فصلنامه پژوهش‌های حفاظتی-امنیتی
دانشگاه جامع امام حسین(علیه‌السلام)
سال یازدهم، شمارهٔ ۴۳ (پاییز ۱۴۰۱)صص ۱۵۸- ۱۲۳

ارائه یک ساز و کار هوشمند برای تشخیص چهره مرتبطین نیروهای مسلح در نقاط قرمز با استفاده از الگوریتم Big-GAK

● محسن رب دوست ●
کارشناس ارشد هو ش مصنوعی ، دانشگاه جامع امام حسین (ع)، تهران، ایران (نویسنده مسئول)
● میثم رب دوست ●
دانشجو دکتری مدیریت فناوری اطلاعات ، واحد تهران جنوب، دانشگاه ازاد اسلامی، تهران، ایران
● محمدرضا حسنی آهنگر ●
استاد، دانشکده هوش مصنوعی و علوم شناختی، دانشگاه جامع امام حسین (ع)، تهران، ایران
● محمدجعفری خرمی ●
کارشناس ارشد مدیریت، دانشگاه جامع امام حسین (ع)، تهران، ایران

تاریخ دریافت: ۱۴۰۱/۰۲/۱۰ تاریخ پذیرش: ۱۴۰۱/۰۳/۲۷

چکیده

مطابق سند معماری هوشمند سازی ساحفاهای نیروهای مسلح، کاهش حضور فیزیکی نیروهای عملیاتی با بکارگیری فناوری های نوین و توانمندسازی های دیجیتال در اقدامات عملیاتی مانند: تعقیب و مراقبت، از اهداف سازمان های حفاظت اطلاعات می باشد. یکی از ابزارهای هوشمند برای تحقق این امر، طراحی و پیاده‌سازی سامانه های تشخیص چهره است که به دلیل استفاده از تکنیک‌های مختلف پردازش تصویر به صورت توأمان و شباهت بسیار زیاد تصاویر چهره انسان در بانک‌های اطلاعاتی حجیم، با پیچیدگی‌های فراوانی روبرو شده است. ما برای فائق آمدن بر این مشکلات و ایجاد امکان تشخیص چهره در بانک‌های اطلاعاتی حجیم بالای ۵۰۰ هزار تصویر، ضمن تشریح الگوریتم Big-GAK به ارائه مکانیسم هوشمندی شامل مراحل ذیل خواهیم پرداخت:

۱.استخراج ویژگی ها در تعامل با واحدهای شناختی پیش‌پردازش و کنترل کیفی ۲.دسته بندی با استفاده از الگوریتم های شناختی و الگوریتم ژنتیک ۳. بکارگیری الگوریتم محلی LBP و الگوریتم K-Points جهت کاهش محدوده تصاویر و سپس تشخیص دقیق ۴. صحت سنجی به صورت باناظر و ارجاع به مرحله دوم الگوریتم ژنتیک در صورت عدم موفقیت.

کلید واژگان: استخراج ویژگی، الگوریتم ژنتیک، الگوریتم Big-GAK ،شناسایی چهره، سوژه یابی،

مقدمه

اطلاع یافتن سازمان‌های حفاظت اطلاعات از حضور مرتبطین در اماکن قرمز بخشی از سوژه یابی می‌باشد.

به دلیل کاستی روش‌های مختلف رصد موقعیت مکانی نظیر:

۱-تعویض گوشی و سیم کارت در نقطه یابی از طریق شبکه سلولار تلفن همراه

۲-محدودیت منابع انسانی در تعقیب و مراقبت فیزیکی

۳-محدودیت لجستیکی، همچنین حفاظت و کفایت در نصب ردیاب‌های الکترونیکی

روش کارآمدی در این زمینه و با استفاده از ابزارهای هوش مصنوعی جهت تشخیص چهره با تقاطع‌گیری در دو منبع ذیل می‌باشد.

۱-دوربین‌های گسترده شهری

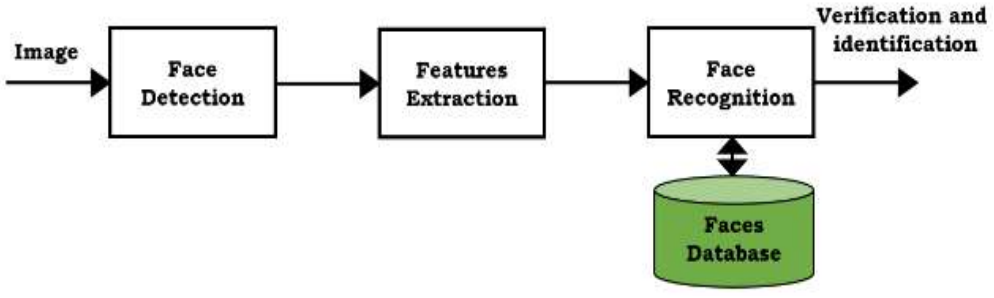
۲-تصاویر مرتبطین و مترددین سازمان‌های حفاظت شونده

ما در این پژوهش با بهره‌گیری از تکنیک‌های پردازش تصاویر در بانک‌های اطلاعاتی حجیم، مکانیسم هوشمندی جهت دستیابی به سرنخ‌های اطلاعاتی ارائه می‌نماییم.

سامانه‌های تشخیص چهره به دلیل کاربردهای فراوان در حوزه‌های صنعتی، امنیتی و خدماتی مورد توجه ویژه مهندسین نرم‌افزار و هوش مصنوعی قرار دارد(Afra S, Alhajj R)

امروزه با تشکیل بانک‌های اطلاعاتی حجیم تصاویر چهره انسان، تقاضاهای روز افزونی جهت تولید سامانه‌های تشخیص چهره با دقّت شناسایی بسیار بالا به وجود آمده است(Zhao F, Li J)

معماری عمومی روش‌های تشخیص چهره کنونی که در شکل (۱) نمایش داده شده، عموماً از سه مرحله اصلی (شناسایی چهره، تشخیص چهره و استخراج ویژگی) تشکیل شده است:



شکل شماره ۱: مکانیزم تشخیص چهره

(Figure-1) Face Recognition method

معایب روش‌های تشخیص چهره کنونی:

۱- عدم بازخورد به مرحله قبل جهت اِصلاح عملکرد

۲- فقدان شناخت کامل به مأموریت سایر قسمت‌ها

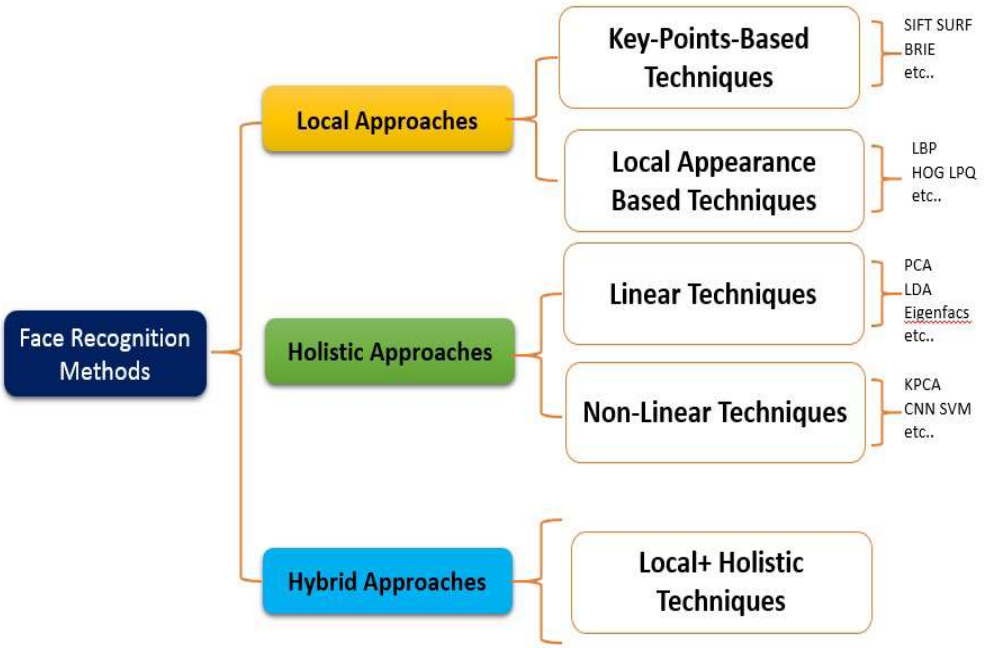
۳- عدم تصمیم‌گیری بر اساس ویژگی‌های اِنحصاری هر بانک تصویر

۴- عدم امکان استفاده از الگوریتم‌ها محلی[1] و عمومی[2] به صورت توأمان و موثّر

۵- عدم وجود مکانیزم کنترل کیفیت جهت تغییر الگوریتم‌های مورد استفاده در هر مرحله

در مکانیزم‌های موجود یا از دسته الگوریتم‌های محلی استفاده می‌نمایند که به دلیل ویژگی این الگوریتم‌ها که دارای قدرت تشخیص و بازدهی بسیارخوب در بانک‌های اطلاعاتی کوچک هستند و یا از دسته الگوریتم‌های عمومی استفاده می‌گردد که دارای دقّت تشخیص متوسط در بانک‌های اطلاعاتی حجیم می‌باشند که در هر دو صورت نتیجه مطلوب حاصل نخواهد شد. جهت آشنایی بیشتر با الگوریتم‌های تشخیص چهره، دسته‌بندی اِجمالی آن‌ها در شکل (۲) نشان داده شده است.

(Beham MP, Roomi SM)



شکل شماره۲: الگوریتم‌های محلی و عمومی تشخیص چهره

(Figure-2) Local and general face recognition algorithms

1. Local Appproaches
2. Holistic Appproaches

عدم ارتباط معنادار سه بخش اصلی سامانه‌های تشخیص چهره و همچنین عدم وجود دانش لازم نسبت به بانک اطلاعات تصاویر مورد استفاده، اکثراً باعث انجام اقدامات و تصمیم گیری‌هایی در مراحل مختلف می‌گردد که گرچه در آن مرحله مفید بوده و منجر به افزایش کارآیی آن بخش می‌گردد، لیکن در مراحل بعدی بعضاً مضر بوده و باعث کاهش کارآیی کلی مکانیزم تشخیص چهره می‌گردد. لذا می‌بایست در ابتدا روشی را پیشنهاد نمود که فاقد ایرادات مذکور بوده و همچنین از تمام الگوریتم‌ها و تکنیک‌های تشخیص چهره و هوش مصنوعی[1] بتوان به صورت موثّر و کارا در این مکانیزم استفاده نمود.

پیشینه پژوهش

در این بخش با توجه به اهمیت موضوع، آشنایی با الگوریتم‌های تشخیص چهره موجود و میزان کارآیی آن‌ها در بانک‌های اطلاعاتی حجیم یک مقاله که به مقایسه کارآیی روش‌های تشخیص چهره تا یک میلیون تصویر می‌پردازد را مورد مطالعه قرار می‌دهیم و سپس به بررسی مقاله دوم که با ارائه یک الگوریتم کارا و ترکیبی موفق به تشخیص چهره با ضریب خطای قابل قبول در ۸۰ میلیون تصویر نموده است اقدام می‌نماییم.

مقاله اول: مقایسه الگوریتم‌های تشخیص چهره در مقیاس‌های مختلف

در ابتدا چهار الگوریتم معروف تشخیص چهره بر روی بانک اطلاعاتی LFW که شامل نزدیک به ۱۰ هزار تصویر است مورد ارزیابی قرار گرفته و سپس این الگوریتم‌ها بر روی مجموعه داده‌ای Flickr شامل ۱میلیون تصویر ارزیابی گردیده است.(Somogyi Z)

الگوی باینری محلی LBP[2]

با استفاده از این روش به عنوان یکی از روش‌های استخراج ویژگی به یک سری اعداد که همان بُردار ویژگی‌های ما هستند دست می‌یابیم و بر اساس میزان اختلاف با عکس کاندید، فرد مورد نظر را تشخیص می‌دهیم[۶]. از ویژگی این روش می‌توان به عدم استفاده از داده آموزشی، سرعت بالا و پیاده‌سازی آسان اشاره نمود که بر روی بانک LFW دارای ۷۰ درصد شناسایی موفق بوده است.

