

Detection of Abnormal Driver Behaviors Using Local Binary Patterns and Histogram of Oriented Gradients



Sahand University
of Technology

DOR:

[20.1001.1.23223146.1404.12.1.2.5](https://doi.org/10.1001.1.23223146.1404.12.1.2.5)

Journal of Nonlinear
Systems in Electrical
Engineering

Vol.12, No.1

Spring and Summer 2025

ISSN: 2322 – 3146

<http://jnsee.sut.ac.ir>

Nima Kaveh Maryan, Hossein Ebrahimnezhad*, Ahmadreza Yeganeh Mehr

Computer Vision Res. Lab., Faculty of Electrical and Computer Engineering,
Sahand University of Technology, Tabriz, Iran

E-mails: kaveh96@sut.ac.ir; ebrahimnezhad@sut.ac.ir; ar_yeganehmehr98@sut.ac.ir

Abstract

Nowadays, abnormal and unconventional driver behaviors have become a serious challenge in society, necessitating fundamental solutions and preventive programs. Lack of vehicle control skills, fatigue, drowsiness, and distraction are among the most common causes of traffic accidents. In computer vision, detecting abnormal driver behaviors is both crucial and challenging. Therefore, our goal is to identify abnormal driver behaviors using computer vision techniques, including driver drowsiness, smoking, and mobile phone usage while driving. In this study, we employ two feature extraction methods: Local Binary Patterns (LBP) and Histogram of Oriented Gradients (HOG) to detect abnormal driver behaviors. The first step involves detecting the driver's face, which is identified using 68 key facial landmarks. The face is then divided into four regions: eyes, mouth, right ear, and left ear. In the first method, feature extraction is performed using the LBP algorithm for each behavior, followed by classification using a Support Vector Machine (SVM). In the second method, feature extraction is conducted using HOG, and the extracted features are classified with an SVM. The results indicate that the HOG-based feature extraction method improves classification accuracy in the SVM model, achieving better performance compared to LBP features.

Keywords

Drowsiness; Driver Safety; Machine Learning; Local Binary Patterns; Histogram of Oriented Gradients

1. Short Introduction

Driver distraction detection is a crucial aspect of advanced driver assistance systems aimed at improving transportation safety. Research has concentrated on analyzing driver eye behavior and head movements, utilizing systems like PERCLOS to measure eye closure. Recent advancements include the application of Convolutional Neural Networks (CNNs) and deep learning models such as ResNet and MobileNetV2 to identify distraction patterns. Additionally, studies have tackled challenges in monitoring driver behavior, including visibility obstructions and lighting conditions. Computer vision-based methods have been proposed to effectively detect drowsiness and distraction in drivers.

2. Proposed Work

The proposed method for detecting the driver's status and state involves several steps. First, face detection is performed as the main preprocessing step. After detecting the face, key facial points are extracted, marking the boundaries of facial components such as the eyes, eyebrows, nose, mouth, and jaw. Mapping these points onto the driver's face image allows for precise segmentation of various facial regions. To assess the driver's facial condition and analyze features related to fatigue and distraction, detailed segmentation of areas including the left eye, right eye, left eyebrow, right eyebrow, nose, mouth, left ear, right ear, and jaw is conducted.

Once the facial regions are separated, image features need to be extracted for analyzing and classifying driver behavior. In this phase, two methods, Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP), are used for feature extraction. After extracting features from each facial region, these features are classified using a Support Vector Machine (SVM) algorithm. With the extracted features and the SVM model, different driver states are identified to determine if the driver is experiencing fatigue or distraction.

To evaluate the effectiveness of the proposed method, classification accuracy results for the two feature extraction methods, HOG + SVM and LBP + SVM, are compared. The average accuracy of 88.5% for the HOG feature extraction method with SVM classification demonstrates superior performance due to its ability to extract structural facial features for detecting driver fatigue and distraction. Overall, HOG + SVM performs significantly better than LBP + SVM, which achieves average accuracy of 61.5%, indicating lower efficiency in detecting various facial states.

Although the proposed method relies on manually extracted features like HOG + SVM, it has limitations compared to deep learning-based approaches. Deep networks can learn abstract and complex features, showing better performance under varying lighting conditions, different angles, and random driver movements through techniques such as data augmentation and adaptive learning. Furthermore, hybrid models with advanced settings like multi-objective optimization and adaptive learning enhance detection accuracy. For instance, CNN-LSTM models can analyze temporal changes in driver behavior, while the HOG + SVM method only analyzes static facial features.

3. Conclusion

In conclusion, this study addresses the critical challenge of detecting driver fatigue and distraction to enhance road safety. Our proposed image processing-based system, which includes face detection, facial component segmentation, and driver status classification, achieved an accuracy of 88.5% with the HOG + SVM method, outperforming LBP + SVM at 61.5%. While HOG + SVM is a fast and cost-effective solution, it has limitations under severe environmental changes compared to deep learning approaches. Advanced models like ViT and CNN + LSTM offer higher accuracy but are less suitable for embedded systems due to their high computational requirements. Future research should explore hybrid models that integrate traditional feature extraction with neural networks to improve the accuracy and robustness of driver distraction detection systems.

تشخیص رفتارهای غیر طبیعی رانندگان از طریق آموزش الگوهای باینری محلی و هیستوگرام گرادیان جهت دار

نیما کاوه مریان^۱، حسین ابراهیم نژاد^{۲*}، احمد رضا یگانه مهر^۳

^۱ کارشناسی ارشد مهندسی برق- مخابرات، آزمایشگاه تحقیقاتی بنیایی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی تبریز (سهند)، n_kaveh96@sut.ac.ir

^۲ نویسنده مسؤؤل، استاد گروه برق- مخابرات، آزمایشگاه تحقیقاتی بنیایی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی تبریز (سهند)، ebrahimnezhad@sut.ac.ir

^۳ کارشناسی ارشد مهندسی برق- مخابرات، آزمایشگاه تحقیقاتی بنیایی کامپیوتر، دانشکده مهندسی برق و کامپیوتر، دانشگاه صنعتی تبریز (سهند)، ar_yeganehmehr98@sut.ac.ir



دانشگاه صنعتی سهند

DOR:

[20.1001.1.23223146.1404.12.1.2.5](https://doi.org/10.1001.1.23223146.1404.12.1.2.5)

نشریه سالانه علمی-فنی در مهندسی برق

دوره ۱۲ - شماره ۱

بهار و تابستان ۱۴۰۴

صفحات ۳۶ الی ۵۶

ISSN: 2322-3146

<http://jnsee.sut.ac.ir>

تاریخ پذیرش: ۱۴۰۳/۱۲/۲۲

تاریخ بازنگری: ۱۴۰۳/۶/۲۰

تاریخ ارسال: ۱۴۰۳/۴/۲۷

چکیده

واژه‌های کلیدی

امروزه رفتارهای غیرطبیعی و نامتعارف رانندگان به یک چالش جدی در جامعه تبدیل شده است که ضروری است برای پیشگیری از آن، راهکارها و برنامه‌های اساسی در نظر گرفته شود. کمبود مهارت در کنترل خودرو، خستگی، خواب آلودگی و حواس پرتی از جمله بیشترین دلایل وقوع تصادفات در کشور است. در بنیایی کامپیوتر، تشخیص رفتارهای غیرطبیعی رانندگان بسیار مهم و چالش برانگیز می‌باشد. براین اساس هدف ما تشخیص رفتارهای غیرطبیعی رانندگان با استفاده از بنیایی کامپیوتر است که شامل خواب آلودگی راننده، استعمال دخانیات و صحبت کردن با تلفن همراه می‌باشد. در این تحقیق، سعی شده از دو روش آموزش ویژگی‌های باینری محلی (LBP) و هیستوگرام گرادیان جهت‌دار (HOG) برای تشخیص رفتارهای غیرطبیعی رانندگان استفاده شود. اولین گام تشخیص چهره فرد مورد نظر می‌باشد. برای این امر، چهره فرد با ۶۸ نقطه حیاتی تشخیص داده و در ادامه چهره را به چهار بخش چشم، دهان، گوش راست و گوش چپ تقسیم بندی شده است. در روش اول برای هر کدام از وضعیت‌ها با استفاده از الگوریتم الگوهای باینری محلی استخراج ویژگی انجام شده و در ادامه با استفاده از SVM طبقه‌بندی صورت گرفت. همچنین در روش دوم با هیستوگرام گرادیان جهت‌دار استخراج ویژگی گردیده و ویژگی‌های مربوطه به طبقه‌بندی SVM اعمال شده است. نتایج نشان می‌دهد که ویژگی هیستوگرام گرادیان جهت‌دار مورد استفاده شده در این کار دقت ارزیابی را در طبقه‌بندی SVM بیشتر کرده است و به نتایج بهتری نسبت به ویژگی الگوهای باینری محلی دست یافته است.

خواب آلودگی،

حواس پرتی،

استخراج ویژگی،

ایمنی راننده،

یادگیری ماشین،

الگوهای باینری محلی،

هیستوگرام گرادیان جهت‌دار.

۱- مقدمه

سالانه تعداد زیادی تصادف در سراسر جهان اتفاق می افتد و این امر باعث از دست رفتن جان بسیاری از افراد می شود. خواب آلودگی و حواس پرتی یکی از علل های اصلی تصادفات منجر به مرگ بشمار می آیند، که طبق گزارش سازمان امنیت ترافیک بزرگراه ملی^۱ (NHTSA) در آمریکا، در سال ۲۰۲۰ حواس پرتی جان ۳۱۴۲ نفر و همچنین خواب آلودگی جان ۶۹۷ نفر را گرفته است [۱]. حواس پرتی باعث کاهش آگاهی و تصمیم گیری رانندگان شده در نتیجه پاسخ به رویدادهای رانندگی با تاخیر و اختلال در سرعت همراه می شود. رانندگان حواس پرت، کنترل وسیله نقلیه خود را از دست داده و با وسایل نقلیه یا اشیاء ثابت تصادف می کنند. اداره ملی ایمنی ترافیک بزرگراه طی گزارشی سه رفتار رانندگان، شامل حرف زدن با تلفن همراه، ارسال پیامک و استعمال دخانیات را که باعث حواس پرتی رانندگان برای مدت طولانی از جاده می شود، عامل تصادفات گزارش کرده است [۲]. انجام این رفتار در حین رانندگی بر عملکرد رانندگان اثر می گذارد. به همین دلیل آن ها زمان کمتری را صرف بررسی آینه ها، چراغ راهنمایی و تقاطع ها می کنند. تصادفات وسایل نقلیه موتوری به دلیل حواس پرتی رانندگی به طور پیوسته در سال های اخیر افزایش یافته است و محققان، کارشناسان را بر آن داشت تا برای کاهش حوادث و رانندگی ایمن تحقیقات گسترده ای در سال های گذشته تا به امروز انجام دهند. پیشرفت های سریع در رایانه ها و فناوری های بی سیم، دستگاه های پیچیده و متنوعی را ساخته است که افراد می توانند در حین رانندگی از آنها استفاده کنند. با این حال، استفاده از چنین دستگاه هایی در حین رانندگی هم می تواند حواس راننده را از تمرکز بر رانندگی منحرف کند و هم می تواند در خودروهای هوشمند در مواقع حواس پرتی راننده، هشدارهایی مبنی بر خطر و عدم توجه راننده به او نشان دهد و در صورت عدم توجه به هشدارها، کنترل خودرو را به حالت خودکار در بیاورد.

