

# تشخیص چهره با داده‌های ناقص توسط شبکه عصبی کانولوشنی عمیق (DCNN)

فرناز حسینی<sup>\*</sup>، استادیار، الهه طیب زاده لمر<sup>۲</sup>، کارشناس ارشد، سید مهدی میرکازمی نیارق<sup>۳</sup>، استادیار

<sup>۱</sup> گروه مهندسی کامپیوتر، دانشگاه فنی و حرفه‌ای، تهران، ایران. [f-hoseini@tvu.ac.ir](mailto:f-hoseini@tvu.ac.ir)

<sup>۲</sup> کارشناس ارشد موسسه آموزش عالی شهریار، گروه مهندسی کامپیوتر، آستارا، ایران. [elatabib2@shahriar.ac.ir](mailto:elatabib2@shahriar.ac.ir)

<sup>۳</sup> استادیار گروه برق، واحد آستارا، دانشگاه آزاد اسلامی، آستارا، ایران. [mirkazemi.seyedmehdi@iau.ac.ir](mailto:mirkazemi.seyedmehdi@iau.ac.ir)

چکیده: چهره انسان یک شیء 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 فرایند تشخیص چهره را با مشکلات مختلفی روبرو سازند. ارائه روشی که بتواند این مشکلات را تا حد قابل قبولی رفع کند، می‌تواند گستره‌ی کاربرد تکنیک‌های 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، پیش پردازش 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 بدست 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.

واژه‌های کلیدی: تشخیص چهره، شبکه عصبی کانولوشنی، الگوریتم بهینه‌سازی زنبور عسل، داده‌های ناقص، شبکه عصبی کانولوشنی عمیق.

# Face Recognition with Incomplete Data by Deep Convolutional Neural Network (DCNN)

Farnaz Hoseini <sup>1\*</sup>, Assistant Professor, Elahe Tabibzade Lamar <sup>2</sup>, MA, Seyed Mehdi Mirkazemi Niarag <sup>3</sup>,  
Assistant Professor

<sup>1\*</sup> Department of Computer Engineering, Technical and Vocational University (TVU), Tehran, Iran.  
[f-hoseini@tvu.ac.ir](mailto:f-hoseini@tvu.ac.ir)

<sup>2</sup> Master of Shahriar Institute of Higher Education, Department of Computer Engineering, Astara, Iran.  
[elatabib2@shahriar.ac.ir](mailto:elatabib2@shahriar.ac.ir)

<sup>3</sup> Assistant Professor of Electrical Department, Astara Branch, Islamic Azad University, Astara, Iran.  
[mirkazemi.seyedmehdi@iau.ac.ir](mailto:mirkazemi.seyedmehdi@iau.ac.ir)

**Abstract:** The human face is a dynamic entity influenced by various factors that give rise to different facial expressions. Face recognition algorithms encounter challenges such as inherent and random factors causing facial appearance variations, incomplete data in the database, database size, differences in image dimensions, and changes in facial expressions. Addressing these challenges can expand the application range of facial recognition techniques. In this study, we propose a method that utilizes a deep convolutional neural network to enhance face recognition in the presence of incomplete data. The proposed method consists of several distinct steps. Firstly, primary data is selected and extracted from the database, followed by preprocessing the information through filtering, histogram transformation, and edge detection. Subsequently, crucial facial landmarks are extracted for each image. The output of this step serves as input to the bee optimization algorithm, which facilitates the selection of relevant features and optimizes them for recognition. Finally, a deep convolutional neural network is employed for face recognition, encompassing training and testing stages. We conducted simulations in the MATLAB environment to evaluate the proposed method and assess using accuracy, correctness, and criteria coverage criteria. The results demonstrated an accuracy of 96.11%, indicating improved face recognition compared to recent works and cost reductions in the overall recognition process.

**Keywords:** Face Recognition, Convolutional Neural Network, Honey-bee Optimization Algorithm, Incomplete Data, Deep Convolutional Neural Network.

\* Farnaz Hoseini, [f-hoseini@tvu.ac.ir](mailto:f-hoseini@tvu.ac.ir)

## ۱. مقدمه

تشخیص چهره تجزیه و تحلیل الگوهایی است صورت یک فرد 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 معمولاً برای اسکن چهره توسط فاصله نقشه‌برداری میان ویژگی‌های مهم چهره تصویربرداری‌های حرارتی و آنالیز بافت‌های سطحی در پوست استفاده می‌شوند [۱]. سیستم‌های 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 کامپیوتر 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 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 از جلوی صورت در دسترس نباشد. تصاویر چهره ناقص که اغلب از دوربین‌های مدار بسته می‌آیند، شامل این موارد هستند [۴]. از این رو، مسئله 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 تأثیر 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، چشم-ها، 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 کامپیوترها 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 کامپیوتر و انسان بسیار مفید است؛ به عنوان مثال برادری تعیین هویت افراد [۷]. تشخیص چهره در حوزه بینایی ماشین نیز در کاربردهای سیستم‌های نظارتی امنیتی و سیستم‌های مانیتورینگ امنیتی و موارد دیگر استفاده می‌شود [۸]. در دنیای تکنولوژی رویکرد تشخیص امر مهمی بوده و کاربردهای بسیاری دارد؛ به عنوان مثال در کاربردهای پروفایلینگ تجاری، اهداف مختلف نظارتی مانند نظارت بر تصاویر و ویدئوها [۹]، کاربردهای مانیتورینگ و سیستم‌های تعاملی بین انسان و کامپیوتر [۱۰]. در این مطالعه، آزمایشی جدید برای تست و 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.

