



کویر تایر

اولین تولیدکننده تایر در ایران

پیش‌بینی ماشین در لبه‌ی خط تولید تایر

پایش بلادرنگ رفتار اپراتور با پردازش ویدئو و یادگیری ماشین

سید محمدرضا جلالیان شهری^۱ - سید هاشم محتشمی^۲ - سیدعلیرضا سیادت^۳
کارشناس واحد سیستم‌های هوشمند^۱ - مدیر واحد سیستم‌های هوشمند^۲ - مدیر اداره فوا^۳ ، کویر تایر
۰۹۳۶۶۹۳۰۱۲۴ سید محمدرضا جلالیان شهری - Mr.jalalian.ieee@gmail.com

چکیده

کیفیت تولید تایر وابستگی زیادی به دقت و نحوه عملکرد اپراتورهای خط تولید دارد و هرگونه خطای انسانی می‌تواند منجر به افزایش ضایعات و کاهش بهره‌وری شود. در این پژوهش، یک سامانه هوشمند مانیتورینگ رفتار اپراتور تایرسازی تک‌مرحله‌ای، طراحی و پیاده‌سازی شده است که بر پایه فناوری پردازش ویدئو عمل می‌کند. در این سامانه، تصاویر ویدیویی ثبت‌شده از محیط کاری اپراتورها پس از انجام پیش‌پردازش‌هایی مانند کاهش نویز و انتخاب ناحیه موردنظر، وارد مرحله تحلیل شدند. به‌منظور شناسایی حرکات و الگوهای رفتاری صحیح، از روش‌های شناخته‌شده در پردازش تصویر نظیر «جریان نوری و تخمین وضعیت بدن» برای استخراج ویژگی‌ها و از الگوریتم‌های یادگیری ماشین مانند «شبکه‌های عصبی کانولوشنی و ماشین بردار پشتیبان» برای طبقه‌بندی استفاده گردید. در مرحله بهره‌برداری، سامانه رفتار روزانه اپراتورها را با الگوی آموزش‌دیده مقایسه کرده و انحرافات احتمالی را شناسایی می‌کند. بررسی نتایج نشان داد که این سامانه با دقت بالای تشخیص، می‌تواند به عنوان ابزاری کاربردی در کاهش خطای انسانی، ارتقای کیفیت محصول و بهبود بهره‌وری خطوط تولید در صنعت تایر به کار گرفته شود.

واژه‌های کلیدی: مانیتورینگ رفتار، پردازش ویدئو، Optical Flow، CNN، خط تولید تایر، هوشمندسازی

محور مقاله: فناوری‌های تولید- هوشمند سازی فرایند تولید

مقدمه

صنعت تایرسازی به عنوان یکی از صنایع راهبردی کشور، نیازمند بالاترین سطح کیفیت در فرآیند تولید است. در این میان، عملکرد صحیح اپراتورهای خط تولید، نقشی کلیدی در کیفیت تایر تولیدی دارد، چرا که کوچک‌ترین خطای انسانی می‌تواند منجر به افزایش ضایعات و کاهش بهره‌وری شود. از این‌رو پایش رفتار اپراتورها در کنار کنترل تجهیزات، به یکی از اولویتهای اصلی مدیران تولید تبدیل شده است.

پیشرفت‌های اخیر در حوزه پردازش ویدئو و یادگیری ماشین امکان توسعه سامانه‌های هوشمند برای پایش رفتار انسانی را فراهم کرده است. در این سامانه‌ها، ابتدا ویدیوهای ثبت‌شده تحت پیش‌پردازش (کاهش نویز، برش ناحیه موردنظر و نرمال‌سازی) قرار گرفته و سپس ویژگی‌های حرکتی با روش‌هایی مانند Optical Flow استخراج می‌شوند. در ساده‌ترین حالت، جریان نوری یا Optical Flow به صورت جایجایی بردار شدت روشنایی بین دو فریم متوالی تعریف می‌شود[۱]:

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad (1)$$

که در آن $I(x, y, t)$ شدت روشنایی پیکسل در مکان و زمان مشخص است. این معادله پایه بسیاری از الگوریتم‌های تحلیل حرکت در ویدئو است.

برای طبقه‌بندی رفتار صحیح و نامصحیح اپراتور نیز از مدل‌های یادگیری ماشین مانند شبکه‌های عصبی کانولوشنی (CNN) و ماشین بردار پشتیبان (SVM) استفاده می‌شود. در حالت کلی، SVM به دنبال یافتن یک ابرصفحه (Hyperplane) است که بیشترین فاصله را بین داده‌های دو کلاس برقرار کند و این ابرصفحه با (فرمول ۲) تعریف می‌شود:

$$b + x \cdot w = f(x) \quad (2)$$

که در آن w بردار وزن و b بایاس است.

