یازدهمین کنگره ملی سراسری
فناوریهای نوین در حوزه توسعه پایدار ایران
11th National Congress of
the New Technologies in Sustainable Development of Iran
senaconf.ir

مقایسه و روند پیشرفت روش های تشخیص سن با استفاده از یادگیری عمیق

محمدرضا جهانبخش(نویسنده مسئول)[۱] ، زهرا اشراقی [۲]، لیلی سلیمیان[۳]

[۱]کارشناس ارشد مهندسی شیمی دانشگاه شیراز، شیراز m.r.jahanbakhsh@shirazu.ac.ir

[۲] دانشجوی کارشناسی مهندسی شیمی جهاد دانشگاهی اصفهان، اصفهان Zahraeshraghi99@gmail.com

[۳] کارشناس ارشد مهندسی کامپیوتر دانشگاه آزاد لنجان، لنجان Leili.salimian@gmail.com

چکیده

سن همیشه ویژگی مهم هویتی بوده است. همچنین عامل مهمی در تعامل اجتماعی است. وضعیت قرارگیری صورت، چین و چروک و صوت صدا از عناصری هستند که پیش بینی سن کاربر را تسهیل می کنند. برآورد سن از طریق تجزیه و تحلیل عددی کاربردهای بالقوه بسیاری از جمله توسعه رابط های هوشمند انسان و ماشین و بهبود ایمنی و حفاظت در بخشهای مختلف مانند حمل و نقل ، امنیت و دارو را پیدا می کند. در بسیاری از آثار ، محققان به خصوص به ویژگی های صورت علاقه مندند تا سن را تخمین بزنند. پیشرفت های اخیر در هوش مصنوعی (AI) و به ویژه تکنیک های یادگیری عمیق (DL) انگیزه های استفاده از این روش ها را برای تخمین سن افزایش می دهد. همچنین استفاده از الگوریتم‌های مختلف مثل الگوریتم ژنتیک برای بهینه‌سازی حجم اطلاعات مرسوم است.

واژه‌های کلیدی

تشخیص سن، هوش مصنوعی، یادگیری عمیق، الگوریتم ژنتیک، بهینه سازی

۱ – مقدمه

در سال ۱۹۵۰ ، "آزمون تورینگ" به عنوان توضیحی مناسب درباره چگونگی انجام استدلال شناختی رایانه توسط انسان پیشنهاد شد. به عنوان یک زمینه تحقیقاتی، هوش مصنوعی در زیرشاخه های تحقیقی خاص تر تقسیم می شود. به عنوان مثال: پردازش زبان طبیعی (NLP) می تواند تجربه نوشتن در کاربردهای مختلف را افزایش دهد. تقسیم بندی کلاسیک در NLP ترجمه ماشینی است که به عنوان ترجمه بین زبان ها شناخته می شود. الگوریتم های ترجمه ماشینی منجر به کاربردهای مختلفی شده اند که ساختار دستور زبان و همچنین اشتباهات املایی را در نظر می گیرند. علاوه بر این ، هنگامی که رایانه پیشنهاد تغییراتی در نویسنده یا ویرایشگر دارد، مجموعه ای از کلمات و واژگان مربوط به موضوع اصلی به طور خودکار به عنوان منبع اصلی استفاده می شود[۱].

هوش مصنوعی (AI) به عنوان هوش استفاده شده توسط ماشین ها رویکردی در یادگیری و استدلال انسان بوده است. اخیراً، یادگیری ماشین و داده کاوی به مرکز توجه و محبوب ترین موضوعات در بین جامعه تحقیقاتی تبدیل شده است. در طول سالها ، پایگاه داده با اهداف آماری جمع آوری شده است. منحنی های آماری می توانند گذشته و حال را برای پیش بینی رفتارهای آینده توصیف کنند. با این وجود ، در طول دهه های گذشته تنها تکنیک ها و الگوریتم های کلاسیک برای پردازش این داده ها استفاده شده است، در حالی که بهینه سازی آن الگوریتم ها می تواند منجر به خودآموزی موثر شود. تصمیم گیری بهتر می تواند بر اساس مقادیر موجود ، معیارهای متعدد و روشهای پیشرفته آماری اجرا شود. بنابراین ، یکی از مهمترین کاربردهای این بهینه سازی دارو است ، جایی که علائم ، علل و راه حل های پزشکی پایگاه داده های بزرگی را ایجاد می کنند که می توانند برای پیش بینی درمان های بهتر استفاده شوند. از آنجا که ML[1] طیف گسترده ای از تحقیقات را پوشش می دهد ، روشهای زیادی ایجاد شده است. خوشه بندی ، شبکه بایزین[۲] ، یادگیری عمیق[۳] و یادگیری تصمیمی درختی[۴] تنها بخشی از رویکردها است. بررسی زیر عمدتاً بر یادگیری عمیق، مفاهیم اساسی آن، کاربردهای گذشته و امروزی در زمینه های مختلف تمرکز دارد. علاوه بر این، چندین شکل به تصویر کشیده توسعه سریع تحقیقات یادگیری عمیق از طریق انتشارات در سال های اخیر در پایگاه های علمی ارائه شده است[۱].

مفهوم یادگیری عمیق (DL) برای اولین بار در سال ۲۰۰۶ به عنوان زمینه جدیدی از تحقیقات در یادگیری ماشین ظاهر شد. اولین بار به عنوان سلسله مراتبی در دانشگاه به عنوان رشته قابل یادگیری شناخته شد و معمولاً شامل بسیاری از زمینه های تحقیقاتی مربوط به تشخیص الگو بود. یادگیری عمیق به طور عمده دو عامل اساسی را در نظر می گیرد: پردازش غیرخطی در چند لایه یا مراحل و یادگیری کنترل شده یا بدون کنترل. پردازش غیرخطی در چندین لایه به الگوریتمی اشاره دارد که در آن لایه فعلی خروجی لایه قبلی را به عنوان ورودی می گیرد. سلسله مراتب در بین لایه ها ایجاد می شود تا اهمیت داده ها را مفید یا غیر مفید ارزیابی کند. از طرف دیگر، یادگیری نظارت شده و بدون نظارت با برچسب هدف کلاس داده مرتبط است، در دسترس بودن آن به معنای یک سیستم کنترل شده است، در حالی که عدم وجود آن به معنی یک سیستم بدون کنترل است[۲].

