

A TECHNIQUE OF ASSESSING STUDENT ENGAGEMENT BASED ON EYE CLOSE/OPEN BEHAVIOR

Tran Thanh Phuong¹, Tran Van Lang¹, Do Nang Toan^{2*}

¹Lac Hong University, ²Vietnam National University, Hanoi

ARTICLE INFO	ABSTRACT
Received: 21/4/2021 Revised: 31/5/2021 Published: 31/5/2021	Automated learning analysis is becoming an essential topic in the educational area, which needs systems to monitor the learning process and to measure the students engagement to help for teachers more rational adjustments in teaching strategies and learning environment. This paper proposes a technique to estimate the degree of sleepiness base on open/ close state of the eye to assess student concentration during in class. Input data for algorithm is captured by the camera in class and the evaluation analysis is performed on each student during learning. The technique has been setted up and effective for medium-sized theoretical classes in engineering disciplines.
KEYWORDS	
Student engagement Eye-Blink Detection Drowsiness Detection Engagement Recognition Human Behavior Analysis	

MỘT KỸ THUẬT ĐÁNH GIÁ SỰ TẬP TRUNG CỦA SINH VIÊN DỰA TRÊN TRẠNG THÁI ĐÓNG/MỞ CỦA MẮT

Trần Thanh Phương¹, Trần Văn Lăng¹, Đỗ Năng Toàn^{2*}

¹Trường Đại học Lạc Hồng, ²Đại học Quốc gia Hà Nội

THÔNG TIN BÀI BÁO	TÓM TẮT
Ngày nhận bài: 21/4/2021 Ngày hoàn thiện: 31/5/2021 Ngày đăng: 31/5/2021	Phân tích học tập tự động đang trở thành một chủ đề thiết yếu trong lĩnh vực giáo dục, trong đó cần các hệ thống giám sát quá trình học tập và đo lường mức độ tham gia của sinh viên để giúp cho giáo viên có những điều chỉnh hợp lý hơn về chiến lược giảng dạy và môi trường học tập. Bài báo này đề xuất một kỹ thuật ước lượng mức độ buồn ngủ dựa trên trạng thái đóng/mở của mắt để đánh giá sự tập trung của sinh viên trong giờ học. Dữ liệu đầu vào của thuật toán được ghi nhận bởi camera trong lớp học và việc phân tích đánh giá được thực hiện trên từng sinh viên trong suốt quá trình học tập. Kỹ thuật đã được cài đặt thử nghiệm và tỏ ra hiệu quả đối với các lớp học lý thuyết với quy mô trung bình của các ngành học kỹ thuật.
TỪ KHÓA	
Sự tham gia của sinh viên Phát hiện đóng mở mắt Phát hiện ngủ gật Nhận dạng sự chú ý Phân tích hành vi	

DOI: <https://doi.org/10.34238/tnu-jst.4406>

* Corresponding author. Email: donangtoan@gmail.com

1. Giới thiệu

Phân tích sự tập trung trong học tập hiện đang trở thành một chủ đề quan trọng trong cộng đồng giáo dục, trong đó cần có các hệ thống hiệu quả để theo dõi quá trình học tập và cung cấp phản hồi cho giáo viên. Những tiến bộ gần đây trong cảm biến thị giác và phương pháp thị giác máy tính cho phép theo dõi tự động hành vi và trạng thái cảm xúc của người học, trong đó:

(i) **Hành vi** được xác định từ các dấu hiệu như: giờ tay phát biểu, tư thế đầu và tư thế hình thể [1]-[5].

(ii) **Cảm xúc** gồm có: thích thú, mệt mỏi, bối rối,... được xác định từ biểu hiện trên khuôn mặt [6]-[8].

