

Advancements in Video-Based Human Activity Recognition: A Comprehensive Review of Methods, Challenges, and Applications

Mohammad Sabouri *

Department of Informatics, Bioengineering,
Robotics and Systems Engineering, University of
Genoa: Genoa, IT.

Abstract

Human Activity Recognition (HAR) using video data has emerged as a pivotal research domain due to its transformative potential in areas such as surveillance, healthcare monitoring, and human-computer interaction. This technology aims to automatically recognize and classify human behaviors and movements from sequences of images or videos by utilizing computer vision and machine learning techniques. However, challenges such as the diversity of human movements, lighting conditions, camera angles, frame rates, image quality and the need for real-time processing have complicated the development of effective systems in this field. This review provides a detailed examination of current HAR methodologies, highlighting both unimodal and multimodal approaches, and assessing their strengths and limitations. Particular emphasis is placed on real-time processing, which is critical for practical deployment in dynamic environments. Furthermore, this study explores a diverse set of real-world applications, discusses prevailing challenges—such as occlusions, noisy data, and computational constraints—and proposes considerations for future development. By addressing key obstacles and mapping out current trends, this paper aims to guide researchers and practitioners toward more robust, adaptive, and scalable HAR solutions.

Keywords: video colorization, deep learning diffusion models, evaluation metrics

Received: 04/April/2025

Accepted: 27/August/2025

eISSN: 3060-6144

ISSN: 2980-8936

* Corresponding Author: m.sabouri@shirazu.ac.ir

بررسی جامع روش‌های شناسایی فعالیت‌های انسانی از روی پردازش ویدیو

محمد صبوری* | دپارتمان رباتیک دانشگاه جنوا - ایتالیا.

چکیده

شناسایی فعالیت انسانی (Human Activity Recognition - HAR) از داده‌های ویدیویی، به یکی از حوزه‌های پرکاربرد و در حال رشد در علوم رایانه تبدیل شده است که در زمینه‌هایی چون نظارت امنیتی، پایش سلامت، تحلیل رفتار و تعامل انسان و رایانه نقش کلیدی ایفا می‌کند. این فناوری تلاش می‌کند تا با بهره‌گیری از تکنیک‌های بینایی ماشین و یادگیری ماشین، رفتارها و حرکات انسان را به صورت خودکار از روی توالی تصاویر یا ویدیوها شناسایی و طبقه‌بندی کند. با این حال، چالش‌هایی نظیر تنوع در حرکات انسانی، شرایط نوری، زوایای دوربین، نرخ فریم، کیفیت تصویر و نیاز به پردازش بلادرنگ، توسعه سامانه‌های مؤثر در این حوزه را پیچیده ساخته است. این مقاله مروری، به بررسی جامع روش‌های متداول در HAR شامل روش‌های تک‌وجهی و چندوجهی و همچنین، نقاط قوت و ضعف هر کدام می‌پردازد. در ادامه، کاربردهای عملی این فناوری در حوزه‌های مختلف از جمله امنیت و سلامت مورد تحلیل قرار گرفته و در نهایت، مسیرهای پژوهشی آینده و توصیه‌هایی برای محققان و توسعه‌دهندگان این حوزه ارائه می‌گردد. هدف این مطالعه، فراهم‌سازی چشم‌اندازی روشن از وضعیت فعلی و فرصت‌های پیش رو در حوزه شناسایی فعالیت انسانی مبتنی بر ویدیو است.

کلیدواژه‌ها: شناسایی فعالیت انسانی، بینایی ماشین، یادگیری ماشین، تحلیل ویدیویی، پردازش بلادرنگ، تعامل انسان و رایانه، پایش سلامت، نظارت امنیتی

۱- مقدمه

شناسایی فعالیت انسانی (Human Activity Recognition یا HAR) از طریق داده‌های ویدیویی در سال‌های اخیر توجه بسیاری از پژوهشگران را به خود جلب کرده است. این تکنولوژی به دلیل قابلیت‌های گسترده‌ای که در حوزه‌هایی مانند نظارت امنیتی، مراقبت‌های بهداشتی و تعامل انسان و رایانه دارد، به یکی از موضوعات برجسته در حوزه بینایی ماشین و یادگیری ماشین تبدیل شده است. فرآیند شناسایی و طبقه‌بندی رفتارها و حرکات انسانی از روی ویدیو، به دلیل پیچیدگی ذاتی و تنوع بالای حرکات انسانی، همواره یک چالش فنی محسوب می‌شود (Poppe, 2010).

