



دانشگاه کاشان  
University of Kashan

مجله محاسبات نرم

SOFT COMPUTING JOURNAL

تارنمای مجله: [scj.kashanu.ac.ir](http://scj.kashanu.ac.ir)



## تشخیص چهره با داده‌های ناقص توسط شبکه عصبی کانولوشنی عمیق<sup>✧</sup>

فرناز حسینی<sup>۱</sup>، استادیار، الهه طبیب زاده لمر<sup>۲</sup>، کارشناسی ارشد، سید مهدی میرکاظمی نیارق<sup>۳</sup>، استادیار

<sup>۱</sup> گروه مهندسی کامپیوتر، دانشگاه فنی و حرفه‌ای، تهران، ایران.

<sup>۲</sup> گروه مهندسی کامپیوتر، موسسه آموزش عالی شهریار، آستارا، ایران.

<sup>۳</sup> گروه برق، واحد آستارا، دانشگاه آزاد اسلامی، آستارا، ایران.

### اطلاعات مقاله

### چکیده

#### تاریخچه مقاله:

دریافت ۲۳ فروردین ماه ۱۴۰۲

پذیرش ۱۵ مهر ماه ۱۴۰۲

#### کلمات کلیدی:

تشخیص چهره

شبکه عصبی کانولوشنی

الگوریتم بهینه‌سازی زنبور عسل

داده‌های ناقص

شبکه عصبی کانولوشنی عمیق

چهره انسان یک شی ثابت نیست و فاکتورهای متنوعی که منجر به نمایش‌های مختلف چهره می‌شوند، وجود دارند. در الگوریتم‌های تشخیص چهره، فاکتورهای ذاتی و تصادفی که باعث ایجاد اختلاف در ظاهر چهره می‌شوند، وجود داده‌های ناقص در پایگاه داده‌ها، حجم پایگاه داده‌ها، اختلاف در ابعاد تصاویر ذخیره شده و تغییر حالات چهره می‌توانند فرآیند تشخیص چهره را با مشکلات مختلفی روبرو سازند. ارائه روشی که بتواند این مشکلات را تا حد قابل قبولی رفع کند، می‌تواند گستره کاربرد روش‌های تشخیص چهره را افزایش دهد. در این مطالعه برای بهبود تشخیص چهره با داده‌های ناقص از یک شبکه عصبی کانولوشنی عمیق استفاده شده است. روش پیشنهادی از چند گام مختلف تشکیل شده است. در مرحله اول پس از انتخاب و استخراج داده اولیه از پایگاه داده، پیش‌پردازش اطلاعات با اعمال فیلتر، تبدیل هیستوگرام و تشخیص لبه انجام می‌شود. در مرحله دوم استخراج نقاط مهم برای هر تصویر انجام می‌شود. خروجی این مرحله به عنوان ورودی به الگوریتم بهینه‌سازی زنبور عسل داده شده تا برای انتخاب ویژگی‌های موثر و شناخت آنها در راستای بهینه‌سازی استفاده گردد. در نهایت از یک شبکه عصبی کانولوشنی عمیق در دو گام آموزش و آزمون برای تشخیص چهره استفاده شده است. روش پیشنهادی در محیط متلب شبیه‌سازی شده و توسط معیارهای دقت، صحت و پوشش مورد ارزیابی قرار گرفته است. نتایج به دست آمده با دقت ۹۶/۱۱٪ بیانگر بهبود تشخیص چهره نسبت به سایر کارهای انجام شده در سال‌های اخیر و همچنین کاهش هزینه در فرآیند تشخیص انجام شده می‌باشد.

© ۱۴۰۲ نویسندگان. مقاله با دسترسی آزاد تحت مجوز CC-BY

### ۱. مقدمه

به دست می‌آید و جز یکی از مهم‌ترین فناوری‌های بیومتریک به حساب می‌آید که قادر است افراد را از یکدیگر تشخیص دهد و یا آنها را شناسایی نماید. سیستم‌های تشخیص چهره دارای راه‌کارها و فناوری‌های گوناگون در پردازش تصویر بوده و به طور معمول برای اسکن چهره توسط فاصله نقشه‌برداری میان ویژگی‌های مهم چهره تصویربرداری‌های حرارتی و تحلیل بافت‌های سطحی در پوست استفاده می‌شوند [۱]. سیستم‌های

تشخیص چهره تجزیه و تحلیل الگوهایی است صورت یک فرد

✧ نوع مقاله: پژوهشی

\* نویسنده مسئول

پست(های) الکترونیک: [f-hoseini@tvu.ac.ir](mailto:f-hoseini@tvu.ac.ir) (حسینی)

[elatabib2@shahriar.ac.ir](mailto:elatabib2@shahriar.ac.ir) (طیب زاده لمر)

[mirkazemi.seyedmehdi@iau.ac.ir](mailto:mirkazemi.seyedmehdi@iau.ac.ir) (میرکاظمی نیارق)

تشخیص چهره دارای کاربردهای فراوانی هستند که از جمله مرسوم‌ترین کاربردهای آن می‌توان به شناسایی افراد و برآورده کردن اهداف امنیتی اشاره کرد [۲]. تشخیص چهره مبتنی بر رایانه یک روش قابل اطمینان است که در بسیاری از سناریوهای کنترل دسترسی از آن استفاده می‌شود [۳]. به این ترتیب که تشخیص چهره یا احراز هویت به طور عمده با استفاده از داده‌های کامل از تصاویر جلوی صورت انجام می‌شود. اگرچه ممکن است در عمل این مورد پیش بیاید، اما موقعیت‌های متعددی وجود دارد که ممکن است تصاویر کامل از جلوی صورت در دسترس نباشد. تصاویر چهره ناقص که اغلب از دوربین‌های مدار بسته می‌آیند، شامل این موارد هستند [۴]. از این رو، مساله تشخیص چهره مبتنی بر رایانه با استفاده از اطلاعات جزئی یا داده‌های ناقص به عنوان شاخصی مهم هنوز هم تا حد زیادی یک حوزه تحقیق ناشناخته است. شناسایی چهره یک فرآیند پیچیده و دشوار است که بیشتر تحت تاثیر تغییرات روشنایی محیط و نور در چهره می‌باشد. در ارتباط بین انسان و ماشین نیازمند تعامل هستیم و ماشین باید هنگام ارتباط چهره انسان را تشخیص دهد، لذا داشتن سیستمی که بتواند چهره انسان را حتی با داده‌های ناقص درست تشخیص دهد، ضروری است. به طور معمول ابتدا تصویر ثابت به سیستم وارد شده و بعد از آشکارسازی ناحیه صورت، نواحی ابروها، چشم‌ها، بینی و دهان آشکار می‌گردد و در این نقاط ویژگی صورت استخراج شده و از فاصله بین نقاط به عنوان ویژگی استفاده می‌گردد. شباهت بین حالات مختلف نیز می‌تواند باعث تشخیص چهره گردد [۵]. تشخیص چهره برای انسان‌ها کاری ساده و در عین حال برای رایانه‌ها چالش برانگیز است. تشخیص چهره از روی تصاویر ثابت امروزه بسیار مورد استفاده قرار گرفته و پژوهش‌های زیادی در این مورد انجام می‌شود که باعث دستاوردهای بسیار دیگری نیز شده است [۶]. تشخیص چهره برای ایجاد تعامل بین رایانه و انسان بسیار مفید است؛ به عنوان مثال می‌توان به تعیین هویت افراد اشاره کرد [۷]. تشخیص چهره در حوزه بینایی ماشین نیز در کاربردهای مانند سیستم‌های نظارتی، استفاده می‌شود [۸]. در دنیای فناوری رویکرد تشخیص امر مهمی بوده و کاربردهای

بسیاری دارد؛ به عنوان مثال در کاربردهای پروفایلینگ تجاری، اهداف مختلف نظارتی مانند نظارت بر تصاویر و ویدیوها [۹]، کاربردهای نظارتی و سیستم‌های تعاملی بین انسان و رایانه [۱۰]، نمونه‌هایی از آنها هستند. در این مطالعه، آزمایشی جدید برای آزمون عملکرد یادگیری ماشین<sup>۱</sup> را با استفاده از تصاویر جزئی چهره و دستکاری‌های دیگر در تصاویر چهره مانند چرخش و بزرگنمایی که به عنوان سرنخ‌های آموزش و تشخیص استفاده شده‌اند [۱۱]، مورد بررسی قرار می‌دهیم. با توجه به موفقیت‌های شبکه عصبی کانولوشنی<sup>۲</sup> در تشخیص و دسته‌بندی [۱۲]، این شبکه‌ها ویژگی‌های سطح بالا و مفید را از تصاویر بانک داده‌های بزرگ استخراج می‌کنند. بنابراین استفاده از این شبکه‌ها در زمینه تشخیص چهره که جز الگوهای مهم یادگیری ماشین است، ضروری به نظر می‌رسد [۱۳]. پیشینه پژوهش‌ها نیز نشان می‌دهد که در این شبکه‌ها، معیار دقت بسیار حائز اهمیت بوده است. از دلایل اصلی انجام این پژوهش انتخاب و طراحی شبکه‌ای است که بتواند ضمن داشتن دقت مناسب، از سرعت خوبی نیز برخوردار باشد. در این مطالعه، ابتدا یک بانک داده استاندارد (FDBB)<sup>۳</sup> [۱۴] انتخاب می‌شود، سپس فراخوانی تصاویر بانک داده، به منظور بهبود کیفیت انجام شده و پیش‌پردازش‌هایی بر روی تصاویر اعمال می‌شود. در واقع در این بخش داده‌های تصویری مورد ترمیم و بازسازی قرار می‌گیرند و برای ادامه کار آماده می‌شوند و در صورت عدم قابلیت استفاده حذف می‌شوند. لیست ویژگی‌های کاربردی بخش‌های مختلف چهره بعد از پیش‌پردازش چهره تهیه می‌شود، سپس با فرآیند طبقه‌بندی [۱۵]، تصاویر به طور تصادفی به تصاویر ناقص تبدیل می‌شوند. با تکمیل فرآیند طبقه‌بندی، واحدهای فعالیت نیز استخراج می‌شوند. برای افزایش دقت تشخیص سیستم، با استفاده از الگوریتم بهینه‌سازی زنبور عسل<sup>۴</sup> [۱۶] بهترین ویژگی‌ها برای شناسایی چهره انتخاب می‌شوند. در نهایت از یک شبکه عصبی کانولوشنی عمیق (DCNN)<sup>۵</sup> در دو مرحله آموزش<sup>۱</sup> و

