

آشکارسازی چندتایی صورت انسان در تصاویر اینترنتی با استفاده از ترکیب فیلترهای گابور و شبکه عصبی

رضا محمدیان و محمود محلوچی

صورت دو چشم و دو ابرو وجود دارد که با یکدیگر متقارن بوده و فاصله بین این اجزا می‌تواند به عنوان یک رابطه برای آشکارسازی صورت استفاده شود. در گروه دوم الگوریتم‌ها به دنبال یافتن ویژگی‌هایی از صورت هستند که با وجود تغییر در زاویه دید، ژست صورت یا شرایط روشنایی، همچنان آن ویژگی‌ها وجود داشته باشند که از این نوع ویژگی‌ها می‌توان به رنگ پوست یا اجزای صورت اشاره کرد. در گروه سوم چندین الگوی صورت استاندارد جمع‌آوری شده و میزان ارتباط بین تصویر ورودی و الگوی ذخیره‌شده برای آشکارسازی صورت مورد استفاده قرار می‌گیرد. در گروه چهارم نیز مانند گروه سوم از الگوی صورت برای آشکارسازی آن استفاده می‌شود با این تفاوت که در گروه سوم چندین الگوی استاندارد جمع‌آوری می‌شود در حالی که در روش‌های گروه چهارم از روی چندین نمونه تصویر آموزشی الگوهای صورت، آموزش داده می‌شوند. در آشکارسازی صورت عموماً روش‌های مبتنی بر ظاهر کارایی بالاتری را نسبت به دیگر روش‌ها از خود نشان داده‌اند [۴] و [۵]. الگوهای آشکارسازی صورت در روش‌های مبتنی بر ظاهر با استفاده از الگوریتم‌های یادگیری ماشین به دست می‌آیند که از آن جمله می‌توان به SVM، آداپوست، بی‌زین و شبکه‌های عصبی اشاره کرد [۶]. برای آموزش الگوهای آشکارساز صورت، ویژگی‌های مختلفی توسط الگوریتم‌های یادگیری ماشین مورد استفاده قرار می‌گیرد. در [۷] تا [۹] از توصیف‌گر هار استفاده شده است. در [۱۰] و [۱۱] از توصیف‌گر SURF و در [۱۲] از توصیف‌گر LBP استفاده شده است. مرجع [۱۳] توصیف‌گر MB-LBP را با دو توصیف‌گر هار و SURF ترکیب کرده و آشکارسازی صورت را انجام داده است. نوع دیگری از توصیف‌گرها که برای آشکارسازی اشیای مختلف از جمله صورت بسیار قدرتمند می‌باشد فیلترهای گابور است. در [۱۴] تا [۱۷] برای آشکارسازی صورت و در [۱۸] برای آشکارسازی چین چروک پوست در ناحیه صورت از ویژگی‌های گابور استفاده شده است. ویژگی‌های گابور علی‌رغم این که از دقت بالایی برخوردار هستند ولی به دلیل داشتن حجم محاسبات زیاد، کمتر در کاربردهای برخط قابل استفاده می‌باشد. مرجع [۱۹] ویژگی‌های گابور را بر روی GPU پیاده‌سازی نموده و به کمک این روش توانسته است محدودیت استفاده از فیلترهای گابور برای پردازش‌های برخط را برطرف سازد. نویسندگان در [۲۰] به آشکارسازی صورت با استفاده از ترکیب فیلترهای گابور و شبکه عصبی از نوع پیش‌خور به روش بازگشتی در مجموعه تصاویر تک‌چهره با پس‌زمینه‌های ساده و پیچیده پرداخته‌اند.

این مقاله به آشکارسازی چندتایی صورت انسان در تصاویر اینترنتی با پس‌زمینه‌های پیچیده به کمک ترکیب ویژگی‌های انرژی فیلتر گابور و شبکه عصبی می‌پردازد. تمرکز روش ارائه‌شده در این مقاله بر کاهش خطای آشکارسازی نمونه‌های غیر صورت به عنوان صورت در تصاویر با کیفیت متوسط به بالا می‌باشد. از جمله دلایل انتخاب این تمرکز در روش پیشنهادی را می‌توان به کاربرد این نوع روش‌ها در موضوع بسیار بااهمیت

چکیده: این مقاله روشی نوین برای آشکارسازی چندتایی صورت انسان از نمای روبه‌رو در تصاویر اینترنتی با پس‌زمینه‌های پیچیده را با تمرکز بر کاهش خطای آشکارسازی نمونه‌های غیر صورت به عنوان صورت بیان می‌کند. در روش ارائه‌شده از ترکیب شبکه عصبی چندلایه پیش‌خور با روش آموزش بازگشتی و ویژگی انرژی فیلتر گابور در حوزه فرکانس استفاده شده است. در روش پیشنهادی این مقاله با معرفی پیش‌پردازشی جدید برای افزایش کیفیت ویژگی انرژی گابور، انجام دو مرحله پایش بر روی تصاویر ورودی و خروجی و همچنین استفاده از سه شاخص شناسایی اجزای صورت در خروجی انرژی گابور، خطای آشکارسازی نمونه‌های غیر صورت به عنوان صورت به شدت کاهش یافته است. در این مقاله پایگاه تصاویر جدیدی به نام RFD از تصاویر اینترنتی جمع‌آوری شده است که دارای ۵۸۳ تصویر صورت غیر تکراری و ۹۹۶۱ تصویر غیر صورت در اندازه 168×192 است. دقت الگوریتم پیشنهادی در آشکارسازی صورت در این پایگاه تصاویر 88.16% و خطای آشکارسازی آن تنها ۴۸ مورد معادل 0.48% است. این در حالی است که الگوریتم ویولا جونز در این پایگاه تصاویر دارای ۱۲۴ خطای آشکارسازی بوده و در نتیجه خطای آشکارسازی الگوریتم پیشنهادی نزدیک به $2/5$ برابر بهتر از الگوریتم ویولا جونز می‌باشد.

کلیدواژه: انرژی گابور، پروجکشن عمودی، فیلتر میانه، ویژگی‌های صورت.

۱- مقدمه

آشکارسازی صورت در تصویر یکی از روش‌های پایه‌ای برای ایجاد فعل و انفعالی طبیعی بین انسان و کامپیوتر می‌باشد [۱]. هدف از آشکارسازی چندتایی صورت در تصویر این است که یک ماشین تعیین نماید آیا در یک تصویر صورتی وجود دارد و اگر وجود دارد تعداد و محل هر یک از آنها را مشخص نماید [۲]. آشکارسازی صورت بخشی اساسی در کاربردهایی مانند شناسایی و تعیین هویت افراد از روی چهره، جستجو و بازیابی یک چهره در تصاویر، دسته‌بندی چهره‌ها و ... بوده و دقت آن تأثیر مستقیمی در کیفیت این نوع کاربردها دارد [۳]. وجود و بروز این نوع کاربردها در پردازش تصویر باعث شده است که از گذشته تاکنون مبحث آشکارسازی صورت از اهمیت بالایی برخوردار بوده و روش‌های مختلفی برای آن ارائه شود. روش‌های ارائه‌شده برای آشکارسازی صورت را می‌توان در چهار گروه روش‌های مبتنی بر دانش، روش‌های مبتنی بر ویژگی‌های ثابت، روش‌های تطبیق قالب و روش‌های مبتنی بر ظاهر دسته‌بندی نمود. در گروه اول آشکارسازی صورت با بررسی آماری بر روی روابط موجود در بین ویژگی‌های صورت انجام می‌شود. برای مثال در

این مقاله در تاریخ ۱ بهمن ماه ۱۳۹۳ دریافت و در تاریخ ۱ تیر ماه ۱۳۹۴ بازنگری شد.

رضا محمدیان، گروه مکترونیک، واحد کاشان، دانشگاه آزاد اسلامی، کاشان، (email: mohammadian_reza@hotmail.com).

محمود محلوچی، گروه مخابرات، واحد کاشان، دانشگاه آزاد اسلامی، کاشان، (email: mmahlouji@yahoo.com).

ساختار مقاله به صورت زیر پیاده‌سازی شده است: در بخش دوم نحوه آشکارسازی چندتایی صورت با استفاده از فیلترهای گابور و شبکه‌های عصبی بیان می‌شود. در بخش سوم روش‌های کاهش خطای آشکارسازی شناسایی نمونه‌های غیر صورت آورده شده است. در بخش چهارم عملکرد الگوریتم پیشنهادی مورد بررسی قرار گرفته و نتایج آن آمده در بخش پنجم نیز نتیجه‌گیری آورده شده است.

۲- آشکارسازی چندتایی صورت با استفاده از فیلترهای گابور و شبکه عصبی

آشکارسازی چندتایی صورت در روش پیشنهادی با استفاده از ترکیب فیلترهای گابور و شبکه عصبی انجام می‌شود، لذا در این بخش ابتدا به معرفی فیلتر گابور و نحوه مقاردهی پارامترهای آن پرداخته شده و سپس روش آشکارسازی صورت در تصاویر چندچهره بیان می‌شود.

