

دریافت مقاله: 92/12/24

پذیرش مقاله: 93/4/15

ارائه روش جدید دو مرحله‌ای برای تخمین هوشمند سن افراد توسط تصاویر چهره

سمیه نوازی^{1*}، عباس رسولزادگان²

¹ کارشناس ارشد کامپیوتر، دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران، ایران

navazi.s84@gmail.com

² استادیار، دانشکده مهندسی، گروه مهندسی کامپیوتر، دانشگاه فردوسی مشهد، مشهد، ایران

rasoolzadegan@um.ac.ir

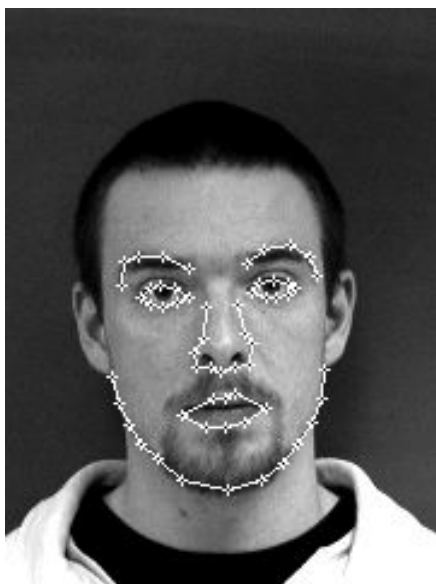
چکیده: تخمین سن هوشمند افراد از روی تصاویر چهره، به دلیل نقش کلیدی در بسیاری از موارد کاربرد، نظیر تحلیل رفتار مشتریان در یک سیستم هوش تجاری، به یکی از موضوعات تحقیقاتی مهم در حوزه بینایی ماشین و تشخیص الگو تبدیل شده است. تخمین سن، فرایندی است که طی آن، تصویر چهره فرد توسط کامپیوتر، پردازش شده و سن او براساس معیار سال، برآورد می‌شود. فرایند تخمین سن از دو بخش اصلی تشکیل می‌شود: 1. توصیف چهره؛ 2. یادگیری تابع تخمین سن. تاکنون، روش‌های متعددی جهت بهبود یک یا هر دو بخش فرایند تخمین سن، ارائه شده است. در این مقاله، یک روش جدید برای بهبود بخش یادگیری تابع تخمین سن ارائه می‌شود. در روش پیشنهادی، توصیف چهره در دو مرحله و با استفاده از روش‌های BIF (Bio-Inspired Features) و PLS (Partial Least Squares) انجام می‌شود. در ادامه، سن فرد توسط یک روش دو مرحله‌ای جدید که بر مبنای شبکه عصبی (Multi-Layer Perception) MLP پیاده‌سازی شده است، تخمین زده می‌شود. به منظور ارزیابی، روش پیشنهادی بر روی دو پایگاه داده معروف FG-NET و MORPH2 اجرا شده است. نتایج حاکی از برتری روش پیشنهادی (54% دقت بیشتر بر روی پایگاه داده FG-NET و 78% دقت بالاتر بر روی پایگاه داده MORPH2) بر سایر روش‌های موجود در فرایند تخمین سن است.

واژه‌های کلیدی: تخمین سن، تشخیص الگو، شبکه عصبی، PLS، BIF.

1. مقدمه

ویژگی‌هایی که بیشترین قدرت تفکیک‌پذیری را دارند، انجام می‌شود. تاکنون، روش‌های مختلفی برای انجام این کار معرفی شده است [3، 4، 5، 6 و 7]. به‌طور کلی، این ویژگی‌ها به سه دسته تقسیم‌بندی می‌شوند: 1. ویژگی‌های محلی مانند عمق، جهت و اندازه چروک‌های چهره [4]؛ 2. ویژگی‌های عمومی که به ارائه مدل کلی از شکل چهره یا سطوح خاکستری چهره (شامل تغییرات شدت روشنایی در تصویر از سفید به سیاه) می‌پردازند. (همان‌طور که در شکل 1 دیده می‌شود، خطوط جداکننده صورت، بینی، چشم‌ها و لب‌ها نمایانگر یک مدل کلی از چهره‌اند)؛ 3. ویژگی‌های ترکیبی [3].

ویژگی‌های محلی عمدتاً به‌منظور دسته‌بندی گروه سنی مورد استفاده قرار می‌گیرند؛ درحالی‌که ویژگی‌های عمومی در تخمین سن دقیق افراد، کارآمدترند. از آنجاکه هر یک از این دسته ویژگی‌ها، منعکس‌کننده خصوصیات خاصی از چهره هستند، به‌کارگیری ترکیبی از این روش‌ها می‌تواند در بهبود فرایند تخمین سن مؤثر باشد [3].



شکل (1): مدل کلی از شکل چهره

در مرحله یادگیری تابع تخمین سن نیز، رویکردهای متفاوتی وجود دارد. تمامی روش‌های تخمین سن، سن را به‌عنوان یک کلاس در نظر می‌گیرند و سپس آن را به یکی از

امروزه تخمین سن هوشمند و خودکار افراد از روی تصاویر چهره، به یکی از موضوعات تحقیقاتی مهم در حوزه بینایی ماشین و تشخیص الگو تبدیل شده است. تخمین سن به فرایندی گفته می‌شود که طی آن، تصویر چهره فرد توسط کامپیوتر، پردازش و سن او براساس معیار سال، برآورد می‌شود [1 و 2].

