

NGHIÊN CỨU XÂY DỰNG HỆ THỐNG THEO DÕI, GIÁM SÁT TRẠNG THÁI BUỒN NGỦ DỰA TRÊN THỊ GIÁC MÁY TÍNH

Điền Thị Hồng Hà

Trường Đại học Kinh tế - Kỹ thuật công nghiệp, Hà Nội, Việt Nam

ARTICLE INFORMATION TÓM TẮT

Journal: Vinh University
*Journal of Science
Natural Science, Engineering
and Technology*
p-ISSN: 3030-4563
e-ISSN: 3030-4180

Volume: 53

Issue: 3A

***Correspondence:**
dthha@uneti.edu.vn

Received: 26 June 2024

Accepted: 16 August 2024

Published: 20 September 2024

Citation:

*Dien Thi Hong Ha (2024).
Research on developing a
monitoring system for drowsiness
detection based on computer
vision. **Vinh Uni. J. Sci.**
Vol. 53 (3A), pp. 80-90
doi: 10.56824/vujs.2024a075a*

Bài báo này tập trung nghiên cứu xây dựng mô hình hệ thống phát hiện và giám sát trạng thái buồn ngủ khi lái xe dựa trên thị giác máy tính. Sản phẩm có thể ứng dụng trên các phương tiện giao thông hiện nay (như xe ô tô hoặc xe có chức năng tương tự). Nghiên cứu sử dụng một số mô hình phát hiện buồn ngủ mới nhất hiện nay, từ đó thu thập và xử lý dữ liệu, huấn luyện mô hình để đưa ra được các kết quả định lượng. Thông qua đó đánh giá và thực hiện đề xuất cải tiến mô hình cho phù hợp với phương tiện giao thông tại Việt Nam. Kết quả thực nghiệm cho thấy, hệ thống có thể ghi lại hình ảnh thông qua camera, phát hiện khuôn mặt người, dự đoán trạng thái buồn ngủ, sau đó gửi cảnh báo về máy chủ và hiển thị dữ liệu cảnh báo trên giao diện website. Đánh giá cơ bản cho thấy hệ thống hoạt động ổn định và tương đối chính xác theo các yêu cầu thiết kế và lập trình trước đó. Kết quả nghiên cứu này có thể đem sử dụng trên thực tế hoặc dùng làm mô hình thực hành cho sinh viên các ngành trí tuệ nhân tạo và Robot, điều khiển tự động hóa nhằm nâng cao khả năng áp dụng lý thuyết vào giải quyết bài toán trong thực tế.

Từ khóa: Buồn ngủ; giao thông; Robot thông minh; thị giác máy tính; trí tuệ nhân tạo.

1. Giới thiệu

Buồn ngủ (*somnolence*) là một trong những nguyên nhân chính của tai nạn giao thông tại nước ta và các nước trên thế giới hiện nay. Theo đó, buồn ngủ được hiểu là một trạng thái khao khát mãnh liệt được ngủ, hoặc ngủ trong thời gian dài bất thường. Buồn ngủ mang ý nghĩa và nguyên nhân riêng biệt. Có thể là sự ám chỉ đến trạng thái bình thường trước khi ngủ, trạng thái buồn ngủ do rối loạn nhịp sinh học, các điều kiện khi ở trong một trạng thái buồn ngủ do nhịp sinh học rối loạn, hoặc một triệu chứng của các vấn đề sức khỏe khác. Có thể được đi kèm với thờ ơ, suy nhược và thiếu sự nhanh nhẹn tinh thần. Buồn ngủ có thể nguy hiểm khi thực hiện các nhiệm vụ đòi hỏi sự tập trung liên tục, chẳng hạn như lái xe. Khi một người mệt mỏi, có thể gặp phải cơn buồn ngủ. Trong những người bị mất ngủ, thì buồn ngủ có thể tự tiêu tan trong một thời gian ngắn; hiện tượng này được gọi là cơn gió thứ hai và kết quả của chu kỳ bình thường của nhịp sinh học can thiệp vào các quá trình cơ thể thực hiện để chuẩn bị cho bản thân nghỉ ngơi.

OPEN ACCESS

Copyright © 2024. This is an Open Access article distributed under the terms of the [Creative Commons Attribution License \(CC BY NC\)](#), which permits non-commercially to share (copy and redistribute the material in any medium) or adapt (remix, transform, and build upon the material), provided the original work is properly cited.

Hiện nay các hãng sản xuất xe hơi đã và đang thực hiện chú trọng tối đa mức an toàn của khách hàng nhờ ứng dụng các công nghệ hiện đại (như trí tuệ nhân tạo) để giúp khách hàng giảm thiểu các nguy cơ liên quan tới tai nạn giao thông. Một số ví dụ về những hệ thống phát hiện buồn ngủ như: *Driver Alert Control* của hãng Volvo, *Attention Assist* của hãng Mazda, *Driver Drowsiness Detection* của hãng Bosch, *Active Driving Assistant with Attention Assistant* của hãng BMW... Những hệ thống này sẽ giám sát người lái xe (tài xế) và đưa ra cảnh báo cần thiết về tình trạng buồn ngủ của họ. Từ đó người tài xế sẽ xem xét hoặc có biện pháp phù hợp như dừng lại để nghỉ ngơi, giải thoát cơn buồn ngủ. Các giải pháp này có độ chính xác tương đối cao, đã hỗ trợ rất tích cực cho các tài xế khi lái xe và góp phần giảm thiểu tai nạn giao thông gây ra bởi hành động ngủ gật của tài xế. Những hệ thống này tuy có độ chính xác cao, nhưng chỉ dừng lại ở mức theo dõi và cảnh báo cho người sử dụng. Trong trường hợp nếu người tài xế cố tình phớt lờ lời cảnh báo, mọi thứ có thể trở nên vô nghĩa. Đặc biệt, những công ty trong lĩnh vực giao thông vận tải và cho thuê xe, bởi vì phương tiện điều khiển không thuộc sở hữu của người điều khiển phương tiện thì tình trạng phớt lờ này lại càng dễ xảy ra hơn. Hơn nữa, đa số các xe có trọng tải lớn chạy đường dài hiện nay tại Việt Nam đều chưa lắp đặt các hệ thống cảnh báo tình trạng buồn ngủ tối ưu.