با ظهور پیشرفت‌های قابل توجه در یادگیری عمیق، حوزه شناسایی فعالیت انسانی مبتنی بر ویدیو به سرعت در حال رشد است و نقش مهمی در صناعی همچون خانه‌های هوشمند، سامانه‌های پایش سلامت و سیستم‌های امنیتی ایفا می‌کند. این سامانه‌ها با تحلیل خودکار حرکات و رفتارهای انسانی، کاربردهای متنوع و روزافزونی یافته‌اند. یکی از الزامات کلیدی برای اثربخشی این سامانه‌ها در شرایط واقعی، پردازش بلادرنگ است که به سیستم‌ها امکان می‌دهد رفتار انسان‌ها را در زمان واقعی پایش و تحلیل کنند.

این مقاله مروری با هدف ارائه یک چشم‌انداز جامع و عمیق از وضعیت فعلی حوزه شناسایی فعالیت انسانی تنظیم شده و محورهای زیر را دنبال می‌کند:

مرور مطالعات پیشین: ارائه خلاصه‌ای دقیق از تحقیقات پیشین در زمینه شناسایی فعالیت انسانی مبتنی بر ویدیو، به همراه معرفی مطالعات کلیدی، روش‌شناسی‌های رایج و یافته‌های برجسته.

بررسی کاربردها: تحلیل کاربردهای متنوع شناسایی فعالیت انسانی مبتنی بر ویدیو از جمله در حوزه‌های سلامت، تحلیل عملکردهای ورزشی، امنیت و دیگر حوزه‌های صنعتی و بررسی روند تکامل این کاربردها در طول زمان.

شناسایی چالش‌ها: معرفی و بررسی چالش‌ها و محدودیت‌های موجود در رویکردهای فعلی شناسایی فعالیت انسانی مبتنی بر ویدیو، از جمله مسائلی همچون دقت، مقیاس‌پذیری و پیاده‌سازی در محیط‌های واقعی.

ارائه راهکار و توصیه‌ها: ارائه پیشنهادها و راهنمایی‌های عملی برای پژوهشگران و توسعه‌دهندگان که قصد دارند در این زمینه به فعالیت تحقیقاتی یا کاربردی بپردازند.

با توجه به رشد روزافزون و کاربردهای حیاتی شناسایی فعالیت انسانی مبتنی بر ویدیو در آینده نزدیک، مطالعات جامع و مروری مانند این مقاله می‌توانند نقش مهمی در هدایت پژوهش‌ها و پروژه‌های عملی ایفا کنند.