تخمین سن کاربردهای فراوانی در حوزه‌های مختلف دارد؛ برای مثال، می‌توان از تخمین سن در نرم‌افزارهایی استفاده کرد که به انجام تحلیل‌های آماری بر روی افراد، به‌منظور بهبود کار و تجارت یا ارائه خدمات می‌پردازند. یک تحلیل آماری در ژاپن مشخص کرد که قربانیان قلب در انتقال وجه نقد، معمولاً افراد با سن بالای 65 سال هستند؛ بنابراین، می‌توان یک سیستم تخمین سن را روی دستگاه‌های خودپرداز نصب کرد تا در هنگام حضور فردی با سن بیش از 65 سال، از طریق یک سیستم هشدار، افراد مربوط را آگاه کنند تا مراقبت لازم را از این‌گونه افراد به عمل آورند. به‌عنوان کاربردی دیگر، می‌توان به نرم‌افزارهایی اشاره کرد که به‌صورت خودکار، تصاویر افراد را برای موتورهای جست‌وجو دسته‌بندی می‌کنند. می‌توان از یک سیستم تخمین سن، به‌منظور کنترل دسترسی افراد به صفحات وب، براساس سن آن‌ها استفاده کرد. در مثالی دیگر، می‌توان از یک تخمین‌زننده سن در ورودی دستگاه‌های بازی در یک شهر بازی استفاده کرد تا در هنگام ورود یک فرد با سن غیرمجاز، آژیر خطر را به‌کار اندازد [1 و 2].

با وجود کاربردهای فراوان تخمین سن، تاکنون روشی که قادر به تخمین سن انسان با دقت زیاد باشد، معرفی نشده است. دقیق‌ترین روش‌های تخمین سن، با خطایی حدود 3 تا 4 سال به تخمین سن افراد می‌پردازند [1 و 2].

اگر فرایند تخمین سن را به‌صورت یک مسئله دسته‌بندی ویژگی‌ها در نظر بگیریم، وجود دو بخش توصیف چهره و یادگیری تابع تخمین سن، برای آن ضروری است [2 و 3]. انتخاب روش مناسب برای توصیف چهره که یکی از مراحل مهم در فرایند تخمین سن است، معمولاً از طریق استخراج

2. مروری بر کارهای گذشته

به طور کلی، سیستم تخمین سن از دو بخش توصیف چهره و یادگیری تابع تخمین سن تشکیل شده است. در ادامه، فعالیت‌های انجام شده در حوزه هریک از این دو بخش، به صورت مجزا بررسی می‌شوند.

1.2. توصیف چهره

اولین روش تخمین سن انسان در سال 1994، براساس دیدگاه زیستی از چهره، طراحی و ارائه شد [2]. این روش طی فرایندی دو مرحله‌ای، تصاویر چهره را به سه دسته کودک، جوان و پیر تقسیم‌بندی می‌کرد. در مرحله اول، براساس محاسبه نرخ فاصله در نقاط مختلف صورت، با کمک علم آنتوپومتری چهره [10]، تصاویر افراد بالغ و نابالغ تفکیک می‌شوند. در مرحله دوم، از تحلیل چروک‌های صورت، به منظور تقسیم‌بندی تصاویر افراد بالغ به دو دسته جوان و پیر استفاده می‌شود. این روش نسبت به جهت و زاویه تصویر حساس است.

بسیاری از مقالات در زمینه تخمین سن، از روش آماری AAM [5] به منظور توصیف چهره استفاده کرده‌اند [9]. [11 و 12]. روش AAM دو مدل از سطوح خاکستری و شکل کلی چهره ارائه می‌کند. استفاده از AAM برای توصیف چهره باعث می‌شود که مشخصه‌های سنی در مدل سطوح خاکستری و مشخصه‌های مربوط به نژاد و جنسیت، در مدل شکل نمایان شوند [11].

در [3] متدی ارائه شده که به منظور توصیف چهره، از یک روش سلسله‌مراتبی برای استخراج ویژگی، استفاده می‌کند. در این روش در مرحله اول، پارامترهای شکلی و ظاهری مدل AAM به عنوان ویژگی‌های عمومی استخراج می‌شوند. در مرحله دوم، ویژگی‌های پوست و چروک به عنوان ویژگی‌های محلی، توسط فیلترهای گابور استخراج می‌شوند؛ سپس از ترکیب ویژگی‌های عمومی و محلی، یک بردار ویژگی تشکیل می‌گردد. این روش به دلیل استفاده از فیلترهای گابور، تا اندازه‌ای در مقابل تغییرات ناشی از جهت و زاویه صورت مقاوم است.

روش‌های تخمین سن تک‌مرحله‌ای [6 و 8] یا تخمین سن سلسله‌مراتبی [3 و 7]، تخمین می‌زنند.

با توجه به اینکه طیف سنی افراد، محدوده وسیعی از اعداد را شامل می‌شود (مثلاً بین 0 تا 90) یادگیری تابع تخمین سن با استفاده از رویکرد تک‌مرحله‌ای معمولاً با خطای زیاد همراه است؛ در حالی که در رویکرد سلسله‌مراتبی برچسب سنی، برای تصاویری که از قبل دسته‌بندی شده‌اند، در یک محدوده سنی خاص جست‌وجو می‌شود. در صورتی که دسته‌بندی اولیه با دقت فراوانی انجام شود، خطا به میزان چشمگیری کاهش می‌یابد. رویکردهای سلسله‌مراتبی به دلیل اینکه نیازمند یادگیری چندین تابع تخمین سن به ازای گروه‌های سنی مختلف‌اند، با پیچیدگی‌های زیادی همراه هستند. از سوی دیگر، بیشتر روش‌های معرفی شده برای تخمین سن، از یک پایگاه داده جهت ایجاد و آزمایش مدل استفاده می‌کنند و معمولاً در مواجهه با تصاویری خارج از پایگاه داده مورد استفاده، دچار خطایی بزرگ می‌شوند؛ بنابراین، نبودن مدلی تعمیم‌پذیر که دقت زیادی در برخورد با تصاویر متنوع، از لحاظ گروه‌های سنی، جنسیتی، نژاد، شرایط نورپردازی و غیره داشته باشد، از دیگر مسائل باز در حوزه تخمین سن است.