روش های تشخیص حواس پرتی راننده را می توان بر اساس نوع داده آنها دسته بندی کرد [۳]. منابع داده برای تشخیص حواس پرتی راننده عبارتند از: (۱) نشانگرهای مبتنی بر رفتار خودرو یا رانندگی، (۲) فیزیولوژیکی: سیگنال های بیولوژیکی راننده، (۳) متغیرهای رفتاری (رفتارهای بصری رانندگی) یا متغیرهای فیزیکی راننده براساس ویژگی های بصری، مانند تصاویر یا فیلم های گرفته شده از راننده و (۴) اقدامات ترکیبی که برای شناسایی حالت راننده برای کمک به هدایت ایمن خودرو و عدم بروز تصادف ضروری می باشد. به همین دلیل سیستمی که بتواند خواب آلودگی و حواس پرتی راننده را تشخیص دهد، از اهمیت بالایی برخوردار می باشد. روش های ارائه شده در تشخیص خواب آلودگی و حواس پرتی راننده به دو دسته بصری و غیربصری تقسیم بندی می شوند. در چند دهه گذشته محققان زیادی در زمینه روش های غیربصری تحقیق نموده اند. در روش غیر بصری از سیگنال هایی همچون EEG^2 ، EMG^3 ، EOG^4 و ECG^5 استخراج ویژگی می شود [۴]، که دقت بالایی دارد اما برای به دست آوردن این ویژگی ها باید الکترودهایی به بدن راننده متصل شود که این امر برای فرد خوشایند نیست. در روش بصری رفتار راننده از جمله صحبت کردن با تلفن همراه، استعمال دخانیات و خواب آلودگی راننده مورد نظارت قرار می گیرد و سپس با استفاده از پردازش تصویر و بینایی کامپیوتر، ویژگی های لازم استخراج می شوند. استخراج ویژگی از تصاویر شامل روش های متعددی است که بسته به کاربرد موردنظر می توان از آن ها بهره برد. در پردازش تصویر کلاسیک، تکنیک هایی مانند، تبدیل ویژگی مستقل از مقیاس (SIFT)^۶ و (SURF)^۷ به منظور شناسایی نقاط کلیدی و توصیفگرهای مقاوم در برابر تغییرات مقیاس و چرخش استفاده می شوند. همچنین، ویژگی های مبتنی بر گرادیان نظیر هیستوگرام گرادیان جهت دار (HOG)^۸ برای توصیف ساختارهای لبه ای و تشخیص اشیا کاربرد دارد. یکی دیگر از

¹ National Highway Traffic Safety Administration (NHTSA)

² Electroencephalography

³ Electromyography

⁴ Electrooculography

⁵ Electrocardiography

⁶ Scale-Invariant Feature Transform

⁷ Speeded-Up Robust Features

⁸ Histogram of Oriented Gradients

روش‌های مهم، الگوی باینری محلی (LBP)^۱ است که با بررسی الگوهای بافتی بر اساس مقایسه شدت روشنایی پیکسل‌های مجاور، ویژگی‌های مقاومی را در برابر تغییرات روشنایی و چرخش استخراج می‌کند و در کاربردهایی مانند تشخیص چهره و دسته‌بندی تصاویر بافتی مورد استفاده قرار می‌گیرد. در روش‌های مدرن‌تر، شبکه‌های عصبی پیچشی (CNN) به‌عنوان استخراج‌کننده‌های خودکار ویژگی، توانایی یادگیری نمایش‌های انتزاعی و سلسله‌مراتبی از تصویر را دارند که در کاربردهایی مانند دسته‌بندی، تشخیص اشیا و بازیابی تصویر عملکرد برتری ارائه می‌دهند. در حوزه بینایی ماشین و ردیابی، Optical Flow به تحلیل حرکت پیکسل‌ها در دنباله‌ای از فریم‌ها کمک می‌کند و در کاربردهایی مانند تشخیص فعالیت و تحلیل ویدئو به کار می‌رود. همچنین، مدل‌های مبتنی بر تبدیل‌ها مانند Wavelet Transform برای استخراج ویژگی‌های فرکانسی-مکانی استفاده می‌شوند که در پردازش تصاویر پزشکی و فشرده‌سازی داده‌ها مفید هستند. در نهایت، روش‌های یادگیری عمیق نظیر ViT^۲، به‌ویژه در پردازش تصاویر پیچیده، توانسته‌اند با یادگیری وابستگی‌های طولانی‌مدت در تصویر، عملکرد قابل توجهی در وظایفی مانند طبقه‌بندی کهکشان‌ها و تشخیص الگوهای پیچیده ارائه دهند. ما در این پژوهش از ۲ روش استخراج ویژگی شامل، (۱) هیستوگرام گرادیان جهت‌دار (HOG)، (۲) الگوی باینری محلی (LBP)، استفاده نموده‌ایم و با ارائه راهکاری، خواب‌آلودگی و عوامل منجر به حواس‌پرتی رانندگان را با بررسی رفتارهای چشم، صحبت با تلفن و کشیدن سیگار تشخیص می‌دهیم. ساختار تحقیق به این صورت است، که بخش اول شامل مقدمه و مفاهیم اولیه در مورد تشخیص حواس‌پرتی و خستگی رانندگان می‌باشد. بخش دوم، به مرور مختصری بر کارهای انجام شده در مورد رفتارهای غیرنرمال رانندگان پرداخته است. در بخش سوم، روش پیشنهادی مورد بحث قرار خواهد گرفت. در بخش چهارم، نتایج به‌دست آمده از روش انجام شده مورد ارزیابی قرار می‌گیرد و در بخش پنجم، نتیجه‌گیری کلی از تحقیق انجام می‌گیرد.

۲- کارهای مربوطه

تشخیص حواس‌پرتی راننده یکی از حوزه‌های مهم در سیستم‌های پیشرفته کمک‌راننده (ADAS) و افزایش ایمنی حمل‌ونقل است که تاکنون مطالعات متعددی در این زمینه انجام شده است. این پژوهش‌ها عمدتاً بر تحلیل رفتار چشمی و حرکات سر، استفاده از یادگیری عمیق برای پردازش تصاویر، پردازش سیگنال‌های بیولوژیکی و ترکیب چندمنظوره داده‌ها متمرکز بوده‌اند. یکی از مهم‌ترین روش‌های تشخیص حواس‌پرتی، تحلیل رفتار چشمی راننده است. مطالعه‌ای توسط جای و همکاران، سیستم PERCLOS را معرفی کرد که درصد بسته بودن چشم‌ها را اندازه‌گیری کرده و به‌عنوان شاخصی برای تشخیص خواب‌آلودگی و حواس‌پرتی به کار گرفته شد [۵]. مطالعه دیگری توسط ویستنه و همکاران، نشان داد که تحلیل مسیر نگاه راننده و ارتباط آن با محیط جاده‌ای می‌تواند نقش مهمی در شناسایی عدم تمرکز داشته باشد [۶]. روش‌های مبتنی بر یادگیری ماشین مانند SVM و Random Forest نیز برای تحلیل الگوهای چشم و سر راننده به کار گرفته شده‌اند [۷]. با پیشرفت بینایی کامپیوتر، مطالعاتی از شبکه‌های عصبی پیچشی (CNN) و مدل‌های یادگیری عمیق برای تشخیص الگوهای پیچیده چهره و حرکات سر بهره برده‌اند. پژوهش ابولناگا و همکاران، مدلی مبتنی بر ResNet را ارائه داد که ویژگی‌های صورت راننده را برای تشخیص توجه وی به مسیر استخراج می‌کرد [۸]. در یک مطالعه دیگر، باهتی و همکاران، از مدل MobileNetV2 برای تشخیص استفاده از تلفن همراه حین رانندگی استفاده کردند و نشان دادند که این روش عملکرد بالایی در مقایسه با مدل‌های کلاسیک دارد [۹]. علاوه بر این، اخیراً استفاده از ویژن ترانسفورمر (ViT) در تحلیل تصاویر رانندگان مورد بررسی قرار گرفته است که دقت بالاتری در تشخیص الگوهای غیرعادی نشان داده است [۱۰]. برخی مطالعات از سیگنال‌های EEG و ECG برای شناسایی وضعیت ذهنی راننده استفاده کرده‌اند. پژوهش جانگ و همکاران، نشان داد که