1. Artificial Intelligence
2. Local binary Pattern

مدل بیزین مشترک Joint Bayesian model

مدل بیزین مشترک، هر تصویر را به صورت جمع دو متغیر گوسین "x=μ+ε" در نظر می‌گیرد به طوری که μ ارائه‌گر یک شخصیت است و ε میزان تغییر درون‌فردی است. برای اینکه دو چهره x_2 و x_1 متعلّق به یک شخصیت یکسانی است احتمال‌های P(x_1 ,x_2) و P(x_1 ,x_2 |H_1) P(H_2)| محاسبه می‌شود. H_1 فرض اینکه دو چهره مشابه یکدیگر هستند و H_2 فرض اینکه دو چهره مخالف هم هستند. این توزیع‌ها می‌توانند به صورت یک توزیع نرمال نیز نوشته شوند به طوری که اجازه طراحی یک آزمون موثر ماکسیمم لایکلی هود (روش آماری برای برآورد کردن پارمترهای یک مدل آماری) را ایجاد می‌کنند.(Fredj HB) این روش از دادهِ آموزش استفاده می‌کند و ۸۹ درصد میزان شناسایی در LFW را بدست می‌آورد.

نرم‌افزار visionLabs

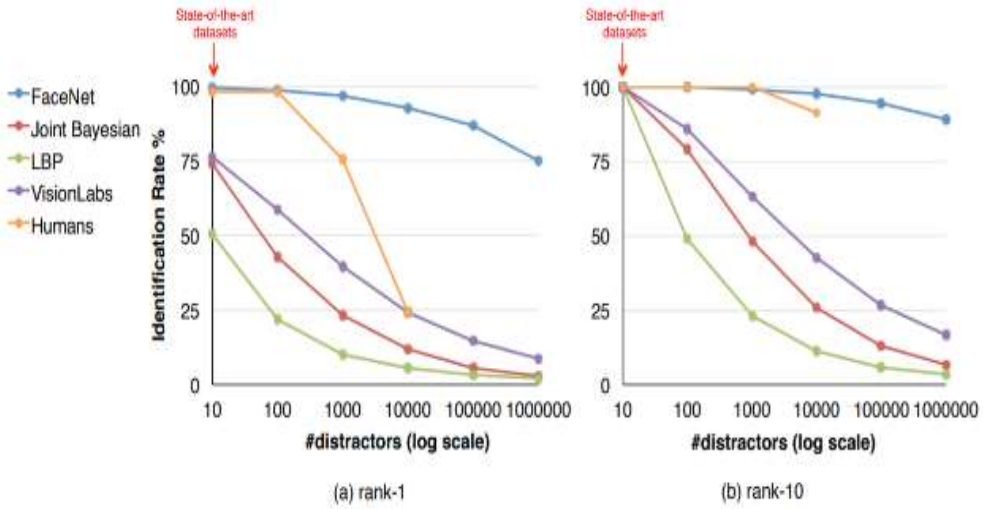
این نرم‌افزار نیز با استفاده از شبکه عصبی[1] یادگیری عمیق و همچنین توجه ویژه به موضوع شناسایی صورت، ترازبندی تصویر و استخراج ویژگی‌ها بنا شده است و با ۶ میلیون تصویر که مربوط به ۴۶۰ هزار نفر است آموزش دیده است(Chen C). همچنین برای محاسبه تشابه جفت‌های تصاویر از فاصله اقلیدسی نیز استفاده می‌نماید. این نرم‌افزار توانسته تشخیص چهره را با نِرخ ۹۳ درصد بر روی مجموعه داده‌ایی LFW بدست آورد.

گوگل FaceNet

گوگل FaceNet یکی از بهترین عملکرد را در بین چندین الگوریتم مورد مقایسه دارد و از روش‌های یادگیری عمیق[2] استفاده می‌نماید. این الگوریتم با استفاده از ۲۰۰ میلیون عکس از ۸ میلیون نفر آموزش دیده است در LFW ۹۹.۶ درصد به دست می‌یابد.

نتیجه: همانطور که در شکل شماره (۳) مشاهده می‌نمایید انسان توانایی تشخیص افراد را در بانک‌های اطلاعاتی بیش از ۱۰۰۰ تصویر را نداشته و فقط الگوریتم‌هایی که مبتنی بر یادگیری عمیق می‌باشند و با بانک داده‌ای چند میلیونی آموزش دیده‌اند، دارای نتایج قابل قبولی در بانک‌های اطلاعاتی حجیم می‌باشند و الگوریتم‌هایی که از رویکردهای محلی استفاده می‌نمایند با افزایش تعداد نمونه‌ها کیفیت شناسایی آن‌ها به صورت تصاعدی کاهش پیدا نموده است.

1. Neural Network
2. Deep Learning



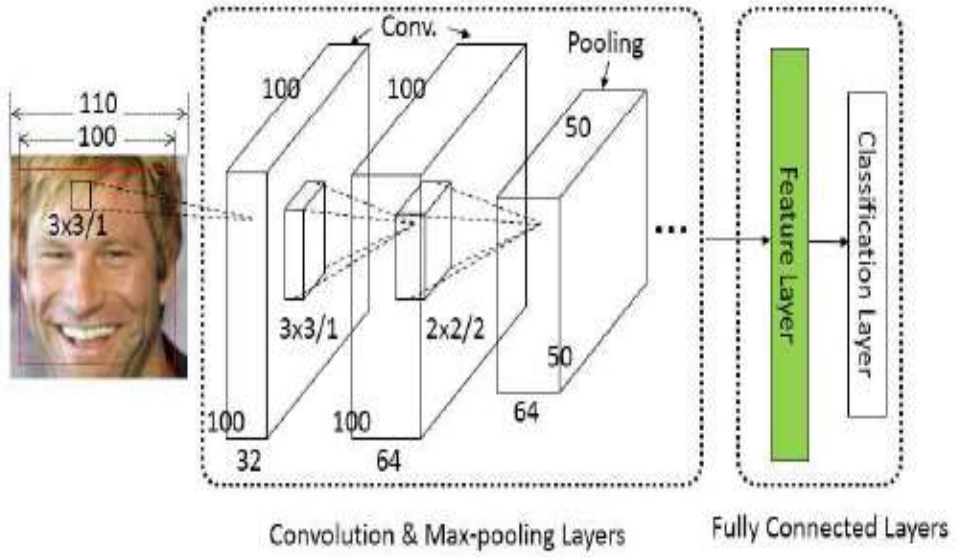
شکل شماره۳:نتایج حاصله از مقایسه الگوریتم‌ها

(Figure-3) Results from the comparison of algorithms

مقاله دوم: ارائه یک مکانیزم تشخیص چهره ترکیبی و هوشمند

این مقاله ابتدا به دو چالش عمده تشخیص چهره در بانک‌های اطلاعاتی حجیم می‌پردازد و چالش اصلی در شناسایی افراد را از دست دادن دقّت جستجو و افزایش پیچیدگی محاسباتی با اندازه مجموعه داده‌ای می‌داند که برای فائق آمدن بر مسئله اصلی که تعداد تصاویر بانک اطلاعاتی است از یک چهارچوب جستجوی چهره بازگشتی استفاده می‌نماید و در واقع مسئله جستجو را به دو مرحله تجزیه می‌نماید: (۱) یک مرحله فیلتر سریع، که از روش تقریب برای برگرداندن یک لیست کوتاه از نامزدها استفاده می‌کند که این کار را با استفاده از شبکه عصبی کانولوشن[1] که در شکل شماره (۴) ساختار آن نمایش داده شده است انجام می‌دهد و (۲) یک مرحله تنظیم مجدد، که لیست نامزدها را با سرعت کمتری مرتّب می‌کند، اما عمل مقایسه دوبه‌دو، در نتیجه جستجو دقیق‌تر ارائه می‌شود.

1. Convolution Neural Network



شکل شماره۴:ساختار شبکه عصبی پیچشی

(Figure-4) Convolution Neural Network

ساختار شبکه کانولوشن عمیق:

۱- لایه‌های کانولوشن

۲- لایه‌های نمایشی ویژگی

۳- یک معماری بسیار عمیق (در کل ۱۰ لایه کانولوشن) و فیلترهایی با پشتیبانی کوچک (۳× ۳) در نظر گرفته شده است. فیلترهای کوچک، تعداد پارامترهای یادگیری را کاهش می‌دهند و معماری بسیارعمیق غیرخطی بودن شبکه را افزایش می‌دهند. بر اساس این فرض اساسی که تصاویر چهره معمولاً روی یک منیفولد [1] با ابعاد کم قرار دارند، شبکه بُردار ویژگی‌های ۳۲۰بُعدی را تولید می‌کند.

ساختار کلی الگوریتم جستجو:

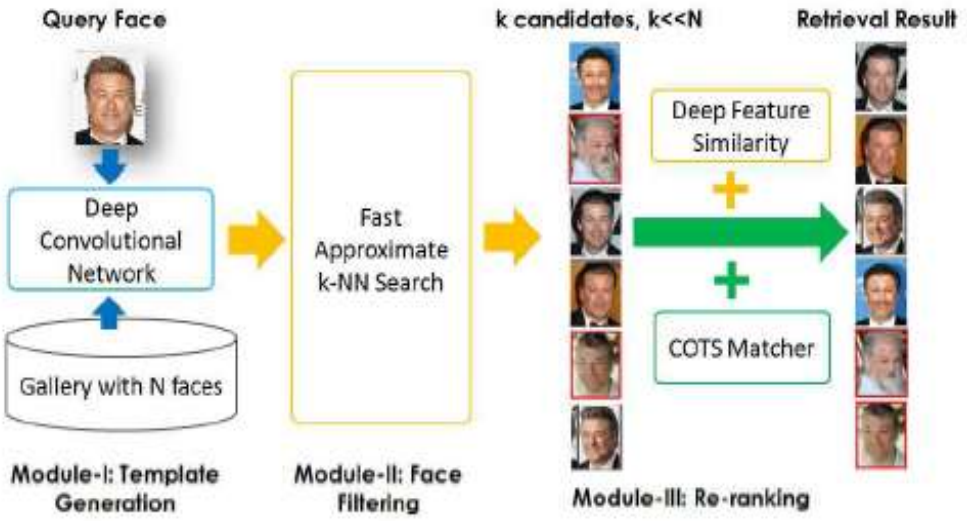
یک تصویر برای جستجو به عنوان ورودی داده می‌شود. سیستم جستجوی چهره کمک می‌کند که بتوان k برترین تصویر (یعنی به تعداد k عدد تصویر که بیشترین شباهت را دارد) در گالری تصاویر پیدا شود. برای اینکه بتوان جستجوی صورت را در تعداد تصاویر حجیم انجام داد از ساختارهای

1. Manifold

جستجوی چهره بازگشتی[1] استفاده می‌شود(Zhang R). شکل (۵) معماری جستجوی چهره ارائه شده شامل سه مرحله اصلی را تشریح می‌کند:

۲) مدل فیلتر کردن چهره، در واقع چهره‌ای که به عنوان ورودی داده شده را با تصاویر موجود در گالری مقایسه می‌کند. این فرآیند را با استفاده از مقداردهی‌های انجام شده بدست می‌آورد تا در نهایت k بالاترین نامزد های مشابه را بدست آورد .(k ≪ N)

۳) مدل رتبه‌بندی مجدد که امتیاز شباهت ویژگی‌های عمیق را با امتیازات از یک چهره COTS (نوعی مطابقت دهنده) ترکیب می‌کند تا ترتیب جدیدی از نامزدهای k ایجاد کند.