Nghiên cứu này thực hiện đề xuất giải pháp cải tiến các thiếu sót trên, đó là sẽ xây dựng một hệ thống phát hiện và cảnh báo buồn ngủ có kết nối điện toán đám mây [1]-[3]. Nếu có cảnh báo, hệ thống sẽ gửi tín hiệu về máy chủ để lưu trữ kết quả. Điều này nếu kết hợp với trách nhiệm pháp lý khi gây ra tai nạn giao thông sẽ phân nào điều chỉnh hành vi của người điều khiển phương tiện khi gặp tình trạng buồn ngủ.

Trong bài báo này sử dụng phương pháp dùng các kỹ thuật phát hiện dựa trên hiện tượng sinh lý của con người để phát hiện trạng thái buồn ngủ và được thực hiện theo hai cách sau [4]-[9]: Cách thứ nhất: Theo dõi, đo lường sự thay đổi trong các tín hiệu sinh lý, chẳng hạn như sóng não, nhịp tim và nháy mắt (intrusive techniques: kỹ thuật xâm nhập). Để áp dụng kỹ thuật này các cảm biến điện sẽ được gắn trực tiếp vào cơ thể của người lái xe, do đó gây khó chịu và mất tập trung. Ngoài ra, lái xe trong thời gian dài sẽ dẫn đến đổ mồ hôi và làm giảm khả năng hoạt động trên các thiết bị cảm biến. Cách thứ hai: Sử dụng Thị giác máy tính để đo lường những thay đổi về cơ thể vật lý (non-intrusive techniques: kỹ thuật không xâm nhập) như tư thế chùng xuống, vị trí nghiêng đầu, trạng thái nhắm/ mở mắt, miệng của người lái xe. Kỹ thuật này là rất phù hợp với điều kiện lái xe thực tế vì không cần tác động trực tiếp vào cơ thể người lái xe mà thông qua máy quay (camera) để phát hiện những thay đổi trên khuôn mặt người. Biện pháp này hứa hẹn là một trong những phương pháp tốt nhất với ưu điểm là việc không có tác động vật lý lên cả người và xe, mang lại sự tiện lợi và hiệu quả. Hướng thiết kế và lập trình điều khiển của sản phẩm bài báo sẽ sử dụng cách thứ hai vừa trình bày. Mục tiêu dự kiến sau khi hoàn thành của nghiên cứu bao gồm: Phát triển một hệ thống đề xuất theo dõi và giám sát (phát hiện) người lái xe buồn ngủ trong thời gian thực. Nghiên cứu đã phát triển một hệ thống đơn giản gồm 5 phần, cụ thể là: (a) Ghi lại hình ảnh camera. (b) Phát hiện khuôn mặt. (c) Dự đoán trạng thái, phát hiện buồn ngủ. (d) Gửi cảnh báo về máy chủ. (e) Hiện thử dữ liệu cảnh báo lên website.

2. Công cụ thực hiện

Để thực hiện thành công nghiên cứu này, nhóm tác giả đã sử dụng các công cụ cốt lõi đó là [2]-[5]: Ngôn ngữ lập trình *Python* dùng để viết nội dung chương trình; Thư viện

Open CV dùng để tận dụng các hàm và thư viện sẵn có (như *Scikit-learn*, *TensorFlow*, *Keras*, *Numpy*) hỗ trợ nhận diện khuôn mặt người trong thị giác máy tính; Hệ điều hành *Linux Ubuntu* dùng để cài đặt chương trình thử nghiệm; *Google Colab* dùng để viết và thực thi đoạn mã python thông qua trình duyệt. Một số thiết bị phần cứng được sử dụng kết hợp đó là: *Camera* dùng để quay video tài xế; Máy tính nhúng *Jetson Nano* được thiết kế dành riêng cho các nhà phát triển, sản xuất và nghiên cứu sản phẩm ứng dụng trí tuệ nhân tạo.

3. Xây dựng hệ thống

3.1. Thu thập dữ liệu

Đối với một bài toán sử dụng Deep Learning, dữ liệu đầu vào của bài toán là vô cùng quan trọng. Khi ta có một dữ liệu tốt, đủ nhiều, bao gồm hết những trường hợp mà model sẽ gặp phải trong thực tế, thì model Deep Learning này sẽ hoạt động rất tốt trong thực tế ngược lại, không thể mong muốn model hoạt động với một độ chính xác cao trong khi bản thân dữ liệu đầu vào rất tệ. Vì vậy, khi giải quyết những bài toán sử dụng Deep Learning, các kỹ sư có kinh nghiệm đều thu thập, xử lý, điều chỉnh dữ liệu một cách kỹ càng trước khi huấn luyện mô hình [6]-[8].

Trạng thái buồn ngủ của con người thực chất không phải là một khoảnh khắc, mà là cả một giai đoạn. Ở đó con người sẽ đi từ trạng thái không tỉnh táo, buồn ngủ, tiền ngủ rồi mới đến ngủ hẳn. Phân tích xem con người có đang ở trạng thái buồn ngủ hơn là một bài toán phức tạp, ngay cả với một con người bởi vì nó có rất nhiều biểu hiện và trường hợp khác nhau. Hiện tại, vì đề tài đang sử dụng bài toán phân loại ảnh, tức là hệ thống sẽ chụp từng khoảnh khắc, liên kết những khoảnh khắc này lại với nhau từ đó đưa ra dự đoán rằng một người có buồn ngủ hay không.

