

## شناسایی چهره با استفاده از تنظیم دقیق شبکه های کانولوشنی عمیق و رویکرد یادگیری انتقالی

راضیه راستگو<sup>۱</sup> و کوروش کیانی<sup>۲\*</sup>

اطلاعات مقاله	چکیده
دریافت مقاله: ۱۳۹۷/۰۸/۰۴ پذیرش مقاله: ۱۳۹۷/۱۲/۱۴	یادگیری عمیق، یکی از رویکردهای مورد توجه در یادگیری ماشین می باشد که شامل معماری های مهمی می باشد. شبکه کانولوشنی عمیق، یکی از معماری های مورد توجه در یادگیری عمیق می باشد که در پردازش های مربوط به تصاویر دیجیتالی کاربرد فراوانی پیدا کرده است. در این پژوهش، شبکه کانولوشنی Alexnet، به منظور شناسایی چهره در عکس های ورودی، مورد استفاده قرار گرفته است. تنظیم دقیق مدل از قبل تعلیم داده شده ی Alexnet، با تبدیل لایه های کاملا متصل به لایه های کانولوشنی و اعمال فیلتر های مناسب، انجام شده است. استفاده از برش های مختلف عکس ورودی و نیز افزایش تعداد لایه های کانولوشنی به منظور استخراج خصوصیت های با سطح بالاتر به همراه فیلترهای مناسب در مدل های پیشنهادی مورد توجه قرار گرفته است. به منظور تجسم اعمال فیلترها در لایه های مختلف، از روش کانولوشن معکوس استفاده شده است. از دو پایگاه داده ی Caltech face و LFW به منظور نشان دادن نتایج، استفاده شده است. پس از پردازش های لازم بر روی پایگاه داده های مورد استفاده، نتایج به دست آمده از شبکه Alexnet، قبل و بعد از تنظیم دقیق، مورد بررسی قرار گرفته است. نتایج بررسی، حاکی از بهبود عملکرد شبکه، تحت عملیات انجام شده، می باشد.
واژگان کلیدی: یادگیری عمیق، شبکه کانولوشنی عمیق، تنظیم دقیق، شناسایی چهره، یادگیری انتقالی، کانولوشن معکوس.	

### ۱- مقدمه

شبکه های عصبی مصنوعی از سال ۱۹۴۰ تاکنون مورد استفاده قرار گرفته اند [۱]. امروزه شبکه های عصبی مصنوعی کاربرد زیادی را در زمینه های مختلف پیدا کرده اند [۲-۶]. از جمله ی این کاربردها میتوان به این موارد اشاره کرد: تولید سیگنال های مصنوعی زلزله [۷]، مدلسازی انرژی ضربه ی فولادهای مرتبه ای [۸]، تخمین پتانسیل تحکیم خاک های رسی [۹]، شبیه سازی و طبقه بندی کیفیت توان [۱۰]، پردازش طبیعی زبان [۱۱]، و بهبود کارایی شبکه های هوشمند [۲ و ۸].

شبکه های عصبی مصنوعی عمیق، گونه ای از شبکه های عصبی هستند که تعداد لایه ها و پارامترهای بیشتری

نسبت به شبکه های عصبی مصنوعی معمول دارند [۱]. در سالهای اخیر، این گونه از شبکه های عصبی، کاربردهای زیادی را در زمینه تشخیص الگو و یادگیری ماشین داشته اند [۹].

یادگیری عمیق، حوزه ی جدیدی از یادگیری ماشین می باشد که شامل مجموعه ای از الگوریتم های یادگیری ماشین برای آموزش شبکه های مصنوعی عمیق می باشد [۷]. یادگیری عمیق، قادر به استخراج خودکار ویژگی های عمیق و سطح بالا می باشد. این نوع از یادگیری، از پایگاه داده های بسیار بزرگ برای آموزش شبکه، استفاده می کند [۸]. تعدادی از مدل های از قبل تعلیم شده، تحت یادگیری عمیق، وجود دارد که دسترسی کاربران را به این خصوصیت

\* پست الکترونیک نویسنده مسئول: Kourosh.kiani@semnan.ac.ir

۱. دانشجوی دکتری هوش مصنوعی، دانشگاه سمنان

۲. استادیار، دانشکده مهندسی برق و کامپیوتر، دانشگاه سمنان

- (Pooling)، لایه softmax تشکیل شده اند [۱۷].
- ماشین بولتزمان محدود شده: اگر ارتباطات میان نرون ها در ماشین بولتزمان را محدود کنیم، به گونه ای از شبکه های عصبی عمیق میسریم که ماشین بولتزمان محدود شده نام دارد [۱۸].
- شبکه های باور عمیق: گونه ای از شبکه های عصبی عمیق هستند که از قرار دادن تعداد زیادی ماشین بولتزمان محدود شده، به دست می آیند [۱۹].
- شبکه رمز گذار خودکار: این شبکه های مصنوعی عمیق، به منظور تولید مجدد ورودی های شبکه به کار می روند [۱۷].
- شبکه های عصبی بازگشتی: هنگامی که با داده های پشت سر هم در بازه های زمانی سرو کار داریم، شبکه های عصبی بازگشتی کارایی بسیار خوبی را از خود نشان می دهند [۱۷].