شکل شماره ۵– معماری بهینه جستجوی تصویر

(Figure-5) Optimal image search architecture

به صورت خلاصه می‌توان گفت: با استفاده از شبکه عصبی به تولید الگو می‌پردازیم. سپس با استفاده از روش نزدیکترین همسایه، نمونه‌های کاندید را انتخاب می‌کنیم و با استفاده از COTS شبیه‌ترین تصاویر را به عنوان کاندید انتخاب می‌نماییم(Zhang R).

1. Recursion

کالبد شکافی مفهومی

استخراج ویژگی[1]

عملکرد اصلی این مرحله، استخراج ویژگی‌های چهره‌ای است که در مرحله شناسایی چهره پیدا شده است. در این مرحله می‌بایست اِلمان‌های موجود در صورت را استخراج نمود و با استفاده از دسته‌بندی آنها، بُردار ویژگی‌های تصویر را تشکیل داد(Agarwal L).

تشخیص چهره[2]

در این مرحله می‌بایست با روش‌های کارآمد به مقایسه تشابه ویژگی‌های استخراج شده تصویر کاندید و تصاویر بانک اطلاعات بپردازیم و بر اساس بیشترین میزان تشابه تصمیم‌گیری نماییم.

روشbig-gak

این الگوریتم به منظور استخراج جامعه هدف از یک جامعه بزرگ که دارای نزدیک‌ترین ویژگی‌ها به سوژه اصلی است، با استفاده از الگوریتم کی پونت، سوژه را شناسایی می‌کند.

شناسایی چهره[3]

اولین گام، تشخیص وجود یا عدم وجود چهره انسان در تصویر است که این موضوع با ویژگی‌های کمّی و کیفی تصویر ارتباط مستقیم دارد. بدین نحو که ممکن است با نور، چرخش و یا خرابی تصویر نتوان صورت انسان را شناسایی نمود. پس منظور ما از شناسایی چهره، پیدا کردن صورتی در تصویر است که حداقل ویژگی‌های چهره انسان را داشته باشد(Agarwal L).

شناسایی چهره همواره یکی از موضوعات مورد مطالعه در علم پردازش تصویر بوده است. اگرچه این کار برای سیستم شبکه عصبی مغز انسان کار ساده‌ای است، لیکن در پیاده‌سازی نرم‌افزاری آن با چالش‌های متعددی روبرو هستیم. در حال حاضر چهار روش اصلی برای شناسایی چهره ارائه شده است(Young AW):

1. Feature Extraction
2. Face Recognition
3. Face Identification

روش‌های شناسایی چهره:

روش‌های مبتنی بر شناخت و آگاهی

این روش‌ها بر پایه مجموعه‌ای از قوانین و شناخت انسان از چهره انسان بنا شده‌اند. مثلاً یک چهره دارای چه اجزایی و با چه ویژگی‌هایی است.

روش‌های مبتنی بر ویژگی‌های تغییر ناپذیر

این روش‌ها برای شناسایی چهره، از ویژگی‌های استخراج شده از تصاویر بانک اطلاعات چهره انسان استفاده می‌نمایند.

روش‌های مبتنی بر قالب یا الگو

این روش‌ها با استفاده از تعیین میزان همبستگی قالب‌های از قبل تعریف شده چهره با تصویر، چهره انسان را شناسایی می‌نمایند.

روش‌های مبتنی بر ظهور یا نمود

این روش‌ها با استفاده از تکنیک‌های یادگیری ماشین نسبت به استخراج ویژگی‌ها و سپس شناسایی چهره اقدام می‌نمایند.

در کلیه روش‌های فوق، دلایل متعددی باعث افزایش و یا کاهش درصد موفقیت آن‌ها می‌گردد، لیکن می‌توان گفت بهترین روش، طرحی است که کلیه این روش‌های مذکور را به نحوی با یکدیگر ترکیب کند که از مزایای آن بهرمند شویم. می‌توان گفت تنها مکانیزمی که کلیه مزایای روش‌های فوق الذکر را دارد الگوریتم ویولاجونز است.

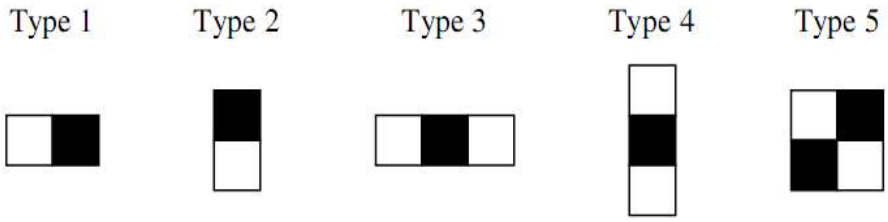
الگوریتم ویولا جونز:

این الگوریتم پرکاربردترین و قوی‌ترین الگوریتم شناسایی چهره انسان است که در سال‌های ۲۰۰۱ تا ۲۰۰۴ توسط Paul Viola و Michael Jones منتشر شد که از چهار بخش اصلی زیر تشکیل می‌شود (Hirzi MF):

۱- ویژگی های‌هار ۲- آدابوست ۳- انتگرال تصویر ۴- ویژگی آبشاری

۱- ویژگی های‌هار

برای استخراج ویژگی‌ها از قالب‌های مربّع و مستطیل به شکل (۶) که همان تابع مقیاس یا موجک‌های‌هار می‌باشند استفاده می‌گردد.



شکل شماره۶–موجک‌های هار

(Figure-6) Haar Wavelet

الگوریتم ویولا جونز از یک پنجره [1]با اندازه دلخواه شما مثلاً ۱۹ در ۱۹ برای طی این مراحل و قرار دادن مستطیل‌ها بر روی صورت استفاده می‌کند که با در نظر گرفتن تعداد و اندازه‌های مختلف این ویژگی‌ها (مستطیل‌ها) با یک حساب سرانگشتی به این نتیجه می‌رسیم که برای محاسبه نتیجه نهایی برای هر یک از این پنجره‌های ۱۹ در ۱۹ نیازمند انجام هزار بار محاسبه است که منطقاً تعداد بالایی است و هزینه و زمان زیادی را برای هر عکس خواهد گرفت. در صورت استفاده از تصاویری با عرض و طول یکسان که توانی از عدد ۲ باشند می‌توانیم از تبدیل فوریه سریع [2]یا FFT استفاده نماییم(Hirzi MF).

۲- آدابوست

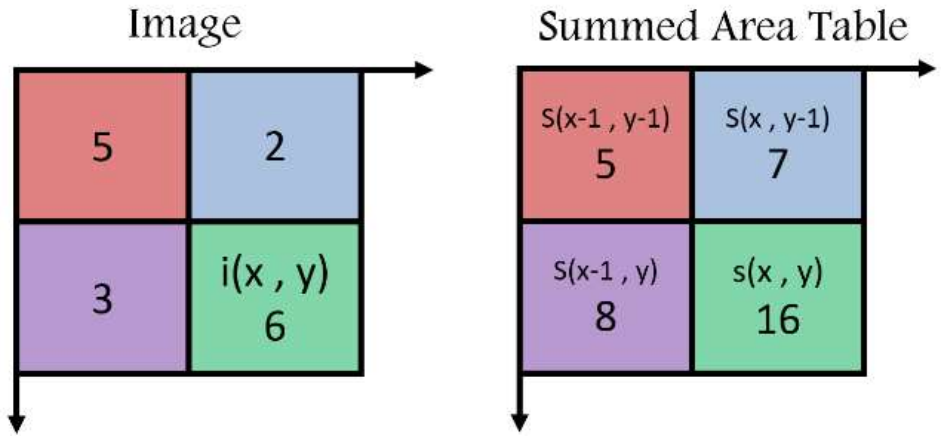
آدابوست یک الگوریتم یادگیری ماشین ۲ کاناله و از خانواده الگوریتم‌های تقویت‌کننده[3] است که وظیفه آن پیدا کردن ویژگی‌های کاربردی از میان هزاران ویژگی استخراج شده است. بعد از شناسایی ویژگی‌های مختلف، از ترکیب حساب شده‌ای از آن برای برآورد و تصمیم‌گیری در مورد هر پنجره استفاده می‌شود و مشخص می‌گردد که آیا هر یک از آن‌ها در خود دارای صورت هستند یا خیر. آدابوست نسبت به داده‌های نویزی[4] و پَرت[5] حساس است. بنابرین در مرحله یادگیری و تعیین ویژگی‌ها فقط در بانک اطلاعات تصویر صورت ثبت می‌شود که بتوان بدون نویز استفاده نمود تا ویژگی‌های

1. Window
2. Fast Fourier Transform
3. Boosting
4. Noise
5. Outliers

استخراج شده دارای کیفیت مناسبی باشند. این الگوریتم کلاس‌بندی نسبت به مشکل بیش‌برازش[1] از بیشتر الگوریتم‌های یادگیری برتری دارد و استفاده از نمونه‌های مجهول بیشتر باعث افزایش دقّت آن می‌شود که این در تشکیل بانک اطلاعات اولیّه به منظور آموزش بسیار تاثیرگذار است(Hirzi MF).

۳- انتگرال تصویر

تصویر یکپارچه، یا جدول محدوده مجتمع[2]، به منظور ارزیابی سریع‌تر ویژگی‌ها (که در بخش ابتدایی معرفی شد) استفاده می‌شود. همانطور که در تصویر زیر می‌بینید، در یک تصویر یکپارچه مقدار پیکسل در مکان برابر با جمع مقادیر پیکسل‌های بالا و چپ پیکسل x,y است(Hirzi MF).



شکل شماره۷:یکپارچه‌سازی پیکسل‌ها در تصویر

(Figure-7) Summed Area Table

۴- ویژگی آبشاری

هر پنجره ۱۹ در ۱۹ پیکسلی نیاز به پردازش ۲۰۰ ویژگی دارد که از اجرای الگوریتم آدابوست به دست می‌آید. تصوّر کنید که شما یک تصویر با ابعاد ۶۴۰ در ۴۸۰ دارید، شما نیاز دارید تا این ابعاد را به مربع‌های ۱۹ در ۱۹ پیکسلی تقسیم کرده و با پردازش هر بخش (در نظر داشته باشید که هر مربع ۲۰۰ پردازش لازم دارد) تشخیص دهید که آیا چهره‌ای در تصویر وجود دارد یا خیر؟ . اینجاست که Cascading وارد شده و این امکان را می‌دهد که این فرآیند را سریع‌تر و مفیدتر انجام دهید. در این روش ما ۲۰۰ ویژگی هر مربع ۱۹ در ۱۹ را به دسته بندی‌های مختلف تقسیم می‌کنیم. بدین گونه که

1. Over Fitting
2. Summed Area Table

به عنوان مثال: ۲ ویژگی در دسته‌بندی اوّلی، ۹ ویژگی در دسته بندی دوم، ۷ ویژگی در دسته‌بندی سوم و به همین روال ادامه دارد(Hirzi MF).

دلایل استفاده از الگوریتم ویولا جونز:

۱- استفاده از آدابوست که نسبت به بیش‌برازش نسبت به اکثر الگوریتم‌های یادگیری ماشین برتری دارد.

۲- آموزش دسته الگوریتم‌های بوستینگ با استفاده از افزایش نمونه‌های مجهول، می‌تواند باعث افزایش دقّت در آن‌ها شود.

۳- استفاده از تکنیک محدوده مجتمع باعث افزایش سرعت شناسایی می‌گردد.

۴- استفاده از ویژگی آبشاری برای افزایش سرعت.