Trong nghiên cứu này có tham khảo tới bộ dữ liệu sau đây:

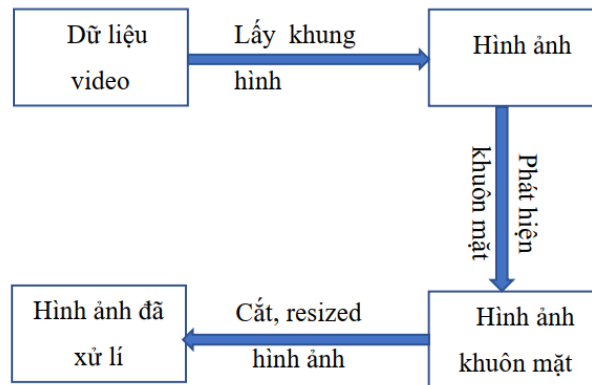
- Bộ dữ liệu về cơn buồn ngủ trong đời thực của Đại học Texas tại Arlington (**UTA Real-Time Drowsiness Datasets**) được tạo ra cho nhiệm vụ phát hiện buồn ngủ nhiều giai đoạn, nhằm mục tiêu không chỉ các trường hợp cực đoan và dễ nhìn thấy, mà còn cả các trường hợp tinh vi khi các biểu hiện vi mô tế nhị là yếu tố phân biệt. Tập dữ liệu của UTA-RLDD là tập dữ liệu buồn ngủ thực tế lớn nhất cho đến nay. Bộ dữ liệu này bao gồm khoảng 30 giờ video RGB của 60 người tham gia khỏe mạnh.

- Bộ dữ liệu **Driver Drowsiness Detection Datasets (DDDD)** sử dụng tập dữ liệu video về tình trạng buồn ngủ của người lái xe do Phòng thí nghiệm Thị giác Máy tính NTHU thu thập.

- **Bộ dữ liệu mô phỏng cho người Việt:** UTA-RLDD và DDDD đều là hai nguồn dữ liệu vô cùng quý giá và đầy đủ, bao hàm hết những trường hợp trong thực tế, gồm rất nhiều sắc tộc. Tuy nhiên, phân phối sắc tộc của 2 nguồn dữ liệu trên có ít người châu Á (với đặc điểm mắt nhỏ hơn) và đặc biệt là người Việt Nam. Vậy nên, nghiên cứu đã xây dựng riêng một bộ dữ liệu của những người Việt Nam (dựa trên kỹ thuật của hai bộ dữ liệu trên) để hệ thống có thể phù hợp hơn với đặc điểm của người Việt và có thể sử dụng cho người Việt.

3.2. Xử lý dữ liệu

Những dữ liệu đã thu thập được đều là dữ liệu dạng video, vì thế cần phải có những công đoạn xử lý dữ liệu thô trở thành dữ liệu đầu vào cho bài toán phân loại [9]-[10]. Các công đoạn xử lý video được tóm tắt theo sơ đồ như trong Hình 1 dưới đây:



Hình 1: Quá trình xử lý dữ liệu video

Trong đó:

- **Lấy khung hình:** Video thực chất là một chuỗi liên tiếp các hình ảnh liên tiếp nhau, được hiển thị một cách nối tiếp nhau. Mỗi video có một tốc độ khung hình nhất định gọi là FPS (Frame per second), tốc độ này đại diện cho số lượng khung hình xuất hiện trên một giây. Tốc độ khung hình càng cao thì độ mượt của video càng tốt. Thông thường FPS phải lớn hơn 25 thì mắt con người sẽ không thấy bị giật. Bởi vì bài toán sử dụng hình ảnh, vì vậy ta cần phải lấy hình ảnh tài xế từ những video trong nguồn dữ liệu. Việc lấy khung hình được thực hiện qua thư viện OpenCV trong Python.

- **Phát hiện khuôn mặt:** Bởi vì hình ảnh toàn cảnh của tài xế có quá nhiều đặc trưng (feature) không cần thiết, nên cần thiết phải có một model phát hiện khuôn mặt để có thể trích xuất khuôn mặt của tài xế. Với hình ảnh đầu vào là hình ảnh toàn cảnh của tài xế, mô hình sẽ trả về kết quả $(x1, y1, x1, x2)$ với $(x1, y1)$ là tọa độ cặp điểm bên trái phía dưới của ô dự đoán và $(x2, y2)$ là tọa độ điểm bên phải trên cùng của ô dự đoán.

- **Cắt, thay đổi kích thước hình ảnh:** Model dự đoán khuôn mặt đã vẽ bounding box khuôn mặt rất chính xác. Và bounding box này thường có hình dạng là hình chữ nhật, tuy nhiên có một vấn đề nảy sinh là hình ảnh đầu vào của Model phân loại là ảnh vuông $(224*224*3)$. Nếu ta cắt ngay ảnh hình chữ nhật rồi thay đổi kích thước về $224*224*3$ thì hình sẽ bị méo, và trở nên không thực tế. Vì vậy ta cần phải biến bounding box của khuôn mặt thành hình vuông bằng cách viết một hàm trong Python.

- **Xóa hình giống nhau:** Ở bước trên, kỹ thuật skip frame đã được áp dụng khi lấy khung hình từ video, tuy nhiên những hình ảnh nhận được sau đó vẫn có khá nhiều ảnh trùng nhau. Vì vậy cần thiết phải thực hiện xóa hình giống nhau để giảm bớt số lượng hình ảnh trong nguồn dữ liệu nhưng vẫn giữ được chất lượng hình ảnh.