¹ Local Binary Patterns

² Vision Transformer

سیگنال‌های EEG می‌توانند برای شناسایی کاهش هوشیاری و حواس‌پرتی مورد استفاده قرار گیرند [۱۱]. همچنین، مطالعه‌ای توسط لین و همکاران، به بررسی تغییرات ضربان قلب (HRV) پرداخت و نشان داد که کاهش تمرکز راننده با تغییر در الگوی HRV قابل تشخیص است [۱۲، ۱۳، ۱۴]. محققان برای تعیین میزان خستگی و خواب‌آلودگی راننده با بررسی و استخراج ویژگی در زمان‌های هوشیاری و خواب‌آلودگی بر سیگنال‌های الکتروانسفالوگرام (EEG) و الکتروکاردیوگرام (ECG) توانستند سطح هوشیاری فرد را تعیین کنند و در نتیجه حالت خواب‌آلودگی را در مراحل اولیه تشخیص بدهند. سیگنال EEG یک وسیله فیزیولوژیکی غیرتهاجمی برای اندازه‌گیری فعالیت مغز است و نزدیک‌ترین رابطه را با خواب‌آلودگی دارد [۱۵]. به‌ویژه، در شرایط مشابه، EEG بهتر از هر سیگنال فیزیولوژیکی دیگر عمل می‌کند [۱۶]. سیگنال EEG دارای باندهای فرکانسی متفاوتی است. بررسی این سیگنال در هنگام خواب‌آلودگی نشان داده است که باندهای آلفا و تتا تغییرات گسترده‌ای نسبت به سایر باندهای فرکانسی دارند. همچنین افزایش تغییرات قدرت در باند فرکانسی تتا و کاهش آن در باند فرکانسی آلفا، یکی از نشانه‌های خواب‌آلودگی محسوب می‌شود [۱۷]. در روش بصری با استفاده از بینایی کامپیوتر، حرکات رانندگان مورد نظارت قرار می‌گیرد. برای این امر نیاز به نصب دوربین در داخل وسیله نقلیه برای بررسی رفتارهای رانندگان وجود دارد. این رفتارها شامل حرکات صورت مانند پلک زدن مداوم، چشمان باز و بسته، خمیازه کشیدن، تکان دادن یا تکان ندادن سر است [۱۸]. البته در این رویکرد می‌توان به چالش‌هایی همچون پوشاندن چشم به وسیله عینک آفتابی یا عینک معمولی، پوشاندن دهان توسط ماسک یا دست، زاویه دید راننده، شرایط نوری، حالت موی راننده و همچنین داشتن نقاب یا کلاه اشاره کرد [۱۹]. در این روش برای نظارت بر رانندگان می‌توان از رفتارهای مختلف برای تشخیص خواب‌آلودگی و حواس‌پرتی که یکی از علل اصلی تصادفات مرگبار است، استفاده کرد. استخراج ویژگی از چشم‌ها و رفتار رانندگان مانند صحبت کردن با تلفن همراه و استعمال دخانیات از جمله این روش‌ها است. در پژوهشی، سلیمان و همکاران [۲۰] از یک روش مبتنی بر یادگیری ماشین برای تشخیص خواب‌آلودگی استفاده کردند. در این روش، دوربینی متصل به داشبورد تصاویری از چهره راننده اخذ کرده، سپس با الگوریتم Viola-Jones و ویژگی‌های Haar چهره را تشخیص داده و ویژگی‌های چشم و دهان را استخراج کردند. در ادامه از روش درخت تصمیم‌گیری برای طبقه‌بندی حالت باز و بسته چشم استفاده شد. مبین عاریف و همکاران [۲۱] از روش پردازش تصویر زمان واقعی برای تشخیص خواب‌آلودگی راننده و سیستم هشدار مبتنی بر نرخ پلک زدن استفاده کردند. عبدالمالک مجاهد و همکاران [۲۲] یک سیستم مانیتورینگ چهره را بر اساس یک توصیفگر بافت چهره فشرده ارائه کردند که می‌تواند متمایزترین ویژگی‌های خواب‌آلودگی را پوشش دهد. در اینجا، چشم‌های هر چهره با استفاده از الگوریتم گروه درختان رگرسیون (ERT) شناسایی شدند که یک الگوریتم قوی و بسیار کارآمد برای محلی‌سازی نشانه‌های چهره محسوب می‌شود. سپس نمایش چندسطحی PML برای استخراج ویژگی‌های محلی هر بلوک در هر سطح با استفاده از توصیفگرهای HOG، COV و LBP انجام شد. داده‌ها با تجزیه و تحلیل مؤلفه اصلی (PCA) کاهش یافتند و در نهایت، طبقه‌بندی با استفاده از ماشین بردار پشتیبان (SVM) انجام شد. ریموند پتوچا و همکاران [۲۳] با استفاده از بینایی کامپیوتر، تلاش کردند رانندگان حواس‌پرت را در مجموعه داده‌ای شامل هفت رفتار مختلف حواس‌پرتی شناسایی کنند. در این تحقیق، ویژگی‌های بینایی کامپیوتر شامل ترکیبی از هیستوگرام گرادبان‌های جهت‌دار (HOG) و توصیفگرهای SIFT همراه با طبقه‌بندی SVM مورد استفاده قرار گرفت. کیم [۲۴] روشی مبتنی بر الگوی باینری محلی (LBP) و SVM برای تشخیص دود در دوربین‌های نظارتی ارائه کرد. شائوچی و همکاران [۲۵] نیز روشی مبتنی بر یادگیری عمیق برای تشخیص فعالیت‌های منجر به حواس‌پرتی راننده ارائه دادند که در آن، تصاویر راننده شامل صحبت کردن با تلفن و پیامک زدن توسط دوربینی متصل به داشبورد وسیله نقلیه جمع‌آوری شد. سیستم مورد نظر از شبکه ResNet-50 که یک CNN پیشرفته است، برای تشخیص حواس‌پرتی رانندگان بهره برد. بررسی مطالعات گذشته نشان می‌دهد که روش‌های مبتنی بر پردازش تصویر و یادگیری عمیق عملکرد بالایی در تشخیص حواس‌پرتی راننده دارند، اما ترکیب این روش‌ها با داده‌های سیگنال‌های زیستی و حسگرهای داخل خودرو می‌تواند دقت سیستم‌های تشخیص را افزایش دهد. به‌ویژه، استفاده از شبکه‌های

عصبی عمیق و مدل‌های مبتنی بر ویژن ترنسفورمر (ViT) رویکردی نوین در این حوزه محسوب می‌شود که می‌تواند در آینده به بهبود سیستم‌های پیشرفته کمک‌رساننده منجر شود.

۳- روش پیشنهادی

ما در این بخش به توضیح روش پیشنهادی خواهیم پرداخت. در این روش تلاش شده است با تمرکز بر استخراج دقیق ویژگی‌های چهره و ترکیب تکنیک‌های مؤثر پردازش تصویر به بهبود دقت سیستم‌های پایش رانندگان کمک شود. پیش از پرداختن به نحوه پیاده‌سازی روش، لازم است به توضیحاتی در مورد روش‌ها و الگوریتم‌های مورد استفاده پرداخته شود، تا ضمن آشنایی با آنها، بیان روش پیشنهادی راحت‌تر گردد و از ابهامات در این خصوص جلوگیری شود. در ادامه توضیحاتی در خصوص الگوی باینری محلی (LBP)، هیستوگرام گرادیان جهت دار (HOG) و ماشین بردار پشتیبان (SVM) داده شده است و سپس به تشریح روش به طور گسترده پرداخته شده است.

۳-۱- الگوی باینری محلی (LBP)

الگوهای باینری محلی، یا به اختصار LBP ها، برای اولین بار توسط تیمو اوچالا و همکارانش معرفی شد [۲۶]. برخلاف ویژگی‌های بافت Haralick که یک نمایش کلی از بافت را بر اساس ماتریس هم‌وقوع سطح خاکستری محاسبه می‌کنند، الگوهای باینری محلی (LBP) در عوض نمایش محلی بافت را محاسبه می‌کنند. این نمایش محلی با مقایسه هر پیکسل با همسایگی مقادیر پیکسل اطراف آن انجام می‌شود [۲۷، ۲۸]. اولین مرحله در ساخت توصیفگر LBP، تبدیل تصویر به مقیاس خاکستری است. برای هر پیکسل در تصویر خاکستری، همسایگی به اندازه r اطراف پیکسل مرکزی انتخاب می‌شود. به این صورت که تصویر به قسمت‌های یکسان تقسیم گردیده و برای هر قسمت و تمامی پیکسل‌های موجود با توجه به اختلاف پیکسل‌های همسایه با پیکسل مرکزی یک ماتریس ایجاد می‌گردد. همانطور که بیان شد پیکسل مرکزی مورد نظر در برابر پیکسل‌های همسایه آن مورد بررسی قرار می‌گیرد. اگر شدت پیکسل مرکزی بیشتر از همسایه‌اش باشد، مقدار آن را ۱ و اگر کوچکتر باشد عدد صفر جای‌گذاری می‌گردد. سپس به صورت پاد ساعت گرد اعداد هر ماتریس که به صورت باینری می‌باشد، خوانده می‌شود. در ادامه عدد دهدهی مربوط به عدد باینری بدست می‌آید. برای همه پیکسل‌ها این کار انجام می‌گردد تا در نهایت همه اعداد دهدهی که از m ناحیه بدست آمده‌اند، یک ماتریس هیستوگرام $m \times n$ را برای هر تصویر به وجود می‌آورند.

۳-۲- هیستوگرام گرادیان جهت دار (HOG)

هیستوگرام گرادیان جهت دار (HOG) یک توصیفگر ویژگی است که توسط ناونیت دالال و بیل تریگر مطرح شده است [۲۹]، اغلب برای استخراج ویژگی‌ها از داده‌های تصویر استفاده می‌شود. این توصیفگر بر این ایده استوار است که ظاهر و شکل یک شی در یک تصویر را می‌توان با استفاده از توزیع گرادیان‌های شدت یا جهت‌های لبه‌ها در نواحی محلی تصویر توصیف کرد. هیستوگرام گرادیان جهت دار (HOG) برخلاف توصیفگر ویژگی‌های لبه، قادر است جهت لبه را نیز ارائه دهد. این کار با استخراج گرادیان جهت‌دار در ناحیه‌های مختلف تصویر انجام شده است. برای بدست آوردن ماتریس هیستوگرام به این صورت عمل می‌شود که ابتدا تصاویر به قطعات کوچک متصل به هم تقسیم می‌شود که به آن‌ها سلول گفته می‌شود. سپس برای هر پیکسل درون یک سلول، گرادیان شدت یا جهت لبه محاسبه می‌شود. این اطلاعات برای ساخت یک هیستوگرام از جهت‌های گرادیان در هر سلول استفاده

می‌شود. در نهایت، توصیفگر HOG با ترکیب هیستوگرام‌های همه سلول‌ها ایجاد می‌شود. برای بهبود دقت و کارایی، هیستوگرام‌های محلی سلول‌ها را می‌توان با در نظر گرفتن یک ناحیه بزرگ‌تر در تصویر به نام بلوک و محاسبه یک مقیاس شدت کلی برای آن بلوک نرمال‌سازی کرد. سپس از این مقیاس برای نرمال کردن مقادیر هیستوگرام در سلول‌های داخل آن بلوک استفاده می‌شود. این فرآیند نرمال‌سازی کنتراست، باعث می‌شود توصیفگر HOG نسبت به تغییرات روشنایی و سایه در تصویر مقاوم‌تر شود. این اطلاعات برای بازنمایی و تشخیص اشیاء در تصاویر مفید هستند و در کاربردهای مختلفی مانند تشخیص عابر پیاده، تشخیص چهره و دسته‌بندی تصاویر استفاده می‌شوند.

۳-۳- ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان (SVM)، یک الگوریتم نظارت‌شده یادگیری ماشین است که برای مسائل طبقه‌بندی و رگرسیون قابل استفاده است. با این حال از آن بیشتر در مسائل طبقه‌بندی استفاده شده است. در الگوریتم ماشین بردار پشتیبان (SVM)، هر نمونه داده به عنوان یک نقطه در فضای n بعدی روی نمودار پراکنندگی داده‌ها ترسیم می‌شود، n تعداد ویژگی‌هایی است که یک نمونه داده دارد و مقدار هر ویژگی مربوط به داده‌ها، یکی از مؤلفه‌های مختصات نقطه روی نمودار را مشخص می‌کند. الگوریتم ماشین بردار پشتیبان (SVM) سعی می‌کند یک سطح یا فضای جداکننده پیدا کند که داده‌های دو کلاس مختلف را با بیشترین فاصله از یکدیگر جدا کند، به این فاصله حاشیه گفته می‌شود. اگر خط جداکننده حاشیه کمی داشته باشد، احتمال طبقه‌بندی نشدن برخی داده‌ها وجود دارد. معمولاً نمونه‌ها به نحوی در فضای ویژگی پخش شده‌اند که طبقه‌بندی خطی قادر به تفکیک آنها نمی‌باشد. در نتیجه ماشین بردار پشتیبان (SVM) از روشی که به آن کرنل گفته می‌شود، استفاده می‌کند. در این روش در واقع توابعی وجود دارند که فضای ورودی بعد پایین را دریافت کرده و آن را به فضای بعد بالاتر تبدیل می‌کنند. این تبدیل، یک مسئله غیر قابل جداسازی را به مسئله قابل جداسازی مبدل می‌کند. به این توابع، کرنل گفته می‌شود. انواع کرنل‌ها عبارتند از کرنل چند جمله‌ای، کرنل گاوسی، کرنل تابع پایه شعاعی و کرنل سیگموئید. در اینجا از کرنل تابع پایه شعاعی استفاده شده که رابطه‌ی آن به صورت زیر می‌باشد:

$$K(x, \hat{x}) = \exp(-\gamma \|x - \hat{x}\|^2) \quad (1)$$

و برای $\gamma > 0$ ، γ با σ نسبت معکوس دارد. همچنین گاهی اوقات γ با استفاده از پارامتر زیر استفاده می‌شود

$$\gamma = 1/2\sigma^2 \quad (2)$$

که در آن واریانس و شباهت یا نزدیکی X و X' بردار ویژگی، محاسبه شده است.