## ۲-۲- یادگیری انتقالی

یادگیری انتقالی، گونه ای از یادگیری می باشد که در طی آن، از اطلاعات و پارامترهای ذخیره شده در آموزش های قبلی شبکه استفاده می کنیم [۲۰]. این امر موجب می گردد تا بتوانیم از شبکه های عمیقی که قبلاً آموزش داده شده اند در کاربردهای متفاوتی استفاده کنیم. این نوع یادگیری، کاربردهای زیادی در زمینه استخراج خصوصیت ها، تنظیم دقیق شبکه و نیز استفاده از مدل های از قبل تعلیم شده دارد. از آنجایی که آموزش شبکه های عمیق، به امکانات سخت افزاری پیشرفته همچون GPU نیاز دارد و همچنین زمان آموزش این شبکه ها طولانی می باشد، استفاده از یادگیری انتقالی بسیار مفید و مقرون به صرفه خواهد بود.

## ۲-۳- مدل های از قبل تعلیم شده

مدل های از قبل تعلیم شده، دسته ای از شبکه های عصبی عمیق هستند که با پایگاه داده های عظیم و با استفاده از سخت افزارهای قدرتمند، آموزش داده شده اند و خصوصیات استخراج شده توسط آنها و نیز پارامترهای شبکه، تحت عنوان یک مدل ذخیره می شوند تا دیگران بتوانند بدون نیاز مجدد به آموزش شبکه، از نتایج آن استفاده کنند. تعدادی مدل از قبل تعلیم شده وجود دارد که توسط شرکت های تجاری بزرگ حمایت شده اند و توسط کاربران زیادی مورد استفاده قرار می گیرند. از

های سطح بالای شبکه، بدون نیاز مجدد به آموزش شبکه، فراهم می سازد. تعدادی از این مدل ها، در بخش ۲-۳ معرفی شده اند. استفاده از این مدل ها، مفهومی به نام یادگیری انتقالی و تنظیم دقیق را مطرح کرده است که شامل استفاده از مدل های از قبل تعلیم شده در کاربردی جدید و مجزای از کاربرد آموزش داده شده می باشد. کاربردهای فراوانی برای یادگیری عمیق و معماری های مختلف آن وجود دارد [۱۱-۱۴]. از جمله پرکاربردترین آنها می توان به کاربردهای دسته بندی داده های مختلف اشاره کرد. شبکه کانولوشنی Alexnet، یکی از معروفترین مدل های موجود، برای دسته بندی داده ها می باشد [۱۵]. این مدل، قادر به دسته بندی داده های مربوط به پایگاه داده ی ImageNet [۱۶]، می باشد. ما در این پژوهش، این مدل را مورد استفاده قرار داده ایم.

در این پژوهش، تأثیر یادگیری انتقالی و نیز تنظیم دقیق مدل های از قبل تعلیم شده، بر روی پایگاه داده های LFW و Caltech مورد بررسی قرار گرفته است. مدل Alexnet، به منظور نشان دادن نتایج، مورد استفاده قرار گرفته است. این پژوهش به این صورت سازماندهی شده است: بخش دوم، شامل معرفی کلی شبکه های عصبی عمیق و یادگیری عمیق می باشد که معماری های مهم یادگیری عمیق و نیز مدل های از قبل تعلیم شده، معرفی شده اند. روش ارزیابی در بخش ۳ مطرح شده است. بخش های ۴ و ۵ نیز شامل نتیجه های به دست آمده از روش ارزیابی شده و نتیجه گیری می باشند.

## ۲- یادگیری عمیق

یادگیری عمیق، یکی از رویکردهای مورد توجه در یادگیری ماشین می باشد که در طی سال های اخیر، مورد توجه زیادی قرار گرفته است. پردازش تصاویر دیجیتال، یکی از حوزه های پرکاربرد در یادگیری عمیق می باشد. در ادامه این بخش، به معرفی مهمترین معماری های یادگیری عمیق، مفهوم یادگیری انتقالی و نیز مدل های از قبل تعلیم شده می پردازیم.

## ۲-۱- مهمترین معماری های یادگیری عمیق

معماری های گوناگونی برای یادگیری عمیق وجود دارد که مهم ترین آنها عبارتند از:

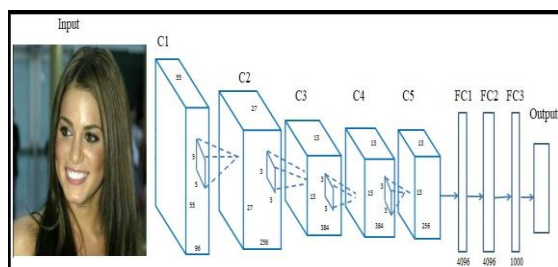
- شبکه عصبی کانولوشنی: این شبکه ها از تعدادی لایه-ی کانولوشن، لایه فعالسازی (ReLU)، لایه ادغام

فیلترهای مختلف در لایه‌ها را ببینیم و بتوانیم از این فیلتر-ها برای تشخیص نواحی مختلف در عکس استفاده کنیم.