۲- بررسی کارهای پیشین

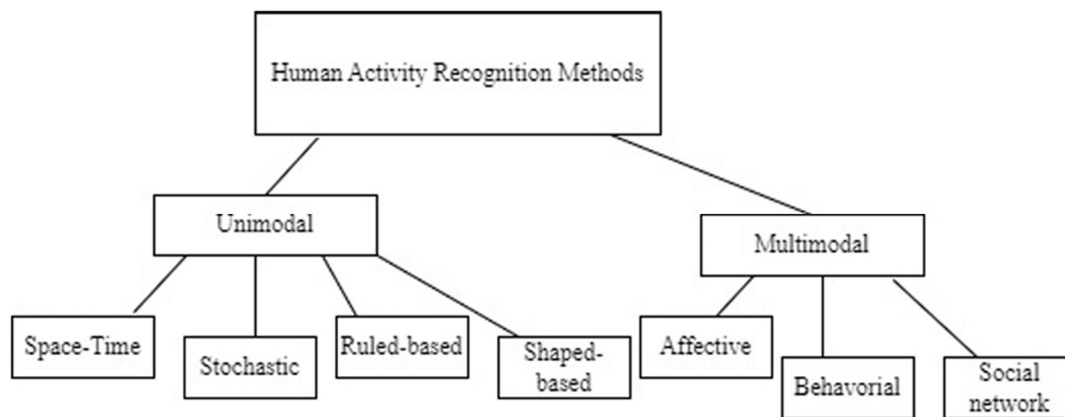
شناسایی فعالیت انسانی مبتنی بر ویدیو یک روش رایجی است که با استفاده از توالی تصاویر، به صورت خودکار فعالیت‌های انسانی را شناسایی و طبقه‌بندی می‌کند. پیشرفت‌های اخیر در حوزه بینایی ماشین و یادگیری ماشین، موجب رشد قابل توجه این حوزه شده‌اند و کاربردهای آن در نظارت تصویری، پایش سلامت و تعامل انسان و رایانه به وضوح قابل مشاهده است (Siddiqi et al., 2014). با این حال، سامانه‌های شناسایی فعالیت انسانی مبتنی بر ویدیو در شناسایی و دسته‌بندی مؤثر رفتارها با چالش‌هایی جدی مواجه‌اند که ناشی از پیچیدگی حرکات انسانی است. عواملی همچون شرایط نوری، زاویه دوربین و عناصر محیطی، باعث افزایش پیچیدگی فرایند شناسایی می‌شوند و عملکرد سیستم را تحت تأثیر قرار می‌دهند.

پردازش بلادرنگ یکی از چالش‌های کلیدی در این زمینه است؛ زیرا نیازمند تحلیل سریع داده‌های ویدیویی ورودی است. تنوع و پیچیدگی حرکات انسانی که می‌تواند تحت تأثیر عواملی مانند سن، توانایی‌های جسمی، فرهنگ

و عادات قرار گیرد، نیاز به سیستم‌های شناسایی فعالیت انسانی مبتنی بر ویدیو منعطف و تطبیق‌پذیر را دوچندان می‌کند تا بتوانند تفسیری دقیق از حرکات ارائه دهند (Sharma et al., 2022).

کیفیت داده‌های تصویری مورد استفاده نیز تأثیر مستقیمی بر دقت شناسایی دارد. ویدیوهایی با وضوح پایین یا نویز زیاد، مانع از استخراج صحیح ویژگی‌ها و تمایز بین فعالیت‌ها می‌شوند (Pareek & Thakkar, 2020). نرخ فریم ویدیو نیز برای ثبت پویایی زمانی حرکات انسانی حیاتی است و نرخ فریم ناکافی می‌تواند منجر به از بین رفتن اطلاعات و کاهش توانایی در شناسایی حرکات ظریف شود (Sheng et al., 2023)؛ بنابراین، استفاده از ویدیوهای باکیفیت و نرخ فریم مناسب، برای دستیابی به دقت بالا در شناسایی فعالیت انسانی مبتنی بر ویدیو ضروری است.

پردازش بلادرنگ در کاربردهای امنیتی و سلامت، از اهمیت بالایی برخوردار است. شناسایی سریع فعالیت‌های مشکوک در محیط‌های امنیتی و تشخیص زودهنگام ناهنجاری‌ها در حوزه سلامت، نیازمند واکنش‌های سریع است؛ اما این نوع پردازش‌ها با پیچیدگی محاسباتی همراه هستند و توسعه سیستم‌های شناسایی فعالیت انسانی مبتنی بر ویدیو مؤثر و پاسخ‌گو، مستلزم الگوریتم‌هایی بهینه است که بین بهره‌وری محاسباتی و دقت بلادرنگ تعادل ایجاد کنند. علاوه بر این، شناسایی فعالیت‌های انسانی برای دستیابی به بینش‌های رفتاری و ارتباطات انسانی نیز اهمیت دارد (Vrigkas et al., 2015). با وجود این، شناسایی فعالیت‌ها از دنباله‌های ویدیویی یا تصاویر، هنوز با چالش‌هایی نظیر شلوغی پس‌زمینه و انسداد جزئی سوژه‌ها روبه‌رو است.



