

瞳検出技術を用いた動画学習支援システム

An Eye Pupil Detection-Based Review Support System for Video-based Learning

DUAN ZHENZHONG^{1*} SHAN JUNJIE² 西原陽子²

¹ 立命館大学大学院 情報理工学研究科

¹ Graduate School of Information Science and Engineering, Ritsumeikan University

² 立命館大学 情報理工学部

² College of Information Science and Engineering, Ritsumeikan University

Abstract: 近年、動画学習を利用する機会が増えている一方、対面授業で勉強するより学習者は集中力を維持しにくくなる傾向がある。本研究では、デバイスの正面カメラでユーザの視聴状況を計測し、学習動画への視聴状況を可視化するグラフを提供する動画学習の支援システムを提案する。20名の被験者に動画学習を実施させ、事後テストの正解率により提案システムを評価した結果、実験群の平均正解率は93%、対照群の平均正解率は82%となった。

1 はじめに

近年、コロナウイルスの流行により、世界各地ではオンライン学習が実施されている[1][2][3]。オンライン学習の一種として、「動画学習」が時間や場所などの制限がないため、学習者に人気があって幅広く利用されている[4][5]。

しかし、動画学習にはデメリットがある。伝統的な対面授業と違い、教師やクラスメイトが近くにいないため、学習者は外部からの妨害で集中が途切れやすくなる。さらに、長時間の学習動画の場合、学習内容は多岐にわたるため、学習者個人の記憶のみで各自の復習すべき時間帯を把握するのも容易ではないと考えられる。結果として、復習時間を増やしても、期待される復習効果が得られない場合もたくさん発生する[6]。

そこで、本研究は瞳検出技術を用いて、動画学習の利用者の復習を支援するシステムを提案する。提案システムは、学習者が学習動画を視聴すると同時に瞳を検出し、その学習者の視聴状況¹を計測する。計測した視聴状況を学習動画の時間軸と整合し、ユーザの視聴状況を反映する視聴状況図を描画する。ユーザは視聴状況図から学習動画の復習すべき具体的な時間帯の情報を簡単に把握できる。これにより、提案システムを通して、ユーザは動画学習への復習効率を向上させる効果が期待される。

本研究では、以下3つの内容を行った：

(1) 人間の瞳を検出するためのYOLOv5モデルを作成した。

(2) 瞳の検出結果により、ユーザの視聴状況を計測及び学習動画の時間軸に合わせて視聴状況図を描画する復習支援システムを実装した。

(3) 被験者実験を実施して、提案システムの動画学習への復習支援効果を確認した。

2 関連研究

本章では、学習支援に関する既存研究を述べ、本研究の位置付けを行う。

オンライン学習の利用者の学習効果と満足度に関して、Ikhsanらの研究がある[7]。この研究では、Structural Equation Modelingを用いて、オンライン授業を受ける大学生の学習の効果と満足度の決定要因を探索した。潜在的な決定要因として、教師からの促進、技術的な支援、カリキュラムの構造、教師からのフィードバック、自己啓発、仲間からの協力があり、6つの側面から調査を行った。結果として、教師からの促進、技術的な支援、教師からのフィードバック、自己啓発は学生の学習効果に影響を与えることが示された。この研究から、オンライン学習を受ける学習者にとって、技術的な支援は重要で有益であることが示された。本研究では技術的な支援として、瞳検出技術を用いた復習すべき時間帯を学習者に知らせ、その効果を検証する。

Yehらは、学習者を監視する研究を行った[8]。この研究では、Dlib1を用いて瞳検出を行い、まばたきを指標にすることで学習者の不正行為を検出し、瞳の角度

*連絡先：立命館大学情報理工学研究科

〒567-8570 大阪府茨木市岩倉町 2-150

¹本文では視聴状況を「ユーザは学習動画を視聴しているかどうか」とする。

を用いて学習者が適切な学習態度や行動を示しているかどうかを判断する。Yeh らの研究では、デバイスのカメラを利用して瞳検出を行い、学習者のオンライン学習を監視した。瞳検出の結果からユーザの視聴状況を分析できることを示した。他に、Abdulkader らは、オンライン学習を利用する学生の授業に対する集中力を分析する研究を行った [9]。このシステムは顔認識アルゴリズムとディープラーニングモデルである「DLIP」を用いて、対象の目と口の特徴を抽出する。しかし、この2つの研究の目的は学習者の学習状態を教師にフィードバックすることであるため、教師が存在するライブ配信授業を前提としていた。また、学習者を監視することで教師の授業を支援するため、学習者に直接的な支援を提供しなかった。この2つの研究と異なり、本研究は教師が存在しない動画学習を対象とし、学習者に視聴状況図を提供することで、学習者に直接的な支援を提供する。