<sup>1</sup> Machine Learning

<sup>2</sup> Convolutional Neural Network (CNN)

بر شبکه‌های عصبی کانولوشنی با بهینه‌سازی توسط الگوریتم بهینه‌سازی زنبورعسل مورد بحث قرار گرفته است. در بخش چهارم نتایج بدست آمده از روش پیشنهادی ارائه شده است. نتیجه گیری و پیشنهادات آتی نیز در بخش پنجم ارائه شده 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 [۱۶] سال ۲۰۲۲ یک شبکه عصبی کانولوشنی چند وظیفه‌ای (MTCNN)<sup>۸</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 پیش پردازش‌های لازم بر روی داده‌ها انجام می‌شوند. به منظور افزایش تعداد تصاویر در مجموعه آموزش از تکنیک افزایش داده<sup>۹</sup> استفاده شده است. با آماده‌سازی تصاویر رمزگذاری<sup>۱۰</sup> شده، از یک شبکه عصبی کانولوشنی چند وظیفه‌ای (MTCNN) برای آموزش و یافتن نواحی صورت استفاده می‌شود. اعتبارسنجی و ارزیابی نتایج توسط معیارهای دقت (Accuracy)، صحت (Precision) و پوشش (Recall) انجام شده است. 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.

بنابراین استفاده از این شبکه‌ها در زمینه‌ی تشخیص چهره که جز الگوهای مهم یادگیری ماشین است، ضروری به نظر می‌رسد [۱۳]. پیشینه‌ی پژوهش‌ها نیز نشان می‌دهد که در این شبکه‌ها به 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 (Fddb)<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 قطع‌بندی، واحدهای فعالیت نیز استخراج می‌شوند. برای افزایش دقت تشخیص سیستم، با استفاده از الگوریتم بهینه‌سازی زنبورعسل<sup>۴</sup> [۲۱] بهترین ویژگی‌ها برای شناسایی چهره انتخاب می‌شوند. در نهایت از یک شبکه عصبی کانولوشنی عمیق (DCNN)<sup>۵</sup> در دو مرحله آموزش<sup>۶</sup> و آزمون<sup>۷</sup> (تست) 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 تکنیک‌ها و 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

<sup>3</sup> <http://vis-www.cs.umass.edu/fddb>

<sup>4</sup> Honey-bee Optimization Algorithm

<sup>5</sup> Deep Convolutional Neural Network

<sup>6</sup> Train

<sup>7</sup> Test

<sup>8</sup> Multi Task Convolutional Neural Network

<sup>9</sup> Data Augmentation

<sup>10</sup> Encoding Images

در [۱۹] سال ۲۰۱۶ 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>۱۴</sup> توسط شبکه 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 ساده و یک شبکه پیشنهادی منطقه‌ای (RPN) پیچیده‌تر را نشان می‌دهد. آزمایش‌ها روی معیارهای تشخیص چهره (دقت، صحت و پوشش) مزایای تمرین مشترک را تأیید می‌کنند. مدل پیشنهادی قادر است چهره را در تصویر با داده‌های ناقص توسط شبکه عصبی کانولوشنی آبشاری با دقت ۹۱/۰۲٪ درصد بر روی مجموعه داده FDDDB تشخیص دهد.

در [۲۰] سال ۲۰۱۵ از یک آشکارساز چهره متراکم مبتنی بر شبکه عصبی کانولوشنی عمیق (DCNN) 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 اشیاء 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

در [۱۷] سال ۲۰۲۱ یک روش تشخیص چهره مبتنی بر شبکه عصبی کانولوشنی عمیق (DCNN) با استفاده از استراتژی بخش جعبه‌های لنگر یا انکر باکس‌ها<sup>۱۱</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 FDDDB 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٪ درصد بر روی مجموعه داده FDDDB تشخیص دهد.

در [۱۸] سال ۲۰۲۱ یک شبکه عصبی کانولوشنی مبتنی بر منطقه (RCNN)<sup>۱۲</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>۱۳</sup> شناسایی می‌شوند. نتایج آزمایش‌ها و مقایسات نشان می‌دهد که روش تشخیص چهره در این پژوهش بر روی مجموعه داده FDDDB 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 ۹۱/۰۰٪ درصد بر روی مجموعه داده FDDDB تشخیص دهد.

<sup>11</sup> Anchor Boxes

<sup>12</sup> Region-based Convolutional Neural Network

<sup>13</sup> Region Proposal Network (RPN)

شبکه‌ها در مورد تشخیص چهره که یکی از الگوهای مهم یادگیری ماشین است، به نظر ضروری می‌رسد. هم چنین ادبیات پژوهش‌های بررسی شده در بخش ۲ نشان می‌دهد که در حوزه‌ی دسته‌بندی به المان 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 داده‌ی استاندارد انتخاب شده و بعد از فراخوانی تصاویر، جهت بهبود کیفیت، پیش پردازش‌هایی بر روی تصاویر صورت می‌گیرد. در این بخش داده‌های معیوب نیز مورد ترمیم و بازسازی قرار می‌گیرند و در صورت عدم قابلیت استفاده حذف می‌شوند. همچنین در مرحله نخست پس از پیش پردازش، لیست خصوصیات کاربردی اجزای چهره تهیه شده و با استفاده از قطعه بندی، واحدهای فعالیت مشخص شده و جهت افزایش دقت تشخیص سیستم، از الگوریتم بهینه‌سازی زنبور عسل برای شناسایی چهره، بهترین ویژگی مؤثر انتخاب می‌شود. در نهایت از یک شبکه عصبی کانولوشنی عمیق برای فرآیند یادگیری و تشخیص چهره استفاده می‌شود.

