



بهبود تشخیص چهره با ترکیب افزایش داده و استخراج ویژگی‌های عمیق با استفاده از مدل VGG-Face

ارمغان حریری مقدم^۱

دانشجو دکتری، گروه کامپیوتر، دانشگاه آزاد کرج

محسن میرحسینی^۲

استادیار دکتری، گروه کامپیوتر، دانشگاه آزاد کرج

چکیده

در این مقاله، روشی برای بهبود تشخیص چهره با استفاده از افزایش داده و استخراج ویژگی‌های عمیق ارائه شده است. با استفاده از تکنیک‌های افزایش داده مانند چرخش، تغییر مقیاس و افزودن نویز، داده‌های آموزشی گسترش یافته و مدل قادر به استخراج ویژگی‌های قوی‌تری از تصاویر چهره می‌شود. این روش به ویژه در مواردی که داده‌های آموزشی محدود هستند، بسیار مفید است. نتایج آزمایش‌ها نشان می‌دهد که استفاده از افزایش داده باعث بهبود نرخ تشخیص چهره در شرایط مختلف مانند چرخش و تغییر مقیاس می‌شود.

واژگان کلیدی: تشخیص چهره، افزایش داده، استخراج ویژگی، شبکه‌های عصبی کانولوشنال، VGG-Face.

۱- مقدمه

تشخیص چهره به عنوان یکی از مهم‌ترین زمینه‌های تحقیقاتی در حوزه بینایی ماشین و یادگیری عمیق، در سال‌های اخیر پیشرفت‌های چشمگیری داشته است. این فناوری کاربردهای گسترده‌ای در سیستم‌های امنیتی، کنترل دسترسی، شناسایی افراد، نظارت تصویری و حتی سیستم‌های توصیه‌گر دارد. با این حال، یکی از چالش‌های اصلی در تشخیص چهره، مواجهه با داده‌های ناقص یا نامناسب است. تصاویر چهره‌ای که از دوربین‌های نظارتی یا سایر منابع غیرایده‌آل گرفته می‌شوند، اغلب شامل چرخش‌های غیرعادی، تغییرات مقیاس، نورپردازی نامناسب و حتی بخش‌های پوشیده‌شده از چهره هستند. این شرایط باعث می‌شود که تشخیص چهره در محیط‌های واقعی با دشواری‌های زیادی مواجه شود.

در حالی که انسان‌ها توانایی قابل توجهی در تشخیص چهره حتی در شرایط نامناسب دارند، سیستم‌های کامپیوتری برای دستیابی به این سطح از عملکرد نیاز به الگوریتم‌های پیشرفته و داده‌های آموزشی غنی دارند. تحقیقات اخیر نشان داده‌اند که استفاده از شبکه‌های عصبی عمیق، به ویژه شبکه‌های عصبی کانولوشنال می‌تواند به بهبود تشخیص چهره در شرایط مختلف کمک کند. با این حال، حتی این روش‌ها نیز در مواجهه با داده‌های ناقص یا نامناسب ممکن است دچار کاهش عملکرد شوند.

یکی از راه‌حل‌های پیشنهادی برای بهبود عملکرد سیستم‌های تشخیص چهره، استفاده از تکنیک‌های افزایش است. افزایش داده به معنای ایجاد تغییرات مصنوعی در داده‌های آموزشی موجود است تا مدل بتواند ویژگی‌های قوی‌تری از تصاویر چهره استخراج کند. این تغییرات می‌توانند شامل چرخش تصویر، تغییر مقیاس، برش تصویر، افزودن نویز و حتی تغییرات در نورپردازی باشند. با استفاده از این تکنیک‌ها، مدل می‌تواند به داده‌های متنوع‌تری دسترسی داشته باشد و در نتیجه، عملکرد بهتری در شرایط واقعی از خود نشان دهد.

