

Face Recognition in Images Using Viola_Jones Method and Image Texture Analysis

Mahdi Hariri^{1*}, Narmineh Heydarzadeh²

1. Assistant Professor, Department of Electrical and Computer Engineering, University of Zanzan, Zanzan, Iran. *Corresponding Author, ma.hariri@znu.ac.ir
2. MS Student, Department of Electrical and Computer Engineering, Zanzan Branch, Islamic Azad University, Zanzan, Iran.

Abstract

Introduction: Face recognition is one of the most important biometric technologies for people identification, also used in access control. Face recognition is one of the important steps before identity recognition. Usually, one method is used to detect the presence of faces in images. Still, in this research, to increase the accuracy of detection, the combination of two methods is used to improve the detection performance: Viola-Jones algorithm and the matching of image components and texture with face and skin components. In the first step, we use the Viola-Jones algorithm to detect the facial features. In the next step, the features of the eye and nose tissues are analyzed with regression neural networks, and facial features are recognized better according to the facial features. In this research, the texture features of the right and left eyes and the nose of the face are used to increase the matching accuracy. We have used the faces of the FDD-Fold dataset to evaluate the proposed method. Comparing the performance of this method with the RCNN deep network method with a much smaller number of training data, we reached an accuracy of 96.36%, more than the deep learning network. This method gives good results in systems with limited computing ability and average amount of data.

The face recognition system is one of the biometric identification systems and one of the most important technologies for people identification, which is also used in access control. Face identification is one of the few biometric methods that, with the advantages of high accuracy and low level of human intervention, is used in cases such as information security, law enforcement and monitoring, traffic control, and registration in attendance systems. This method creates more convenience and development with fewer requirements. then, this method has received more attention during the last twenty years.

Face detection is a local binary classification problem that shows the presence of faces in the given image using boxes surrounding them. Although the Viola-Jones method is less accurate than modern methods such as convolutional neural networks; Its much lower efficiency and training parameters compared to the millions of parameters of a typical CNN result in faster training, better accuracy with limited data, and its use in devices with limited computing power such as cameras and mobile phones. The innovation of this method is matching the geometric pattern of the edges to identify the presence of the face in the image, along with matching the skin texture. This method seems to be faster and more accurate than the previous ones.

Method: In this research, in the first step, we use Viola-Jones, one of the optimal face recognition algorithms in the image, to detect facial components. In the next step, we use the adaptation of the general shape of facial parts such as eyes, and match the textures in the image with the predicted texture for human skin, to improve the recognition performance and increase the recognition accuracy, in such a way that the regression neural networks examine the eye and nose tissue characteristics and according to the characteristics of the facial tissue, the facial components are recognized by the regression neural network. The investigated features in the texture include minimum and maximum color intensity, mean and median, and variance of the image. The data is given to the regression neural network for training. Here Remarkable thing is matching the overall shape of the human head and face, and in the next step matching the overall shape of the facial parts such as the eyes to improve the accuracy of the presented method. We also use the matching of textures in the image with the texture predicted for human skin to further improve the accuracy of the program's performance.

Keywords: Face recognition, Viola-Jones algorithm, Integral Image, Regression Neural Network, Eye and Nose Matching, Accuracy.

تشخیص چهره در تصویر با استفاده از روش ویولا-جونز و تحلیل بافت تصویر

دوره چهارم، تابستان ۱۴۰۲
شماره دوم، صص: ۱-۱۰

تاریخ دریافت: ۱۴۰۲/۰۱/۲۹
تاریخ پذیرش: ۱۴۰۲/۰۳/۱۱

مهدی حریری^{۱*}، نریمینه حیدرزاده^۲

۱- استادیار، گروه مهندسی برق و کامپیوتر، دانشگاه زنجان، زنجان، ایران. (نویسنده مسئول) ma.hariri@znu.ac.ir
۲- کارشناسی ارشد، مهندسی کامپیوتر، واحد زنجان، دانشگاه آزاد اسلامی، زنجان، ایران. na_heydarzadeh@yahoo.com

چکیده: شناسایی چهره از مهمترین فناوری‌های بیومتریک برای شناسایی افراد است که در کنترل دسترسی هم استفاده دارد. تشخیص چهره یکی از مراحل مهم قبل از شناسایی هویت می باشد. معمولاً برای تشخیص وجود چهره در تصاویر از یک روش استفاده شده است، اما در این تحقیق برای افزایش صحت تشخیص از ترکیب دو روش الگوریتم ویولا-جونز و تطبیق اجزا و بافت تصویر با اجزا چهره و پوست برای بهبود عملکرد تشخیص استفاده می شود. در مرحله اول با استفاده از الگوریتم ویولا-جونز به تشخیص اجزای صورت می پردازیم و در مرحله بعدی با شبکه‌های عصبی رگرسیون ویژگی‌های بافت چشم و بینی را مورد بررسی قرار داده و با تطبیق ویژگی‌های بافت صورت، اجزاء صورت بهتر تشخیص داده می شوند. در این تحقیق از ویژگی‌های بافت مربوط به چشم راست و چپ و بینی در صورت برای افزایش دقت تطبیق استفاده می شود. برای ارزیابی روش پیشنهادی از چهره‌های مجموعه داده FDD-Fold استفاده کرده ایم. در مقایسه عملکرد این روش با روش شبکه عمیق RCNN با تعداد خیلی کمتر داده‌های آموزشی نسبت به آن به صحت ۹۶٫۳۶٪، بیشتر از شبکه یادگیری عمیق رسیدیم. این روش در سیستم‌های با توانایی محاسباتی محدود و تعداد داده متوسط نتیجه مطلوبی می دهد.

واژه‌های کلیدی: تشخیص چهره، الگوریتم ویولا-جونز، تصویر یکپارچه، شبکه عصبی رگرسیونی، تطبیق چشم و بینی، صحت تشخیص.

۱. مقدمه و بیان مسئله

فن آوری تشخیص چهره^۱ یکی از انواع روش‌های سیستم‌شناسایی بیومتریک محسوب می‌شود و از مهمترین تکنولوژی‌های تشخیص و شناسایی افراد است که در کنترل دسترسی نیز مورد استفاده قرار می‌گیرد. با وجود سیستم‌های بیومتریک قابل اعتمادی مانند تشخیص اثر انگشت و عنبیه چشم، سیستم تشخیص چهره، کاربر پسندتر بوده و بدون تماس کامل عضوی از بدن با سیستم یا حتی اطلاع او عملیات تشخیص و شناسایی انجام می‌گیرد. فن آوری بازشناسی چهره یکی از معدود روش‌های بیومتریک می‌باشد که با دارا بودن مزایای دقت بالا و سطح پایین دخالت فرد، در مواردی مانند امنیت اطلاعات، اجرا و نظارت بر قانون، کنترل و ثبت تردد در سیستم‌های حضور و غیاب مورد استفاده قرار می‌گیرد. این روش با ملزومات کمتر سهولت و توسعه بیشتری در روند تشخیص ایجاد می‌کند. به همین دلیل این فناوری در طی بیست سال گذشته در عرصه‌های صنعتی و علمی مورد توجه قرار گرفته است.