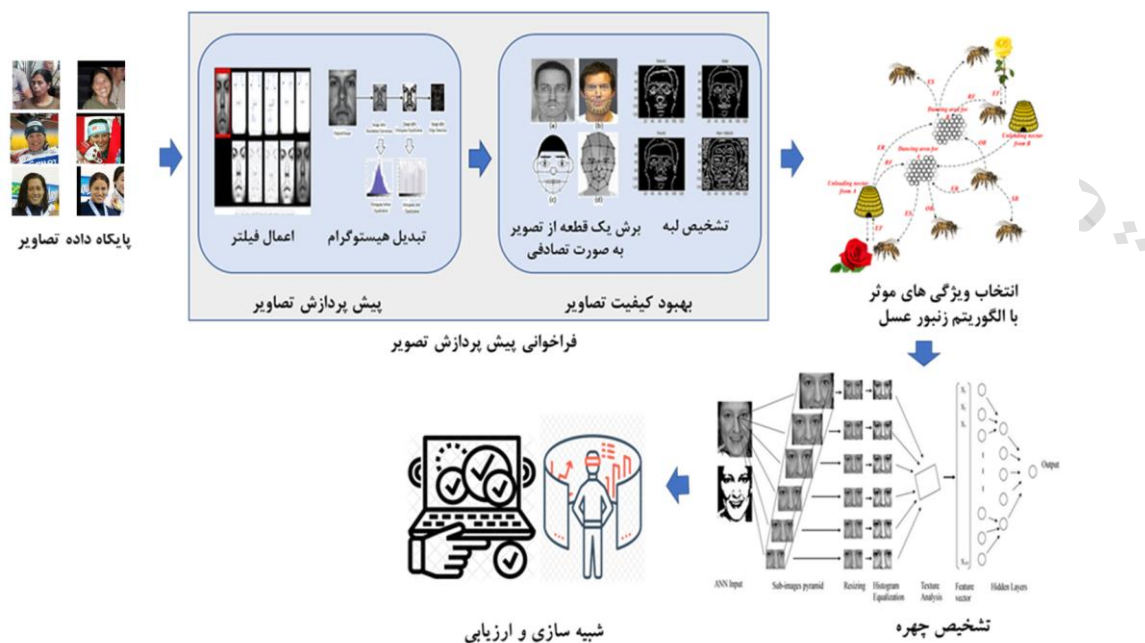
مطابق با هدف اصلی تحقیق که بهبود تشخیص چهره به کمک داده‌های ناقص می‌باشد، روش پیشنهادی از چند گام مختلف تشکیل شده است. به این طریق که پس از انتخاب و استخراج داده‌ی اولیه به کمک بانک داده‌ی منتخب اقدام به پیش پردازش تصاویر می‌شود. در این مرحله اعمال فیلتر، تبدیل هیستوگرام، برش قطعات تصویر و تشخیص لبه انجام می‌گردد. در مرحله‌ی بعد استخراج نقاط مهم در تصویر انجام می‌شود. خروجی گام قبلی به عنوان ورودی به الگوریتم زنبور عسل داده می‌شود تا انتخاب ویژگی‌های مؤثر و شناخت آن‌ها در راستای بهینه‌سازی فرایند تشخیص انجام گیرد. در نهایت از شبکه عصبی کانولوشنی عمیق برای تشخیص چهره استفاده می‌شود. این

نمونه‌گیری بهتر و تکنیک‌های پیچیده‌تر افزایش داده‌ها بهبود داد. ارزیابی‌ها روی مجموعه داده معیار تشخیص چهره محبوب Fddb نشان می‌دهد که آشکارساز چهره متراکم پیشنهادی توانسته دقت  $0.85/0.0$  درصد بر روی این مجموعه داده را بدست آورد.

### ۳. روش پیشنهادی

امروزه تشخیص چهره مبتنی بر کامپیوتر یک مکانیسم 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 روش تشخیص چهره از اطلاعات جزئی صورت استفاده شده است. این مسئله را با استفاده از آزمایش‌های جدید برای تست عملکرد یادگیری ماشین، تصاویر جزئی چهره و دستکاری‌های دیگر در تصاویر چهره مانند چرخش و زوم، که به عنوان سرخ‌های اصلی آموزش و تشخیص استفاده می‌شوند، مورد بررسی قرار می‌دهیم. با توجه به اینکه روش شبکه‌های عصبی کانولوشنی در دسته بندی و تشخیص اشیاء خوب عمل می‌کند، واضح است که این شبکه‌ها ویژگی‌های سطح بالا و مؤثری را می‌تواند از تصاویر بانک داده‌ی بزرگ استخراج کند. بنابراین، بهره‌گیری از این

تشخیص در دو مرحله آموزش و آزمون انجام می‌پذیرد. برای درک بهتر رویکرد پیشنهادی خلاصه‌ای از روند کلی روش



شکل (۱): خلاصه‌ای از روند کلی روش پیشنهادی

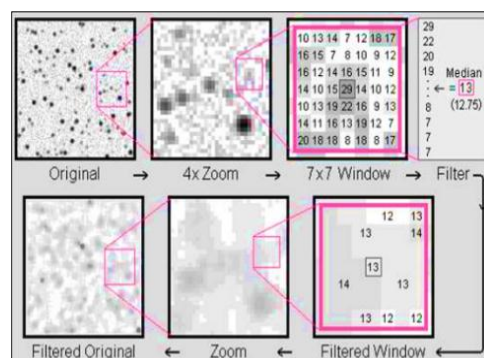
می‌باشد. مقدار میانه در مقایسه با مقدار میانگین، شاخص مرکزی قویتری است و به راحتی تحت تاثیر مقادیر پراکنده در میان پیکسل‌های موجود در محدوده همسایگی قرار نمی‌گیرد. در این مقاله نخست به مرتب سازی مقدار پیکسل‌های موجود در همسایگی که به صورت صعودی انتخاب شده‌اند، پرداخته می‌شود و پس از معین نمودن میانه 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.