۳-۴- ساختار روش پیشنهادی

روش پیشنهادی در چند گام اقدام به تشخیص وضعیت و حالت راننده می‌کند که در ادامه به تشریح بخش‌های مختلف آن می‌پردازیم.

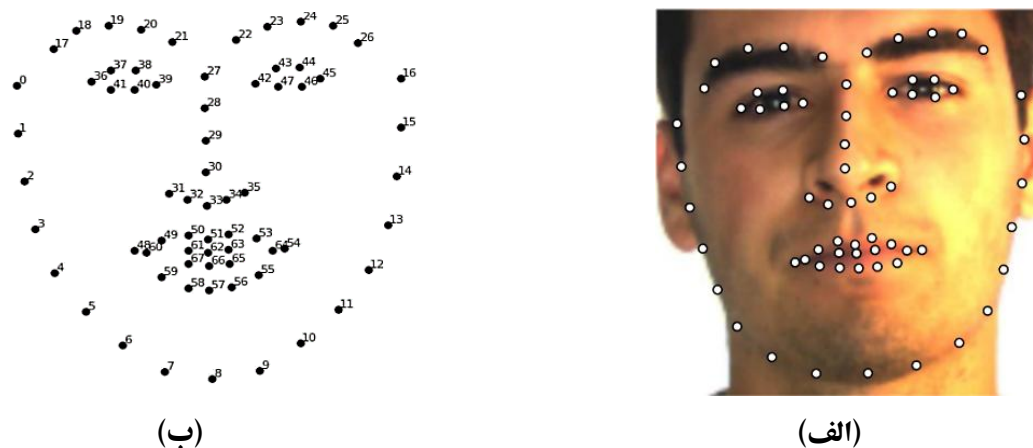
۳-۴-۱- تشخیص چهره در تصاویر رانندگان

در مرحله نخست از روش پیشنهادی، تشخیص چهره راننده به عنوان پیش‌پردازش اصلی انجام می‌شود. برای این کار، از الگوریتم کاظمی و سالیوان [۳۰] که در کتابخانه Dlib پیاده‌سازی شده است، استفاده شده که یکی از دقیق‌ترین و سریع‌ترین الگوریتم‌ها برای تشخیص چهره محسوب می‌شود. این روش مبتنی بر مدل جنگل‌های تصمیم‌گیری تجمعی (Ensemble of Regression Trees) بوده و امکان محلی‌سازی نقاط کلیدی چهره را با دقت بالا فراهم می‌کند. مزیت استفاده از این الگوریتم، سرعت بالای اجرا و عملکرد

مناسب آن حتی در شرایطی مانند تغییرات نور، زوایای مختلف صورت و وجود موانعی مانند عینک یا ماسک جزئی است.

۳-۴-۲- استخراج نقاط کلیدی چهره

پس از تشخیص چهره، الگوریتم کاظمی و سالیوان از مدل ۶۸ نقطه‌ای برای استخراج مختصات نقاط کلیدی چهره استفاده می‌کند. این نقاط کلیدی، مرزهای اصلی اجزای صورت را مشخص کرده و شامل بخش‌هایی مانند چشم‌ها، ابروها، بینی، دهان و فک می‌شود. نگاشت این نقاط روی تصویر چهره راننده، امکان بخش‌بندی دقیق نواحی مختلف صورت را فراهم می‌کند. در شکل ۱، نمونه‌ای از این نقاط کلیدی روی چهره نمایش داده شده است که نشان می‌دهد هر بخش از چهره دارای شاخص‌های مختصات مشخصی است. تجسم ۶۸ مختصات (x, y) برجسته چهره را می‌توان در شکل ۱ (ب) مشاهده کرد.



شکل ۱: شناسایی چهره با کتابخانه Dlib. الف) نمونه‌ای از آشکارسازی ۶۸ مختصات برجسته چهره، ب) تجسم ۶۸ مختصات برجسته چهره.

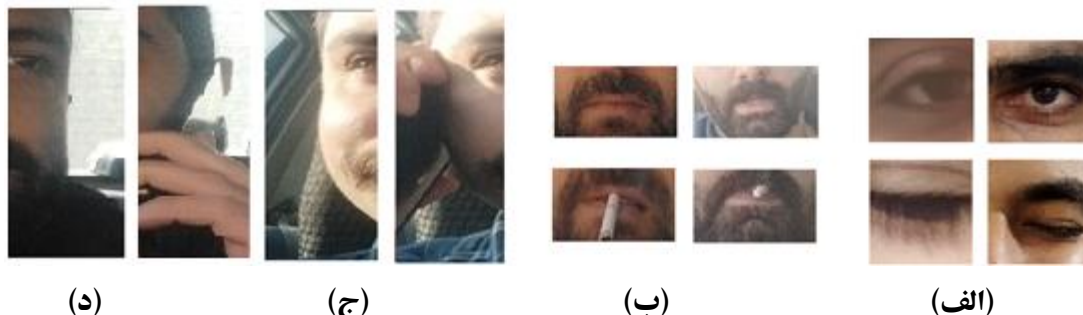
جدول ۱، مقادیر مربوط به بخش‌بندی صورت بر اساس نقاط کلیدی مدل ۶۸ نقطه‌ای Dlib

بخش صورت	محدوده نقاط کلیدی	افست (پیکسل)	توضیحات
چشم چپ	نقاط ۳۶ تا ۴۱	-	شامل پلک بالا و پایین
چشم راست	نقاط ۴۲ تا ۴۷	-	شامل پلک بالا و پایین
ابروی چپ	نقاط ۱۷ تا ۲۱	-	ناحیه ابروی چپ
ابروی راست	نقاط ۲۲ تا ۲۶	-	ناحیه ابروی راست
بینی	نقاط ۲۷ تا ۳۵	-	شامل پل بینی و سوراخ‌های بینی
دهان	نقاط ۴۸ تا ۶۷	-	شامل لب بالا و پایین
گوش چپ	نقاط ۱۷ و ۸	۷۵ پیکسل	تخمین لبه گوش چپ
گوش راست	نقاط ۲۵ و ۸	۷۵ پیکسل	تخمین لبه گوش راست
فک	نقاط ۰ تا ۱۶	-	شامل خط فک و چانه

۳-۴-۳- بخش‌بندی نواحی مختلف صورت

به منظور بررسی وضعیت چهره راننده و تحلیل ویژگی‌های مرتبط با خستگی و حواس‌پرتی، بخش‌بندی دقیق نواحی مختلف چهره انجام می‌شود. برای این منظور، نقاط کلیدی استخراج‌شده از مرحله قبل مورد استفاده قرار گرفته و نواحی به صورت مجزا مطابق جدول ۱ تعیین می‌شوند.

جدول ۱ مشخص می کند که هر بخش صورت چگونه توسط نقاط کلیدی مدل مشخص شده و نواحی اضافی (مانند گوش‌ها) با مقدار آفست ۷۵ پیکسل استخراج شده‌اند. بخش‌بندی به ما این امکان را می‌دهد که هر یک از نواحی چهره را به‌طور جداگانه تحلیل کرده و ویژگی‌های خاص هر بخش را استخراج کنیم. شکل ۲ نشان‌دهنده نحوه جداسازی این نواحی بر اساس نقاط کلیدی و آفست مشخص شده است.



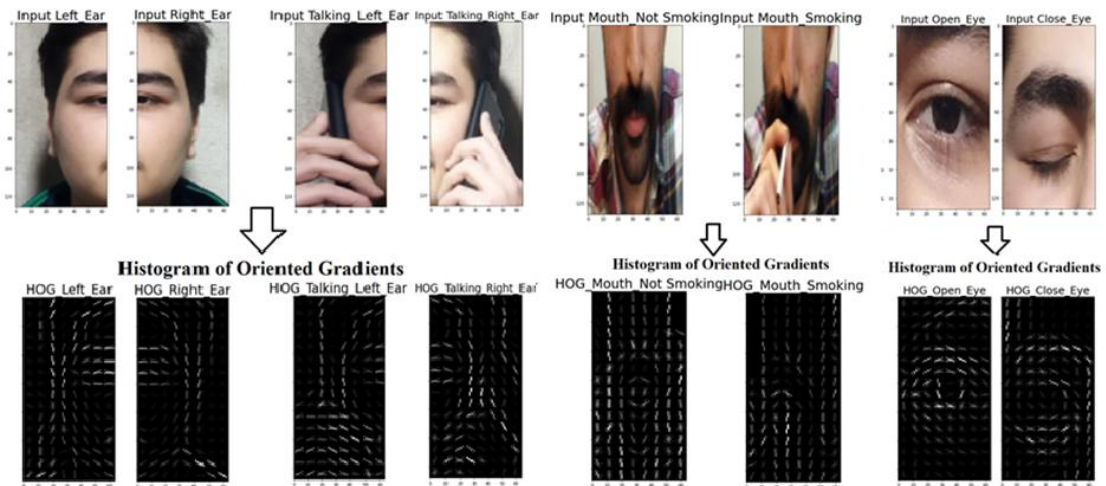
شکل ۲: نمونه جداسازی نواحی مختلف صورت. الف) جداسازی چشم راست و چپ، ب) جداسازی دهان، ج) جداسازی گوش چپ، د) جداسازی گوش راست.

۳-۴-۴- استخراج ویژگی‌های تصویری

پس از جداسازی نواحی مختلف چهره، لازم است ویژگی‌های تصویری استخراج شوند تا امکان تحلیل و طبقه‌بندی رفتار راننده فراهم گردد. در این مرحله، دو روش هیستوگرام گرادیان جهت‌دار (HOG) و الگوی باینری محلی (LBP) برای استخراج ویژگی‌ها مورد استفاده قرار می‌گیرند. قبل از استخراج ویژگی توسط HOG، به عنوان پیش پردازش عکس‌ها به اندازه 128×256 تبدیل می‌شوند. در الگوریتم HOG، تصویر ابتدا به بلوک‌های کوچک تقسیم شده و برای هر بلوک، گرادیان‌های مربوط به تغییرات شدت پیکسل‌ها در جهت‌های مختلف محاسبه و در یک هیستوگرام ذخیره می‌شود. این ویژگی‌ها به‌طور خاص برای شناسایی تغییرات در ساختار چهره، مانند افتادگی پلک‌ها یا تغییر در وضعیت دهان، مفید هستند. برای تشخیص خستگی و حواس‌پرتی رانندگان، ابتدا برای تمامی کلاس‌ها با استفاده از هیستوگرام استخراج ویژگی صورت می‌گیرد. برای هر تصویر آموزشی و تست، ویژگی‌های بدست آمده از روش HOG به صورت یک ماتریس $1 \times n$ می‌باشد که بسته به نوع تصویر ویژگی‌های متفاوتی بدست می‌آید. در شکل ۳ ویژگی‌های استخراج شده از تصاویر و جهت لبه‌ها نشان داده شده است.