تجزیه و تحلیل پژوهش

جامعه آماری:

یکی از نکات مهم و تاثیرگذار در طراحی سامانه تشخیص چهره که بر روی انتخاب الگوریتم‌های مورد استفاده در تشخیص چهره بسیار موثر است ویژگی‌های منحصر به فرد هر مجموعه داده‌ای است که می‌توان آن را در چهار بخش مورد تجزیه و تحلیل قرار داد(Nagrani A):

ویژگی‌های مجموعه داده‌ای:

الف) تعداد تصاویر

ب) تعداد شخصیت‌های منحصر به فرد

د) نسبت تعداد تصاویر رنگی به سیاه و سفید

ه) میانگین اندازه تصاویر

در صورت تعیین دقیق ۴ ویژگی اصلی مجموعه داده‌ای تصاویر، می‌توان به صورت موثّرتر نسبت به چیدمان الگوریتم‌های مورد استفاده به منظور تشخیص چهره اقدام نمود. در مجموعه داده‌ای که ما پیشنهاد داده‌ایم هر تصویر مربوط به یک شخصیّت است. اندازه مجموعه داده‌ای مورد آزمایش ۵۱۰ هزار تصویر است که نِرخ تعداد شخصیّت‌ها در این بانک اطلاعاتی نسبت به تعداد تصاویر برابر با ۱ است.

تصاویر مجموعه داده‌ای ما شامل ۵۱۰ هزار تصویر پرسنلی است که با استفاده از الگوریتم شناسایی چهره مورد استفاده، تقریباً ۷۰ درصد صورت‌ها در این تصاویر قابل استخراج است و سایر

ویژگی‌های صورت که در الگوریتم پیشنهادی تشخیص چهره استفاده می‌گردد، بر اساس قدرت الگوریتم و نحوه پیاده‌سازی آن، درصدهای مختلف بازدهی دارند که این به دلایل مختلفی مانند: داشتن عینک، گذاشتن کلاه، داشتن ماسک تنفسی بر روی صورت، داشتن مقنعه و روسری برای تصاویر چهره بانوان، داشتن چشم‌بند، گذاشتن هدبند بر روی پیشانی و لباس‌های محلی مخصوص سر و هر چیزی که باعث بشود قسمتی ازچهره فرد معلوم نباشد وابسته است(Nagrani A).

جدول شماره۱:اطلاعات بانک اطلاعاتی تصویر

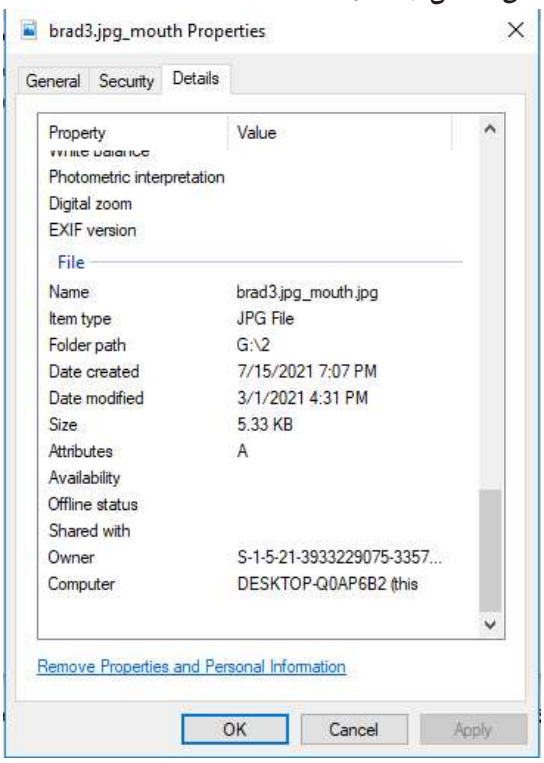
تعداد تصاویر	۵۱۰ هزار
تعداد شخصیت‌های منحصر به فرد	۵۱۰ هزار
نسبت تعداد تصاویر رنگی به سیاه و سفید	۷۳
میانگین اندازه تصاویر	۱۰۰×۱۵۰

استخراج ویژگی‌های تصویر

از یک تصویر که شامل صورت انسان است می‌توان حجم زیادی از اطلاعات ارزشمند را شناسایی نمود، لیکن در بیشتر روش‌های استخراج ویژگی، صرفاً به ویژگی‌هایی از قبیل: مدل صورت، چشم، دهان و دماغ و نحوه جاگیری آن‌ها در صورت پرداخته می‌شود که با یک بررسی اِجمالی می‌توان گفت این ویژگی‌ها تقریباً ۲۰ درصد از کل اطلاعات موجود در تصویر را تشکیل می‌دهند(Zhang Q).

ما در این مقاله با یک نگاه جدید به استخراج ویژگی‌های تصاویر شامل چهره انسان می‌پردازیم و با توجه به ماهیت آن‌ها ویژگی‌های تصویر چهره را بر اساس نوع اطلاعات قابل استخراج می‌توانیم به پنج سطح تقسیم می‌نماییم:

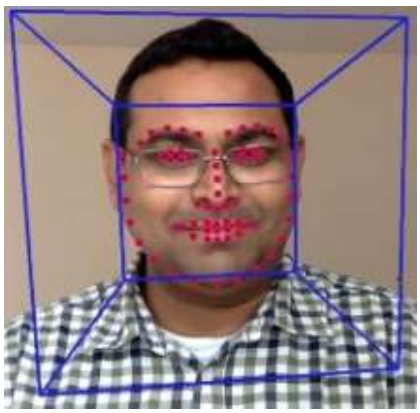
سطح ۱: اطلاعات عکس، شامل زمان و مکان



شکل شماره۸:اطلاعات زمان و مکان تصویر

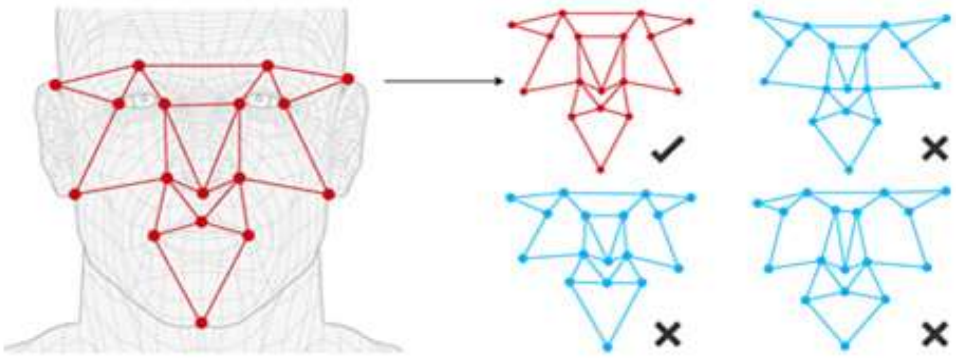
(Figure-8) Image time and place information

سطح ۲: مدل هندسی سر و شانه و فاصله ذوزنقه‌ای



شکل شماره۹:مدل هندسی سر و شانه

(Figure-9) Geometric model of head and shoulders



شکل شماره ۱۰:فاصله ذوزنقه ای

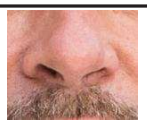
(Figure-10) Trapezoidal distance

سطح ۳: سن، جنسیّت و حالت چهره

شکل شماره ۱۱:سن، جنسیّت و حالت چهره

(Figure-11) Age, gender and facial features

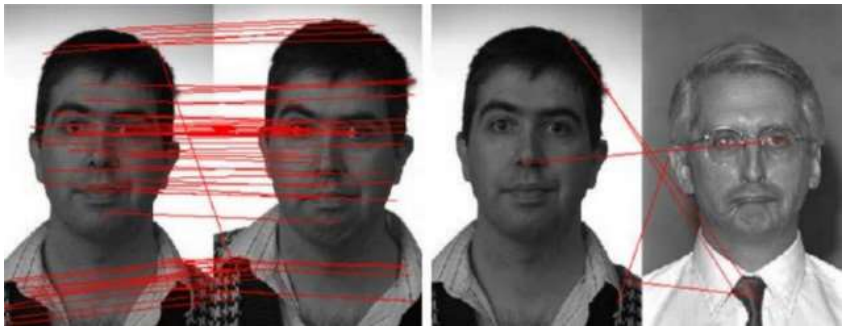
سطح ۴: ویژگی‌های اختصاصی صورت انسان شامل: مدل چشم، ابرو، بینی، دهان، مو، ریش و سبیل، همچنین رنگ پوست صورت، رنگ و حجم موی سر

			
چشم راست	بینی	چشم چپ	دهان
بادامی	مدل یونانی	بادامی	مدل لب عریض
رنگ سبز	سبیل دارد	رنگ سبز	سبیل دارد

شکل شماره ۱۲:استخراج ویژگی

(Figure-12) Feature extraction

سطح ۵: شامل نقاط کلیدی صورت انسان که توسط الگوریتم نقاط کلیدی استخراج می‌شود.



شکل شماره۱۳: استخراج نقاط کلیدی

(Figure-13) Extract key points

بعد از استخراج ویژگی‌های جدول شناختی، استخراج ویژگی که شامل اطلاعات زیر است تشکیل گردد:

۱- نوع ویژگی استخراج شده

۲- گروه‌بندی ویژگی استخراج شده

۳- ترتیب‌بندی گروه‌ها نسبت به یکدیگر

۴- درصد کارآیی الگوریتم استفاده شده در استخراج ویژگی

۵- تعداد نمونه‌های قرار گرفته در هر گروه

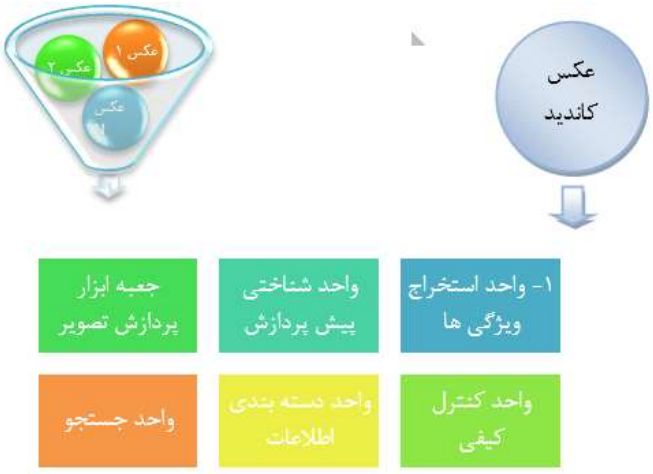
جدول شماره۲:جدول ویژگی‌های تصویر

رنگ مو	سفید	بلوند	بلوند تیره	قهوه‌ای	مشکی
۸۳ درصد	۱	۲	۳	۴	۵
	۱۰۰۰	۶۰۰۰	۲۴۰۰۰	۳۷۳۰۰۰	۱۱۰۰۰
سن تقویمی ۶۵ درصد	نوزاد	کودک	جوان	بزرگسال	سالمند
	۱	۲	۳	۴	۵
	۱۰۰۰	۱۰۰۰	۲۵۵۰۰۰	۹۶۰۰۰	۱۲۰۰۰
مدل لب ۵۳ درصد	زخیم	نازک	غنچه	عریض	
	۱	۲	۳	۴	
	۲۲۷۰۰۰	۵۳۰۰۰	۲۰۰۰	۴۳۰۰۰	
حجم مو ۵۲ درصد	تاس	کم مو	متوسط	پرمو	
	۱	۲	۳	۴	
	۲۰۰۰	۳۰۰۰۰	۱۸۲۰۰۰	۴۶۰۰۰	