فیلتر میانه<sup>۱۵</sup> را تحت عنوان روشی پیشرفته برای پیش‌پردازش می‌شناسند. این فیلتر قادر به حفظ جزئیات یک تصویر با دقت بالایی می‌پردازد. در اطراف هر پیکسل در یک تصویر دیجیتال، محدوده‌ای مشخص می‌شود. مقادیر کل پیکسل‌های آن محدوده بر اساس یک فلوجارت مشخص آنالیز می‌شوند و سپس مقادیر اصلی هر پیکسل با مقدار حاصل از آنالیز 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 معمولاً برای کاهش نویز در تصاویر استفاده می‌گردد و تا حدی عملکرد مشابه با فیلتر میانگین<sup>۱۶</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

<sup>15</sup> Median

<sup>16</sup> Mean

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 تکنیک تشخیص لبه کنی در نرم افزار 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. هم چنین زنبورهای ناظر ما بین 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 فلوچارت 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 وسیله‌ی 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 آغاز به استخراج غذا از منبع پیدا شده می‌کنند. زنبور کارگر حامل شهد به کندو بازگشته و بار شهد خود را تخلیه می‌کند. پس از تخلیه شهد، قادر است مستقیماً به منبع کشف شده‌ی خود برگردد یا داده‌هایی که درباره منبع غذایی خود دارد را به وسیله‌ی اجرای یک رقص گردون در ناحیه رقص به اشتراک دیگران بگذارد. حال اگر یک منبع غذایی خالی شود، زنبورهای کارگر تبدیل به دیده بان شده و به طور تصادفی مشغول به جست و جوی منابع



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 مشخصه‌ی 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 مخصوصاً 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 است. سه نوع مختلف از ناپیوستگی در سطح خاکستری وجود دارد که عبارتند از نقطه، خط و لبه. ماسک‌های فضایی می‌توانند برای تشخیص هر سه نوع ناپیوستگی در یک تصویر به کار روند. تکنیک تشخیص لبه کنی<sup>۱۷</sup> یکی از



این مرحله بر روی مقادیر به دست آمده، تابع فعالساز غیرخطی اعمال خواهد شد. این مطالعه از تابع یک‌سوساز  $ReLU^{21}$  به عنوان تابع غیرخطی استفاده می‌کند. 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.

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

نهایتاً 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 پارامترهای 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 آن ادغام حداکثری یا  $\max$  pooling می‌باشد. این ادغام بیشترین مقدار پیکسل بین پیکسل‌های منتخب را به عنوان داده‌ی 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 اپاک<sup>۲۲</sup> یا دور بوده و بعد از این تعداد

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 (مقدار تناسب) می‌باشد. هم چنین 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 اینطور 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>۱۸</sup>، لایه پولینگ<sup>۱۹</sup> یا ادغام و لایه تماماً متصل<sup>۲۰</sup>. در روش پیشنهادی تصاویر بدست 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 کانال می‌شوند. هر لایه کانولوشنی از سه زیر لایه تشکیل شده است. زیرلایه اول تصویر پیش‌پردازش شده با کرنل‌های کانولوشنی از پیش تعیین شده کانال 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 زیر لایه غیرخطی منتقل می‌شود. در

<sup>21</sup> Rectified Linear Unit (ReLU)

<sup>22</sup> Epoch

<sup>18</sup> Convolution Layer

<sup>19</sup> Pooling Layer

<sup>20</sup> Fully Connected Layer

این جدول اندازه فیلتر، گام<sup>۲۳</sup>، پد<sup>۲۴</sup> و حذف<sup>۲۵</sup> در هر لایه و همچنین حجم و تعداد مؤلفه های پد در هر لایه به نمایش گذاشته می شود.

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

شماره لایه	فیلد گام	پد	حذف	حجم	پارامترها
۰	INPUT	-	-	-	-
۱	CONV -> RELU	۵*۵	۲	۱	۸۳۲
۲	CONV -> RELU	۳*۳	۱	۱	۹,۲۴۸
۳	MP	۳*۳	۲	۱	۱۸,۴۹۶
۴	CONV -> RELU	۳*۳	۱	۱	۳۶,۹۲۸
۵	CONV -> RELU	۳*۳	۱	۲	۵۵,۳۹۲
۶	MP	۳*۳	۲	۲	۸۳,۰۴۰
۷	FC-> RELU	۵*۵	۱	۶	۲,۴۰۱
۸	Soft max	-	-	-	-

\* Conv: Convolutional Layer

\* MP: Max Pooling Layer

\* ReLU (Rectified Linear Unit): Activation Function

\* FC: Fully Connected Layer

#### ۴. نتایج تجربی

در این بخش به بیان نتایج بدست آمده از روش پیشنهادی، معرفی بانک داده، محیط شبیه سازی و نمایش نمودارها و تصاویر مربوط به خروجی مدل پیشنهادی پرداخته شده است.