ویژگی‌های استخراج شده توسط الگوریتم در چهار حالت برای آموزش به SVM داده می‌شوند. تصاویر مربوطه برچسب گذاری شده و SVM توسط تصاویر آموزش ساخته می‌شود و در ادامه طبقه‌بندی تصاویر آزمایشی توسط SVM صورت می‌پذیرد. در الگوریتم HOG با در نظر گرفتن ۹ جهت گیری^۱ و سلول 8×8 سعی گردیده عوامل حواس‌پرتی و خستگی در رانندگان تشخیص داده شود. شکل ۶ نمونه‌هایی از تشخیص درست و غلط در روش پیشنهادی را نشان می‌دهد. در شکل ۴ نمونه تشخیص‌های صحیح و نادرست از حالت‌های راننده نشان داده شده است.

¹ orientations



(د)

(ج)

(ب)

(الف)

شکل ۳: استخراج ویژگی هیستوگرام گرادیان جهت دار (HOG). الف) استخراج ویژگی چشم راست و چپ، ب) استخراج ویژگی دهان، ج) استخراج ویژگی گوش چپ، د) استخراج ویژگی گوش راست



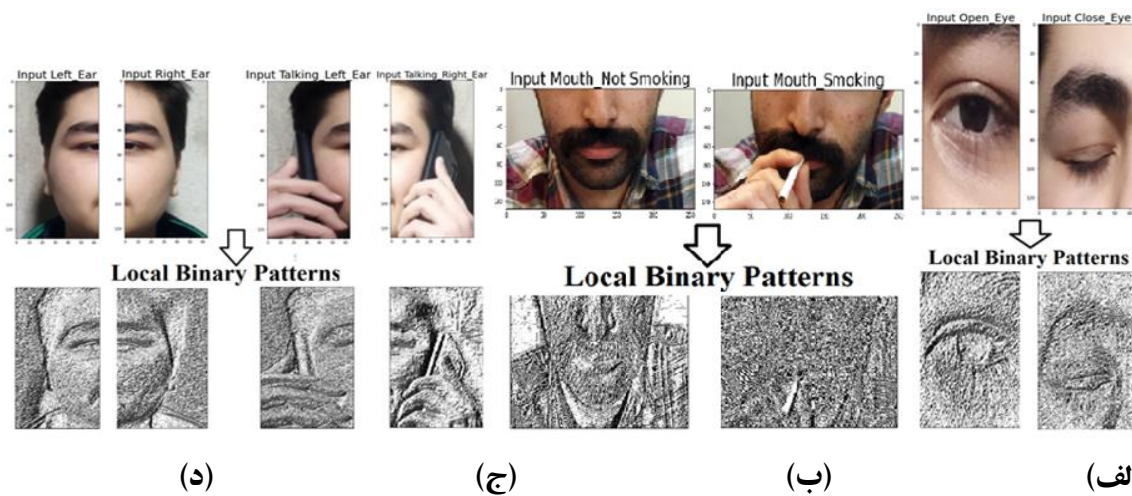
(الف)



(ب)

شکل ۴: نتایج حاصل از استخراج ویژگی و طبقه‌بندی SVM برای تشخیص خستگی و حواس پرتی رانندگان، الف) نمونه‌هایی که درست تشخیص داده شده، ب) نمونه‌هایی که اشتباه تشخیص داده شده.

از روش الگوی باینری محلی LBP، که یکی دیگر از الگوریتم‌های متداول برای تحلیل بافت تصاویر است، در تشخیص ناهنجاری‌ها و تغییرات ظاهری چهره راننده استفاده شده است. در این روش، هر پیکسل از تصویر بر اساس مقدار شدت روشنایی پیکسل‌های همسایه‌اش کدگذاری شده و یک مقدار باینری برای آن محاسبه می‌شود. ویژگی‌های استخراج شده توسط LBP به‌ویژه برای تحلیل تغییرات ظاهری ناشی از خستگی و خواب‌آلودگی، مانند چین و چروک‌های پلک و افتادگی گوشه‌های دهان، بسیار موثر هستند. در شکل ۵ عملکرد الگوریتم بر تمامی کلاس‌ها نشان داده شده است. همانطور که در شکل ۵ نشان داده شده است، بخش‌های چهره شناسایی شده توسط Dlib از مجموعه آموزش استخراج گردیده است. در ادامه تصاویر مربوطه برچسب گذاری شده و از آن‌ها برای ساخت و آموزش SVM روش پیشنهادی استفاده نمودیم. مطابق مراحل الگوریتم HOG، به عنوان پیش پردازش عکس‌ها به اندازه 128×256 تبدیل شده و وارد الگوریتم با چهار کلاس برای آموزش و طبقه بندی می‌شود. در الگوریتم $LBP = 64$ P پیکسل بر روی دایره‌ای به شعاع $R = 4$ حول پیکسل در نظر گرفته شده است. نتایج حاصل از نمونه‌های داده‌های آزمایش در شکل ۶ برای تشخیص صحیح و نادرست حواس‌پرتی و خستگی در رانندگان نشان داده شده است.



شکل ۵: ویژگی‌های استخراج شده توسط الگوریتم الگوی باینری محلی (LBP). الف) استخراج ویژگی چشم راست و چپ، ب) استخراج ویژگی دهان، ج) استخراج ویژگی گوش چپ، د) استخراج ویژگی گوش چپ

۳-۴-۵- طبقه‌بندی رفتار راننده با ماشین بردار پشتیبان (SVM)

پس از استخراج ویژگی‌های مربوط به هر ناحیه از چهره، این ویژگی‌ها به الگوریتم ماشین بردار پشتیبان (SVM) برای طبقه‌بندی داده می‌شوند. SVM یکی از الگوریتم‌های یادگیری نظارت شده است که در دسته‌بندی داده‌ها به دسته‌های مجزا بسیار کارآمد عمل می‌کند. در اینجا، مدل SVM با استفاده از ویژگی‌های استخراج شده از روش‌های HOG و LBP، وضعیت‌های مختلف راننده را تشخیص داده و برچسب گذاری می‌کند.

۳-۴-۶- تشخیص خستگی و حواس‌پرتی راننده

با استفاده از ویژگی‌های استخراج شده و مدل SVM، وضعیت‌های مختلف راننده شناسایی شده و مشخص می‌شود که آیا راننده دچار خستگی یا حواس‌پرتی شده است یا خیر. در این بخش، کلاس‌های مختلف رفتاری راننده شامل چشم باز/ بسته، سیگار کشیدن، نگاه کردن به چپ یا راست، و صحبت کردن با تلفن همراه مورد بررسی قرار می‌گیرند. نتایج نشان می‌دهد که روش پیشنهادی قادر است با دقت بالایی وضعیت‌های مختلف را شناسایی کرده و هشدارهای لازم را ارائه دهد. در بخش نتایج عملکرد روش پیشنهادی به طور کامل بررسی و مورد بحث قرار می‌گیرد.



(الف)



(ب)

شکل ۶: نتایج LBP برای تشخیص خستگی و حواس پرتی رانندگان، (الف) نمونه‌هایی که درست تشخیص داده شده، (ب) نمونه‌هایی که اشتباه تشخیص داده شده.

الگوریتم جداسازی نواحی مختلف صورت

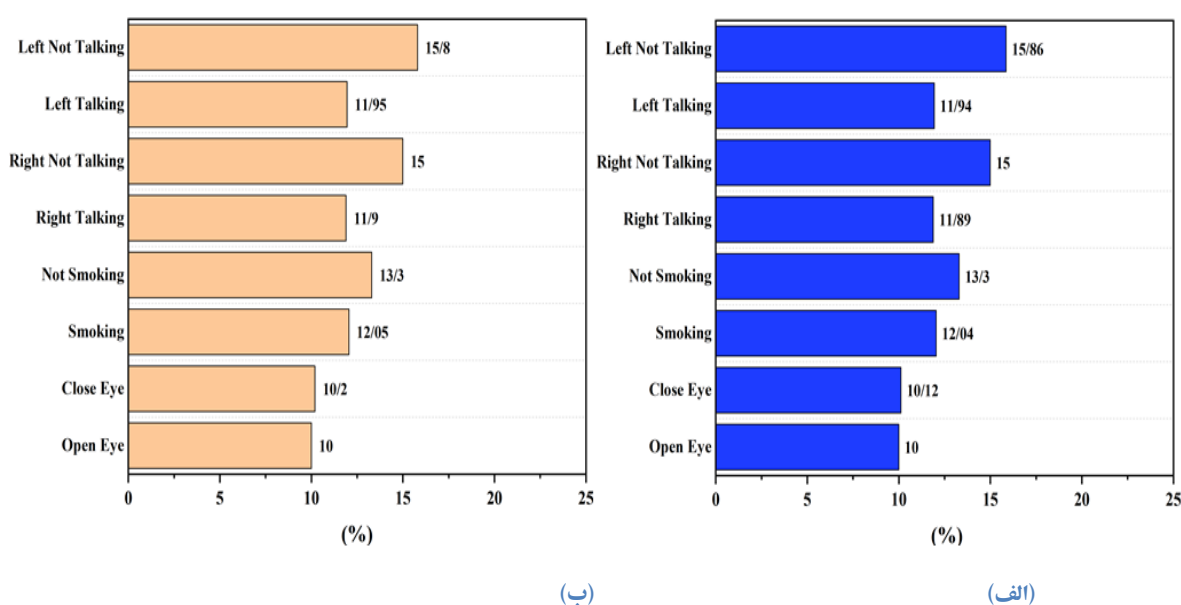
ورودی: تصویر چهره راننده

خروجی: نواحی مختلف چهره (چشم‌ها، بینی، دهان و ...)

۱. بارگذاری مدل تشخیص چهره
۲. دریافت تصویر ورودی و تبدیل آن به سطح خاکستری
۳. اجرای الگوریتم تشخیص چهره روی تصویر
۴. برای هر چهره تشخیص داده شده انجام دهید:
 - ۴.۱. استخراج ۶۸ نقطه کلیدی چهره با استفاده از مدل محلی سازی
 - ۴.۲. استخراج نواحی اصلی چهره:
 - استخراج نقاط مربوط به چشم چپ
 - استخراج نقاط مربوط به چشم راست
 - استخراج نقاط مربوط به بینی
 - استخراج نقاط مربوط به دهان
 - ۴.۳. اعمال آفست‌های مشخص برای جداسازی نواحی چهره
 - ۴.۴. ذخیره یا نمایش نواحی جداسازی شده
۵. پایان

۳-۵- مجموعه داده‌ها

برای ایجاد یک سیستم که خستگی و حواس پرتی رانندگان را تشخیص دهد، در گام نخست نیازمند داده‌هایی از راننده هستیم که منجر به حواس پرتی و خواب‌آلودگی او می‌شود. اکثر پژوهش‌ها از داده‌های منحصر به خودشان و همچنین داده‌هایی که متناسب با زاویه دید ما نبودند استفاده نموده‌اند و اکثر کلاس‌های مورد نظر آن‌ها با کلاس‌های بررسی شده‌ی ما متفاوت بود. در نهایت، مجموعه داده‌ای شامل خواب‌آلودگی، استعمال دخانیات و صحبت با تلفن همراه که با استفاده از تکنیک‌های داده‌افزایی، برهم زدن مقیاس، زاویه چرخش و موقعیت شکل گوش ارتقا پیدا کردند، فراهم گردیده است. تعداد ۲۵۰۶۸ عکس جمع‌آوری گردیده تا به کمک آنها سیستم و طبقه‌بندی بر روی این داده‌ها آموزش داده شود. این مجموعه به دو بخش آموزش و آزمایش تقسیم شده است، بخش آموزش شامل ۲۰۰۵۴ تصویر و بخش آزمایش شامل ۵۰۱۴ تصویر می‌باشد. اطلاعات مربوط به تعداد عکس‌های بخش آموزش و آزمایش به ترتیب برای هر وضعیت در شکل ۷ نشان داده شده است.