<sup>۱</sup> Machine Learning

<sup>۲</sup> Convolutional Neural Network (CNN)

<sup>۳</sup> <http://vis-www.cs.umass.edu/fddb>

<sup>۴</sup> Honey-bee Optimization Algorithm

<sup>۵</sup> Deep Convolutional Neural Network

آزمون<sup>۲</sup> برای تشخیص چهره استفاده می‌شود. عملکرد روش ارائه شده توسط معیارهای دقت (Accuracy)، صحت (Precision) و پوشش (Recall) مورد ارزیابی قرار گرفته است. ساختار این مقاله به این شکل است که در بخش دوم، مروری بر روش‌های مختلف در تشخیص چهره با داده‌های ناقص یا اطلاعات جزئی مربوط به سال‌های اخیر مورد بحث و بررسی قرار گرفته است. در بخش سوم، روش پیشنهادی با ارائه یک معماری عمیق مبتنی بر شبکه‌های عصبی کانولوشنی با بهینه‌سازی توسط الگوریتم بهینه‌سازی زنبور عسل بیان شده است. در بخش چهارم، نتایج به دست آمده از روش پیشنهادی ذکر شده است. نتیجه‌گیری و پیشنهادات آتی نیز در بخش پنجم آمده است.

## ۲. پیشینه تحقیق

در این بخش مطالعات اخیر انجام گرفته در حیطه موضوع مورد مطالعه در سال‌های اخیر مورد بحث و بررسی قرار می‌گیرد. در مرجع [۱۷]، یک شبکه عصبی کانولوشنی چند وظیفه‌ای (MTCNN)<sup>۳</sup> برای تشخیص چهره با اطلاعات جزئی صورت پیشنهاد شد. در این پژوهش برای آموزش شبکه عصبی کانولوشنی از مجموعه داده Fddb استفاده شده است. روند اجرای پژوهش به این شکل است که ابتدا تصاویر مجموعه آموزش از بانک داده استخراج شده، سپس پیش‌پردازش‌های لازم بر روی داده‌ها انجام می‌شوند. به منظور افزایش تعداد تصاویر در مجموعه آموزش از روش افزایش داده<sup>۴</sup> استفاده شده است. با آماده‌سازی تصاویر رمزگذاری<sup>۵</sup> شده، از یک شبکه عصبی کانولوشنی چند وظیفه‌ای برای آموزش و یافتن نواحی صورت استفاده می‌شود. نتایج ارزیابی نشان داد که چارچوب پیشنهادی در این پژوهش به دقت طبقه‌بندی ۹۲/۰۶٪ دست یافته است. این رویکرد، تشخیص زود هنگام و طبقه‌بندی صورت با داده‌های جزئی را تسهیل می‌بخشد.

در مرجع [۱۸]، یک روش تشخیص چهره مبتنی بر شبکه عصبی کانولوشنی عمیق (DCNN) با استفاده از استراتژی بخش جعبه‌های لنگر یا انکر باکس‌ها<sup>۶</sup> ارائه شد. در این پژوهش تعداد جعبه‌ها و مقیاس‌ها محدود شده و تنها تشخیص شکل صورت از داده‌های ناقص مورد بررسی قرار گرفته است. مدل پیشنهادی بر روی مجموعه داده‌های معیار تشخیص چهره پرطرفدار و چالش برانگیز Fddb مورد آموزش و آزمون قرار گرفته است. آزمایش‌های گسترده‌ای با معیارهای ارزیابی مختلف بر روی خروجی‌های مدل پیشنهادی انجام شده است. نتایج نشان داد که مدل پیشنهادی قادر است چهره را در تصویر با دقت ۹۴/۰۸٪ بر روی مجموعه داده Fddb تشخیص دهد.

در مرجع [۱۹]، یک شبکه عصبی کانولوشنی مبتنی بر منطقه (RCNN)<sup>۷</sup> برای تشخیص چهره در تصاویر پیچیده با انسداد جزئی صورت استفاده شده است. هدف از این پژوهش بهبود تشخیص چهره در تصاویری با داده‌های ناقص یا اطلاعات جزئی از جمله نور ناهموار، پدیده گم شدن و تشخیص کاذب چهره است. روش پیشنهادی تعداد زیادی فریم پیشنهادی چهره را تولید می‌کند که توسط یک شبکه پیشنهادی منطقه‌ای (RPN)<sup>۸</sup> شناسایی می‌شوند. نتایج آزمایش‌ها و مقایسات نشان می‌دهد که روش تشخیص چهره در این پژوهش بر روی مجموعه داده Fddb می‌تواند به طور موثر از تشخیص نادرست تصاویر کلی و تشخیص نادرست چهره‌های متعدد تحت انسداد جزئی و روشنایی ناهموار جلوگیری کرده و در عین حال دارای دقت تشخیص و استحکام بالایی باشد. روش پیشنهادی قادر است چهره را با دقت ۹۱٪ بر روی مجموعه داده Fddb تشخیص دهد.

در مرجع [۲۰]، از یک شبکه عصبی کانولوشنی آبشاری (Cascaded CNN) با داده‌های ناقص برای تشخیص چهره استفاده شده است. در این پژوهش یک رویکرد آموزش مشترک برای دستیابی به بهینه‌سازی حداکثری انتها به انتها<sup>۹</sup> توسط شبکه

<sup>۶</sup> Anchor Boxes

<sup>۷</sup> Region-based Convolutional Neural Network

<sup>۸</sup> Region Proposal Network (RPN)

<sup>۹</sup> Endo to End

<sup>۱</sup> Train

<sup>۲</sup> Test

<sup>۳</sup> Multi Task Convolutional Neural Network

<sup>۴</sup> Data Augmentation

<sup>۵</sup> Encoding Images

عصبی کانولوشنی آبشاری پیشنهاد شده است. نتایج نشان می‌دهد که الگوریتم انتشار برگشتی مورد استفاده در آموزش شبکه عصبی کانولوشنی می‌تواند به طور طبیعی در آموزش شبکه عصبی کانولوشنی آبشاری نیز مورد استفاده قرار گیرد. نتایج خروجی چگونگی آموزش مشترک بر روی شبکه عصبی کانولوشنی ساده و یک شبکه پیشنهادی منطقه‌ای پیچیده‌تر را نشان می‌دهد. آزمایش‌ها روی معیارهای تشخیص چهره (دقت، صحت و پوشش) مزایای تمرین مشترک را تایید می‌کنند. مدل پیشنهادی قادر است چهره را در تصویر با داده‌های ناقص توسط شبکه عصبی کانولوشنی آبشاری با دقت  $91/02\%$  بر روی مجموعه داده FDDDB تشخیص دهد.

در مرجع [۲۱]، از یک آشکارساز چهره متراکم مبتنی بر شبکه عصبی کانولوشنی عمیق استفاده شده است که نیازی به حاشیه‌نویسی ندارد و قادر است چهره‌ها را در طیف وسیعی از جهت‌گیری‌ها با داده‌های ناقص بر اساس شبکه‌های عصبی کانولوشنی عمیق تشخیص دهد. روش پیشنهادی دارای حداقل پیچیدگی بوده و بر خلاف سایر روش‌های اخیر تشخیص اشیا در یادگیری عمیق، به اجزای اضافی مانند تقسیم‌بندی، رگرسیون جعبه مرزی یا طبقه‌بندی‌کننده‌های مختلف نیاز ندارد. روش پیشنهادی قادر به تشخیص چهره‌ها با داده‌های ناقص از زوایای مختلف است و می‌تواند تا حدودی انسداد را کنترل کند. نمونه‌های مثبت در مجموعه آموزشی و امتیازات آشکارساز چهره پیشنهادی نشان می‌دهد که عملکرد روش پیشنهادی را می‌توان با استفاده از استراتژی‌های نمونه‌گیری بهتر و روش‌های پیچیده‌تر افزایش داده‌ها بهبود داد. ارزیابی‌ها روی مجموعه داده FDDDB نشان می‌دهد که آشکارساز چهره متراکم پیشنهادی توانسته دقت  $85\%$  بر روی این مجموعه داده را به دست آورد.