در این مقاله، به منظور مقابله با چالش‌های مطرح شده، یک روش دومرحله‌ای برای بهبود مرحله یادگیری تابع تخمین سن ارائه شده است؛ به طوری که از یک سو، محدودیت روش‌های موجود در پوشش طیف وسیع سنی را مرتفع ساخته و از سوی دیگر، پیچیدگی روش‌های سلسله‌مراتبی را ندارد. روش یادشده علاوه بر تصاویر پایگاه داده ساخت مدل FG-NET و MORPH2 بر روی سایر تصاویر نیز، دقت فراوانی دارد.

ادامه مقاله به این صورت تنظیم شده است که در بخش دوم، مروری بر کارهای گذشته انجام خواهد شد. بخش سوم حاوی توضیحات کامل در مورد روش پیشنهادی است. در بخش چهارم، آزمایش‌ها و نتایج حاصل از آن‌ها ارائه می‌شوند و در نهایت در بخش پنجم، نتیجه‌گیری حاصل از انجام این تحقیق مطرح می‌شود.

2.2. یادگیری تابع تخمین سن

در مرحله یادگیری تابع تخمین سن نیز، به‌طور کلی، دو رویکرد تک‌مرحله‌ای و سلسله‌مراتبی دنبال شده است. رویکرد تک‌مرحله‌ای به یافتن برچسب سنی از میان سایر برچسب‌های سنی می‌پردازد [8 و 10]. حال آنکه در رویکرد سلسله‌مراتبی، ابتدا تصاویر به گروه‌های سنی، جنسیتی و غیره تقسیم می‌شوند و سپس برچسب سنی در محدوده کوچک‌تری جست‌وجو می‌شود [3 و 7]. به دلیل وسعت محدوده سنی، یافتن برچسب سنی دقیق توسط یک رویکرد تک‌مرحله‌ای، دشوار و با خطای زیاد همراه است. در رویکرد سلسله‌مراتبی، به دلیل وجود گروه‌های سنی، جنسیتی و غیره و نیاز به آموزش مدل مجزا برای هر گروه، پیچیدگی افزایش می‌یابد و این موضوع باعث کاهش محبوبیت این روش‌ها می‌شود.

3. روش پیشنهادی

در این مقاله، یک روش جدید برای بهبود مرحله یادگیری تابع تخمین سن ارائه شده است. روش ارائه‌شده، در دو مرحله با استفاده از شبکه عصبی MLP به تخمین سن افراد می‌پردازد. شکل (2) فرایند روش پیشنهادی را نمایش می‌دهد.

1.3. توصیف چهره

همان‌طور که پیش از این نیز بیان شد، گام نخست در فرایند تخمین سن، توصیف چهره از طریق انتخاب ویژگی‌های مناسب است. این ویژگی‌ها باید حاوی اطلاعاتی باشند که علاوه بر ایجاد بیشترین خاصیت تفکیک‌پذیری، نسبت به تغییراتی مانند چرخش و مقیاس نیز حساس نباشند.

در این مقاله، برای به‌دست‌آوردن توصیف مناسب از چهره، از روشی که در [8] پیشنهاد شده بود، استفاده می‌شود. در این روش، ابتدا ویژگی‌های چهره با استفاده از روش BIF استخراج می‌شود. ویژگی‌های استخراج‌شده با استفاده از روش BIF در برابر تغییرات ناشی از چرخش و زاویه حساس نیستند؛ سپس به‌منظور کاهش ابعاد بردار ورودی و ایجاد بردارهای حاوی ویژگی‌هایی با بیشترین قدرت تفکیک‌پذیری،

روش بعدی با نام AGES¹ [9] شناخته می‌شود که در آن، تصاویر مختلفی از سنین مختلف افراد خاصی در یک الگوی سنی مرتب می‌شود؛ سپس هریک از تصاویر موجود در الگوی سنی، به شکل مجزا توسط روش AAM توصیف می‌شوند. در صورت نبودن تصویر برای یک یا چند سن مختلف از یک فرد، این تصویر توسط الگوریتم تکراری EM² از روی سایر تصاویر موجود ساخته می‌شود. به‌منظور تخمین سن یک تصویر آزمایشی با این روش، ابتدا الگوی سنی مشابه آن تصویر و سپس مکان تصویر که متناظر با سن تصویر است، در آن الگوی سنی تعیین می‌شود. در عمل، در صورت نبودن تصاویر زیاد در یک الگوی سنی، تصاویر ساخته‌شده دارای وضوح و دقت کافی نبوده و همین موضوع کارایی روش را کاهش می‌دهد.

به‌منظور بهبود روش‌های مذکور، روش‌های کاهش ابعاد غیرخطی مبتنی بر یادگیری منیفلد، مورد استفاده قرار گرفته‌اند؛ از جمله روش‌های منیفلد استفاده‌شده برای تجسم الگوی سنی می‌توان به OLPP³ [13]، MFA⁴ [14] و LSDA⁵ [15] اشاره کرد. این روش‌ها برای کاهش ابعاد، جهتی را انتخاب می‌کنند که داده‌های مشابه نزدیک یکدیگر و به‌دور از داده‌های غیرمشابه قرار بگیرند. دو محدودیت عمده این روش‌ها نیاز به مجموعه آموزش بزرگ و حساسیت به تغییرات در جهت و زاویه چهره است. برای حذف محدودیت دوم، می‌توان به جای اعمال روش منیفلد بر تصاویر خام از چهره، این روش را بر روی ویژگی‌های استخراج‌شده از چهره با استفاده از روش BIF [16] اعمال کرد [2، 7، 12، 17، 18 و 19]. علاوه بر روش‌های یادشده، روش‌های کاهش ابعاد دیگری نیز در فرایند تخمین سن مورد استفاده قرار گرفته‌اند که از میان آن‌ها روش PLS [8] موفقیت بیشتری نسبت به سایر روش‌ها در تجسم داده‌ها در ابعاد پایین‌تر دارد.