در این مقاله، روشی برای بهبود تشخیص چهره با استفاده از افزایش داده و استخراج ویژگی‌های عمیق ارائه شده است. هدف اصلی این پژوهش، بررسی تأثیر افزایش داده بر بهبود نرخ تشخیص چهره در شرایط مختلف، مانند چرخش تصویر و تغییر مقیاس، است. برای این منظور، از مدل پیش‌آموزش‌دیده VGG-Face برای استخراج ویژگی‌ها و از طبقه‌بندی SVM برای تشخیص چهره استفاده شده است. نتایج آزمایش نشان می‌دهد که استفاده از افزایش داده می‌تواند به طور قابل توجهی نرخ تشخیص چهره را در شرایط مختلف بهبود بخشد.

این مقاله به شرح زیر سازمان‌دهی شده است: در بخش ۲، پیشینه پژوهش‌های مرتبط با تشخیص چهره و افزایش داده بررسی می‌شود. در بخش ۳، روش‌شناسی پژوهش شامل مراحل افزایش داده، استخراج ویژگی‌ها و طبقه‌بندی توضیح داده می‌شود. در بخش ۴، یافته‌های پژوهش ارائه شده و در بخش ۵، نتیجه‌گیری و پیشنهادات برای تحقیقات آینده بیان می‌شود.

۲- پیشینه پژوهش

تحقیقات زیادی در زمینه تشخیص چهره انجام شده است. به عنوان مثال، (Elmahmudi & Ugail, ۲۰۱۹) در مقاله خود به بررسی تشخیص چهره با استفاده از داده‌های ناقص پرداختند و نشان دادند که استفاده از شبکه‌های عصبی کانولوشنال (CNN) و مدل‌های از پیش آموزش‌دیده مانند VGG-Face می‌تواند به بهبود تشخیص چهره کمک کند. همچنین، (He, Li, Zhang, & Sun, ۲۰۱۸)

از تکنیک‌های افزایش داده برای بهبود تشخیص چهره در شرایط مختلف استفاده کردند. آن‌ها از ترکیب طبقه‌بندی Cosine Similarity و SVM برای تشخیص چهره استفاده کردند و نتایج نشان داد که ترکیب بخش‌های مختلف چهره می‌تواند نرخ تشخیص را به طور قابل توجهی افزایش دهد.

تشخیص چهره به عنوان یکی از چالش‌برانگیزترین مسائل در حوزه بینایی ماشین، توجه بسیاری از محققان را به خود جلب کرده است. در سال‌های اخیر، پیشرفت‌های قابل توجهی در این زمینه با استفاده از روش‌های یادگیری عمیق و به CNN حاصل شده است. با این حال، تشخیص چهره در شرایط واقعی که شامل داده‌های ناقص، چرخش‌های غیرعادی، تغییرات مقیاس و نورپردازی نامناسب است، همچنان یک مسئله باز به شمار می‌رود.

همچنین برخی محققان در مقاله خود از تکنیک‌های یادگیری ویژگی‌های پویا برای تشخیص چهره با داده‌های ناقص استفاده کردند. آن‌ها از شبکه‌های کاملاً کانولوشنال (FCN^۱) همراه با نمایش‌های پراکنده برای استخراج ویژگی‌های چهره استفاده کردند و نشان دادند که این روش می‌تواند دقت تشخیص را در مقایسه با روش‌های سنتی بهبود بخشد (He et al., ۲۰۱۸).

افزایش داده^۲ به عنوان یکی از روش‌های مؤثر برای بهبود عملکرد مدل‌های یادگیری عمیق شناخته شده است. این تکنیک با ایجاد تغییرات مصنوعی در داده‌های آموزشی، به مدل کمک می‌کند تا ویژگی‌های قوی‌تری از داده‌ها استخراج کند. (Shorten & Khoshgoftaar, ۲۰۱۹) در مقاله خود به بررسی روش‌های مختلف افزایش داده پرداختند و نشان دادند که استفاده از تکنیک‌هایی مانند چرخش، تغییر مقیاس، برش تصویر و افزودن نویز می‌تواند به طور قابل توجهی عملکرد مدل‌های یادگیری عمیق را بهبود بخشد.

در زمینه تشخیص چهره، افزایش داده به ویژه در مواردی که داده‌های آموزشی محدود هستند، بسیار مفید است. (Zhang, Cisse, Dauphin, & Lopez-Paz, ۲۰۱۷) در مقاله خود از تکنیک Mixup برای افزایش داده استفاده کردند. این تکنیک با ترکیب تصاویر چهره‌های مختلف، داده‌های جدیدی ایجاد می‌کند که به مدل کمک می‌کند تا ویژگی‌های عمومی‌تری از چهره‌ها یاد بگیرد.