تکرار، آموزش شبکه به حالت بهینه رسیده است. در انتهای شبکه به منظور نگاشت نهایی ویژگی های کسب شده، از لایه ای کاملاً متصل استفاده می شود. یک لایه ای که کاملاً متصل است، دقیقاً مانند یک لایه کانولوشنی می باشد با این اختلاف که تعاملات پراکنده در اینجا انجام نمی پذیرد و درست همانند شبکه های عصبی مرسوم میان این لایه و لایه ی قبل از آن یک اتصال کامل صورت می گیرد. خروجی آخرین لایه یک بردار یک بعدی بوده و تعداد مؤلفه های این بردار با تعداد کلاس های دسته بندی برابر می باشد. در مدل پیشنهادی تعداد لایه ها برابر با تعداد ویژگی های تصویر چهره و کلاس خروجی برابر با تصویر شناسایی شده می باشد. در حقیقت این لایه عمل دسته بندی را انجام می دهد. در کل ساختارهای کانولوشنی آخرین لایه کاملاً متصل به یک لایه softmax متصل می شود. لایه softmax در حقیقت عمل طبقه بندی در شبکه های کانولوشنی را انجام می دهد. این لایه مشتمل بر تعدادی نرون مساوی با تعداد کلاس های مسئله دسته بندی بوده و به منظور نگاشت نهایی ویژگی و اعمال دسته بندی مورد استفاده قرار می گیرد. در این مقاله جهت تنظیم پارامترهای شبکه و یا همان آموزش شبکه، از نتایج خروجی برای محاسبه میزان خطای شبکه استفاده می شود. از این رو خروجی شبکه را توسط یک تابع خطا با پاسخ صحیح مقایسه کرده و به این شکل میزان خطا را محاسبه می کنیم. نحوه محاسبه خطای تجربی توسط عبارت (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 پیش بینی اشتباه  $\hat{Z}$  به جای  $Z$  معین می نماید. در مرحله بعدی با توجه به مقدار خطای محاسبه شده مرحله پس انتشار شروع می شود. پس از تکرار شدن تعداد مناسبی از این مرحله ها، آموزش شبکه پایان می یابد. جدول ۱ جزئیات معماری شبکه عصبی کانولوشنی عمیق پیشنهادی را نمایش می دهد. در

<sup>23</sup> Stride

<sup>24</sup> Zero padding

<sup>25</sup> Drop-out

## ۱.۴. بانک داده

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 داده‌ی مورد استفاده و تغییرات تصویر را نمایش می‌دهد.



شکل (۳): نمونه‌هایی از تصاویر بانک داده Fddb

## ۲.۴. محیط شبیه‌سازی

روش پیشنهادی توسط نرم‌افزاری متلب نسخه Matlab R2016a در محیط ویندوز<sup>۲۷</sup> و توسط یک رایانه استاندارد با پردازنده Intel(R) Core(TM) i7-5500U CPU@2.40GHz و رم ۸ گیگابایت پیاده‌سازی شد.

## ۳.۴. معیارهای ارزیابی

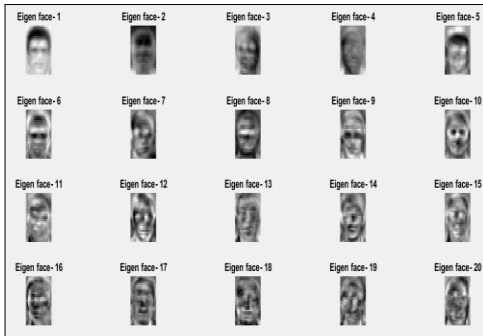
در این مطالعه از معیارهای دقت (Accuracy)، صحت (Precision) و پوشش (Recall) و برای ارزیابی عملکرد مدل عمیق پیشنهادی استفاده شده است. معیار دقت به معنای تعداد نمونه‌های درست طبقه‌بندی شده نسبت به کل نمونه‌ها می‌باشد. معیار صحت مشخص می‌کند که الگوریتم چند درصد «بله»‌هایش درست است؛ یعنی تمرکز اصلی آن، بر روی صحت تشخیص‌های «بله» توسط فلوچارت است. هم‌چنین تأکید اصلی معیار پوشش بر عکس معیار صحت روی داده‌هایی است که

مجموعه داده‌ها و معیار تشخیص چهره Fddb<sup>۲۶</sup> مجموعه‌ای از چهره‌های برجسب‌گذاری شده از تصاویر چهره مختلف در یک بانک داده وسیع می‌باشد. این بانک داده حاوی حاشیه نویسی 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، اندازه‌ی چهره و پوشش می‌باشد. معمولاً تصویرهای موجود در این بانک داده، به سه دسته‌ی 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 به همین جهت تصاویر مورد استفاده در این پژوهش، در دسته‌ی 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 یا تست 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 فرایند آموزش اولیه دخالتی ندارند. نکته‌ای در شبیه‌سازی انجام شده باید به آن اشاره کرد استخراج

<sup>27</sup> Windows

<sup>26</sup> Face Detection Data Set and Benchmark

برش یک قطعه از تصویر به صورت تصادفی با تطابق و استخراج چهره‌های مشابه با ویژگی‌های موردنظر بررسی می‌شود. به کمک شبکه عصبی کانولوشنی عمیق پیشنهادی آموزش و آزمون، در دو مرحله‌ی مجزا، فرآیند تشخیص را تکمیل می‌کند. 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 تست 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. 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

واقعاً «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 نمونه‌ی مذکور 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.  $TP^{28}$  (مثبت حقیقی): تعداد پیش‌بینی‌های صحیح مربوط به همان کلاس را نشان می‌دهد،  $TN^{29}$  (منفی حقیقی): تعداد درست پیش‌بینی شده به غیر از کلاس فعلی می‌باشد،  $FP^{30}$  (مثبت کاذب): نیز تعداد پیش‌بینی نادرست مربوط به کلاس‌های دیگر را نشان می‌دهد و  $FN^{31}$  (منفی کاذب): تعداد پیش‌بینی نادرست مربوط به همان کلاس را نشان می‌دهد. برای ارزیابی فاکتورهای مورد بررسی توسط مدل‌های پیشنهادی، متغیرهای وابسته برای معیار دقت توسط عبارت (۳)، صحت توسط عبارت (۴) و پوشش توسط عبارت (۵) تعریف و محاسبه می‌شوند.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

#### ۴.۴. خروجی‌های روش پیشنهادی

مدل ارائه شده به کمک ابزار متلب توسعه داده شده است. 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 پیش‌پردازش و اعمال فیلتر بر روی چهره مورد نظر انجام شده است. همچنین

<sup>28</sup> True Positive

<sup>29</sup> True Negative

<sup>30</sup> False Positive

<sup>31</sup> False Negative

۳ نشان داده شده است. برای تکرارهای ۱۵۰ و ۲۰۰ نیز همین روند تکرار شده است.