۲-۱ فیلتر گابور

فیلتر گابور به کمک ترکیب یک موج سینوسی و تابع گوسین به دست می‌آید. نوع دوبعدی فیلتر گابور به صورت گسترده در پردازش تصویر مخصوصاً برای آشکارسازی و شناسایی ناحیه صورت استفاده می‌شود. خروجی فیلتر گابور به شکل اعداد مختلط بوده و برای تولید ویژگی‌های گابور می‌توان از بخش حقیقی، موهومی یا مختلط تابع تولید فیلتر گابور استفاده کرد. اندازه شکل مختلط خروجی فیلتر گابور به انرژی گابور معروف بوده و با استفاده از آن بهتر می‌توان اجزای صورت را نشان داد. در این مقاله از (۱) برای تولید فیلتر گابور دوبعدی استفاده شده است

$$g(x, y, \lambda, \sigma, \gamma, \theta, \varphi) = g_{\text{Real}}(x, y, \lambda, \sigma, \gamma, \theta, \varphi) + jg_{\text{Imaginary}}(x, y, \lambda, \sigma, \gamma, \theta, \varphi)$$

$$g_{\text{Real}}(x, y, \lambda, \sigma, \gamma, \theta, \varphi) = \frac{\gamma}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{x_r^2 + \gamma^2 y_r^2}{2\sigma^2}\right) \cos(2\pi f x_r + \varphi)$$

$$g_{\text{Imaginary}}(x, y, \lambda, \sigma, \gamma, \theta, \varphi) = \frac{\gamma}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{x_r^2 + \gamma^2 y_r^2}{2\sigma^2}\right) \sin(2\pi f x_r + \varphi)$$

که در آن متغیرهای x_r و y_r از طریق (۲) به دست می‌آید

$$\begin{aligned} x_r &= x \cos \theta + y \sin \theta \\ y_r &= -x \sin \theta + y \cos \theta \end{aligned} \quad (2)$$

که σ انحراف استاندارد تابع گوسین دوبعدی و θ میزان چرخش محورهای تابع گوسین دوبعدی (میزان چرخش فیلتر گابور) است. φ فاز افست بوده و تقارن تابع گابور را نشان می‌دهد و γ نرخ شکل ظاهر در تابع گابور دوبعدی می‌باشد. در این مقاله برای تولید ویژگی انرژی گابور پارامترهای تابع تولید فیلتر گابور مطابق (۳) مقاردهی شده است [۲۰]. شکل ۱ خروجی فیلتر گابور در روش پیشنهادی را بر روی یک نمونه تصویر از پایگاه RFD نشان می‌دهد

$$f = \frac{\pi}{\gamma}, \quad \sigma = 0.4, \quad \gamma = 1, \quad \theta = \frac{\pi}{\gamma}, \quad \varphi = 0 \quad (3)$$

۲-۲ آشکارسازی چندتایی صورت در روش پیشنهادی

در الگوریتم آشکارسازی صورت در تصاویر چندچهره لازم است ابتدا یک تصویر بزرگ به زیرتصاویر در اندازه 168×192 تبدیل شود. ساده‌ترین شکل تبدیل یک تصویر بزرگ به زیرتصاویر کوچک‌تر استفاده از روش



(الف)



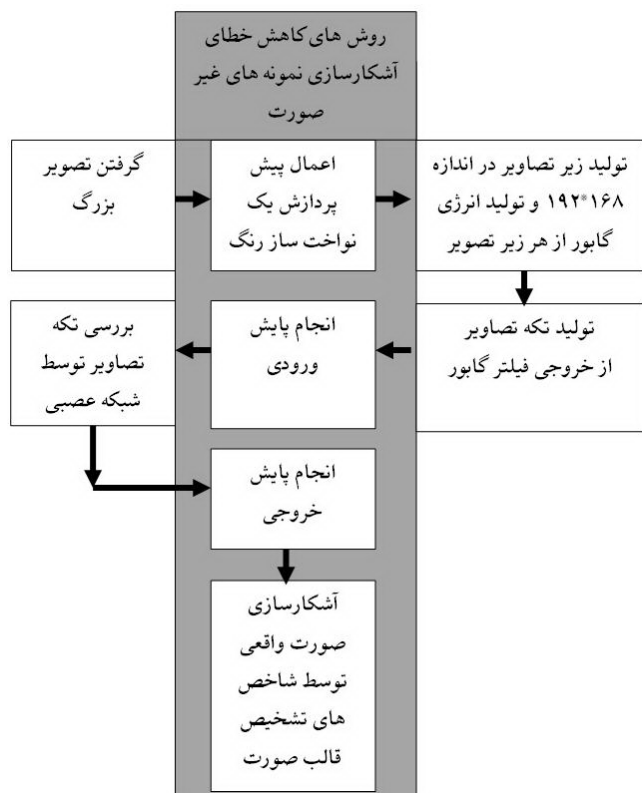
(ب)

شکل ۱: خروجی فیلتر گابور در روش پیشنهادی در یک نمونه تصویر از پایگاه RFD، (الف) تصویر اصلی و (ب) خروجی فیلتر گابور.

کاهش مصرف انرژی دانست. برای مثال در [۲۱] سیستمی ارائه شده است که به کمک روش آشکارسازی صورت، وجود بیننده را در مقابل تلویزیون بررسی می‌کند، در صورتی که برای ۵ تا ۱۰ دقیقه بیننده‌ای را شناسایی نکند مقدار درخشندگی صفحه نمایشگر را کاهش داده و با این روش ۳۰٪ در مصرف انرژی صرفه‌جویی نموده است یا در [۲۲] استفاده از آشکارسازی صورت را برای ذخیره انرژی در کامپیوتر پیاده‌سازی نموده است. در این نوع کاربردها لازم است تا آشکارسازی صورت با حداقل خطای آشکارسازی نمونه‌های غیر صورت انجام پذیرد تا خطا موجب هدررفت انرژی نگردد.

کاهش خطای آشکارسازی نمونه‌های غیر صورت به عنوان صورت در روش پیشنهادی در چهار مرحله صورت می‌گیرد: در مرحله اول پیش از شروع فرایند آشکارسازی، پیش‌پردازش یکنواخت‌ساز رنگ بر روی تصویر اعمال می‌شود. در مرحله دوم بر روی تکه تصاویر ورودی بر اساس احتمال صورت‌بودن پایش ورودی صورت می‌گیرد. در مرحله سوم در بین تکه تصاویر آشکارشده به عنوان صورت (در یک زیرتصویر) بر اساس مقدار خروجی شبکه عصبی در ازای هر تکه تصویر پایش خروجی انجام می‌گیرد. در مرحله چهارم بر روی نمونه تصاویر آشکارشده سه شاخص تشخیص اجزای صورت اعمال می‌گردد.

از آنجایی که مقدار پارامترهای فیلتر گابور در روش پیشنهادی برای آشکارسازی چهره در تصاویر با کیفیت متوسط به بالا تنظیم شده است لذا بر خلاف [۷]، [۱۰]، [۱۳]، [۱۵] و [۱۷] امکان ارزیابی روش پیشنهادی بر روی مجموعه تصاویر پایگاه CMU+MIT وجود نداشته و از پایگاه تصاویر جدیدی با نام RFD استفاده شده است. در این مقاله به منظور بررسی دقت عملکرد الگوریتم پیشنهادی مقایسه‌ای بین نتایج حاصل از روش ارائه‌شده و نتایج حاصل از الگوریتم ویولا جونز (V.J.) [۲۳] نیز آورده شده است.



شکل ۵: فرایند آشکارسازی صورت در تصاویر چندچهره با استفاده از فیلترهای گابور و شبکه عصبی در الگوریتم پیشنهادی. بخش های نشان داده شده در مستطیل خاکستری رنگ به منظور کاهش خطای آشکارسازی به الگوریتم پیشنهادی اضافه شده است.

شکل ۴ تعدادی از زیرتصاویر تولیدشده از نمونه تصویر شکل ۱ را نشان می دهد. در شکل ۵ فرایند آشکارسازی صورت در تصاویر چندچهره آورده شده است. مطابق این فرایند ابتدا یک تصویر بزرگ به زیرتصاویر در اندازه 192×168 تبدیل شده و از هر یک از زیرتصاویر تولیدشده با استفاده فیلتر گابور خروجی انرژی گابور به دست آورده می شود. خروجی گابور به دست آمده به تکه تصاویر در اندازه 60×60 تبدیل شده و سپس هر یک از این تکه تصاویر جهت بررسی وجود صورت در آنها به شبکه عصبی داده می شود. در صورتی که مقدار خروجی شبکه عصبی در ازای یک تکه تصویر از مقدار آستانه 0.85 بیشتر باشد آن تکه تصویر به عنوان صورت شناخته شده و اگر مقدار شبکه عصبی کمتر از مقدار آستانه باشد آن تکه تصویر به عنوان صورت شناخته نمی شود. به منظور کاهش خطای آشکارسازی نمونه های غیر صورت به عنوان صورت نیز مجموعه ای از روش ها در مراحل مختلف الگوریتم فوق اضافه شده که این مجموعه توسط مستطیلی خاکستری رنگ مشخص شده است. در بخش ۳ به تشریح روش های کاهش خطای آشکارسازی نمونه های غیر صورت به عنوان صورت پرداخته شده است.

۳- روش های کاهش خطای آشکارسازی نمونه های غیر صورت به عنوان صورت

کاهش خطای آشکارسازی نمونه های غیر صورت به عنوان صورت در چهار مرحله صورت می گیرد: در مرحله اول پیش از شروع فرایند آشکارسازی، پیش پردازش یکنواخت ساز رنگ بر روی تصویر اعمال می شود. در مرحله دوم بر روی تکه تصاویر ورودی بر اساس احتمال صورت بودن پایش ورودی صورت می گیرد. در مرحله سوم در بین تکه تصاویر آشکار شده به عنوان صورت (در یک زیرتصویر) بر اساس مقدار



(الف)



(ب)

شکل ۲: قرارگرفتن صورت در دو تکه تصویر در هنگام تولید زیرتصاویر از یک تصویر بزرگ در روش ماتریسی، (الف) در هنگام تقسیم عمودی و (ب) در هنگام تقسیم افقی.



شکل ۳: جداسازی زیرتصاویر در اندازه 192×168 از یک تصویر بزرگ در روش پیشنهادی.



شکل ۴: تعدادی از زیرتصاویر تولیدشده از یک نمونه تصویر از پایگاه RDF.

ماتریسی است که در آن از ابتدای تصویر بزرگ زیرتصاویر کوچک تر به صورت پیاپی جدا شوند ولی در این روش ممکن است بخشی از صورت در زیرتصویر قبلی قرار گرفته و بخش دیگر صورت نیز در زیرتصویر بعدی قرار گیرد. شکل های ۲- الف و ۲- ب به ترتیب قرارگرفتن صورت در دو تکه تصویر در هنگام تقسیم عمودی و در هنگام تقسیم افقی را در روش ماتریسی نشان می دهد.