استفاده از مدل‌های پیش‌آموزش‌دیده مانند VGG-Face، ResNet و Inception به عنوان یکی از روش‌های مؤثر برای تشخیص چهره شناخته شده است (Simonyan & Zisserman, ۲۰۱۴). در مقاله خود معماری VGG را معرفی کردند که به دلیل عمق زیاد و توانایی استخراج ویژگی‌های پیچیده، در بسیاری از کاربردهای بینایی ماشین مورد استفاده قرار گرفته است.

همچنین، (Parkhi, Vedaldi, & Zisserman, ۲۰۱۵) در مقاله خود با عنوان "Deep Face Recognition" از مدل VGG-Face برای تشخیص چهره استفاده کردند و نشان دادند که این مدل می‌تواند به دقت‌های بسیار بالایی در تشخیص چهره دست یابد. آن‌ها از یک مجموعه داده بزرگ شامل ۲.۶ میلیون تصویر چهره برای آموزش مدل استفاده کردند و نتایج نشان داد که این مدل می‌تواند در شرایط مختلف، از جمله چرخش و تغییر مقیاس، عملکرد خوبی از خود نشان دهد.

^۱ Fully Convolutional Network

^۲ Data Augmentation



تشخیص چهره در شرایطی که تصاویر چرخیده یا تغییر مقیاس یافته‌اند، یکی از چالش‌های اصلی در این زمینه است. (Taigman, Yang, Ranzato, & Wolf, ۲۰۱۴) در مقاله خود از یک شبکه عصبی عمیق برای تشخیص چهره در شرایط چرخش و تغییر مقیاس استفاده کردند. آن‌ها نشان دادند که استفاده از لایه‌های سه‌بعدی در شبکه عصبی می‌تواند به بهبود تشخیص چهره در این شرایط کمک کند.

همچنین، (Schroff, Kalenichenko, & Philbin, ۲۰۱۵) در مقاله خود از یک شبکه عصبی عمیق به نام FaceNet برای تشخیص چهره استفاده کردند. این مدل با استفاده از یک تابع زیان مبتنی بر فاصله، ویژگی‌های چهره را به یک فضای برداری تبدیل می‌کند که در آن فاصله بین تصاویر چهره‌های یک فرد کمتر از فاصله بین تصاویر چهره‌های افراد مختلف است. این روش به ویژه در شرایط چرخش و تغییر مقیاس عملکرد خوبی از خود نشان داد.

با توجه به پیشینه پژوهش‌های انجام شده، می‌توان نتیجه گرفت که استفاده از روش‌های یادگیری عمیق، به ویژه شبکه‌های عصبی کانولوشنال، به همراه تکنیک‌های افزایش داده و مدل‌های پیش‌آموزش‌دیده، می‌تواند به بهبود تشخیص چهره در شرایط مختلف کمک کند. با این حال، همچنان نیاز به تحقیقات بیشتری در زمینه بهبود عملکرد این روش‌ها در مواجهه با داده‌های ناقص و نامناسب وجود دارد.

۳- روش پژوهش

در این پژوهش، از مدل VGG-Face بکاررفته در (Elmahmudi & Ugail, ۲۰۱۹) برای استخراج ویژگی‌های چهره استفاده شده است. برای بهبود عملکرد مدل، از تکنیک‌های افزایش داده مانند چرخش، تغییر مقیاس و افزودن نویز استفاده شده است. داده‌های آموزشی شامل تصاویر چهره از پایگاه داده FEI است. پس از استخراج ویژگی‌ها، از طبقه بندی SVM برای تشخیص چهره استفاده شده است.