- **Gán nhãn:** Việc xác định trạng thái của một người thông qua hình ảnh là việc vô cùng khó khăn, cho dù là đối với con người bởi vì mỗi người có một đặc điểm khuôn mặt khác nhau, đặc biệt là đặc điểm mắt. Quy tắc gán nhãn được thực hiện trong nghiên cứu này như sau: Một người được xem là đang buồn ngủ khi có một hoặc những biểu hiện sau: Ngáp; Mắt nhắm hờ; Có cuống thâm ở mắt, mắt mở nhỏ; Không tỉnh táo. Một người được xem là đang ngủ khi có một hoặc những biểu hiện sau: Nhắm mắt; Cúi đầu không thấy mắt; Ngâng đầu và nhắm mắt. Một người được xem là đang tỉnh táo khi không có những biểu hiện trên. Nếu một hình ảnh có thể thuộc hai lớp buồn ngủ hoặc tỉnh táo hay buồn ngủ và ngủ, thì hình ảnh đó sẽ được xếp ở trạng thái thấp hơn.

- **Chuyển đổi thành ảnh xám:** Bởi vì nguồn dữ liệu video thô gồm có cả ảnh màu, ảnh xám và ảnh hồng ngoại. Vì thế cần thiết phải có một bước chuyển hình ảnh thành ảnh xám để quy toàn bộ dữ liệu về hệ ảnh xám.

3.3. Thực hiện bài toán phân loại dữ liệu

3.3.1. Chuẩn hóa và làm giàu dữ liệu

Sau khi tiến hành xử lý dữ liệu, dữ liệu cho bài toán phân loại trạng thái khuôn mặt như sau [10]-[12]:

- **Tập huấn luyện (Train sets):** 8464 ảnh với 3 lớp: 3672 ảnh lớp tỉnh táo, 2804 ảnh lớp buồn ngủ, 1988 ảnh lớp ngủ. Tập thẩm định (Validation): 1253 ảnh với ba lớp: 484 ảnh lớp tỉnh táo, 436 ảnh lớp buồn ngủ, 333 ảnh lớp ngủ.

- **Làm giàu dữ liệu:** Bởi vì dữ liệu của bài toán bao gồm 8464 bức ảnh, khá ít so với một bài toán Deep Learning thông thường. Vì thế cần phải có sự làm giàu dữ liệu để tăng cường số dữ liệu hiện có. Kỹ thuật làm giàu dữ liệu là kỹ thuật mà ở đó chúng ta tạo thêm những điểm dữ liệu có ích cho bài toán dựa trên dữ liệu ban đầu.

3.3.2. Xây dựng mô hình

Mô hình phát hiện buồn ngủ của hệ thống được xây dựng dựa trên mô hình **MobileNet V2** với kỹ thuật *Fine tuning* và sự hỗ trợ của *Framework Tensorflow Keras*. Mô hình gồm hai phần: Head Model và End Model với các thông số chi tiết như Hình 2.

Layer (type)	Output Shape	Param #
mobilenetv2_1.00_224 (Functional)	(None, 7, 7, 1280)	2257984
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1280)	0
dense (Dense)	(None, 512)	655872
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 256)	131328
dropout_1 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 3)	771

Total params: 3,045,955
Trainable params: 3,011,843
Non-trainable params: 34,112

Hình 2: Thông số của mô hình đề xuất

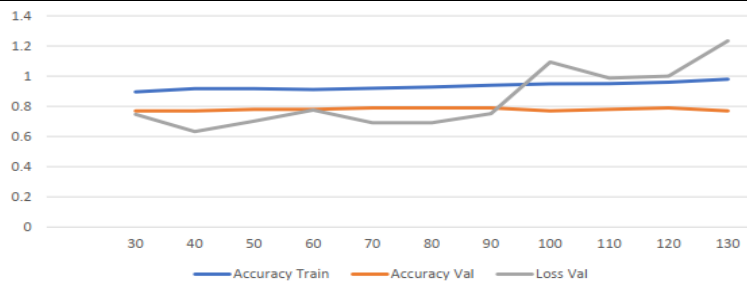
3.2.3. Huấn luyện mô hình

Sau khi hoàn thành việc xây dựng mô hình, mô hình được huấn luyện với toàn bộ dữ liệu đã được làm giàu. Những thông số trong lúc huấn luyện như sau: Thuật toán tối ưu: Adam, learning_rate (tốc độ học) là 0.00003; Hàm mất mát: Categorical_crossentropy; Ma trận đánh giá: Accuracy; Số epoch mỗi lần train: 10 [13]-[15].

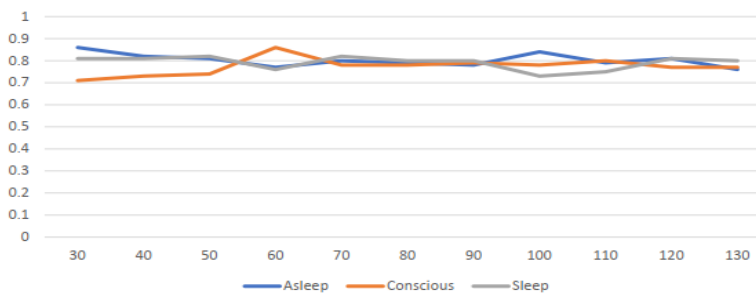
Dựa vào Bảng 1 và các biểu đồ Hình 3, 4, 5 ta có thể thấy ở 70 epoch thì mô hình đã đạt được những tiêu chí mà nghiên cứu đã nêu ra ở trên cũng như có sự cân bằng giữa các thông số trong cả ba lớp. Việc chọn ra một mô hình có sự cân bằng giữa những thông số của cả ba lớp cũng giúp cho mô hình hoạt động tốt trong việc đưa ra cảnh báo cũng như không gửi cảnh báo sai quá nhiều.