#### ۴- نتایج به دست آمده

در این بخش، نتایج به دست آمده از تنظیم دقیق، تجسم و نمایش لایه‌های کانولوشنی، میزان بهبود دقت شبکه مورد بررسی قرار گرفته است.

همان گونه که در بخش‌های قبل بیان شد، مدل Alexnet، یک شبکه‌ی کانولوشنی شامل ۵ لایه‌ی کانولوشن، سه لایه-ی کاملاً متصل و تعدادی لایه‌ی پولینگ، فعالسازی (ReLU) و Dropout می‌باشد. این مدل در مجموع، ۲۵ لایه دارد. معماری این مدل در شکل (۱) نشان داده شده است.



شکل ۱- معماری لایه‌های مدل Alexnet.

در ابتدا، مدل از قبل تعلیم داده شده‌ی Alexnet را بدون تغییری در ساختار لایه‌های آن، به کار می‌گیریم و برای دسته‌بندی داده‌های مربوط به پایگاه داده‌های LFW و Caltech Faces به کار می‌بریم. پایگاه داده‌ی LFW شامل ۱۳۲۳۳ عکس [۲۲] و پایگاه داده‌ی Caltech Faces شامل ۴۵۰ عکس [۲۳] می‌باشد. به دلیل استفاده از این CPU، تنظیم دقیق مدل مورد نظر، با استفاده از این پایگاه داده‌های نسبتاً بزرگ، به زمان زیادی نیاز دارد. در صورت استفاده از GPU، زمان مورد نظر به میزان قابل توجهی کاهش می‌یابد.

پس از انتخاب پایگاه داده‌ها، ما ابتدا پایگاه داده‌های مورد نظر را برای استفاده به عنوان داده‌های ورودی شبکه، آماده کردیم. این آماده‌سازی شامل تقسیم عکس‌های پایگاه داده دورن پوشه‌هایی می‌باشد که هر پوشه شامل عکس‌های یک شخص خاص می‌باشد. بنابراین، به تعداد اشخاص مختلف در هر پایگاه داده، پوشه خواهیم داشت. پایگاه داده‌های مورد استفاده، شامل عکس‌هایی از اشخاص مختلف، در حالت‌ها و زاویه‌های مختلف می‌باشند. پس از آن، عملیات

جمله‌ی این مدل‌ها می‌توان به موارد زیر اشاره کرد:

- Alexnet: یک شبکه عصبی کانولوشنی عمیق می‌باشد که شامل ۵ لایه‌ی کانولوشن، سه لایه کاملاً متصل و تعدادی لایه‌های فعالسازی و پولینگ می‌باشد. این مدل، برنده ی رقابت سالیانه‌ی طبقه‌بندی عکس‌ها با استفاده از پایگاه داده‌های بزرگ در سال ۲۰۱۴ می‌باشد [۱۵]. ما در این پژوهش از این مدل استفاده می‌کنیم.
- VGG-16: این مدل شامل ۱۶ لایه می‌باشد که دارای دقت مناسبی بر روی پایگاه داده‌های بزرگ می‌باشد. این مدل از کتابخانه‌ی caffe برای آموزش و تنظیم دقیق استفاده می‌کند [۲۱].
- VGG-19: مدل ارتقا یافته‌ی VGG-16 می‌باشد که تعداد لایه‌های بیشتری دارد. این مدل، بهبود اندکی نسبت به VGG-16 دارد [۲۱].

#### ۳- روش ارایه شده

در این بخش، به ارایه‌ی جزئیات روش ارایه شده می‌پردازیم. مدل از قبل تعلیم شده‌ی Alexnet، مورد استفاده قرار گرفته است. ابتدا این مدل، تنظیم دقیق شده است تا برای کاربرد مورد نظر قابل استفاده باشد. برای این منظور، لایه‌ی اول تا شانزده را بدون تغییر مورد استفاده می‌دهیم. لایه‌ی هفده تا بیست و پنج را تغییر می‌دهیم. به این ترتیب که تعدادی از لایه‌های کاملاً متصل را به لایه کانولوشنی تبدیل می‌کنیم. تنظیم صحیح تعداد فیلترها و پنجره‌ی فیلتر، تأثیر زیادی در نتایج به دست آمده دارد. ما از طریق امتحان کردن مقادیر گوناگون، بهترین مقادیر را برای کاربرد مورد نظر اعمال کردیم.

لایه‌های کانولوشنی، قابلیت تجسم و نمایش اعمال نتایج به دست آمده در این لایه‌ها بر روی داده‌های ورودی را دارند. این قابلیت به ما کمک می‌کند تا بتوانیم نتایج به دست آمده از اعمال فیلترها در لایه‌های کانولوشنی را ببینیم. ما از روش کانولوشن معکوس، به منظور نمایش نتایج به دست آمده از لایه‌های کانولوشنی بر روی داده‌های ورودی، استفاده کرده‌ایم.