یادگیری عمیق به معنای تجزیه و تحلیل لایه انتزاعی و روش های سلسله ای و متوالی است. با این حال ، می توان آن را در بسیاری از برنامه های واقعی زندگی استفاده کرد. به عنوان مثال، در پردازش تصویر دیجیتال؛ رنگ آمیزی تصویر در مقیاس خاکستری تصویری که به صورت دستی توسط کاربرانی انجام می‌شد که مجبور بودند هر رنگ را براساس نظر خود انتخاب کنند. با استفاده از یک الگوریتم یادگیری عمیق، رنگ آمیزی به طور خودکار توسط کامپیوتر انجام می شود. به همین ترتیب ، صدا را می توان با استفاده از شبکه های عصبی تکرار شونده (RNN[5]) به عنوان بخشی از روش های یادگیری عمیق ، به یک فیلم بی صدا درامز اضافه کرد. یادگیری

---

[1] Machine Learning
[2] Bayesian Network
[3] Deep learning
[4] Decision Tree Learning
[5] Recurrent Neural Networks

عمیق را می توان به عنوان روشی برای بهبود نتایج و بهینه سازی زمان پردازش در چندین فرآیند محاسباتی درک کرد. در زمینه پردازش زبان طبیعی ، از روش های یادگیری عمیق برای ایجاد زیرنویس تصویر و ایجاد دست خط استفاده شده است. برنامه های زیر در پردازش تصویر دیجیتال خالص، دارو و بیومتریک دسته بندی شده اند[۳].

در دهه های اخیر ، با افزایش روزافزون نیاز به خودکارسازی سیستم های شناسایی و نظارت، تجزیه و تحلیل چهره انسان به بیشترین موضوعات مورد توجه در زمینه های دید رایانه‌ای مانند تشخیص چهره، طبقه بندی جنسیت و تشخیص حالت چهره تبدیل شده است. در این چارچوب ، تحقیقات در زمینه تکنیک های تخمین سن پیشرفت‌های خوبی داشته است[۴].

پردازش تصویر صورت یک چارچوب تحقیقاتی است که در دهه های اخیر توسط بسیاری از محققان مورد مطالعه قرار گرفته است. تشخیص چهره[۵]، حالت چهره[۶]، طبقه بندی جنسیتی[۷]، تشخیص چهره[۸]و استخراج ویژگی های صورت [۹] نتایج تحقیقات در این زمینه است که مقاله بستانفرد و همکاران[۸] دارای دیتابیس برای چهره‌های ایرانیان است. علیرغم این ، طبقه بندی گروه سنی صورت یکی از کاربردهاست که متناسب با اهمیت نظری و عملی به اندازه کافی مورد توجه قرار نگرفته است. این کاستی به سه دلیل شامل (I) افزایش تعداد کلاسها ، (II) عدم دقت حتی توسط ارزیابی‌های انسانی و (III) عدم وجود مجموعه داده های بزرگ و مناسب است[۱۰]. اخیراً ، سن انسان در بسیاری از برنامه ها مانند کنترل دسترسی مبتنی بر سن، تخمین سن در تحقیقات جنایی، بازاریابی هدفمند با انطباق با سن، مشتری الکترونیکی، مدیریت روابط، سرگرمی برای سیستم های توصیه شده، کنترل دسترسی امنیتی، بیومتریک، نشانه مهمی در مورد کاربر مورد نظر است. تبلیغات دقیق ، نظارت هوشمند ، کنترل دسترسی به اینترنت ، طبقه بندی قومیت، تجزیه و تحلیل رسانه های اجتماعی ، پزشکی ، روانشناسی و تعامل رایانه ای انسان (HCI) همه از کاربردهای این تکنولوژی هستند[۴].

در حقیقت، چالش جدید HCI اکنون نه تنها ایجاد تعامل طبیعی و ساده با رابطها بلکه انطباق با شناخت، توانایی های حسی و فرهنگی کاربر است. به منظور بهبود قابلیت اطمینان و سهولت، این رابطهای تعاملی جدید به صورت خودکار به یادگیری و افزایش دقت می پردازند و خود را با توانایی‌ها و ویژگی های طبیعی کاربر سازگار می‌کنند[۴]. در نتیجه، برای بهبود کیفیت تجربه ، رابط باید حداکثر اطلاعات مربوط به کاربر را جمع آوری کند. در این چارچوب ، تخمین سن به عنوان شاخص مهم هویت کاربر تلقی می شود. در واقع، سن؛ کانال اصلی اطلاعات غیرکلامی است که می تواند به محققان در این حوزه برای بهبود عملکرد و کارایی رابط های جدید کمک کند.

به طور کلی ، برآورد خودکار سن توسط ماشین در کاربردهایی مفید است که هدف آن تعیین سن یک فرد بدون شناسایی قبلی است. تخمین سن از تصاویر صورت کار پیچیده ای است: اولاً ، یک مشخصه سنی می تواند به عنوان یک کلاس فردی در نظر گرفته شود، بنابراین ما در مورد تعداد مهمی کلاس صحبت می کنیم. در مرحله دوم ، از نظر یادگیری، این مشکل بسیار دشوارتر از مشکلات تشخیص چهره و تخمین جنسیت است و طبقه بندی کننده های باینری که قبلا تعریف شده‌اند، ممکن است مستقیماً اعمال نشوند. همچنین، دو فرد مختلف می توانند متفاوت رشد کنند، زیرا روند پیری نه تنها توسط ژن های فرد تعیین می شود، بلکه توسط بسیاری از عوامل خارجی مانند سلامتی، سبک زندگی، محیط و شرایط تعیین می شود. مشکل دیگر این است که از بزرگسالی تا پیری، محسوس ترین تغییر پیر شدن پوست (تغییر در بافت) و تغییر در شکل ادامه چین و چروک است. سرانجام ، زنان و مردان می توانند سن متفاوت داشته باشند. به طور کلی، برآورد سن می تواند به عنوان یک طبقه بندی یا یک مسئله رگرسیون بسته به اینکه آیا گروه های سنی به عنوان کلاس در نظر گرفته می شوند یا سن به عنوان یک متغیر پیوسته در نظر گرفته شود[۱۱].