در این پژوهش در مرحله اول با یکی از الگوریتم‌های بهینه شناسایی چهره در تصویر یعنی الگوریتم ویولا-جونز به تشخیص اجزای صورت پرداخته می‌شود و در مرحله بعد از تطبیق شکل کلی اجزای صورت مانند چشم‌ها و تطبیق بافت‌های موجود در تصویر با بافت پیش‌بینی شده برای پوست انسان، در بهبود عملکرد تشخیص و افزایش صحت شناسایی استفاده می‌کنیم به این صورت که شبکه‌های عصبی رگرسیون و ویژگی‌های بافت چشم و بینی را بررسی کرده و با تطبیق ویژگی‌های بافت صورت، اجزای صورت توسط شبکه عصبی رگرسیونی تشخیص داده می‌شوند. ویژگی‌های مورد بررسی در بافت شامل حداقل و حداکثر شدت رنگ، میانگین و میانه و واریانس تصویر است. داده‌ها برای آموزش به شبکه عصبی رگرسیونی داده می‌شوند.

تشخیص چهره یک مسئله کلاسه‌بندی دوتایی به صورت محلی است که حضور چهره‌ها را در تصویر داده شده با استفاده از جعبه‌های احاطه کننده^۲ دور آن‌ها نشان می‌دهد. هرچند روش ویولا-جونز نسبت به روش‌های مدرن مانند شبکه‌های عصبی پیچشی صحت کمتری دارد؛ کارایی و پارامترهای آموزشی خیلی کمتر آن در مقایسه با میلیون‌ها پارامتر یک CNN نمونه موجب سرعت بیشتر آموزش، صحت بهتر با داده‌های محدودتر و استفاده از آن در ابزارهای محاسباتی محدود نظیر دوربین‌های عکاسی و موبایل‌ها می‌شود.

در این تحقیق سعی بر ابداع روشی برای بهبود دقت تشخیص حضور چهره انسان در تصویر داریم و آنچه مورد توجه است تطبیق شکل کلی سر و صورت انسان و در مرحله بعدی تطبیق شکل کلی اجزای صورت مانند چشم‌ها است تا صحت روش ارائه شده، بهبود یابد. از تطبیق بافت‌های موجود در تصویر با بافت پیش‌بینی شده پوست انسان نیز برای بهبود بیشتر دقت عملکرد برنامه استفاده می‌کنیم. در این تحقیق برآنیم که با توجه به روش‌های گوناگون طبقه‌بندی و مدل‌سازی ویژگی‌های

تصویر، با الگوریتمی کارآمد تشخیص محل حضور چهره‌ها در تصویر را بهبود بخشیم.

نوآوری این روش استفاده از تطبیق الگوی هندسی لبه‌ها برای شناسایی محل حضور چهره در تصویر، در کنار تطبیق بافت پوست می‌باشد و به نظر می‌آید این روش می‌تواند سریعتر و دقیق‌تر از روش‌های قبلی باشد. همچنین تشخیص، تنها با پردازش یک تصویر ثابت انجام می‌پذیرد و نیازی به بررسی الگوهای حرکتی نیز نمی‌باشد. در این تحقیق از الگوریتم ویولا-جونز - که در آن با استفاده از ماشین آداپوست ویژگی‌های نامطلوب در تشخیص مناسب حذف شده‌اند- برای تشخیص ناحیه صورت و اجزا و برای بهبود نتایج از شبکه عصبی رگرسیونی عمومی^۳ با توجه به بافت تصویر برای افزایش دقت تشخیص استفاده شده است.

۲. پیشینه پژوهش

تشخیص حضور چهره انسان در تصویر به دلیل نیاز به راه‌حل‌های جامع، سریع و مقاوم در برابر تغییرات صحنه تصویر و مسائل جانبی آن، موضوع ساده‌ای نیست و اغلب روش‌هایی که تاکنون برای آن ارائه شده‌اند علی‌رغم تلاش‌های بسیار، از ویژگی‌های مذکور فاصله داشته و دارای عملکردی کند بوده‌اند که نیازمند محاسبات پیچیده و زمانبرند. تلاش برای ارائه الگوریتم‌های بهتر تا جایی ادامه داشته است که حتی تعقیب حرکت‌ها در سری تصاویر نیز برای بهبود نتایج پیشنهاد شده است [۱]. در این مسیر استفاده از روش‌های معمولی و ساده بر اساس جستجوی رنگ پوست و مو مانند فیلتر کردن رنگ‌ها با روش کالمن [۲] و یا فیلتر کردن دانه‌ای رنگ [۳] و حتی فیلترینگ فازی رنگ‌های تصویر [۴] کارآمد نبوده‌اند. روش تعقیب حرکت یادشده در ترکیب با فیلتر رنگ به این ترتیب عمل می‌کند که حرکت مناطق حاوی رنگ‌های شبیه به رنگ پوست در تصاویر متوالی تعقیب شده و نسبت به تطبیق پروفایل حرکت آن با حرکت انسان اقدام می‌شود [۵] و حتی برای بهبود این روش از سیستم‌های تصویر برداری استریو با بیش از یک دوربین نیز استفاده شده است [۶]. از تطبیق هیستوگرام رنگ‌های صورت نیز در مرجع [۷] برای تشخیص چهره در تصویرگیری سه بعدی با چند دوربین استفاده شده است. روش‌های مبتنی بر تطبیق رنگ و حرکت به دلیل احتمال بالای بروز اشتباه، ناشی از عدم توجه به ویژگی‌های مشخص‌تر صورت مانند اجزاء و نمای کلی آن در کنار رنگ و بافت دیگر چندان مورد توجه نیستند.

از این روش‌های دیگر تشخیص حضور چهره در تصویر مانند تطبیق هیستوگرام گرادیان‌ها^۴ پیشنهاد شده‌اند [۸]. که در آن‌ها از تطبیق گرادیان مناطق مختلف تصویر با گرادیان تصویر صورت از روبرو استفاده می‌شود. نقطه ضعف این نوع روش‌ها نیز نیاز به دید کامل از روبرو و بار محاسباتی نسبتاً سنگین آن‌ها می‌باشد. این روش‌ها نیز در برابر مسائلی مانند استفاده از عینک و کلاه و یا تصاویر غیر روبروی صورت آسیب‌پذیر هستند. مطالعات مانپان از اطلاعات بافت پوست

انسان‌ها علی‌رغم رنگ‌های مختلف آن‌ها، برای شناسایی صورت مورد توجه قرار گرفته‌است [۹]. نکته قابل توجه اینکه پوست تمامی اجزاء بدن دارای بافت تقریباً یکسانی است و تکیه بر این ویژگی به تنهایی باعث بالارفتن امکان اشتباه می‌شود.