پوست صورت ۹۲درصد	سفید	صورتی	قهوه ای	قهوه‌ای تیره	مشکی
	۱	۲	۳	۴	۵
	۹۰۰۰	۹۲۰۰۰	۳۲۰۰۰۰	۲۳۰۰۰	۱۵۰۰۰
فاصله ذوزنقه ای ۸۱ درصد	کلاس ۱	کلاس ۲	کلاس ۳	کلاس ۴	
	۱	۲	۳	۴	
	۲۹۰۰۰	۴۱۰۰۰	۳۱۸۰۰۰	۴۶۰۰۰	
حالت چهره ۵۳ درصد	نرمال	خوشحال	ناراحت	عصبانی	ترس
	۱	۲	۳	۴	۵
	۲۲۳۰۰۰	۲۳۰۰۰	۴۰۰۰۰	۲۰۰۰	۱۵۰۰

تشخیص چهره انسان با استفاده از الگوریتمBig-GAK

همانطور که در ابتدای مقاله توضیح داده شد، مشکل اصلی روش‌ها و مکانیزم کنونی تشخیص چهره، ضعف در معماری روش‌های تشخیص چهره است که موجب گردیده نتوان از کلیه امکانات و تکنیک‌های موثر در تشخیص چهره به صورت کارا استفاده نمود. ما با ارائه الگوریتم تشخیص چهره BIG-GAK این اِمکان را فراهم آورده‌ایم که با استفاده از تکنیک‌های پردازش تصویر، هوش مصنوعی و علوم شناختی بتوان یک مکانیزم بهینه و اِنعطاف‌پذیر را طراحی نماییم. این الگوریتم بر اساس تعامل چند واحد مستقل طراحی و پیاده‌سازی شده است:

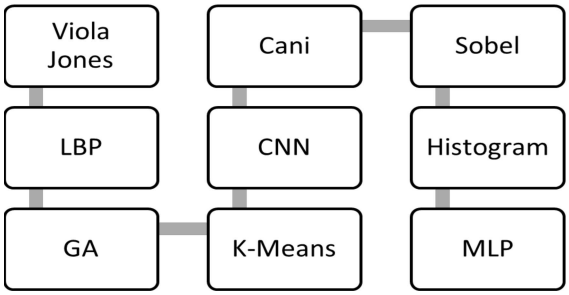


شکل شماره۱۴–اجزاء تشکیل دهنده معماری الگوریتم big-GAk

(Figure-14) Architectural components of the big-GAk algorithm

۱- جعبه ابزار هوشمند پردازش تصویر

این واحد نسبت به ارائه الگوریتم‌های استخراج ویژگی‌های سطح یک تا پنج تصویر، برابر نظر و گام‌هایی که در واحد استخراج ویژگی تعیین می گردد اقدام می‌نماید.



شکل شماره۱۵: مراحل استخراج ویژگی

(Figure-15) Feature extraction steps

۲- واحد استخراج ویژگی

این واحد نسبت به ارائه مکانیزم استخراج ویژگی‌های سطح یک تا پنج تصویر اقدام می‌نماید و برای آماده‌سازی تصاویر جهت استخراج هر یک از ویژگی‌ها از واحد شناختی پیش‌پردازش استفاده می‌نماید.

این بخش بر اساس ویژگی‌های الگوریتم انتخابی از جعبه ابزار پردازش تصویر، پیش‌پردازش‌های مورد نیاز همان الگوریتم را که باعث افزایش کارآیی واحد استخراج ویژگی می گردد صرفاً انجام می‌دهد و کارآیی آن توسط واحد کنترل کیفی ارزیابی شده و نتیجه جهت تصمیم‌گیری به واحدهای استخراج ویژگی و پیش‌پردازش منعکس می گردد.

به عنوان مثال؛ پیش‌پردازش رفع نویز که باعث بهبود عملکرد الگوریتم شناسایی صورت انسان (ویولا جونز) می‌گردد در استفاده از الگوریتم استخراج لبه تصاویر باعث کاهش کارآیی این الگوریتم می‌گردد. پس در هر مرحله، در صورت انجام پیش‌پردازشی که برای الگوریتم مورد استفاده فعلی تاثیر مخرب دارد از تصویر آماده شده صَرف نظر می‌گردد و تصویر اصلی مجدداً انتخاب و پیش‌پردازش بر روی آن انجام می‌گردد.

۳- واحد شناختی پیش‌پردازش

واحد شناختی پیش‌پردازش، تغییر و اصلاح ویژگی‌های کیفی وکمّی تصاویر با استفاده از قواعد شناختی است. کارآیی سیستم‌های تشخیص چهره و استخراج ویژگی‌های تصویر به شدّت به پارامترهایی مانند: ابعاد، رنگ، میزان نوردهی، مگاپیکسل دوربین عکاسی، نویز تصاویر، فاصله شخص از دوربین عکاسی، رنگی و یا سیاه سفید بودن تصاویر اصلی وابسته است.

این تصوّر که اصلاح و یا عدم اصلاح و یکسان‌سازی ویژگی‌های مذکور بر روی کلیه تصاویر بانک اطلاعات موجب بهبود کارآیی الگوریتم‌های تشخیص چهره می‌شود، اشتباه است. لذا بر اساس اینکه هر بخش از واحد استخراج ویژگی‌ها چه ویژگی را و به وسیله چه الگوریتمی استخراج می‌نماید، می‌بایست اصلاحات و تغییرات بر روی عکس اولیه اِعمال و در اختیار واحد استخراج ویژگی‌ها قرار گیرد. ضمناً این واحد بر اساس کیفیت استخراج ویژگی‌ها، که از واحد کنترل کیفی دریافت می‌نماید نسبت به اصلاح چگونگی تغییر و اصلاح ویژگی‌های کیفی و کمّی اقدام می‌نماید.

۴- واحد کنترل کیفی

این واحد، کیفیت ویژگی‌های استخراج شده را به صورت با ناظر و بدون ناظر ارزیابی نموده و درصد موفقیت را در استخراج هر یک از ویژگی‌ها در اختیار واحدهای پیش‌پردازش، استخراج ویژگی و دسته‌بندی جهت بهینه‌گزینی قرار می‌دهد. این واحد برای هر الگوریتمی دارای جدول شناختی است که در آن پیش پردازش‌ها و الگوریتم‌های جایگزین تعیین شده است.

به عنوان مثال: در استخراج ویژگی رنگ عدسی چشم، نیازمند به استفاده از الگوریتم‌های تشخیص لبه برای جداسازی قرنیه چشم می‌باشیم که بهترین تکنیک، استفاده از مشتق‌گیری سوبل[1] عمودی است و در صورت عدم کارآیی، می‌توان از الگوریتم استخراج لبه شبکه عصبی CNN جایگزین نمود.

۵- واحد شناختی دسته‌بندی ویژگی‌ها

این واحد از دو بخش دسته‌بندی و بهینه‌گزینی تشکیل شده است. وظیفه بخش اول، دسته‌بندی ویژگی‌های استخراج شده به صورت با ناظر و شناختی است و وظیفه بخش دوم، بهینه‌گزینی است که این کار می‌بایست بر روی جدول دسته‌بندی ویژگی‌ها که شامل نوع ویژگی، گروه‌ها و تعداد هر گروه و درصد کیفیت استخراج ویژگی است. با استفاده از الگوریتم‌های دسته مسائل بهینه‌سازی می‌توان کوچک‌ترین جامعه با بیشترین احتمال موفقیت را شناسایی نمود که ما در این قسمت از الگوریتم ژنتیک[2] (راهکار کوله پشتی صفر و یک[3]) استفاده می‌نماییم.

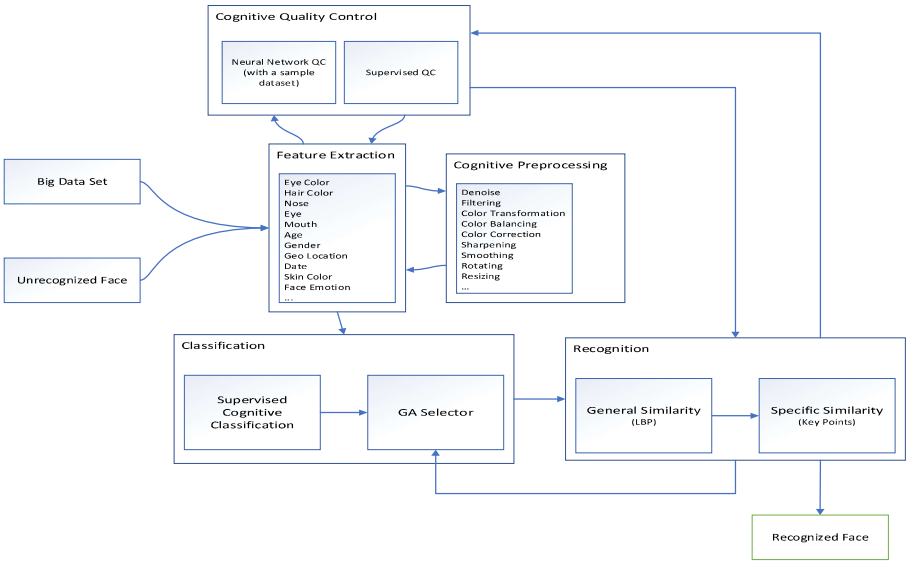
1. Sobel
2. Genetic Algoritm
3. Knapsack Problem

روش حل مسئله کوله پشتی صفر و یک با الگوریتم ژنتیک

یکی از بهترین روش‌ها برای حل دسته مسائل بهینه‌سازی، الگوریتم ژنتیک است. این الگوریتم در مسائلی که فضای جستجوی بزرگی داشته و یا دارای فضای فرضیه پیچیده‌ای باشند به نحوی که تاثیر اَجزاء آن در فرضیه کلی ناشناخته باشد مناسب است.

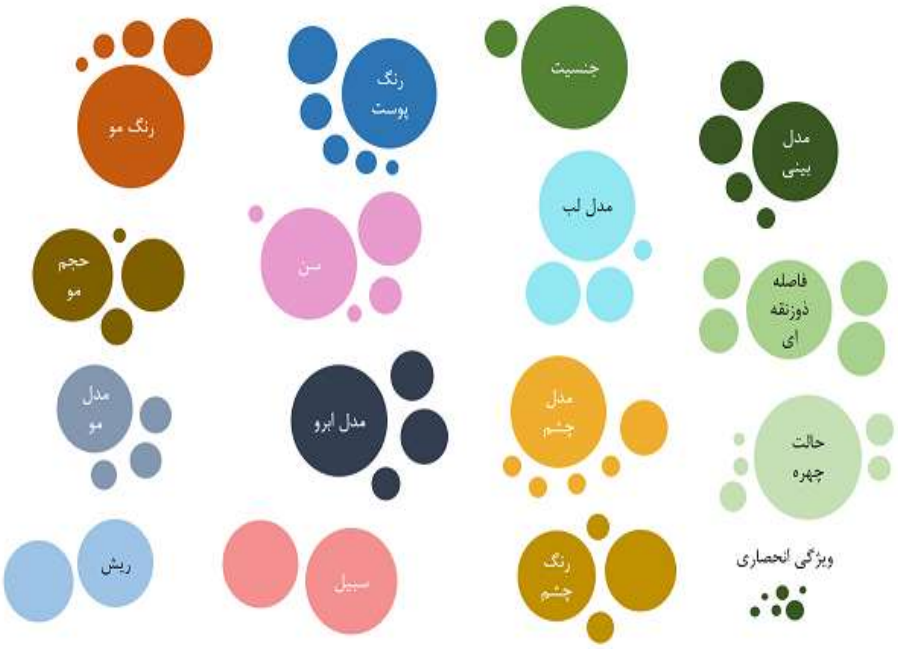
یکی از مسائل معروف و مهم بهینه‌سازی که می‌تواند دارای فضای فرضیه بزرگ و پیچیده باشد مسئله کوله پشتی صفر و یک است که در آن می‌خواهیم از میان تعداد محدودی انتخاب‌ها، اشیائی را انتخاب نماییم که ارزش انتخاب‌های ما بیشینه و وزن کوله پشتی ما از حداکثر وزنی که کوله پشتی می‌تواند تحمّل نماید بیشتر نباشد.