یکی از این سیستم ها توانایی استفاده به صورت جهانی و فارغ از نژاد را به عنوان ورودی داشت. مزیت دیگر نسبت به سیستم قبلی این است که روند آموزش در ELM واقعاً سریع است. از طریق آزمایشات، مشخص شد که ELM نه تنها یک شروع کننده خوب برای درک وضعیت موجود در حوزه کاری بلکه یک بهینه ساز خوب است که در مدت زمان کوتاهی تعمیم منطقی و عملکرد قابل اعتماد را ارائه می دهد[۱۱].

از دهه ۱۹۷۰ ، محققان هوش مصنوعی در تلاشند تا اسرار نحوه کار مغز انسان را کشف کنند به این امید که روزی امکان تکثیر واقعی آن و انتقال آن به ماشین آلاتی فراهم شود که به اندازه انسان هوشمند و خودمختار شوند. این سناریو به محققان اجازه می دهد تا ماشین های هوشمندتر را طراحی کنند[۱۲].

در حقیقت ، قبل از سال ۲۰۰۶ ، ما نمی دانستیم که چگونه یک معماری عمیق را آموزش دهیم: بهینه سازی تکراری به سمت حداقل محلی با کیفیت پایین همگرا است. شبکه های عصبی با بیش از دو لایه پنهان که به طور تصادفی ایجاد شده اند، نتایج بدتری نسبت به شبکه های معمولی می دهند. شبکه های عصبی عمیق (DNN) چیزی بیش از شبکه های چند لایه با معماری کلاسیک نیستند، آنها چندین لایه پنهان دارند. در واقع ، حتی اگر قضیه های تقریب در اواخر دهه ۱۹۸۰ ادعا کنند که یک لایه پنهان منفرد برای تقریب هر عملکرد به اندازه کافی منظم کافی است، چیزی برای احراز اولویت اجرای یادگیری از طریق انتشار در شبکه هایی با چندین لایه پنهان وجود ندارد[۱۳].

تحقیقات مربوط به سن از دو نقطه نظر بررسی شده است: طبقه بندی سن صورت[۱۴] و شبیه سازی سنی صورت[۱۵]. دسته اول شامل آموزش و مجموعه های آزمایشی برای تولید مدلی است که می تواند سن چهره مشخصی را در تصویر طبقه بندی کند. دسته دوم اثرات پیری را شبیه سازی می کند، که به ایجاد تصاویر پیشرفته یا گذری از سن فرد کمک می کند.

همانطور که گفته شد، طبقه بندی گروه سنی از نظر تئوری و عملی مهم است. استخراج ویژگی های مناسب برای شناسایی . انتخاب صحیح مانند پارامترهای رشد صورت و محیط چین و چروک می تواند در کاربردهایی مانند بیان چهره و تشخیص چهره مبتنی بر ویژگی ها به کار رود[۱۶]. علاوه بر این ، تعیین پارامترهای پیری که در ظاهر چهره تأثیر می گذارند، طیف گسترده ای از برنامه ها در نرم افزار تجاری و امنیتی مانند اجرای قوانین محدودیت سنی هستند. اگرچه آثار قبلی سهم خود را در این زمینه داشته اند، اما به نظر می رسد که آنها دو مشکل عمده دارند:

۱- مجموعه داده های آنها در دسترس نیستند یا اطلاعات دقیق سنی ندارند.

۲- الگوریتم های طبقه بندی افراد را در گروه های سنی طبقه بندی کرده اند که فاقد نظم مناسب هستند یا فقط بر روی سنین خاصی تمرکز دارند.

اهمیت فوق الذکر ما را بر آن می دارد که الگوریتمی جدید برای طبقه بندی گروه سنی ارائه دهیم که عاری از معایب قبلی باشد. برای مثال روش پیشنهادی در مقاله دهشیبی و بستانفرد[۸] ارائه شده تصاویر صورت را به یکی از چهار گروه سنی منظم طبقه بندی می کند که شامل چهار مرحله اصلی است که تقریبا برای کل مطالعات در این زمینه مشترک است[۱۰].

(۱) پیش پردازش: در این مرحله ، ناحیه صورت به صورت دستی برش داده می شود تا ضرورت شرایط لیوان را از بین ببرد. سپس ، از فرایندی برای بهبود تصاویر برای مراحل بعدی استفاده می شود.

(۲) مکان یابی ویژگی های صورت: در این مرحله ، یک روش طرح ریزی جدید برای مکان یابی ویژگی های صورت ارائه شده است. پس از تعیین محل صورت (چشم ، بینی ، دهان ، چانه ، کناره های صورت و بالای سر) ، مجموعه ای از نسبت ها برای تشخیص چهره های نابالغ از بزرگسالان تعریف می شود.

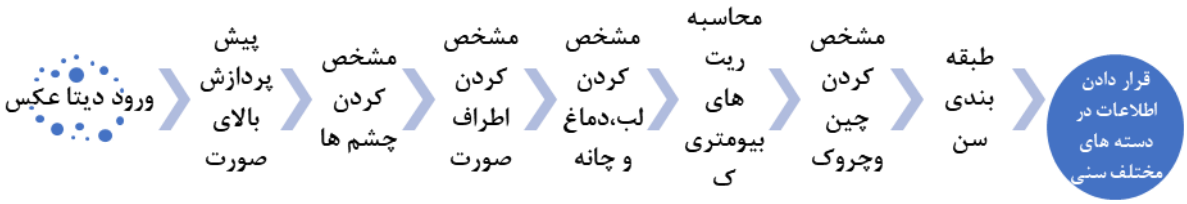
(۳) استخراج محیط چین و چروک: چین و چروک اثر قابل توجهی در ظاهر صورت دارد و طبقه بندی سن را آسان تر می کند. در این مرحله ، بر اساس پارامترهای سنجش صورت، چین و چروک های پیشانی، گوشه چشم و گونه ها دریافت می شوند. سپس ، تراکم چین و چروک در هر محیط محاسبه می شود.

(۴) طبقه بندی گروه سنی: در این مرحله ، یک بسط ANN فوروارد[6] ساخته می شود ، که از ویژگی های محاسبه شده در مراحل قبلی استفاده می کند. این شبکه با داده های ۲۹۸ تصویر ارتقا یافته و با ۲۰۰ تصویر معمولی شده است[۱۰].