پژوهش حاضر با هدف ترکیب این ابزارهای فنی با نگاه مدیریتی طراحی شده است تا علاوه بر تشخیص خودکار انحرافات رفتاری، بتواند به کاهش خطای انسانی، افزایش کیفیت محصول و ارتقای بهره‌وری خطوط تولید تایر کمک نماید.

بخش تجربی:

فرآیند طراحی و پیاده‌سازی سامانه مانیتورینگ رفتار اپراتور تایرسازی شامل پنج گام اصلی بود که در بلوک‌دیگرام شکل ۱ نشان داده شده است.

۱. دریافت داده (Data Acquisition)

در این پژوهش از دوربین‌های مداربسته موجود در محیط خط تولید برای جمع‌آوری داده استفاده شد. انتخاب این روش، علاوه بر کاهش هزینه نصب تجهیزات جدید، امکان استفاده از زیرساخت‌های موجود کارخانه را فراهم کرد. ویدیوهای ثبت فعالیت روزانه اپراتورها، ورودی اولیه سیستم محسوب شدند.

۲. پیش‌پردازش ویدئو (Video Preprocessing)
ویدیوهای خام کارخانه دارای مشکلاتی همچون تغییرات نور محیط، نویز حرکتی و وجود بخش‌های غیرمرتبط بودند. برای آماده‌سازی داده‌ها، مراحل زیر انجام شد:

- کاهش نویز با استفاده از فیلترهای گاوسی و میانه (Gaussian & Median Filters).
- برش ناحیه موردنظر (ROI Cropping) به‌منظور تمرکز بر حرکات دست و بدن اپراتور.
- نرمال‌سازی روشنایی جهت کاهش اثر تغییرات نور محیط.

این مرحله سبب شد کیفیت داده‌ها افزایش یافته و خروجی مراحل بعدی دقیق‌تر باشد.

۳. استخراج ویژگی‌ها (Feature Extraction)

برای شناسایی رفتار اپراتور، ویژگی‌های حرکتی و وضعیتی استخراج شدند

- تحلیل حرکت بین فریم‌ها (Optical Flow)، رابطه‌ی پایه جریان نوری به صورت (فرمول ۱) می‌باشد.

این معادله تغییرات شدت روشنایی پیکسل‌ها بین دو فریم متوالی را نشان می‌دهد.

- تخمین موقعیت لندام بدن اپراتور (Pose Estimation): تخمین موقعیت لندام بدن اپراتور (مثلاً دست‌ها و سر) برای ارزیابی درست‌بودن وضعیت بدنی.
- هیستوگرام حرکت: (Motion Histogram) شمارش و تحلیل شدت حرکات در بازه‌های زمانی کوتاه.

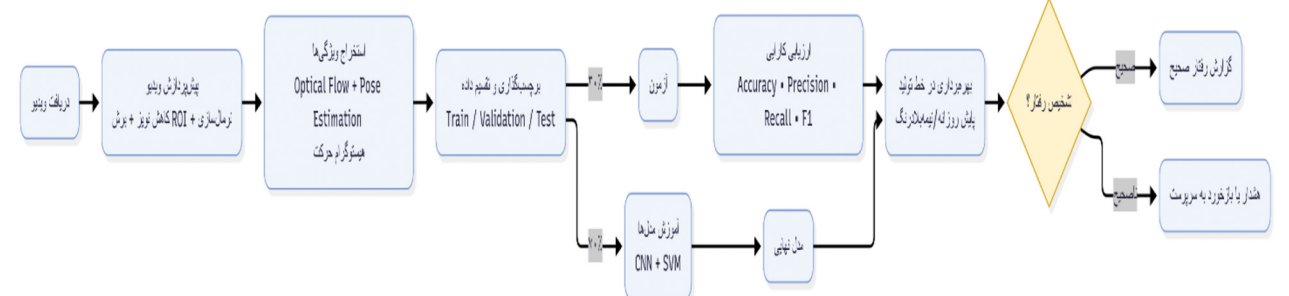
۴. آموزش مدل (Model Training)

برای طبقه‌بندی رفتار صحیح و نامصحیح، دو مدل مورد استفاده قرار گرفت:

- شبکه عصبی کانولوشنی (CNN): مناسب برای یادگیری مستقیم از توالی فریم‌ها و الگوهای تصویری.
 - ماشین بردار پشتیبان (SVM): مناسب برای دسته‌بندی داده‌های ویژگی‌محور.
- مدل‌ها با داده‌های برچسب‌گذاری‌شده آموزش دیدند. داده‌ها به نسبت ۷۰٪ آموزش و ۳۰٪ آزمون تقسیم شدند. مدل SVM ابرصفحه‌ای به شکل (فرمول ۲) برای جداسازی کلاس‌ها ایجاد کرد.