از جمله روش‌های دیگر نیز که عملکرد نسبتاً بهتری داشته‌است استفاده از تطبیق الگو برای تشخیص چهره است [۱۰]. در مقاله اکبری و حسینی، روشی مبتنی بر یادگیری عمیق جهت تشخیص چهره و تعیین نقاط شاخص چهره ارائه گردیده است. بدین منظور از شبکه‌های کانولوشنی چندمنظوره پی‌درپی استفاده شده است [۱۱]. دو موضوع تشخیص چهره و تعیین نقاط شاخص چهره همواره تحت تأثیر شرایط، سن، زوایا و شدت نور بوده‌اند.

در این تحقیق با ترکیب روش‌های بافت‌شناسی و تطبیق الگوی هندسی لبه‌های تصویر با الگوی از پیش تعریف شده سعی در شناسایی بهتر و مؤثرتر تصاویر دارای چهره و محل حضور چهره در تصاویر خواهیم داشت. روش تطبیق الگوی هندسی لبه‌ها، برای شناسایی محل حضور چهره در تصویر، در کنار تطبیق بافت پوست تاکنون مورد بررسی قرار نگرفته‌است و به نظر می‌آید این روش می‌تواند کاراتر از روش‌های قبلی باشد. همچنین تشخیص، تنها با پردازش یک تصویر انجام می‌پذیرد و نیازی به بررسی الگوهای حرکتی نیز نمی‌باشد.

۳. مبانی و تعاریف پایه‌ای

هر جسم در تصویر به صورت الگویی به هم پیوسته از رنگ یا بافت و با شکل مشخص قابل شناسایی است، این به معنای تکرار مقادیری ثابت یا نزدیک به هم در محدوده‌ای از تصویر می‌باشد، حال اگر از تصویر مشتق‌گیری کنیم به ماتریسی می‌رسیم که در نواحی با تکرار پیکسل‌های مشابه دارای مقدار صفر و در نواحی تغییر الگوی پیکسل‌ها، لبه‌های اجسام موجود در تصاویر، دارای مقادیر غیر صفر است. این مشتق‌گیری منجر به ظاهر شدن خطوطی در تصویر مشتق می‌گردد که در برخی موارد می‌توانند معرف مرزهای قرارگیری جسم در تصویر باشند و متصل نمودن این لبه‌ها می‌تواند منجر به تشکیل مرز اشیاء شود. لازم به ذکر است که روش‌های مبتنی بر استخراج لبه‌ها در

مورد تصاویری با لبه‌های متعدد مناسب نبوده و برای تصاویری با الگوهای اندک مناسب‌ترند. شناسایی لبه‌ها معمولاً با مشتق‌گیری از ماتریس خاکستری ممکن می‌باشد به این ترتیب که به‌ازای پیکسل‌هایی که میزان آن‌ها با پیکسل‌های همسایه برابر است صفر و به‌ازای پیکسل‌های دارای همسایگان تغییر یافته، مقدار عددی متناسب با شدت تغییر قراردادده می‌شود، لبه‌یابی در تصاویری می‌تواند مفید باشد که میزان تغییرات پیکسل‌های قرار گرفته در هر جسم اندک باشد و بدین ترتیب مشتق‌گیری منجر به آشکار شدن پیکسل‌های تشکیل‌دهنده مرز میان اجسام می‌شود [۱۲].

تطبیق الگو: در صورتی که ارائه شی با توجه به اجزا آن انجام شود این سؤال پیش می‌آید که کدام قسمت‌ها باید ارائه شوند. آیا انتخاب تعدادی

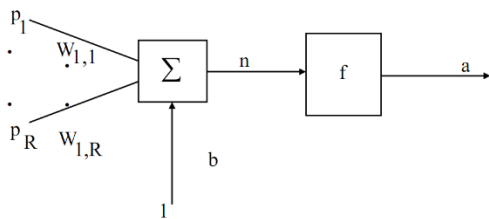
از ویژگی‌های برجسته و مهم شی مثلاً چشم‌ها، دهان و بینی کفایت یا باید اطلاعات تصویری موجود به مجموعه‌هایی کوچک تقسیم شده و از آن‌ها برای ارائه استفاده نمود؟

همجوشی ویژگی‌ها: به این معنی است که ابتدا تعدادی ویژگی‌های مهم و چشمگیر شیء مورد نظر را در نظری‌گیریم که بسته به شرایط ممکن است رنگ، بافت، مشخصات هندسی و یا موارد دیگر باشد. برای هر جسم یک بردار ویژگی‌ها با مقادیر عددی نرمالیزه شده ایجاد می‌کنیم. ایجاد این بردار ویژگی‌هاطوری که یک بردار مرجع داشته باشیم و بعد برای هر جسم استخراج شده از تصویر همین بردار را تشکیل داده و با بردار مرجع مقایسه کنیم همان‌طور که در [۱۳] بیان شده‌است، به‌عنوان همجوشی ویژگی‌ها شناخته می‌شود. در مواردی که تنها ویژگی‌های ساده مانند رنگ یا اندازه مد نظر است می‌توان به‌جای تشکیل بردار ویژگی‌ها، برای هر ویژگی در مورد اجسام موجود در صحنه یک هیستوگرام تشکیل داد و میزان شباهت را از روی آن اندازه‌گیری کرد.

بافت تصویر: بافت یا الگو عبارت است از روش قرارگیری پیکسل‌هایی با ارزش مشخص در کنار هم به‌صورت یک سلول که تکرار این سلول باعث به‌وجود آمدن بافت مشخصی برای برخی اجسام مانند جاده یا پوست یا هر تصویر دارای بافت مشخص می‌گردد. با شناسایی بافت مربوط به جسم مورد نظر می‌توان قابی برای شناسایی جسم طراحی کرد و در هر فریم آن را روی تصویر حرکت داده و تصویر را فیلتر کرد.

۴. شبکه عصبی

۱.۴. مدل ریاضی یک نرون با چند ورودی: نرون کوچکترین واحد پردازشگر اطلاعات است، که اساس عملکرد شبکه‌های عصبی را تشکیل می‌دهد. شکل (۱) مدل نرون با بیش از یک ورودی را نشان می‌دهد. اسکالرها p و a به ترتیب ورودی و خروجی می‌باشند. بدین ترتیب خروجی نرون با معادله زیر تعریف می‌شود.