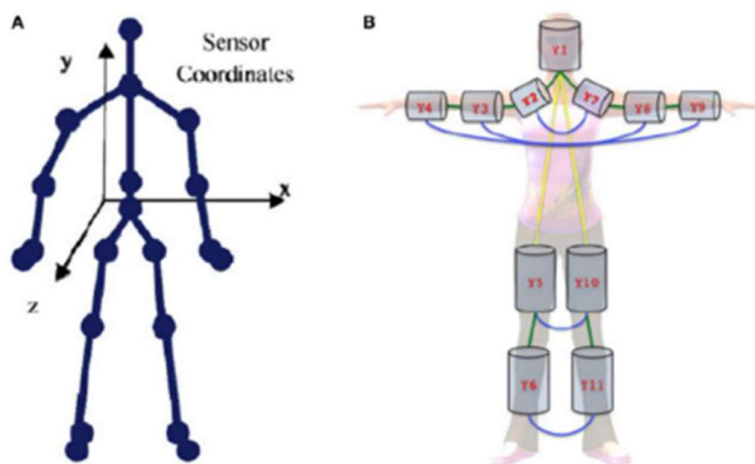
شکل ۱. روش‌های فعلی شناسایی فعالیت انسانی

در شکل ۱، روش‌های فعلی شناسایی فعالیت انسانی نمایش داده شده‌اند. این روش‌ها به دو دسته اصلی تقسیم می‌شوند: تک‌وجهی (Unimodal) و چندوجهی (Multimodal). هر کدام از این دسته‌ها شامل زیرروش‌های متعددی هستند. در این بخش، به بررسی جامع هر یک از این روش‌ها می‌پردازیم.

۱. تک‌وجهی: در روش‌های تک‌وجهی HAR، شناسایی فعالیت‌ها تنها بر اساس یک نوع داده (معمولاً داده

تصویری یا ویدیویی) انجام می‌گیرد. این تکنیک‌ها معمولاً با استخراج ویژگی‌های بصری از تصاویر یا ویدیوها و استفاده از الگوریتم‌های طبقه‌بندی، نوع فعالیت را تشخیص می‌دهند (Kong et al., 2014). با اینکه این روش‌ها برای تشخیص فعالیت‌های مبتنی بر حرکت مناسب‌اند؛ اما شناسایی فعالیت صرفاً بر اساس حرکت چالش‌برانگیز است (Vrigkas et al., 2015). بزرگ‌ترین مشکل در اینجا حفظ پیوستگی حرکات در جریان ویدیو است، به‌ویژه زمانی که فعالیت به‌صورت مداوم یا غیر یکپارچه رخ می‌دهد. این دسته شامل چهار روش اصلی است: فضای-زمانی، احتمالاتی، مبتنی بر قواعد و مبتنی بر شکل.