Các kết quả nghiên cứu cho thấy, vấn đề phân tích sự tập trung của sinh viên khi tham gia học tập rất đa dạng và được nhiều nhà nghiên cứu quan tâm và chủ yếu tập trung vào phân tích hành vi và cảm xúc, như: phân tích biểu hiện cảm xúc trên khuôn mặt, phân tích chuyển động mắt, phân tích chuyển động đầu, tư thế hình thể v.v.. Những nghiên cứu liên quan đến hành vi, có thể kể đến, Divjak và Bischof [9] đã phân tích dựa vào hành vi chuyển động đầu và mắt, Bidwell và Fuchs [10] đã đo lường mức độ tương tác của sinh viên bằng hệ thống quan sát tự động hành vi của mắt, họ đề xuất đưa ra 8 danh mục hành vi riêng biệt, nhưng chỉ có thể phân loại sinh viên là "tham gia" hoặc "không tham gia". Krithika [11] sử dụng chuyển động của mắt và đầu để kiểm tra sự tập trung của học sinh và tạo ra cảnh báo về mức độ tập trung thấp.

Những nghiên cứu liên quan đến cảm xúc, có thể kể đến, Turabzadeh và cộng sự [12] dựa trên nhận dạng cảm xúc khuôn mặt trong thời gian thực, sử dụng thuật toán mẫu nhị phân cục bộ (Local Binary Point - LBP), Kamath, Biswas và Balasubramania [13] sử dụng thuật toán phát hiện khuôn mặt Vilola Jones để phân tích hình ảnh đầu vào và sau đó là biểu đồ phân định hướng (HOG, Sharma và cộng sự đã đề xuất một hệ thống thời gian thực, dựa trên những biểu hiện cảm xúc trên khuôn mặt, để kiểm tra sự tham gia của học sinh trong bối cảnh học tập điện tử, tự động điều chỉnh nội dung theo mức độ tập trung của học sinh, bằng cách phân tích cảm xúc của học sinh và đưa ra ba mức độ tập trung khác nhau (cao, trung bình và thấp) [14].

Trong bài báo này, chúng tôi đề xuất một kỹ thuật đánh giá sự tập trung của sinh viên trong giờ học, là sự kết hợp của cả việc phân tích hành vi và cảm xúc với một trạng thái cảm xúc đặc biệt là "buồn ngủ", nhờ ước lượng mức độ buồn ngủ dựa trên trạng thái đóng/mở của mắt. Phần còn lại của bài báo được thể hiện như sau: Phần 2 trình bày một số nghiên cứu liên quan. Phần 3 là kỹ thuật đề xuất. Tiếp theo là thử nghiệm và cuối cùng là kết luận về kỹ thuật đề xuất.

2. Cơ sở lý thuyết của thuật toán

Mức độ buồn ngủ được ước lượng dựa trên các yếu tố khác nhau, chẳng hạn như tỷ lệ khung hình và thời lượng đóng của mắt. Ứng với mỗi mức độ sẽ có những biểu hiện, hành vi và thời lượng mắt đóng tương ứng. Ngoài ra, yếu tố mắt đóng nhanh hay chậm cũng được chúng tôi quan tâm xem xét. Đối với người thức giấc/tỉnh táo thì mắt đóng và mở rất nhanh. Đối với người trong tình trạng mệt mỏi, buồn ngủ thì mắt có xu hướng đóng chậm [15].