یکی دیگر از قابلیت‌های شبکه‌ی کانولوشنی، امکان نمایش heat-map داده‌های ورودی، به منظور مشخص کردن اشیای موجود در عکس، می‌باشد. ما این کار را برای لایه‌های کانولوشنی انجام داده‌ایم تا بتوانیم نتیجه‌ی اعمال

های پیشنهادی اول و دوم، ۱۶ لایه نخست را بدون تغییر انتقال داده و یک یا دو لایه از لایه های کاملا متصل را به لایه کانولوشنی تبدیل می کنیم. با این تفاوت که در این مدل ها، ورودی های شبکه، عکس-های برش داده شده هستند که از هر عکس اصلی در ورودی شبکه، پنج برش مورد استفاده قرار گرفته است. نواحی پیکسلی مربوط به گوشه سمت راست بالا، گوشه سمت چپ بالا، گوشه سمت راست پایین، گوشه سمت چپ پایین و در نهایت، ناحیه پیکسلی مرکزی عکس ورودی، مورد استفاده قرار گرفته است. جزئیات بیشتر در رابطه به نحوه برش عکس ورودی، در مرجع [۱۴] قابل مشاهده می باشد. شکل های (۴) و (۵)، مدل های پیشنهادی سوم و چهارم را نشان می دهد. اضافه کردن لایه های کانولوشنی، سبب استخراج خصوصیات سطح بالاتر شده و منجر به افزایش دقت تشخیص می شود.

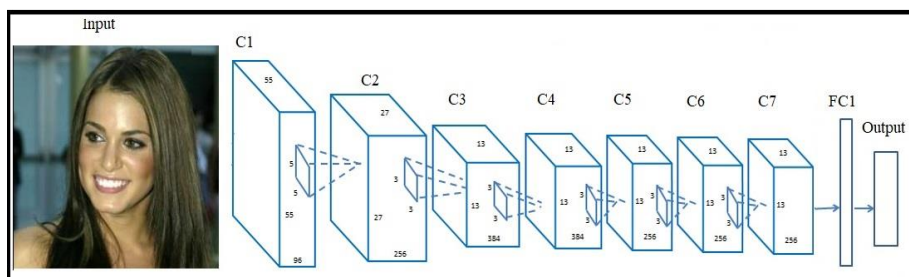
نتایج به دست آمده از تنظیم دقیق شبکه، حاکی از بهبود عملکرد شبکه می باشد. جدول ۱ نتایج به دست آمده را برای دو حالت، قبل از تنظیم دقیق و بعد از آن، نشان می دهد.

پیش پردازش را انجام می دهیم که شامل تنظیم اندازه ی صحیح عکس ها، به منظور ورود به شبکه ی مورد نظر، نیز می باشد.

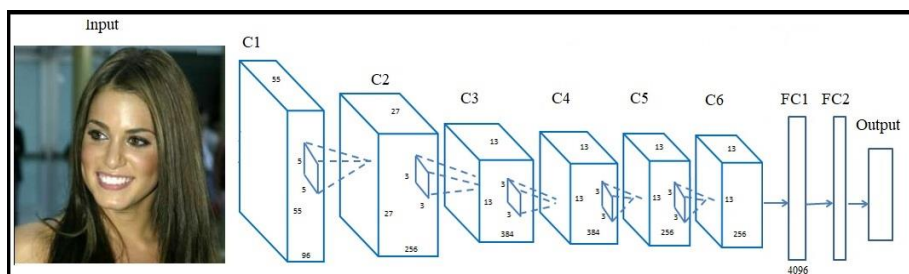
پس از آماده سازی پایگاه داده ها، عملیات تنظیم دقیق را به دو روش و تحت عنوان دو مدل پیشنهادی، به صورت زیر، بر روی لایه های شبکه انجام می دهیم:

- در مدل های پیشنهادی اول و دوم، ۱۶ لایه ی اول را بدون تغییر، انتقال می دهیم. به همین دلیل است که نام یادگیری انتقالی را به کار می برند. از لایه ی ۱۷ به بعد را تنظیم دقیق می کنیم. در مدل اول، اولین لایه ی کاملا متصل را به لایه های کانولوشنی تبدیل می کنیم و دولایه کاملا متصل باقیمانده را نیز پس از ششمین لایه کانولوشنی قرار می دهیم. در این صورت، مدل پیشنهادی شامل شش لایه کانولوشنی و دو لایه کاملا متصل می باشد. شکل (۲)، مدل ارائه شده توسط ما را نشان می دهد. در مدل دوم، دو لایه از لایه های کاملا متصل را به لایه کانولوشنی تبدیل می کنیم. شکل (۳)، معماری مدل را نشان می دهد. در این دو مدل، ورودی های شبکه، عکس های معمولی و دستکاری نشده در دیتاست هستند.

