



دانشگاه صنعتی سهند

DOR:

20.1001.1.23223146.1400.8.2.2.3

نشریه علمی-پژوهشی دانش فنی

دوره ۸ - شماره ۲

پاییز و زمستان ۱۴۰۰

صفحات ۴ الی ۱۸

ISSN: 2322-3146

http://journals.sut.ac.ir/jnsee

مکان یابی نقاط راهنمای چهره با مقاردهی اولیه نقاط راهنما از طریق آموزش ویژگیهای الگوی باینری محلی و هیستوگرام گرادیان جهت‌دار

امین عسگری^۱ و حسین ابراهیم نژاد^۲

^۱ کارشناسی ارشد مهندسی برق- مخابرات، آزمایشگاه تحقیقاتی بینایی کامپیوتر، دانشکده مهندسی برق، دانشگاه صنعتی سهند، تبریز، ایران، a_asgari95@sut.ac.ir

^۲ نویسنده مسئول، استاد گروه برق- مخابرات، آزمایشگاه تحقیقاتی بینایی کامپیوتر، دانشکده مهندسی برق، دانشگاه صنعتی سهند، تبریز، ایران، ebrahimnezhad@sut.ac.ir

(تاریخ دریافت مقاله: ۱۳۹۹/۰۸/۰۶ تاریخ پذیرش مقاله: ۱۴۰۰/۰۸/۲۲)

چکیده

واژه‌های کلیدی

نقاط راهنمای چهره،

تنظیم چهره،

مقاردهی اولیه،

ویژگی تفاوت پیکسل،

رگرسیون حالت آبخاری قوی.

چهره نقش مهمی در برقراری ارتباط بصری ایفا می‌کند. با نگاه به چهره، انسان می‌تواند به طور خودکار بسیاری از پیام‌های غیرکلامی، مانند هویت، قصد و احساسات انسان را استخراج کند. در بینایی کامپیوتر، برای استخراج خودکار اطلاعات چهره، مکان‌یابی نقاط کلیدی چهره معمولاً یک مرحله کلیدی است و بسیاری از روش‌های تحلیل چهره بر روی آشکارسازی دقیق این نقاط برجسته ساخته می‌شوند. مکان‌یابی و تنظیم نقاط راهنمای چهره در تصاویر دارای انسداد، یک کار بسیار مهم و چالش برانگیز در بسیاری از کاربردها می‌باشد. در این تحقیق، یک روش جامع برای مقاردهی اولیه نقاط راهنمای چهره از طریق آموزش ویژگی‌های باینری محلی (LBP^۱) و هیستوگرام گرادیان جهت‌دار (HOG^۱) ارائه می‌شود. همچنین، یک روش آشکارسازی نقاط راهنمای چهره با استفاده از رگرسیون حالت آبخاری قوی نقاط راهنما که به صورت ویژگی‌های تفاوت پیکسل نقاط راهنما مشخص می‌شود، معرفی می‌شود. ابتدا، با استفاده از آنالیز همبستگی الگوی باینری محلی (LBP) و سپس با استفاده از هیستوگرام گرادیان جهت‌دار، ویژگی‌های چهره‌های آموزشی بدست می‌آید. با استفاده از این ویژگی‌های تصاویر آموزشی، نقاط راهنمای بهینه برای تصویر تست تخمین زده می‌شود. در مرحله تست، با توجه به مقاردهی اولیه تصویر، از انتخاب ویژگی مناسب برای تصویر استفاده می‌گردد تا سرعت انجام فرآیند بیشتر شود. یعنی تعداد مراحل با توجه به انتخاب ویژگی بهتر برای هر تصویر کمتر شود. سپس برای تنظیم چهره، از رگرسیون حالت آبخاری قوی استفاده می‌شود و یک اصل محلی برای یادگیری ویژگی‌های نقاط راهنما به کار گرفته می‌شود. اصل محلی کمک می‌کند تا مجموعه‌ای از ویژگی‌های باینری بسیار متمایزکننده برای نقاط راهنمای چهره به صورت مستقل یاد گرفته شود. ویژگی‌های باینری محلی بدست آمده برای یادگیری مشترک رگرسیون حالت آبخاری برای خروجی نهایی مورد استفاده قرار می‌گیرد. نتایج نشان می‌دهد که مقاردهی اولیه مورد استفاده در این کار دقت ارزیابی را در رگرسیون حالت آبخاری بیشتر کرده است و به نتایج بهتری نسبت به مقاردهی اولیه تصادفی دست یافته است.



