

ディープラーニングを利用した顔画像による着席状態の集中度解析

Shen Jinjian[†] 平石広典[‡]足利大学大学院 工学研究科 情報・生産工学専攻^{†‡}

1. はじめに

近年、コロナの影響で対面授業からオンライン授業の割合が多くなっている。オンライン授業中の学生の状態の把握は難しいのが現状である。集中や散漫等の状態を確認できれば、授業内容で補足説明が必要な場合や、学生の興味の状態などを把握することが可能となり、授業を効率的に実施することができるものと考えられる。

本研究では、利用者の顔画像を元にディープラーニングを利用した着席者の集中度の解析を試みる。その際に、利用者の集中度の基準となるデータとして、簡易脳波センサと座圧センサを利用した。脳波センサによって利用者の注意や散漫の状態を検出することが可能であり、座圧センサによって利用者の疲労の状態を検出することができる[1]。本研究では、各センサによって得られた集中度をもとに、利用者の着席時の顔画像を収集し、ディープラーニングによる分類を行なった。

2. 簡易型脳波センサと座圧センサ

本研究では簡易型脳波センサとして NeuroSky 社によって開発された MindWave Mobile を利用した(図1左)。このセンサでは、8種類の脳波データ(Delta/Theta/Low Alpha/High Alpha/Low Beta/High Beta/Low Gamma/Mid Gamma)を収集することができる。また、独自のモジュールにより、集中度とリラックス度を0~100の数値で取得することが可能である。それぞれのデータは1秒毎に収集することが可能である。



図1: 簡易型脳波センサ(左)と座圧センサ(右)

座圧センサとして SR ソフトビジョンを利用した(図1右)。この座圧センサでは、専用のソフトウェア(図2)を利用して圧力分布に加え、圧力値、面圧中心位置をリアルタイムに可視化することができる。データを収集する頻度は自由に設定することが可能である。

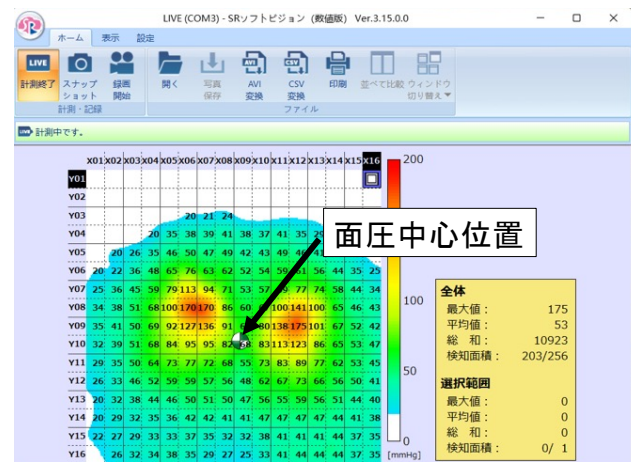


図2: 座圧センサの専用ソフトウェア

3. データ収集

本研究では1人の被験者(26歳、男子留学生)に対して2回の実験を行なった。被験者は脳波センサを装着し、椅子の上に座圧センサを敷いた状態で、30分のドラマ(アメリカン・ホラー・ストーリー)を視聴してもらい、その時のセンサデータと顔画像を収集した。座圧データに関しては5秒に1回のデータを収集するように設定し、顔画像も5秒毎にデータを保存するように設定した。図4は1回目の実験の結果であり、図5は2回目の実験の結果を示している。それぞれ上のグラフは脳波データによる集中度であり、下のグラフは座圧データの1分毎の標準偏差の値を示している。座圧データに関しては、値が大きいほど疲れていることを意味する[1]。

図4の1回目の実験結果を見ると、時間が経過するにしたがって、脳波による集中度は下がる傾向があり、それに伴って、座圧データの値も大きなものが見られ、疲れる傾向が見られる。一方、図5の2回目の実験結果では、1回目とは逆に、徐々に集中度が高くなる傾向が見られ、また、座圧データの値は、徐々に小さくなる傾向がみられる。つまり、1回目の実験では、最初に集中し、徐々に疲れて集中が減少するパターンであり、2回目の実験では、最初はあまり集中

Analysis of concentrating level based on face mage using deep learning

Shen jinjian, Ashikaga university[†]

Hironori Hiraishi, Ashikaga university[‡]