این مراحل در شکل(۱- ۱) خلاصه شده است.

---

[6] Forward



شکل(۱- ۱): نمودار بلوک سیستم تشخیص سن

۲- مروری بر تحقیقات گذشته

کوون و لوبو [۱۴] برای اولین بار روی مسئله طبقه بندی سنی کار کردند. آنها تصاویر صورت در مقیاس خاکستری را در سه گروه سنی دسته بندی کردند. میزان شناسایی کودکان زیر ۶۸٪ است. به دلیل استفاده از توابع انرژی و ماری، روش آنها برای پردازش در زمان واقعی کارآمد نیست. روش آنها در یک پایگاه داده نسبتاً کوچک (۴۷ تصویر با وضوح بالا) که پارامتر های کمی برای تجزیه و تحلیل چین و چروک داشت، آزمایش شد. این پایگاه داده کوچک ارزیابی صحیح نتایج را دشوار کرده بود.

هایاشی و همکاران [۴] سنین طبقه بندی شده در ده سال. ابتدا مناطق پوست را از تصاویر صورت استخراج می کنند. سپس ، برای افزایش چین و چروک ها ، یکسان سازی هیستوگرام انجام می شود و برای استخراج چین و چروک صورت ، از انتقال (Hough) استفاده می شود. سرانجام، یک جدول جستجو برای طبقه بندی سن و جنسیت ایجاد می شود. اگرچه آنها دامنه سن منظمی را برای طبقه بندی تعریف کردند، اما سیستم تخمین سن آنها در ۳۰۰ تصویر فقط ۲۷٪ دقیق بود. با وجود این واقعیت که تغییرات عمده در صورت تا ۱۵ سالگی رخ می دهد، گروه سنی "۱۵" زیر "در روش خود نادیده گرفته بودند[۷].

قبل از اینکه یادگیری عمیق رسماً به عنوان یک رویکرد تحقیقاتی جدید ظاهر شود، برخی از کاربردها در زمینه شناخت الگو از طریق پردازش لایه‌ای انجام شده بود. در سال ۲۰۰۳ ، یک مثال جالب با استفاده از فیلتر کردن نقاط و بسط بایزین ایجاد شد. کاربرد این نرم افزار به این شکل بود که می توانست با مشاهده فقط یک تصویر صورت نیمه بریده شده، چهره شخص را تشخیص دهد، بنابراین یک کامپیوتر می تواند تصویر صورت را از یک صورت بریده شده بازسازی کند[۱۷].

بعداً در سال ۲۰۰۶، الگوریتم گرید(حریص) و سلسله مراتب در یک برنامه با قابلیت پردازش ارقام دست نویس ترکیب شدند. تحقیقات اخیر از یادگیری عمیق به عنوان اصلی ترین ابزار برای پردازش تصویر دیجیتال استفاده کرده است. به عنوان مثال ، استفاده از شبکه عصبی (CNN[7]) برای تشخیص عنبیه می تواند موثرتر از استفاده از سنسورهای عنبیه مرسوم باشد. اثربخشی CNN می تواند تا دقت ۹۹.۳۵ ٪ برسد. امروزه تشخیص مکان تلفن همراه به کاربر اجازه می دهد تا یک آدرس مشخص را بر اساس یک تصویر مشخص کند. یک الگوریتم تحت نظارت حفظ معانی هش (Deep Hashing) (SSPDH) در مقایسه با Visual Hash Bit (VHB) و Space - Saliency Finger Selecting (SSFS) بهبود قابل توجهی را نشان داده است. دقت SSPDH حتی ۷۰٪ کارآمدتر است[۱۸].

سرانجام، یکی دیگر از کاربردهای قابل توجه در پردازش تصویر دیجیتال، با استفاده از روش یادگیری عمیق، تشخیص چهره است. گوگل ، فیس بوک و مایکروسافت دارای مدل های منحصر به فرد یادگیری عمیق چهره هستند. اخیراً شناسایی با استفاده از تصویر صورت با تعیین سن و جنسیت به عنوان پارامترهای اولیه به شناسایی خودکار تغییر یافته است. به عنوان مثال Sighthound Inc یک الگوریتم شبکه عصبی کانولوشن را آزمایش کرد که قادر است نه تنها سن و جنسیت ، بلکه حتی احساسات را تشخیص دهد. علاوه بر این ، یک سیستم قوی برای تعیین دقیق سن و جنسیت شخص از یک تصویر با استفاده از یک معماری یادگیری عمیق چند کاره ایجاد کرد[۱۹, ۲۰].

[7] Convolutional Neural Networks

در سال ۲۰۰۲ ، لانیت[۲۱] از مدلهای ظاهر فعال (AAM[8]) برای نمایش سن استفاده کرد ، طبق نتایج آزمایش، بعضی از قسمتهای صورت و بخصوص نواحی اطراف چشم منبع قابل توجهی از اطلاعات را برای وظیفه تخمین سن خودکار ارائه می دهند. در سال ۲۰۰۷ گنگ و همکاران [۲۲] از طبقه بندی سنی MORPH و FG-Net برای تجزیه و تحلیل خطی تفکیکی (LDA[9]) و KNN (k-نزدیک ترین همسایگی) استفاده کردند. روشی به نام AGES که در این کار ارائه شده است از زیر فضای نماینده برای مدل سازی الگوی پیری استفاده می کند. در سال ۲۰۱۵ پونتس جی کی وهمکاران [۲۳] همچنین از مجموعه داده های AAM، MORPH و FG-Net برای پیش بینی سنین چهار گروه بین ۱ تا ۷۷ سال استفاده می کنند. آنها چارچوبی از AAM ، الگوهای باینری محلی (LBP[10]) و موجک گابور را برای استخراج ویژگی های بسیار تبعیض آمیز ارائه دادند. یک سال بعد دات تین و همکاران[۲۴] از الگوی باینری محلی چند سطحی (WMLBP[11]) مبتنی بر سیستم منطقی فازی برای استخراج ویژگی ها و ماشین بردار پشتیبانی (SVM) برای طبقه بندی سنین ۶۳۷ شرکت کننده از ۱۸ تا ۹۳ سال استفاده کردند. پیر بیشترین عیب AAM در کارهای قبلی تأثیر از بین بردن بینی در نواحی مو، زمینه و غیر یکنواخت بر عملکرد آنها بود. آنها در مقاله خود روشی جدید مبتنی بر WMLBP پیشنهاد کردند تا از این معایب معمول جلوگیری شود. از سال ۲۰۰۵ ، عبارت DL به یادگیری ماشین وارد شد. DL مجموعه ای از یادگیری خودکار روش ها است که سعی در مدل سازی با سطح بالایی از انتزاع داده ها دارد. بسیاری از محققان از عملکرد DL در تجزیه و تحلیل صورت برای پس زدن و تخمین سن کاربر استفاده کردند[۲۵]. ایوان هورتا و همکاران [۲۶]در کار خود چارچوبی را بر اساس دو مرحله ارائه دادند:

۱- در مرحله اول آنها HOG[12] ، LBP و (SURF[13]) را برای استخراج ویژگی ها استفاده کردند.