شکل ۱: مدل نرون شبکه عصبی با بیش از یک ورودی [۱۴]

مجموعه سیناپس‌های $w_{1,i}$ عناصر ماتریس وزن w را تشکیل می‌دهند. در این حالت w یک بردار سطری با عناصر $w_{1,j}$ و $j=1, \dots, R$ ، تعداد عناصر بردار ورودی است. هر عنصر از بردار ورودی p در عنصر متناظرش از w ضرب می‌شود. نرون، یک جمله بایاس b دارد که با حاصل ضرب ماتریس وزن w با بردار ورودی p جمع می‌شود. ورودی خالص n ، مطابق رابطه ۱ محاسبه می‌شود:

برای تخمین خروجی y از زوج داده های ورودی خروجی برای آموزش با ناظر y ، این شبکه آموزش از توابع چگالی احتمالی گوسی طبق رابطه ۳ با تعیین پارامتر انحراف معیار σ آن‌ها استفاده می‌نماید.

$$\bar{y} = \frac{\sum_{p=1}^P y_p \exp(-d_p^2/2\sigma^2)}{\sum_{p=1}^P \exp(-d_p^2/2\sigma^2)} \quad 3$$

در این رابطه آموزشی با ناظر P زوج آموزشی شامل P بردار ورودی n بعدی X_{1p} تا X_{np} و P خروجی y_p متناظر با آن داریم. d_p مقدار فاصله بین بردار ورودی جاری p و امین بردار ورودی در مجموعه آموزش است [۱۴، ۱۵]. σ شیب تغییرات تابع گوسی را نشان می‌دهد. این شبکه در کاربردهای با ورودی زیاد، حجم زیادی از حافظه را در برمی‌گیرد؛ لذا در این تحقیق سعی شده با انتخاب ویژگی‌های مهم با روش آدابوست حجم زیادی از حافظه اشغال نشود.

۵. الگوریتم ویولا جونز

یافتن چهره در تصویر یکی از موضوعات مورد مطالعه در علوم کامپیوتر و هدف آن تشخیص حضور چهره در عکسی است که به کامپیوتر داده می‌شود. از دشواری‌های تشخیص برای کامپیوترها، تفاوت عکس‌ها در مقیاس، محل، نقطه دید، نور و انسداد..... می‌باشد.

چارچوب تشخیص چهره ویولا جونز یک چارچوب تشخیص اشیاء در تصویر است که توسط پاول ویولا و مایکل جونز ارائه شده است. این چارچوب با قابلیت تشخیص انواع اشیاء، در ابتدا به منظور تشخیص چهره توسعه داده شده بود و اولین الگوریتم تشخیص چهره بلادرنگ به شمار می‌رود. مقاله آن‌ها مشتمل بر سه مفهوم مهم می‌باشد اولین آن‌ها استفاده از تصویر یکپارچه^۸ است که به کمک آن امکان محاسبه سریع ویژگی‌های مورد نیاز برای تشخیص در زمان ثابت فراهم می‌گردد. ماژول دوم بر استفاده از الگوریتم مبتنی بر تقویت تطبیقی، آدابوست^۹، به منظور انتخاب مجموعه‌ای کوچک و کارا از میان مجموعه بزرگی از ویژگی‌های تصویری تخصیص دارد و نوآوری سوم آن‌ها ترکیب بهینه مجموعه‌ای از الگوریتم‌های طبقه‌بند آماری به فرم یک معماری آبخاری^{۱۰} برای حذف سریع قسمت‌های پس‌زمینه و بی‌اهمیت تصویر می‌باشد تا بدینوسیله الگوریتم‌های پیچیده‌تر طبقه‌بندی تنها روی قسمت‌های مهم‌تر تصویر تمرکز نمایند. در ادامه این مفاهیم اساسی به اجمال بیان می‌شوند [۱۴].

۱.۵. کلاسه‌بندی کننده ویژگی‌های هار

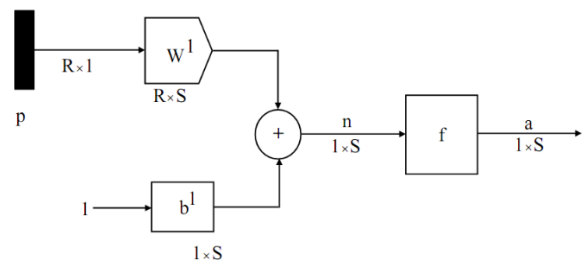
این کلاسه‌بند به زبان ساده یک دسته بند ادراکی دودویی^{۱۱} می‌باشد که برای تصویر ورودی $I(x,y)$ با استفاده از فیلترهای هار $W(x,y)$ و حدآستانه b طبق رابطه ۴ عمل می‌کند.

$$n = \sum_{i=1}^R p_i w_{1,i} + b = W p + b \quad 1$$

در نهایت، با در نظر گرفتن ضرب بردارها خروجی نرون طبق رابطه خواهد بود:

$$a = f(Wp + b) \quad 2$$

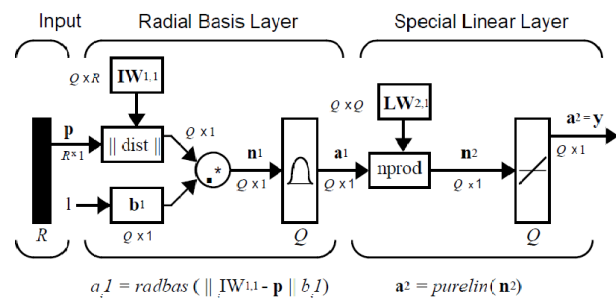
۲.۴. شبکه تک‌لایه: یک شبکه تک‌لایه با نرون عصبی در شکل ۲ نشان داده شده است. باید توجه داشت که هر یک از ورودی‌ها به همه نرون‌های لایه متصل شده است. ماتریس W نیز در این حالت دارای S سطر، به تعداد نرون‌ها و R ستون می‌باشد. همان‌گونه که در شکل مشاهده می‌شود لایه‌ها شامل ماتریس وزن، جمع‌کننده‌ها، بردار بایاس b (دارای عنصر S) تابع تبدیل f هستند [۱۴].



شکل ۲: فرم فشرده، ماتریسی، شبکه تک‌لایه با S نرون

۳.۴. شبکه عصبی رگرسیون عمومی GRNN

شبکه عصبی GRNN در واقع یک روش تخمین توابع غیرخطی است. این شبکه، که در سال ۱۹۹۰ توسط اسپچت^۵ معرفی شد، برای تخمین توابع با زوج داده‌های ورودی و خروجی استفاده می‌شود. که شامل یک لایه توابع پایه شعاعی^۶ و یک لایه خطی ویژه می‌باشد. این شبکه یک تخمین‌زن غیر پارامتری است که برخلاف تخمین‌گرهای معمولی کاهش‌دهنده خطا به تعیین پارامترهای زیاد یا فرض تابع معینی در ابتدا کار ندارد و شیب تغییرات تابع شعاعی آن تنها پارامتری است که در ابتدا تعیین می‌گردد. شکل ۳ ساختار یک شبکه عصبی رگرسیون عمومی را نشان می‌دهد.