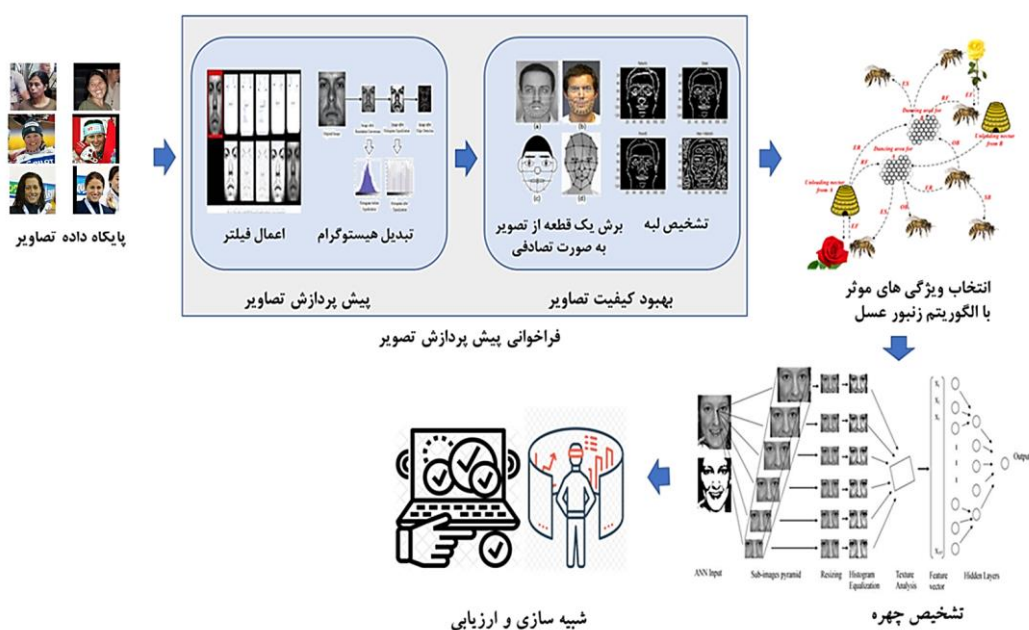
### ۳. روش پیشنهادی

امروزه تشخیص چهره مبتنی بر رایانه یک رویکرد بالغ و قابل اطمینان است که به طور عمده برای بسیاری از سناریوهای کنترل دسترسی مورد استفاده قرار می‌گیرد. به این ترتیب که تشخیص چهره یا احراز هویت عمدتاً با استفاده از داده‌های کامل از

تصاویر جلوی صورت انجام می‌شود. اگرچه ممکن است در عمل این مورد پیش بیاید، اما موقعیت‌های متعددی وجود دارد که ممکن است تصاویر کامل جلوی صورت در دسترس نباشد. تصاویر چهره ناقص که اغلب از دوربین‌های مدار بسته می‌آیند، شامل این موارد هستند. با توجه به این که به طور ذاتی انسان‌ها و رایانه‌ها در تشخیص چهره و احراز هویت متفاوت هستند، باید جالب و جذاب باشد که بدانند یک رایانه زمانی که با یک چالش تشخیص چهره روبرو می‌شود، چگونه به اجزای مختلف صورت توجه می‌کند. در این مطالعه، پاسخ به این پرسش بررسی شده که در آن برای روش تشخیص چهره از اطلاعات جزئی صورت استفاده شده است. این مساله را با استفاده از آزمایش‌های جدید برای آزمون عملکرد یادگیری ماشین، تصاویر جزئی چهره و دستکاری‌های دیگر در تصاویر چهره مانند چرخش و بزرگنمایی، که به عنوان سرنخ‌های اصلی آموزش و تشخیص استفاده می‌شوند، مورد بررسی قرار می‌دهیم. با توجه به اینکه روش شبکه‌های عصبی کانولوشنی در دسته‌بندی و تشخیص اشیا خوب عمل می‌کند، واضح است که این شبکه‌ها ویژگی‌های سطح بالا و موثری را می‌تواند از تصاویر بانک داده بزرگ استخراج کند. بنابراین، بهره‌گیری از این شبکه‌ها در مورد تشخیص چهره که یکی از الگوهای مهم یادگیری ماشین است، به نظر ضروری می‌رسد. همچنین ادبیات پژوهش‌های بررسی شده در بخش ۲ نشان می‌دهد که در حوزه دسته‌بندی به معیار دقت به طور ویژه توجه شده است. در نتیجه، انتخاب و طراحی شبکه‌ای عمیق نیز یک امر واجب و ضروری به نظر می‌رسد تا ضمن داشتن دقت مناسب، قادر به حفظ سرعت خوبی نیز باشد. در این تحقیق، ابتدا یک بانک داده استاندارد انتخاب شده و بعد از فراخوانی تصاویر، جهت بهبود کیفیت، پیش‌پردازش‌هایی بر روی تصاویر صورت می‌گیرد. در این بخش داده‌های معیوب نیز مورد ترمیم و بازسازی قرار می‌گیرند و در صورت عدم قابلیت استفاده، حذف خواهند شد. همچنین در مرحله نخست پس از پیش‌پردازش، لیست خصوصیات کاربردی اجزای چهره تهیه شده و با استفاده از قطعه‌بندی، واحدهای فعالیت مشخص شده و جهت افزایش دقت تشخیص سیستم، از الگوریتم بهینه‌سازی

استخراج نقاط مهم در تصویر انجام می‌شود. خروجی گام قبلی به عنوان ورودی به الگوریتم زنبور عسل داده می‌شود تا انتخاب ویژگی‌های موثر و شناخت آنها در راستای بهینه‌سازی فرآیند تشخیص انجام گیرد. در نهایت از شبکه عصبی کانولوشنی عمیق برای تشخیص چهره استفاده می‌شود. این تشخیص در دو مرحله آموزش و آزمون انجام می‌پذیرد. برای درک بهتر رویکرد پیشنهادی خلاصه‌ای از روند کلی روش پیشنهادی در شکل (۱) نمایش داده شده است.

زنبور عسل برای شناسایی چهره، بهترین ویژگی موثر انتخاب می‌شود. در نهایت از یک شبکه عصبی کانولوشنی عمیق برای فرآیند یادگیری و تشخیص چهره استفاده می‌شود. مطابق با هدف اصلی تحقیق که بهبود تشخیص چهره به کمک داده‌های ناقص می‌باشد، روش پیشنهادی از چند گام مختلف تشکیل شده است. به این طریق که پس از انتخاب و استخراج داده اولیه به کمک بانک داده منتخب اقدام به پیش‌پردازش تصاویر می‌شود. در این مرحله اعمال فیلتر، تبدیل هیستوگرام، برش قطعات تصویر و تشخیص لبه انجام می‌گردد. در مرحله بعد



شکل (۱): خلاصه‌ای از روند کلی روش پیشنهادی

برای کاهش نویز در تصاویر استفاده می‌گردد و تا حدی عملکرد مشابه با فیلتر میانگین<sup>۲</sup> دارد؛ با این تفاوت که در مقایسه با فیلتر میانگین جهت حفظ جزئیات مفید تصویر عملکرد بهتری دارد. لازم به ذکر است که منظور از میانه، مقدار متوسط نیست؛ بلکه مقدار میانه از مقدار نیمی از پیکسل‌ها بزرگتر و از مقدار نیمی دیگر کوچکتر می‌باشد. مقدار میانه در مقایسه با مقدار میانگین، شاخص مرکزی قویتری است و به راحتی تحت تاثیر مقادیر پراکنده در میان پیکسل‌های موجود در محدوده همسایگی قرار نمی‌گیرد. در این مقاله نخست به مرتب‌سازی مقدار پیکسل‌های

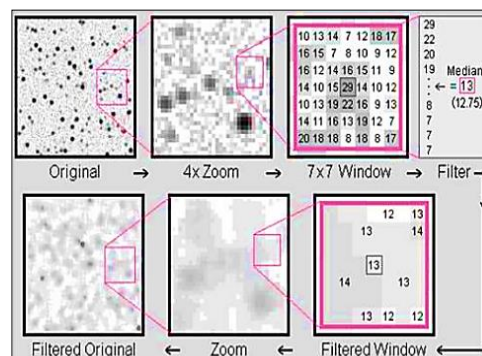
فیلتر میانه<sup>۱</sup> تحت عنوان روشی پیشرفته برای پیش‌پردازش شناخته می‌شود. این فیلتر قادر به حفظ جزئیات یک تصویر با دقت بالا است. با استفاده از آن در اطراف هر پیکسل در یک تصویر دیجیتال، محدوده‌ای مشخص می‌شود که مقادیر کل پیکسل‌های آن محدوده بر اساس یک فلوچارت مشخص تحلیل می‌شوند و سپس مقادیر اصلی هر پیکسل با مقدار حاصل از تحلیل روی پیکسل‌های محدوده همجوار جایگزین می‌شوند. روی پیکسل‌های هر تصویر، این محدوده حرکت کرده و برای هر پیکسل این رویه تکرار می‌شود. از فیلتر میانه به طور معمول

<sup>2</sup> Mean

<sup>1</sup> Median

وجود دارد که عبارتند از نقطه، خط و لبه. ماسک‌های فضایی می‌توانند برای تشخیص هر سه نوع ناپیوستگی در یک تصویر بکار روند. رویکرد تشخیص لبه کنی<sup>۱</sup> یکی از روش‌های استاندارد تشخیص لبه می‌باشد. این روش، به دلیل یافتن لبه‌ها با ارزش‌های مختلف بدون تخریب ویژگی‌های لبه در تصویر، نسبت به روش‌های دیگر عملکرد بهتری دارد. در این مطالعه، برای تشخیص لبه ابتدا برش قسمتی از تصویر به صورت تصادفی انجام شده (جهت ایجاد داده‌های ناقص)، سپس روش تشخیص لبه کنی در نرم‌افزار متلب برای لبه‌یابی اعمال شده است. در مدل پیشنهادی، از الگوریتم بهینه‌سازی زنبور عسل برای یافتن ویژگی‌های بهینه استفاده شده است. کلونی زنبور عسل مصنوعی دارای سه نوع زنبور کارگر، ناظر و دیده‌بان است. زنبورهای کارگر به گردآوری غذا و آوردن آن به کندو از یک منبع غذایی خاص تمرکز دارند. همچنین زنبورهای ناظر مابین زنبورهای کارگر می‌گردند تا تشخیص دهند یک منبع غذایی همچنان دارای ارزش است یا خیر و زنبورهای دیده‌بان نیز منابع غذایی جدید را جستجو می‌کنند. در فلوجارت زنبور عسل، یک منبع غذایی را به عنوان حالتی در فضای جست‌وجو تعریف کرده و در ابتدا تعداد منابع غذایی با تعداد زنبورهای موجود در کندو برابری می‌کند. کیفیت منابع غذایی به وسیله میزان تابع هدف در آن موقعیت (مقدار تناسب) مشخص می‌شود. برخورد هوشمندانه ناپایدار زنبورهای عسل در مدل پیشنهادی را می‌توان در چند گام خلاصه کرد. زنبورها در تلاش هستند تا به صورت اتفاقی در فضا به دنبال منابع غذایی خوب (مقدار تناسب) باشند. پس از پیدا کردن یک منبع غذایی، آنها به زنبورهای کارگر تبدیل شده و به استخراج غذا از منبع پیدا شده مشغول می‌شوند. زنبور کارگر حامل شهد به کندو بازگشته و بار شهد خود را تخلیه می‌کند. پس از تخلیه شهد، قادر است مستقیماً به منبع کشف شده خود برگردد یا داده‌هایی که درباره منبع غذایی خود دارد را به وسیله اجرای یک رقص گردون در ناحیه رقص با دیگران به اشتراک بگذارد. حال اگر یک منبع غذایی خالی شود، زنبورهای کارگر تبدیل به دیده‌بان شده و به طور تصادفی مشغول به