جدول (۲). نمونه هایی از نتایج تصویری بدست آمده توسط شبکه عصبی

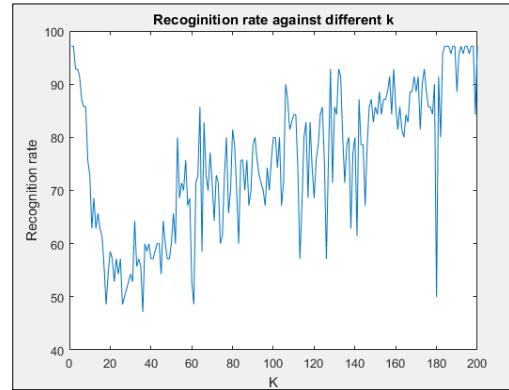
کانولوشنی عمیق (DCNN) پیشنهادی

نوع داده ها (کامل/ناقص)	خروجی تشخیص چهره توسط شبکه عصبی کانولوشنی عمیق	نمونه تصویر
کامل		
ناقص		
کامل		
ناقص		
کامل		
ناقص		
کامل		
ناقص		

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

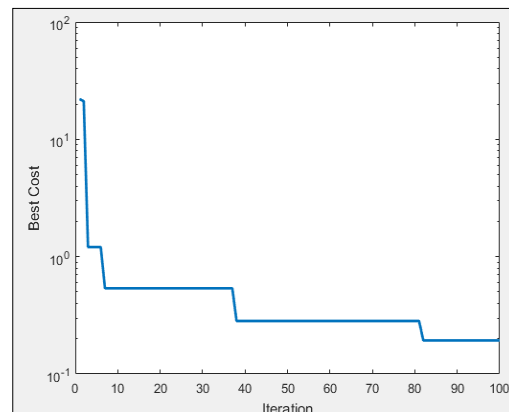
تکرار (Iteration)	هزینه (Cost)	تکرار (Iteration)	هزینه (Cost)	تکرار (Iteration)	هزینه (Cost)
1	0.19313	34	0.28303	67	0.53704

پنهان در شبکه عصبی کانولوشنی عمیق می باشند. همانطور که مشخص است با افزایش تعداد  $k$  نرخ شناسایی نیز روند صعودی داشته است. می توان گفت وضعیت نوسانی رو به بهبود برای نرخ تشخیص چهره تشکیل شده است.



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

در شکل ۷ نشان داده شده است که با افزایش تعداد تکرار در انتخاب بهینه ویژگی ها جهت تشخیص توسط الگوریتم بهینه سازی زنبور عسل هزینه تمام شده یک روند نزولی در پیش گرفته است.



شکل (۷): نمایش نسبت تعداد تکرار به هزینه

جدول ۲ نمونه هایی از چند تصویر با داده ها کامل و ناقص و نتایج حاصل از تشخیص چهره توسط شبکه عصبی کانولوشنی عمیق پیشنهادی را نمایش می دهد. در ادامه ی این بخش به دلیل تشابه خروجی های مورد نظر در تکرارهای دیگر، نمودارهای خروجی نشان داده نشده است. اما در یک قالب جدولی کلی میزان هزینه به تعداد تکرار ۱۰۰ در سناریوی اول توسط جدول

خروجی بدست 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 بخاطر برش تصادفی تصویر و شناسایی و استخراج شاخص‌های مهم در تصویر اولیه است، اما کلیت در یک بازه عددی و رویکرد مشخص شده حاصل می‌شود و بعبارتی در هر اجرا عیناً نتایج مندرج در این جدول دیده نخواهد شد و داده‌ها و خروجی‌ها با کمترین تلورانس نسبت به همدیگر حاصل می‌شوند.

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

تکرار (Iteration)	هزینه (Cost)	دقت (%) (Accuracy)	صحت (%) (Precision)	پوشش (%) (Recall)
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

مهم‌ترین یافته‌ی پژوهش حاضر بدین شرح است که با عمیق کردن شبکه عصبی کانولوشنی و کنترل بیش‌برازش<sup>۳۲</sup> می‌توان به دقت بالاتری رسید. نتایج بدست 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 مقایسه‌ی عادلانه، تنها معیار دقت (Accuracy) مورد بررسی قرار گرفت. مقالات مورد بررسی در مراجع مورد مقایسه در این جدول

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 Model is: 90.0899				100	0.53704

با توجه به نتایج بدست آمده، میانگین کل از معیارهای ارزیابی دقت (Accuracy)، صحت (Precision) و پوشش (Recall) و برای ارزیابی عملکرد مدل عمیق پیشنهادی توسط سه سناریوی بررسی شده و خلاصه‌ی نتایج در جدول ۴ نشان داده شده است.

<sup>32</sup> Overfitting

همگی از یک مدل مبتنی بر شبکه‌های عصبی کانولوشنی با معماری‌های مختلف بر روی بانک داده FDDDB استفاده کرده‌اند.