Bảng 1: Những thông số của mô hình trong quá trình huấn luyện

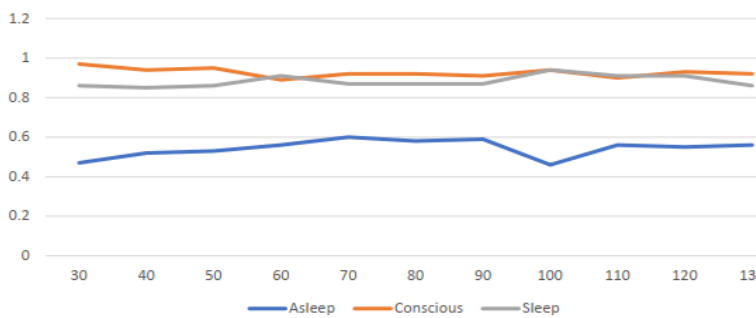
Epoch	Accuracy Train	Accuracy Val	Loss Val	Precision			Recall		
				Asleep	Conscious	Sleep	Asleep	Conscious	Sleep
30	0,896	0,77	0,747	0,86	0,71	0,81	0,47	0,97	0,86
40	0,917	0,77	0,633	0,82	0,73	0,81	0,52	0,94	0,85
50	0,917	0,78	0,702	0,81	0,74	0,82	0,53	0,95	0,86
60	0,912	0,78	0,776	0,77	0,86	0,76	0,56	0,89	0,91
70	0,920	0,79	0,691	0,80	0,78	0,82	0,60	0,92	0,87
80	0,928	0,79	0,692	0,79	0,78	0,80	0,58	0,92	0,87
90	0,940	0,79	0,752	0,78	0,79	0,80	0,59	0,91	0,87
100	0,950	0,77	1,094	0,84	0,78	0,73	0,46	0,94	0,94
110	0,951	0,78	0,987	0,79	0,80	0,75	0,56	0,90	0,91
120	0,960	0,79	1,000	0,81	0,77	0,81	0,55	0,93	0,91
130	0,98	0,77	1,234	0,76	0,77	0,80	0,56	0,92	0,86



Hình 3. Biểu đồ Accuracy Train/Val và Loss Val trong quá trình huấn luyện



Hình 4: Biểu đồ Precision của mỗi lớp trong quá trình huấn luyện



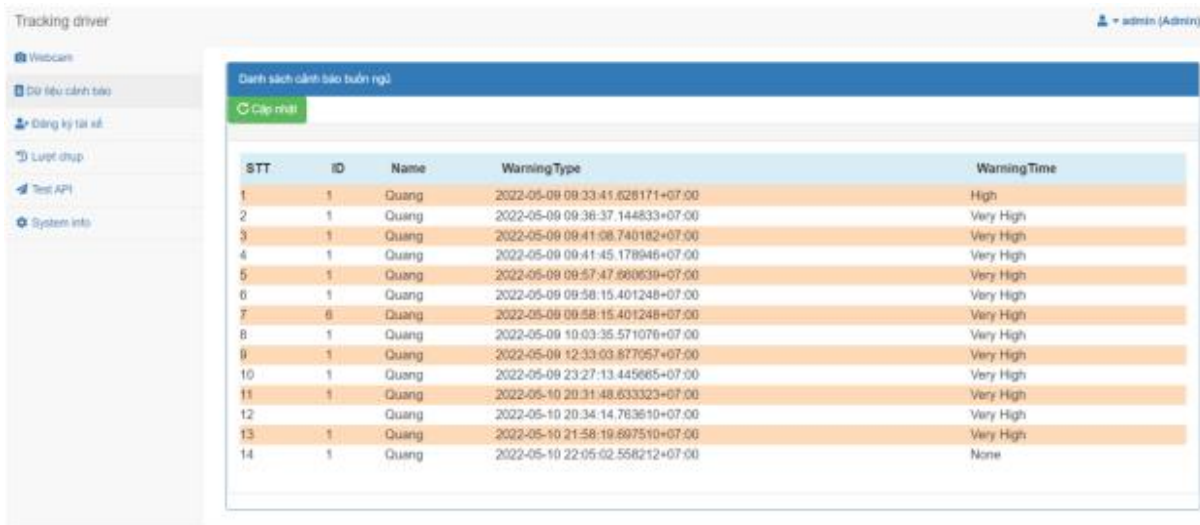
Hình 5: Biểu đồ Recall của mỗi lớp trong quá trình huấn luyện

3.2.4. Đề xuất thuật toán phát hiện buồn ngủ

Sau khi tiến hành huấn luyện rất nhiều mô hình và chọn ra được một mô hình tốt nhất thì độ chính xác của mô hình đã cải thiện lên khoảng 80%. Trạng thái bị dự đoán sai nhiều nhất đó chính là lớp tỉnh táo. Vì thế, không thể dựa hoàn toàn vào từng khoảnh khắc (frame) để quyết định một người có buồn ngủ hay không. Cần nhiều hơn những frame liên tục nhau để quyết định điều này. Thuật toán dưới đây giúp cải thiện các thiếu sót đã nêu:

- Bước 1: Đặt biến $Score=0$
- Bước 2: Thực hiện lấy ba hình ảnh khuôn mặt tài xế liên tục, mỗi hình ảnh cách nhau khoảng 0.12s.
- Bước 3: Dự đoán ba hình ảnh này cùng một lúc:
 - Output là lớp buồn ngủ: $Score=score+1$
 - Output là lớp ngủ: $Score=score+2$
- Bước 4: Tiến hành so sánh Score:
 - $Score \leq 3$: Low warning.
 - $3 < Score < 4$: Medium warning.
 - $4 < Score < 5$: High warning, tiến hành gửi cảnh báo cho Server.
 - $5 \leq Score < 6$: Very high warning, tiến hành gửi cảnh báo cho Server.
- Bước 5: Reset score=0, nghỉ 0.5s và quay lại bước 1.