Hình 1. Thời gian mắt mở đến khi mắt đóng hoàn toàn

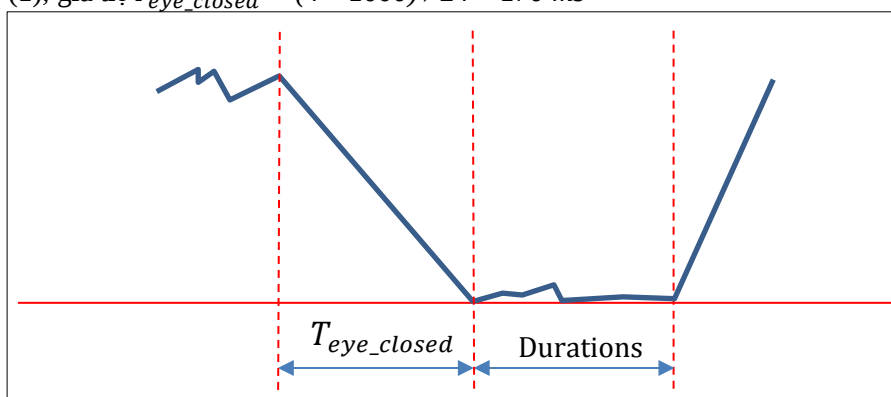
Trong hình 1, dựa trên dữ liệu Talking Face¹ với máy ảnh ghi hình là 24fps, trạng thái mắt đóng hoàn toàn xảy ra tại frame thứ tư. Do đó thời gian đóng mắt hoàn toàn (hình 2) là:

$$T_{eye_closed} = \frac{N_{frame} * 1000}{fps} \quad (1)$$

Trong đó:

- N_{frame} là số lượng khung hình tính từ khung có trạng thái mắt mở đến khung có trạng thái mắt đóng hoàn toàn (trong trường hợp này là 4).
- fps (frames per second) là số lượng khung hình trên mỗi giây, giá trị này thay đổi tùy vào máy ghi hình (trong trường hợp này là 24).

Nên theo (1), giá trị $T_{eye_closed} = (4 * 1000) / 24 \approx 170 \text{ ms}$



Hình 2. Minh họa thời gian mắt mở đến khi mắt đóng hoàn toàn

Trước khi thực hiện ước lượng mức độ buồn ngủ, bước tiền xử lý của chúng tôi là loại bỏ những đoạn video có mắt đóng nhanh ($T_{eye_closed} \leq 170\text{ms}$) xem công thức (1) và thời lượng mắt đóng ngắn (duration < 1s), những đoạn này được xem xét là tập trung. Chúng tôi chỉ quan tâm đến những đoạn video có mắt đóng chậm ($T_{eye_closed} > 170\text{ms}$) và thời lượng mắt đóng dài (duration $\geq 1\text{s}$) (xem bảng 1).

Bảng 1. Sự tương quan giữa mức độ buồn ngủ và hành vi [15]

Mức độ
Thức giấc/Tỉnh táo (Không buồn ngủ): Thời lượng mắt đóng <1s Người thức giấc/tỉnh táo có các hành vi và biểu cảm như: Sắc mặt bình thường; đóng/mở mắt nhanh; các cử động/cử chỉ của cơ thể xảy ra không thường xuyên.
Hơi buồn ngủ: Thời lượng mắt đóng 1-2s Khi người trở nên hơi buồn ngủ, các hành vi khác nhau có thể được thể hiện. Những hành vi này được gọi là cách cư xử, có thể bao gồm dụi mắt hoặc mắt, gãi, sắc mặt giảm và di chuyển không yên trên ghế. Những hành động này có thể được coi là biện pháp đối phó với cơn buồn ngủ. Chúng xảy ra trong các giai đoạn trung gian của buồn ngủ. Không phải tất cả các cá nhân đều thể hiện cách cư xử trong các giai đoạn trung gian. Một số cá nhân có vẻ trầm lắng hơn, đóng mắt chậm hơn, sắc mặt của họ có thể giảm, thiếu hoạt động rõ ràng, họ có thể nhìn chằm chằm vào một vị trí cố định.
Ngủ gật: Thời lượng mắt đóng 2-3s hoặc lâu hơn Khi người trở nên ngủ gật, thường xảy ra tình trạng mí mắt khép lâu hơn. Điều này thường đi kèm với sự di chuyển của mắt lên hoặc sang một bên. Mắt không tập trung thích hợp, thiếu linh hoạt, sắc mặt giảm, thiếu hoạt động rõ ràng, có thể có các chuyển động bị cô lập (hoặc ngắt quãng) lớn.
Cực kỳ buồn ngủ: Thời lượng mắt đóng $\geq 4\text{s}$ Người trong tình trạng này thường có biểu hiện mí mắt đóng kéo dài (4 giây trở lên) và thời gian thiếu hoạt động kéo dài tương tự. Có thể có những chuyển động bị ngắt quãng lớn.