شکل ۷: نمودار توزیع داده‌های آموزش و آزمایش (الف) توزیع داده‌های آموزشی بر اساس وضعیت، (ب) توزیع داده‌های آزمایش بر اساس وضعیت.

در شکل ۷ - الف) میزان درصد هر کدام از وضعیت‌های موجود برای بخش آموزش نمایش داده شده است. همانطور که مشاهده می‌شود بیشتر داده‌ها مربوط به صحبت با تلفن همراه و کمترین داده‌ها متعلق به چشم باز و بسته می‌باشد. از آن‌جا که بخش صحبت با تلفن همراه به دو بخش گوش راست و گوش چپ تقسیم می‌شود، لذا توزیع داده‌های آزمایش در شکل مربوط بیشتر است. همچنین در شکل ۷ - ب) میزان درصد هر کدام از وضعیت‌های موجود برای بخش آموزش نمایش داده شده است. در این شکل همانند بخش قبلی، بیشتر داده‌ها مربوط به صحبت با تلفن همراه و کمترین داده‌ها متعلق به چشم باز و بسته می‌باشد. در شکل ۸ حالت‌های مختلف حواس پرتی نشان داده شده است.

حالت‌های مختلف راننده	نمونه تصاویر حالت‌های راننده
حالت عادی	
حالت خواب آلودگی	
حالت استعمال دخانیات	
حالت مکالمه با گوشی در دست چپ	
حالت مکالمه با گوشی در دست راست	

شکل ۸: حالت‌های مختلف حواس پرتی شامل استعمال دخانیات، مکالمه با تلفن همراه و خواب‌آلودگی راننده

۴- نتایج

برای بررسی کارایی روش پیشنهادی، نتایج دقت طبقه‌بندی برای دو روش استخراج ویژگی HOG + SVM و LBP + SVM مقایسه شده‌اند. جدول ۱ نشان‌دهنده دقت به دست آمده در تشخیص وضعیت‌های مختلف است. معیار دقت صحت کلاس‌ها در واقع از تقسیم تعداد کل فریم‌های صحیح مثبت و صحیح منفی بر کل فریم‌ها محاسبه شده است. سپس، از دقت تمامی چهار وضعیت میانگین گرفته شده تا دقت کل برای هر الگوریتم بدست آید. معیار دقت در معادله ۲ نشان داده شده است، که در آن TP تعداد نمونه‌های درست مثبت، TN تعداد نمونه‌های درست منفی، FP تعداد نمونه‌های اشتباه مثبت و FN تعداد نمونه‌های اشتباه منفی می‌باشد.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100 \quad (1)$$

جدول ۲: مقایسه دقت داده‌های آموزش و آزمایش در تشخیص حالت‌ها.

حالت‌ها روش‌ها	چشم باز - بسته	سیگار	گوشی - راست	گوشی - چپ	میانگین (دقت)
LBP + SVM	٪۶۱	٪۶۰	٪۶۲	٪۶۳	۶۱٪/۵
HOG + SVM	٪۸۵	٪۹۰	٪۹۰	٪۸۹	۸۸٪/۵

بر این اساس نتایج حاصل از استخراج ویژگی با استفاده از الگوریتم هیستوگرام گرادیان جهت‌دار (HOG) و الگوریتم الگوی باینری محلی (LBP) با طبقه‌بندی ماشین بردار پشتیبان (SVM) در جدول ۲ بیانگر این حقیقت است که استخراج ویژگی با استفاده از الگوریتم HOG نسبت به استخراج ویژگی الگوریتم LBP در وضعیت چشم باز یا بسته، استعمال دخانیات و صحبت با تلفن همراه به نتایج قابل قبولی دست یافته است. میانگین دقت ٪۸۸.۵ در روش استخراج ویژگی HOG با طبقه‌بندی SVM بهترین عملکرد را به دلیل توانایی بیشتر در استخراج ویژگی‌های ساختاری چهره، برای تشخیص خستگی و حواس‌پرتی در رانندگان را داشته است. روش HOG + SVM به طور کلی عملکرد بسیار بهتری نسبت به LBP + SVM دارد. دقت روش LBP + SVM در محدوده ٪۶۰ تا ٪۶۳ است، که نشان‌دهنده کارایی کمتر این روش در تشخیص حالات مختلف چهره است.

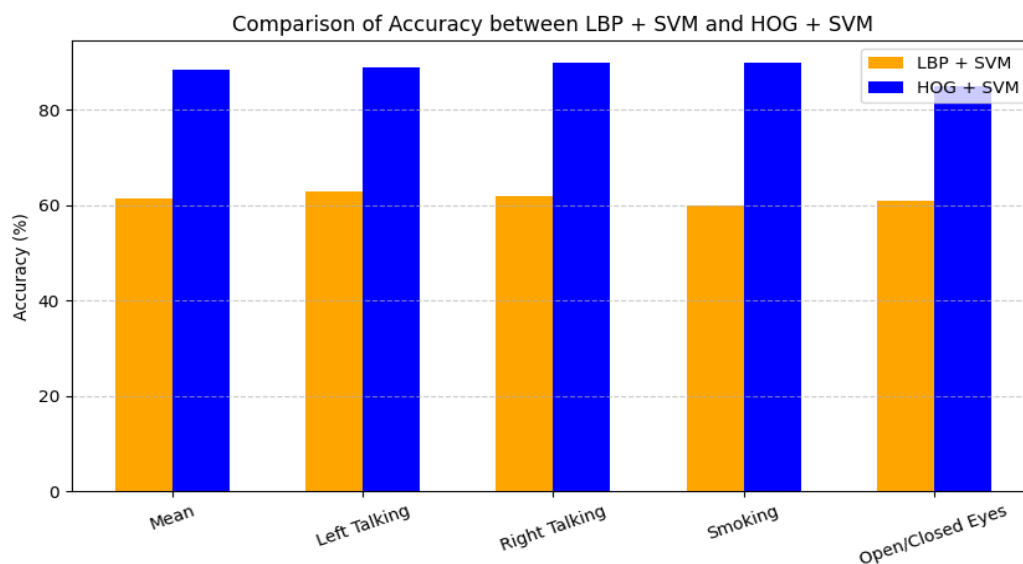
روش HOG + SVM در همه دسته‌ها دقت بالای ٪۸۵ را دارد و برای گوشی راست به دقت ٪۹۰ رسیده است. بهترین عملکرد HOG + SVM برای تشخیص استفاده از گوشی سمت راست (٪۹۰) و سیگار کشیدن (٪۹۰) بوده است. ضعیف‌ترین عملکرد LBP + SVM در تشخیص سیگار کشیدن (٪۶۰) مشاهده می‌شود. ما همچنین جدول ۳ (ارزیابی) را برای Precision (تعداد پیش‌بینی‌های درست از یک کلاس تقسیم بر کل پیش‌بینی‌های انجام شده برای آن کلاس)، Recall (تعداد پیش‌بینی‌های درست از یک کلاس تقسیم بر کل داده‌های واقعی آن کلاس)، F1-Score (میانگین هارمونیک Precision و Recall)، محاسبه نمودیم. بر اساس اطلاعات این جدول، روش HOG + SVM به طور قابل توجهی عملکرد بهتری نسبت به LBP + SVM دارد، به ویژه در تشخیص سیگار و گوشی راست که Precision و Recall بالاتری دارند. مدل LBP + SVM دقت متوسطی

دارد، اما نسبت به HOG + SVM کارایی پایین تری دارد. در هر دو مدل، تشخیص چشم باز و بسته کمترین مقدار دقت را دارد، که نشان دهنده دشواری در شناسایی این ویژگی است.

جدول ۳: مقادیر ارزیابی مدل‌ها در تشخیص حالت‌ها

میانگین	چشم باز - بسته	سیگار	گوشی راست	گوشی چپ	معیارهای ارزیابی (روش)
۰.۶۱	۰.۶۱	۰.۶۰	۰.۶۲	۰.۶۱	Precision (LBP + SVM)
۰.۶۱۵	۰.۶۱	۰.۶۰	۰.۶۲	۰.۶۳	Recall (LBP + SVM)
۰.۶۱۲	۰.۶۱	۰.۶۰	۰.۶۲	۰.۶۲	F1-Score (LBP + SVM)
۰.۸۸۵	۰.۸۵	۰.۹۰	۰.۹۰	۰.۸۹	Precision (HOG + SVM)
۰.۸۸۲	۰.۸۵	۰.۹۰	۰.۹۰	۰.۸۸	Recall (HOG + SVM)
۰.۸۸۳	۰.۸۵	۰.۹۰	۰.۹۰	۰.۸۸۵	F1-Score (HOG + SVM)

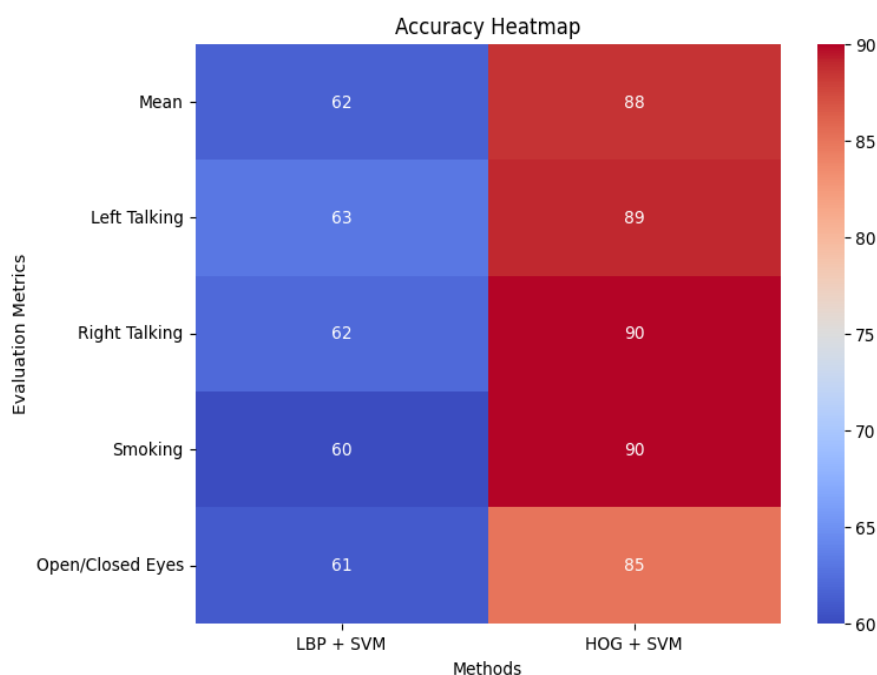
نمودار ارائه شده در شکل ۹، مقایسه‌ای بین دقت دو روش استخراج ویژگی یعنی LBP + SVM و HOG + SVM را در تشخیص حالات مختلف راننده نمایش می‌دهد. محور افقی شامل پنج وضعیت مختلف شامل، میانگین دقت، صحبت کردن با گوشی در سمت چپ، صحبت کردن با گوشی در سمت راست، سیگار کشیدن، و وضعیت باز یا بسته بودن چشم‌ها است. محور عمودی نیز مقدار دقت را به صورت درصد نمایش می‌دهد.