موجود در همسایگی که به صورت صعودی انتخاب شده‌اند، پرداخته می‌شود و پس از معین نمودن میانه پیکسل‌های محدوده همسایگی، مقدار پیکسل مطلوب با این مقدار میانه جایگزین می‌گردد. حال چنانچه تعداد این پیکسل‌ها در محدوده همسایگی زوج باشند، میانگین دو مقدار میانه، تحت عنوان میانه استفاده می‌شود. در شکل (۲) مثالی از این روش نشان داده شده است.



شکل (۲): عملکرد فیلتر میانه

از تبدیل هیستوگرام به عنوان روشی رایج برای بهبود ظاهر تصویر استفاده می‌شود. با فرض اینکه تصویری که در اختیار است، در ابتدا تاریک است، در نتیجه هیستوگرام تصویر مذکور، به سوی انتهای پایینی مقیاس خاکستری اریب شده و کل جزئیات تصویر در انتهای تاریک هیستوگرام کمپرس می‌شوند. بنابراین اگر بتوان سطوح خاکستری در انتهای تاریک را به طریقی کشید که توزیع در هیستوگرام یکنواخت‌تر شود، آنگاه تصویر واضح‌تری به دست خواهد آمد. عملیات کشش هیستوگرام در طیف کلی پیکسل‌ها (صفر تا ۲۵۵) را تبدیل هیستوگرام انجام می‌دهد و می‌تواند بنا به مشخصه هیستوگرام یا اصلاح آن، تصویر جدیدی را کسب کند. تشخیص لبه یک روش بنیادی جهت تقسیم‌بندی تصاویر می‌باشد. روش‌های تشخیص لبه، تصویر اصلی را تبدیل به لبه‌هایی از تصویر می‌کنند که از میزان تغییرات خاکستری در تصویر مفیدتر است. در پردازش تصویر به خصوص در بینایی ماشین، تشخیص لبه برای جایابی متغیرهای مهم از یک سطح خاکستری در تصویر و نیز تشخیص خصوصیات فیزیکی و ژئومتریک اشیای در صحنه مورد استفاده قرار می‌گیرد. سه نوع مختلف از ناپیوستگی در سطح خاکستری

<sup>۱</sup> Canny

عنوان تابع غیرخطی استفاده می‌شود. به طور کلی در شبکه‌های عصبی کانولوشنی عمیق از تابع یک‌سوساز نسبت به دیگر توابع غیرخطی بیشتر استفاده می‌شود، زیرا این تابع محاسبه ساده‌ای دارد و منابع محاسباتی را زیاد درگیر نمی‌کند؛ همچنین تجربه نشان داده است که استفاده از این تابع منجر به رسیدن به دقت‌های قابل قبولی می‌شود. رابطه (۱) تعریف تابع یک‌سوساز را نمایش می‌دهد.

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{o.w.} \end{cases} \quad (1)$$

در نهایت مقادیر به دست آمده جهت انجام عملیات ادغام در اختیار زیرلایه سوم قرار می‌گیرد. در زیرلایه ادغام چکیده آماری پیکسل‌های همجوار پیکسل مرکزی، با مقادیر پیکسل مرکزی در پنجره ادغام جایگزین می‌شوند. اعمال ادغام منجر به پایداری ویژگی‌ها و کاهش حساسیت نسبت به تغییرات ناخواسته می‌شود. ابعاد داده خروجی در این زیرلایه می‌تواند با توجه به مقدار متغیرهای مرحله قبل و صفرگذاری مساوی با ابعاد داده ورودی باشد و یا ابعاد متفاوتی داشته باشد. به همین دلیل است که اگر در مرحله ادغام کاهش ابعاد صورت پذیرد، می‌توان کاربرد دیگر این زیرلایه را حفظ ویژگی‌های برتر و حذف ویژگی‌های با اهمیت پایین تلقی کرد. نوع‌های مختلفی از ادغام وجود دارند که محبوب‌ترین آنها ادغام حداکثری یا max pooling می‌باشد. این ادغام بیشترین مقدار پیکسل بین پیکسل‌های منتخب را به عنوان داده ارزشمند انتخاب می‌کند. حال به طور کامل یک لایه کانولوشنی روی تصویر پیش‌پردازش شده اعمال گشته است. تعداد تکرارها در مدل پیشنهادی به ترتیب ۱۰۰، ۱۵۰ و ۲۰۰ اپیاک<sup>۵</sup> یا دور بوده و بعد از این تعداد تکرار، آموزش شبکه به حالت بهینه رسیده است. در انتهای شبکه به منظور نگاشت نهایی ویژگی‌های کسب شده، از لایه کاملاً متصل استفاده می‌شود. یک لایه‌ای که کاملاً متصل است، دقیقاً مانند یک لایه کانولوشنی می‌باشد با این اختلاف که تعاملات پراکنده در اینجا انجام نمی‌پذیرد و درست همانند شبکه‌های عصبی مرسوم میان این لایه و لایه قبل از آن یک اتصال کامل

جست‌وجوی منابع غذایی می‌شوند. زنبورهای ناظر در کندو منتظر می‌مانند تا زنبورهای کارگر را در منابع غذایی گردآوری کرده‌شان مورد نظارت قرار داده و از میان منابع غذایی موجود با بیشترین سود، یک منبع خوب را انتخاب کنند. انتخاب منابع غذایی متناسب با کیفیت آن منبع (یعنی مقدار تناسب) می‌باشد. همچنین رفتار اصلی یک زنبور کارگر استخراج غذا از یک منبع غذایی می‌باشد به طوری که زنبورهای کارگر تا مرحله خالی شدن منبع کار می‌کنند. در مرحله پیاده‌سازی، این عمل را می‌توان به عنوان تولید موقعیت‌های جدید در نزدیکی محلی که زنبورهای کارگر در حال کار هستند، قلمداد کرد و این طور ارزیابی کرد که آیا این محل جدید میزان بهتری غذا آماده می‌کند یا خیر. شبکه‌های عصبی کانولوشنی از جمله روش‌های یادگیری عمیق می‌باشند که در آنها چند لایه با روشی قدرتمند آموزش می‌بینند. یک شبکه عصبی کانولوشنی از سه لایه اصلی تشکیل شده که عبارتند از: لایه کانولوشن<sup>۱</sup>، لایه پولینگ<sup>۲</sup> یا ادغام و لایه تماماً متصل<sup>۳</sup>. در روش پیشنهادی تصاویر به دست آمده از مرحله قبل به اندازه استاندارد تغییر یافته و به ورودی لایه اول کانولوشنی داده می‌شوند. در مرحله دوم، کرنل‌ها با کرنل‌های کانولوشنی لایه اول کانوالو می‌شوند. هر لایه کانولوشنی از سه زیرلایه تشکیل شده است. زیرلایه اول تصویر پیش‌پردازش شده با کرنل‌های کانولوشنی از پیش تعیین شده را کانوالو می‌کند که مقدار و ابعاد این کرنل‌ها با یکدیگر متفاوت می‌باشند. بسته به تعداد صفرگذاری تصویر، ابعاد خروجی یا هم‌اندازه تصویر ورودی هستند و یا کوچکتر و بزرگتر می‌باشند. می‌توان در این بخش از چند کرنل کانولوشنی استفاده کرد. عملگر کانولوشن، عملگری است که بر روی داده‌ها به صورت محلی عمل کرده و تغییرناپذیر با جابجایی است. خروجی زیرلایه اول، بعد از زیرلایه اول کانولوشنی به زیرلایه غیرخطی منتقل می‌شود. در این مرحله بر روی مقادیر به دست آمده، تابع فعال‌ساز غیرخطی اعمال خواهد شد. این مطالعه از تابع یک‌سوساز ReLU<sup>۴</sup> به

<sup>1</sup> Convolution Layer

<sup>2</sup> Pooling Layer

<sup>3</sup> Fully Connected Layer

<sup>4</sup> Rectified Linear Unit (ReLU)

<sup>5</sup> Epoch

تصاویر مربوط به خروجی مدل پیشنهادی، پرداخته شده است.

