

Title	オンライン動画講義におけるエンゲージメント推定を向上させるためのスキップ移動平均の時系列データ処理への適用
Author(s)	鄭, 羨文
Citation	
Issue Date	2025-03
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/19925
Rights	
Description	Supervisor: 長谷川 忍, 先端科学技術研究科, 博士

氏 名	ZHENG Xianwen		
学 位 の 種 類	博士（情報科学）		
学 位 記 番 号	博情第 547 号		
学 位 授 与 年 月 日	令和 7 年 3 月 21 日		
論 文 題 目	Application of Skipped Moving Average in Time-Series Data Processing to Enhance Engagement Estimation in Online Video-based Lectures		
論 文 審 査 委 員	長谷川 忍	北陸先端科学技術大学院大学	教授
	池田 心	同	教授
	岡田 将吾	同	教授
	謝 浩然	同	准教授
	柏原 昭博	電気通信大学	教授

論文の内容の要旨

Online learning, which gained traction in the 1980s, has expanded access to education by offering flexible, interactive, and effective learning opportunities through web-based platforms and learning management systems. The success of online learning is heavily influenced by factors such as self-motivation, student engagement, and interaction between students and instructors. Increased engagement can boost self-motivation and foster more effective interactions, ultimately enhancing the quality and experience of online education. Therefore, improving learner engagement is crucial in overcoming challenges like the digital divide, limited face-to-face interaction, and issues related to self-motivation.

However, in engagement estimation research, due to factors like the Hawthorne effect, existing public datasets often suffer from class imbalance, with relatively few data points representing low engagement levels. This imbalance presents a significant challenge in accurately training and validating machine learning models for engagement estimation. We introduce an original preprocessing approach called "Skipped Moving Average," which not only preserves the integrity of the original video data but also captures its temporal dynamics and variations to address the imbalance issue.

First, to enrich the existing computer vision features and better interpret learners' facial and body language during online learning, we adopted a series of features that can represent facial and body information. Additionally, to further enhance our input features, we experimented with features such as standard deviation and extreme values. We then introduced our proposed Skipped Moving Average data processing method, which includes selecting an appropriate skipping window based on the current data distribution, as well as how to reasonably choose oversampled data sequences using cosine similarity. We also experimented with different normalization methods to evaluate their effectiveness in processing video sequence data.

In the experimental phase, we divided the work into two major parts. Experiment 1 used LSTM and LSTM-FCN models to verify whether the proposed SMA preprocessing method could address the issue of imbalance in the current video sequence data. Ultimately, the combination of Skipped Moving Average Oversampling and Standard Deviation for training and validation produced the best outcomes. For engagement estimation with different labels, it achieved Recall/Precision/F1 scores of 0.462/0.157/0.234 for the low label,

0.449/0.504/0.475 for the high label, and 0.456/0.501/0.477 for the very high label. To further validate our proposed method, we also compared it with the SMOTE oversampling method, which further demonstrated the superiority of our approach.

In Experiment 2, we used transfer learning to verify that the proposed SMA data processing method could be applied to different datasets. The three datasets used in the experiment had varying sample time spans and were quite irregular. Our proposed method achieved Recall/Precision/F1 scores of 0.635/0.720/0.675 for the low engagement label, which is an improvement of nearly 0.25 in the F1 score compared to the results before applying transfer learning.

In this study, we tackled the challenge of class-imbalanced time-series video data in the context of engagement estimation and detection by introducing a novel approach: Skipped Moving Average oversampling. This approach not only mitigates the effects of class imbalance but also preserves the continuity and authenticity of the time-series data, leading to more precise and consistent results in engagement detection.

Keyword: emotional engagement estimation, time-series data, oversampling, online learning, class imbalances data

論文審査の結果の要旨

本論文は、学習者の低エンゲージメントの推定精度向上を目指し、時系列ビデオデータにおけるクラス不均衡問題を解決するためのオーバーサンプリング手法を提案することである。特にオンライン学習は、対面授業に比べて学習者のインタラクションが少なくエンゲージメントが下がりやすいことが指摘されている。こうした環境では、学習者の低エンゲージメント状態を早期に検知することが求められるが、公開されているエンゲージメントデータセットは低エンゲージメントクラスのデータが少ないというクラス不均衡の問題があり、モデルの訓練と検証が困難である。本論文では、スキップ移動平均（SMA）を用いたオーバーサンプリング手法が提案された。この手法は、動画フレームを一定間隔でスキップしながら特徴量を平均化することで、ビデオ時系列データの時間的な特性を保持しつつ、ノイズを低減しながらデータ量を増やすことが可能であり、クラス不均衡の緩和が期待できる。また、感情的なエンゲージメントの分析に焦点を当て、顔の表情だけでなく、体の動きも考慮した包括的な特徴量設計を採用している。これらの特徴量には、頭部の姿勢、顔の回転角度、目の情報、口の形、体の動き、顔とスクリーンの距離などが含まれている。評価として、公開データセットである DAiSEE を用いて、SMA オーバーサンプリングをランダム移動平均(RMA)や SMOTE といった既存のオーバーサンプリング手法と比較し、LSTM および LSTM-FCN モデルを用いて性能を評価した。その結果、SMA オーバーサンプリングが、他の手法よりも高いエンゲージメント推定精度を達成することが示された。また、DAiSEE で訓練したモデルを、転移学習を通じて他のデータセットに適用する実験も行い、提案手法が異なるデータセットにも適応可能であることを示した。提案手法は、既存の手法と比較して、データの時間的特性を保持し、ノイズの影響を低減する点で優れており、転移学習でも適用可能であることから、オンライン学習環境における学習者のエンゲージメント評価をより効果的に行うことが期待できる。

以上、本論文は、学習者のエンゲージメント推定を向上させるためのデータ処理方法の提案について述べたものであり、時系列データのクラス不均衡問題への対処と、実用的なエンゲージメント

推定モデル開発から，学術的に貢献するところが大きい．よって博士（情報科学）の学位論文として十分価値あるものと認めた．