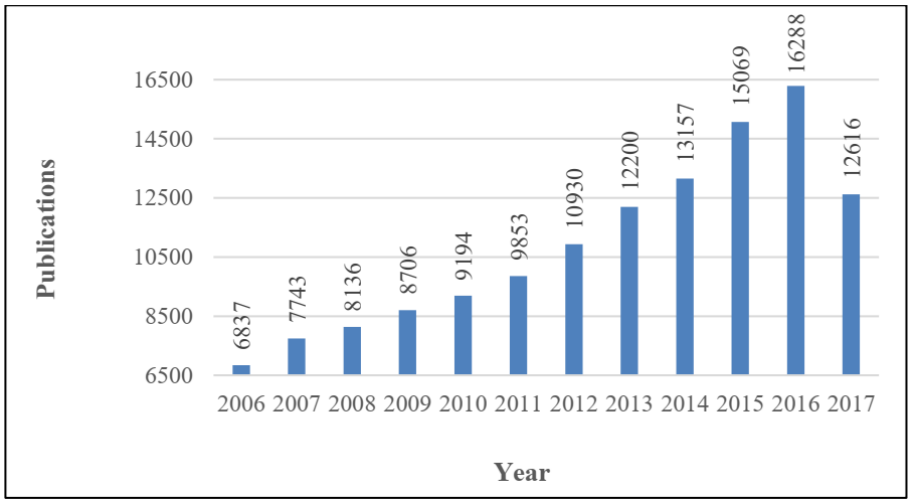
۲- مرحله دوم آنها از شبکه عصبی (CNN) برای ترکیب انواع لایه ها (لایه های کانولوشن و استخر) برای استخراج ویژگی های مهم و لایه های دیگر (عصبی محلی و کاملاً متصل) برای ارتباط دادن اطلاعات جهانی و یادگیری داده ها استفاده کردند.

یوان دونگ و همکاران [۲۷] برای استخراج ویژگی ها برای پیش بینی سن تصویر صورت، از شبکه عصبی عمیق انحرافی (DCNN[14]) استفاده کردند. آنها از استراتژی یادگیری انتقال برای برچسب زدن تصویر چهره استفاده کردند. پل رودروگز و همکاران [۲۵]روش جدیدی برای شناخت سن و جنسیت ارائه دادند. هدف افزایش عملکرد CNN و مقاومت در برابر تغییر شکل تصویر بود. آنها خط لوله ای از یک شبکه توجه را ارائه دادند که آموزنده ترین نقاط اتصال را در تصویر با وضوح پایین تخمین می زند. نتایج آزمایش استحکام این روش در بهم ریختگی و تغییر شکل را نشان داده است. هائو لیو و همکاران [۲۸] برای تخمین سن کاربر از معماری (ODFL[15]) استفاده کردند. آنها از DCNN برای یادگیری ویژگی ها مستقیماً از پیکسل های تصویر استفاده کردند. نتایج آزمایش تأثیر این روش را نشان داد. شیائو لانگ وانگ و همکاران[۲۹] چارچوب جدیدی را برای تخمین سن مبتنی بر تصویر ارائه دادند ، آنها از یک ویژگی پیری یادگیری بدون کنترل از طریق کدگذاری پراکنده کانولوشن استفاده کردند. برای بهبود استحکام این روش ، از استاندارد استخراج مشتق بر روی نقشه ویژگی ها استفاده شد. نتایج نمایان ای موضوع بودند که خروجی تولید شده توسط کدگذاری پراکنده کانولوشن، نتایج تجربی نتایج بهتری نسبت به طرح های سنتی نشان داد. زکریا قواقنه و همکاران [۳۰] الگوریتمی جدید برای طبقه بندی سن و جنسیت انسان ارائه دادند که تصاویر گفتاری و چهره را با استفاده از CNN برای استخراج و طبقه بندی ویژگی ها بررسی می کند.

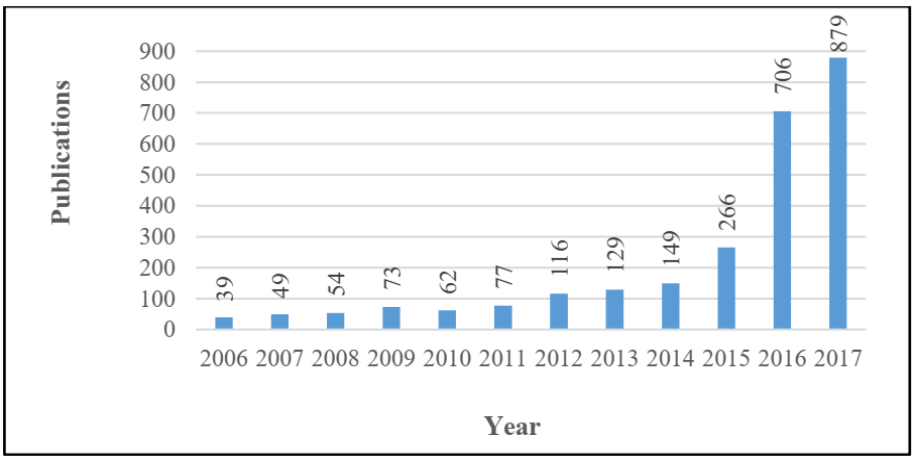
---

[8] Active Appearance Models
[9] Linear Discriminant Analysis
[10] Local Binary Patterns
[11] Weighted Multi-level Local Binary Pattern
[12] Histogram of Oriented Gradients
[13] Speeded Up Robust Features
[14] Deep Convolutional NeuralNetwork
[15] Ordinal Deep Features Learning