جدول (۱): جزئیات معماری شبکه عصبی کانولوشنی عمیق پیشنهادی

| لایه        | فیلتر | گام | پد | حذف | حجم      | متغیرها |
|-------------|-------|-----|----|-----|----------|---------|
| INPUT       | -     | -   | -  | -   | ۵۲×۵۲×۱  | -       |
| CONV → RELU | ۵×۵   | ۲   | ۲  | ۱   | ۲۶×۲۶×۳۲ | ۸۳۲     |
| CONV → RELU | ۳×۳   | ۱   | ۱  | ۱   | ۲۶×۲۶×۳۲ | ۹,۲۴۸   |
| MP          | ۳×۳   | ۲   | ۱  | ۱   | ۱۳×۱۳×۶۴ | ۱۸,۴۹۶  |
| CONV → RELU | ۳×۳   | ۱   | ۱  | ۱   | ۱۳×۱۳×۶۴ | ۳۶,۹۲۸  |
| CONV → RELU | ۳×۳   | ۱   | ۲  | ۲   | ۱۳×۱۳×۶۴ | ۵۵,۳۹۲  |
| MP          | ۳×۳   | ۱   | ۲  | ۲   | ۱۳×۱۳×۹۶ | ۸۳,۰۴۰  |
| FC → RELU   | ۵×۵   | ۱   | ۶  | ۳   | ۱۳×۱۳×۱  | ۲,۴۰۱   |
| Softmax     | -     | -   | -  | -   | ۵۲×۵۲×۱  | -       |

\* Conv: Convolutional Layer

\* MP: Max Pooling Layer

\* ReLU (Rectified Linear Unit): Activation Function

\* FC: Fully Connected Layer

#### ۱.۴. بانک داده

مجموعه داده‌ها و معیار تشخیص چهره <sup>۴</sup>FDDDB مجموعه‌ای از چهره‌های برجسب‌گذاری شده از تصاویر چهره مختلف در یک بانک داده وسیع می‌باشد. این بانک داده حاوی حاشیه‌نویسی برای ۵۱۷۱ چهره در مجموعه‌ای از ۲۸۴۵ تصویر است که تصاویر آن دارای دو وضوح متفاوت با مقادیر ۳۶۳×۴۵۰ و ۲۲۹×۴۱۰ هستند. مجموعه داده شامل طیف وسیعی از چالش‌ها، از جمله زوایای مختلف و گاهی پیچیده، چهره‌های خارج از فوکوس و چهره‌هایی با وضوح پایین است. از هر تصویر دو نمونه سطح خاکستری و رنگی در این بانک داده قرار دارد. چهره‌های موجود در این بانک داده، پراکندگی زیاد در جهت صورت، پوشش و اندازه دارند. به غیر از برجسب مکانی، هر چهره در تصویر، تعدادی برجسب دیگر هم برای وضوح، اندازه چهره و پوشش دارد. به طور معمول تصویرهای موجود در این بانک داده، به سه دسته آسان، متوسط و سخت بر اساس برجسب‌های اضافی ذکر شده، طبقه‌بندی می‌گردند. در این تحقیق، تعدادی از چهره‌های این بانک داده حذف شده است و

صورت می‌گیرد. خروجی آخرین لایه یک بردار یک بعدی بوده و تعداد مولفه‌های این بردار با تعداد کلاس‌های دسته‌بندی برابر می‌باشد. در مدل پیشنهادی تعداد لایه‌ها برابر با تعداد ویژگی‌های تصویر چهره و کلاس خروجی برابر با تصویر شناسایی شده می‌باشد. در حقیقت این لایه عمل دسته‌بندی را انجام می‌دهد. در کل ساختارهای کانولوشنی آخرین لایه کاملاً متصل به یک لایه softmax متصل می‌شود. لایه softmax در حقیقت عمل طبقه‌بندی در شبکه‌های کانولوشنی را انجام می‌دهد. این لایه مشتمل بر تعدادی نورون مساوی با تعداد کلاس‌های مساله دسته‌بندی بوده و به منظور نگاشت نهایی ویژگی و اعمال دسته‌بندی مورد استفاده قرار می‌گیرد. در این مقاله، جهت تنظیم متغیرهای شبکه و یا همان آموزش شبکه، از نتایج خروجی برای محاسبه میزان خطای شبکه استفاده می‌شود. از این رو خروجی شبکه را توسط یک تابع خطا با پاسخ صحیح مقایسه کرده و به این شکل میزان خطا را محاسبه می‌کنیم. نحوه محاسبه خطای تجربی توسط رابطه (۲) نمایش داده شده است.

$$l(w) = \frac{1}{n} \sum_{i=1}^n l(Z_i, f(X_i, w)) \quad (2)$$

در این عبارت  $l(Z_i, \hat{Z})$  یک تابع تلفات است که مقدار جریمه را در زمان پیش‌بینی اشتباه  $\hat{Z}$  به جای  $Z$  معین می‌نماید. در مرحله بعدی با توجه به مقدار خطای محاسبه شده مرحله پس انتشار شروع می‌شود. پس از تکرار شدن تعداد مناسبی از این مرحله‌ها، آموزش شبکه پایان می‌یابد. جدول (۱) جزئیات معماری شبکه عصبی کانولوشنی عمیق پیشنهادی را نمایش می‌دهد. در این جدول اندازه فیلتر، گام<sup>۱</sup>، پد<sup>۲</sup> و حذف<sup>۳</sup> در هر لایه و همچنین حجم و تعداد مولفه‌های پد در هر لایه نشان داده شده است.

#### ۴. نتایج تجربی

در این بخش به بیان نتایج به دست آمده از روش پیشنهادی، معرفی بانک داده، محیط شبیه‌سازی و نمایش نمودارها و

<sup>۱</sup> Stride

<sup>۲</sup> Zero padding

<sup>۳</sup> Drop-out

<sup>۴</sup> Face Detection Data Set and Benchmark

به همین جهت تصاویر مورد استفاده در این پژوهش، در دسته سخت جای دارند. بهبود تشخیص چهره بر روی تصاویر با داده‌های ناقص توسط شبکه عصبی کانولوشنی عمیق (DCNN) در اینجا با توجه به گام‌های مطرح شده در روش پیشنهادی به صورت مرحله به مرحله انجام شده است. از مجموعه تصاویر انتخاب شده از این بانک داده ۷۰ درصد برای آموزش، ۳۰ درصد برای آزمون و بخشی از تصاویر به صورت تصادفی برای اعتبارسنجی انتخاب شده‌اند که این تعداد شامل ۲۰۰ تصویر است. تصاویری که جهت آزمون شبکه کنار گذاشته می‌شوند در فرآیند آموزش اولیه دخالتی ندارند. نکته‌ای که در شبیه‌سازی انجام شده باید به آن اشاره کرد، استخراج ویژگی‌های موثر در بهبود کیفیت تصویر برای تشخیص بهینه با استفاده از الگوریتم بهینه‌سازی زنبور عسل و تعداد تکرارها است که باعث شده با افزایش تعداد تکرار که بالطبع هزینه بالاتر می‌رود، بهینگی محلی و قحطی‌زدگی اتفاق نیفتد. این امر در نهایت به آموزش و یادگیری بهینه‌تری توسط شبکه عصبی کانولوشنی عمیق منجر می‌شود. در ادامه، شکل (۳) نمونه‌های از تصاویر بانک داده مورد استفاده و تغییرات تصویر را نمایش می‌دهد.



شکل (۳): نمونه‌هایی از تصاویر بانک داده Fddb

## ۲.۴. محیط شبیه‌سازی

روش پیشنهادی توسط نرم‌افزار متلب نسخه R2016a در محیط ویندوز و توسط یک رایانه استاندارد با پردازنده Intel Core i7-5500U 2.40GHz و رم ۸ گیگابایت پیاده‌سازی شده است.

## ۳.۴. معیارهای ارزیابی

در این مطالعه، از معیارهای دقت (Accuracy)، صحت (Precision) و پوشش (Recall) برای ارزیابی عملکرد مدل عمیق پیشنهادی استفاده شده است. معیار دقت به معنای تعداد نمونه‌های درست طبقه‌بندی شده نسبت به کل نمونه‌ها می‌باشد.

معیار صحت مشخص می‌کند که الگوریتم چند درصد «بله» هایش درست است؛ یعنی تمرکز اصلی آن، بر روی صحت تشخیص‌های «بله» توسط فلوچارت است. همچنین تاکید اصلی معیار پوشش برعکس معیار صحت روی داده‌هایی است که واقعا «بله» بوده‌اند. در تمام معیارهای ارزیابی فوق، اگر نمونه مثبت باشد و تحت عنوان مثبت هم طبقه‌بندی شده باشد، مثبت حقیقی محسوب خواهد شد و اگر نمونه به صورت منفی طبقه‌بندی شود، منفی کاذب می‌باشد. حال اگر نمونه منفی بوده و به عنوان منفی نیز طبقه‌بندی شود، آنگاه منفی حقیقی محسوب می‌شود و اگر این نمونه به صورت مثبت طبقه‌بندی شود، مثبت کاذب خواهد بود. مثبت حقیقی (TP)<sup>۱</sup>، تعداد پیش‌بینی‌های صحیح مربوط به همان کلاس را نشان می‌دهد، منفی حقیقی (TN)<sup>۲</sup>، تعداد درست پیش‌بینی شده به غیر از کلاس فعلی می‌باشد، مثبت کاذب (FP)<sup>۳</sup>، نیز تعداد پیش‌بینی نادرست مربوط به کلاس‌های دیگر را نشان می‌دهد و در نهایت منفی کاذب (FN)<sup>۴</sup>، تعداد پیش‌بینی نادرست مربوط به همان کلاس را نشان می‌دهد. برای ارزیابی فاکتورهای مورد بررسی توسط مدل‌های پیشنهادی، معیار دقت توسط رابطه (۳)، صحت توسط رابطه (۴) و پوشش توسط رابطه (۵) تعریف و محاسبه می‌شوند.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (۳)$$

$$Precision = \frac{TP}{TP + FP} \quad (۴)$$

$$Recall = \frac{TP}{TP + FN} \quad (۵)$$

## ۴.۴. خروجی‌های روش پیشنهادی

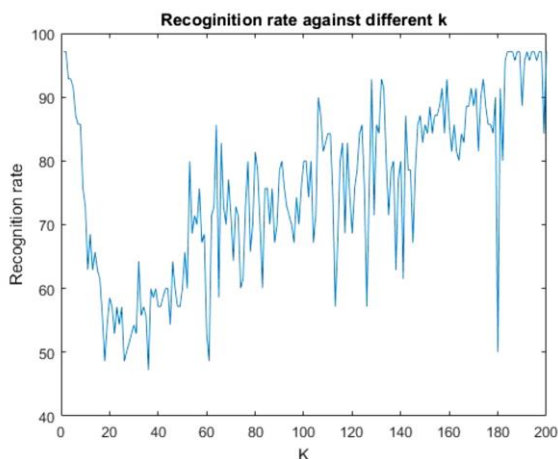
در سناریوی اول تعداد تکرار ۱۰۰ و تعداد زنبورها ۳۰ عدد در نظر گرفته شده است. در سناریوی‌های بعدی تعداد تکرار به ترتیب ۱۵۰ و ۲۰۰ در نظر گرفته شده است. تا این مرحله پیش-پردازش و اعمال فیلتر بر روی چهره مورد نظر انجام شده است. همچنین برش یک قطعه از تصویر به صورت تصادفی با تطابق