Dựa vào mức độ và hành vi ở bảng 1, chúng tôi xây dựng luật xác định mức độ buồn ngủ như sau:

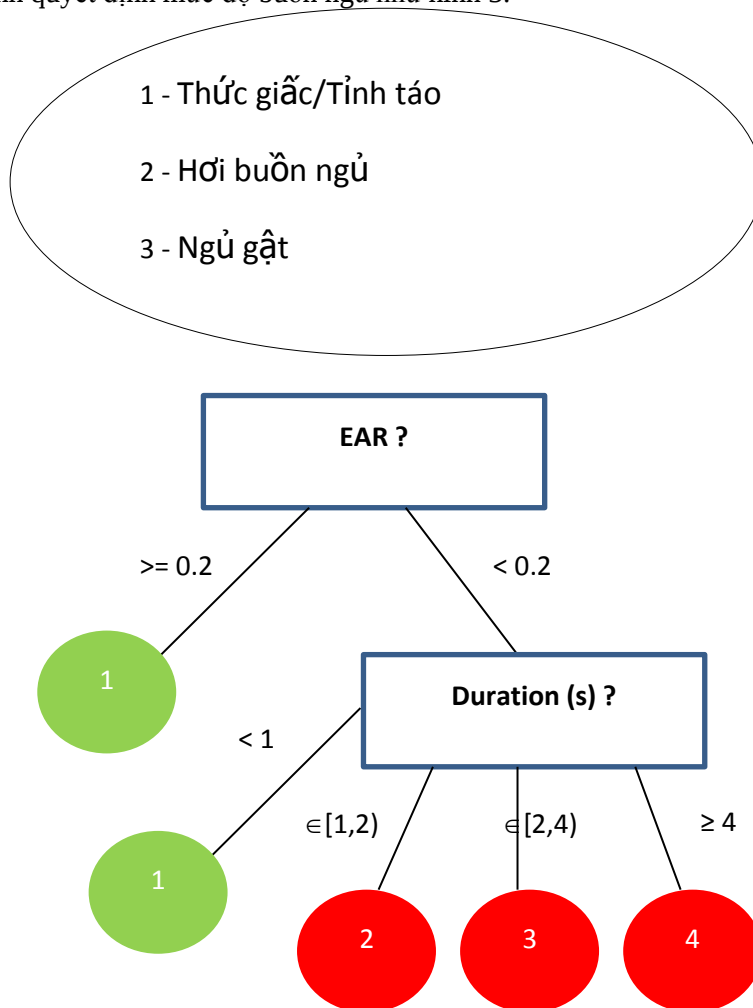
¹ https://personalpages.manchester.ac.uk/staff/timothy.f.cootes/data/talking_face/talking_face.html

Bảng 2. Luật xác định mức độ buồn ngủ

Mức độ	Điều kiện
Thức giấc/Tỉnh táo	$EAR > 0.2$
Hơi buồn ngủ	$EAR < 0.2$ AND $Duration \in [1s, 2s)$
Ngủ gật	$EAR < 0.2$ AND $Duration \in [2s, 4s)$
Cực kỳ buồn ngủ	$EAR < 0.2$ AND $Duration \geq 4s$

Trong đó, EAR là mức độ mở của mắt được tính tỷ lệ chiều rộng và chiều dài của hình chữ nhật chứa mắt.