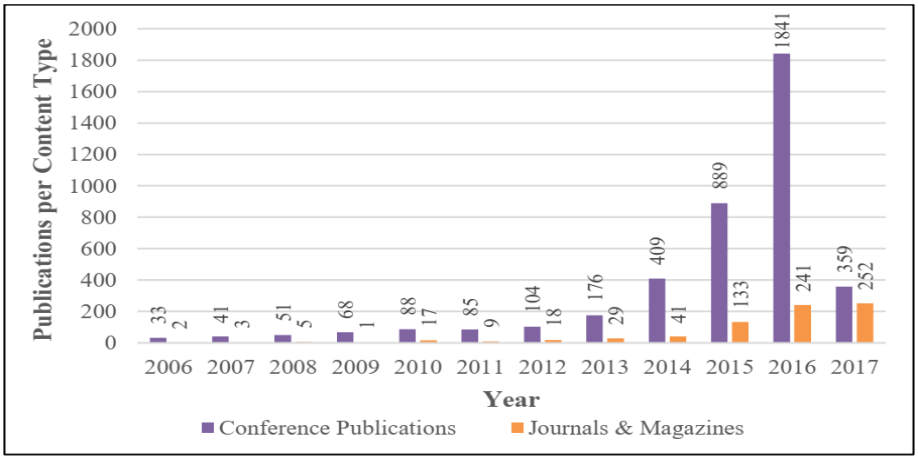
۲-۱- تجزیه و تحلیل انتشار در هر سال



شکل(۲- ۱) : رشد تعداد نشریات در Deep Learning ، پایگاه داده Sciencedirect (ژانویه ۲۰۰۶ - ژوئن ۲۰۱۷)[۱]



شکل(۲- ۲): رشد تعداد نشریات در Deep Learning ، پایگاه داده Springer (ژانویه ۲۰۰۶ - ژوئن ۲۰۱۷) [۱]



شکل(۲- ۳): رشد انتشارات در Deep Learning از پایگاه داده IEEE. (ژانویه ۲۰۰۶ تا ژوئن ۲۰۱۷) [۱]

همان طور که از نمودار شکل(۲- ۱) تا شکل(۲- ۳) مشخص است رشد مقالات بر اساس یادگیری عمیق چشم گیر بوده‌است.

۳- روش کار و مکانیسم عملکرد روش انتخابی

در حالی که شناخت بیشتر تغییرات صورت ، مانند هویت ، بیان و جنسیت ، به طور گسترده مورد بررسی قرار گرفته است ، تخمین سن اتوماتیک به ندرت مورد بررسی قرار گرفته است. برخلاف سایر تغییرات صورت ، تغییرات پیری چندین ویژگی منحصر به فرد دارد که تخمین سن را به عنوان یک کار چالش برانگیز معرفی می کند. ایده اصلی این است که الگوی پیری ، که به عنوان دنباله ای از تصاویر چهره یک فرد خاص به ترتیب زمان مرتب شده است، با ساخت یک فضای خالی نماینده ، مدل سازی شود. الگوی پیری مناسب برای یک تصویر صورت که قبلاً دیده نشده است، توسط توزیع در فضای زیر تعیین می شود که می تواند با حداقل خطای بازسازی ، تصویر صورت را بازسازی کند، در حالی که موقعیت تصویر صورت در آن الگوی پیری، سن آن را نشان می دهد. در آزمایشات ، سن AGES و انواع آن با روش های محدود برآورد سن موجود (WAS و AAS) و برخی از روش های طبقه بندی کاملاً تثبیت شده (kNN، BP، C4.5 و SVM) مقایسه می شود. علاوه بر این ، مقایسه ای با توانایی درک انسان بر روی سن انجام می شود. جالب است بدانید که عملکرد AGES نه تنها به طور قابل توجهی بهتر از تمام الگوریتم های دیگر است، بلکه قابل مقایسه با ناظران انسانی است.

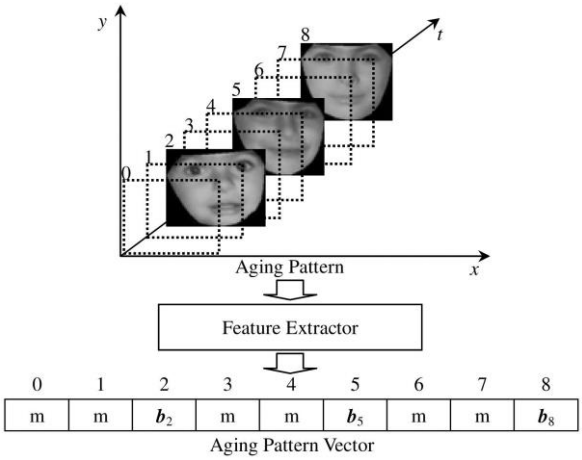
برآورد بسیار دشوار است. این مشکل اکنون به دلیل انتشار عمومی پایگاه داده پیری FG-NET تا حدی برطرف شده است. با این وجود ، هر موضوع در این پایگاه داده فقط در چند سنی دارای تصاویر چهره است ، به عنوان مثال ، مجموعه داده ها از نظر الگوهای پیری بسیار ناقص است. خوشبختانه ، یک پایگاه داده "کامل" پیر شدن صورت غیر ضروری است زیرا انسان همچنین یاد می گیرد سن صورت را از الگوی پیری ناقص درک کند. بنابراین ، الگوریتم یادگیری اعمال شده در الگوهای پیری باید بتواند داده های کاملاً ناقص را کنترل کند. دوم ، نقشه برداری از نمونه ها (تصاویر چهره) به برچسب های کلاس (سن) منحصر به فرد نیست، اما به طور پیچیده ای به عوامل شخصی بستگی دارد. بنابراین ، انتخاب یک الگوی پیری مناسب برای یک چهره خاص ، یک مرحله مهم در تخمین سن می شود. سوم ، مجموعه برچسب های کلاس (سن) یک مجموعه کاملاً مرتب شده است. هر سن دارای یک رتبه منحصر به فرد در توالی زمانی است. پس از انتخاب الگوی پیری مناسب برای یک تصویر صورت خاص ، "موقعیت" صورت در آن الگوی پیری به طور منحصر به فردی سن آن را تعیین می کند. مطابق با دو ویژگی آخر ، تخمین سن اتوماتیک باید حداقل شامل دو مرحله اصلی باشد: مرحله ۱ تعیین الگوی پیری مناسب برای یک صورت خاص و مرحله ۲ یافتن موقعیت صورت در آن الگوی پیری است.