جدول (۵): مقایسه دقت شبکه عصبی کانولوشنی عمیق (DCNN)

پیشنهادی با سایر کارهای انجام شده در سال‌های اخیر

شماره مرجع مقاله	سال چاپ	بانک داده	روش کار	دقت (%)
[۱۶]	۲۰۲۲	FDDDB	MTCNN	۹۲,۰۶
[۱۷]	۲۰۲۱	FDDDB	DCNN	۹۴,۰۸
[۱۸]	۲۰۲۱	FDDDB	RCNN	۹۱,۰۰
[۱۹]	۲۰۱۶	FDDDB	Cascaded CNN	۹۱,۰۲
[۲۰]	۲۰۱۵	FDDDB	DCNN	۸۵,۰۰
[#]	۲۰۲۳	FDDDB	DCNN	۹۶,۱۱

## ۵. بحث

در این مطالعه برای تشخیص چهره بر روی داده‌ها ناقص از یک مدل عمیق مبتنی بر شبکه‌های عصبی کانولوشنی استفاده شد. فلوچارت کلی روند پیشنهادی در شکل ۱ نمایش داده شد. در فرآیند پیش‌پردازش استخراج لیست ویژگی‌های کاربردی از اجزای چهره انجام شده و با اتکا بر تکنیک‌های 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. روند اعمال عملکرد میانه در الگوریتم پیشنهادی در شکل ۲ نشان داده شده است. 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 بدست آمده، مدل پیشنهادی مبتنی بر CNN عمیق با بهره‌گیری از پنج لایه همگشت، دو لایه پولینگ یا ادغام و به دنبال آن لایه‌ی softmax منجر به تشخیص مناسب محل چهره شد. نتایج تصویری بدست 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 عمیق پیشنهادی نسبت به مدل‌های عمیق معرفی شده در بخش پیشینه (جدول ۵) بهتر و دقیق‌تر عمل کرده و از کارایی مناسب‌تری نسبت به این مدل‌ها برخوردار است. اگر چه 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 داده‌ی اولیه از پایگاه داده، پیش‌پردازش‌های لازم بر روی تصاویر انجام می‌شود. خروجی این مرحله به عنوان ورودی به الگوریتم بهینه‌سازی کرم شب تاب 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 نرم افزار متلب برای پیاده سازی استفاده شده است. روش

این حوزه از مسائل باید چهره‌ها در مقیاس‌ها و حالت‌های مختلف از تصاویر در زوایای متفاوت از افراد را مورد بررسی قرار داد. در روش‌های 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 می‌شود. بنابراین با استفاده از روش‌هایی با کارایی مناسب نسبت به حجم بسیار زیاد داده‌های در دسترس، کاربردهای این حوزه قابل توسعه است. در این مطالعه هدف اصلی بهبود تشخیص چهره بر روی تصاویر با داده‌های ناقص توسط یک شبکه عصبی کانولوشنی عمیق (DCNN) بود. در روش پیشنهادی از یک رویکرد ترکیبی برای محقق کردن هدف ذکر شده، استفاده شد. به این صورت که ابتدا تصویر اولیه در قالب یک پایگاه داده خام آنالیز و بررسی شده و تکنیک پیش‌پردازش داده با اعمال فیلتر و تبدیل هیستوگرام تصاویر اعمال گردید. در گام بعدی با برش یک قطعه از تصویر به صورت تصادفی و بهره‌مندی از تکنیک تشخیص لبه اقدام به بهبود کیفیت تصویر ورودی شد. در مرحله بعدی به کمک الگوریتم بهینه‌سازی زنبور عسل انتخاب ویژگی‌های موثر برای مطابقت و تشخیص صحیح با داده‌های پایگاه داده انجام شد. در مرحله آخر به کمک شبکه عصبی کانولوشنی، انتخاب لایه‌های بهینه یادگیری و تست 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 آنالیز واقع شد. دقت نهایی با تکرار ۲۰۰ برای سناریوی سوم ۹۶،۱۱ درصد بدست آمد. نتایج بدست آمده بیانگر این موضوع است که روش ارائه شده بهبود مورد نظر را محقق کرده است. به طور خلاصه نتایج حاصل از این تحقیق عبارتند از بهبود عملیات تشخیص چهره بر روی تصاویر با داده‌های ناقص توسط شبکه‌های عصبی با کمترین هزینه، استخراج و کنکاش ویژگی‌های اصلی برای بهبود در تشخیص چهره به کمک الگوریتم

پیشنهادی با استفاده از نرم افزار متلب پیاده‌سازی شده و نتایج پیاده‌سازی نشان دهنده‌ی عملکرد مناسب و افزایش دقت در فرآیند تشخیص می‌باشد. در روش ارائه شده به دلیل انتخاب بهینه‌ی ویژگی‌ها، میزان دقت تشخیص ۹۶،۱۱٪ بدست آمده که نسبت به سایر روش‌های ارائه شده در سال‌های اخیر بالاتر است. علاوه 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 هر چه بیشتر عملکرد روش پیشنهادی می‌توان از ترکیب سایر الگوریتم‌های فراابتکاری برای بهینه‌سازی و از ظرفیت‌های پردازنده گرافیکی (GPU<sup>۳۳</sup>) برای افزایش سرعت در روند آموزش مدل‌های عمیق استفاده کرد.

## ۶. نتیجه‌گیری و پیشنهادات

اکنون با توسعه علوم در حوزه‌ی هوش مصنوعی، تعامل بین ماشین و انسان بیشتر شده است. با توجه به اینکه آنالیز اطلاعات در مغز انسان، با سرعت و دقت بسیار بالا قابل بررسی است. بنابراین پژوهشگران جهت برقراری ارتباط مؤثر با دنیای واقعی، حوزه‌های مطالعاتی کثیری از جمله بینایی، ماشین، یادگیری ماشین، رباتیک، پردازش گفتار و تشخیص چهره را به وجود آورده‌اند. پژوهشگران مرتبط با بینایی ماشین برای آنالیز مناسب اطلاعات مرتبط، نیاز به تشخیص و تحلیل اجسام در تصویر و ویدئو دارند. این تشخیص و طراحی با هدف‌های متفاوتی انجام شده است. تشخیص و شناسایی چهره یکی از اساسی‌ترین و پرچالش‌ترین مسائل در بینایی ماشین می‌باشد. تشخیص و شناسایی چهره به علت تفاوت در رنگ، زاویه، حالت، شدت روشنایی و اندازه نسبی می‌تواند باعث ایجاد چالش‌های بزرگی در این زمینه شود که از این رو روش‌ها و الگوریتم‌های قوی برای اتخاذ ویژگی‌های ساختاری از تصاویر، مورد نیاز است. در