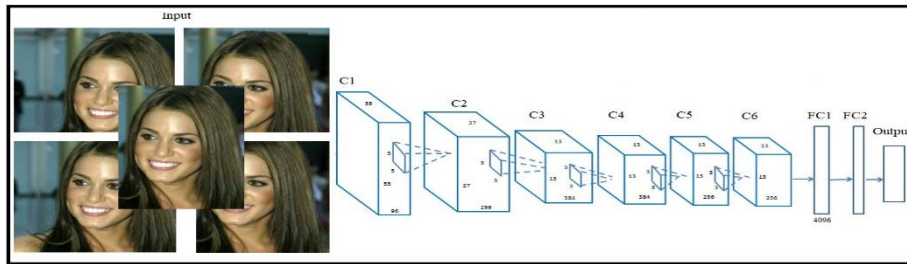
- در مدل های پیشنهادی سوم و چهارم، همانند مدل



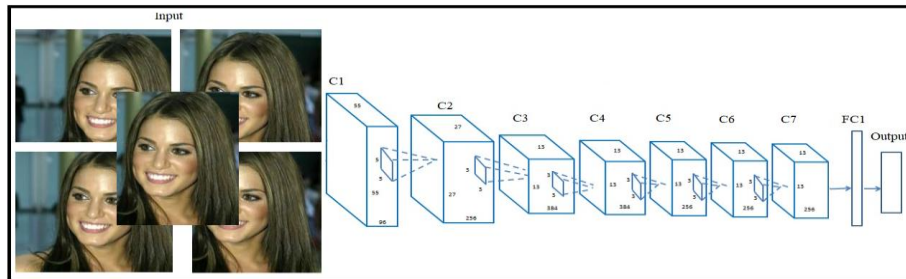
شکل ۲- معماری مدل پیشنهادی اول شامل شش لایه کانولوشن و دو لایه کاملا متصل



شکل ۳- معماری مدل پیشنهادی دوم شامل هفت لایه کانولوشنی و یک لایه کاملا متصل



شکل ۴- معماری مدل پیشنهادی سوم شامل شش لایه کانولوشنی و دو لایه کاملاً متصل با ورودی های برش داده شده



شکل ۵- معماری مدل پیشنهادی چهارم شامل هفت لایه کانولوشنی و یک لایه کاملاً متصل با ورودی های برش داده شده

لایه‌های کانولوشنی و تمام متصل و نیز نوع داده ورودی این مدل‌ها را نشان می‌دهد. تفسیر نتایج به دست آمده از تنظیم دقیق شبکه، با استفاده از پایگاه داده‌های LFW و Caltech، در واقع همان توانایی شناخت افراد مختلف توسط شبکه می‌باشد. از آنجایی که هر پوشه در پایگاه داده-ی آماده شده توسط ما، شامل عکس‌های یک شخص خاص می‌باشد، پس مفهوم دسته بندی در اینجا، همان شناخت افراد مختلف می‌باشد.

به منظور ارزیابی بهتر مدل‌های ارائه شده، نتایج به دست آمده از مدل‌های پیشنهادی را با تعدادی از مدل‌های موجود در زمینه تشخیص چهره مقایسه نموده‌ایم. نتایج حاصل از مقایسه دقیق‌ترین مدل پیشنهادی توسط ما، مدل چهارم، با سایر روش‌ها بر روی پایگاه داده LFW در جدول شماره ۳ قابل مشاهده می‌باشد. همان گونه در این جدول قابل مشاهده می‌باشد، مدل پیشنهادی چهارم، دقت تشخیص بهتری نسبت به سایر مدل‌ها دارد. استفاده از برش‌های مختلف عکس ورودی، نه تنها منجر به افزایش تعداد داده-های ورودی می‌شود؛ بلکه سبب افزایش پایداری سیستم جهت مواجهه با شرایطی می‌شود که ورودی کامل در دسترس نمی‌باشد و یا بخش‌هایی از آن دچار تخریب شده است. پس از انجام آزمایش‌های مختلف جهت تعیین تعداد مناسب برش‌ها، پنج برش مختلف از عکس ورودی در نظر گرفته می‌شود.

جدول ۱- نتایج موبوط به دقت شبکه ی Alexnet قبل از تنظیم دقیق و پس از آن

مدل	دقت بر روی پایگاه داده LFW	دقت بر روی پایگاه داده Caltech Faces
مدل Alexnet اصلی	۸۶/۲۷	۸۶/۶۷
مدل پیشنهادی اول	۹۶/۱۲	۹۷/۰۴
مدل پیشنهادی دوم	۹۷/۳۶	۹۸/۶۸
مدل پیشنهادی سوم	۹۸/۰۱	۹۸/۸۷
مدل پیشنهادی چهارم	۹۹/۰۲	۹۹/۳۴