برای حل این مشکل در روش پیشنهادی همیشه زیرتصویر بعدی از نیمه زیرتصویر قبلی شروع می شود و با این روش احتمال این که تصویر صورتی از دست داده شود بسیار پایین می آید. شکل ۳ روش جداسازی زیرتصاویر از یک تصویر بزرگ را در دو راستای عمودی و افقی در روش پیشنهادی نشان می دهد. در این شکل مستطیل سیاه رنگ نشان دهنده زیرتصویر قبلی و مستطیل سفید رنگ نشان دهنده زیرتصویر بعدی است. رابطه (۴) تعداد زیرتصاویر تولیدشده از یک تصویر بزرگ در اندازه $A \times B$ را با استفاده از این روش نشان می دهد

$$\left(\left(\frac{A}{192} \times 2 \right) - 1 \right) \times \left(\left(\frac{B}{128} \times 2 \right) - 1 \right) \quad (4)$$



(الف)



(ب)

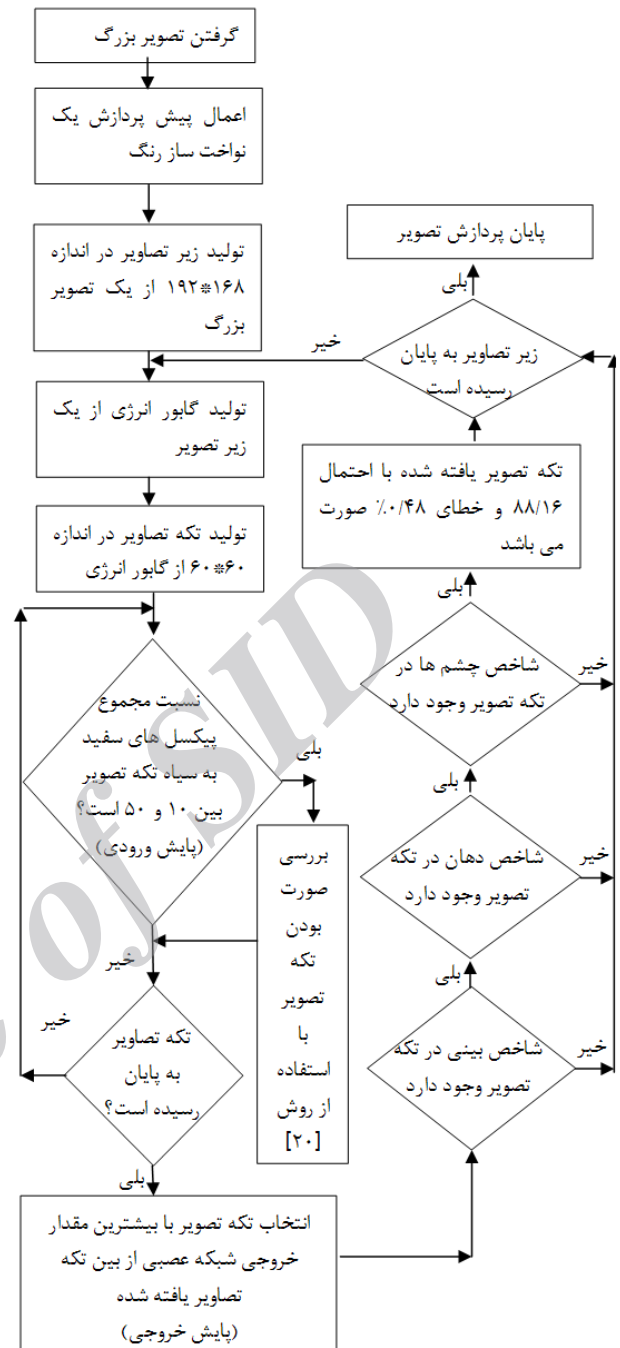


(ج)

شکل ۷: تأثیر پیش‌پردازش یکنواخت‌ساز رنگ در خروجی فیلتر گابور، (الف) تصویر ورودی، (ب) خروجی فیلتر گابور قبل از انجام پیش‌پردازش و (ج) خروجی فیلتر گابور بعد از انجام پیش‌پردازش.

پیش‌پردازش یکنواخت‌ساز رنگ نشان می‌دهد. همان طور که در این تصویر مشخص است خروجی فیلتر گابور دارای نویزهای زیادی می‌باشد. با توجه به این مطلب که هر کجا در تصویر ورودی تغییر رنگی رخ دهد فیلتر گابور مرز بین آن دو ناحیه را مشخص می‌کند می‌توان نتیجه گرفت که این نویزها در اثر تغییر رنگ‌های کوچکی در تصویر که با چشم قابل رؤیت نیستند به وجود آمده‌اند. این تغییر رنگ‌ها می‌توانند با استفاده از نرم‌افزارهای ویرایش عکس ایجاد شده باشند. در روش پیشنهادی برای حذف این نویزها یک فیلتر میانه با ماتریس 3×3 بر روی تصاویر اعمال می‌شود. انجام این پیش‌پردازش باعث یکنواخت‌تر شدن رنگ تصاویر گردیده و در نتیجه این نوع نویزها از خروجی فیلتر گابور حذف می‌شوند. شکل ۷-ج خروجی فیلتر گابور را بعد از اعمال پیش‌پردازش یکنواخت‌ساز رنگ نشان می‌دهد. همان طور که مشخص است میزان نویزها در تصویر ج نسبت به تصویر ب به شدت کاسته شده است.

به منظور بررسی تأثیر عملکرد این پیش‌پردازش در آشکارسازی صورت، ۵ تصویر ۳ چهره به صورت تصادفی از پایگاه تصاویر RFD انتخاب شده و عمل آشکارسازی صورت قبل و بعد از پیش‌پردازش یکنواخت‌ساز رنگ بر روی آنها اعمال می‌شود. مطابق جدول ۱ از مجموع ۱۵ تصویر صورت قبل از انجام پیش‌پردازش یکنواخت‌ساز رنگ ۹ تصویر صورت آشکار شده و ۳ تصویر غیر صورت نیز به عنوان صورت شناسایی می‌شود در حالی که بعد از انجام پیش‌پردازش یکنواخت‌ساز رنگ تمامی ۱۵ تصویر صورت آشکار شده و یک تصویر غیر صورت به عنوان صورت شناسایی می‌شود. لذا در روش پیشنهادی قبل از انجام فرایند آشکارسازی صورت پیش‌پردازش یکنواخت‌ساز رنگ بر روی تصاویر اعمال می‌شود.

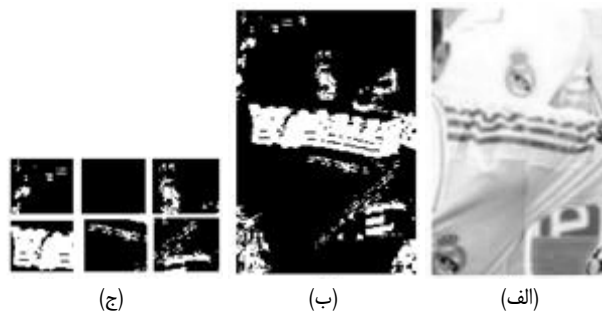


شکل ۶: نمودار بلوکی آشکارسازی صورت در روش پیشنهادی همراه با چگونگی اعمال مراحل کاهش خطای آشکارسازی.

خروجی شبکه عصبی در ازای هر تکه تصویر پیش‌خروجی انجام می‌گیرد. در مرحله چهارم بر روی نمونه تصاویر آشکار شده سه شاخص تشخیص اجزای صورت اعمال می‌گردد. در این بخش به معرفی هر یک از چهار مرحله فوق پرداخته می‌شود. شکل ۶ نمودار بلوکی آشکارسازی صورت در روش پیشنهادی را همراه با چگونگی اعمال مراحل کاهش خطای آشکارسازی به تفصیل نشان می‌دهد.

۳-۱ بهبود کیفیت ویژگی انرژی گابور با استفاده از پیش‌پردازش یکنواخت‌ساز رنگ

در یک بررسی تجربی مشخص شد که تصاویر اینترنتی دارای نویزهایی هستند که در ظاهر مشخص نبوده ولی وقتی فیلتر گابور در آنها اعمال می‌شود باعث پایین آمدن کیفیت خروجی فیلتر گابور می‌شوند. شکل ۷-ب خروجی فیلتر گابور را بر روی یک نمونه تصویر قبل انجام



شکل ۸: تأثیر پایش ورودی در یک نمونه زیرتصویر از پایگاه RFD، (الف) زیرتصویر ورودی، (ب) خروجی فیلتر گابور و (ج) بعضی از تکه تصاویر جداشده از خروجی فیلتر گابور که در مرحله پایش ورودی حذف می‌شوند.

شکل ۹: تأثیر پایش خروجی بر روی دو زیرتصویر از پایگاه RFD، (۱) زیرتصویر اصلی، (۲) آشکارسازی صورت قبل از انجام عمل پایش، (۳) تکه تصاویر آشکارشده به عنوان صورت توسط شبکه عصبی همراه با مقدار خروجی شبکه عصبی در ازای هر تکه تصویر و (۴) آشکارسازی صورت بعد از انجام عمل پایش.

جدول ۱: مقایسه نتایج آشکارسازی صورت قبل از انجام پایش پردازش یکنواخت‌ساز رنگ و بعد از اجرای آن.