Áp dụng luật bảng 2 chúng tôi định nghĩa 4 lớp thể hiện mức độ buồn ngủ theo cấp độ từ 1 – 4 và hai thuộc tính cần xem xét là **EAR** và **Duration** (s). Dựa vào thông tin trên, chúng tôi xây dựng mô hình quyết định mức độ buồn ngủ như hình 3.

**Hình 3.** Mô hình ước lượng mức độ buồn ngủ

3. Kỹ thuật đánh giá sự tập trung

3.1. Đánh giá sự tập trung dựa vào mức độ buồn ngủ

Dựa vào bảng 3, chúng tôi xây dựng thang độ đánh giá sự tập trung cho hệ thống của chúng tôi gồm 3 mức độ.

Bảng 3. Mức độ tập trung

ID	Mức độ buồn ngủ	Mức độ tập trung
1	Thức giấc/Tỉnh táo	Tập trung cao
2	Hơi buồn ngủ	Tập trung thấp
3	Ngủ gật Cực kỳ buồn ngủ	Hoàn toàn mất tập trung

Do đây là hệ thống phân tích đánh giá sự tập trung của sinh viên, nên nhiệm vụ cần được thực hiện cả một quá trình, có nghĩa là hệ thống phải xem xét và đánh giá sinh viên từ lúc bắt đầu video cho đến kết thúc video. Không giống như các hệ thống khác như phát hiện ngủ gật trong lúc lái xe chẳng hạn, khi phát hiện tài xế mất tập trung, hệ thống cảnh báo và đưa ra kết luận tại thời điểm đó. Sở dĩ chúng tôi đặt ra điều kiện này bởi các lý do:

(i) Giả sử tại khung hình $frame_i$ sinh viên có thể mất tập trung nhưng sang đến khung hình $frame_k$ thì sinh viên tập trung trở lại. Do đó, đánh giá sự tập trung của sinh viên dựa vào khung hình hiện tại là không phù hợp

(ii) Trong quá trình học tập, sinh viên có thể có nhiều mức độ tập trung ở những thời điểm khác nhau. Vì vậy, dựa vào kết quả phân tích hiện tại để đánh giá cả quá trình học tập là không phù hợp.

(iii) Sinh viên tập trung trong suốt quá trình học tập, nhưng tại khung hình cuối cùng có hành vi chớp mắt, hoặc trường hợp ngược lại. Do đó, hệ thống đánh giá dựa vào khung hình cuối cùng là thiếu chính xác.

Chính vì vậy, chúng tôi đề xuất mô hình đánh giá sự tập trung bằng cách tính tổng thời gian của mỗi trạng thái (xem bảng 2), sau đó lập tỷ lệ mỗi trạng thái so với tổng thời lượng của video. Trạng thái nào chiếm tỷ lệ cao nhất thì sẽ quyết định mức độ tham gia của người học trong suốt quá trình học tập. Đây là phương pháp đề xuất của chúng tôi mang tính chủ quan, tuy nhiên kết quả từ hệ thống đưa ra rất phù hợp với những gì chúng tôi mong muốn đạt được.

Cách đánh giá được trình bày chi tiết như sau:

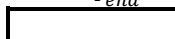
Giả sử quá trình tham gia học tập của sinh viên được mô phỏng như hình 4, gồm có nhiều trạng thái và xảy ra ở những thời điểm khác nhau:

**Hình 4. Mô phỏng trạng thái của quá trình học tập**

Trong đó

- T_{begin} : Thời gian bắt đầu học

- T_{end} : Thời gian kết thúc học

 1 - Thức giấc/Tỉnh táo

 2 - Hơi buồn ngủ

 3 - Ngủ gật

 4 - Cực kỳ buồn ngủ

Gọi t_1 là tổng thời gian của mức độ 1 và t_2, t_3, t_4 là tổng thời gian tương ứng cho các mức độ còn lại.

Gọi r_1, r_2, r_3, r_4 là các tỷ lệ tương ứng của các mức độ t_1, t_2, t_3, t_4 .