بعد از تعریف پارامترهای ورودی مسئله شامل: وزن، ارزش اشیاء و حداکثر وزنی که کوله پشتی می‌تواند تحمّل کند، می‌بایست تعداد جامعه هدف اولیه که مقادیر آن به صورت تصادفی تولید می‌شوند را تعیین نمود و بعد از ارزیابی وزن، ارزش این جامعه و انتخاب بهترین جواب‌های تولید شده در هر مرحله نسبت به اِعمال قوانین پایه‌ای الگوریتم ژنتیک مثل ترکیب و جهش اِقدام نموده و جامعه جدیدی را تولید می‌نمائیم. این کار را برابر تعداد دفعات محدودی انجام می‌دهیم که بعد از چندین مرحله تکرار به یک فضای حالت و جمعیتی با حالت بهینه که حداکثر ارزش جمع‌آوری شده است، اما شرط حداکثر وزن قابل تحمّل کوله پشتی رعایت شده باشد می‌رسیم.



شکل شماره۱۶: مکانیزم انتخاب ویژگی‌ها

(Figure-16) Feature selection mechanism



شکل شماره۱۷:جدول ویژگی‌ها

(Figure-17) Features table

متغییرها

v ارزش موارد همان درصد موفقیت استخراج ویژگی‌ها است.
ω وزن موارد همان تعداد تصاویر هر گروه از ویژگی‌های استخراج شده است.
W حداکثر تعداد تصاویری که الگوریتم LBP در آن عملکرد خوبی دارد و از کارهای مرتبط استفاده شده است که در این کار ۰۰۰۱ در نظر گرفته شده است.

υ_i: ارزش مورد i اُم

ω_i: وزن مورد i اُم

W: وزن کل قابل حمل

x_i: متغیر باینری ← آیا مورد i اُم انتخاب شده است یا نه؟

فرمول‌های (۱)، (۲) و (۳) مربوط به تابع تناسب است:

$$\max \sum_i [\upsilon_i x_i] \quad (۱)$$

$$\sum_i [\omega_i x_i] \le W \quad (۲)$$

$$x_i \in \{0,1\} \quad (۳)$$

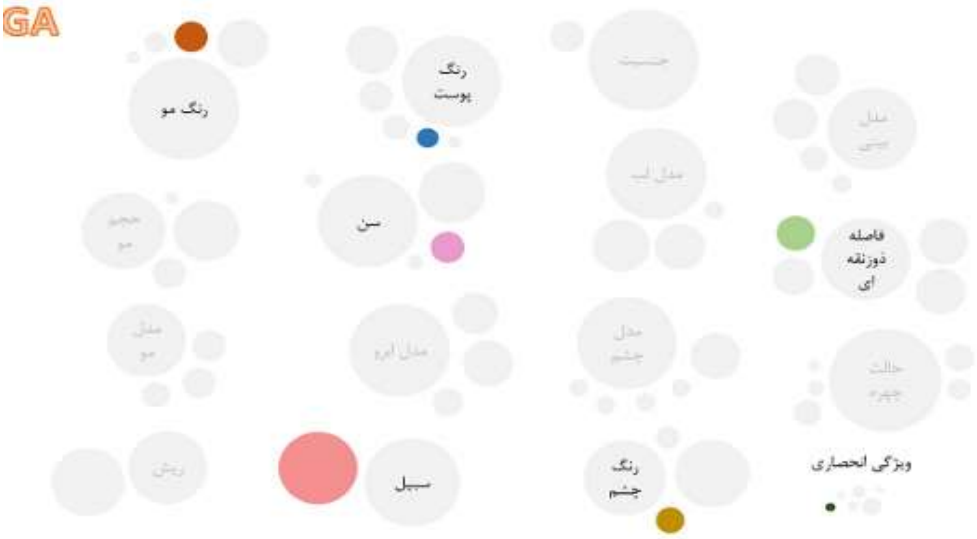
فرمول‌های (۴)، (۵) و (۶) مربوط به تابع جریمه است:

$$\upsilon\{g\leq g_0\}=\max(g/g_0-1,0) \quad (۴)$$

$$\upsilon\{g\leq g_0\}=\max(1-g/g_0,0) \quad (۵)$$

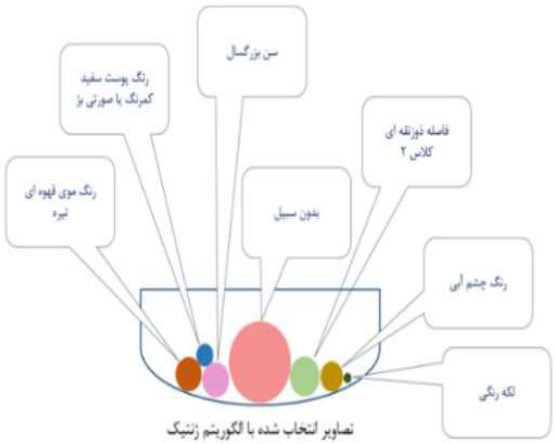
$$\upsilon\{g\leq g_0\}=|g/g_0-1| \quad (۶)$$

در شکل زیر نمونه‌های برگزیده با استفاده از الگوریتم ژنتیک را مشاهده می‌نمایید:



شکل شماره ۱۸: نمونه‌های برگزیده با استفاده از الگوریتم ژنتیک

(Figure-18) Selected samples using genetic algorithm



شکل شماره۱۹: جامعه هدف منتخب

(Figure-19) Selected target community

در پایان این مرحله جامعه هدف منتخب با کمترین هزینه و بالاترین احتمال موفقیت با ۳۰ بار تکرار محاسبه و آماده برای عملیات محاسبه تشابه عمومی می‌شود.

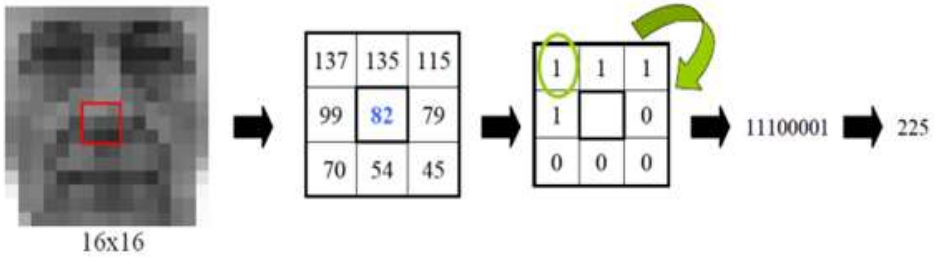
۶- واحد انتخاب‌کننده (جستجوگر)

این مرحله، مرحله پایانی و مهم‌ترین مرحله الگوریتم BIG-GAK است، زیرا مرحله تشخیص هویّت فرد کاندید در این مرحله انجام می‌شود.

دسته الگوریتم‌های شناسایی عمومی، تعداد زیادی هستند که انتخاب بهترین آن‌ها با توجه به بانک اطلاعات و تست با ناظر امکان‌پذیر می‌شود، همچنین ماهیت‌های این الگوریتم‌ها به نحوی است که به دنبال ویژگی‌های خاص و تمایز دهنده چهره انسانی نمی‌گردند، بلکه چهارچوب اصلی چهره فرد را به تصویر می‌کشند. پس با استفاده از آن‌ها افراد مشابه عمومی را پیدا می‌نماییم و در قسمت بعد با استفاده از الگوریتم‌هایی که به دنبال ویژگی‌های خاص در چهره افراد می‌گردند نسبت به تشخیص دقیق کاندیدای مورد نظر اقدام می‌نماییم.

مرحله ۱: محاسبه تشابه عمومی با استفاده از الگوریتم LBP

الگوی باینری محلی در ابتدا به عنوان توصیف‌گر بافت عمومی معرفی شد. این عملگرد با مقایسه همسایگی ۳ در ۳ هر پیکسل از یک تصویر با خود آن پیکسل و در نظر گرفتن نتیجه به صورت یک عدد دودویی، به آن پیکسل یک برچسب نسبت می‌دهد. در نشریات مختلف، مقادیر دایره‌ای به صورت صفر و یک و در جهت عقربه‌های ساعت یا عکس آن خوانده می‌شود. در ادامه نتیجه به صورت عدد باینری خواهد بود که از بالا سمت چپ به صورت ساعت‌گَرد بدست می‌آید که در شکل زیر نشان داده شده است:



شکل شماره ۲۰: الگوی باینری محلی بر روی چهره انسان

(Figure-20) Local binary pattern on human face

به بیان دیگر برای یک پیکسل دلخواه LBP(x_c,y_c) به صورت مجموعه‌ی منظمی از مقایسه‌های مقادیر پیکسلی، بین پیکسل مرکزی و پیکسل‌های همسایگی مقایسه می‌شوند. برچسب دهدهی نهایی نتیجه شده از این کلمه ۸ بیتی به صورت فرمول شماره (۷) بیان می‌شود:

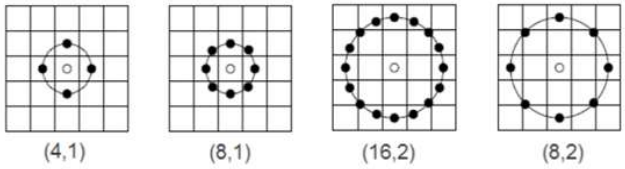
$$LBP(x_c,y_c)=\sum_(n=0)^7 ⟦s(I_n-I_c)2^n⟧ \qquad (۷)$$

به این صورت که I_c مربوط به ارزش خاکستری پیکسل مرکزی(x_c,y_c) و I_n ارزش خاکستری از ۸ پیکسل اطراف آن و تابع s(x) به صورت فرمول شماره (۸) تعریف می‌شود:

$$s(x)=\{■(1 \quad \text{if } x\geq 0@0 \quad \text{otherwise})┤ \qquad (۸)$$

عملگر الگوی باینری محلی برای بررسی بافت‌ها در مقیاس‌های مختلف گسترش داده شده است، به گونه‌ای که از همسایگی‌هایی با اندازه مختلف استفاده می‌کند. این کار با استفاده از دایره همسایگی و درون‌یابی خطی از مقادیر پیکسل هر شعاع و تعداد نمونه‌ها در همسایگی قابل انجام است.

شکل زیر چند نمونه از نقاط نمونه بَرداری‌های مختلف و شعاع آن‌ها را نشان می‌دهد، (P,R) به این معنی است که P نقطه با شعاع جهت نمونه بَرداری انتخاب شده است:



شکل شماره۲۱:نقاط نمونه بَرداری‌های مختلف

(Figure-21) Different sampling points

دلیل اینکه چرا ۴ نقطه عمودی و افقی انتخاب شده این است که چهره، شامل لبه‌های افقی و عمودی بیشتری نسبت به قطرهاست.

در هنگام محاسبه عملیات پیکسل در همسایگی N×Nدر مرز تصویر، بخشی از این پوشش که از لبه تصویر بیرون است غیر فعال می‌گردد. در چنین موقعیت‌هایی تکنیک‌های پَهن کردن مختلفی از قبیل پَهن کردن تصویر به صورت صفر کردن و تکرار عناصر مرزی یا روش آینه‌ای استفاده می‌شود. با وجود این در مورد عملگر الگوی باینری محلی، مرز بحرانی تعریف شده توسط شعاع R از عملگر دایره‌ای با استفاده از روش پَهن کردن حل نمی‌شود. به جای آن عملگر از پیکسل (R.R)تصویر شروع به کار می‌کند. مزیّت هیستوگرام نهایی الگوی باینری محلی این است که تحت تاثیر مرزها

نخواهد بود. اگرچه اندازه برچسب تصویر الگوی باینری محلی به اندازه (Width-R)×(Height-R) پیکسل کاهش پیدا خواهد کرد.