آشکارسازی قبل از پایش پردازش یکنواخت‌ساز		آشکارسازی بعد از پایش پردازش یکنواخت‌ساز		تعداد زیرتصاویر بدون صورت	تعداد زیرتصاویر شامل صورت
غلط	صحیح	غلط	صحیح		
۳	۹	۱	۱۵	۱۳	۱۵

کامل صورت داشته و لذا در الگوریتم پیشنهادی با در نظر گرفتن این مطلب از بین تمامی تکه تصاویری که دارای خروجی شبکه عصبی بیش از مقدار آستانه می‌باشند تنها یک تکه تصویر با بالاترین مقدار خروجی شبکه عصبی به عنوان صورت انتخاب می‌شود. در این مقاله این فرایند تحت عنوان پایش تکه تصاویر آشکارشده بر اساس مقدار خروجی شبکه عصبی (پایش خروجی) معرفی شده است. پایش خروجی در بین تکه تصاویری که شبکه عصبی به عنوان صورت می‌شناسد انجام شده و باعث می‌شود در زیرتصاویری که صورتی در آنها وجود دارد کامل‌ترین قالب صورت آشکارسازی شده و در زیرتصاویری که هیچ صورتی در خود ندارند در صورتی که تکه تصاویری به خطا آشکارسازی شده باشند تنها یکی از خطاها به خروجی منتقل گردد. شکل ۸ تأثیر پایش خروجی را بر روی دو نمونه تصویر نشان می‌دهد. ستون دوم از شکل ۹ نتیجه آشکارسازی صورت قبل از اعمال پایش خروجی را نشان می‌دهد. در ستون سوم از این شکل تکه تصاویر شناسایی شده به عنوان صورت همراه با مقدار خروجی شبکه عصبی در ازای هر یک از این تکه تصاویر نشان داده است. ستون چهارم نیز نتیجه آشکارسازی صورت در روش پیشنهادی بعد از اعمال پایش خروجی را نشان می‌دهد. همان طور که مشخص است در نمونه تصویر شکل ۹-الف شبکه عصبی سه تکه تصویر را به عنوان صورت تشخیص داده که شباهت زیادی به قالب صورت دارند ولی مقدار خروجی شبکه عصبی در ازای تکه تصویر صورت از مقدار خروجی شبکه عصبی در ازای دو تکه تصویر دیگر بالاتر است، لذا با استفاده از روش پایش خروجی صورت واقعی آشکار شده و نمونه‌های خطا حذف می‌شوند. در شکل ۹-ب نیز هیچ صورتی در زیرتصویر وجود ندارد و با وجود این که شبکه عصبی دو تکه تصویر را به خطا به عنوان صورت آشکار کرده است به کمک اعمال پایش خروجی تنها یک خطا به خروجی منتقل می‌شود. جدول ۲ تأثیر این روش را در کاهش آشکارسازی نمونه‌های غیر صورت به عنوان صورت در ۵ نمونه تصویر چندچهره از پایگاه RFD نشان می‌دهد.

۲-۳ پایش بر اساس احتمال صورت بودن (پایش ورودی)

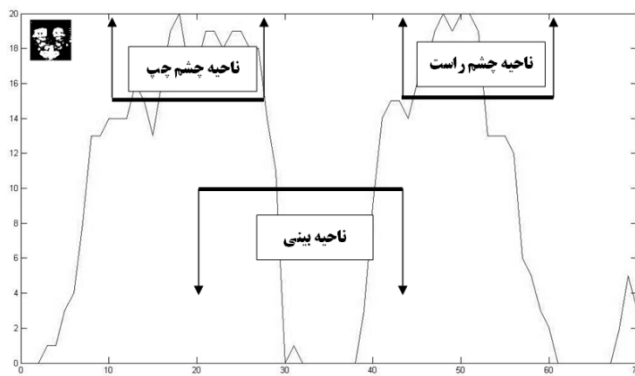
در بررسی تکه تصاویر تولیدشده از یک زیرتصویر با پس‌زمینه پیچیده مشخص می‌شود که در بعضی از این تکه تصاویر بخش عمده‌ای از خروجی فیلتر گابور به صورت سیاه یا سفید بوده و احتمال این که صورت در آنها وجود داشته باشد بسیار کم است. از بررسی آماری بر روی ۱۰۰ تکه تصویر شامل قالب کامل صورت مشخص می‌شود که مطابق (۵) در یک تکه تصویر صورت در اندازه $n \times n$ مانند T نسبت مجموع پیکسل‌های سفید به کل پیکسل‌های آن بین ۱۰ و ۵۰ می‌باشد

$$10 \leq \frac{\sum_{i,j=1}^n T(i,j)}{n \times n} \leq 50 \quad (5)$$

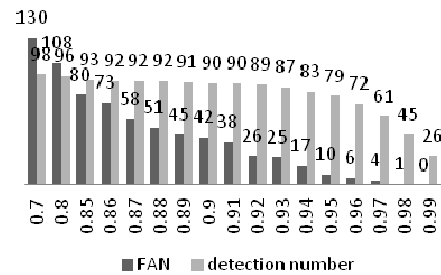
لذا در صورتی که در یک تکه تصویر مقدار این نسبت کمتر از ۱۰ و بیشتر از ۵۰ باشد صورتی در آن وجود ندارد. در این مقاله این فرایند تحت عنوان پایش بر اساس احتمال صورت بودن (پایش ورودی) معرفی شده است. پایش ورودی قبل از ارسال یک تکه تصویر به شبکه عصبی انجام شده و علاوه بر کاهش احتمال خطای آشکارسازی باعث افزایش سرعت الگوریتم پیشنهادی نیز می‌شود. شکل ۸ بعضی از تکه تصاویر حذف شده در مرحله پایش ورودی را نشان می‌دهد.

۳-۳ پایش بر اساس مقدار خروجی شبکه عصبی (پایش خروجی)

در تصاویر با پس‌زمینه‌های پیچیده این احتمال وجود دارد که تکه تصاویری که شباهت زیادی به صورت دارند شبکه عصبی را دچار خطا کرده و به عنوان صورت شناسایی شوند. در یک بررسی تجربی مشخص شد که در بین تکه تصاویر آشکارشده در یک زیرتصویر، تکه تصویری که بالاترین خروجی شبکه عصبی را دارد بیشترین تطابق را نسبت به قالب



شکل ۱۲: تعیین ناحیه بینی و دو چشم در پروجکشن عمودی نیمه بالای خروجی انرژی گابور از یک تکه تصویر صورت.



شکل ۱۰: آشکارسازی صورت در ۱۷ تصویر چندچهره شامل ۱۰۰ صورت از پایگاه RFD در آستانه‌های مختلف.



شکل ۱۱: خروجی فیلتر گابور در چند نمونه تکه تصویر صورت.

جدول ۲: تأثیر روش پایش خروجی بر کاهش آشکارسازی نمونه‌های غیر صورت به عنوان صورت در چند نمونه تصویر چندچهره از پایگاه RFD.

آشکارسازی قبل از پایش		آشکارسازی بعد از پایش		تعداد زیرتصاویر بدون صورت	تعداد زیرتصاویر شامل صورت
شناسایی غلط	شناسایی صحیح	شناسایی غلط	شناسایی صحیح		
۵	۱۵	۱	۱۵	۱۳	۱۵

آشکار شده به ۱ مورد رسیده ولی صورت‌های صحیح آشکار شده نیز به ۴۵ مورد کاهش می‌یابد.

با توجه به این نتیجه می‌بایست راهکاری پیاده‌سازی کرد که بدون افزایش سطح آستانه بتوان نمونه‌های غیر صورتی را که شبکه عصبی به عنوان صورت آشکارسازی کرده است را تشخیص داده و حذف نمود. در الگوریتم پیشنهادی حل این مشکل به کمک سه شاخص شناسایی اجزای صورت در خروجی فیلتر گابور انجام می‌گیرد. این سه شاخص شامل شاخص بینی، شاخص دهان و شاخص چشم‌ها است. در ادامه هر یک از این سه شاخص معرفی شده و سپس نحوه تأثیر شاخص‌های مذکور در کاهش خطای آشکارسازی مورد بررسی قرار می‌گیرد.

۳-۴-۱- شاخص بینی

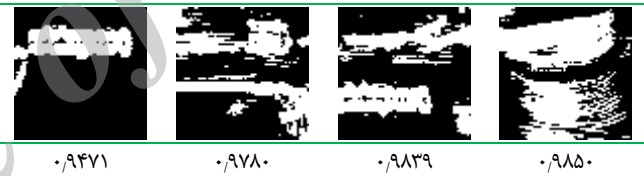
در شکل ۱۱ خروجی فیلتر گابور در چند نمونه تکه تصویر صورت آورده شده است. با دقت در این شکل می‌توان دریافت در نیمه بالای خروجی فیلتر گابور در تکه تصویر صورت بین دو ناحیه چشم تعداد پیکسل‌های سفید کمتر از محدوده راست و چپ آن می‌باشد. این ناحیه همان محل وجود بینی در تصویر صورت است.

با اجرای پروجکشن عمودی با استفاده از تابع جمع [۲۴] در نیمه بالای خروجی فیلتر گابور در تکه تصویر صورت ناحیه بینی شناسایی می‌شود. منظور از پروجکشن عمودی جمع مقادیر صفر و یک پیکسل‌های هر ستون از تصویر سیاه و سفید است. خروجی پروجکشن عمودی ماتریسی است $n \times n$ که n ، تعداد ستون‌های تصویر است. شکل ۱۲ نتیجه حاصل از اجرای پروجکشن عمودی بر روی نیمه بالای خروجی فیلتر گابور در یک تکه تصویر صورت را نشان می‌دهد.

با بررسی پروجکشن عمودی نیمه بالای ۳۸ تصویر صورت مختلف شاخص بینی به صورت (۶) تعریف می‌شود که در آن منظور از V.P. پروجکشن عمودی است

$$\begin{aligned} \exists x \in (20, 45) &\rightarrow V.P.(x) < 10 \\ \exists x \in (10, 30) &\rightarrow V.P.(x) \geq 16 \\ \exists x \in (40, 60) &\rightarrow V.P.(x) \geq 16 \end{aligned} \quad (6)$$

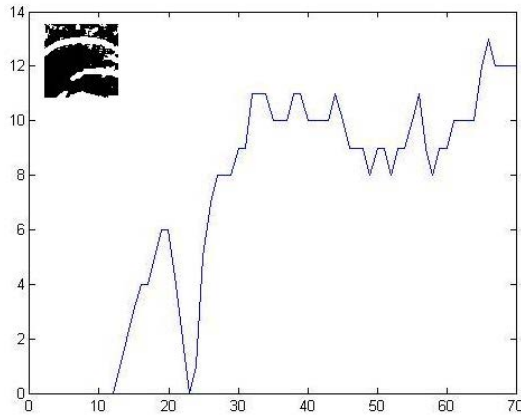
جدول ۳: چند نمونه تکه تصاویر غیر صورت که بعد از پیش‌پردازش یکنواخت‌سازی رنگ و انجام عملیات پایش به خطا توسط شبکه عصبی به عنوان صورت آشکار شده‌اند. در پایین هر تکه تصویر مقدار خروجی شبکه عصبی در ازای آن تکه تصویر آورده شده است.