1. AGing Pattern Subspace
2. Expectation Maximization
3. Orthogonal Locality Preserving Projection
4. Marginal Fisher Analysis
5. Locality Sensitive Discriminant Analysis

بردارهای امتیاز² توصیف کرد و اینکه بردارهای امتیاز از طریق اعمال PCA بر روی بردارهای داده به دست می‌آیند، فرایند PLS را بررسی می‌کنیم. در ساده‌ترین حالت، می‌توان مدل PLS را شامل ارتباط خارجی هر یک از مجموعه‌ها به صورت مجزا و نیز ارتباط درونی میان دو مجموعه، در نظر گرفت. ارتباط خارجی برای هر یک از این دو مجموعه، از طریق اعمال PCA به دست می‌آید (رابطه 1) و هدف PLS توصیف این بردارها همراه با یافتن ارتباط آن‌ها به یکدیگر است.

$$\begin{aligned} X &= TP^T + E \\ Y &= UQ^T + F \end{aligned} \quad (1)$$

در رابطه (1)، T و U ماتریس‌هایی $(n \times p)$ از بردار امتیاز استخراج شده‌اند. ماتریس $(N \times p)$ به نام P و ماتریس $(M \times p)$ به نام Q نیز، ماتریس‌های بارگذاری³ هستند. همچنین ماتریس‌های $N \times n$ و $M \times n$ با نام‌های E و F نیز مشخص‌کننده ماتریس‌های خطا⁴ می‌باشند. در اینجا هدف این است که در حد امکان، Y را توسط بردارهای امتیاز توصیف کنیم و برای انجام این کار، باید تا حد امکان، $|F|$ را کوچک کرد. با این روش، می‌توان هم‌زمان با کوچک شدن $|F|$ اطلاعات مفیدی را در راستای یافتن ارتباط میان X و Y نیز به دست آورد. این ارتباط را می‌توان از طریق یافتن رگرسیون خطی U در مقابل T برای هر مؤلفه به دست آورد. این مدل ساده خطی را می‌توان به صورت زیر نمایش داد:

$$\bar{u} = b_h t_h \quad (2)$$

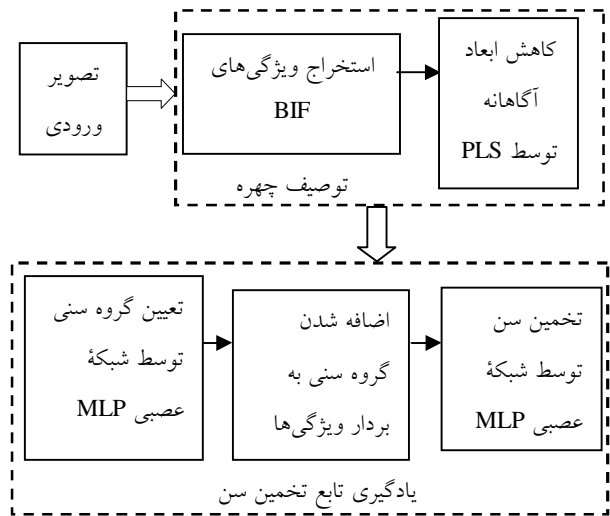
که b_h نقش ضرایب رگرسیون را ایفا می‌کند و برابر $b_h = \frac{\hat{u}_h t_h}{t_h t_h}$ است. ساده‌ترین حالت مدل PLS در فرم الگوریتمی خود، شامل مراحل¹ است که در جدول (1) بیان شده‌اند.

به منظور ارتقای ارتباط درونی میان این دو مجموعه، باید هر مجموعه اطلاعاتی از دیگری را داشته باشد؛ بنابراین، می‌توان در مرحله دوم الگوریتم‌های جدول (1)، t را با u و u را با t جایگزین کرد. اگر مجموعه Y تنها شامل یک متغیر باشد، مراحل مرتبط با آن، با قراردادن $q = 1$ حذف

روش کاهش ابعاد آگاهانه PLS بر روی ویژگی‌های حاصل از روش BIF اعمال می‌شود و یک بردار با 40 ویژگی به دست می‌آید.

1.1.3. روش استخراج ویژگی BIF

روش استخراج ویژگی BIF برگرفته از پردازش بصری در ناحیه قشر چشم است. در این روش، مشابه با ساختار قشر چشم، مدلی معرفی می‌شود که متشکل از تعدادی لایه‌های متناوب به نام‌های ساده (S) و پیچیده (C) است. این لایه‌ها پیچیدگی افزایشی واحدهای سلولی از قشر اولیه بینایی تا قشر تحتانی چشم را مدل می‌کنند. لایه اول (لایه ساده S1) توسط اعمال فیلتر گابور بر روی تصویر ورودی ایجاد می‌شود. لایه دوم (لایه پیچیده C1) از اعمال عملگر ماکزیمم بر روی لایه اول ایجاد می‌شود. به منظور استفاده از روش استخراج ویژگی BIF در تخمین سن، استفاده از همین دو لایه کفایت می‌کند [8].