Chương trình của hệ thống phát hiện buồn ngủ sẽ được nhúng trên máy tính Jetson Nano. Sau đó thực hiện việc gửi cảnh báo buồn ngủ về máy chủ và thực hiện hiển thị thông tin lên website để người dùng có thể quản lý dữ liệu. Dữ liệu cảnh báo mỗi khi tài xế buồn ngủ gồm bốn trường sau: Tên tài xế (Name), mã định danh tài xế (ID), loại cảnh báo (WarningType) và thời gian cảnh báo (WarningTime) như trong Hình 6.



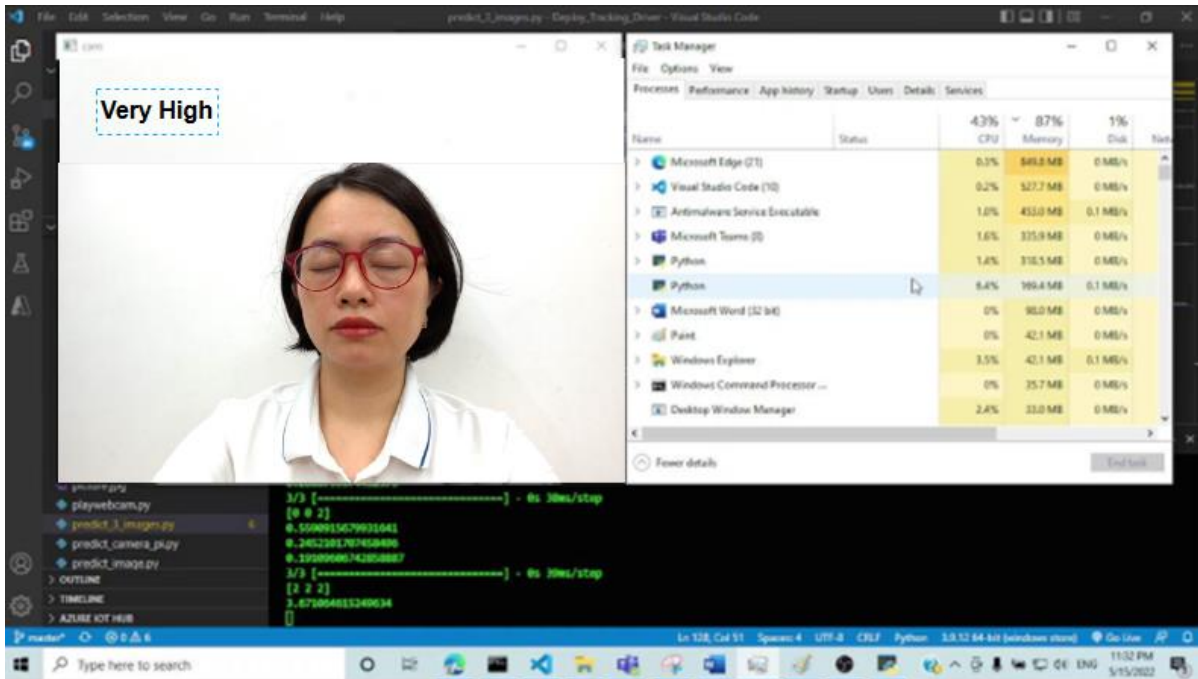
STT	ID	Name	WarningType	WarningTime
1	1	Quang	2022-05-09 09:33:41.628171+07:00	High
2	1	Quang	2022-05-09 09:36:37.144833+07:00	Very High
3	1	Quang	2022-05-09 09:41:08.740182+07:00	Very High
4	1	Quang	2022-05-09 09:41:45.178946+07:00	Very High
5	1	Quang	2022-05-09 09:57:47.690639+07:00	Very High
6	1	Quang	2022-05-09 09:58:15.401248+07:00	Very High
7	6	Quang	2022-05-09 09:58:15.401248+07:00	Very High
8	1	Quang	2022-05-09 10:03:35.571076+07:00	Very High
9	1	Quang	2022-05-09 12:33:00.877057+07:00	Very High
10	1	Quang	2022-05-09 23:27:13.445665+07:00	Very High
11	1	Quang	2022-05-10 20:31:48.633323+07:00	Very High
12		Quang	2022-05-10 20:34:14.763610+07:00	Very High
13	1	Quang	2022-05-10 21:58:19.697510+07:00	Very High
14	1	Quang	2022-05-10 22:05:02.558212+07:00	None

Hình 6: Thông tin về dữ liệu cảnh báo buồn ngủ được hiển thị trên giao diện website