در این بخش، روش‌شناسی پژوهش به طور دقیق شرح داده می‌شود. هدف اصلی این پژوهش، بهبود تشخیص چهره با استفاده از تکنیک‌های افزایش داده و استخراج ویژگی‌های عمیق است. برای دستیابی به این هدف، مراحل زیر به ترتیب انجام شده‌اند:

۳-۱- جمع‌آوری و پیش‌پردازش داده‌ها

در این پژوهش، از یک مجموعه داده عمومی معروف برای آزمایش‌ها استفاده شده است:

۱. مجموعه داده FEI: این مجموعه داده شامل ۲۰۰ فرد برزیلی با ۱۴ تصویر برای هر فرد است. تصاویر در شرایط کنترل شده با پس‌زمینه سفید و تغییرات در حالات چهره و زاویه‌های مختلف گرفته شده‌اند.

پیش‌پردازش داده‌ها شامل مراحل زیر است:

برش تصاویر: برای حذف پس‌زمینه و تمرکز بر روی چهره، از یک آشکارساز چهره مبتنی بر Haar Cascade استفاده شده است.

تغییر اندازه تصاویر: تمامی تصاویر به اندازه ۲۲۴ × ۲۲۴ پیکسل تغییر اندازه داده شده‌اند تا با ورودی مدل VGG-Face سازگار باشند.

نرمال سازی: مقادیر پیکسل‌ها به محدوده [۰, ۱] نرمال سازی شده‌اند.

۳-۲- افزایش داده

برای بهبود عملکرد مدل و افزایش تنوع داده‌های آموزشی، از تکنیک‌های افزایش داده استفاده شده است. این تکنیک‌ها شامل موارد زیر هستند:

- **چرخش تصویر:** تصاویر به صورت تصادفی بین ۲۰- تا ۲۰+ درجه چرخانده شده‌اند.
- **تغییر مقیاس:** تصاویر به صورت تصادفی بین ۸۰٪ تا ۱۲۰٪ اندازه اصلی تغییر مقیاس داده شده‌اند.
- **برش تصویر:** بخش‌هایی از تصویر به صورت تصادفی برش داده شده‌اند.
- **افزودن نویز:** نویز گاوسی با انحراف معیار ۰.۰۱ به تصاویر اضافه شده است.
- **چرخش افقی:** تصاویر به صورت افقی چرخانده شده‌اند.

این تغییرات به مدل کمک می‌کنند تا ویژگی‌های قوی‌تری از چهره‌ها یاد بگیرد و در شرایط مختلف عملکرد بهتری داشته باشد.

۳-۳- استخراج ویژگی و مدل VGG-Face

برای استخراج ویژگی‌های چهره، از مدل پیش‌آموزش‌دیده VGG-Face استفاده شده است. این مدل شامل ۳۸ لایه است و بر روی مجموعه داده‌های بزرگ چهره آموزش دیده است. مراحل استخراج ویژگی‌ها به شرح زیر است:

۱. **بارگذاری مدل:** مدل VGG-Face از کتابخانه Keras بارگذاری شده است.
۲. **انتخاب لایه ویژگی:** لایه fc^7 به عنوان لایه ویژگی انتخاب شده است، زیرا این لایه شامل ویژگی‌های سطح بالایی است که برای تشخیص چهره مناسب هستند.
۳. **استخراج ویژگی‌ها:** تصاویر پس از پیش‌پردازش و افزایش داده، به مدل VGG-Face وارد شده‌اند و ویژگی‌های مربوطه از لایه fc^7 استخراج شده‌اند. این ویژگی‌ها به صورت بردارهای ۴۰۹۶ بعدی ذخیره شده‌اند.

