “YOLO-face: a real-time face detector.” The Visual Computer: International Journal of Computer Graphics, vol. 37, pp. 805-813, (2021).
- [4] X. Guo, S. Li, J. Yu, J. Zhang, J. Ma, L. Ma, W. Liu and H. Ling, “PFLD: A practical facial landmark detector.” arXiv preprint arXiv:1902.10859 (2019).
- [5] T. Soukupova and J. Cech, “Eye blink detection using facial landmarks,” 21st computer vision winter workshop, (2016).