Tỷ lệ của các mức độ được xác định:

$$\begin{aligned}
 r_1 &= \frac{t_1}{T_{begin} - T_{end}} & r_2 &= \frac{t_2}{T_{begin} - T_{end}} \\
 r_3 &= \frac{t_3}{T_{begin} - T_{end}} & r_4 &= \frac{t_4}{T_{begin} - T_{end}}
 \end{aligned} \tag{2}$$

Mức độ tập trung được xác định như sau

$$x = \text{MAX}(r_1, r_2, r_3, r_4) = \begin{cases} \text{Tập trung cao, } x = r_1 \\ \text{Tập trung thấp, } x = r_2 \\ \text{Hoàn toàn mất tập trung, } x = r_3 | r_4 \end{cases} \tag{3}$$

3.2. Thuật toán đánh giá sự tập trung

Algorithm 1: đánh giá mức độ tập trung của sinh viên

Input: DATA – Video sinh viên tham gia trong lớp học

Output: LIST_STUDENT – Danh sách đánh giá mức độ tập trung của sinh viên

Khởi tạo giá trị mảng lưu thời gian đóng/mở mắt của từng sinh viên: EBD $\leftarrow \{\}$

while not at end of DATA **do** // đọc dữ liệu

frame $\leftarrow$ DATA.get_frame() // trích xuất khung hình

faces $\leftarrow$ detect_face(frame) // phát hiện các khuôn mặt trong khung hình

for face_id **in** faces

blink_duration $\leftarrow$ detect_blink(face_id, pre_blink) // phát hiện đóng mở mắt

// cập nhật thời gian của các mức độ buồn ngủ khi xảy ra đóng mắt

if $1 \leq \text{blink_duration} < 2$ **then**

EBD_{face_id}.dg_2 += blink_duration // hơi buồn ngủ

else if $2 \leq \text{blink_duration} < 4$ **then**

EBD_{face_id}.dg_3 += blink_duration // ngủ gật

else if $\text{blink_duration} \geq 4$ **then**

EBD_{face_id}.dg_4 += blink_duration // cực kỳ buồn ngủ

end if

end for

end while

for each data **in** EBD

dg_2 $\leftarrow$ data_{face_id}.dg_2 // tổng thời gian hơi buồn ngủ

dg_3 $\leftarrow$ data_{face_id}.dg_3 // tổng thời gian ngủ gật

dg_4 $\leftarrow$ data_{face_id}.dg_4 // tổng thời gian cực kỳ buồn ngủ

dg_1 = TOTAL_TIME_VIDEO – (dg_2 + dg_3 + dg_4) // bình thường

degree $\leftarrow$ MAX (dg_1, dg_2,
dg_3 + dg_4) // sự tập trung
dựa trên mức độ buồn ngủ cao
nhất

degree $\leftarrow$ LIST_STUDENT [data.face_id]