روش های مبتنی بر عملکرد پیری ، تخمین سن را به عنوان یک مشکل طبقه بندی معمول در نظر می گیرند: داده ها تصاویر چهره هستند ، هدف برچسب های سن آنها است. با توجه به ویژگی شخصی شده ، هر تصویر باید دارای یک برچسب دیگر به غیر از سن برچسب سن آن باشد ، یعنی شناسه هویت شخصی آن. اگر قرار است با تکنیک های نظارتی مانند LDA مسئله حل شود ، پس الگوریتم

باید با داده های چند پارامتری سروکار داشته باشد ، که تنها مسئله یادگیری ماشین است. از طرف دیگر ، اگر همه این برچسب ها بتوانند در نمایش داده ها ادغام شوند ، مسئله چند برچسب می تواند به یک مسئله یادگیری بدون نظارت تبدیل شود. بنابراین ، ما نمایشی از داده ها را به نام الگوی پیری پیشنهاد می دهیم که اساس آن AGES است. یک تعریف رسمی به شرح زیر ارائه می شود[۳۱]:

تعریف ۱. الگوی پیری دنباله ای از تصاویر چهره شخصی است که به ترتیب زمانی مرتب شده اند. کلمات کلیدی "شخصی" و "زمان" هستند. تمام تصاویر صورت در الگوی پیری باید از یک شخص تهیه شده و طبق زمان آنها طبقه بندی شوند. الگوی پیری نشان داده شده در شکل ۱ را به عنوان مثال در نظر بگیرید. در امتداد محور t ، به هر سن (۰-۸) یک موقعیت اختصاص داده شده است.

اگر تصاویر چهره برای سنین خاص (۲ ، ۵ و ۸) موجود باشد ، آنها در موقعیت های مربوطه پر می شوند. در غیر این صورت ، موقعیت ها خالی می مانند. اگر تمام موقعیت ها پر شود ، الگوی پیری را الگوی پیری کامل می نامند. در غیر این صورت ، این یک الگوی پیری ناقص نامیده می شود. قبل از اینکه الگوی پیری بیشتر پردازش شود ، تصاویر صورت موجود در آن ابتدا به بردارهای ویژگی تبدیل می شوند. بدیهی است که پیری فرآیندی است که هم به شکل و هم به فرم صورت مربوط می شود. بنابراین ، از مدل Appearance به عنوان استخراج کننده ویژگی استفاده می شود ، که مزیت اصلی آن این است که ویژگی استخراج شده ، هم شکل و هم شدت خصوصیات تصاویر چهره را با هم ترکیب می کند. در شکل۱ نمونه ای از بردار ی از الگوی پیری آورده شده است ، جایی که b2، b5 و b8 به ترتیب بردارهای مشخصه تصاویر چهره را در سنین ۲ ، ۵ و ۸ سال نشان می دهند[۳۱].



شکل(۳- ۱): برداری از الگوی پیری. سن (۰-۸) در بالا سمت چپ به موقعیت های مربوطه و بالاتر از بردارهای ویژگی مربوطه مشخص می شوند. قسمتهای از دست رفته در بردار الگوی پیری با "m" مشخص شده اند[۳۱].

با نشان دادن الگوهای پیری به این روش ، دو پارامتر age1 و id1 به طور طبیعی و بدون هیچ پیش فرضی در داده ها ادغام می شوند. هر الگوی پیری حاوی یک شناسه است، هر سن در الگوی پیری قرار می گیرد و موقعیت با توجه به زمان مرتب می شود. در نتیجه ، از خصوصیات شخصی و زمانی می توان به خوبی استفاده کرد. تا زمانی که از الگوی پیری به خوبی نمونه برداری شود ، می توان مدل مناسبی از الگوی پیری را آموخت و روند یادگیری بدون نظارت است. با این حال ، این دو چالش دیگر را به همراه دارد: ۱) در طول آموزش ، الگوریتم یادگیری اعمال شده در الگوهای پیری باید بتواند نمونه های آموزش بسیار ناقص را مدیریت کند، و ۲) در طول تخمین سن بر روی داده های آزمون ، مناسب ترین الگوی پیری ، و همچنین مناسب ترین موقعیت در آن الگوی پیری ، باید برای تصویر چهره ناشناخته انتخاب شود. بخش بعدی عمدتا به این دو مشکل می پردازد.

۴- نتیجه گیری

افراد می توانند انواع مختلفی از اطلاعات مفید را از تصویر چهره مانند هویت، جنسیت، بیان، سن تقریبی و غیره استخراج کنند. استخراج خودکار اکثر اطلاعات به طور گسترده مورد بررسی قرار گرفته است. با این حال، تاکنون، در مورد تخمین سن اتوماتیک، کار نسبتاً کمی وجود دارد، علیرغم این واقعیت که این مسئله به تنهایی یک مسئله جالب است و همچنین یک مسئله فرعی چالش برانگیز در کارهایی مانند تشخیص چهره است. رفتار و اولویت افراد در سنین مختلف متفاوت است، این نشان دهنده کاربردهای بالقوه گسترده ای برای تخمین سن اتوماتیک است. در مقایسه با سایر تغییرات صورت ، اثرات پیری سه ویژگی منحصر به فرد را نشان می دهد:

۱. پیشرفت پیری غیرقابل کنترل است. هیچ کس به میل خود نمی تواند پیری را پیشبرد یا به تأخیر بیندازد. روش پیری کند و غیر قابل برگشت است.

۲. الگوهای پیری شخصی. افراد مختلف از طرق مختلف پیر می شوند. الگوی پیری هر فرد توسط ژن های وی و همچنین بسیاری از عوامل خارجی مانند سلامتی ، سبک زندگی ، شرایط آب و هوایی و غیره تعیین می شود.