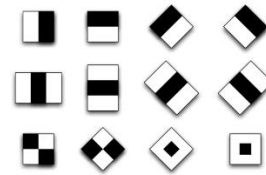


شکل ۳: شبکه عصبی رگرسیون عمومی [۱۵]

این شبکه شبیه شبکه توابع پایه شعاعی و دارای یک لایه خطی ویژه است که متوسط وزندهار خروجی‌های لایه پایه شعاعی را محاسبه می‌کند. این خاصیت خطی بودن آن موجب همگرایی روش آموزش آن می‌شود. این شبکه با توجه به الگوریتم آموزش یک‌بارگذر خود آموزش سریعی داشته و برای کاربردهای بلادرنگ مناسب است.

$$f_{w,b}(I) = \begin{cases} 1, & \text{if } \sum_{x,y} w(x,y)I(x,y) + b > 0 \\ 0, & \text{else} \end{cases} \quad ۴$$

با توجه به ویژگی‌های اجزای چهره، مثلاً قرار گرفتن پل بینی بصورت روشن بین دو ناحیه تیره چشم‌ها یا قرار گرفتن ناحیه تیره چشم‌ها بین نواحی روشن پیشانی و گونه‌ها، قاب‌های W به صورت ویژه با مقادیر دودویی، نواحی سیاه و سفید، به صورت متقارن طراحی می‌شوند. شکل (فیلترها) نمونه‌ای از فیلترها در الگوریتم ویولاجونز را نشان می‌دهد. طبیعت بسیار ساده این فیلترها بار محاسباتی اعمال آن‌ها را روی تصویر بسیار کم می‌کند.



شکل ۴. نمونه‌ای از قاب‌های هار مورد استفاده در الگوریتم [۱۶]

هر مستطیل با قرار گرفتن در بخش‌های مختلف صورت، در طی مراحل متعدد با سایزهای مختلف شروع به محاسبه می‌کند و از کم کردن پیکسل‌های زیربخش‌های سیاه از جمع پیکسل‌های زیرنقاط سفید و مقایسه با استانه b، خروجی کلاسه‌بند به دست می‌آید. الگوریتم ویولا-جونز از یک پنجره 24×24 و قرار دادن این مستطیل‌ها با تعداد و سایزهای مختلف روی صورت استفاده می‌نماید که برای هر پنجره بیش از ۱۵۰ هزار محاسبه نیاز دارد که هزینه و زمان محاسباتی زیادی را صرف می‌کند. به طور متوسط فقط ۱٪ از تمام زیر پنجره‌ها (چهره‌ها) مثبت هستند.

۱.۱.۵. تصویر یکپارچه: جدول محدوده مجتمع^{۱۲}، یک روال محاسبه داده برای تولید سریع مجموعه مساحت‌های سطوح مستطیلی دلخواه از یک جدول یا یک تصویر در حوزه پردازش تصویر و بینایی ماشینی می‌باشد که به آن در حوزه پردازش تصویر، تصویر یکپارچه گفته می‌شود. از این تصویر به ویژه در چارچوب تشخیص ویولاجونز، به منظور ارزیابی سریعتر مشخصه‌ها^{۱۳} استفاده می‌شود. در این تصویر مقدار ذخیره‌شده برای هر پیکسل (x,y) همان‌طور که در رابطه ۵ می‌بینید برابر با جمع مقادیر پیکسل‌های بالا و سمت چپ آن پیکسل در تصویر اصلی به علاوه مقدار خود پیکسل است [۱۷].

$$I(x,y) = i(x,y) + i(x-1,y) + i(x,y-1) + i(x-1,y-1) \quad ۵$$

حسن بزرگ تصویر یکپارچه این است که در آن تعداد عملیات به اندازه‌ی همسایگی وابستگی ندارد؛ در حالی که در روال معمول تعداد عملیات با افزایش اندازه قاب به صورت تصاعدی افزایش می‌یابد.

۲.۵. ماژول تقویت تطبیقی Ada Boost

آدابوست یک الگوریتم یادگیری ماشینی می‌باشد که وظیفه آن پیدا کردن ویژگی‌های کاربردی از میان ویژگی‌های زیاد است، که دسته‌بند

این ویژگی‌ها به عنوان دسته‌بند ضعیف معرفی می‌شوند. آدابوست به عنوان دسته‌بند قوی تعدادی دسته‌بند ضعیف را با هم ترکیب می‌کند. الگوریتم ویولاجونز یک سری از کلاسه‌بندی‌کننده‌های آدابوست را به عنوان یک زنجیره فیلتری با هم ترکیب می‌کند که هر کلاسه‌بند طبق رابطه ذیل، رابطه (۳) از تعدادی دسته‌بند ضعیف تشکیل شده است.

$$F(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) + \dots \quad ۶$$

$\alpha_1, \alpha_2, \alpha_3$ و وزن‌های ویژگی‌ها هستند. با این روش ده‌ها هزار ویژگی به کمتر از ۲۵۰۰ ویژگی کاهش می‌یابند که موجب صرفه‌جویی زیادی در زمان و بار محاسباتی می‌شود.

۳.۵. معماری آبشاری

طبقه‌کننده‌های پیچیده که دقت تشخیص بالایی دارند به اندازه کافی سریع نیستند که به صورت بلادرنگ اجرا شوند، به همین دلیل دسته‌بندی‌کننده‌های ساده‌تر به صورت سلسله مراتبی در یک آبشار به منظور ایجاد دسته‌بندی‌کننده پیچیده منتهی با سرعت تشخیص خیلی بیشتر مرتب شده‌اند، که در آن هر دسته‌بندی‌کننده متوالی تنها توسط نمونه منتخبی که از دسته‌بندی‌کننده‌های قبلی انتخاب شده است آموزش می‌بیند. بنابراین آبشار فرم درختی به خود می‌گیرد. اگر در هر مرحله‌ای در آبشار، دسته‌بندی‌کننده قاب تحت بازرسی، را رد کند، پردازش بیشتری اجرا نشده و جستجوگر به سراغ قاب بعدی می‌رود.

در مورد تشخیص چهره‌ها، اولین دسته‌بندی‌کننده در آبشار، به عنوان عملگر توجه، تنها از دو ویژگی برای کسب یک نرخ ۰٪ برای نرخ رد اشتباه قاب‌های دارای تصویر و نرخ کمتر از ۵۰٪ برای این تشخیص نادرست قاب‌های فاقد چهره استفاده می‌کند. تاثیر این دسته‌بندی‌کننده کاهش تقریباً نیمی از دفعاتی است که کل آبشار ارزیابی می‌شود. در آبشار، هر مرحله شامل یک دسته‌بندی‌کننده با ویژگی‌های مخصوص به خود است. لذا همه ویژگی‌ها در چند مرحله گروه‌بندی می‌شوند که در هر مرحله تعداد مشخصی از ویژگی‌ها مورد بررسی قرار می‌گیرند.

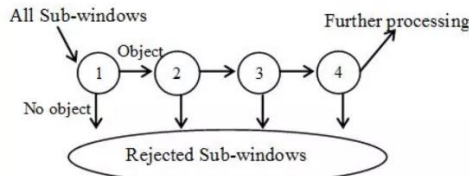
کار دسته‌بند در هر مرحله تعیین این است که آیا یک قاب مشخص قطعاً فاقد چهره است یا ممکن است دارای چهره باشد. این قاب اگر فاقد چهره باشد بلافاصله کنار گذاشته می‌شود. در این ساختار تمام ویژگی‌ها در چند مرحله گروه‌بندی می‌شوند که در آن هر مرحله دارای تعداد خاصی از ویژگی‌ها است. وظیفه هر مرحله این است که تعیین کند: آیا یک قاب معین روی تصویر اصلی، قطعاً فاقد چهره است یا ممکن است دارای صورت فردی باشد. اگر این قاب قطعاً فاقد چهره تشخیص داده شد، کنار گذاشته می‌شود.