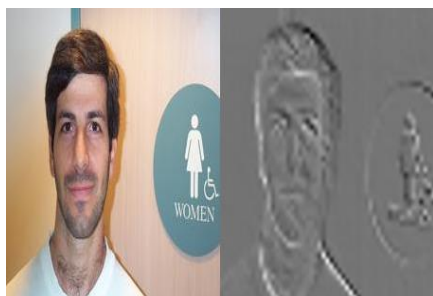
مدل‌های ارائه شده در تعداد لایه‌های کانولوشنی، تعداد لایه‌های تمام متصل، و نیز در داده ورودی با هم تفاوت دارند. به منظور بررسی تأثیر تعداد لایه‌های کانولوشنی در استخراج ویژگی‌های سطح بالاتر و نیز آنالیز تأثیر استفاده از برش‌های مختلف عکس ورودی، مدل‌های پیشنهادی مورد بررسی قرار گرفته‌اند و نتایج تجربی آنها بر روی دو پایگاه داده عمومی گزارش شده است. بررسی نتایج، حاکی از بهبود نتایج در اثر افزایش تعداد لایه‌های کانولوشنی می‌باشد که منجر به استخراج ویژگی‌های سطح بالاتر از عکس ورودی در برش‌های مختلف عکس می‌شود. جدول ۲، تعداد

خروجی مربوط به یک کانال عکس در لایه کانولوشن اول می باشد. نقاط روشن تر مربوط به نقاطی هستند که مقدار فعالسازی بیشتری نسبت به سایر نقاط دارند.



شکل ۷- بیشترین میزان تابع فعالسازی در کانال های لایه ی کانولوشن اول

بعضی از مربعات موجود در شکل (۷)، حاوی اطلاعات خاصی نیستند و به صورت خاکستری هستند. شکل (۸)، فعالسازی مربوط به یکی از این کانال ها را در ابعاد بزرگتری نشان می دهد.



شکل ۸- عکس ورودی و خروجی مربوط به تابع فعالسازی یکی از کانال ها

همچنین می توانیم کانالی را که حاوی بیشترین میزان تابع فعالسازی می باشد را به صورت شکل (۶) نمایش دهیم. همان گونه که در شکل (۵) مشخص است، این شکل شامل لبه های عکس می باشد و به اصطلاح، روی لبه ها فعال می باشد.

در شبکه های کانولوشنی عمیق، هر چه به سمت لایه ی خروجی نزدیکتر می شویم، خصوصیت های استخراج شده از شبکه، سطح بالاتر و عمیق تر می باشند. به عنوان مثال،

جدول ۲- جزییات مدل های پیشنهادی

ورودی	تعداد لایه تمام متصل	تعداد لایه کانولوشنی	مدل
عکس کامل	۲	۶	مدل اول
عکس ورودی	۱	۷	مدل دوم
پنج برش از عکس ورودی	۲	۶	مدل سوم
پنج برش از عکس ورودی	۱	۷	مدل چهارم

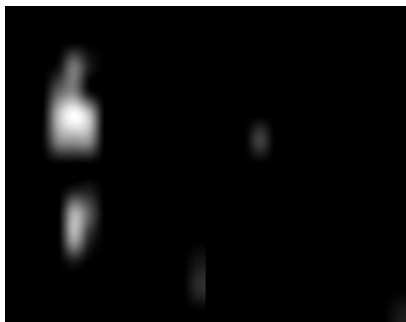
جدول ۳- مقایسه دقت تشخیص مدل پیشنهادی با سایر مدل ها

مدل	دقت مدل
DeepFace [۲۴]	۹۷/۳۵
VGGface [۲۵]	۹۸/۹۵
Light-CNN [۲۶]	۹۸/۸۰
مدل پیشنهادی ما	۹۹/۰۲

به منظور تجسم نتایج به دست آمده از اعمال فیلترها در لایه های کانولوشنی، تأثیر این فیلترها بر داده های ورودی را بررسی می کنیم. به این منظور، از روشی به نام کانولوشن معکوس استفاده می کنیم. خروجی این روش، داده های ورودی در لایه اول شبکه می باشد که توسط فیلترهای لایه ی کانولوشنی، تحت تأثیر قرار گرفته اند. مزیت این گونه نمایش، شناسایی فیلترهای گوناگون برای کاربرد مختلف، بر اساس نتیجه ی آنها بر روی داده های ورودی، می باشد. شکل (۶)، خروجی حاصل از لایه ی کانولوشن اول را نشان می دهد. همان گونه که در شکل مشخص است، خروجی این لایه شامل استخراج ویژگی های سطح پایین، از جمله لبه ها، می باشد.



شکل ۶- خروجی حاصل از کانولوشن معکوس مربوط به لایه ی کانولوشن اول هر کدام از مربعات کوچک در شکل (۷)، نشان دهنده ی



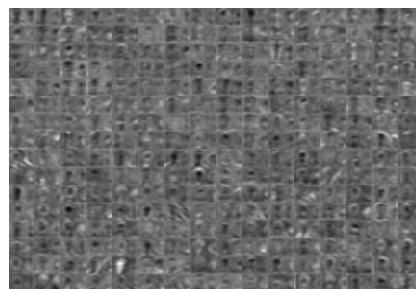
شکل ۱۱- خروجی حاصل از اعمال تابع فعالسازی بر روی لایه کانولوشنی پنجم (Heat-map تصویر)

با مقایسه‌ی شکل‌های (۱۰) و (۱۱) متوجه می‌شویم که شکل (۹)، خصوصیت‌های مربوط به عکس ورودی را به صورت واضح تری مورد توجه و تأکید قرار می‌دهد.