- structural similarity index measure and principal component analysis," *Journal of Intelligent & Fuzzy Systems*, vol. 42, 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 Computing Surveys*, vol. 54, 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 Recognition*, vol. 122, pp. 108-308, 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)," *Theoretical Computer Science*, 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," *Learning, Media and Technology*, vol. 45, pp. 115-128, 2020, doi: 10.1080/17439884.2020.1686014.
- [10] Z. Liao, R. D. J. Samuel, and S. Krishnamoorthy, "Computer vision for facial analysis using human-computer interaction models," *International Journal of Speech Technology*, vol. 25, pp. 379-389, 2022, doi: 10.1007/s10772-021-09953-6.
- [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," *Computers and electronics in agriculture*, vol. 152, pp. 261-268, 2018, doi: 10.1016/j.compag.2018.07.012.
- [12] H. N. Vu, M. H. Nguyen, and C. Pham, "Masked face recognition with convolutional neural networks and local binary patterns," *Applied Intelligence*, vol. 52, pp. 5497-5512, 2022, doi: 10.1007/s10489-021-02728-1.
- [13] ز. شریفی مهرجرد، ه. مومنی، ح. ادبی اردکانی، «مروری بر الگوریتم‌های یادگیری ماشین جهت تشخیص اوتیسم با استفاده از سیگنال EEG»، محاسبات نرم، شهریور ۱۴۰۲، doi: 10.22052/SCJ.2023.248522.1110
- بهینه‌سازی زنبور عسل، استفاده از شبکه عصبی کانولوشنی عمیق جهت بهبود تشخیص چهره با داده‌های ناقص، بهبود ضریب اطمینان تشخیص چهره بر روی تصاویر با داده‌های ناقص در سیستم‌های امنیتی و افزایش کیفیت تصویر با پیش‌پردازش بهبود یافته. از جمله محدودیت‌های مهم در این تحقیق پیچیدگی حل مساله بود که با اتکا بر چند روش و تکنیک مختلف پیش‌برده شد تا این پیچیدگی به حداقل برسد. از جمله فعالیت‌ها و کارهای آتی نیز می‌توان به موارد پیش‌رو اشاره کرد. بررسی روش‌هایی که امکان تبدیل تشخیص چهره با داده‌های ناقص را به مدل دارای قابلیت اجرا صرف نظر از دخالت انسان و به صورت کاملا اتوماتیک را فراهم آورد، همچنین استخراج بهترین و مؤثرترین ویژگی‌ها برای تشخیص چهره با استفاده از بهینه‌سازی فرا ابتکاری 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.

## مراجع

[۱] ف. حسینی، ا. شاه بهرامی، «آشکارسازی چهره غیر وابسته به چرخش با استفاده از مدل مبتنی بر بخش»، علوم رایانشی، جلد ۱، شماره ۱، صفحه ۵۰-۴۵، تابستان ۱۳۹۵.

[2] H.O. Ikromovich, and B.B. Mamatkulovich, "Facial Recognition Using Transfer Learning in the Deep CNN," *Open Access Repository*, vol. 4, 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," *Applied Sciences*, vol. 4, pp. 16-48, 2022, doi: 10.3390/app12104816.

[4] K. Chumachenko, A. Iosifidis, and M. Gabbouj, "Self-attention fusion for audiovisual emotion recognition with incomplete data," *26th International Conference on Pattern Recognition (ICPR)*, 2022, pp. 2822-2828, doi: 10.1109/ICPR56361.2022.9956592.

[5] G., Rajeswari and P. Ithaya Rani, "Face occlusion removal for face recognition using the related face by

- [14] V. Jain and E. Learned-Miller, "Fddb: A benchmark for face detection in unconstrained settings," *UMass Amherst technical report*, vol. 2, pp. 1-11, 2010.
- [۱۵] م. افتخاریان، ح. نودهی، «بهبود تشخیص و طبقه‌بندی توده‌های سرطان سینه بر اساس یادگیری عمیق و روش‌های پردازش تصویر»، محاسبات نرم، آذر ۱۴۰۱، doi: 10.22052/SCJ.2023.246701.1086.
- [16] M.T. Islam, T. Ahmed, A. R. Rashid, T. Islam, M. S. Rahman, and M. T. Habib, "Convolutional Neural Network Based Partial Face Detection," in *IEEE 7th International conference for Convergence in Technology (I2CT)*, Apr 2022, pp. 1-6, doi: 10.1109/ICONAT57137.2023.10080063.
- [17] 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 and Cognitive Computing*, vol. 6, pp. 1-9, 2022, doi: 10.3390/bdcc6010009.
- [18] H. Yan, X. Wang, Y. Liu, Y. Zhang, and H. Li, "A new face detection method based on Faster RCNN," *Journal of Physics: Conference Series*, vol. 1754, pp. 102-209, 2021, doi: 10.3390/app12104816.
- [19] H. Qin, J. Yan, X. Li, and X. Hu, "Joint training of cascaded CNN for face detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 3456-3465. doi: 10.1088/1742-6596/1754/1/012209.
- [20] S. S. Farfade, M. J. Saberian, and L. J. Li, "Multi-view face detection using deep convolutional neural networks," in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, Jun 2015, pp. 643-650, doi: 10.1145/2671188.2749408.
- [۲۱] ر. بهمنش، ن. مجمع، «الگوریتم فرا ابتکاری نفرون-۲ (NOA-2)، جهت حل مسائل بهینه‌سازی»، محاسبات نرم، فروردین ۱۴۰۲، doi: 10.22052/SCJ.2023.248427.1104.