شکل (2): روش پیشنهادی مقاله

2.1.3. روش کاهش ابعاد آگاهانه PLS

تنظیمات عمومی الگوریتم PLS به مدل کردن ارتباط میان دو مجموعه داده متعلق به دو فضای N و M بعدی می‌انجامد که به صورت $X \in \mathbb{R}^N$ و $Y \in \mathbb{R}^M$ نمایش داده می‌شوند. مدل PLS بر اساس مشخصه‌های الگوریتم NIPALS¹ ساخته شده است. در اینجا با قبول این واقعیت که می‌توان ماتریس داده را توسط

1. NonLinear Iterative Partial Least Squares

2. Score Vectors
3. Loading
4. Residuals

این حالت، به دلیل اینکه جست‌وجو برای یافتن برچسب سنی، در محدوده کوچک‌تری صورت می‌پذیرد، نتایج بهتری حاصل می‌شود؛ اما این روش دو ضعف عمده دارد:

1. در صورتی که تخمین اولیه دقیق نباشد، تصاویر در گروه‌های نامناسب قرار می‌گیرند و تخمین ثانویه آن‌ها توسط دسته‌بندی‌کننده نامناسب انجام می‌شود و نتایج حتی از حالت تک‌مرحله‌ای هم ضعیف‌تر می‌شوند؛ مثلاً ممکن است تصویر یک فرد جوان به اشتباه در گروه افراد میان‌سال قرار گیرد؛ بنابراین تخمین سن وی با خطای زیادی همراه است.

2. استفاده از دسته‌بندی‌کننده‌های مختلف، پیچیدگی بیشتری ایجاد می‌کند.

برای حل چالش‌های مطرح‌شده، در این مقاله، روش جدیدی پیشنهاد شده است. روش پیشنهادی از یک سو، با در پیش‌گرفتن یک رویکرد دو مرحله‌ای، مشکل گستردگی برچسب‌های سنی در یک رویکرد تک‌مرحله‌ای را حل می‌کند و از سوی دیگر، فارغ از پیچیدگی‌های روش‌های سلسله‌مراتبی موجود است. روش پیشنهادی به این ترتیب عمل می‌کند که در ابتدا یک دسته‌بندی کلی روی داده‌ها انجام می‌دهد و آن‌ها را به گروه‌های سنی با فاصله 4 سال، تقسیم می‌کند؛ بنابراین، برای تصاویر پایگاه داده‌های MORPH2 که محدوده سنی بین 16 تا 68 سال است، 13 گروه سنی و برای پایگاه داده‌های FG-NET با محدوده سنی 0 تا 69 سال، 17 گروه سنی به دست می‌آید. سپس این گروه‌های سنی همانند یک ویژگی به بردار ویژگی‌ها افزوده می‌شود و تخمین سن دقیق برای بردار ویژگی حاصل، محاسبه می‌شود. تابع تخمین سن در هر دو مرحله، توسط شبکه عصبی MLP، تقریب زده می‌شود. در این روش، تنها دو مدل برای تخمین سن تمامی تصاویر ایجاد می‌شود؛ لذا ایراد پیچیدگی به دلیل تعدد مدل در روش سلسله‌مراتبی برطرف می‌شود.

علت تقسیم‌بندی تصاویر به بازه‌های 4 ساله، این است که در صورتی که بردار ویژگی به دست آمده از مرحله توصیف چهره را طی یک فرایند تک‌مرحله‌ای دسته‌بندی کنیم، خطای تخمین سن حدود 4 سال خواهد بود. از سوی دیگر، تعداد گروه‌های سنی مناسب با روش سعی و خطا و به صورت تجربی هم

می‌شوند. این الگوریتم معمولاً به سرعت همگرا می‌شود و مؤلفه‌های مجموعه X و Y را مشخص می‌کند.

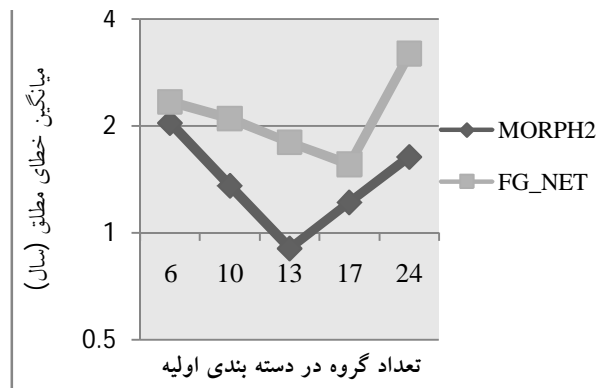
جدول (1): مراحل الگوریتم PLS [8]

مرحله	برای مجموعه X	برای مجموعه Y
1	$t = X^T u$	$u = Y^T t$
2	$p' = t^T X / (t^T t)$	$q' = u^T Y / (u^T u)$
3	$p'_{new} = p'_{old} / \ p'_{old}\ $	$q'_{new} = q'_{old} / \ q'_{old}\ $
4	$t = X p / p^T p$	$u = Y q / q^T q$
5	اگر t در مرحله دوم و چهارم برابر بودند، متوقف و در غیر این صورت، به مرحله دوم برو	اگر u در مرحله دوم و چهارم برابر بودند، متوقف و در غیر این صورت به مرحله دوم برو

2.3. یادگیری تابع تخمین سن

دیدیم که می‌توان فرایند یادگیری تابع تخمین سن را به یکی از دو روش تک‌مرحله‌ای یا سلسله‌مراتبی انجام داد. در حالت تک‌مرحله‌ای، فرایند تخمین سن توسط یک دسته‌بندی‌کننده انجام می‌شود و برچسب سنی تصویر مورد نظر، از میان تمامی برچسب‌های سنی برگزیده می‌شود. در مقابل، در روش سلسله‌مراتبی، از بیش از یک دسته‌بندی‌کننده برای تخمین سن استفاده می‌شود. عموماً انجام این فرایند به این صورت است که در مرحله اول، از یک (چند) دسته‌بندی‌کننده استفاده می‌شود تا تصاویر به گروه‌های سنی، جنسیتی، نژاد و غیره تقسیم شوند و سپس در مرحله دوم، برای هر گروه، دسته‌بندی‌کننده ویژه‌ای طراحی می‌شود تا به تخمین دقیق‌تری از سن پردازد؛ برای مثال، می‌توان تصاویر را به گروه‌های زنان سفیدپوست جوان، زنان سفیدپوست میان‌سال، زنان سیاه‌پوست جوان، مردان سفیدپوست جوان، مردان سفیدپوست میان‌سال، مردان سفیدپوست پیر، مردان سیاه‌پوست جوان، مردان سیاه‌پوست میان‌سال و مردان سیاه‌پوست پیر، تقسیم‌بندی کرد و سپس برای هر گروه، یک تابع تخمین سن مجزا تقریب زد. در

محاسبه شده است که درستی این مسئله را تصدیق می‌کند (شکل 3).