ویژگی‌های معماری آبخاری

- به صورت میانگین تنها ۰.۰۱٪ زیر پنجره‌ها دارای چهره هستند بنابراین تعداد زیادی از قاب‌ها و بخش بزرگی از تصویر که فاقد چهره است، کنار گذاشته می‌شود.
- زمان محاسباتی یکسانی برای همه قاب‌ها سپری می‌شود.
- بیشتر زمان باید بر زیر پنجره‌های مثبت سپری شود.
- یک دسته‌بندی‌کننده دو ویژگی می‌تواند نرخ تشخیص ۱۰۰٪ را با نرخ پیدا کردن نادرست 50٪ برای قاب بدون چهره به دست آورد.

• دقت این نوع طبقه‌بندی قابل افزایش بوده و نرخ تشخیص بهتری می‌تواند کسب کند. ارزیابی طبقه‌بندی‌کننده قوی تولید شده توسط فرآیند یادگیری می‌تواند سریعاً انجام شود، اما با افزایش طبقات و دقت تشخیص طبقه‌بندی‌کننده، سرعت تشخیص کاهش می‌یابد و تشخیص از حالت بلادرنگ فاصله می‌گیرد. معماری آبخاری پیامدهای جالبی برای عملکرد دسته‌بندی‌کننده‌ها دارد. فعال‌سازی هر دسته‌بندی‌کننده فقط به رفتار موارد قبلی آن بستگی دارد. لذا نرخ تشخیص نادرست چهره برای قاب‌های فاقد آن رفته‌رفته در طول ایشار کاهش می‌یابد.

با استفاده از این روش می‌توانیم بعد از پردازش هر دسته‌بندی حدسیات خود را درباره وجود یا عدم وجود چهره در آن دسته‌بندی اعلام کنیم. با این روش ما می‌توانیم زودتر بخش‌هایی که اطمینان داریم بخشی از صورت در آن وجود ندارد را حذف کنیم. تنها قایی که تمام از تمام بخش‌بندی‌ها می‌گذرد حاوی چهره خواهد بود. شکل ۶ رفتار این دسته‌بندی‌کننده را نشان می‌دهد.



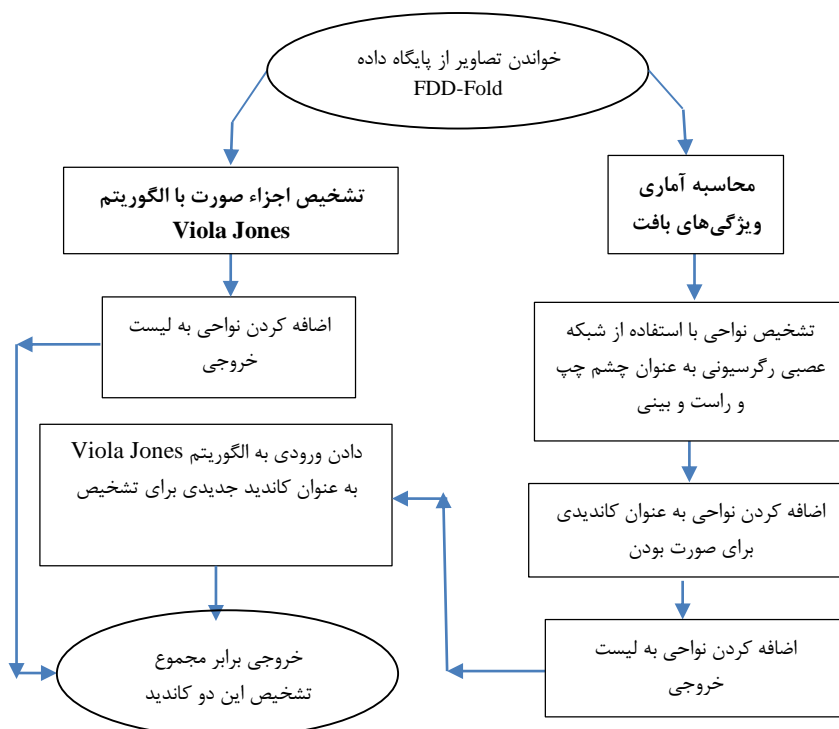
شکل ۵: طبقه‌بند آبخاری

۵. پیاده‌سازی و اجرا

الگوریتم ویولاجونز از سریعترین و دقیقترین الگوریتم‌های تشخیص چهره می‌باشد که مشتمل بر تبدیل تصاویر ورودی به تصاویر یکپارچه، استخراج ویژگی‌های تصویر با فیلترهای هار^۴، انتخاب ویژگی‌های مناسب با ماژول تقویت تطبیقی، آداپوست، و فرآیند توالی آبخاری برای تسهیل و تسریع کار کلاسه‌بندها می‌باشد.

در این مقاله برای تشخیص اجزای صورت هم از الگوریتم Viola Jones و هم ویژگی‌های بافت اجزای صورت استفاده خواهد شد تا دقت روش مورد استفاده بیشتر شود.

شکل ۳ روندنمای مراحل تشخیص چهره با تلفیق شبکه عصبی رگرسیون برای تشخیص بافت چهره و الگوریتم Viola Jones را برای پیدا کردن اجزای چهره در تصویر نشان می‌دهد.



شکل ۶: روندنمای مراحل تشخیص روش پیشنهادی با الگوریتم Viola Jones و شبکه عصبی رگرسیونی

۵، ۱. مجموعه داده‌های FDDB

نرخ تشخیص درست منفی (TN)^{۲۰} یعنی تشخیص چهره بصورت اشتباه برای ناحیه‌ای که چهره نیست. تشخیص نادرست مثبت (FP)^{۲۱} تشخیص عدم وجود چهره در ناحیه‌ای که فاقد چهره است

تشخیص نادرست منفی (FN)^{۲۲} برای تشخیص نادرست مبنی بر وجود چهره برای ناحیه‌ای که فاقد چهره است. از آنجا که قدرت تشخیص چهره الگوریتم برای ما دارای اهمیت است از بین معیارهای مختلف سنجش اعتبار ماتریس درهم‌شدگی، معیار سنجش میزان

روش	داده های آموزشی	صحت تشخیص
شبکه رگرسیونی (GRNN)	۸۰۰	٪۸۹،۲۹
شبکه پس انتشار (BPNN)	۸۰۰	٪۶۳،۵۰
شبکه (RCNN)	۸۰۰	٪۷۹،۳۳

صحت تشخیص، طبق رابطه ۶ است یعنی نسبت کل تعداد چهره، چه این چهره‌ها توسط سیستم تشخیص داده شوند یا نشوند، به کل نواحی مورد بررسی که دارا یا فاقد چهره باشند.

