

عنوان مقاله:

ارائه یک معماری یادگیری عمیق برای شناسایی اعمال انسان در ویدئو

محل انتشار:

پنجمین کنفرانس پردازش سیگنال و سیستم‌های هوشمند (سال: 1398)

تعداد صفحات اصل مقاله: 6

نویسندگان:

مهدی سوادی حسینی - دانشجو کارشناسی ارشد، دانشگاه تربیت مدرس،

فواد قادری - استادیار دانشکده مهندسی برق و کامپیوتر، آزمایشگاه تعامل انسان و کامپیوتر، دانشگاه تربیت مدرس،

خلاصه مقاله:

شناسایی اعمال انسان در ویدئو با توجه به کاربردهای بسیاری که در زمینه های مختلف از جمله امنیت، سلامت، مدیریت هوشمند شهر و ساختمان و ... دارد، طی چند سال اخیر توجه زیادی را به خود جلب کرده است. از سوی دیگر، رویکردهای یادگیری عمیق مانند شبکه های عصبی پیچشی (CNN ها) و شبکه های عصبی بازگشتی (RNN ها) در زمینه های مختلف توانسته اند نتایج بسیار خوبی را کسب کنند. با این وجود، علی رغم تغییرات زمانی قابل توجه در محتوای ویدئو، تاکنون شبکه های عصبی بازگشتی در شناسایی اعمال انسان ضعیف تر از روش های مبتنی بر شبکه های پیچشی عمل کرده اند. در این پژوهش با معرفی یک رویکرد سلسله مراتبی زمانی در دو سطح محلی و سراسری، با استفاده از شبکه های پیچشی متورم دو جریانی که از شبکه های عصبی پیچشی دو بعدی معروف برای دسته بندی تصاویر ایجاد شده اند و همچنین استفاده از پشته ای از لایه های واحدهای بازگشتی گیت دار (GRU) توانسته ایم رویکردی جدید برای شناسایی اعمال انسان در ویدئو ارائه دهیم. در این رویکرد ابتدا ویژگی های فضایی-زمانی را برای هر دو جریان فضایی و زمانی به صورت محلی با استفاده از شبکه های پیچشی سه بعدی متورم (I3D) فضایی و زمانی استخراج می کنیم که این کار باعث تبدیل دنباله طولانی قاب های یک ویدئو به یک دنباله کوتاه تر و حاوی اطلاعات مفیدتر می شود. این دنباله کوتاه شده را به شبکه ای از GRU ها می دهیم و در نهایت نیز نتایج دو جریان را جمع می کنیم. برای تجمیع نتایج، یک لایه جدید با نام میانگین وزن دهی شده را معرفی می کنیم که اهمیت هر جریان را در فرایند آموزش به صورت خودکار فرا می گیرد. ارزیابی ها نشان دهنده نتایجی قابل قبول برای دو مجموعه داده HMDB51 و UCF101 هستند. روش پیشنهادی موجب 1.6 درصد بهبود در صحت دسته بندی نمونه های مجموعه داده پرچالش HMDB51 نسبت به نتایج گزارش شده بهترین روش موجود گردیده است.

کلمات کلیدی:

شبکه های عصبی پیچشی متورم، واحد بازگشتی گیت دار، شناسایی اعمال، معماری دو جریانی.

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/983120>