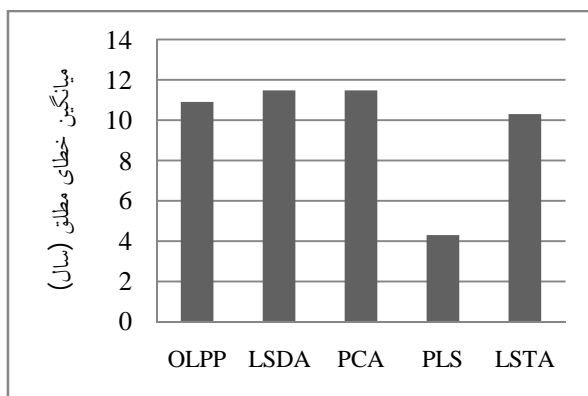
شکل (3): یافتن تعداد گروه‌های سنی مناسب

در هنگام استفاده از هریک از این دو پایگاه داده، 20% تصاویر برای آزمون و مابقی به منظور آموزش مدل مورد استفاده قرار گرفتند. تمامی تصاویر استفاده شده جهت انجام هریک از مراحل ساخت و آزمایش مدل، ابتدا به مدل خاکستری تبدیل و سپس، به منظور جلوگیری از ایجاد سربار محاسباتی توسط سایر اطلاعات اضافی موجود در تصویر، مانند حاشیه و پس‌زمینه و سرعت‌بخشیدن به زمان پردازش، قسمت درونی چهره در اندازه 128×128 از تصویر جدا می‌شود.

به منظور ارزیابی کیفیت نتایج و مقایسه آن با سایر روش‌ها، از معیار ارزیابی میانگین خطای مطلق استفاده شده است.

1.4. آزمایش اثر روش کاهش ابعاد بر دقت تخمین سن

در این آزمایش، ابتدا با اعمال روش BIF، ویژگی‌هایی از تصاویر استخراج شدند. در ادامه، بعد بردار ویژگی‌های استخراج شده با روش BIF توسط روش‌های مختلف کاهش ابعاد، کم شد. یادگیری تابع تخمین سن در یک مرحله و توسط شبکه عصبی MLP با 10 نرون و دو لایه، انجام شد. شکل (4) نتایج حاصل از این آزمایش را نمایش می‌دهد.



شکل (4): بررسی اثر روش‌های مختلف کاهش ابعاد بر دقت تخمین سن

با توجه به شکل (4)، مشخص است که روش PLS نسبت به سایر روش‌ها موفقیت بیشتری در کاهش ابعاد داشته است. علت این اتفاق با توجه به توضیحات بخش 2.1.3 پیش‌بینی‌شدنی بود؛ زیرا روش PLS در هنگام انتخاب ویژگی‌هایی با بیشترین قدرت تفکیک‌پذیری، ارتباط میان بردار ورودی و خروجی (بردار برجسته‌های سنی) را مد نظر قرار

تاکنون، بیشتر روش‌های تخمین سن، از دسته‌بندی‌کننده SVM به منظور دسته‌بندی تصاویر و تخمین سن استفاده کرده‌اند. در این مقاله، از شبکه عصبی MLP به منظور دسته‌بندی بردارهای ویژگی حاصل از تصاویر استفاده شده است. علت این انتخاب یکی دقت بیشتر شبکه عصبی MLP در مقایسه با سایر دسته‌بندی‌کننده‌ها (آزمایش 3.4) و دیگری سرعت دسته‌بندی بیشتر در مقایسه با سایر دسته‌بندی‌کننده‌ها می‌باشد.

4. آزمایش‌ها و نتایج

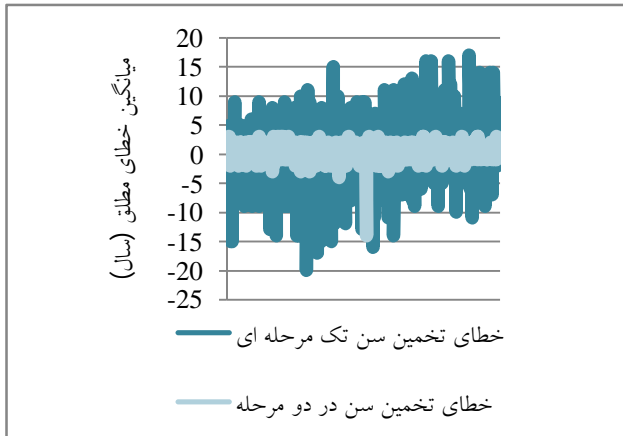
به منظور آزمون روش پیشنهادی و نیز مشخص شدن برتری آن نسبت به روش‌های مشابه، از تصاویر دو پایگاه داده MORPH2 و FG-NET استفاده شده است. پایگاه داده FG-NET حاوی 1002 تصویر از افراد در محدوده سنی 0 تا 69 است. جدول (2)، حاوی مشخصات پایگاه داده MORPH2 است.

جدول (2): مشخصات پایگاه داده MORPH2

نژاد	زن و مرد	زن	مرد
سیاه	42560	5757	36803
سفید	10600	2601	7999
اسپانیایی	1751	100	1651
آسیایی	159	13	146
هندی	57	14	43
سایر	5	2	3
مجموع	55132	8487	46645

عبارت است از: 54% بر روی پایگاه داده FG-NET و 78% بر روی پایگاه داده MORPH2.

با توجه به نتایج به دست آمده در این آزمایش، برتری از لحاظ دقت، با استفاده از روش پیشنهادی مشخص می‌شود.