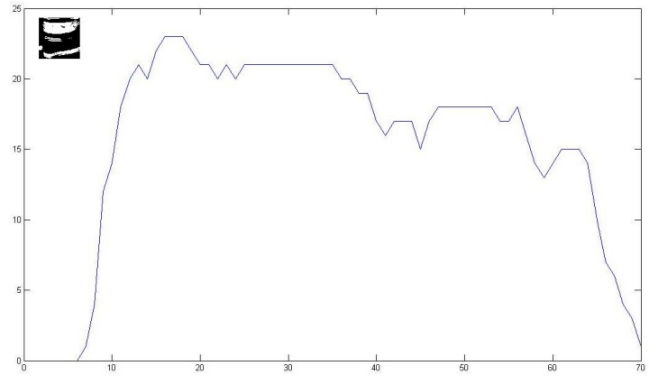
۳-۴-۲- شاخص‌های شناسایی اجزای کامل صورت در خروجی فیلتر گابور

در ادامه فرایند کاهش آشکارسازی نمونه‌های غیر صورت به عنوان صورت تکه تصاویری که صورتی در آنها وجود نداشته و بعد از اعمال پیش‌پردازش یکنواخت‌سازی رنگ تصویر و عملیات پایش هنوز به عنوان صورت آشکارسازی شده‌اند مورد بررسی قرار می‌گیرند. از بررسی‌های انجام‌شده مشخص می‌شود که شبکه عصبی در روش پیشنهادی بر این اساس آموزش یافته است که در صورت وجود دو باند سفیدرنگ در بالا و پایین یک تکه تصویر، با احتمال بالایی آن تکه تصویر را به عنوان صورت شناسایی کرده و در صورت وجود یک باند سفیدرنگ در بالای تصویر نیز با احتمال کمتری آن را به عنوان صورت شناسایی می‌کند. جدول ۳ چند نمونه از این نوع تکه تصاویر را همراه با مقدار خروجی شبکه عصبی در مورد آنها نشان می‌دهد.

در صورتی که بخواهیم این گونه موارد خطا را با افزایش سطح آستانه حذف کنیم موجب کاهش شدید نمونه‌های صحیح صورت خواهد شد. برای نشان دادن این مطلب ۱۷ تصویر چند چهره شامل ۱۰۰ صورت از پایگاه RFD انتخاب شده و آشکارسازی صورت با استفاده از آستانه‌های مختلف در آنها صورت می‌گیرد. نمودار شکل ۱۰ نتیجه حاصل از این بررسی را نشان می‌دهد. همان طور که از این نمودار مشخص است در آستانه ۰.۸۵ تعداد صورت‌های صحیح شناسایی شده ۹۳ مورد و موارد به غلط آشکار شده ۸۰ مورد است در حالی که در آستانه ۰.۹۸ موارد به غلط



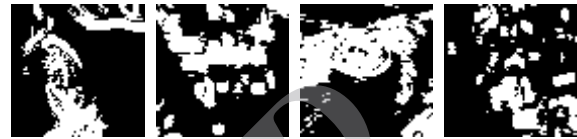
شکل ۱۶: پروجکشن عمودی نیمه پایین خروجی فیلتر گابور از یک تکه تصویر غیر صورت.



شکل ۱۳: پروجکشن عمودی نیمه بالای خروجی فیلتر گابور از یک تکه تصویر غیر صورت.



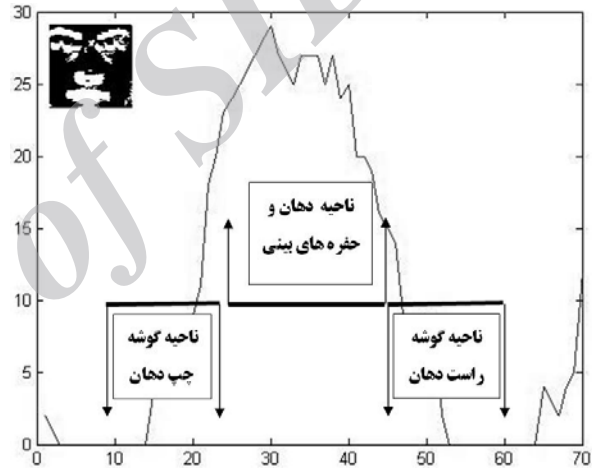
شکل ۱۷: چند نمونه خطای آشکارسازی که شاخص دهان شناسایی می‌کند.



شکل ۱۴: چند نمونه خطای آشکارسازی که شاخص بینی شناسایی می‌کند.



شکل ۱۸: چند نمونه خطای آشکارسازی که شاخص چشمها شناسایی می‌کند.



شکل ۱۵: تعیین ناحیه دهان در پروجکشن عمودی نیمه پایین خروجی انرژی گابور از یک تکه تصویر صورت.

$$\begin{aligned} \exists x_{1-1} \in (25, 45) &\rightarrow V.P.(x) > 10 \\ \exists x_{1-5} \in (10, 25) &\rightarrow V.P.(x) < 10 \\ \exists x_{1-5} \in (45, 60) &\rightarrow V.P.(x) < 10 \end{aligned} \quad (7)$$

مطابق (۷) یک تکه تصویر صورت است اگر در بین پیکسل‌های ۲۵ تا ۴۵ حداقل ۱۰ نقطه وجود داشته باشد که مقدار پروجکشن عمودی در آن بزرگ‌تر از ۱۰ بوده و همچنین در بین پیکسل‌های ۱۰ تا ۲۵ و ۴۵ تا ۶۰ آن نیز حداقل ۵ نقطه وجود داشته باشد که مقدار پروجکشن عمودی در آنها کمتر از ۱۰ باشد. شکل ۱۶ پروجکشن عمودی حاصل از نیمه پایین یک تکه تصویر غیر صورت را نشان می‌دهد. در این شکل نیمه پایین تکه تصویر مطابق الگوی شاخص دهان نمی‌باشد. در شکل ۱۷ بعضی از نمونه‌های خطای شناسایی شده توسط شاخص دهان آورده شده است.

۳-۳-۳ شاخص چشمها

در یک بررسی آماری بر روی ناحیه دو چشم چپ و راست در خروجی فیلتر گابور از ۳۸ تصویر صورت مختلف مشخص می‌شود که در ناحیه چشم چپ (۸) و در ناحیه چشم راست (۹) برقرار است

$$22/8 \leq \frac{\sum_{i=1}^{n/r} \sum_{j=1}^{n/r} T(i, j)}{n^2/4} \times 100 \leq 57 \quad (8)$$

$$22/8 \leq \frac{\sum_{i=(n/r)+1}^n \sum_{j=1}^{n/r} T(i, j)}{n^2/4} \times 100 \leq 57 \quad (9)$$

که در آن T یک تکه تصویر در اندازه $n \times n$ می‌باشد. مطابق (۸) و (۹) چشم راست و چپ بین ۲۲/۸ تا ۵۷ درصد از ناحیه خود را اشغال کرده‌اند و بنابراین اگر مقادیری خارج از این الگو در تصویری وجود داشته باشد آن نمونه صورت نخواهد بود. شکل ۱۸ چند نمونه از خطاهایی که با استفاده از این شاخص تشخیص داده شده است را نشان می‌دهد.

مطابق (۶) یک تکه تصویر صورت است اگر در بین پیکسل‌های ۱۰ تا ۳۰ و ۴۰ تا ۶۰ یک نقطه وجود داشته باشد که مقدار پروجکشن عمودی در آن بزرگ‌تر مساوی ۱۶ بوده و همچنین در بین پیکسل‌های ۲۰ تا ۴۵ نقطه‌ای وجود داشته باشد که مقدار پروجکشن عمودی در آن کمتر از ۱۰ باشد. شکل ۱۳ پروجکشن عمودی حاصل از نیمه بالای یک تکه تصویر غیر صورت را نشان می‌دهد. این شکل مطابق الگوی شاخص بینی نمی‌باشد. در شکل ۱۴ بعضی از نمونه‌های خطای شناسایی شده توسط شاخص بینی آورده شده است.

۳-۳-۲ شاخص دهان

با دقت در شکل ۱۱ می‌توان دریافت پیکسل‌های سفید در نیمه پایین این تصاویر نشان‌دهنده ناحیه دهان است. برای به دست آوردن الگوی دهان، پروجکشن عمودی در نیمه پایین خروجی فیلتر گابور در یک تکه تصویر صورت اجرا می‌شود. شکل ۱۵ نتیجه حاصل از اجرای پروجکشن عمودی را بر روی نیمه پایین خروجی فیلتر گابور در یک نمونه تصویر از صورت نشان می‌دهد.

با بررسی پروجکشن عمودی نیمه پایین ۳۸ تکه تصویر صورت مختلف شاخص دهان به صورت (۷) تعریف می‌شود

جدول ۴: تأثیر سه شاخص در آشکارسازی نمونه‌های صحیح صورت و کاهش تعداد خطای آشکارسازی نمونه‌های غیر صورت.