۲. روش فضای-زمانی: این روش بر اساس ویژگی‌های فضا-زمان یا تطابق مسیر حرکت فعالیت‌ها را شناسایی می‌کند. فرض بر این است که فعالیت در یک فضای سه‌بعدی فضا-زمان تعریف می‌شود که حاصل اتصال فریم‌های دوبعدی در طول زمان است (Wang et al., 2013). این روش‌ها مجموعه‌ای از ویژگی‌ها یا مسیرهای حرکتی را از ویدیو استخراج می‌کنند؛ اما نسبت به نویز و انسداد حساس هستند و برای فعالیت‌های پیچیده مناسب نیستند (Beddiar et al., 2020).
۳. روش احتمالاتی: تکنیک‌هایی مانند مدل‌های مخفی مارکوف (HMM) و میدان‌های تصادفی شرطی پنهان (HCRF) برای شناسایی فعالیت با استفاده از ویژگی‌هایی مانند موقعیت، سرعت و حرکت به کار رفته‌اند (Mendoza & Blanca 2008). این روش‌ها رفتار انسان را به صورت دنباله‌ای تصادفی مدل می‌کنند. اگرچه HMM برای فعالیت‌های ساده مناسب است؛ اما فرض استقلال شرطی بین ویژگی‌ها همیشه برقرار نیست. CRF دقت بالاتری دارد؛ ولی نیازمند محاسبات بیشتری است (Vrigras et al., 2015).
۴. روش مبتنی بر قواعد: در این روش، هر فعالیت به صورت مجموعه‌ای از قواعد یا ویژگی‌های ابتدایی مدل می‌شود که رویداد را توصیف می‌کنند. اگرچه این روش امکان ساخت مدل توصیفی خوبی را برای فعالیت انسانی فراهم می‌کند، توسعه قواعد و تحلیل ویدیوهای طولانی یا فعالیت‌های پیچیده در آن چالش‌برانگیز است (Beddiar et al., 2020).
۵. روش مبتنی بر شکل: در این روش، اجزای بدن انسان به صورت دقیق از ویدیو استخراج شده و در حالت دوبعدی به شکل نواحی مستطیلی و در سه‌بعدی به شکل ساختارهای حجمی مدل می‌شوند. هر فعالیت با استفاده از هیستوگرام‌های وضعیت بدن (pose primitives) توصیف می‌شود (Vrigras et al., 2015). با وجود پیشرفت در ابزارهای تخمین وضعیت بدن، اجرای بلادرنگ این روش‌ها هنوز به دلیل انسداد، شلوغی پس‌زمینه و تغییرات نوری یا لباس دشوار است (Beddiar et al., 2020).



شکل ۲. استخراج اسکلت بدنی و ویژگی‌های بدنی

۶. چندوجهی: در روش‌های HAR چندوجهی، از ترکیب چند نوع داده (نه فقط تصویر) برای شناسایی فعالیت‌ها استفاده می‌شود (Sun et al., 2009). این داده‌ها می‌توانند از حسگرهایی نظیر دوربین RGB-D، حسگرهای مادون قرمز، دوربین‌های حرارتی، حسگرهای حرکتی و غیره به دست آیند. استفاده از چند منبع داده باعث

می‌شود مشکلاتی مانند نوسانات نوری، انسداد یا شلوغی پس‌زمینه بهتر مدیریت شوند و دقت شناسایی افزایش یابد (Yadav et al., 2021).

۷. روش عاطفی: رایانش عاطفی^۱ تلاش می‌کند احساسات انسان را از طریق ویژگی‌هایی مانند حالات چهره، حرکات بدن، تغییرات فیزیولوژیکی و گفتار مدل‌سازی کند. یکی از چالش‌های اصلی در این حوزه، برچسب‌گذاری دقیق داده‌های احساسی به دلیل تفاوت‌های فردی و وجود احساسات هم‌زمان است (Liu et al., 2011). تکنیک‌های مورد استفاده شامل مدل‌های ترکیبی، رگرسیون و یادگیری عمیق برای ادغام حالت‌های مختلف و استخراج ویژگی‌های مرتبط هستند. شکاف معنایی بین ویژگی‌های سطح پایین و مفاهیم سطح بالای احساسی، فرآیند طبقه‌بندی را دشوار می‌سازد.

۸. روش رفتاری: شناسایی رفتارهای انسانی از دنباله‌های ویدیویی، روشی پیچیده در بینایی ماشین است که برای تحلیل وضعیت روانی، شخصیتی یا اجتماعی افراد به کار می‌رود (Candamo et al., 2010). این روش‌ها با ترکیب ویژگی‌های بصری و صوتی، تلاش می‌کنند احساسات و حالات ذهنی افراد را درک کنند. انتخاب ویژگی مناسب در این روش‌ها نیاز به آزمون و خطای زیادی دارد و معمولاً از ترکیب اطلاعات بصری و عاطفی استفاده می‌شود (Vrigkas et al., 2015).