شکل (5): مقایسه خطای تخمین سن هر تصویر از مجموعه آزمایشی در دو حالت تخمین سن تک مرحله‌ای و دو مرحله‌ای

3.4. آزمایش مقایسه قدرت دسته‌بندی کننده‌های مختلف

این آزمایش نیز بر روی زیرمجموعه 4000 تایی از پایگاه داده‌های MORPH2، به این ترتیب انجام شد که ویژگی‌های تصاویر توسط روش‌های ترکیبی BIF+PCA و BIF+PLS استخراج شدند و مرحله یادگیری تابع تخمین سن بر روی این ویژگی‌ها توسط دسته‌بندی کننده‌های مختلف انجام شد. نتایج حاصل از این بررسی در شکل (6)، به تصویر کشیده شده است. همان‌طور که از تصویر مشخص است، دسته‌بندی کننده شبکه عصبی نتایج بهتری از لحاظ دقت، نسبت به سایر دسته‌بندی کننده‌ها دارد.

4.4. آزمایش بررسی تعمیم پذیری روش پیشنهادی

به منظور انجام این آزمایش، پایگاه داده FG-NET به عنوان مجموعه آموزشی در نظر گرفته شد. در اینجا به منظور بررسی قدرت تعمیم پذیری روش پیشنهادی، یک مجموعه آزمون حاوی 20 تصویر از افراد در سنین 1 تا 65 سال توسط نویسندگان مقاله جمع آوری شد. تصاویر این مجموعه نیز به مدل خاکستری تبدیل شدند و قسمت درونی چهره در اندازه

می‌دهد (روش کاهش ابعاد آگاهانه).

2.4. آزمایش قدرت روش پیشنهادی در بهبود دقت تخمین سن

به منظور انجام این آزمایش، پس از انجام مرحله توصیف چهره، ابتدا یک دسته‌بندی کلی توسط شبکه عصبی MLP روی بردارهای ورودی انجام شد و داده‌ها به گروه‌های سنی با فاصله چهار سال تقسیم شدند. در ادامه، عدد گروه سنی به بردار ویژگی‌ها اضافه شد و دوباره دسته‌بندی توسط شبکه عصبی MLP انجام گرفت. نتایج این آزمایش در جدول (3) نمایش داده شده است.

جدول (3): بررسی کارایی روش پیشنهادی

پایگاه داده	توصیف چهره	تعداد ویژگی‌ها	میانگین خطای مطلق برای دسته‌بندی کننده اولیه	میانگین خطای مطلق (سال)
MORPH2	BIF + PCA	201	0.2	0.9694
MORPH2	BIF + PLS	41	0.2	0.9056
FG-NET	BIF + PCA	201	0.2	2.3
FG-NET	BIF + PLS	41	0.2	1.56

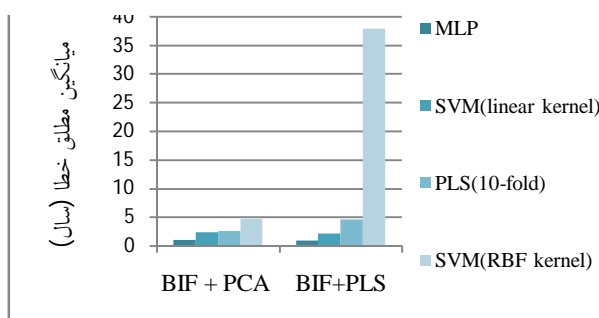
علت بهبود نتایج را می‌توان با توجه به شکل (5)، تفسیر کرد. در صورتی که تخمین سن طی یک مرحله انجام شود، بازه بروز خطا بسیار گسترده و در محدوده 20- تا 16 سال قرار می‌گیرد؛ اما با به کارگیری روش دو مرحله‌ای پیشنهاد شده، محدوده بروز خطا به دلیل محدود شدن بازه جست‌وجوی برچسب سنی کوچک‌تر شده و در نتیجه، میانگین مطلق خطا کاهش چشمگیری دارد.

از مقایسه نتایج جدول (3) با بهترین نتایج به دست آمده تاکنون، بر روی دو پایگاه داده MORPH2 و FG-NET (به ترتیب 3.4 سال [20] و 4.15 سال [8]) مشخص می‌شود که روش پیشنهادی موفقیت چشمگیری در بهبود دقت فرایند تخمین سن دارد؛ بنابراین، میزان بهبود روش پیشنهادی نسبت به بهترین روش‌های معرفی شده تا پیش از این مقاله، به ترتیب

آزاد جمع‌آوری شده توسط نویسندگان مقاله، مورد آزمایش قرار گرفت و نتایج رضایت‌بخشی (کاهش خطا تا 1.45 سال براساس معیار میانگین مطلق خطا) به همراه داشت؛ لذا می‌توان گفت روش ارائه شده علاوه بر اینکه دقت زیاد و پیچیدگی بسیار کمی در مقایسه با سایر روش‌ها دارد، برخلاف بیشتر روش‌های تخمین سن معرفی شده تاکنون، قابلیت تعمیم‌پذیری فراوانی نیز دارد. همچنین روش معرفی شده، برای ساخت مدل تعمیم‌پذیر، به تعداد زیادی از تصاویر نیاز ندارد. به دلیل اینکه تمامی تصاویر استفاده شده برای این روش، به مدل خاکستری تبدیل شده و در اندازه 128×128 ، استفاده می‌شوند، این روش مستقل از ویژگی‌هایی مانند رنگ و اندازه تصویر است.