<sup>۱</sup> True Positive

<sup>۲</sup> True Negative

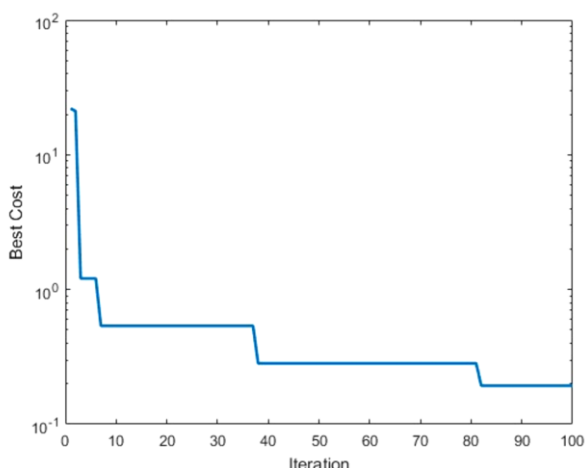
<sup>۳</sup> False Positive

<sup>۴</sup> False Negative



شکل (۶): نرخ تشخیص در برابر K لایه مختلف

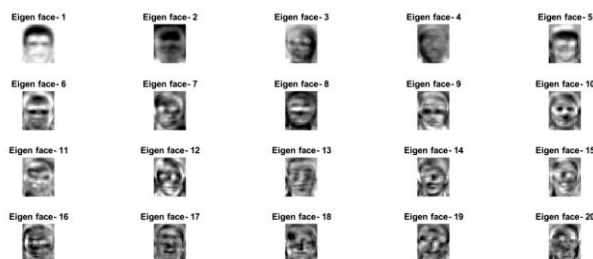
در شکل (۷) نشان داده شده است که با افزایش تعداد تکرار در انتخاب بهینه ویژگی‌ها جهت تشخیص توسط الگوریتم بهینه‌سازی زنبور عسل هزینه تمام شده، یک روند نزولی در پیش گرفته است.



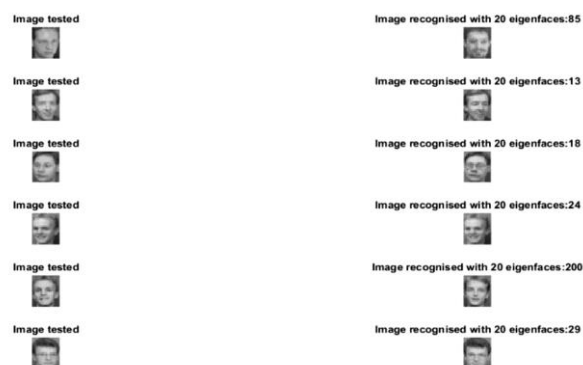
شکل (۷): نمایش نسبت تعداد تکرار به هزینه

جدول (۲) نمونه‌هایی از چند تصویر با داده‌ها کامل و ناقص و نتایج حاصل از تشخیص چهره توسط شبکه عصبی کانولوشنی عمیق پیشنهادی را نمایش می‌دهد. در ادامه این بخش به دلیل تشابه خروجی‌های مورد نظر در تکرارهای دیگر، نمودارهای خروجی نشان داده نشده است، اما در قالب یک جدول کلی میزان هزینه در تعداد تکرار ۱۰۰ در سناریوی اول توسط جدول (۳) نشان داده شده است. برای تکرارهای ۱۵۰ و ۲۰۰ نیز همین روند تکرار شده است.

و استخراج چهره‌های مشابه با ویژگی‌های مورد نظر، بررسی می‌شود. به کمک شبکه عصبی کانولوشنی عمیق پیشنهادی آموزش و آزمون، در دو مرحله مجزا، فرآیند تشخیص را تکمیل می‌کنند. در شکل (۴) و (۵) بخشی از این دو فرآیند به ترتیب نشان داده شده است.



شکل (۴): بخش از فرآیند آموزش (TRAIN)



شکل (۵): بخش از فرآیند آزمون (TEST)

همان‌طور که در شکل (۵) مشخص شده است، تصاویر آزمون شده با بانک داده مورد بررسی توسط ویژگی‌های مختلف بررسی شده و از دید شباهت به صورت تعدادی مشخص در نظر گرفته شده‌اند. نکته‌ای که در اینجا مهم است، عمق شبکه عصبی کانولوشنی عمیق پیشنهادی است. به عبارتی، با توجه به عمق لایه‌های پنهان و برش انجام شده از تصویر اولیه شناسایی و بررسی به کمک فیلتراسیون و دیگر مراحل تعریف شده به خوبی انجام شده است. در ادامه در شکل (۶)، نرخ شناسایی با  $k$  لایه مختلف نشان داده شده است. این تعداد لایه‌ها همان نورون‌ها و لایه‌های پنهان در شبکه عصبی کانولوشنی عمیق هستند. همان‌طور که مشخص است با افزایش  $k$ ، نرخ شناسایی نیز روند صعودی داشته است. می‌توان گفت وضعیت نوسانی رو به بهبود برای نرخ تشخیص چهره تشکیل شده است.

است، اما کلیت در یک بازه عددی و رویکرد مشخص شده، حاصل می‌شود و به عبارتی در هر اجرا دقیقاً نتایج مندرج در این جدول دیده نخواهد شد و داده‌ها و خروجی‌ها با کمترین میزان اختلافی نسبت به همدیگر حاصل می‌شوند.

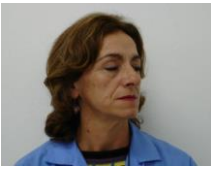
جدول (۳): خروجی هزینه برای سناریوی اول با تعداد تکرار ۱۰۰

| تکرار                                                   | هزینه   | تکرار | هزینه   | تکرار | هزینه   |
|---------------------------------------------------------|---------|-------|---------|-------|---------|
| 1                                                       | 0.19313 | 34    | 0.28303 | 67    | 0.53704 |
| 2                                                       | 0.19313 | 35    | 0.28303 | 68    | 0.28303 |
| 3                                                       | 0.19313 | 36    | 0.28303 | 69    | 0.28303 |
| 4                                                       | 0.19313 | 37    | 0.28303 | 70    | 0.28303 |
| 5                                                       | 0.19313 | 38    | 0.28303 | 71    | 0.53704 |
| 6                                                       | 0.19313 | 39    | 0.28303 | 72    | 0.53704 |
| 7                                                       | 0.19313 | 40    | 0.28303 | 73    | 0.53704 |
| 8                                                       | 0.19313 | 41    | 0.28303 | 74    | 0.53704 |
| 9                                                       | 0.19313 | 42    | 0.28303 | 75    | 0.53704 |
| 10                                                      | 0.19313 | 43    | 0.28303 | 76    | 0.53704 |
| 11                                                      | 0.28303 | 44    | 0.28303 | 77    | 0.53704 |
| 12                                                      | 0.19313 | 45    | 0.28303 | 78    | 0.53704 |
| 13                                                      | 0.19313 | 46    | 0.28303 | 79    | 0.53704 |
| 14                                                      | 0.19313 | 47    | 0.28303 | 80    | 0.53704 |
| 15                                                      | 0.19313 | 48    | 0.28303 | 81    | 0.53704 |
| 16                                                      | 0.19313 | 49    | 0.28303 | 82    | 0.53704 |
| 17                                                      | 0.19313 | 50    | 0.28303 | 83    | 0.53704 |
| 18                                                      | 0.19313 | 51    | 0.28303 | 84    | 0.53704 |
| 19                                                      | 0.19313 | 52    | 0.28303 | 85    | 0.53704 |
| 20                                                      | 0.19313 | 53    | 0.28303 | 86    | 0.53704 |
| 21                                                      | 0.28303 | 54    | 0.28303 | 87    | 0.53704 |
| 22                                                      | 0.28303 | 55    | 0.28303 | 88    | 0.53704 |
| 23                                                      | 0.28303 | 56    | 0.28303 | 89    | 0.53704 |
| 24                                                      | 0.28303 | 57    | 0.28303 | 90    | 0.53704 |
| 25                                                      | 0.28303 | 58    | 0.28303 | 91    | 22.0285 |
| 26                                                      | 0.28303 | 59    | 0.28303 | 92    | 21.1192 |
| 27                                                      | 0.28303 | 60    | 0.28303 | 93    | 1.2069  |
| 28                                                      | 0.28303 | 61    | 0.53704 | 94    | 1.2069  |
| 29                                                      | 0.28303 | 62    | 0.53704 | 95    | 1.2069  |
| 30                                                      | 0.28303 | 63    | 0.53704 | 96    | 1.2069  |
| 31                                                      | 0.28303 | 64    | 0.53704 | 97    | 0.53704 |
| 32                                                      | 0.28303 | 65    | 0.53704 | 98    | 0.53704 |
| 33                                                      | 0.28303 | 66    | 0.53704 | 99    | 0.53704 |
| The best Accuracy obtained by DCNN<br>Model is: 90.0899 |         |       |         | 100   | 0.53704 |