۹. شبکه اجتماعی: تعاملات اجتماعی که نقش مهمی در زندگی روزمره دارند، نیازمند تطبیق رفتار بر اساس حضور سایر افراد هستند. این تعاملات در زمینه‌هایی مانند رویدادهای اجتماعی و ورزشی قابل بررسی هستند و معمولاً با استفاده از گراف‌ها، اطلاعات زمینه‌ای و ویژگی‌های چندوجهی مدل می‌شوند. یادگیری عمیق نیز در تشخیص احساسات و رویدادهای اجتماعی نقش مؤثری ایفا کرده است (Vrigkas et al., 2015). با این حال، مدل‌سازی دقیق این تعاملات به دلیل مشکلاتی مانند انسداد، الگوهای حرکتی متقابل و تعداد متغیر افراد مشارکت‌کننده، همچنان پیچیده باقی مانده است (Beddiar et al., 2020).

۳- ابزارهای مفید حوزه شناسایی فعالیت انسانی

در جدول ۱، یک تحلیل جامع از ابزارهای مختلف شناسایی فعالیت انسانی ارائه شده است. هر یک از این ابزارها دارای ویژگی‌ها، مزایا و محدودیت‌های خاص خود هستند که بسته به نیازهای پروژه، می‌توان از آن‌ها بهره برد.

۱. MediaPipe ابزاری است که قابلیت تخمین وضعیت بدن به صورت بلادرنگ در حالت دویبعدی و سه‌بعدی را فراهم می‌کند و برای کاربردهایی نظیر واقعیت افزوده و ردیابی حرکات دست و چهره بسیار مناسب است. با این حال، پشتیبانی آن در پلتفرم‌های غیر گوگلی محدود است.

۲. OpenPose به‌طور خاص در شناسایی نقاط کلیدی چندین نفر به صورت هم‌زمان و تشخیص فعالیت‌های انسانی عملکرد بسیار خوبی دارد؛ اما برای اجرای بهینه به سخت‌افزارهای قدرتمند (مانند GPUهای پیشرفته) نیاز دارد.

۳. PaddleDetection مجموعه‌ای متنوع از قابلیت‌ها مانند شناسایی اشیاء، سگمنتیشن نمونه‌ای، سگمنتیشن معنایی و تشخیص چهره را ارائه می‌دهد؛ اما استفاده از آن مستلزم آشنایی با فریم‌ورک PaddlePaddle است.

۴. ActionAI ابزارهای اختصاصی برای شناسایی حرکات و تحلیل رفتار ارائه می‌دهد؛ اما اطلاعات محدودی از آن به صورت عمومی در دسترس است و نیازمند کسب مجوز تجاری است.

۵. OpenCV به عنوان یک راهکار متن باز، قابلیت‌های گسترده‌ای برای پردازش تصویر و ویدیو در اختیار توسعه‌دهندگان قرار می‌دهد؛ اما در مقایسه با فریم‌ورک‌های تخصصی، از پشتیبانی محدودتری در زمینه یادگیری عمیق برخوردار است.

این مقایسه به پژوهشگران و توسعه‌دهندگان کمک می‌کند تا با در نظر گرفتن نیازهای خاص پروژه خود، ابزار مناسب در حوزه شناسایی فعالیت انسانی را با دقت انتخاب کنند.