کل زیرتصاویر ۴۴۲، زیرتصاویر بدون صورت ۳۹۰ و زیرتصاویر شامل صورت ۵۲ مورد	تعداد آشکارسازی صحیح صورت	تعداد خطای آشکارسازی غیر صورت
آشکارسازی قبل از اعمال سه شاخص	۵۲	۶۳
آشکارسازی بعد از اعمال سه شاخص	۴۷	۹

جدول ۵: بررسی تأثیر شاخص‌های شناسایی اجزای صورت در کاهش تعداد خطای آشکارسازی به صورت جداگانه.

شاخص	بینی	لب	چشم راست	چشم چپ
کاهش تعداد خطا	۳۳	۳۳	۱۸	۲۴
کاهش تعداد خطا به تنهایی	۷	۱۰	۴	۱
کاهش تعداد صورت	۱	۳	۳	۲
کاهش تعداد صورت به تنهایی	۰	۲	۱	۰

جدول ۶: مقایسه‌ای بین پایگاه تصاویر RFD و CMU+MIT.

	CMU+MIT	RFD
تعداد تصاویر	۱۳۰	۱۰۰
تعداد چهره	۵۰۷	۵۸۳
کیفیت تصاویر	متوسط به پایین	متوسط به بالا
مجموع پیکسل‌های تشکیل‌دهنده تصاویر	حدود ۳۰ میلیون	بیش از ۹۰ میلیون

۴-۱ نحوه تولید و مشخصات پایگاه تصاویر RFD

برای بررسی عملکرد الگوریتم پیشنهادی در این مقاله پایگاه تصاویر جدیدی به نام RFD شامل ۱۰۰ تصویر چندچهره از اینترنت با بیش از ۹۰ میلیون پیکسل جمع‌آوری می‌شود. این مجموعه تصاویر جدید با استفاده از روش بیان‌شده در بخش ۲-۲ به ۱۱۰۱۵ زیرتصاویر در اندازه 168×192 تبدیل شده که در ۹۹۶۱ زیرتصویر آن صورتی وجود نداشته و در ۵۸۳ زیرتصویر آن نیز چهره غیر تکراری وجود دارد. با توجه به این که در بیش از ۹۰٪ از زیرتصاویر تولیدشده در این پایگاه تصاویر صورتی وجود ندارد لذا این پایگاه از پس‌زمینه زیادی برخوردار بوده و در نتیجه احتمال آشکارسازی نمونه‌های غیر صورت به عنوان صورت در آن زیاد می‌باشد. در جدول ۶ مقایسه‌ای بین مشخصات پایگاه RFD و CMU+MIT آورده شده و جدول ۷ مشخصات تصاویر پایگاه RFD را نشان می‌دهد.

۴-۲ بررسی نتایج حاصل از اجرای روش پیشنهادی در پایگاه تصاویر RFD

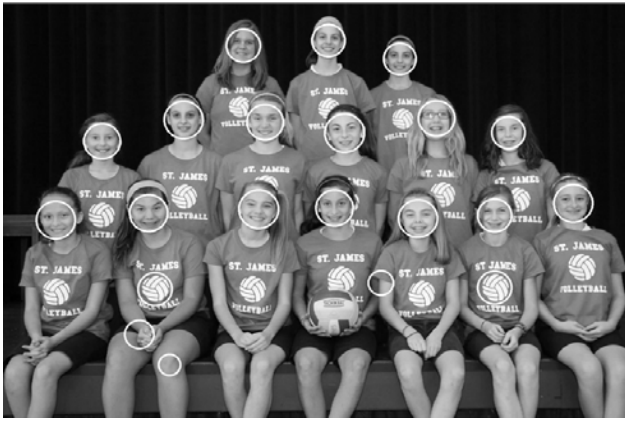
به منظور بررسی عملکرد روش پیشنهادی، آشکارسازی صورت بر روی پایگاه تصاویر RFD توسط الگوریتم پیشنهادی و الگوریتم V.J. انجام می‌شود. جدول ۸ نتایج حاصل از اجرای این دو الگوریتم را نشان می‌دهد. همان‌طور که مشخص است الگوریتم پیشنهادی آشکارسازی نمونه‌های صحیح صورت را با دقت ۸۸٫۱۶٪ انجام داده و از بین ۹۹۶۱ زیرتصویر بدون صورت تنها در ۴۸ مورد دچار خطای آشکارسازی شده و این در حالی است که الگوریتم V.J. در این پایگاه تصاویر با دقت ۹۹٫۱۴٪ نمونه‌های صحیح صورت را آشکارسازی نموده و تنها ۱۲۴ مورد خطای آشکارسازی دارد. این اعداد نشان‌دهنده آن است که روش پیشنهادی اگرچه ۱۱٪ در شناسایی نمونه‌های صحیح صورت نسبت به الگوریتم V.J. دقت کمتری دارد ولی در شناسایی نکردن نمونه‌های غیر صورت به عنوان صورت بیش از ۶۰٪ از الگوریتم ویولا جونز موفق‌تر عمل کرده است. شکل ۱۹ نتایج حاصل از آشکارسازی صورت در روش پیشنهادی و روش V.J. را بر روی دو نمونه تصویر از پایگاه RFD نشان می‌دهد. تصاویر الف و ج خروجی الگوریتم V.J. بوده و تصاویر ب و د خروجی الگوریتم پیشنهادی می‌باشد. در تصویر الف الگوریتم V.J.، ۱۸ مورد نمونه غیر صورت را به عنوان صورت آشکارسازی کرده در حالی که در تصویر ب روش پیشنهادی هیچ خطایی در آشکارسازی نمونه غیر صورت نداشته است. در تصویر ج روش V.J. از ۱۶ چهره ۱۶ مورد را آشکارسازی کرده و

۴-۳-۴ بررسی تأثیر شاخص‌های شناسایی اجزای صورت در کاهش خطای آشکارسازی

در این بخش ابتدا به صورت تصادفی ۸ تصویر چندچهره از مجموعه تصاویر پایگاه RFD انتخاب شده و با استفاده از الگوریتم تولید زیرتصاویر بیان‌شده در بخش ۲-۲ به ۴۴۲ زیرتصویر در اندازه 168×192 تبدیل شده که در ۳۹۰ زیرتصویر آن صورتی وجود ندارد. جدول ۴ نتایج حاصل از اجرای الگوریتم آشکارسازی صورت بدون استفاده از سه شاخص و همراه با استفاده از سه شاخص را در این مجموعه تصاویر نشان می‌دهد. همان‌طور که مشخص است الگوریتم آشکارسازی صورت بدون استفاده از سه شاخص در آستانه ۰٫۸۵، ۵۲ چهره را به درستی آشکار کرده و ۶۳ خطای آشکارسازی دارد. با اجرای الگوریتم آشکارسازی صورت با استفاده از سه شاخص خطای آشکارسازی از ۶۳ به ۹ مورد کاهش یافته و ۵ تصویر سالم صورت نیز به اشتباه حذف می‌شوند. بنابراین در این مجموعه تصاویر موفقیت سه شاخص در کاهش خطای آشکارسازی ۸۵٫۷٪ بوده است. جدول ۵ تأثیر هر یک از سه شاخص در به دست آمدن نتایج جدول ۴ را به تفکیک نشان می‌دهد. همان‌طور که مشخص است شاخص بینی با حذف ۳۳ مورد خطا تنها باعث حذف یک صورت صحیح شده و بهترین کارایی را در بین شاخص‌های دیگر داشته و این در حالی است که از این ۳۳ خطا ۷ مورد آن را به تنهایی شناسایی کرده است. شاخص دهان نیز از موفقیت بالایی برخوردار بوده ولی باعث حذف دو نمونه صورت شده است.

۴-۴ بررسی عملکرد الگوریتم پیشنهادی

در این قسمت به بررسی عملکرد الگوریتم پیشنهادی در پایگاه تصاویر RFD و مقایسه نتایج آن با دیگر روش‌ها پرداخته می‌شود. برای این منظور در بخش اول به نحوه تولید پایگاه تصاویر RFD و بیان مشخصات آن پرداخته و در بخش دوم به بررسی عملکرد الگوریتم پیشنهادی در این پایگاه تصاویر پرداخته می‌شود. در قسمت سوم به تحلیل خطا در روش پیشنهادی پرداخته و در قسمت چهارم مقایسه‌ای بین نتایج حاصل از اجرای الگوریتم پیشنهادی و دیگر روش‌ها آورده شده است.



(ج)



(الف)



(د)



(ب)

شکل ۱۹: مقایسه عملکرد روش پیشنهادی و روش V.J. در دو تصویر از پایگاه RFD، (الف) و (ج) خروجی روش V.J. و (ب) و (د) خروجی روش پیشنهادی.

جدول ۸: مقایسه نتایج حاصل از اجرای الگوریتم پیشنهادی و V.J. از پایگاه تصویر RFD.

نام روش	تعداد	درصد (%)	خطای آشکارسازی در ۹۹۶۱ زیرتصویر بدون صورت	تعداد	درصد (%)
پیشنهادی	۵۱۴	۸۸٫۱۶	۴۸	۰٫۴۸	
ویولا جونز	۵۷۸	۹۹٫۱۴	۱۲۴	۱٫۲۴	

نشان می‌دهد. دسته چهارم خطاهایی است که خروجی انرژی گابور در آنها با خروجی فیلتر گابور در نمونه صورت بسیار شبیه است. شکل ۲۳ چند نمونه از این نوع خطاها را نشان می‌دهد.

در بررسی دلایل عدم آشکارسازی صحیح نمونه‌های صحیح صورت می‌توان به وجود عینک در چهره افراد اشاره کرد. وجود عینک باعث می‌شود تا حجم ناحیه چشم در خروجی فیلتر گابور از اندازه طبیعی بیشتر شده و این عامل باعث می‌شود تا شاخص تعریف‌شده برای چشم آن را به عنوان نمونه غیر صورت شناسایی نماید. در شکل ۲۴ یک نمونه از خطای عدم آشکارسازی چهره‌های با عینک توسط روش پیشنهادی آورده شده است.