end for

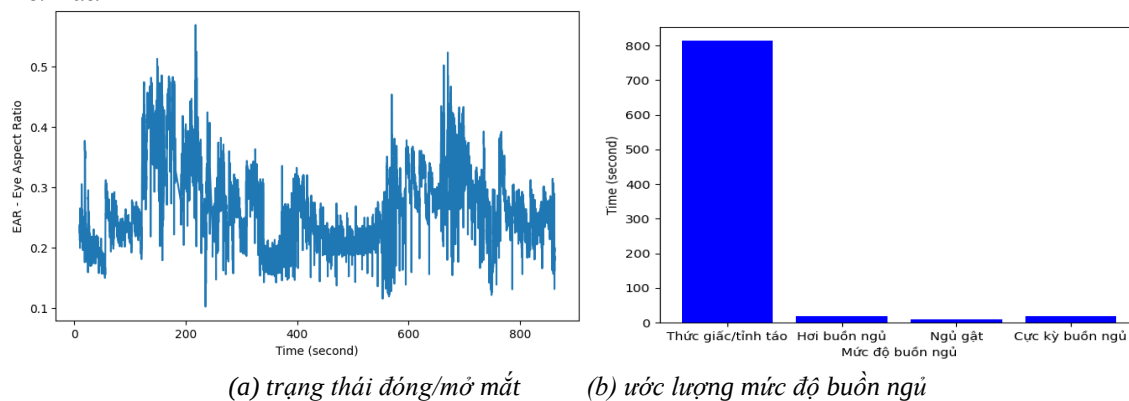
*) Độ phức tạp của thuật toán

- Chi phí xác định thời gian đóng/mở mắt là $T_1 = O(n^2)$
- Chi phí xác định sự tập trung là : $T_2 = O(n)$
- Vậy tổng chi phí thực hiện là: $T = T_1 + T_2 = O(n^2)$

4. Thử nghiệm

Chúng tôi thực nghiệm kiểm tra các mục tiêu chúng tôi đã đặt ra như: ước lượng mức độ buồn ngủ dựa trên hành vi đóng/mở mắt và cuối cùng là đánh giá sự tập trung dựa trên mức độ buồn

ngủ. Video quay giờ học của một lớp sinh viên ngành kỹ thuật tại trường Đại học Lạc Hồng, tổng độ dài của video là 862.64s. Hình 5 là kết quả ước lượng mức độ buồn ngủ dựa trên hành vi đóng mở/mắt.



Hình 5. Kết quả ước lượng mức độ buồn ngủ của sinh viên

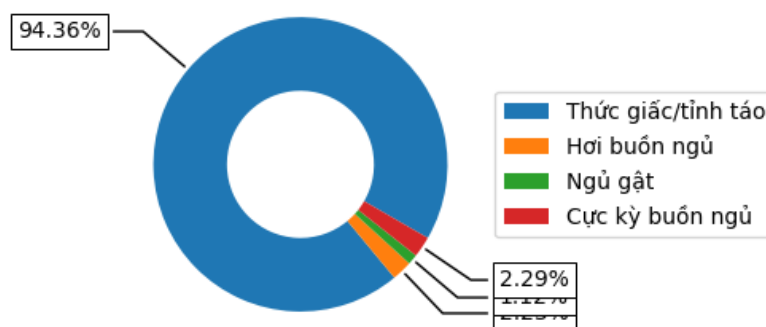
Thử nghiệm trong trường hợp này, chúng tôi đánh giá người học với tư thế ngồi, hướng đầu thẳng và hướng nhìn về phía giáo viên.

Trong đó các chỉ số được hệ thống đo chi tiết từ hình 6 như sau:

- Thức giấc/tỉnh táo: 813,99s chiếm 94,36%
- Hơi buồn ngủ: 19,22s chiếm 2,23%
- Ngủ gật: 9,65s chiếm 1,12%
- Cực kỳ buồn ngủ: 19,78s chiếm 2,29%

Giá trị cao nhất của các mức độ là 94,36% nên dựa vào bảng 3 hệ thống kết luận đối tượng tập trung cao trong giờ học xem hình 6.

Đánh giá: Tập trung cao



Hình 6. Kết quả đánh giá sự tập trung của sinh viên

5. Kết luận

Đánh giá chất lượng dạy và học luôn là vấn đề đặt ra trong đào tạo. Việc đánh giá chất lượng giảng dạy, một cách gián tiếp có thể thông qua việc đánh giá sự tập trung của người học, nhờ đó có thể giúp cho giáo viên có những điều chỉnh hợp lý để tăng hiệu quả giảng dạy của mình.

Trong bài báo này chúng tôi đề xuất một kỹ thuật đánh giá sự tập trung của sinh viên trong giờ học nhờ ước lượng mức độ buồn ngủ dựa trên trạng thái đóng/mở của mắt. Kỹ thuật đã được cài đặt thử nghiệm và tỏ ra hiệu quả đối với các lớp học lý thuyết với quy mô trung bình của các ngành học kỹ thuật, với người học với tư thế ngồi, đầu thẳng và hướng nhìn về phía bục giảng.