با توجه به اینکه روش پیشنهادی، از بردار ویژگی با ابعاد پایین (1×40) برای یادگیری تابع تخمین سن استفاده می‌کند، می‌توان گفت سربار محاسباتی این روش، در مقایسه با بیشتر روش‌های پیشین، در زمینه تخمین سن که از بردارهایی با ابعاد بالاتر استفاده می‌کنند، بسیار کمتر است. سرعت تخمین سن برای یک تصویر جدید با استفاده از روش پیشنهادی، حدود دو دقیقه است. (با توجه به شواهد حاکی از پیاده‌سازی روش‌های مشابه توسط نویسندگان مقاله، سرعت روش پیشنهادی کمتر از سرعت روش‌های مشابه نیست؛ هرچند که هیچ‌یک از روش‌های مشابه، به صورت مستند، اشاره‌ای به سرعت روش پیشنهادی خود نکرده‌اند.)

128×128 از تصویر جدا شد. سپس از مدل ایجاد شده با استفاده از روش پیشنهادی، بر روی مجموعه آموزشی به منظور تخمین سن تصاویر جمع‌آوری شده در مجموعه آزمون استفاده شد. دقت تخمین سن در این حالت، به 1.45 سال رسید. دقت به دست آمده مشخص‌کننده قدرت تعمیم‌پذیری فراوان روش پیشنهادی است.



شکل (6): مقایسه دقت دسته‌بندی‌کننده‌های مختلف در تخمین سن

5. نتیجه‌گیری

در این مقاله، یک روش دومارحله‌ای برای تخمین سن افراد از روی تصاویر چهره آنان ارائه شد. روش ارائه شده بر روی تصاویر دو پایگاه داده معروف FG-NET و MORPH2 اجرا شد و نتیجه بسیار بهتری از سایر روش‌های اجرا شده بر روی این دو پایگاه داده به دست آمد؛ به طوری که دقت حاصل از تخمین سن با استفاده از روش پیشنهادی بر روی پایگاه داده FG-NET، 54% و بر روی پایگاه داده MORPH2، 78% است. همچنین مدل به دست آمده از طریق اجرای این روش، بر روی پایگاه داده FG-NET، بر روی تصاویر یک پایگاه داده

مراجع

- [1] Fu Y., Guo G., Hough T. S., "Age Synthesis and Estimation via Faces: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 (11), pp.1955-1976, 2010.
- [2] Guo G., "Human Age Estimation and Sex Classification," *Video Analytics for Business Intelligence*, Springer-Verlag Berlin Heidelberg, 409, pp. 101-126, 2012.
- [3] Choi S., Lee Y., Lee S., Park K., Kim J., "Age Estimation Using a Hierarchical Classifier Based on Global and Local Facial Features," *Pattern Recognition*, 44 (6), pp. 1262-1281, 2011.
- [4] Kwon Y., Lobo N., "Age Classification from Facial Images," *Computer Vision and Image Understanding (CVIU)*, 74 (1), pp. 1-121, 1999.
- [5] Cootes T., Edwards F., Taylor C., "Active Appearance Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23

- (6), pp. 681-685, 2001.
- [6] Lanitis A., Draganova C., Christodoulou C., "Comparing Different Classifiers for Automatic Age Estimation," *IEEE Transactions on Systems, Man, and Cybernetics*, 24 (4), pp. 621-628, 2002.
- [7] Guo G., Fu Y., Dyer Ch., Y, Huang T., Fellow L., "Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression," *IEEE Transactions on Image*, 17 (7), pp. 1178-1188, 2008.
- [8] Guo G., "Simultaneous Dimensionality Reduction and Human Age Estimation via Kernel Partial Least Squares Regression," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Colorado Springs, CO, USA, pp. 657- 664, 2011.
- [9] Geng X., Zhou Z. H., Zhang Y., Li G., and Dai H., "Learning from Facial Aging Patterns for Automatic Age Estimation," *ACM International Conference on Multimedia*, New York, NY, USA, pp. 307 – 316, 2006.
- [10] Farkas L., "Anthropometry of the Head and Face", *Raven Press*, New York. 1994.
- [11] Ricanek K., Wang Y., Chen C., Simmons S. J., "Generalized Multi-Ethnic Face Age Estimation," *IEEE 3rd International Conference on Biometrics, Theory, Applications, and Systems*, Washington, DC, pp. 1- 6, 2009.
- [12] Fu Y., Huang T., "Human Age Estimation with Regression on Discriminative Aging Manifold," *IEEE Transactions on Multimedia*, 10 (4), pp. 578-584, 2008
- [13] Cai D., He X., Han J., "Orthogonal Laplacian Faces for Face Recognition," *IEEE Transactions on Image*, 15 (11), pp. 3608-3614, 2006.
- [14] Yen S., Yu D., Zhang B., and Zhang H., "Graph Embedding: A General Framework for Dimensionality Reduction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 (1), pp. 40 – 51, 2007.
- [15] Cai D., He X., Zhou K., "Locality Sensitive Discriminant Analysis," *International Journal of Human-Computer Studies*, 7, pp. 708–713, 2007.
- [16] Riesenhuber M., Poggio T., "Hierarchical Models of Object Recognition in Cortex," *Nature Neuroscience*, 2 (11), pp. 1019-1025, 1999.
- [17] Guo G., Mu G., Fu Y., Dyer C., and Huang, T., "A Study on Automatic Age Estimation on a Large Database," *IEEE 12th International Conference on Computer Vision*, Kyoto, Japan, pp. 1986-1991, 2009.
- [18] Guo G., Mu G., Fu Y., Dyer C., Huang T., "Locally Adjusted Robust Regression for Human Age Estimation," *IEEE Workshop on Application of Computer Vision*, Copper Mountain, CO, USA, pp. 1- 6, 2008.
- [19] Guo G., Mu G., "Human Age Estimation: What is the Influence Across Race and Gender?" *Computer Vision and Pattern Recognition Workshop (CVPRW)*, San Francisco, CA, 2010. pp. 71-78, 2010.
- [20] Thukral T., Mitra K., Chellappa R. A., "Hierarchical Approach for Human Age Estimation," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, pp. 1529 -1532, 2012.