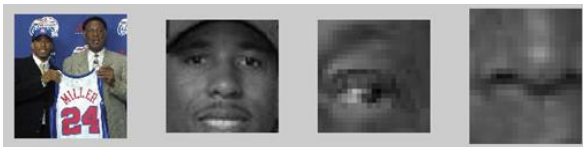
$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad ۶$$

۶. نتیجه‌گیری

در شکل‌های ۹ و ۸ نتیجه تشخیص اجزا چهره با الگوریتم ویولاجونز برای تصاویر حاوی یک چهره تنها و تصویر مرکب از اجزاء مختلف همراه با چهره نشان داده شده است.



شکل ۸: اجزا مهم صورت، دو چشم و بینی، تشخیص داده شده با الگوریتم ویولاجونز برای یک صورت



شکل ۹: نتیجه الگوریتم Viola Jones در تشخیص اجزای چهره برای بهبود دقت الگوریتم تشخیص از ویژگی‌های بافت صورت نیز استفاده می‌شود. ویژگی‌های مورد استفاده شامل حداقل و حداکثر شدت رنگ، میانگین و میانه و واریانس تصویر می‌باشند. داده‌ها برای آموزش به شبکه عصبی رگرسیونی داده می‌شوند. برای تشخیص اجزای صورت شبکه عصبی رگرسیونی با تعدادی از داده‌های مجموعه داده FDD-Fold آموزش داده می‌شود. ٪۷۰ داده‌ها برای آموزش و ٪۳۰ آن‌ها برای آزمون استفاده می‌شوند.

برای ارزیابی دقت روش پیشنهادی نخست مقایسه‌ای بین دقت تشخیص با روش‌های شبکه عصبی رگرسیونی عمومی، شبکه عصبی پس انتشار خطا و شبکه عصبی پیچشی ناحیه محور RCNN [۲۰] برای ۸۰۰

مجموعه داده و محک تشخیص چهره (FDDB) مجموعه‌ای از چهره‌های برجسته خورده برای مسائل تشخیصی اخذ تصاویر چهره در شرایط کنترل نشده است. این مجموعه شامل ۵۱۷۱ چهره برجسته خورده از مجموع ۲۸۴۵ تصویر انتخابی از مجموع داده wild می‌باشد [۱۸]. نمونه‌ای از این تصاویر در شکل ۳ نشان داده شده است.



شکل ۷: نمونه‌ای از تصاویر مجموعه داده FDDB همراه با نواحی چهره از پیش تعیین شده [۱۸]

در روش پیشنهادی ابتدا یک بردار ویژگی مرجع برای شکل هندسی کل سر، مکان قرارگیری اجزای صورت و نیز بافت پوست با الگوریتم ویولاجونز ارائه می‌گردد. در مرحله بعدی با بررسی لبه‌های موجود در تصویر ارائه شده، به دنبال الگوهای قابل تطبیق با شکل کلی سر می‌گردیم و سپس در نواحی‌ای از تصویر که دارای لبه‌هایی مشابه شکل کلی سر انسان می‌باشند، دو پارامتر بافت پوست و نحوه قرارگیری اجزای صورت نیز بررسی می‌شوند تا بتوانیم نسبت به تشخیص صحیح محل حضور چهره انسان در تصویر تصمیم‌گیری کنیم.

برای افزایش سرعت در تشخیص وجود چهره، بافت تصویر در نواحی مختلف با بافت پوست مقایسه می‌شود. برای بهبود دقت تشخیص، لبه‌های تصویر مورد نظر با یک تصویر محتوی لبه‌های تصویر چهره مقایسه می‌شوند.

دقت تفکیک پذیری تصویر نیز باید در حدی باشد که الگوها و بافت‌های چهره قابل استخراج باشند. مطابق تحقیقات انجام شده در مقاله [۱۹] حداقل این مقدار باید ۱۰۲۴ پیکسل بر میلیمتر مربع باشد.

۵، ۲. معیار ارزیابی

برای ارزیابی صحت و میزان درستی روش پیشنهادی از ماتریس درهم‌شدگی^{۲۳} در جهت تشخیص درست و نادرست چهره در تصویر و معیارهای مختلف ارزیابی آن نظیر صحت^{۲۴} استفاده می‌کنیم تا در صد درستی تشخیص را محاسبه کنیم.

اگر حضور چهره در تصویر را به عنوان درست و عدم حضور چهره در محل تشخیص را به عنوان نادرست و تشخیص چهره توسط سیستم تشخیصی را مثبت و عدم تشخیص چهره را منفی لحاظ نماییم پارامترهای رابطه عبارتند از تشخیص درست مثبت (TP)^{۲۵} نرخ تشخیص درست چهره‌ای که در تصویر وجود دارد.

ورودی از مجموعه FDD-Fold داشته‌ایم که نتایج در جدول ۱ درج شده‌اند.

جدول ۱: مقایسه صحت تشخیص شبکه عصبی رگرسیون عمومی، شبکه عصبی پس‌انتشار خطا و شبکه RCNN برای ۸۰۰ داده

نام روش	تعداد داده‌های آموزشی	در صد تشخیص
شبکه پیشنهادی	۳۰۰۰	۹۶٫۳۶٪
شبکه عصبی پس‌انتشار	۳۰۰۰	۷۸٫۳۲٪
شبکه عمیق RCNN	۳۰۰۰	۹۲٫۳۷٪

مشاهده می‌شود که برای داده‌های به نسبت کم شبکه عصبی رگرسیونی صحت تشخیص بیشتری نسبت به دو شبکه دیگر دارد. در مرحله بعدی صحت شبکه پیشنهادی که تلفیق الگوریتم ویولاجونز و شبکه رگرسیونی است را با دو شبکه دیگر در جدول (۱) مقایسه می‌نماییم همچنین برای افزایش دقت تشخیص شبکه یادگیری عمیق تعداد داده‌های آموزش را به ۳۰۰۰ عدد می‌رسانیم که نتایج در جدول (۲) ذکر شده‌اند.

جدول ۲: مقایسه صحت تشخیص شبکه پیشنهادی با شبکه عصبی پس‌انتشار و شبکه عمیق RCNN برای ۳۰۰۰ داده

ملاحظه می‌شود که با افزایش تعداد داده‌های آموزش صحت تشخیص همه روش‌ها، مخصوصاً شبکه یادگیری عمیق افزایش یافته‌است اما تلفیق الگوریتم ویولاجونز با شبکه رگرسیونی موجب نرخ تشخیص و صحت بیشتر نسبت به سایر روش‌ها و شبکه یادگیری عمیق شده‌است. این تحقیق نشان می‌دهد هر چند که شبکه‌های یادگیری عمیق به مدد لایه‌های متعدد قدرت تشخیص بیشتری ارائه‌نموده و ما را از استخراج ویژگی‌های الگوهای مورد توجه بی‌نیاز می‌نمایند اما این قدرت، ارتباط مستقیم با افزایش تعداد داده‌های آموزشی و به تبع آن افزایش قابلیت‌های سیستم پردازشگر دارد که هر دو نیاز به افزایش زمان و هزینه دارند در حالی که با افزودن ویژگی‌ها و تلفیق سیستم‌های شناسایی حتی با تعداد داده‌های متوسط، نه خیلی زیاد، با سیستم‌های پردازش متعارف مرسوم نیز می‌توان به صحت بیشتری دست یافت.