Raca らは、対面授業中にクラス全員の集中状況を監視し、教師にフィードバックできるシステムを提案した [10]。このシステムを利用することで、教師はリアルタイムで学生がどの程度授業を受けているかを把握し、授業のスケジュールを調整することができる。Abdulrahman らは、コンピュータビジョン技術を用いて、学生の顔を特定し、対面授業に集中しているかどうかを判断して、教師に示せるシステムを提案した [11]。この2つのシステムと異なり、本研究の提案システムは対面授業ではなく、オンライン学習の動画学習に注目し、直接学習者に復習支援の視聴状況図を提供する。

3 提案システム

本章では提案システムについて説明する。

3.1 提案システムの流れ

図1に提案システムの流れを示す。はじめに、提案システムを起動し、デバイスの正面カメラでユーザの学習動画を視聴する時の顔画像を撮る。次に、撮った顔画像に対して、再訓練されたYOLOv5モデルで瞳検出を行い、検出結果を記録する。システムは検出結果からユーザの視聴状況を判断する。学習動画の視聴が終了した後、ユーザの学習動画を視聴している時間帯と視聴していない時間帯を特定し、ユーザが理解しやすいように、視聴状況図を可視化する。

3.2 顔画像からの瞳検出

顔画像から瞳検出を行う方法を説明する。提案システムではYOLOv5を用いて瞳検出を行った。

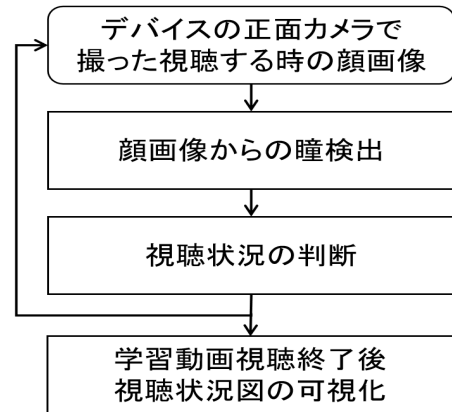


図1: 提案システムの流れ

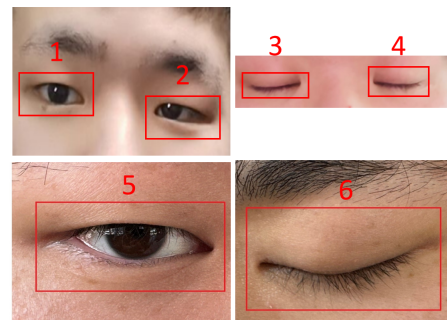


図2: 教師データの例

YOLOv5 公式サイトで公開されたモデルは人間の瞳を検出できないので、再訓練する必要がある。教師データとして、「開いている両目 (1, 2)」、「閉じている両目 (3, 4)」、「開いている片目 (5)」、「閉じている片目 (6)」(図2)、4種類の目の画像各248枚を用意し、手作業でラベリングした。各画像のラベルは目が開いているか閉じているかを示す0または1の数値である。目が開いていると1、目が閉じていると0とする。加えて、瞳の位置情報を画像にアノテーションとして付与した。位置情報として4つの数値があり、それぞれは画像の幅と高さを基に標準化された瞳の中心点のx座標、中心点のy座標、瞳の範囲を示す枠の幅と枠の高さである。画像に含まれる目の数に応じて、記録された目のラベルと位置情報の数も変化する(表1)。

YOLOv5は事前訓練モデルを5つ提供している。それぞれは「YOLOv5n」、「YOLOv5s」、「YOLOv5m」、「YOLOv5l」、「YOLOv5x」といい、検出精度が順に高くなるが、処理速度が順に遅くなる。本研究では、処理速度と検出精度を両立させる「YOLOv5m」モデルを用いて、教師データで再訓練を行った。その際に設定したパラメータを表2に示し、再訓練の結果を表3に示す。再訓練されたモデルを用いることにより、精度の高い瞳検出が可能になる。