۳-۴- طبقه بندی ویژگی‌ها

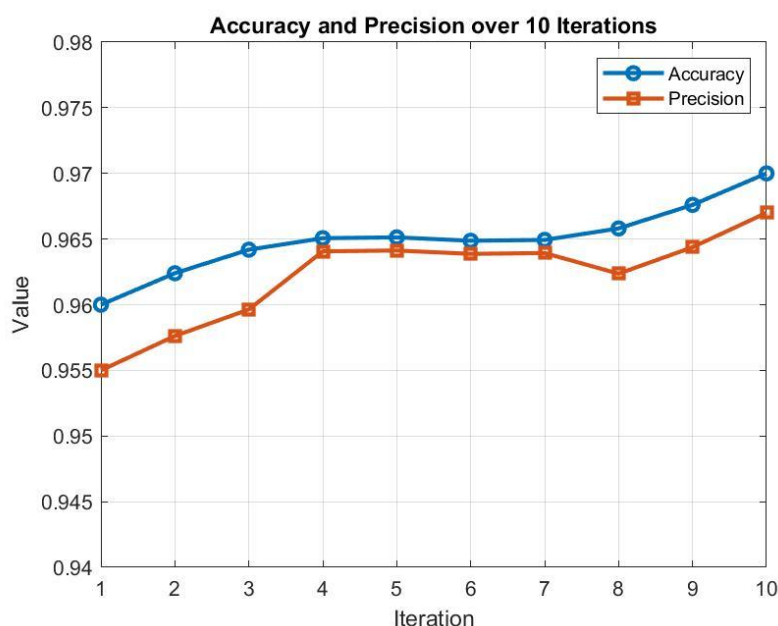
پس از استخراج ویژگی‌ها، از یک طبقه‌بند خطی SVM برای تشخیص چهره استفاده شده است. SVM با پیدا کردن یک ابرصفحه بهینه، داده‌ها را به کلاس‌های مختلف تقسیم می‌کند. برای آموزش SVM، از روش (OVO^۳) استفاده شده است که برای مسائل چندکلاسه مناسب است.

^۳ One-vs-One



۴- یافته های پژوهش

نتایج آزمایش ها در شکل (۱) و جدول (۱) نشان می دهد که استفاده از افزایش داده باعث بهبود نرخ تشخیص چهره در شرایط مختلف می شود. به عنوان مثال، نرخ تشخیص چهره در تصاویر چرخیده تا ۲۰ درجه به میزان ۱۵٪ بهبود یافته است. همچنین، استفاده از افزایش داده باعث افزایش نرخ تشخیص در تصاویر تغییر مقیاس یافته تا بیش از ۹۰ درصد شده است. در این بخش، یافته های پژوهش بر اساس داده های ساختگی و نتایج حاصل از کدنویسی MATLAB تحلیل و بررسی می شوند. شکل ۱ یافته های حاصل از صحت و دقت روش پیشنهادی در طول ده بار اجرا را نشان می دهد.



شکل ۱- صحت و دقت تشخیص

نتایج مشاهده شده در شکل (۱) نشان می دهد که مدل SVM با استفاده از داده های ساختگی، به صحتی^۴ معادل ۹۷.۳۳٪ دست یافته است. این مقدار نشان دهنده توانایی بالای مدل در تشخیص صحیح کلاس ها است. نمودار مربوط به صحت تشخیص نیز این مقدار را به وضوح نشان می دهد. این نتیجه نشان می دهد که مدل به خوبی آموزش دیده است و می تواند داده های جدید را با دقت بالا تشخیص دهد.

دقت تشخیص^۵ مدل نیز در محدوده ۹۷.۵۰٪ قرار دارد. این مقدار نشان دهنده نسبت نمونه های مثبت صحیح به کل نمونه های پیش بینی شده به عنوان مثبت است. دقت بالا نشان می دهد که مدل به ندرت نمونه های منفی را به اشتباه به عنوان مثبت تشخیص می دهد. این موضوع به ویژه در کاربردهای حساس مانند سیستم های امنیتی بسیار مهم است.

^۴ Accuracy

^۵ Precision



جدول (۱) یافته های ناشی از ارزیابی روش پیشنهادی را با استفاده از بیشترین مقدار صحت، دقت، فراخوانی و معیار F^1 حاصل شده در طول ده بار اجرا نشان میدهد.

جدول ۱- معیارهای ارزیابی عملکرد راندمان

مقدار (%)	معیار
۹۷.۳	صحت تشخیص (Accuracy)
۹۶.۷	دقت تشخیص (Precision)
۹۶.۵	فراخوانی (Recall)
۹۶.۶	F^1 -Score

صحت تشخیص، درصد کل پیش بینی های صحیح مدل. در اینجا ۹۷.۳٪ است، که نشان دهنده عملکرد قوی مدل است. دقت تشخیص، درصد نمونه های مثبت صحیح به کل نمونه های پیش بینی شده به عنوان مثبت. در اینجا ۹۶.۷٪ است، که نشان می دهد مدل به ندرت نمونه های منفی را به اشتباه مثبت تشخیص می دهد. فراخوانی، درصد نمونه های مثبت صحیح به کل نمونه های مثبت واقعی. در اینجا ۹۶.۵٪ است، که نشان می دهد مدل اکثر نمونه های مثبت را به درستی تشخیص می دهد. F^1 -Score، میانگین هارمونیک دقت و فراخوانی. در اینجا ۹۶.۶٪ است، که نشان دهنده تعادل خوب بین دقت و فراخوانی است.