۴-۴ مقایسه الگوریتم پیشنهادی با دیگر روش‌ها

همان‌طور که در مقدمه نیز اشاره شد اغلب مقالات ارزیابی روش خود را بر روی مجموعه تصاویر پایگاه CMU+MIT که از تصاویر با کیفیت متوسط به پایین برخوردار است انجام می‌دهند. از آنجایی که پارامترهای فیلتر گابور در روش پیشنهادی برای عملکرد در تصاویر با کیفیت متوسط به بالا تنظیم شده است لذا امکان ارزیابی آن در پایگاه CMU+MIT وجود نداشته و در نتیجه نمی‌توان مقایسه مستقیمی بین روش پیشنهادی

۵ خطا در آشکارسازی نمونه غیر صورت به عنوان صورت دارد در حالی که در تصویر د روش پیشنهادی از ۱۶ چهره ۱۵ مورد را آشکارسازی نموده و تنها یک مورد خطا در آشکارسازی نمونه غیر صورت دارد.

۴-۳ تحلیل خطاها در روش پیشنهادی

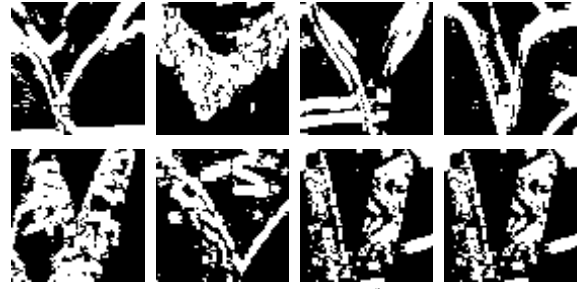
تحلیل خطاها در روش پیشنهادی شامل تحلیل دلایل خطای آشکارسازی نمونه غیر صورت به عنوان صورت و تحلیل دلایل خطای عدم آشکارسازی نمونه‌های صحیح صورت در روش پیشنهادی می‌باشد. در بررسی ۴۸ مورد خطای رخ داده در آشکارسازی نمونه‌های غیر صورت به عنوان صورت می‌توان تقسیم‌بندی زیر را انجام داد. یک دسته خطاها مربوط به آشکارسازی تصاویر هفتگی شکل مانند یقه در لباس‌های یقه هفت، پارچه آویز مدال بازیکنان در مسابقات و همچنین کراوات‌هایی است که به صورت آزاد بسته شده‌اند. شکل ۲۰ چند نمونه از این نوع خطاها را نشان می‌دهد.

دسته دوم مربوط به بعضی حروف انگلیسی مانند M و J است که به صورت هنری نوشته شده باشند. شکل ۲۱ چند نمونه از این نوع خطاها را نشان می‌دهد.

دسته سوم مربوط به نواحی است که دارای پس‌زمینه چمن بوده یا چیدمان در هم ریخته‌ای دارند. شکل ۲۲ چند نمونه از این نوع خطاها را



شکل ۲۳: نمونه‌هایی از خطای آشکارسازی نمونه غیر صورت به عنوان صورت در روش پیشنهادی (گروه تصاویر چهره‌مانند).



شکل ۲۰: نمونه‌هایی از خطای آشکارسازی نمونه غیر صورت به عنوان صورت در روش پیشنهادی (گروه تصاویر هفتی‌شکل).



شکل ۲۱: نمونه‌هایی از خطای آشکارسازی نمونه غیر صورت به عنوان صورت در روش پیشنهادی (گروه بعضی حروف انگلیسی).



شکل ۲۲: نمونه‌هایی از خطای آشکارسازی نمونه غیر صورت به عنوان صورت در روش پیشنهادی (گروه تصاویر چمنی‌شکل).



شکل ۲۴: نمونه‌هایی از خطای عدم آشکارسازی نمونه صحیح صورت در روش پیشنهادی (تصویر فرد با عینک).

جدول ۹: مقایسه نتایج حاصل از الگوریتم V.J. در منابع مختلف بر روی پایگاه CMU+MIT و پیش‌بینی عملکرد هر یک از آنها در پایگاه RFD.

تعداد خطا در RFD	تعداد خطا در CMU+MIT	آشکارسازی در CMU+MIT (%)	پایه‌سازی روش V.J.
۸۲۶	۱۰۰	۹۳	[۲۶]
۱۴۰۵	۱۷۰	۹۴	[۱۷]
۴۱۳	۵۰	۹۱٫۷	[۲۵]
۱۲۴	۱۵	۵۷	این مقاله

ترتیب دارای بیش از $۱۴۰۵ \approx ۱۷۰/۱۵ \times ۱۲۴$ ، $۴۱۳ \approx ۵۰/۱۵ \times ۱۲۴$ و $۸۲۶ \approx ۱۰۰/۱۵ \times ۱۲۴$ مورد خطا در آشکارسازی صورت خواهد بود. این در حالی است که روش پیشنهادی تنها ۴۸ مورد خطا در این پایگاه تصویر دارد.

مقایسه‌ای دیگر را می‌توان بر اساس پیچیدگی تصاویر در دو پایگاه CMU+MIT و RFD انجام داد. همان‌طور که بیان شد روش V.J. پایه‌سازی شده در این مقاله در پایگاه CMU+MIT دارای ۱۵ مورد خطای آشکارسازی بوده در حالی که در پایگاه RFD دارای ۱۲۴ مورد خطای آشکارسازی می‌باشد. حال اگر بپذیریم که پیچیدگی تصاویر در پایگاه RFD بیش از $۸ \approx ۱۲۴/۱۵$ برابر پایگاه CMU+MIT می‌باشد بنابراین مطابق جدول ۱۰ می‌توان پیش‌بینی کرد که در صورت اجرای روش‌های ارائه‌شده در [۱۷]، [۲۵] و [۲۶] بر روی پایگاه RFD هر یک از این روش‌ها به ترتیب دارای ۱۶۰، ۲۰۰ و ۳۲ خطای آشکارسازی در دقت حدود ۹۰٪ خواهند شد و این در حالی است که الگوریتم پیشنهادی تنها ۴۸ مورد خطا در پایگاه RFD دارد.

۵- نتیجه‌گیری

این مقاله به آشکارسازی صورت در تصاویر چندچهره با استفاده از ترکیب فیلترهای گابور و شبکه عصبی پرداخت و موفق شد با حفظ نسبی دقت آشکارسازی نمونه‌های صحیح صورت از میزان خطای آشکارسازی نمونه‌های غیر صورت به عنوان صورت نیز به شدت بکاهد. کاهش خطای

روش ارائه‌شده در سایر مقالات انجام داد ولی می‌توان آنها را به صورت غیر مستقیم با یکدیگر مقایسه نمود. برای این منظور در ابتدا به بررسی نتایج حاصل از آشکارسازی صورت در تصاویر چندچهره در چند مقاله مختلف پرداخته و در ادامه به مقایسه آنها به صورت غیر مستقیم با روش پیشنهادی می‌پردازیم.

مراجع [۱۷]، [۲۵] و [۲۶] روش خود را در پایگاه تصاویر CMU+MIT مورد ارزیابی قرار داده و در این راه روش خود را با نتایج حاصل از چندین روش دیگر از جمله الگوریتم V.J. مقایسه کرده‌اند. مطابق جدول ۹ مشخص است که نتایج ارائه‌شده از اجرای الگوریتم V.J. بر روی پایگاه تصاویر CMU+MIT در این سه مرجع به ترتیب ۹۴٪، ۹۱٫۷٪ و ۹۳٪ است در حالی که تعداد خطای آشکارسازی در آنها به ترتیب ۱۷۰، ۵۰ و ۱۰۰ مورد می‌باشد. دلیل این تفاوت را می‌توان در روش‌های مختلف پایه‌سازی الگوریتم V.J. در این سه منبع دانست. این در حالی است که الگوریتم V.J. که در این مقاله پایه‌سازی شده است در پایگاه CMU+MIT دارای دقت آشکارسازی صورت ۵۷٪ و ۱۵ مورد خطای آشکارسازی می‌باشد. این بدان معنی است که الگوریتم V.J. مورد استفاده در این مقاله نسبت به روش‌های مورد استفاده در سه مرجع مذکور به ترتیب بیش از $۱۱ \approx ۱۷۰/۱۵$ ، $۳/۳ \approx ۵۰/۱۵$ و $۶ \approx ۱۰۰/۱۵$ برابر از توانایی بالاتری برای شناسایی نکردن نمونه‌های غیر صورت برخوردار است. حال با این توضیح می‌توان پیش‌بینی کرد وقتی الگوریتم V.J. پایه‌سازی شده در [۱۷]، [۲۵] و [۲۶] در پایگاه RFD اجرا شود به

جدول ۱۰: مقایسه نتایج حاصل از آشکارسازی صورت در منابع مختلف بر روی پایگاه CMU+MIT و پیش‌بینی تعداد خطای آشکارسازی هر یک از این روش‌ها در پایگاه RFD.