表 1: 図 2 の画像に対応する目のラベルと位置情報

目の ID	ラベル	x 座標	y 座標	幅	高さ
1	1	0.20	0.42	0.34	0.29
2	1	0.78	0.26	0.33	0.30
3	0	0.19	0.57	0.25	0.21
4	0	0.68	0.49	0.32	0.21
5	1	0.53	0.48	0.85	0.55
6	0	0.49	0.36	0.88	0.64

表 2: YOLOv5m の再訓練のパラメータ

パラメータ	数値
GPU Memory	7.2 GB
Batch Size	16
Epochs	60
Learning Rate	0.01

3.3 視聴状況の判断

瞳検出の結果から、視聴状況を判断する。瞳検出は 1/30 秒ごとに行われる。はじめに瞳検出の結果を 4 種類に分ける。

- (1) 瞳が検出されなかった
- (2) 閉じている両目が検出された
- (3) 開いている両目が検出された
- (4) 開いている片目が検出された

(1) 瞳が検出されなかった、(2) 閉じている両目が検出された場合は、ユーザは学習動画を視聴していないとみなす。反対に、(3) 開いている両目が検出された場合は、ユーザは学習動画を視聴しているとみなす。(4) 開いている片目が検出された場合は、片目をつぶって学習動画を視聴する学習者が現実には存在しにくいと考えられる。瞳検出モデルの動作を確認したところ、ユーザの目が髪の毛やメガネなどに遮られるなどして、瞳検出に失敗することがあった。そこで本研究では(4)の場合は、ユーザは学習動画を視聴しているとみなす。

3.4 視聴状況の可視化

ユーザの復習すべき学習動画の時間帯を直観的に示すため、本研究では、学習動画の時間軸に合わせてユーザの視聴状況を可視化する。

時刻 $t=0$ を学習動画視聴開始時刻とする。縦軸に 3 つの値を設け、一番下を(1) 瞳が検出されなかった、または(2) 閉じている両目が検出された、真ん中を(4) 開いている片目が検出された、一番上を(3) 開いている両目が検出されたとする。1/30 秒ごとに動画の終了まで検出結果をプロットする。

表 3: YOLOv5m の再訓練の結果

指標	数値
Bounding Box Regression Loss	0.01533
Objectness Loss	0.004425
Classification Loss	0.02435
Mean Average Precision 50	0.952

続いて、時間軸を幅 T ごとに区切り、 T の範囲内で閉じている両目が検出されたあるいは瞳が検出されなかった割合を算出する。割合が事前に設定した閾値 V 以上であれば、該当の T の範囲を復習すべき時間帯と判断し、点線で囲う。本研究では、 T を 5 秒に、 V を 80% に設定している。

図 3 に視聴状況図の例を示す。青い点が閉じている両目が検出された、あるいは瞳が検出されなかったところで、赤い点が開いている片目が検出されたところ、緑の点が開いている両目が検出されたところになる。灰色の点線で囲われた部分が、ユーザが復習すべき時間帯を示す。

4 提案システムの評価実験

20 名の大学生に被験者の協力を依頼し、提案システムの評価実験を行った。実験目的は、提案システムの動画学習における復習支援の効果を検証することである。

4.1 評価実験の手順

P1) 被験者を実験群と対照群の両群に分ける。

P2) 被験者に学習動画を視聴させる。学習動画を視聴していない状況を再現するため、動画の再生の間に、被験者のスマートフォンに妨害メッセージを送って、返信を要求する。

P3) 動画学習が終わったら、被験者は 3 分間の復習をする。実験群は提案システムを利用し、提供された視聴状況図を参考にしながら復習する。対照群は提案システムを利用せず、個人の記憶に基づいて復習する。