できず、徐々に集中していくパターンである。

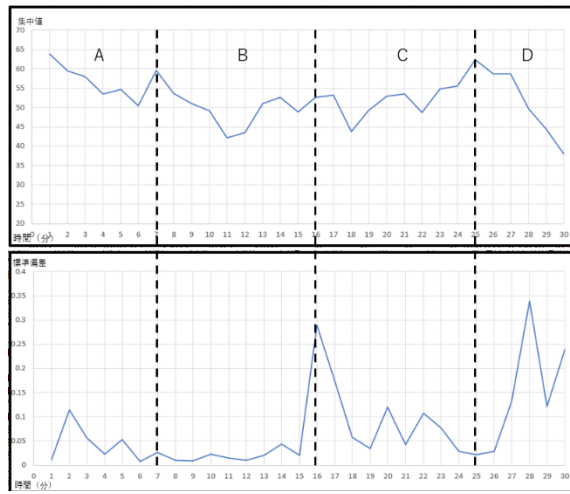


図4: 1回目の脳波データ(上)と座圧データ(下)

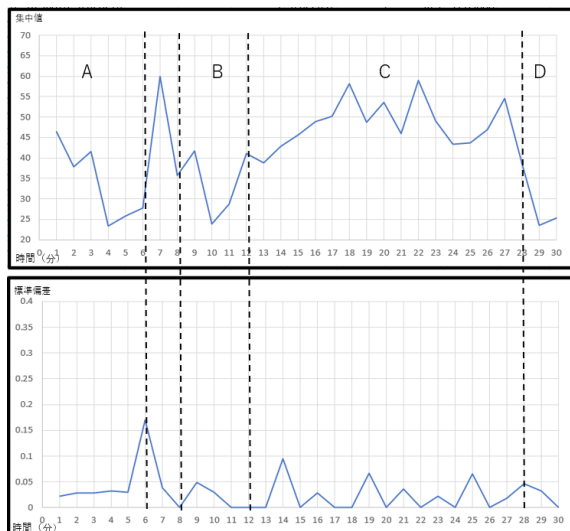


図5: 2回目の脳波データ(上)と座圧データ(下)

次に、脳波データと座圧データを元に、収集した顔画像を、ディープラーニングによる分析を行うために、集中状態と散漫状態に分類する。ここで、2回目の実験において、A区間(0-7分)は集中していると同時に疲れてない状態であり、B区間(7-16分)はやや疲れが見られる状態である。C区間(16-25分)は集中しているが疲れた状態であり、D区間(25-30分)は、集中も途切れ、疲れも見られる状態を表している。したがって、A区間は疲れもなく完全に集中した状態であり、逆に、D区間は完全に疲れていて集中が途切れ状態であるため、A区間の顔画像データ(84枚)を集中状態とし、D区間の顔画像データ(60枚)を散漫状態として分類を行なった。

2回目の実験では、C区間(12-28分)は集中していると同時に疲れてない状態であり、A(0-6分)、B(8-12分)、D(28-30分)区間は、集中も途切れ、疲

れも見られる状態を表しているため、C区間の顔画像データ(192枚)を集中状態とし、A, B, D区間の顔画像データ(108枚)を散漫状態として分類した。

4. ディープラーニングによる分析

本研究では機械学習ソフトウェアの Weka [2]を利用して、1回目の実験の顔画像を利用して分類モデルの生成を行ない、2回目のデータを利用して、その分類モデルの評価を実施した。Wekaでは、様々なディープラーニングのモデルを利用することができるが、本研究では、実行可能であった3つのモデルを利用して評価を実施した。

表1は、それぞれのモデルの分類精度を示している。LeNet は非常に精度が高く、FaceNet と SqueezeNet の精度は低い結果となった。FaceNet と SqueezeNet の精度の低い原因は、識別の特徴量が多いため、データ量が少なくいために、学習の実行が十分でなかったことが考えられる。

表1: 実験結果

モデル	分類精度
LeNet	99.12
FaceNet	57.68
SqueezeNet	52.62

5. まとめ

実験中の着席者の集中度の状態を計測することを目指して、脳波センサと座圧センサを利用し、顔画像の分析を行なった。今回の実験において、精度が低い原因は以下の3つの要因が考えられる。

本研究では、実験の容易さから30分のドラマを視聴する実験を行なった。しかし、オンライン授業等を考えた場合には30分は短い。今後は実際のオンライン授業を想定し、90分程度のデータを分析を実施する。また、実験における実験者の数を増やし、集中度の識別を目標として実験をする。

6. 参考文献

- [1] Li Yuchuan, 平石広典, 機械学習による座圧センサを利用した着席者の疲労状態の解析, 情報処理学会第83回全国大会, Vol.2, pp.491-492, 2021.3.
- [2] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.
- [3] Florian Schroff Dmitry Kalenichenko James Philbin "FaceNet: A Unified Embedding for Face Recognition and Clustering" arXiv:1503.03832v3[cs.CV] 17 Jun 2015