اگر در این مرحله کیفیت نتیجه نهایی که به مرحله بعد داده می‌شود خوب ارزیابی نگردد قسمت کنترل کیفی از جدول شناختی خود استفاده خواهد نمود.

جدول شماره۴: جدول شناختی جستجو تشابه عمومی

۱	LBP	الگوی باینری محلی Local Binary Pattern
۲	HOG	هیستوگرام گرادیان
۳	TF	جدا کننده فیشر آموزش دیده Fischer Trained
۴	PCA	تحلیل مولفه‌های اصلی Principal Component Analysis
۵	SVM	ماشین بُردار پشتیبان
۶	NN	شبکه عصبی مصنوعی Neural Network
۷	Naïve Bayes	شبکه کلاس‌بندهای نایو بیز
۸	Hidden Markov	مدل مارکوف مخفی
۱	LBP	الگوی باینری محلی Local Binary Pattern
۲	HOG	هیستوگرام گرادیان
۳	TF	جدا کننده فیشر آموزش دیده Fischer Trained
۴	PCA	تحلیل مولفه‌های اصلی Principal Component Analysis
۵	SVM	ماشین بُردار پشتیبان
۶	NN	شبکه عصبی مصنوعی Neural Network
۷	Naïve Bayes	شبکه کلاس‌بندهای نایو بیز
۸	Hidden Markov	مدل مارکوف مخفی

در این مرحله با استفاده از الگوریتم الگوی باینری محلی به استخراج ویژگی‌های صورت پرداخته و با استفاده از روش حداقل مربعات میزان تشابه را محاسبه می‌نماییم و بعد از محاسبه تشابه بین

تصویر کاندید با بانک تصاویر برگزیده شده توسط قسمت به گزینی واحد دسته بندی، ۱۰ درصد از بالاترین تشابهات، انتخاب و به واحد بررسی تشابه خصوصی تحویل می‌گردد.

مرحله ۲: محاسبه تشابه خصوصی با استفاده از الگوریتم Key-Points

دسته الگوریتم‌های نقاط کلیدی[1] خود نیز می‌تواند ویژگی‌های مختلفی را استخراج نماید که این ویژگی‌ها به شرح ذیل می‌باشند که کلیه این ویژگی‌ها می‌بایست توأمان از کلیه تصاویر منتخب استخراج گردند و جمع تعداد ویژگی‌های استخراج شده به همراه اِعمال ضرایب تعیین کننده میزان تشابه خصوصی است.

مهمترین دلایل انتخاب الگوریتم نقاط کلیدی:

۱-چرخش تصویر: یکی از دلایل عمده در کاهش کیفیت واحد استخراج ویژگی، موضوع چرخش عکس است که این اَمر باعث کاهش شدید قدرت الگوریتم‌ها در استخراج ویژگی می‌گردند و روش‌های مختلف نقاط کلیدی به این موضوع حساس نیستند.

۲-حالت تصویر: مقدار کمی تغییر در حالت تصویر موجب کاهش شدید بازدهی الگوریتم‌های استخراج ویژگی می‌گردد که الگوریتم نقاط کلیدی در این خصوص کارآیی بسیار مناسبی دارد.

۳-تعداد تصاویر: بیشتر الگوریتم‌های تشخیص چهره با افزایش تعداد تصاویر کارآیی آن‌ها به شدت کاهش می‌یابد. تنها الگوریتم‌های مبتنی بر شبکه عصبی در حجم بالا دچار مشکل نمی‌گردند که آن‌ها نیز در موضوع حالت تصویر و چرخش بازدهی خوبی ندارند.

انواع روش‌های پیاده‌سازی شده برای کشف نقاط کلیدی تصویر چهره انسان

روش SIFT

تبدیل ویژگی بی‌تفاوت به اَبعاد یا Scale Invariant Feature Transform (SIFT) یک توصیف‌گر تصویر است که برای اِنطباق و تشخیص بر مبنای تصویر مورد استفاده قرار می‌گیرد. توصیف‌گر SIFT نسبت به حرکت‌های انتقالی، چرخش و تغییرات ابعاد بی‌تفاوت است و در مقابله با تبدیل‌های چشم‌انداز (perspective) معمول و تغییرات نور قدرتمند است. بر حسب تجربه، توصیف‌گر SIFT در دنیای واقعی، برای انطباق تصویر و شناسایی اجسام بسیار کارا عمل کرده است(Dubey S).

1. Key-points

روش SURF

توصیف‌گر SURF برای استخراج نقاط جذاب (یا همان نقاط کلیدی) بسیار شبیه به SIFT عمل می‌کند. یکی از تفاوت‌های SURF نسبت به SIFT در این است که به جای تقریب‌های مشتق بر موجک‌های هار بنا شده است(Winnicka A).

شکل شماره۲۲: نقاط کلیدی روش SURF

(Figure-22) Key points of SURF method

روش KAZE

KAZE Featuresیک روش تشخیص ویژگی و توصیف است که به صورت کامل در فضای مقیاس غیرخطی (nonlinear scale space) عمل می‌کند. روش‌هایی مانند SIFT و SURF ویژگی‌ها را در فضای مقیاس گوسی (Gaussian scale space) پیدا می‌کنند، لیکن تارسازی گوسی (Gaussian blurring) به مرزهای طبیعی اجسام اهمیت نمی‌دهد و جزئیات و نویز را به میزان یکسان در هنگام تکامل تصویر اصلی در فضای مقیاس (scale space) نرم می‌کند(Khalid NA).

روش BRISK

BRISK نیز یک روش تشخیص ویژگی و توصیف است که نسبت به تغییرات ابعاد و چرخش بی‌تفاوت است. این روش توصیف‌گر، ویژگی‌های تصویر محلی را به وسیله ارتباط (در فضای سیاه سفید) جفت نقطه‌های تصادفی در همسایگی تصویر محلی ایجاد می‌کند و به توصیف‌گر ویژگی دودویی دست می‌یابد. در مقایسه با الگوریتم‌های سنتی، سرعت تطبیق BRISK بیشتر و میزان مصرف حافظه کمتر بوده اما قدرت کمتری نسبت به آن‌ها دارد(Dixit D).

روش HARRIS

Harris روشی برای پیدا کردن گوشه‌ها و تخمین ویژگی‌ها در الگوریتم‌های بینایی رایانه‌ای (com-puter vision) است. این روش سال ۱۹۸۸ برای بهبود روش Moravec معرفی گردید. در مقایسه با روش مذکور، روش Harris به جای انتقال (shift) قطعات به ازای هر ۴۵ درجه، دیفرانسیل مقدار گوشه را همراه با ارجاع به جهت، به صورت مستقیم محاسبه می‌کند که ثابت شده است دقّت بیشتری در تمیز بین لبه‌ها و گوشه را داراست. این روش از آن زمان در الگوریتم‌های مختلف، بهبود داده شده و مورد استفاده قرار گرفته است(Zeghina AO).

شکل شماره ۲۳: نقاط کلیدی روش HARRIS

(Figure-23) Key points of HARRIS method

روش FAST

آزمون ویژگی‌های استخراج شده از بخش شتاب‌گرفته یا Features from accelerated segment test (FAST) یک روش تشخیص گوشه است که می‌توان از آن برای استخراج نقاط ویژگی و بعداً برای دنبال کردن و نگاشت اشیاء در بسیاری از اُمور بینایی رایانه‌ای (computer vision)استفاده کرد.

مهمترین برتری روش FAST بهینگی پردازشی آن است. این روش از بسیاری از روش‌های شناخته شده‌ی استخراج ویژگی مانند تفاوت گوسی (DoG) که در SIFT استفاده می‌شود سریع‌تر است، علاوه بر آن وقتی تکنیک‌های یادگیری ماشین اِعمال شود، کارآیی بیشتری در پردازش و مصرف منابع می‌توان انتظار داشت(Kusnadi A).

شکل شماره ۲۴: نقاط کلیدی روش FAST
(Figure-24) Key points of FAST method

روش MinEigen

این روش، گوشه‌ها را به وسیله الگوریتم مقدار خاصّ کمینه (Minimum eigenvalue algorithm) شناسایی می‌کند.

در جبر خطی، یک بُردار خاص (eigenvector) تبدیل خطی، یک بُردار غیر صفر است که حداکثر به میزان یک فاکتور اِسکالری در صورت اِعمال آن تبدیل بر این بُردار تغییر می‌کند. مقدار خاصّ (eigenvalue) همان فاکتور اِسکالری است که به وسیله آن بُردار خاص تغییر کرده است[۲۰].

شکل شماره ۲۵: نقاط کلیدی روش MinEign
(Figure-25) Key points of MinEign method

در پایان اجرا، تعداد نقاط مطابقت داده شده از اجرای هر روش الگوریتم نقاط کلیدی[1] در ضریب کارآیی آن روش ضرب شده و جمع آن‌ها امتیاز نهایی شناسایی را مشخص می‌کند. ضرایب کارآیی هر کدام از روش‌ها به صورت با ناظر و با داده‌های آموزشی مشخص گردیده است.

1. Key Points

مزایا و معایب مدل پیشنهادی

قابلیت استفاده در بانک اطلاعات حجیم

این مدل برای مجموعه داده‌های حجم بالا طراحی شده است و تا ظرفیت ۱۰ میلیون تصویر در مجموعه داده‌ای کارآیی تضمین شده‌ای نسبت به کارهای انجام شده دارد. برخی از کارهای مطالعه شده حتی برای حجم‌های بیش از یک هزار تصویر در مجموعه داده‌ای، کارآیی خود را از دست می‌دهند.

نسبت تعداد اشخاص به تعداد تصاویر

برخی از سیستم‌های تشخیص چهره صرفاً با یک دسته به خصوصی از مجموعه‌های داده‌ای کار می‌کنند، به طوری که حتماً به ازای هر شخص باید بیش از ۱۰۰ تصویر در مجموعه داده‌ایی موجود باشد. یعنی این نسبت نزدیک به صفر است. اما در مجموعه داده‌ای ما حتی برای نسبت یک هم کارآیی مناسبی از خود نشان می‌دهد. گرچه هراندازه این نسبت به صفر نزدیک‌تر باشد سیستم‌های تشخیص چهره ما بهتر عمل می‌کند، اما در کاربردهای واقعی نیازمند سیستم‌های تشخیص چهره‌ایی هستیم که برای نسبت تعداد اشخاص به تصاویری که برابر یک است نیز کارآیی خوبی داشته باشد.

عدم وابستگی به پارامترهای مختلف تصویر

بسیاری از سیستم‌های تشخیص چهره صرفاً با مجموعه داده‌هایی می‌توانند درست کار کنند که ویژگی‌های به خصوصی از جمله کیفیت تصویر و وضوح بالا داشته باشند، تصاویر دقیقاً از رو به روی فرد باشد، چهره فرد دارای زاویه نباشد، تصویرها در اندازه‌های ثابت باشند و مواردی از این قبیل را رعایت کنند. در حالی که سیستم تشخیص چهره پیشنهادی در تمامی تصاویر با کیفیت‌های مختلف، اندازه‌های متفاوت و... کارآیی خوبی دارد. هر چه کیفیت تصاویر بهتر باشد میزان کارآیی و دقّت سیستم تشخیص چهره افزایش می‌یابد اما در کاربردهای واقعی عدم وابستگی به این گونه از پارامترها، یکی از ملاک‌های مهم سیستم‌های تشخیص چهره است.