P4) 学習動画に関する事後テストを行う。テストの正解率を用いて、提案システムの評価を行う。

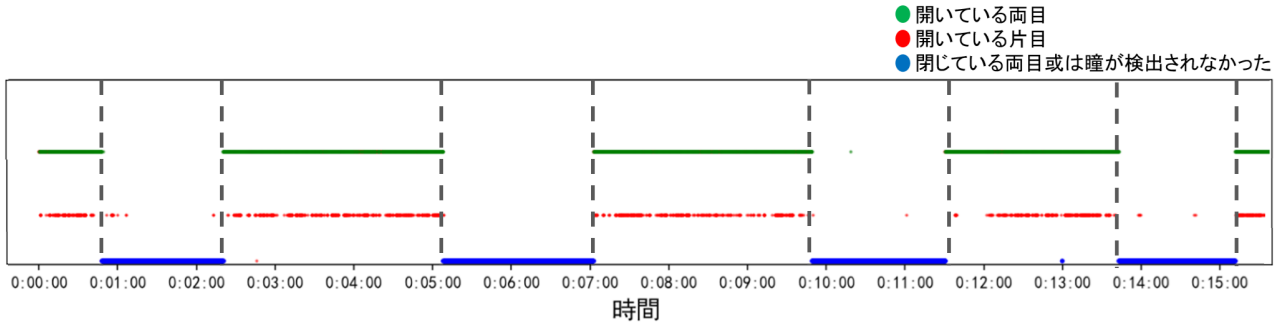


図 3: 視聴状況図の例

表 4: 学習動画のデータ

平均文字数	平均時間	合計文字数	合計時間
128	26 秒	5122	17 分 51 秒

表 5: 妨害メッセージの種類と説明

妨害メッセージ	説明
自己紹介	自己紹介文をスマホで返信させる
動画視聴	送った動画をスマホで視聴させる

4.2 実験手順の詳細

評価実験に用いた学習動画は第一著者が作成したもので、内容は「動物おもしろ雑学集」²と「ほんよま」³2つのウェブサイトを参考にして作成した「動物の雑学40選」であった。それぞれの雑学を説明する際の平均文字数と平均時間、学習動画の合計文字数と合計時間を表4に示す。

評価実験に用いた学習動画は18分程度であったため、最後まで集中力を保てる被験者が居た可能性が高い。そこで、意図的に被験者の学習動画を視聴しない状況を作る必要がある。本実験で採用した方法は、「学習動画の再生の間に、被験者のスマホに妨害メッセージを送って、返信を要求する」であった。妨害メッセージの種類と説明は表5に示す。妨害メッセージは自己紹介文を要求するものと、学習に関係ない動画視聴を要求するものの2つであった。

評価実験の最後に、被験者に学習動画の内容について、選択肢問題40個を出題した。各被験者の事後テストの正解率を以下の3種類に分類した。

(1) FCR: 学習動画を視聴している時間帯の問題の

表6: 両群の各正解率 (FCR: 学習動画を視聴している時間帯の問題の正解率, UCR: 学習動画を視聴していない時間帯の問題の正解率, TCR: 全体正解率, * $p < 0.05$, ** $p < 0.01$)

	FCR	UCR	TCR
実験群	90.2%	93.7%**	92.8%*
対照群	92.3%	67.0%	81.9%

正解率

(2) UCR: 学習動画を視聴していない時間帯の問題の正解率

(3) TCR: 全体正解率

表6に両群の各正解率の数値を示す。

5 考察

両群のFCR (学習動画を視聴している時間帯の正解率) について、実験群は90.2%、対照群は92.3%、両群の平均値の差は2.1%であった ($t = 0.476$, $p = 0.643 > 0.05$)。これは、学習動画を視聴している時、両群の被験者は同等の学習を行なっていることを示している。本研究では、両群の被験者の学習能力は同等であると考えられる。

しかし、学習動画の視聴が妨害されると、両群の被験者の学習成果に差が現れた。両群のUCR (学習動画を視聴していない時間帯の正解率) について、実験群は93.7%、対照群は67.0%、両群の平均値の差は26.7%であった ($t = -5.032$, $p = 0.002 < 0.05$)。この結果から、提案システムは被験者の学習動画を視聴していない時間帯の復習を支援できたことが分かった。

両群のTCR (全体正解率) について、実験群は92.8%、対照群は81.9%、両群の平均値の差は10.9%であった ($t = -2.811$, $p = 0.01 < 0.05$)。この結果から、提案システムを利用した実験群の被験者は動画学習が効果的に行えることが確認された。

²<https://zooing.honpo21.net/> 2024年11月29日アクセス確認。

³<https://sanctuarybooks.jp/webmag/> 2024年11月29日アクセス確認。

この3つの結果から、提案システムはユーザの動画学習の復習に支援を提供することで、学習動画を視聴していない時間の復習が効果的に行え、学習成果を向上させることが確認された。