۵. ارزیابی و بهره‌برداری (Evaluation & Deployment)

در مرحله نهایی، سامانه با ویدیوهای جدید تست شد و معیارهایی مانند دقت (Accuracy)، بازخوانی (Recall) و دقت مثبت (Precision) اندازه‌گیری شدند. نتایج اولیه نشان داد که مدل CNN با دقت بالاتر از ۹۰٪ عملکرد بهتری نسبت به SVM داشت. سامانه به‌صورت پایلوت در کارخانه پیاده‌سازی و برای پایش روزانه اپراتورها به کار گرفته شد. این قابلیت علاوه بر شناسایی انحرافات رفتاری، امکان ارائه گزارش مدیریتی و هشدار فوری به سرپرستان خط را فراهم کرد. بلوک دیگرام خلاصه کار انجام شده در (شکل ۱) مشاهده می‌گردد.



شکل ۱- بلوک دیگرام روش پیشنهادی

نتایج و بحث:

به‌منظور ارزیابی سامانه پیشنهادی، داده‌های ویدیویی چندین اپراتور در بازه‌های زمانی مختلف مورد آزمایش قرار گرفت. معیارهای دقت (Accuracy)، بازخوانی (Recall)، دقت مثبت (Precision) و F1-Score برای دو مدل CNN و SVM محاسبه شد.

جدول ۱- مقایسه عملکرد SVM و CNN

الگوریتم	دقت (Accuracy)	بازخوانی (Recall)	دقت مثبت (Precision)	F1-Score
CNN	٪۹۱	٪۸۹	٪۹۲	٪۹۰
SVM	٪۸۷	٪۸۵	٪۸۸	٪۸۶

همان‌طور که (جدول ۱) نشان می‌دهد، مدل CNN توانسته عملکرد بهتری نسبت به SVM ارائه دهد، به‌ویژه در دقت مثبت که نشان‌دهنده توانایی صحیح در شناسایی رفتار درست اپراتور است.

جدول ۲- نتایج به تفکیک اپراتورها

اپراتور	دقت (Accuracy)	Precision	Recall
اپراتور ۱	٪۹۳	٪۹۴	٪۹۱
اپراتور ۲	٪۸۹	٪۹۰	٪۸۸
اپراتور ۳	٪۸۶	٪۸۵	٪۸۷

نتایج جدول ۲ نشان می‌دهد که دقت مدل در تشخیص رفتار برخی اپراتورها (مانند اپراتور ۱) بالاتر است، در حالی که در دیگر موارد (اپراتور ۳) به دلیل تفاوت‌های حرکتی و شرایط کاری، دقت کمی کاهش یافته است. این موضوع نشان می‌دهد سامانه نیازمند داده‌های متنوع‌تری برای افزایش تعمیم‌پذیری است.

جدول ۳- زمان پاسخ‌گویی سامانه

مدل	زمان پردازش هر فریم (ms)	توان پردازش (FPS)
CNN	۳۸ ms	~۲۶ FPS
SVM	۲۲ ms	~۴۵ FPS

نتایج جدول ۳ نشان می‌دهد اگرچه مدل SVM سرعت پردازش بیشتری دارد، اما به دلیل دقت پایین‌تر، مدل CNN انتخاب بهتری برای کاربرد واقعی است. با این حال، در نسخه‌های آینده می‌توان با بهینه‌سازی ساختار CNN یا استفاده از GPU، سرعت آن را بهبود بخشید.

جدول ۴- ماتریس سردرگمی مدل CNN

	پیش‌بینی: صحیح	پیش‌بینی: نامصحیح
واقعی: صحیح	۲۷۵	۱۸
واقعی: نامصحیح	۲۲	۱۸۵

نتایج جدول ۴ نشان می‌دهد مدل CNN توانسته است بخش عمده رفتار صحیح اپراتورها را درست شناسایی کند (۲۷۵ مورد از

۱۷th Conference And 9th Exhibition Rubber And Related Industries

Iranmall January 6-7 2026

همایش ۹^{ام} نمایشگاه لاستیک و صنایع وابسته

ایران مال ۱۸ تا ۲۱ دی ماه ۱۴۰۴

۲۹۳). همچنین در شناسایی رفتار نامصحیح عملکرد قابل قبولی داشته است (۱۸۵ مورد از ۲۰۷). خطاها (۲۲ و ۱۸ مورد) عمدتاً ناشی از شرایط نوری و حرکات سریع بوده‌اند.