جدول (۴): میانگین خروجی مربوط به معیارهای دقت، صحت و پوشش برای مدل شبکه عصبی کانولوشنی عمیق پیشنهادی

| تکرار | هزینه   | دقت (%) | صحت (%) | پوشش (%) |
|-------|---------|---------|---------|----------|
| ۱۰۰   | ۰/۸۰۷۴۶ | ۹۰/۰۸۹۹ | ۹۰/۳۷۶۱ | ۹۱/۲۴۰۰  |
| ۱۵۰   | ۰/۸۰۸۱۹ | ۹۴/۷۰۸۰ | ۹۳/۰۷۴۰ | ۹۴/۱۵۱۰  |
| ۲۰۰   | ۰/۵۳۲۶۳ | ۹۶/۱۱۵۰ | ۹۵/۶۷۷۰ | ۹۵/۴۱۵۰  |

جدول (۲): نمونه‌هایی از نتایج تصویری به دست آمده توسط شبکه عصبی کانولوشنی عمیق پیشنهادی

| نمونه تصویر                                                                           | خروجی تشخیص چهره                                                                     | نوع داده‌ها |
|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-------------|
|    |    | کامل        |
|    |    | ناقص        |
|    |    | کامل        |
|   |   | ناقص        |
|  |  | کامل        |
|  |  | ناقص        |

با توجه به نتایج به دست آمده، میانگین کل از معیارهای ارزیابی دقت، صحت و پوشش برای ارزیابی عملکرد مدل پیشنهادی توسط سه سناریوی بررسی شده و خلاصه نتایج در جدول (۴) نشان داده شده است. خروجی به دست آمده بیانگر این است که با افزایش تعداد تکرار در راستای تشخیص چهره هزینه کاهش داشته و دقت تشخیص و صحت و پوشش نیز یک روند بهینه و قابل قبولی در پیش گرفته است. لازم به ذکر است که در هر اجرا نتایج بالطبع تفاوت خواهد داشت و این به خاطر برش تصادفی تصویر و شناسایی و استخراج شاخص‌های مهم در تصویر اولیه

مهم‌ترین یافته پژوهش حاضر بدین شرح است که با عمیق کردن شبکه عصبی کانولوشنی و کنترل بیش‌برازش<sup>۱</sup> می‌توان به دقت بالاتری رسید. نتایج به دست آمده در جدول (۴) نشان می‌دهد که مدل عمیق پیشنهادی دارای کارایی مناسبی است. جدول (۵) خلاصه‌ای از عملکرد مدل عمیق پیشنهادی، در مقایسه با سایر کارهای مورد مطالعه در بخش ۲ مربوط به این حوزه در سال‌های اخیر را نمایش می‌دهد. به منظور ایجاد یک مقایسه عادلانه، تنها معیار دقت مورد بررسی قرار گرفته است. مقالات مورد بررسی در مراجع در این جدول همگی از یک مدل مبتنی بر شبکه‌های عصبی کانولوشنی با معماری‌های مختلف بر روی بانک داده FDDB استفاده کرده‌اند.

جدول (۵): مقایسه دقت شبکه عصبی کانولوشنی عمیق پیشنهادی با سایر

| کارهای انجام شده در سال‌های اخیر |            |              |         |
|----------------------------------|------------|--------------|---------|
| مرجع                             | سال انتشار | روش کار      | دقت (%) |
| [۱۷]                             | ۲۰۲۲       | MTCNN        | ۹۲/۰۶   |
| [۱۸]                             | ۲۰۲۱       | DCNN         | ۹۴/۰۸   |
| [۱۹]                             | ۲۰۲۱       | RCNN         | ۹۱/۰۰   |
| [۲۰]                             | ۲۰۱۶       | Cascaded CNN | ۹۱/۰۲   |
| [۲۱]                             | ۲۰۱۵       | DCNN         | ۸۵/۰۰   |
| روش پیشنهادی                     |            | DCNN         | ۹۶/۱۱   |

## ۵. جمع‌بندی

در این مطالعه برای تشخیص چهره بر روی داده‌ها ناقص از یک مدل عمیق مبتنی بر شبکه‌های عصبی کانولوشنی استفاده شد. در فرآیند پیش‌پردازش استخراج لیست ویژگی‌های کاربردی از اجزای چهره انجام شده و با اتکا بر روش‌های قطعه‌بندی و تبدیل تصادفی تصاویر به تصاویر ناقص، یک رویکرد بهینه اتخاذ شده است. در واقع تشخیص چهره مبتنی بر شبکه‌های عصبی کانولوشنی با تصاویر اصلاح شده در محله پیش‌پردازش آغاز شد تا ترسیم دقیق‌تر در تصاویر با وضوح بالا انجام شود. در سناریوی مربوط به آموزش شبکه عصبی کانولوشنی عمیق پیشنهادی از سه ترتیب تکرار مختلف استفاده شد. کار با یک

معماری عمیق برای داشتن عملکرد مینا برای شبکه عصبی کانولوشنی مطابق جزئیات ارائه شده در جدول (۱) آغاز شد. در سناریوی پیشنهادی به ترتیب تکرارهای ۱۰۰، ۱۵۰ و ۲۰۰ بر روی مدل اعمال شد. در جدول (۳) میزان هزینه به تعداد تکرار ۱۰۰ در سناریوی اول نشان داده شد. برای تکرارهای ۱۵۰ و ۲۰۰ نیز همین روند تکرار شده است. با توجه به نتایج به دست آمده، مدل پیشنهادی با بهره‌گیری از پنج لایه همگشت، دو لایه پولینگ یا ادغام و به دنبال آن لایه softmax منجر به تشخیص مناسب محل چهره شد. نتایج تصویری به دست آمده از مدل عمیق پیشنهادی برای دو کلاس ناقص و کامل، در جدول (۲) نشان داده شد. با توجه به مقادیر و نتایج به دست آمده برای معیارهای ارزیابی مطابق جدول (۴) مدل پیشنهادی نسبت به مدل‌های عمیق معرفی شده در بخش پیشینه بهتر و دقیق‌تر عمل کرده و از کارایی مناسب‌تری نسبت به این مدل‌ها برخوردار است (جدول (۵)). اگرچه عمل تشخیص از تصاویر ثابت چهره به درجه بالایی از اطمینان و کارایی رسیده، اما همچنان افزایش دقت تشخیص در تصاویر چهره ناقص نیز یک موضوع مهم و قابل بررسی می‌باشد. در این مطالعه، برای بهبود تشخیص چهره از یک شبکه عصبی کانولوشنی عمیق استفاده شده است. در مرحله اول از روش پیشنهادی پس از انتخاب و استخراج داده اولیه از پایگاه داده، پیش‌پردازش‌های لازم بر روی تصاویر انجام می‌شود. خروجی این مرحله به عنوان ورودی به الگوریتم بهینه‌سازی زنبور عسل داده شده تا برای انتخاب ویژگی‌های موثر و شناخت آنها در راستای بهینه‌سازی استفاده گردد. در نهایت از یک شبکه عصبی کانولوشنی عمیق در دو گام آموزش و آزمون برای تشخیص چهره استفاده شده است. اهداف روش پیشنهادی، کاهش پیچیدگی روند محاسباتی و افزایش دقت در فرآیند تشخیص چهره با استفاده از داده‌های ناقص بوده که در آن از نرم‌افزار متلب برای پیاده‌سازی استفاده شده است. نتایج پیاده‌سازی نشان‌دهنده عملکرد مناسب و افزایش دقت در فرآیند تشخیص می‌باشد. در روش ارائه شده به دلیل انتخاب بهینه ویژگی‌ها، میزان دقت تشخیص ۹۶/۱۱٪ به دست آمده که نسبت به سایر روش‌های ارائه شده در سال‌های اخیر بالاتر است. علاوه

<sup>1</sup> Overfitting

بر این، عمل تشخیص در روش پیشنهادی سریع بوده و به طور کلی قادر است الزامات تشخیص چهره را در زمان واقعی برآورده کند. برای بهبود هرچه بیشتر عملکرد روش پیشنهادی می توان از ترکیب سایر الگوریتم های فراابتکاری برای بهینه سازی و از ظرفیت های پردازنده گرافیکی (GPU)<sup>۱</sup> برای افزایش سرعت در روند آموزش مدل های عمیق استفاده کرد.