۷. کارهای آتی

ملاحظه می‌شود که با افزایش تعداد داده‌های آموزش صحت تشخیص همه روش‌ها، مخصوصاً شبکه یادگیری عمیق افزایش یافته‌است اما تلفیق الگوریتم ویولاجونز با شبکه رگرسیونی موجب نرخ تشخیص و صحت بیشتر نسبت به سایر روش‌ها و شبکه یادگیری عمیق شده‌است. این تحقیق نشان می‌دهد هر چند که شبکه‌های یادگیری عمیق به مدد لایه‌های متعدد قدرت تشخیص بیشتری ارائه‌نموده و ما را از استخراج ویژگی‌های الگوهای مورد توجه بی‌نیاز می‌نمایند اما این قدرت، ارتباط مستقیم با افزایش تعداد داده‌های آموزشی و به تبع آن افزایش قابلیت‌های سیستم پردازشگر دارد که هر دو نیاز به افزایش زمان و هزینه دارند، در حالی که

با افزودن ویژگی‌ها و تلفیق سیستم‌های شناسایی حتی با تعداد داده‌های متوسط، نه خیلی زیاد، با سیستم‌های پردازش متعارف مرسوم نیز می‌توان به صحت بیشتری دست یافت.

محدودیت الگوریتم ویولاجونز در تشخیص تصاویر فقط با نمای کامل (بدون انسداد)، به‌طور مستقیم (نگاه رو به جلو بدون چرخش سر)، حالت عمودی سر، با نور خوب و تمام تصویر با وضوح معین می‌باشد [۱۶]. در این تحقیق چهره از روبرو و بدون چرخش زاویه قرارگیری سر فرض می‌شود، به دلیل استفاده از تطبیق الگوی ریخت‌شناسی سر و صورت، در الگوریتم پایه‌ای ویولاجونز نمای سر و چهره نباید به شکل ناقص رؤیت شوند، وگرنه تضمینی برای موفقیت روش ارائه‌شده، نخواهد بود. در تحقیقات آتی می‌توان بررسی کرد که با چه مقدار دوران و حول کدام محور، می‌توان محل حضور چهره در تصویر را به درستی تشخیص داد. همچنین با تصاویر دارای شرایط نورپردازی مختلف و نماهای گوناگون دیگر تصویربرداری از چهره میزان دقت روش پیشنهادی را در برابر تغییرات نور و نماهای تصاویر چهره بررسی نمود.

References

- [1] S. J. McKenna, S. Gong, and J. Collins, "Face Tracking and Pose Representation," in *BMVC*, 1996, pp. 1-10.
- [2] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, 2000, pp. 142-149.
- [3] K. Nummiaro, E. Koller-Meier, and L. Van Gool, "An adaptive color-based particle filter," *Image and vision computing*, vol. 21, pp. 99-110, 2003.
- [4] R. Paúl, E. Aguirre, M. García-Silvente, and R. Muñoz-Salinas, "A new fuzzy based algorithm for solving stereo vagueness in detecting and tracking people," *International Journal of Approximate Reasoning*, vol. 53, pp. 693-708, 2012.
- [5] F. J. Aherne, N. A. Thacker, and P. I. Rockett, "The Bhattacharyya metric as an absolute similarity measure for frequency coded data," *Kybernetika*, vol. 34, pp. [363]-368, 1998.
- [6] T. Darrell, G. Gordon, M. Harville, and J. Woodfill, "Integrated person tracking using stereo, color, and pattern detection," *International Journal of Computer Vision*, vol. 37, pp. 175-185, 2000.
- [7] D. Grest and R. Koch, "Realtime multi-camera person tracking for immersive environments," in *IEEE 6th Workshop on Multimedia Signal Processing, 2004.*, 2004, pp. 387-390.
- [8] L. R. Cerna, G. Camara-Chavez, and D. Menotti, "Face detection: Histogram of oriented gradients and bag of feature method," in *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (ICCV)*, 2013, p. 1.
- [9] V. Manian and A. Ross, "A texture-based approach to face detection," in *Biometric Consortium Conference (BCC)*, 2002, pp. 342-347.
- [10] P. Bose and S. Bandyopadhyay, "Human Face and Facial Parts Detection using Template Matching Technique," *International Journal of Engineering and Advanced Technology (IJE)*, 9 (4), 2020.

- 2001 IEEE Computer Society Conference on computer vision and pattern recognition. *CVPR 2001*, 2001, pp. 1-1.
- [18] V. Jain and E. Learned-Miller, "Fddb: A benchmark for face detection in unconstrained settings," UMass Amherst technical report 2010.
- [19] B. J. Boom, G. Beumer, L. J. Spreeuwiers, and R. N. Veldhuis, "The effect of image resolution on the performance of a face recognition system," in *2006 9th international conference on control, automation, robotics and vision*, 2006, pp. 1-6.
- [20] X. Sun, P. Wu, and S. C. Hoi, "Face detection using deep learning: An improved faster RCNN approach," *Neurocomputing*, vol. 299, pp. 42-50, 2018.
- [11] Hasani Ahangar, M. (2019). Triple-Sequential Deep Convolutional Neural Networks in Face Recognition, The 5th National Conference of Electrical and Mechatronics Engineering of Iran, Tehran, Iran. [Persian]
- [12] R. C. Gonzalez and R. E. Woods, "Digital image processing 4th edition, global edition," 2018.
- [13] L. I. Kuncheva, *Combining pattern classifiers: methods and algorithms*: John Wiley & Sons, 2014.
- [14] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. De Jesús, "Neural network design. 2nd edition," *Oklahoma: Martin Hagan*, 2014.
- [15] M. H. Beale, M. T. Hagan, and H. B. Demuth, "Neural network toolbox," *User's Guide, MathWorks*, vol. 2, pp. 77-81, 2010.
- [16] Y.-Q. Wang, "An analysis of the Viola-Jones face detection algorithm," *Image Processing On Line*, vol. 4, pp. 128-148, 2014.
- [17] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the*

پی‌نوشت

- ¹³ Feature
- ¹⁴ Haar Filters
- ¹⁵ Face Detection Dataset and Benchmark(FDDB)
- ¹⁶ Ground truth
- ¹⁷ Confusion Matrix
- ¹⁸ Accuracy
- ¹⁹ True Positive (TP)
- ²⁰ True Negative (TN)
- ²¹ False Positive (FP)
- ²² False Negative (FN)
- ¹ Face Recognition
- ² Bounding boxes
- ³ Global Regression Neural Network(GRNN)
- ⁴ Histogram of the Gradients- HOG
- ⁵ Specht
- ⁶ Radial Basis Function
- ⁷ Supervised learning
- ⁸ Integral Image
- ⁹ AdaBoost
- ¹⁰ Cascading
- ¹¹ Perceptron
- ¹² Summed area table