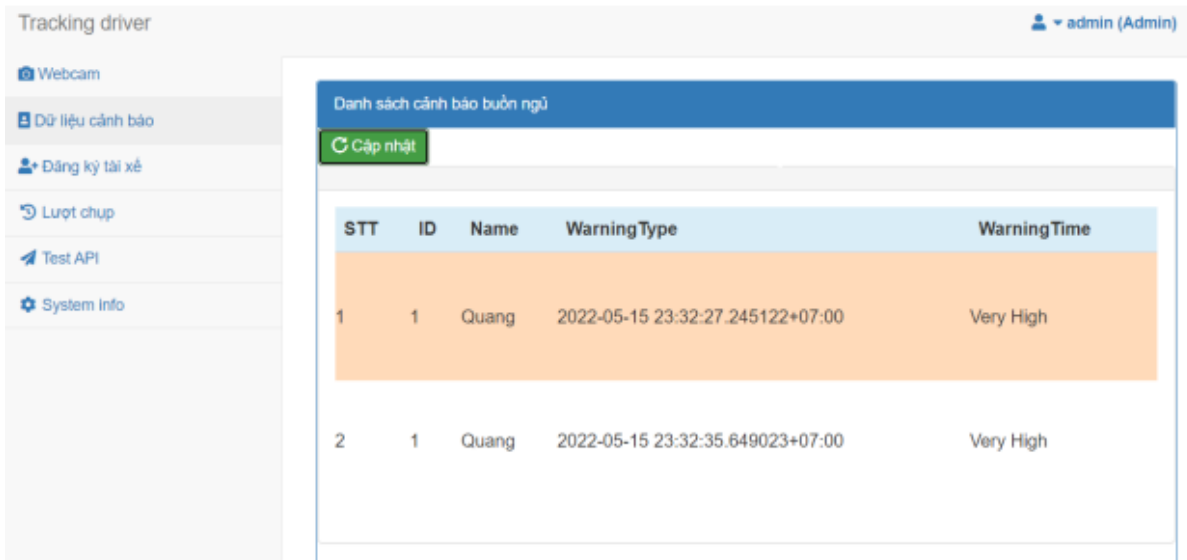
4. Đánh giá và bàn luận

Hệ thống sau khi được xây dựng thành công được chạy thử qua hai phương pháp: Một là: chạy thử bằng webcam mô phỏng việc quay mặt tài xế. Hai là: chạy thử những video từ nguồn dữ liệu thử nghiệm. Kết quả cho thấy: Hệ thống hoạt động trơn tru, không

bị lỗi. Dữ liệu định lượng cho thấy hệ thống phát hiện được hầu hết những trường hợp buồn ngủ ở cả việc mô phỏng và từ video nguồn dữ liệu thử nghiệm. Hệ thống hoạt động theo thời gian thực, phát tín hiệu cảnh báo và gửi cảnh báo vô cùng kịp thời. Kết quả chạy thử nghiệm và thông tin dữ liệu hiển thị thu được thể hiện như trong Hình 7, Hình 8.



Hình 7: Hệ thống phát hiện được buồn ngủ ở thời gian 11:32 phút



Hình 8: Hệ thống gửi cảnh báo về Server và hiển thị lên website

Một số kết quả đánh giá mô hình phát hiện buồn ngủ:

- **Dữ liệu thử nghiệm:** mô hình được lấy từ nguồn dữ liệu DDDD của NTHU. Dữ liệu thử nghiệm gồm có 7 người với 4 hoạt cảnh khác nhau là: Có kính, không kính, có

kính ban đêm và không kính ban đêm. Dữ liệu này gồm 1620 bức ảnh mới: lớp tỉnh táo có 674 bức ảnh, lớp buồn ngủ có 459 bức ảnh, lớp ngủ có 674 bức ảnh. Kết quả hiển thị cho thấy mô hình đã xác định được chính xác các trạng thái buồn ngủ dựa trên dữ liệu đầu vào.

- **Đánh giá bằng độ chính xác và hàm mất mát:** Khi thực hiện đánh giá mô hình dựa trên tập thử nghiệm, độ chính xác và hàm mất mát của mô hình lại rất tốt, độ chính xác (~89,5%) cao hơn nhiều so với độ chính xác trong tập kiểm thử (79%), hàm mất mát của tập kiểm nghiệm cũng là 0.324 thấp hơn so với tập kiểm thử. Độ chính xác của mô hình, đặc biệt là ở trường hợp lớp thứ hai là lớp buồn ngủ, mô hình đem lại dự đoán tốt hơn.

- **Đánh giá mô hình dựa trên Precision và Recall:** đối với bài toán này, chỉ số *Recall* của hai lớp buồn ngủ và ngủ là rất quan trọng. Chỉ số càng cao thì thể hiện mô hình có khả năng không bỏ sót được những trường hợp tài xế đang buồn ngủ hoặc đã ngủ. Đặc biệt chỉ số *Recall* của lớp ngủ là 0,97, tức là gần như mô hình không bỏ sót những trường hợp ngủ nào trong thực tế. Những chỉ số *Precision* và *Recall* của mô hình trên tập thử nghiệm cũng tốt hơn so với tập kiểm thử, điều này là rất tốt.

Thảo luận chung: Sau khi tiến hành thử nghiệm trên tập thử nghiệm, gần như tất cả những chỉ số trên tập thử nghiệm đều tốt hơn trên tập kiểm thử. Điều này trong thực tế hiếm gặp, kết quả này có thể đến từ một số nguyên nhân sau: Tập kiểm thử đa dạng hơn tập thử nghiệm hoặc do Tập kiểm thử có một số dữ liệu bị gán nhãn sai. Về mặt kết quả của mô hình này: nghiên cứu sử dụng mô hình Deep Learning để phát hiện nên có những ưu điểm sau: Độ chính xác cao; Không cần có sự phân tích toán học quá nhiều; Dễ dàng triển khai lên thiết bị với sự hỗ trợ tối đa của các thư viện; Có thể phát hiện được một số trường hợp không thấy mắt, như xoay trái xoay phải hoặc cúi xuống.