#### ۵- نتیجه‌گیری

در این پژوهش، نتایج حاصل از تنظیم دقیق و یادگیری انتقالی مدل از قبل تعلیم شده‌ی Alexnet بر روی پایگاه داده‌های LFW و Caltech ارایه شده است. این نتایج حاکی از بهبود عملکرد و دقت شبکه می‌باشد. تنظیم دقیق شبکه‌ی کانولوشنی Alexnet، با استفاده از تبدیل لایه‌های کاملاً متصل به لایه‌های کانولوشنی و نیز به کارگیری فیلترهای مناسب برای لایه‌های کانولوشنی، از جمله عملیاتی بوده است که منجر به بهبود کارایی و دقت شبکه شده است. استفاده از برش‌های مختلف عکس‌های ورودی، نه تنها موجب افزایش تعداد داده‌های ورودی می‌شود؛ بلکه سبب بهبود کارایی مدل در مواجهه با شرایطی می‌شود که ورودی کامل در دسترس نمی‌باشد و تنها بخش‌هایی از آن در دسترس می‌باشد. همچنین، heat-map خروجی شبکه، به منظور شناسایی افراد موجود در عکس، نشان داده شده است. روش کانولوشن معکوس، به منظور نشان دادن نتایج حاصل از اعمال فیلترها در لایه‌های کانولوشنی نیز مورد استفاده قرار گرفته است.

از آنجایی که داده‌های مربوط به پایگاه داده‌های مورد استفاده، به گونه‌ای پیش پردازش شده‌اند که عکس‌های مربوط به هر شخص، درون یک پوشه قرار گرفته‌اند، مفهوم دسته بندی در اینجا، در واقع همان شناسایی افراد مختلف می‌باشد. مطالعات ما بر روی این کار ادامه دارد و همچنان



شکل ۹- خروجی حاصل از کانولوشن معکوس مربوط به لایه کانولوشنی پنجم

در لایه‌ی کانولوشن پنجم، شاهد خصوصیت‌های مربوط به ویژگی‌های خاصی از داده‌های ورودی، مثلاً مکان چشم‌ها، صورت، مکان یک شی خاص در شکل، می‌باشیم. در شکل (۹)، خروجی مربوط به لایه‌ی کانولوشن پنجم، با استفاده از روش کانولوشن معکوس، نشان داده شده است. به منظور تمرکز بر جزئیات مربوط به خروجی لایه کانولوشن پنجم، خروجی مربوط به کانالی با بیشترین میزان تابع فعالسازی را بررسی می‌کنیم. این خروجی در شکل (۱۰) نشان داده شده است. نقاط روشن در این شکل، نشان دهنده‌ی نقاطی است که میزان فعالسازی بیشتری دارند و نقاط تیره‌تر، نقاط کم‌اهمیت عکس می‌باشند. در واقع، با تجزیه و تحلیل خروجی حاصل از این لایه می‌توان پی برد که این لایه بر چه خصوصیتی بیشتر تمرکز می‌کند.



شکل ۱۰- خروجی حاصل از کانولوشن معکوس مربوط به بیشترین تابع فعالسازی لایه کانولوشنی پنجم

از آنجایی که مقدار تابع فعالسازی در لایه‌ی کانولوشن پنجم شامل مقادیر مثبت و منفی می‌باشد، تابع فعالسازی ReLu بر روی خروجی‌های این لایه اعمال می‌شود تا تمامی مقادیر، مثبت باشند. به این شکل، heat-map تصویر نیز گفته می‌شود. خروجی حاصل از اعمال این تابع فعالسازی بر روی لایه کانولوشنی پنجم در شکل (۱۱) نشان داده شده است.

افراد در شرکت راه حل تیز هوش برتر (هیس)، High Intelligent Solution (HIS) تعریف شده و در حال پیگیری می باشد. ما به سهم خود، از حمایت بی دریغ شرکت هایس و نیز شرکت NVIDIA به خاطر پردازنده گرافیکی Tian XP GPU، تشکر می کنیم.

به دنبال بهبود بیشتر و استفاده ی بهتر از معماری های یادگیری عمیق هستیم.