6 おわりに

本研究では、オンライン学習の一種である動画学習の復習効率問題を注目し、瞳検出技術を用いた動画学習中の視聴状況を表示する学習支援システムを提案した。提案システムは瞳検出を実装することで、ユーザの学習動画への視聴状況を計測し、学習動画の再生時間に合わせる可視化図をユーザに提供する。視聴状況の可視化図を参照にすることにより、ユーザが効率的に復習でき、動画学習の効果を向上させると期待される。

提案システムの有用性を確認するため、20名の被験者を募集し、実際の動画学習を模倣した評価実験を行った。実験の結果として、実験群の学習動画を視聴していない時間帯の問題の正解率は対照群より26.7%高くなった。この結果から、提案システムの利用により、ユーザは限られた時間に学習動画の視聴していない部分をより効率的に復習できることが確認された。

今後の課題として、大学の講義のように長い時間での動画学習における提案システムの効果を検証することが考えられる。

参考文献

- [1] Cathy Mae Toquero: Challenges and opportunities for higher education amid the COVID-19 pandemic: The Philippine context, *Pedagogical Research*, Vol. 5, No. 4, pp. 1–5 (2020)
- [2] Razzaqul Ahshan: A Framework of Implementing Strategies for Active Student Engagement in Remote/Online Teaching and Learning during the COVID-19 Pandemic, *Education Sciences*, Vol. 11, No. 9, pp. 1–24 (2021)
- [3] Murat Ertan Dogan, Tulay Goru Dogan, Aras Bozkurt: The Use of Artificial Intelligence (AI) in Online Learning and Distance Education Processes: A Systematic Review of Empirical Studies, *Applied Sciences*, Vol. 13, No. 5, pp. 3056 (2023)
- [4] Samad, Albertus Laurensius Setyabudhi: Application of Video-Based Learning Media with Applications to Improve Learning Activities of Vocational High School Students, *Technical and Vocational Education International Journal (TAVEIJ)*, Vol. 3, No. 1, pp. 22–27 (2023)
- [5] Samy Cheikh Youssef, Abdullatif Aydin, Alexander Canning, Nawal Khan, Kamran Ahmed, Prokar Dasgupta: Learning Surgical Skills Through Video-Based Education: A Systematic Review, *Surgical Innovation*, Vol. 30, No. 2, pp. 220–238 (2023)
- [6] Jon-Chao Hong, Yue Liu, Yinsheng Liu, Li Zhao: High School Students' Online Learning Ineffectiveness in Experimental Courses During the COVID-19 Pandemic, *Frontiers in Psychology*, Vol. 12, pp. 738695 (2021)
- [7] Ridho Bramulya Ikhsan, Listya Ayu Saraswati, Brian Garda Muchardie, Vional, Andrianto Susilo: The Determinants of Students' Perceived Learning Outcomes and Satisfaction in BINUS Online Learning, *2019 5th International Conference on New Media Studies*, pp. 68–73 (2019)
- [8] ChunHsiu Yeh, Wei-Cheng Shen, Hung-Yu Chi, Chin-En Lin, Jong-Shin Chen: Enhancing Online Learning Monitoring with Novel Image Recognition Method Using Dlib for Eye Feature Detection, *2023 12th International Conference on Awareness Science and Technology*, pp. 340–345 (2023)
- [9] Rasheed Abdulkader, Firas Tayseer Mohammad Ayasrah, Venkata Ramana Gupta Nallagattla, Kamal Kant Hiran, Pankaj Dadheech, Vivekanandam Balasubramaniam, Sudhakar Sengan: Optimizing student engagement in edge-based online learning with advanced analytics, *Array*, Vol. 19, pp. 100301 (2023)
- [10] Mirko Raca, P.Dillenbourg: System for assessing classroom attention, *Proceedings of the 3rd International Learning Analytics and Knowledge Conference*, pp. 265–269 (2013)
- [11] Abdulrahman Renawi, Fady Alnajjar, Medha Parambil, Zouheir Trabelsi, Munkhjargal Gochoo, Sumaya Khalid, Omar Mubin: A simplified real-time camera-based attention assessment system for classrooms: pilot study, *Education and Information Technologies*, Vol. 27, pp. 4753–4770 (2022)