۳. الگوهای پیری داده های زمانی هستند. پیشرفت پیری باید از نظم زمانی پیروی کند. وضعیت چهره در یک سن خاص بر تمام چهره های مسن تأثیر می گذارد، اما روی جوان ترها تأثیر نخواهد گذاشت.

هر یک از این ویژگی ها به مشکلات برآورد سن خودکار کمک می کند. اول ، از آنجا که افراد نمی توانند آزادانه تنوع پیری را کنترل کنند ،دوم جمع آوری اطلاعات کافی آموزش برای سن را انجام می دهد.

منابع:

۱. Vargas, R., A. Mosavi, and R. Ruiz, *Deep learning: a review*. Advances in Intelligent Systems and Computing, 2017.

۲. Mosavi, A. and A.R. Varkonyi-Koczy, *Integration of machine learning and optimization for robot learning*, in *Recent Global Research and Education: Technological Challenges*. 2017, Springer. p. 349-355.

۳. Mohamed, A.-r., G. Dahl, and G. Hinton. *Deep belief networks for phone recognition*. in *Nips workshop on deep learning for speech recognition and related applications*. 2009. Vancouver, Canada.

٤. Zaghbani, S., N. Boujneh, and M.S. Bouhlel, *Age estimation using deep learning*. Computers & Electrical Engineering, 2018. **68**: p. 337-347.

٥. Zhao, W., et al., *Face recognition: A literature survey*. ACM computing surveys (CSUR), 2003. **35**(4): p. 399.٤٥٨-

٦. Kanade, T., J.F. Cohn, and Y. Tian. *Comprehensive database for facial expression analysis*. in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*. 2000. IEEE.

۷. Hayashi, J.-i., et al. *Age and gender estimation from facial image processing*. in *Proceedings of the 41st SICE Annual Conference. SICE 2002*. 2002. IEEE.

۸. Bastanfard, A., M.A. Nik, and M.M. Dehshibi. *Iranian face database with age, pose and expression*. in *2007 International Conference on Machine Vision*. 2007. IEEE.

۹. Viola, P. and M.J. Jones, *Robust real-time face detection*. International journal of computer vision, 2004. **57**(2): p. 137-154.

۱۰. Dehshibi, M.M. and A. Bastanfard, *A new algorithm for age recognition from facial images*. Signal Processing, 2010. **90**(8): p. 2431-2444.

۱۱. Sai, P.-K., J.-G. Wang, and E.-K. Teoh, *Facial age range estimation with extreme learning machines*. Neurocomputing, 2015. **149**: p. 364-372.

.۱۲ Zaghbani, S., N. Boujneh, and M. Bouhlel. *Real time hand gesture recognition using features extraction*. in *International conference ICMV (International Conference on Machine Vision)*. 2017.

.۱۳ Tian, Q. and S. Chen, *Cross-heterogeneous-database age estimation through correlation representation learning.* Neurocomputing, 2017. **238**: p. 286-295.

.۱٤ Kwon, Y.H. and N. da Vitoria Lobo. *Locating facial features for age classification*. in *Intelligent robots and computer vision XII: Algorithms and techniques*. 1993. International Society for Optics and Photonics.

.۱٥ Burt ,D.M. and D.I. Perrett, *Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and colour information.* Proceedings of the Royal Society of London. Series B: Biological Sciences, 1995. **259**(1355): p. 137-143.

.۱٦ Gökberk, B ,.et al., *Learning the best subset of local features for face recognition.* Pattern Recognition, 2007. **40**(5): p. 1520-1532.

.۱۷ Mosavi, A. and A. Vaezipour, *Reactive search optimization; application to multiobjective optimization problems.* Applied Mathematics, 2012. **3**(10A): p. 1572-1582.

.۱۸ Bannister, A., *Biometrics and AI: how face sentinel evolves 13 times faster thanks to deep learning.* September 13, 2016.

.۱۹ Vaezipour, A., A. Mosavi, and U. Seigerroth. *Machine learning integrated optimization for decision making*. in *The 26th EURO-INFORMS Conference, July 1-4 2013, Rome, Italy*. 2013.

.۲۰ Moor, J., *The Turing test: the elusive standard of artificial intelligence*. Vol. 30. 2003: Springer Science & Business Media.

.۲۱ Lanitis, A. *On the significance of different facial parts for automatic age estimation*. in *2002 14th International Conference on Digital Signal Processing Proceedings. DSP 2002 (Cat. No. 02TH8628)*. 2002. IEEE.

.۲۲ Geng, X. and Z.-h. Zhou, *S. Member, K. Smith-miles, and S. Member,"Automatic Age Estimation Based on Facial Aging Patterns,".* IEEE Trans. Pattern Analysis and Machine Intelligence, 2007. **29**(12): p. 2234-2240.

.۲۳ Pontes, J.K., et al., *A flexible hierarchical approach for facial age estimation based on multiple features.* Pattern Recognition, 2016. **54**: p. 34-51.

.۲٤ Nguyen, D.T. and K.R. Park, *Enhanced age estimation by considering the areas of non-skin and the non-uniform illumination of visible light camera sensor.* Expert Systems with Applications, 2016. **66**: p. 302-322.

.۲٥ Rodríguez, P., et al., *Age and gender recognition in the wild with deep attention.* Pattern Recognition, 2017. **72**: p. 563-571.

.۲٦ Huerta, I., et al., *A deep analysis on age estimation.* Pattern Recognition Letters, 2015. **68**: p. 239-249.

.۲۷ Dong, Y., Y. Liu, and S. Lian, *Automatic age estimation based on deep learning algorithm.* Neurocomputing, 2016. **187**: p. 4-10.

.۲۸ Liua, H., et al., *Age and gender recognition in the wild with deep attention.* Pattern Recognit, 2017(October).

.۲۹ Wang, X., et al. *A study of convolutional sparse feature learning for human age estimate*. in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*. 2017. IEEE.

.۳۰ Qawaqneh, Z., A.A. Mallouh, and B.D. Barkana, *Age and gender classification from speech and face images by jointly fine-tuned deep neural networks.* Expert Systems with Applications, 2017. **85**: p. 76-86.

.۳۱ Geng, X., Z.-H. Zhou, and K. Smith-Miles, *Automatic age estimation based on facial aging patterns.* IEEE Transactions on pattern analysis and machine intelligence, 2007. **29**(12): p. 2234-2240.