## ۶. نتیجه گیری و پیشنهادات

اکنون با توسعه علوم در حوزه هوش مصنوعی، تعامل بین ماشین و انسان بیشتر شده است. با توجه به اینکه تحلیل اطلاعات در رایانه ها، با سرعت و دقت بسیار بالا قابل بررسی است، بنابراین پژوهشگران جهت برقراری ارتباط موثر با دنیای واقعی، حوزه های مطالعاتی کثیری از جمله بینایی ماشین، یادگیری ماشین، رباتیک، پردازش گفتار و تشخیص چهره را به وجود آورده اند. پژوهشگران مرتبط با بینایی ماشین برای تحلیل مناسب اطلاعات مرتبط، نیاز به تشخیص اجسام در تصویر و ویدیو دارند. این تشخیص و طراحی با هدف های متفاوتی انجام شده است. تشخیص و شناسایی چهره یکی از اساسی ترین و پرچالش ترین مسائل در بینایی ماشین می باشد. تشخیص و شناسایی چهره به علت تفاوت در رنگ، زاویه، حالت، شدت روشنایی و اندازه نسبی می تواند باعث ایجاد چالش های بزرگی در این زمینه شود. از این رو روش ها و الگوریتم های قوی برای اتخاذ ویژگی های ساختاری از تصاویر، مورد نیاز است. در این حوزه از مسائل باید چهره ها در مقیاس ها و حالت های مختلف از تصاویر در زوایای متفاوت از افراد مورد بررسی قرار گیرد. در روش های متداول برای تشخیص و شناسایی چهره، حجم بالای تصویرها در مجموعه داده مورد جستجو، موجب افت کارایی در شناسایی شده و در نهایت باعث ضعف در کاربردهای امنیتی می شود. بنابراین با استفاده از روش هایی با کارایی مناسب نسبت به حجم بسیار زیاد داده های در دسترس، کاربردهای این حوزه قابل توسعه است. در این مطالعه، هدف اصلی بهبود تشخیص چهره بر روی تصاویر با داده های ناقص

توسط یک شبکه عصبی کانولوشنی عمیق است. در روش پیشنهادی از یک رویکرد ترکیبی برای محقق کردن هدف ذکر شده، استفاده شده است. به این صورت که ابتدا تصویر اولیه در قالب یک پایگاه داده خام تحلیل و بررسی شده و روش پیش پردازش داده با اعمال فیلتر و تبدیل هیستوگرام تصاویر اعمال گردیده است. در گام بعدی با برش یک قطعه از تصویر به صورت تصادفی و بهره مندی از روش تشخیص لبه اقدام به بهبود کیفیت تصویر ورودی شده است. در مرحله بعدی به کمک الگوریتم بهینه سازی زنبور عسل انتخاب ویژگی های موثر برای مطابقت و تشخیص صحیح با داده های پایگاه داده انجام شده است. در مرحله آخر به کمک شبکه عصبی کانولوشنی، انتخاب لایه های بهینه یادگیری و آزمون کننده انجام شده و تشخیص صورت گرفته از دید معیارهای دقت، صحت و پوشش برای سه سناریوی متفاوت با تعداد تکرارهای ۱۰۰، ۱۵۰ و ۲۰۰ مورد تحلیل واقع شده است. دقت نهایی با تکرار ۲۰۰ برای سناریوی سوم ۹۶/۱۱ درصد به دست آمد. نتایج به دست آمده بیانگر این موضوع است که روش ارائه شده بهبود مورد نظر را محقق کرده است. به طور خلاصه نتایج حاصل از این تحقیق شامل بهبود عملیات تشخیص چهره بر روی تصاویر با داده های ناقص توسط شبکه های عصبی با کمترین هزینه، استخراج و کنکاش ویژگی های اصلی برای بهبود در تشخیص چهره به کمک الگوریتم بهینه سازی زنبور عسل، استفاده از شبکه عصبی کانولوشنی عمیق جهت بهبود تشخیص چهره با داده های ناقص، بهبود ضریب اطمینان تشخیص چهره بر روی تصاویر با داده های ناقص در سیستم های امنیتی و افزایش کیفیت تصویر با پیش پردازش بهبود یافته است. از جمله محدودیت های مهم در این تحقیق پیچیدگی حل مساله بود که با اتکا بر چند روش مختلف پیش برده شد تا این پیچیدگی به حداقل برسد. از جمله فعالیت ها و کارهای آینده نیز می توان به موارد پیش رو شامل بررسی روش هایی که امکان تبدیل تشخیص چهره با داده های ناقص را به مدل دارای قابلیت اجرا صرف نظر از دخالت انسان و به صورت کاملاً خودکار را فراهم آورد، همچنین استخراج بهترین و موثرترین ویژگی ها برای تشخیص چهره با استفاده از بهینه-

<sup>۱</sup> Graphics Processing Unit

تعارض منافع: نویسندگان اعلام می‌کنند که هیچ تعارض منافعی ندارند.

سازی فراابتکاری و ترکیب آن با سایر مدل‌های شبکه‌های عصبی عمیق و استفاده از الگوریتم‌های یادگیر برای تشخیص چهره در سیستم‌های مبتنی بر ابر با رویکرد کاهش زمان پاسخ اشاره کرد.

## مراجع

- [1] F. Hosseini, A. Shah Bahrami, M. Hajjarian, and N. Khalili Dizji, "Rotation-independent face detection using segment-based modelm" *Comput. Sci. J.*, vol. 1, no. 1, pp. 1-6, 2016 [In Persian].
- [2] H.O. Ikromovich and B.B. Mamatkulovich, "Facial Recognition Using Transfer Learning in the Deep CNN," *Int. Sci. Res. J.*, vol. 4, no. 3, pp. 502-507, 2023, doi: 10.17605/OSF.IO/NRMK2.
- [3] S. Li and H.J. Lee, "Effective Attention-Based Feature Decomposition for Cross-Age Face Recognition," *Appl. Sci.*, vol. 12, no. 10, p. 4816, 2022, doi: 10.3390/app12104816.
- [4] K. Chumachenko, A. Iosifidis, and M. Gabbouj, "Self-attention fusion for audiovisual emotion recognition with incomplete data," in *26th Int. Conf. Pattern Recognit. (ICPR)*, Montreal, QC, Canada, 2022, pp. 2822-2828, doi: 10.1109/ICPR56361.2022.9956592.
- [5] G. Rajeswari and P.I. Rani, "Face occlusion removal for face recognition using the related face by structural similarity index measure and principal component analysis," *J. Intell. Fuzzy Syst.*, vol. 42, no. 6, pp. 5335-5350, 2022, doi: 10.3233/JIFS-211890.
- [6] H. Du, H. Shi, D. Zeng, X.-P. Zhang, and T. Mei, "The elements of end-to-end deep face recognition: A survey of recent advances," *ACM Comput. Surv.*, vol. 54, no. 10s, pp. 1-42, 2022, doi: 10.1145/3507902.
- [7] G. Jeevan, G.C. Zacharias, M.S. Nair, and J. Rajan, "An empirical study of the impact of masks on face recognition," *Pattern Recognit.*, vol. 122, p. 108308, 2022, doi: 10.1016/j.patcog.2021.108308.
- [8] G. Revathy, K.B. Raj, A. Kumar, S. Adibatti, P. Dahiya, and T.M. Latha, "Investigation of E-voting system using face recognition using convolutional neural network (CNN)," *Theor. Comput. Sci.*, vol. 925, pp. 61-67, 2022, doi: 10.1016/j.tcs.2022.05.005.
- [9] M. Andrejevic and N. Selwyn, "Facial recognition technology in schools: critical questions and concerns," *Learn. Media Technol.*, vol. 45, no. 2, pp. 115-128, 2020, doi: 10.1080/17439884.2020.1686014.
- [10] S. Ren, "Computer vision for facial analysis using human-computer interaction models," *J. Interconnect. Networks*, vol. 22, pp. 1-19, 2022, doi: 10.1142/S0219265921440059.
- [11] Q. Su, N. Kondo, M. Li, H. Sun, D.F. Al Riza, and H. Habaragamuwa, "Potato quality grading based on machine vision and 3D shape analysis," *Comput. Electron. Agric.*, vol. 152, pp. 261-268, 2018, doi: 10.1016/j.compag.2018.07.012.
- [12] H.N. Vu, H.M. Nguyen, and C. Pham, "Masked face recognition with convolutional neural networks and local binary patterns," *Appl. Intell.*, vol. 52, no. 5, pp. 5497-5512, 2022, doi: 10.1007/s10489-021-02728-1.
- [13] Z. Sharifi Mehrjard, H. Momeni, and H. Adabi Ardekani, "A review of machine learning algorithms to diagnose autism using the EEG signal," *Soft Comput. J.*, vol. 13, no. 1, pp. 2-19, 2024, doi: 10.22052/scj.2023.248522.1110 [In Persian].
- [14] V. Jain and E. Learned-Miller, "Fddb: A benchmark for face detection in unconstrained settings," Jan. 2010, [Online]. Available: <http://vis-www.cs.umass.edu/fddb/fddb.pdf>.
- [15] V. Esmacili and M. Mohassel Fegghi, "Real-time Authentication for Electronic Service Applicants using a Method Based on Two-Stream 3D Deep Learning," *Soft Comput. J.*, vol. 11, no. 2, pp. 38-49, 2023, doi: 10.22052/scj.2023.246701.1086 [In Persian].
- [16] R. Behmanesh and N. Majma, "Nephron-2 Meta-Heuristic Algorithm (NOA-2), to Solve Optimization Problems," *Soft Comput. J.*, vol. 11, no. 2, pp. 62-71, 2023, doi: 10.22052/scj.2023.248427.1104 [In Persian].
- [17] M.T. Islam, T. Ahmed, A.B.M.R. Rashid, T. Islam, M.S. Rahman, and M.T. Habib, "Convolutional Neural Network Based Partial Face Detection," in *IEEE 7th Int. Conf. Conver. Technol. (I2CT)*, Pune, India, 2022, pp. 1-6, doi: 10.1109/I2CT54291.2022.9825259.
- [18] D. Garg, P. Jain, K. Kotecha, P. Goel, and V. Varadarajan, "An efficient multi-scale anchor box approach to detect partial faces from a video sequence," *Big Data Cogn. Comput.*, vol. 6, no. 1, p. 9, 2022, doi: 10.3390/bdcc6010009.
- [19] H. Yan, X. Wang, Y. Liu, Y. Zhang, and H. Li, "A new face detection method based on Faster RCNN," *J. Phys. Conf. Ser.*, vol. 1754, p. 012209, 2021, doi: 10.1088/1742-6596/1754/1/012209.
- [20] H. Qin, J. Yan, X. Li, and X. Hu, "Joint training of cascaded CNN for face detection," in *IEEE Conf. Comput. Vision Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 3456-3465, doi: 10.1109/CVPR.2016.376.
- [21] S.S. Farfade, M.J. Saberian, and L.-J. Li, "Multi-view face detection using deep convolutional neural networks," in *Proc. 5th ACM on Int. Conf. Multimedia Retr.*, Shanghai, China, 2015, pp. 643-650, doi: 10.1145/2671188.2749408.