Sahand University
of Technology

DOR:

[20.1001.1.23223146.1400.8.2.2.3](https://doi.org/10.1001.1.23223146.1400.8.2.2.3)

Journal of Nonlinear
Systems in Electrical
Engineering

Vol.8, No.2

Autumn and Winter 2021

ISSN: 2322 – 3146

<http://journals.sut.ac.ir/jnsee>

Facial Landmark Localization by Initializing the Landmark Points through Training of Local Binary Properties and Histogram of Oriented Gradient

Amin Asghari¹ and Hossein Ebrahimnezhad²

¹MSc, Computer Vision Res. Lab, Electrical Engineering Department, Sahand University of Technology, a_asghari95@sut.ac.ir

²**Corresponding Author**, Professor, Computer Vision Res. Lab, Electrical Engineering Faculty, Sahand University of Technology, ebrahimnezhad@sut.ac.ir

ABSTRACT

Keywords

Facial Landmark,
Face Alignment,
Initialization,
Pixel difference feature,
Robust Cascade Pose
Regression.

Face plays an important role in visual communication. By looking at the face, it can be automatically extracted many non-verbal messages, such as identity, intention, and emotion. In computer vision, localization of the key points of the face is usually a key step for automatic extraction of face information, and many facial analysis techniques are built on the precise recognition of these embossed. Facial landmark detection and alignment in images with occlusion is a very important and challenging task in many visual and image processing tasks. In this paper, a comprehensive method for initialization and alignment of facial landmark through training of local binary features (LBP) and histogram orientated gradient (HOG) and a facial landmark detection method using robust cascade pose regression, which are specified as pixel difference features of landmarks, is introduced. At first, by analyzing the correlation of the local binary pattern histogram (LBP) and then by using histogram orientated gradient, the features of the training images are obtained. For the test image using these features the instructional images are estimated as optimal guide points. In the test stage, according to initialization of the image, the selection of the appropriate feature for the image is used to speed up the process, which means the number of steps to be chosen for each image is better. A strong cascade mode regression is then used to adjust the face, and a local principle is applied to learn the features of the guide points. The local principle helps to learn a set of highly distinctive binary features for the face guide points independently; these local binary features are used to jointly learn the cascade mode regression for the final output. The results show that the initialization used in this work has increased the accuracy of the estimation in the cascade state regression and has obtained better results than the random initialization.

۱- مقدمه

در دنیای واقعی، چهره نقش مهمی در برقراری ارتباط بصری ایفا می‌کند. با نگاه کردن به چهره، انسان به طور خودکار می‌تواند بسیاری از پیام‌های غیر واقعی، مانند هویت، قصد و احساسات انسان‌ها را استخراج کند. پردازش تصاویر مربوط به چهره یک روش بسیار کارآمد برای استخراج نقاط ویژه چهره و به دنبال آن تشخیص این نوع از پیام‌ها به شمار می‌آید. انسان برای جمع‌آوری اطلاعات بصری درباره محیط اطرافش نیاز به سیستم بینایی دارد. معمولاً پردازش تصویر به مجموعه‌ای از تکنیک‌های محاسباتی اشاره می‌کند که تصاویر را به عنوان ورودی می‌پذیرد. نتایج پردازش می‌تواند تصاویر جدید و یا اطلاعات استخراج شده از تصاویر ورودی باشد. آشکارسازی نقاط راهنمای چهره که در کاربردهای نظامی و سایر زمینه‌ها کاربرد دارد، یکی دیگر از زمینه‌های بینایی کامپیوتر می‌باشد [۱، ۲]. در بینایی کامپیوتر، برای استخراج خودکار اطلاعات چهره، مکان یابی نقاط کلیدی چهره بعنوان یک مرحله کلیدی وجود دارد و بسیاری از روش‌های تحلیل چهره بر اساس آشکارسازی دقیق نقاط راهنمای برجسته ساخته شده‌اند. تشخیص حالت چهره و الگوریتم‌های تخمین حالت ممکن است به اطلاعات شکل چهره که توسط موقعیت نقاط راهنما ارائه شده است، بسیار وابسته باشد. چهره در تصاویر مختلف دارای شرایط مختلفی است که از آن جمله می‌توان به نحوه قرارگیری چهره، انسداد، حالت فرد شامل عصبانیت، خوشحالی، غمگینی و شدت روشنایی یعنی وضعیت نور در تصویر اشاره کرد.