شکل ۹، مقایسه بین دقت دو روش استخراج ویژگی LBP + SVM و HOG + SVM

نتایج نشان می‌دهند که روش HOG + SVM (میل‌های آبی) در تمامی حالات عملکرد بهتری نسبت به LBP + SVM (میل‌های نارنجی) دارد. به طور خاص، مقدار دقت در روش HOG + SVM برای تمامی حالات نزدیک به ۹۰٪ است، در حالی که روش LBP + SVM دقتی در حدود ۶۰ تا ۶۵ درصد دارد. این اختلاف نشان می‌دهد که ویژگی‌های استخراج شده از طریق HOG (هیستوگرام گرادیان‌های جهت‌دار) حاوی اطلاعات بیشتری برای تشخیص الگوهای صورت و حالات راننده هستند و طبقه‌بند SVM

توانایی بالاتری در دسته‌بندی این ویژگی‌ها دارد. به‌ویژه در وضعیت‌های صحبت کردن با گوشی (چپ و راست) و سیگار کشیدن، روش $HOG + SVM$ دقت بالایی را ارائه کرده است، که این نشان‌دهنده قابلیت بالای این روش در تمایز دقیق ویژگی‌های مرتبط با دست و دهان در هنگام استفاده از گوشی یا سیگار است. در مقابل، روش $LBP + SVM$ احتمالاً به دلیل وابستگی به الگوهای بافتی محلی، اطلاعات مهم را در این حالات از دست داده و عملکرد ضعیف‌تری دارد. همچنین، در بخش میانگین دقت، روش $HOG + SVM$ با اختلاف حدود ۲۵٪ نسبت به $LBP + SVM$ برتری دارد، که تأییدی بر کارایی بالای آن در تشخیص تمامی حالات راننده است. همچنین شکل ۱۰ نقشه حرارتی (Heatmap) ارائه شده، دقت دو روش را در پنج معیار ارزیابی مختلف نمایش می‌دهد. محور عمودی شامل میانگین دقت، صحبت کردن با گوشی در سمت چپ، صحبت کردن با گوشی در سمت راست، سیگار کشیدن، و باز یا بسته بودن چشم‌ها است، در حالی که محور افقی نشان‌دهنده دو روش مورد مقایسه یعنی $LBP + SVM$ و $HOG + SVM$ می‌باشد.



شکل ۱۰، نقشه حرارتی مقایسه دقت دو روش در پنج معیار ارزیابی مختلف

رنگ‌بندی نقشه حرارتی به‌وضوح نشان می‌دهد که روش $HOG + SVM$ (ستون دوم) در تمامی معیارها دقت بالاتری نسبت به $LBP + SVM$ دارد. مقدار دقت برای $LBP + SVM$ در محدوده ۶۰٪ تا ۶۳٪ قرار دارد، که نسبتاً پایین است. در مقابل، مقدار دقت برای $HOG + SVM$ بین ۸۵٪ تا ۹۰٪ است که نشان از برتری قابل توجه این روش دارد. بیشترین اختلاف دقت بین دو روش در تمامی معیارها تقریباً ۲۵٪ تا ۳۰٪ است، که تأیید می‌کند روش $HOG + SVM$ ویژگی‌های غنی‌تر و متمایزتری را برای تشخیص حالات مختلف راننده استخراج می‌کند. بهترین عملکرد $HOG + SVM$ در معیار صحبت کردن با گوشی در سمت راست و سیگار کشیدن مشاهده می‌شود که دقتی معادل ۹۰٪ دارد. این نتیجه نشان می‌دهد که روش HOG قادر است ویژگی‌های مهمی از حرکات دست و دهان را که در این دو حالت تأثیرگذارند، بهتر شناسایی کند. کمترین دقت $HOG + SVM$ در معیار باز یا بسته بودن چشم‌ها دیده می‌شود (۸۵٪). این امر ممکن است به دلیل چالش‌های مربوط به تغییرات نور، زاویه دید، و ویژگی‌های چهره باشد.

در مقابل، روش $LBP + SVM$ در تمامی معیارها عملکرد ضعیف‌تری دارد و کمترین دقت آن در معیار سیگار کشیدن با مقدار ۶۰٪ ثبت شده است، که احتمالاً به دلیل ناکارآمدی LBP در استخراج الگوهای پیچیده از ویژگی‌های چهره و اشیای اطراف (مانند

سیگار) است. در جدول ۴ مقایسه‌ای میان روش‌های آزمایش شده، با سایر روش‌های تشخیص خستگی و حواس‌پرتی رانندگان انجام شده است، که در ادامه به تحلیل مقادیر جدول می‌پردازیم.

جدول ۴: مقایسه روش LBP + SVM و HOG + SVM با روش‌های اخیر در تشخیص خستگی و حواس‌پرتی رانندگان.

روش	دقت (%)	سرعت به پردازش (ms)	نیاز داده زیاد	مناسب برای سیستم‌های تعبیه شده	مقاومت در برابر تغییرات محیطی
LBP + SVM	۶۱.۵	۱۲	کم	بله	ضعیف
HOG + SVM	۸۸.۵	۱۸	متوسط	بله	متوسط
Gabor Filters + SVM [31]	۸۴.۳	۲۲	متوسط	بله	متوسط
ResNet-50 [32]	۹۵.۶	۵۵	زیاد	خیر	قوی
CNN + LSTM [33]	۹۶.۸	۶۰	خیلی زیاد	خیر	خیلی قوی
ViT (Transformers) [34]	۹۷.۵	۸۵	بسیار زیاد	خیر	خیلی قوی
MobileNetV2 [35]	۹۲.۱	۳۰	متوسط	بله	قوی
EfficientNet-B3 [36]	۹۴.۷	۴۰	زیاد	خیر	قوی
YOLOv5 [37]	۹۳.۸	۲۵	متوسط	بله	قوی

در این مقایسه، روش پیشنهادی HOG + SVM همچنان به عنوان یک گزینه سریع و کم‌هزینه با دقت ۸۸.۵٪ شناخته می‌شود که عملکرد بهتری نسبت به LBP + SVM (۶۱.۵٪) و Gabor Filters + SVM (۸۴.۳٪) دارد. این روش در پردازش بلادرنگ و سیستم‌های تعبیه شده کارایی مناسبی دارد، اما نسبت به روش‌های یادگیری عمیق از دقت کمتری برخوردار است. همچنین در برابر تغییرات شدید محیطی مانند نورپردازی متغیر و حرکات تصادفی راننده، عملکردی پایین‌تر از مدل‌های عمیق مانند CNN + LSTM و ViT دارد. در میان مدل‌های یادگیری عمیق، ResNet-50 (۹۵.۶٪)، CNN + LSTM (۹۶.۸٪) و ViT (۹۷.۵٪) دقت بالایی ارائه می‌دهند، اما به دلیل نیاز پردازشی زیاد، مناسب سیستم‌های کم‌قدرت نیستند. این امر باعث می‌شود که روش‌های مبتنی بر شبکه‌های عصبی کمتر مناسب برای پردازش بلادرنگ در خودروهای تجاری باشند، مگر اینکه از سخت‌افزارهای پیشرفته مانند GPU یا TPU استفاده شود. EfficientNet-B3 با دقت (۹۴.۷٪) و MobileNetV2 با (۹۲.۱٪)، از نظر پردازشی بهینه‌تر هستند و گزینه‌هایی مناسب برای دستگاه‌های قابل حمل و خودروهای هوشمند محسوب می‌شوند. YOLOv5 نیز به دلیل سرعت پردازش بالا (25ms) و دقت (۹۳.۸٪)، برای کاربردهای تشخیص چهره و حواس‌پرتی رانندگان در ویدئوهای بلادرنگ بسیار مناسب است. این مدل، برخلاف روش‌های دیگر، در رابطه با تشخیص چند کلاس هم‌زمان (چشم بسته، صحبت با تلفن، سیگار کشیدن و غیره) برتری دارد.

روش پیشنهادی، اگرچه از ویژگی‌های دستی استخراج شده مانند HOG + SVM بهره می‌برد، اما در مقایسه با روش‌های مبتنی بر شبکه‌های عمیق محدودیت‌هایی دارد. شبکه‌های عمیق توانایی یادگیری ویژگی‌های انتزاعی و پیچیده را دارند و با استفاده از تکنیک‌هایی مانند داده‌افزایی و یادگیری تطبیقی، عملکرد بهتری در شرایط متغیر نوری، زاویه‌های مختلف و حرکات تصادفی راننده

نشان می‌دهند. همچنین، مدل‌های ترکیبی با تنظیمات پیشرفته‌ای مانند بهینه‌سازی چندهدفی و یادگیری تطبیقی دقت تشخیص را افزایش می‌دهند. برای مثال، مدل‌های CNN-LSTM قادر به بررسی تغییرات زمانی رفتار راننده هستند، درحالی‌که روش HOG + SVM + تنها ویژگی‌های استاتیک چهره را تحلیل می‌کند. علاوه بر این، ترکیب روش‌های مختلف مانند PCA + HOG + SVM یا استفاده از چند توصیف‌گر ویژگی می‌تواند عملکرد بهتری ارائه دهد. یکی از چالش‌های دیگر روش پیشنهادی، نیاز به تنظیم دقیق پارامترها (مانند اندازه پنجره و تعداد باین‌های هیستوگرام) است، درحالی‌که شبکه‌های عمیق می‌توانند این ویژگی‌ها را به صورت خودکار بهینه‌سازی کنند. در مجموع، روش HOG + SVM یک توازن بهینه بین دقت، سرعت پردازش و هزینه محاسباتی ارائه می‌دهد. این روش می‌تواند به‌عنوان جایگزینی کم‌هزینه و سریع برای سیستم‌های هشداردهنده خستگی راننده در خودروهای هوشمند به کار رود، به‌ویژه در مواردی که منابع سخت‌افزاری محدود هستند. برای تحقیقات آینده، پیشنهاد می‌شود از مدل‌های هیبریدی که ترکیبی از شبکه‌های عمیق و روش‌های کلاسیک استخراج ویژگی هستند، استفاده کنیم، تا ضمن حفظ قابلیت تفسیرپذیری روش‌های کلاسیک، دقت و پایداری در شرایط مختلف افزایش یابد. ترکیب این روش با یادگیری عمیق هیبریدی (مانند HOG + CNN)، یک CNN سبک (مانند MobileNetV2) و یا استفاده از YOLOv5 برای تحلیل ویدیویی می‌تواند دقت را افزایش داده و نقاط ضعف آن را کاهش دهد.