5. Kết luận

Bài báo đã trình bày nghiên cứu về phương pháp thực hiện để xây dựng thành công mô hình sản phẩm theo dõi, giám sát trạng thái buồn ngủ của lái xe (tài xế). Kết quả nghiên cứu cho thấy khả năng hoạt động của hệ thống đã có những ưu điểm sau: Hệ thống hoạt động trơn tru, không bị lỗi. Khả năng cập nhật kịp thời cảnh báo để gửi và lưu thông tin dữ liệu về Database và hiển thị rõ ràng lên website. Thời gian xử lý mỗi ba khung (frame) vào khoảng 1.4s. Bộ vi xử lý CPU hoạt động ở mức 30-35%, với RAM sử dụng là 500MB. Thử nghiệm bằng webcam hoặc thử nghiệm bằng video thì khả năng dự đoán rất tốt, gần như dự đoán được những trường hợp buồn ngủ, đã ngủ và cũng rất ít những trường hợp bị đoán nhầm. Bên cạnh đó hệ thống hoạt động với độ chính xác cao, phát hiện được nhiều trường hợp, bao gồm cả ban đêm. Hệ thống hoạt động được theo thời gian thực. Tốn ít RAM và CPU, vì thế có thể triển khai hệ thống thực tế nếu đầu tư nhiều hơn về thời gian và tiền bạc. Bởi vì có máy chủ, nên hệ thống có thể hoạt động mọi lúc mọi nơi, có thể tích hợp lên xe ô tô trên thực tế.

TÀI LIỆU THAM KHẢO

- [1] M. Jain and C. N. Sowmyarani, "Real-Time Driver Drowsiness Detection using Computer Vision," *International Journal of Engineering and Advanced Technology*, vol. 11, issue 1, pp. 109-113, 2021. DOI: 10.35940/ijeat.A3159.1011121

- [2] R. A. Bhope, "Computer Vision based drowsiness detection for motorized vehicles with Web Push Notifications," In *IEEE 4th International Conference on Internet of Things*, Ghaziabad, India, 2019.
- [3] A. Ranjan, K. Vyas, S. Ghadge, S. Patel and S. S. Pawar, "Driver Drowsiness Detection System Using Computer Vision," *International Research Journal of Engineering and Technology (IRJET)*, 2020.
- [4] B. Yang and F. Wang, "Computer Vision Technology for Monitoring of Indoor and Outdoor Environments and HVAC Equipment: A Review," *Sensors*, vol. 23, issue 13, 2023. DOI: 10.3390/s23136186
- [5] Y. Dang and S. Miao, "A fatigue assessment method based on attention mechanism and surface electromyography," *Int. Things Cyber Phys. Syst.*, vol. 3, pp. 112-120, 2023. DOI: 10.1016/j.iotcps.2023.03.002
- [6] M. Kopaczka and D. Merhof, "A modular system for detection, tracking and analysis of human faces in thermal infrared recordings," *Sensors*, vol. 19, issue 19, 2019. DOI: 10.3390/s19194135
- [7] R. N. Khushaba, S. Kodagoda, S. Lal and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Transactions on biomedical engineering*, vol. 58, issue 1, pp. 121-131, 2010. DOI: 10.1109/TBME.2010.2077291
- [8] C. Craye, M. S. K. A. Rashwan and F. Karray, "A multimodal driver fatigue and distraction assessment system," *International Journal of Intelligent Transportation Systems Research*, vol. 14, pp. 173-194, 2016. DOI: 10.1007/s13177-015-0112-9
- [9] L. F., M. Abulhair, A. D. AlShomrani, M. AL-Garni, A. AL-Mutiry, F. AL-Gamdi, and R. Kalenen, "Using haar classifiers to detect driver fatigue and provide alerts," *Multimedia Tools and Applications*, vol. 71, pp. 1857-1877, 2012.
- [11] G. Li, B. Lee and W. Chung, "Smartwatch-Based Wearable EEG System for Driver Drowsiness Detection," *IEEE Sensors Journal*, vol. 15, issue 12, pp. 7169-7180, 2015. DOI: 10.1109/JSEN.2015.2473679
- [12] M. Sunagawa, S. Shikii, W. Nakai, M. Mochizuki, K. Kusakame and H. Kitajima, "Comprehensive Drowsiness Level Detection Model Combining Multimodal Information," *IEEE Sensors Journal*, vol. 20, issue 7, pp. 3709-3717, 2020. DOI: 10.1109/JSEN.2019.2960158
- [13] F. You, X. and H. Li, "A Real-time Driving Drowsiness Detection Algorithm with Individual Differences Consideration," *IEEE Access*, vol. 7, pp. 179396-179408, 2019. DOI: 10.1109/ACCESS.2019.2958667
- [14] A. Dasgupta, D. Rahman and A. Routray, "A Smartphone-Based Drowsiness Detection and Warning System for Automotive Drivers," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, issue 11, pp. 4045-4054, 2019, DOI: 10.1109/TITS.2018.2879609
- [15] M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas and A. Mahmood, "A Survey on State-of-the-Art Drowsiness Detection Techniques," *IEEE Access*, vol. 7, pp. 61904-61919, 2019. DOI: 10.1109/ACCESS.2019.2914373

ABSTRACT

RESEARCH ON DEVELOPING A MONITORING SYSTEM FOR DROWSINESS DETECTION BASED ON COMPUTER VISION

Dien Thi Hong Ha

University of Economics - Industrial Engineering, Hanoi, Vietnam

Received on 26/6/2024, accepted for publication on 16/8/2024

This article focuses on the research and development of a model for a drowsiness detection and monitoring system while driving based on computer vision. The product can be applied to current vehicles (such as cars or vehicles with similar functions). The study utilizes several state-of-the-art sleep detection models to collect and process data and train the model to provide quantitative results. The model is evaluated through this process, and improvements are proposed to make it suitable for vehicles in Vietnam. Experimental results show that the system can capture images via a camera, detect human faces, predict sleep states, and send alerts to a server, displaying warning data on a website interface. Essential evaluations indicate that the system operates stably and relatively accurately according to the pre-established design and programming requirements. The results of this research can be applied in practical scenarios or serve as a practical model for students in artificial intelligence, robotics, and automation control fields to enhance their ability to apply theoretical knowledge to real-world problem-solving.

Keywords: Drowsiness; traffic; smart robots; computer vision; artificial intelligence.