### تقدیر و تشکر

این پژوهش، به عنوان بخشی از پروژه ی شناسایی چهره

### مراجع

- [1] J. Heaton, "Neural Networks and Deep Learning", Heaton Research, Vol. 3, No. 3, 2015, pp. 1-268.
- [2] R. Rastgoo, V. Sattari-Naeini, "A Neuro-Fuzzy QoS-Aware Routing Protocol for Smart Grids", 22nd Iranian IEEE Conference on Electrical Engineering (ICEE 2014), Shahid Beheshti University, Tehran, Iran, 2014, pp. 1080-1084.
- [3] R. Rastgoo, V. Sattari-Naeini, Tuning Parameters of the QoS-Aware Routing Protocol for Smart Grids Using Genetic Algorithm, International Journal of Applied Artificial Intelligence, Vol. 30, No.1, 2016, pp. 52-76.
- [4] F. Bordbar, R. Rastgoo, M.A. Askarzadeh, M.S. Tavallali, Prediction of Residential Natural Gas Consumption Using Artificial Neural Network, The 9th International Chemical Engineering Congress & Exhibition (IChEC 2015) Shiraz, Iran, 26-28 December, 2015.
- [5] R. Rastgoo, V. Sattari-Naeini, GSOMCR: Multi-Constraint Genetic-Optimized QoS-Aware Routing Protocol for Smart Grids. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, Vol. 42, No. 2, 2018, pp. 185-194.
- [6] R. Rastgoo, V. Sattari-Naeini, Multi-Constraint Optimal Path Finding for QoS-Enabled Smart Grids: A Neuro-Fuzzy Approach, Vol. 4, No. 2, 2017, pp. 47-61.
- [7] ع. سلیمانی ایوری، محمد رضا فدوی امیری و حسین مروی، "تولید سیگنال مصنوعی زلزله به کمک مدلی جدید در فشرده سازی و آموزش شبکه های عصبی مصنوعی"، مجله مدل سازی در مهندسی، دانشگاه سمنان، سال ۱۴، شماره ۴۶، پاییز ۱۳۹۵، صفحه ۷۵-۸۵.
- [8] علی نظری، "مدلسازی انرژی ضربه ی فولادهای مرتبه ای با استفاده از شبکه های عصبی مصنوعی"، مجله مدل سازی در مهندسی، دانشگاه سمنان، دوره ۱۴، شماره ۴۵، تابستان ۱۳۹۵، صفحه ۱۶۲-۱۴۵.
- [9] جواد احدیان و فاطمه بهروزی، "کاربرد سیستم تطبیقی ANFIS در تخمین پتانسیل تحکیم خاکهای رسی"، مجله مدل سازی در مهندسی، دانشگاه سمنان، سال ۱۴، شماره ۴۵، تابستان ۱۳۹۵، صفحه ۳۱-۱۷.
- [10] زهرا مروج و جواد آذرخش، "شبیه سازی و طبقه بندی وقایع کیفیت توان با استفاده از شبکه عصبی"، مجله مدل سازی در مهندسی، دانشگاه سمنان، دوره ۱۳، شماره ۴۱، تابستان ۱۳۹۴، صفحه ۱۴۶-۱۳۷.
- [11] T. Du., V. Shanker, "Deep Learning for Natural Language Processing", Eeci.Udel.edu, 2009, pp. 1-7.
- [12] L. Wang, G. Wang, and D. Sng, "Deep Learning Algorithms with Applications to Video Analytics for A Smart City: A Survey", Computer Vision and Pattern Recognition, 2015, pp. 1-6.
- [13] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning", MIT Press Book, 2016, pp.1-800.
- [14] R. Rastgoo, K. Kiani, and S. Escalera, "Multi-Modal Deep Hand Sign Language Recognition in Still Images Using Restricted Boltzmann Machine", Entropy, Vol. 20, No. 11, 2018, 809; <https://doi.org/10.3390/e20110809>.
- [15] A. Krizhevsky, I. Sutskever, GE. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", Proceeding NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems, USA, pp. 1097-1105, 2012.
- [16] <http://www.image-net.org/>
- [17] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview", Technical Report of Lab IDSIA, 2014, pp. 1-88.



- [18] B. Marlin, K. Swersky, B. Chen, and N. Freitas, "Inductive Principles for Restricted Boltzmann Machine Learning", Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (PMLR), Vol. 9, pp. 509-516, 2010.
- [19] G. E. Hinton, S. Osindero, Y. The, "A Fast Learning Algorithm for Deep Belief Nets", Neural Computation, Vol. 18, 2006, pp. 1527-1554.
- [20] S. Gutstein, Transfer Learning Techniques for Deep Neural Nets, Ph.D Thesis, University of Texas, 2010, pp. 1-120.
- [21] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", Computer Vision and Pattern Recognition, 2015, pp. 1-14.
- [22] <http://vis-www.cs.umass.edu/lfw/>
- [23] <http://www.vision.caltech.edu/html-files/archive.html>
- [24] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, DeepFace: Closing the gap to human-level performance in face verification", In CVPR, 2014, pp.1701–1708.
- [25] O. M. Parkhi, A. Vedaldi, A. Zisserman et al, Deep face recognition, In BMVC, 2015, pp. 1-6.
- [26] X. Wu, R. He, Z. Sun, and T. Tan, "A light-CNN for deep face representation with noisy labels", arXiv preprint arXiv: 1511.02683, 2015.