Trong các nghiên cứu tiếp theo, các hành vi liên quan đến chuyển động gât, lắc của đầu sẽ được chúng tôi quan tâm để làm tăng tính chính xác trong việc đánh giá sự tập trung của người học.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] D. Dinesh, A. Narayanan, and K. Bijlani, "Student analytics for productive teaching/learning," *International Conference on Information Science (ICIS), Kochi, India*, 2016, pp. 97-102.
- [2] A. S. Won, J. N. Bailenson, and J. H. Janssen, "Automatic detection of nonverbal behavior predicts learning in dyadic interactions," *IEEE Trans. Affect Comput*, vol. 5, no. 2, pp. 112-125, 2014.
- [3] P. Goldberg, Ö. Sümer, K. Stürmer, W. Wagner, R. Göllner, P. Gerjets, E. Kasneci, and U. Trautwein, "Attentive or Not? Toward a Machine Learning Approach to Assessing Students' Visible Engagement in Classroom Instruction," *Springer*, vol. 33, pp. 27-49, 2021.
- [4] X. Ma, M. Xu, Y. Dong, and Z. Sun, "Automatic Student Engagement in Online Learning Environment Based on Neural Turing Machine," *International Journal of Information and Education Technology*, vol. 11, no. 3, pp. 107-111, 2021.
- [5] T. S. Ashwin and R. M. R. Guddeti, "Unobtrusive Behavioral Analysis of Students in Classroom Environment Using Non-Verbal Cues," *IEEE Access*, vol. 7, pp. 1-7, 2019.
- [6] N. J. Butko, G. Theodorou, M. Philipose, and J. R. Movellan, "Automated facial affect analysis for one-on-one tutoring applications," In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference On*, 2011, pp. 382-387.
- [7] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan, "The faces of engagement: Automatic recognition of student engagement from facial expressions," *IEEE Trans. Affect. Comput*, vol. 5, no. 1, pp. 86-98, 2014.
- [8] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Trans. Affect. Comput*, vol. 1, no. 1, pp. 18-37, 2010.
- [9] M. Divjak and H. Bischof, "Eye blink based fatigue detection for prevention of computer vision syndrome," *Proceedings of IAPR Conference on Machine Vision Applications (MVA)*, 2009, pp. 350-353.
- [10] J. Bidwell and H. Fuchs, "Classroom analytics: Measuring student engagement with automated gaze tracking," *BehavRes Methods*, 2011.
- [11] L. B. Krithika, "Student emotion recognition system (SERS) for e-learning improvement based on learner concentration metric," *Procedia Computer Science*, vol. 85, pp. 767-776, 2016.
- [12] S. Turabzadeh, H. Meng, R. M. Swash, and M. Pleya, "Facial expression emotion detection for real-time embedded systems," *Technologies*, vol. 6, no. 1, p. 17, 2018.
- [13] A. Kamath, A. Biswas, and V. Balasubramanian, "A Crowdsourced Approach to Student Engagement Recognition in E-learning Environments," *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016, pp. 1-9.
- [14] P. Sharma, M. Esengönül, S. R. Khanal, T. T. Khanal, V. Filipe, and M. J. C. S. Reis, "Student Concentration Evaluation Index in an E-learning Context Using Facial Emotion Analysis," *Technology and Innovation in Learning, TECH-EDU*, 2018, pp. 529-538.
- [15] D. M. Wiegand, J. McClafferty, S. E. McDonald, and R. J. Hanowski, "Development and Evaluation of a Naturalistic Observer Rating of Drowsiness Protocol," *Final Report of the National Surface Transportation Safety Center for Excellence (NSTSCE)*, 2009.