در شکل ۲ نمونه‌ی خروجی پردازش ویدئو مشاهده می‌گردد. آنچه دیده می‌شود یک نمای دوبخشی از همان فریم ویدئو است:

- سمت چپ: فریم خام محیط کاری اپراتور (بدون پردازش).
- سمت راست: همان فریم با خروجی Pose Estimation نمایش نقاط کلیدی و اتصالات اسکلتی اندام فوقانی و تنه.

این تصویر برای کنترل کیفی برچسب‌ها، ارزیابی وضعیت بدنی اپراتور و استخراج ویژگی‌های وضعیتی/زاویه‌ای در فرآیند آموزش مدل‌ها استفاده شده است

کیفیت تشخیص نقاط: تعداد نقاط شناسایی‌شده و پیوستگی اتصالات نشان می‌دهد الگوریتم در این شرایط نوری/زاویه‌ی دوربین پایدار است.

وضعیت بدنی اپراتور: راستای شانه‌ها و آرنج/لچ در نسبت با میز کار یا ابزار قابل ارزیابی‌اند. در فاز تحلیل می‌توان زاویه‌ی مفاصل را از نقاط کلیدی محاسبه کرد (مثلاً زاویه آرنج با سه نقطه‌ی شانه-آرنج-مچ).

کاربری در آموزش/ارزیابی: این تصویر به‌عنوان نمونه‌ی مثبت «وضعیت صحیح» یا نمونه‌ی منفی «وضعیت نامصحیح» در دیتاست استفاده می‌شود و برای بازبینی انسانی (کیفیت‌سنجی برچسب‌ها) مناسب است.

محدودیت‌ها: در پوشیدگی (occlusion) یا محوشدگی حرکت ممکن است چند نقطه از بین بروند.



شکل ۲- نمونه‌ی خروجی پردازش ویدئو در ایستگاه تایرسازی تک‌مرحله‌ای 507. (چپ) فریم خام، (راست) خروجی Pose Estimation با نمایش نقاط کلیدی و اتصالات اسکلتی.

جمع‌بندی

نتایج نشان داد مدل CNN در شاخص‌های Precision و F1 از SVM بهتر است و برای کاربرد صنعتی پایش اپراتور مناسب‌تر است؛ این یعنی کاهش هشدارهای اشتباه و اعتمادپذیری بیشتر برای سرپرستان خط. تفاوت عملکرد بین اپراتورها (جدول ۲) عمدتاً از زاویه دوربین، شرایط نوری و پوشیدگی دست‌آ بدن ناشی می‌شود و با تنوع‌دادن داده‌ی آموزشی و نرمال‌سازی روشنایی قابل بهبود است. از نظر فنی، تلفیق ویژگی‌های حرکتی مثلاً Optical Flow های به‌روزتر مانند RAFT با نشانه‌های وضعیتی بدن (BlazePose/MoveNet) و یک مدل زمانی سبک (SlowFast/TSM) می‌تواند هم دقت و هم پایداری را بالا ببرد [۱و۵].

همچنین مرورهای اخیر نشان می‌دهند مدل‌های سبک جدید مثل (MoveNet) برای بلادرنگ روی لبه مناسب‌اند و با داده‌های متنوع، تعمیم‌پذیری بهتری دارند [۶]

سامانه‌ی پیشنهادی با تکیه بر دوربین‌های موجود کارخانه و تحلیل خودکار ویدئو، ابزار مؤثری برای کاهش خطای انسانی، یکنواختی اجرای فرایند و بهبود کیفیت است. برای استقرار گسترده، پیشنهاد می‌شود: (۱) استانداردسازی زوایان دوربین، (۲) گزارش‌های دوره‌ای برای آموزش هدفمند اپراتورها، (۳) بهینه‌سازی محاسبات GPU یا مدل‌های سبک جهت بلادرنگ پایداری.

[1] Z. Teed and J. Deng, "RAFT: Recurrent All-Pairs Field Transforms for Optical Flow," ECCV, 2020. ECVA
[2] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, "BlazePose: On-device Real-time Body Pose Tracking," arXiv:2006.10204, 2020. arXiv:1
[3] R. Völz and N. Li, "Next-Generation Pose Detection with MoveNet and TensorFlow.js," TensorFlow Blog, May 2021; see also "MoveNet on TF-Hub," 2024. TensorFlow Blog/TensorFlow
[4] C. Feichtenhofer, H. Fan, J. Malik, and K. He, "SlowFast Networks for Video Recognition," ICCV, 2019. CVF Open Access
[5] J. Lin et al., "TSM: Temporal Shift Module for Efficient Video Understanding," ICCV, 2019. CVF Open Access
[6] F. Roggio et al., "A Comprehensive Analysis of Machine-Learning Pose Estimation Models," Sensors (open-access via PMC), 2024.