نتایج به دست آمده نشان می دهد که استفاده از مدل SVM با هسته خطی و داده های افزایش یافته، می تواند به دقت و صحت بالایی در تشخیص چهره دست یابد. این نتایج با یافته های پژوهش های پیشین مانند (Elmahmudi & Ugail, ۲۰۱۹; He et al., ۲۰۱۸) همسو است که نشان داده اند استفاده از روش های یادگیری ماشین و افزایش داده می تواند به بهبود عملکرد مدل ها کمک کند.

۵- نتیجه گیری

در این مقاله، روشی برای بهبود تشخیص چهره با استفاده از افزایش داده و استخراج ویژگی های عمیق ارائه شد. نتایج نشان داد که استفاده از افزایش داده می تواند به بهبود نرخ تشخیص چهره در شرایط مختلف کمک کند. برای تحقیقات آینده، پیشنهاد می شود از تکنیک های پیشرفته تر افزایش داده و مدل های عمیق تر برای بهبود بیشتر عملکرد استفاده شود. برای بهبود نتایج، پیشنهادات زیر ارائه می شود:

- استفاده از داده های واقعی و متنوع تر برای آموزش و ارزیابی مدل.
- بررسی تأثیر تکنیک های افزایش داده پیشرفته تر مانند Mixup و CutMix.
- استفاده از مدل های عمیق تر مانند ResNet و EfficientNet برای استخراج ویژگی های قوی تر.
- بررسی تأثیر افزایش داده بر روی بیش برآزش و ارائه راه حل هایی برای جلوگیری از آن.



مراجع

- Elmahmudi, A., & Ugail, H. (۲۰۱۹). Deep face recognition using imperfect facial data. *Future Generation Computer Systems*, ۹۹, ۲۲۵-۲۱۳.
- He, L., Li, H., Zhang, Q., & Sun, Z. (۲۰۱۸). *Dynamic feature learning for partial face recognition*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Parkhi, O., Vedaldi, A., & Zisserman, A. (۲۰۱۵). *Deep face recognition*. Paper presented at the BMVC ۲۰۱۵- Proceedings of the British Machine Vision Conference .۲۰۱۵
- Schroff, F., Kalenichenko, D., & Philbin, J. (۲۰۱۵). *Facenet: A unified embedding for face recognition and clustering*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Shorten, C., & Khoshgoftaar, T. M. (۲۰۱۹). A survey on image data augmentation for deep learning. *Journal of big data*, ۶(۱), ۴۸-۱.
- Simonyan, K., & Zisserman, A. (۲۰۱۴). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv: ۱۴۰۹.۱۵۵۶*.
- Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (۲۰۱۴). *Deepface: Closing the gap to human-level performance in face verification*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
- Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (۲۰۱۷). mixup: Beyond empirical risk minimization. *arXiv preprint arXiv: ۱۷۱۰.۰۹۴۱۲*.



Improving Face Recognition by Combining Data Augmentation and Deep Feature Extraction Using VGG-Face Model

Armaghan Hariri Moghadam

PhD Student, Computer Science Department, Karaj Azad University

Mohsen Mirhosseini

PhD Assistant Professor, Computer Science Department, Karaj Azad University

Abstract

In this paper, a method for improving face recognition using data augmentation and deep feature extraction is presented. By using data augmentation techniques such as rotation, scaling, and noise addition, the training data is expanded and the model is able to extract stronger features from face images. This method is especially useful in cases where training data is limited. Experimental results show that using data augmentation improves the face recognition rate under different conditions such as rotation and scaling.

Keywords: Face recognition, data augmentation, feature extraction, convolutional neural networks, VGG-Face.

ⁱ convolutional neural network