منابع مختلف	آشکارسازی در CMU+MIT (%)	تعداد خطا در CMU+MIT	تعداد خطا در RFD
[۲۶]	۹۶	۱۰۰	۸۰۰
[۲۶]	۹۰	۴	۳۲
[۱۷]	۹۶	۱۶۵	۱۳۲۰
[۱۷]	۹۰	۲۰	۱۶۰
[۲۵]	۹۱٫۷	۵۰	۴۰۰
[۲۵]	۹۰	۲۵	۲۰۰
روش پیشنهادی	-	-	۴۸

- [3] A. Lumini, B. Nanni, and F. Dominio, "Effective and precise face detection based on both gray-level image and depth map," *Applied Computing and Informatics*, Apr. 2014.
- [4] D. Kriegman, M. Yang, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34-58, Jan. 2002.
- [5] S. Kim, S. -K Kim, K. Toh, Y. Ban, and S. Lee, "Face detection based on skin color likelihood," *Pattern Recognition*, vol. 47, no. 4, pp. 1573-1585, Apr. 2014.
- [6] X. Zhang, J. Yan, and Z. Lei, "Face detection by structural models," *Image and Vision Computing*, 2013.
- [7] R. Pereira de Magalhaes and C. Lima, "A new method for Haar-Like features weight adjustment using principal component analysis for face detection," in *Proc. of the 9th Int. Conf. on Systems*, pp. 55-62, 2014.
- [8] M. Gopi Krishna and A. Srinivasulu, "Face detection system on AdaBoost algorithm using Haar classifiers," *International J. of Modern Engineering Research, IJMER*, vol. 2, no. 5, pp. 3556-3560, Sep./Oct. 2012.
- [9] S. Ulukaya, C. E. Erdem, and A. Karaali, "Combining Haar feature and skin color based classifiers for face detection," in *Proc. of IEEE IEEE Int. Conf. on Acoustics, Speech and Signal Processing, ICASSP'11*, pp. 1497-1500, Prague, Czech Republic, 22-27 May 2011.
- [10] T. Wang, J. Li, and Y. Zhang, "Face detection using SURF cascade," in *Proc. IEEE Int. Conf. on Computer Vision Workshops*, pp. 2183-2190, Barcelona, Spain 6-13 Nov. 2011.
- [11] D. Kim and R. Dahyot, "Face components detection using SURF descriptors and SVMs," in *Proc. Int. Machine Vision and Image Processing Conf., IMVIP '08*, pp. 51-56, Portrush, UK, 3-5 Sep. 2008.
- [12] D. Huang *et al.*, "Local binary patterns and its application to facial image analysis: A survey," *IEEE Trans. on Systems, Man, and Cybernetis*, vol. 41, no. 6, pp. 765-781, Nov. 2011.
- [13] Y. Zhu, H. Pan, and L. Xia, "Efficient and accurate face detection using heterogeneous feature descriptors," *Computer Vision and Image Understanding*, vol. 117, no. 1, pp. 12-28, Jan. 2013.
- [14] M. Fische, H. Ekenel, F. Jiang, and E. Shi, "Combining texture and stereo disparity cues for real-time face detection," *Signal Processing: Image Communication*, vol. 28, no. 9, pp. 1100-1113, Oct. 2013.
- [15] A. Shimizu, H. Kobatake, and L. Huang, "Robust face detection using Gabor filter features," *Pattern Recognition Letters*, vol. 26, no. 11, pp. 1641-1649, Aug. 2005.
- [16] S. Prakash Sahu, "Face detection by fine tuning the Gabor filter parameter," *International J. of Computer Science and Information Technologies*, vol. 2, no. 6, pp. 2719-2724, Nov./Dec 2011.
- [17] L. Xiaohua, K. Lam, S. Lansun, and Z. Jiliu, "Face detection using simplified Gabor features and hierarchical regions in a cascade of classifiers," *Pattern Recognition Letters*, vol. 30, no. 8, pp. 717-728, Jun. 2009.
- [18] N. Batool and R. Chellappa, "Fast detection of facial wrinkles based on Gabor features using image morphology and geometric constraints," *Pattern Recognition*, vol. 48, no. 3, pp. 642-658, Mar. 2015.
- [19] M. Gaianu and D. M. Onchis, "Face and marker detection using Gabor frames on GPUs," *Signal Processing*, vol. 96, pp. 90-93, Mar. 2014.

[۲۰] م. محلوجی و ر. محمدیان، "آشکارسازی صورت با استفاده از فیلترهای گابور و شبکه‌های عصبی"، *نشریه مهندسی برق و مهندسی کامپیوتر ایران*، سال ۱۳، شماره ۱-ب، صص. ۱۳-۱، تابستان ۱۳۹۴.

- [21] S. Kaneda, R. Ariizumi, and H. Haga, "Energy saving of TV by face detection," *Pervasive Technologies Related to Assistive Environments*, 2008.

مذکور بر اساس این یافته‌ها محقق گردید: در یک بررسی تجربی مشخص شد که تصاویر اینترنتی عمدتاً دارای تغییر رنگ‌های نامحسوسی در سطح خود هستند که فیلتر گابور نسبت به آنها عکس‌العمل نشان داده و باعث از دست رفتن کیفیت در خروجی فیلتر گابور و بروز خطا در آشکارسازی صورت می‌شوند. برای رفع این مشکل پیش‌پردازش یکنواخت‌ساز رنگ معرفی گردید. بررسی تأثیر این پیش‌پردازش در چند نمونه تصویر چندچهره نشان داد که این پیش‌پردازش می‌تواند نزدیک به ۴۰٪ شناسایی نمونه‌های صحیح صورت را افزایش داده و همچنین بیش از ۶۶٪ نیز خطای آشکارسازی نمونه‌های غیر صورت را کاهش دهد.

از بررسی عملکرد شبکه عصبی در آشکارسازی تکه تصاویر شامل صورت و غیر صورت مشخص گردید که شبکه عصبی در روش پیشنهادی بر این اساس آموزش یافته است که در صورت وجود دو باند سفیدرنگ در بالا و پایین یک تکه تصویر، با احتمال بالایی آن تکه تصویر را به عنوان صورت شناسایی کرده و بر اساس شکل واقعی اجزای صورت تصمیم‌گیری نمی‌کند. این مقاله با استفاده از سه شاخص شناسایی بینی، دهان و چشم موفق شد این نوع خطای شبکه عصبی را نیز اصلاح نماید. این شاخص‌ها که با استفاده از اجرای پروجکشن عمودی بر روی خروجی انرژی گابور به دست می‌آیند توانستند میزان خطای آشکارسازی را به میزان ۸۵٫۷٪ کاهش دهند.

در این مقاله پایگاه تصاویر جدیدی به نام RFD شامل ۱۰۰ تصویر چندچهره با پس‌زمینه‌های پیچیده و کیفیت متوسط به بالا از اینترنت جمع‌آوری گردید. پایگاه جدید دارای ۵۸۳ چهره غیر تکراری و ۱۱۰۱۵ زیرتصویر در اندازه ۱۶۸×۱۹۲ می‌باشد. نتایج اجرای الگوریتم پیشنهادی در پایگاه RFD نشان داد که دقت آشکارسازی نمونه‌های صحیح صورت ۸۸٫۱۶٪ بوده و خطای آشکارسازی نمونه‌های غیر صورت در آن نیز ۰٫۴۸٪ معادل ۴۸ خطا در ۹۹۶۱ زیرتصویر بدون چهره است. این در حالی است که الگوریتم V.J. پیاده‌سازی شده در این مقاله در پایگاه RFD دارای ۱۲۴ خطا می‌باشد.

از بررسی ۴۸ خطای رخ داده در روش پیشنهادی مشخص شد که بخش عمده‌ای از این تعداد خطا به تکه تصاویر هفتی‌شکل مانند پیراهن‌های یقه هفت، گردن‌آویز مدال ورزشکاران یا کراوات‌های آزاد بسته‌شده مربوط می‌شود. این خطا بدان جهت است که در خروجی فیلتر گابور شباهت بالایی بین این نوع تکه تصاویر و قالب صورت وجود دارد.

مراجع

- [1] N. Rathore and D. Chaubey, "A survey on face detection and recognition," *International J. of Computer Architecture and Mobility*, vol. 1, no. 5, 6 pp., Mar. 2013.
- [2] C. Zhang and Z. Zhang, *A Survey of Recent Advances in Face Detection*, Technical Report, Microsoft Research, Jun. 2010.

رضا محمدیان در سال ۱۳۸۳ با رتبه ممتاز اول مدرک کارشناسی مهندسی کامپیوتر گرایش نرم‌افزار و در سال ۱۳۹۳ با رتبه ممتاز اول مدرک کارشناسی ارشد مهندسی مکترونیک خود را از دانشگاه آزاد اسلامی واحد کاشان دریافت نمود. نام‌برده از سال ۱۳۹۴ مشغول به تحصیل در مقطع دکتری در مهندسی کامپیوتر گرایش نرم‌افزار در دانشگاه آزاد اسلامی واحد قم است. ایشان از سال ۱۳۸۵ تاکنون نیز به عنوان مدیرعامل در بخش خصوصی مشغول به کار می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: پردازش تصویر، شبکه‌های عصبی و رباتیک.

محمود محلوجی در سال ۱۳۶۸ مدرک کارشناسی مهندسی برق گرایش مخابرات و در سال ۱۳۷۳ مدرک کارشناسی ارشد مهندسی برق گرایش الکترونیک خود را از دانشگاه صنعتی شریف دریافت نمود. از سال ۱۳۷۴ تاکنون نام‌برده به عنوان عضو هیئت علمی دانشکده برق و کامپیوتر در دانشگاه آزاد اسلامی واحد کاشان به کار مشغول بوده و در سال ۱۳۸۷ موفق به اخذ درجه دکتری در مهندسی برق گرایش مخابرات از دانشگاه آزاد اسلامی واحد علوم و تحقیقات تهران گردید. زمینه‌های تحقیقاتی مورد علاقه ایشان شامل موضوعاتی نظیر پردازش تصویر، بینایی ماشین و شناسایی الگو می‌باشد.

- [22] W. Eric, E Wong, and Tingshao, "Computer power management system based on the face detection," in *Computer Engineering and Networking: Springer International Publishing*, pp. 913-920, 2014.
- [23] P. Viola and M. J. Jones, "Robust real-time face detection," *International J. of Computer Vision*, vol. 57, no. 2, pp. 137-154, May 2004.
- [24] S. Adwan and H. Arof, "Modified integral projection method for eye detection using dynamic time warping," *International J. of Innovation Computing, Information, and Control*, vol. 8, no. 1-A, pp. 187-189, Jan. 2012.
- [25] J. Meynet, V. Popovici, and J. Thiran, "Face detection with boosted Gaussian features," *Pattern Recognition*, vol 40, no. 8, pp. 2283-2291, Aug. 2007.
- [26] S. Shan, X. Chen, S. Yan, and W. Gao, "Locally assembled binary (LAB) feature with feature-centric cascade for fast and accurate face detection," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition, CVPR'08*, 7 pp., Anchorage, AK, USA, 23-28 Jun. 2008.

Archive of SID