۴- کاربردهای شناسایی فعالیت انسانی

کاربردهای شناسایی فعالیت انسانی

۱. نظارت ویدیویی: در فضاهای عمومی، شناسایی فعالیت انسانی می‌تواند برای شناسایی و جلوگیری از جرائمی مانند دزدی، خرابکاری، خشونت و تروریسم مورد استفاده قرار گیرد. به عنوان مثال، پژوهش تیدا و همکاران^۱ (۲۰۱۳) سیستمی مبتنی بر شناسایی فعالیت انسانی پیشنهاد داده است که قادر به تشخیص رفتارهای مشکوک در محیط‌های شلوغ مانند پرسه‌زنی، درگیری و دویدن است. این سیستم می‌تواند مقامات مسئول را مطلع کرده و اطلاعاتی مانند موقعیت و توصیف مظنون را فراهم سازد. استفاده از شناسایی فعالیت انسانی برای ردیابی رفتار عابران و ترافیک در تقاطع‌ها نیز نمونه‌ای از افزایش ایمنی جاده‌ای از طریق این فناوری است. کاستا و همکاران^۲ (۲۰۱۸) سیستمی را معرفی کرده‌اند که می‌تواند انواع تخلفات رانندگی مانند سرعت غیرمجاز را شناسایی و طبقه‌بندی کرده و شواهد معتبری برای پلیس فراهم آورد.
۲. سلامت: بیماران مبتلا به بیماری‌هایی نظیر افسردگی، اوتیسم، پارکینسون و آلزایمر می‌توانند از طریق سامانه‌های شناسایی فعالیت انسانی مورد پایش و حمایت قرار گیرند. به عنوان نمونه، سالاریان و همکاران^۳ (۲۰۰۷) سیستمی طراحی کرده‌اند که با بهره‌گیری از حسگرهای پوشیدنی، میزان و شدت لرزش‌ها در بیماران پارکینسونی را تشخیص داده و آن را درجه‌بندی می‌کند. این سیستم علاوه بر کمک به تنظیم زمان و دوز داروها، می‌تواند بازخوردهایی به بیماران و مراقبین ارائه دهد. همچنین، رودوویچ و همکاران^۴ (۲۰۱۸) سیستمی توسعه داده‌اند که به ارزیابی حرکات بدنی، جهت نگاه و حالات چهره کودکان اوتیستیک در موقعیت‌های اجتماعی می‌پردازد. این سیستم نیز می‌تواند اطلاعات مفیدی برای والدین و درمانگران فراهم سازد.

۵- نتیجه‌گیری

شناسایی فعالیت انسانی مبتنی بر ویدیو و به صورت بلادرنگ، نگاه ما به تحلیل و درک رفتارهای انسانی را دگرگون ساخته است. این فناوری با فراهم‌سازی امکان شناسایی خودکار فعالیت‌های انسانی، ما را از محدودیت‌های وابستگی به اپراتور انسانی رها کرده و احتمال بروز خطاهای انسانی را به‌طور چشم‌گیری کاهش می‌دهد (Kamthe and Patil, 2018). این فناوری پیشرفته سازمان‌ها را قادر می‌سازد تا حجم عظیمی از داده‌های ویدیویی را به صورت مؤثر پایش و تفسیر کنند و فرصت‌های جدیدی برای نظارت و تحلیل کارآمد ایجاد نمایند. شناسایی فعالیت انسانی نه تنها فرایندها را ساده‌تر می‌کند بلکه با کاهش وابستگی به نیروی انسانی، صرفه‌جویی قابل توجهی در هزینه‌ها به همراه دارد. بهره‌گیری از این رویکرد نوین به سازمان‌ها کمک می‌کند تا منابع خود را بهینه تخصیص دهند و تمرکز بیشتری بر

1. Thida et al.
2. Costea et al.
3. Salarian et al.
4. Rudovic et al.

جنبه‌های حیاتی دیگر داشته باشند. درنهایت، اهمیت شناسایی فعالیت انسانی مبتنی بر ویدیو در توانایی آن برای ارتقای ایمنی و امنیت، پایش سلامت و تسهیل تعامل انسان و رایانه نهفته است. این فناوری با پرداختن به چالش‌های مطرح شده و دستیابی به اهداف تعیین شده می‌تواند تأثیر مثبتی بر بخش‌های مختلف جامعه داشته باشد و زمینه‌ساز شکل‌گیری جامعه‌ای ایمن‌تر و کارآمدتر گردد.