به طور کلی نمودار نقشه حرارتی و اطلاعات جداول این تحقیق نشان دادند که روش HOG + SVM به‌طور قابل توجهی از روش LBP + SVM بهتر عمل می‌کند، به‌ویژه در تشخیص حالات پیچیده مانند صحبت کردن با گوشی و سیگار کشیدن. این نتایج اهمیت استفاده از ویژگی‌های گرادینانی در استخراج الگوهای چهره و حرکات را برجسته می‌کند و پیشنهاد می‌شود که برای کاربردهای مرتبط با تشخیص حواس‌پرتی راننده، روش HOG + SVM به‌عنوان گزینه بهینه در نظر گرفته شود و می‌تواند تأثیر مستقیمی بر بهبود دقت طبقه‌بندی داشته باشد.

۵- نتیجه‌گیری

تشخیص خستگی و حواس‌پرتی رانندگان یکی از چالش‌های اساسی در افزایش ایمنی جاده‌ای است. طبق گزارش سازمان امنیت ترافیک بزرگراه ملی (NHTSA) در آمریکا، حواس‌پرتی و خواب‌آلودگی یکی از عوامل تصادفات برون جاده‌ای می‌باشد. حواس‌پرتی باعث کاهش آگاهی و تصمیم‌گیری رانندگان شده، در نتیجه پاسخ به رویدادهای رانندگی با تاخیر و اختلال در سرعت همراه بوده است. به همین منظور نیاز به سیستم تشخیص خواب‌آلودگی و عوامل منجر به حواس‌پرتی احساس می‌شود که بتواند به صورت بلادرنگ در کوتاهترین زمان ممکن عامل حواس‌پرتی راننده را تشخیص و هشدار دهد. در این پژوهش، سیستمی مبتنی بر پردازش تصویر توسعه داده شد که شامل مراحل تشخیص چهره، بخش‌بندی اجزای چهره و طبقه‌بندی وضعیت راننده است. روش‌های مختلفی برای این منظور مقایسه شدند که در این میان، HOG + SVM با دقت ۸۸.۵ درصد، به‌عنوان گزینه‌ای سریع و کم‌هزینه عملکرد بهتری نسبت به LBP + SVM با دقت ۶۱.۵ درصد داشت، اما همچنان در برابر تغییرات شدید محیطی، نسبت به روش‌های مبتنی بر یادگیری عمیق محدودیت‌هایی دارد. مدل‌های عمیقی مانند ViT و CNN + LSTM دقت بالاتری ارائه می‌دهند، اما نیاز پردازشی بالای آن‌ها باعث می‌شود برای سیستم‌های تعبیه‌شده مناسب نباشند. برای تحقیقات آینده، استفاده از مدل‌های هیبریدی که ترکیبی از ویژگی‌های کلاسیک و شبکه‌های عصبی هستند، می‌تواند به بهبود دقت و پایداری سیستم در روش پیشنهادی کمک کند. پیاده‌سازی این سیستم در Python و اجرای آن در Google Colab انجام شد.

- [1] Transportation, U.S.D, "National Highway Traffic Safety Administration". 2022; Available from: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813266>
- [2] Overview of Motor Vehicle Crashes in 2020, "National Highway Traffic Safety Administration". 2018; Available from: <https://www.nhtsa.gov/campaign/distracted-driving>
- [3] H.Veeraraghavan, N. Bird, S. Atev, et al., "Classifiers for driver activity monitoring", Transportation Research Part C: Emerging Technologies, Vol. 15, Issue. (1), pp. 51-67, Jan. 2007.
- [4] A. Azman, Q. Meng, E. A. Edirisinghe, et al., "Eye and mouth movements extraction for driver cognitive distraction detection", in 2012 IEEE Business, Engineering & Industrial Applications Colloquium (BEIAC), Apr. 2012.
- [5] Q. Ji, Z. Zhu, P. Lan, Real-time nonintrusive monitoring and prediction of driver fatigue", IEEE Transactions on Vehicular Technology, Vol. 53, Issue. (4), pp. 1052-1068, Jul.2004.
- [6] F. Vicente, Z. Huang, T. Tran, et al., "Driver gaze tracking and eyes off the road detection system", IEEE Transactions on Intelligent Transportation Systems, Vol. 18, Issue. (4), pp. 837-848, Aug. 2015.
- [7] A. Doshi, M. M. Trivedi, "Head and eye gaze dynamics during visual attention shifts in complex environments", Journal of Vision, Vol. 12, Issue. (2), pp. 9-10, Feb. 2012.
- [8] Y. Abouelnaga, H. M. Eraqi, M. N. Moustafa, "Real-time distracted driver posture classification", Conference: 2018 NIPS Workshop on Machine Learning for Intelligent Transportation SystemsAt: Montreal, Canada, Vol. 20, Issue. (5), pp. 1572-1582. Dec. 2018.
- [9] B. Baheti, S. Gajre, S. Talbar, "Detection of Distracted Driver Using Convolutional Neural Network", Conference: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1-5, Jun. 2018.
- [10] M. U. Khan, T. Mahmood, A. Afsar, "Vision Transformers for Driver Distraction Detection", Neural Computing and Applications, pp. 1-15, Aug. 2022.
- [11] S. Jung, S. Jo, C. H. Im, "Brain activity classification for driver drowsiness detection using deep learning with EEG signals", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 26, Issue. (7), pp. 1239-1245, Feb. 2019.
- [12] C. T. Lin, L. W. Ko, I. F. Chung, et al., "Adaptive EEG-based alertness estimation system by using ICA-based fuzzy neural networks", IEEE Transactions on Circuits and Systems I: Regular Papers, Vol. 53, Issue. (11), pp. 2469-2476. Des. 2006.
- [13] Y. Liu, Y. Wang, C. H. Wu, et al., "Multi-modal sensor fusion for detecting driver distraction", IEEE Transactions on Intelligent Vehicles, Vol. 4, Issue. (2), pp. 179-188, Jul. 2019.
- [14] N. George, S. Ghosh, H. Zhang, et al., "A multi-modal approach to driver distraction detection using deep learning", IEEE Transactions on Intelligent Transportation Systems, Vol. 22, Issue. (3), pp. 1478-1490. Dec. 2021.
- [15] S. K. L. Lal, A. Craig, "A critical review of the psychophysiology of driver fatigue", Biological Psychology, Vol. 55 Issue. (3), pp.173-194, Feb. 2001.
- [16] M. Awais, N. Badruddin, M. Drieberg, "A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability", Sensors, Vol. 17, Issue. (9), pp. 1991, Aug. 2017.
- [17] B. T. Jap, S. Lal, P. Fischer, et al. , "Using EEG spectral components to assess algorithms for detecting fatigue", Expert Systems with Applications, Vol. 36, Issue. (2), pp. 2352-2359. Mar. 2009.
- [18] H. Singh, N. Papanikolopoulos, "Monitoring driver distraction using facial tracking and head pose estimation", IEEE Transactions on Intelligent Transportation Systems, Vol. 16, Issue. (2), pp. 636-650. Dec. 2015.
- [19] R. A. Naqvi, M. Arsalan, G. Batchuluun, et al., "Deep Learning-Based Gaze Detection System for Automobile Drivers Using a NIR Camera Sensor", Sensors, Vol. 18, Issue. (2), pp. 456, Feb. 2018.
- [20] M. Suleman, S. B. Zulifqar, N. Memon, "Real-time driver drowsiness detection using computer vision

- techniques”, International Electrical Engineering Conference, Apr. 2021.
- [21] A. Mobin, M. M. Rahman, M. A. Rahman, ” Real-time image processing-based drowsiness detection using eye aspect ratio and facial landmark points”, International Conference on Computer and Information Technology (ICCIT), pp.245-250, Jul. 2019.
- [22] M. Abdulmalik, A. Rahman, , A. A. Bakar, ” Facial monitoring system for drowsiness detection using machine learning”, Journal of Intelligent & Fuzzy Systems, Vol. 40, Issue. (1), pp.1-11, Feb. 2021.
- [23] R. Petoucha, W. Zhang, M. Rezaei, ” A computer vision-based approach to detect distracted driving behaviors”, IEEE Transactions on Intelligent Vehicles, Vol. 5, Issue. (4), pp. 512-522, Jul. 2020.
- [24] J. H. Kim, S. H. Bae, ” Smoke Detection Method Using Local Binary Pattern Variance in RGB Contrast Image”, Journal of Korea Multimedia Society, Vol. 18, Issue. (10), pp. 1197-1204, Oct. 2015.
- [25] L. Shaoji, , C. Xiong, ” Deep learning-based driver activity recognition for distraction detection”, Neural Computing and Applications, Vol. 33, Issue. (5), pp. 1749-1762, Oct. 2021.
- [26] T. Ojala, M. Pietikainen, T. Maenpaa, ” Multiresolution gray-scale and rotation invariant texture classification with local binary patterns”, IEEE Transactions on pattern analysis and machine intelligence, Vol. 24, Issue. (7), pp. 971-987, Jul. 2002.
- [27] M. Arif , KB. Khattak, K. Fiaz, et al., ” A Real-Time Driver Drowsiness Detection and Warning System Based on an Eye Blinking Rate”, in Intelligent Technologies and Applications: Second International Conference, INTAP 2019, Bahawalpur, Pakistan, pp. 106-117, May. 2020.
- [28] Adrian Rosebrock. Detect eyes, nose, lips, and jaw with dlib, OpenCV, and Python. 2021; Available from: <https://pyimagesearch.com/2017/04/10/detect-eyes-nose-lips-jaw-dlib-opencv-python/>
- [29] N. Dalal, B. Triggs. ” Histograms of oriented gradients for human detection”, in 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), Jul. 2005.
- [30] V. Kazemi, J. Sullivan. ” One millisecond face alignment with an ensemble of regression trees”, in Proceedings of the IEEE conference on computer vision and pattern recognition, Jun. 2014.
- [31] C. Liu, H. Wechsler, ”Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition”, IEEE Transactions on Image Processing, Vol. 11, Issue. (4), Apr. 2002.
- [32] K. He, X. Zhang, S. Ren, J. Sun, ”Deep Residual Learning for Image Recognition”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Dec. 2016.
- [33] J. Donahue, L. A. Hendricks, S. Guadarrama, et al., ” Long-term recurrent convolutional networks for visual recognition and description”, IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 39, Issue. (4), pp. 677 – 691, Apr. 2017.
- [34] A. Dosovitskiy, L. Beyer, A. Kolesnikov, et al., ” An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale”, International Conference on Learning Representations, Oct. 2020.
- [35] M. Sandler, A. Howard, M. Zhu, et al., ” MobileNetV2: Inverted Residuals and Linear Bottlenecks”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Dec. 2018.
- [36] M.Tan, Q. Le, ” EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks”, International Conference on Machine Learning (ICML), Vol. 97, pp. 6105-6114, May. 2019.
- [37] G. Jocher, A. Stoken, A. Chaurasia, et al., "YOLOv5 by Ultralytics." GitHub Repository.