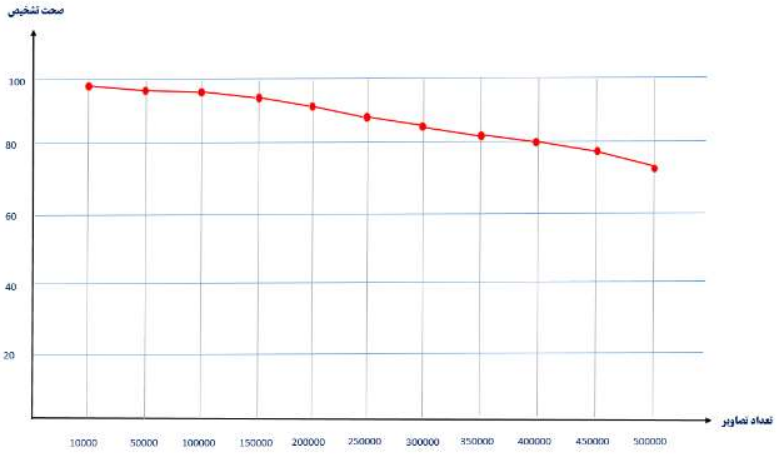
معایب مدل پیشنهادی

تمامی سیستم‌های تشخیص چهره دارای مزایا و معایبی هستند. از جمله معایب سیستم تشخیص چهره پیشنهادی در این مقاله می‌توان به موارد زیر اشاره کرد:

با توجه به استفاده از تکنیک‌های مختلف پردازش تصویر، الگوریتم‌های فازی، شبکه عصبی و الگوریتم ژنتیک، پیاده‌سازی این مدل احتیاج به دانش بالا در زمینه نرم‌افزار و هوش مصنوعی دارد.

نتیجه گیری

مکانیزم BIG-GAK بعد از پیاده‌سازی بر روی بانک تصاویر ۵۱۰ هزار نفری توضیح داده شده در مقاله تست گردیده است که به دلیل جامعیت مکانیزم نتایج بسیار خوبی در خصوص استخراج ویژگی‌ها و نهایتاً تشخیص هویّت افراد داشته است که نتایج آن در شکل زیر قابل مشاهده است:

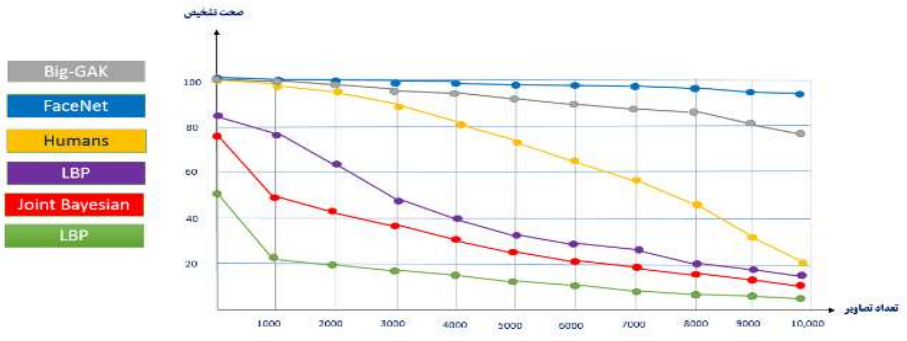


شکل شماره ۲۶: مکانیزم BIF_GAK بر روی ۵۰۰ هزار تصویر

(Figure-26) BIF_GAK mechanism on 500,000 images

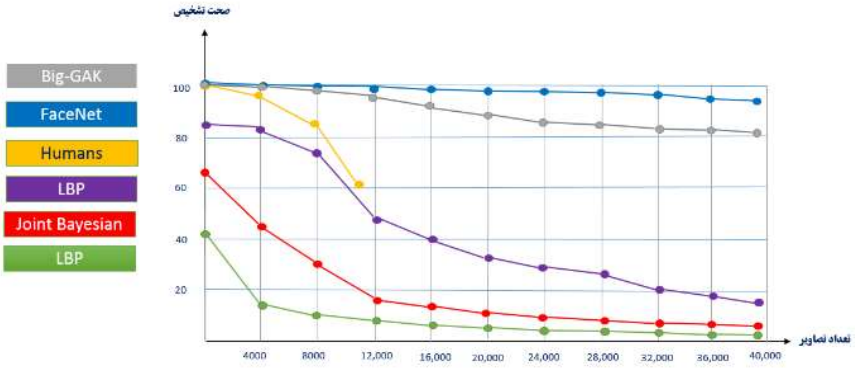
به منظور مقایسه درست بین الگوریتم BIG-GAKو الگوریتم‌های ارائه شده در قسمت‌های کارهای مرتبط بانک اطلاعاتی LFW که شامل ۹۸۰۰ تصویر افراد مشهور با اندازه ۲۵۰ در ۲۵۰ پیکسل است نیز مورد بررسی قرار گرفته است که نتیجه آن به شرح ذیل قابل نمایش است.

همچنین با بانک تصاویر CASIA که شامل ۲ مجموعه ۳۴۰۰۰ و ۴۰۰۰۰ تصویر انسان است و همچنین تصاویر با کیفیت‌های و اندازه‌های مختلفی می‌باشند مورد بررسی قرار گرفته است که نتیجه آن به شرح ذیل قابل نمایش است.



شکل شماره۲۷: مقایسه مکانیزم BIF_GAK با سایر روش‌ها

(Figure-27) Compare BIF_GAK mechanism with other methods



شکل شماره ۲۸: مقایسه مکانیزم BIF_GAK با سایر روش‌ها

(Figure-28) Compare BIF_GAK mechanism with other methods

مکانیزم تشخیص چهره ارائه شده توسط ما یک فرآیند شناختی است که با استفاده از چندین عامل هوشمند و تعامل درست آن‌ها پیاده‌سازی شده است و به دلیل واحد پایه بودن آن و تفکیک وظایف هر قسمت بسیار قابل توسعه و ارتقاء است. همچنین از اکثر سیستم‌های تشخیص چهره که در کارهای مرتبط مورد بررسی قرار گرفته بود کارآیی بهتری دارد اما هنوز از روش FaceNet فاصله داریم. به منظور افزایش کارآیی سیستم تشخیص چهره پیشنهادی مان، راه کارهایی را برای آینده ارائه می‌کنیم تا با تحقیق بر روی آن‌ها در آینده بتوان کارآیی سیستم را بیشتر کرد.

برخی از مهمترین کارهایی که در آینده می‌توان انجام داد عبارتند از:

الف) افزایش تعداد توابع ویژگی در مرحله اول

با افزایش تعداد توابع ویژگی می‌توان خروجی بهتری از کاندایدهای کمتری به عنوان ورودی برای مرحله دوم سیستم تشخیص چهره پیشنهادی‌مان ارائه داد.

ب) افزایش دقّت و کارآیی توابع ویژگی در مرحله اول

محقّقان به طور پیوسته در تلاش هستند که الگوریتم‌های ویژگی بهتر و کارآمدتری را ایجاد کنند. به عنوان مثال: برای تشخیص رنگ چشم انسان در تصویر ممکن است در آینده الگوریتم‌های دقیق‌تری ارائه شود. می‌توان به راحتی آن الگوریتم را جایگزین الگوریتم فعلی کرد و کارآیی و دقّت سیستم تشخیص چهره پیشنهادی‌مان را افزایش داد.

ج)استفاده از توابع جدید نقاط کلیدی در مرحله دوم

الگوریتم‌های مختلفی برای استخراج ویژگی نقاط کلیدی هر چند وقت یکبار ارائه می‌شود که برای کارآیی بهتر سیستم تشخیص چهره پیشنهادی‌مان می‌توان از آن‌ها استفاده کرد.

د) اجرای آزمایش در مقیاس ۱۰ میلیون تصویر

همان طور که قبلاً توضیح دادیم میزان حجم تصاویر مجموعه داده‌ای حائز اهمیت است و با افزایش حجم تصاویر مجموعه داده‌ای معمولاً اُفت‌های شدید یکباره‌ای رُخ می‌دهد. بررسی این موضوع را به کارهای آینده می‌سپاریم.

منابع

1- Afra S, Alhajj R. Early warning system: From face recognition by surveillance cameras to social media analysis to detecting suspicious people. Physica A: Statistical Mechanics and its Applications. 2020 Feb 15؛ 540: 123151.

2-Zhao F, Li J, Zhang L, Li Z, Na SG. Multi-view face recognition using deep neural networks. Future Generation Computer Systems. 2020 Oct 1؛ 111: 375-80.

3- Agarwal L, Mukim M, Sharma H, Bhandari A, Mishra A. Face Recognition Based Smart and Robust Attendance Monitoring using Deep CNN. In2021 8th International Conference on Computing for Sustainable Global Development (INDIACom) 2021 Mar 17 (pp. 699-704). IEEE.

4- Beham MP, Roomi SM. A review of face recognition methods. International Journal of Pattern Recognition and Artificial Intelligence. 2013 Jun 16؛ 27(04): 1356005.

5- Somogyi Z, Somogyi Z. Face Recognition. The Application of Artificial Intelligence: Step-by-Step Guide from Beginner to Expert. 2021: 175-84.

6- Shahid AM, Fraz MM, Shahzad M. Large Scale Face Recognition In the Wild: Technical Challenges and Research Directions. In2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2) 2021 May 20 (pp. 1-7). IEEE.

7- Fredj HB, Bouguezzi S, Souani C. Face recognition in unconstrained environment with CNN. The Visual Computer. 2020 Jan 18: 1-0.

8- Chen C, Surette R, Shah M. Automated monitoring for security camera networks: promise from computer vision labs. Security Journal. 2020 Feb 17: 1-21.

9- Zhang R, Lee EJ. Human Face Recognition Based on improved CNN Model with Multi-layers. Journal of Korea Multimedia Society. 2021؛ 24(5): 701-8.

10- Young AW, Burton AM. Insights from computational models of face recognition: A reply to Blauch, Behrmann and Plaut. Cognition. 2021 Mar 1؛ 208: 104422.

11- Hirzi MF, Efendi S, Sembiring RW. Literature Study of Face Recognition using The Viola-Jones Algorithm. In2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS) 2021 Apr 28 (pp. 1-6). IEEE.

12- Nagrani A, Chung JS, Xie W, Zisserman A. Voxceleb: Large-scale speaker verification in the wild. Computer Speech & Language. 2020 Mar 1؛ 60: 101027..

13- Zhang Q, Li H, Li M, Ding L. Feature extraction of face image based on LBP and 2-D Gabor wavelet transform. Math. Biosci. Eng. 2020 Jan 1؛ 17(2): 1578-92.

14- Dubey S, Meena R. A Review of Face Recognition Using SIFT Feature Extraction.

15- Winnicka A, Kęsik K, Połap D, Woźniak M. SURF Algorithm with Convolutional Neural Network as Face Recognition Technique. InInternational Conference on Artificial Intelligence and Soft Computing 2020 Oct 12 (pp. 95-102). Springer, Cham.

16- Khalid NA, Ahmad MI, Mandeel TH, Isa MN. Palmprint Verification System Using LBP And KAZE Features Detection. In2021 8th International Conference on Computer and Communication Engineering (ICCCE) 2021 Jun 22 (pp. 13-16). IEEE.

17- Dixit D, Parashar S, Gondalia A, Sengupta A, Sivagami M. Facial Identification using Haar Cascading with BRISK. In2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) 2020 Feb 24 (pp. 1-7). IEEE.

18- Zeghina AO, Zoubia O, Behloul A. Face Recognition Based on Harris Detector and Convolutional Neural Networks. InInternational Symposium on Modelling and Implementation of Complex Systems 2020 Oct 24 (pp. 163-171). Springer, Cham.

19- Kusnadi A. Face Detection Keypoints Using Dct And Clahe. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 2021 May 10؛ 12(11): 4365-72.

20- Soualmi A, Alti A, Laouamer L. A novel blind watermarking approach for medical image authentication using MinEigen value features. Multimedia Tools and Applications. 2021 Jan؛ 80(2): 2279-93.