منابع

- Beddiar, D. R., Nini, B., Sabokrou, M., & Hadid, A. (2020). Vision-based human activity recognition: A survey. *Multimedia Tools and Applications*, 79(41–42), 30509–30555.
- Belagiannis, V., Amin, S., Andriluka, M., Schiele, B., Navab, N., & Ilic, S. (2014). 3D pictorial structures for multiple human pose estimation. *CiteSeer X (The Pennsylvania State University)*.
- Candamo, J., Shreve, M., Goldgof, D. B., Sapper, D. B., & Kasturi, R. (2010). Understanding transit scenes: A survey on human behavior-recognition algorithms. *IEEE Transactions on Intelligent Transportation Systems*, 11(1), 206–224.
- Costea, A. D., Ionescu, S. A., Pirlogea, M. I., & Popescu, M. (2018). Traffic violation detection using human action recognition. In *2018 International Symposium on Signals, Circuits and Systems (ISSCS)* (pp. 1–4).
- Kamthe, U. M., & Patil, C. G. (2018). Suspicious activity recognition in video surveillance system. *IEEE Xplore*.
- Kong, Y., Jia, Y., & Fu, Y. (2014). Interactive phrases: Semantic descriptions for human interaction recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(9), 1775–1788.
- Liu, N., Dellandréa, E., Tellez, B., & Chen, L. (2011). Associating textual features with visual ones to improve affective image classification. *HAL Archives Ouvertes*.
- Mendoza, M. A., & Blanca. (2008). Applying space state models in human action recognition: A comparative study. *Lecture Notes in Computer Science* (pp. 53–62).
- Pareek, P., & Thakkar, A. (2020). A survey on video-based human action recognition: Recent updates, datasets, challenges, and applications. *Artificial Intelligence Review*, 54(3), 2259–2322.
- Poppe, R. (2010). A survey on vision-based human action recognition. *Image and Vision Computing*, 28(6), 976–990.
- Rudovic, O., Lee, J., Dai, M., Schuller, B., & Picard, R. W. (2018). Personalized machine learning for robot perception of affect and engagement in autism therapy. *Science Robotics*, 3(19), eaao6760.
- Salarian, A., Russmann, H., Vingerhoets, F. J. G., Dehollain, C., Blanc, Y., Burkhard, P. R., & Aminian, K. (2007). Quantification of tremor and bradykinesia in Parkinson's disease using a novel ambulatory monitoring system. *IEEE Transactions on Biomedical Engineering*, 54(2), 313–322.
- Sharma, V., Gupta, M., Pandey, A. K., Mishra, D., & Kumar, A. (2022). A review of deep learning-based human activity recognition on benchmark video datasets. *Applied Artificial Intelligence*, 36(1).
- Sheng, X., Li, K., Shen, Z., & Xiao, G. (2023). A progressive difference method for capturing visual tempos on action recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(3), 977–987.
- Siddiqi, M., Ali, R., Rana, M., Hong, E.-K., Kim, E., & Lee, S. (2014). Video-based human activity recognition using multilevel wavelet decomposition and stepwise linear discriminant analysis. *Sensors*, 14(4), 6370–6392.
- Sun, X., Chen, M., & Hauptmann, A. (2009). Action recognition via local descriptors and holistic features. *IEEE Xplore*.
- Theodorakopoulos, I., Kastaniotis, D., Economou, G., & Fotopoulos, S. (2014). Pose-based human action recognition via sparse representation in dissimilarity space. *Journal of Visual Communication and Image Representation*, 25(1), 12–23.
- Thida, M., Yong, Y. L., Climent-Pérez, P., Eng, H. L., & Remagnino, P. (2013). A literature review on video analytics of crowded scenes. In *Intelligent Multimedia Surveillance* (pp. 17–36). Springer.

- Vrigkas, M., Nikou, C., & Kakadiaris, I. A. (2015). A review of human activity recognition methods. *Frontiers in Robotics and AI*, 2 (November).
- Wang, H., Kläser, A., Schmid, C., & Liu, C.-L. (2013). Dense trajectories and motion boundary descriptors for action recognition. *International Journal of Computer Vision*, 103(1), 60–79.
- Yadav, S. K., Tiwari, K., Pandey, H. M., & Akbar, S. A. (2021). A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions. *Knowledge-Based Systems*, 223, 106970.

استناد به این مقاله: صبوری، محمد. (۱۴۰۴). بررسی جامع روش‌های شناسایی فعالیت‌های انسانی از روی پردازش ویدیو. فصلنامه پژوهش‌های نوین در شهر هوشمند، ۳(۴)، ۶۲–۷۰.



New Researches in The Smart City is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.