برای استخراج ویژگی‌های یک تصویر چندین روش موجود است که هر کدام به صورت‌های گوناگون ویژگی‌های تصویر را استخراج می‌کنند. تعدادی از روش‌های استخراج ویژگی عبارتند از: (۱) تبدیل ویژگی مستقل از مقیاس، (۲) هیستوگرام گرادبان جهت‌دار (HOG)، (۳) الگوی باینری محلی (LBP)، (۴) ویژگی‌های مقاوم سرعت بخشیده شده (SURF). SURF یک الگوریتم تشخیص ویژگی در بینایی کامپیوتر است که برای تشخیص و توصیف ویژگی‌های محلی در تصویر به کار می‌رود و توسط دیوید لائو در سال ۱۹۹۹ [۳] منتشر شده است. برنامه‌های کاربردی تبدیل ویژگی مستقل از مقیاس شامل تشخیص شیء، نگاشت و جهت‌یابی ربایتیک، دوخت تصویر، مدل‌سازی سه بعدی، تشخیص ژست، ردیابی ویدئو، و غیره است. نقاط کلیدی SIFT اشیاء از مجموعه‌ای از تصاویر مرجع استخراج می‌شوند. هیستوگرام گرادبان جهت‌دار که توسط کادوتا و همکاران [۴] در سال ۲۰۰۹ مطرح شده است، یک توصیفگر ویژگی است که در بینایی کامپیوتر و پردازش تصویر برای هدف آشکارسازی شیء مورد استفاده قرار می‌گیرد. این تکنیک، شمارش جهت‌های گرادبان در بخش‌های محلی یک تصویر است. این روش شبیه به هیستوگرام‌های جهت‌گیری لبه، توصیفگرهای تغییر شکل ویژگی مقیاس ثابت و زمینه‌های شکل است، اما تفاوت آن در این است که بر روی یک شبکه متراکم از سلول‌های فاصله‌یکنواخت محاسبه شده است و برای بهبود دقت، از همپوشانی نرمالیزه سازی کنتراست محلی استفاده شده است. الگوی باینری محلی (LBP) که توسط هوانگ و همکاران ایشان [۵] مطرح شده است، یک عملگر ساده و در عین حال بسیار کارآمد است که پیکسل‌های تصویر را با آستانه‌گذاری همسایگی از هر پیکسل برچسب می‌زند و نتایج را به عنوان اعداد باینری نشان می‌دهد. با توجه به قدرت تمایز و سادگی محاسبات، اپراتور بافت LBP در برنامه‌های کاربردی مختلف به رویکرد رایج تبدیل شده است. برای تشخیص نقاط راهنمای چهره روش‌های نوین و متفاوتی وجود دارد که از آن جمله می‌توان به رگرسیون و شبکه‌های عصبی اشاره کرد. تکنیک‌های بسیاری برای انجام تحلیل رگرسیون توسعه داده شده است. روش‌هایی مانند رگرسیون خطی و رگرسیون حداقل مربعات معمولی به صورت پارامتری هستند، به این ترتیب که تابع رگرسیون از

طرف تعداد محدودی از پارامترهای ناشناخته که از داده‌ها تخمین زده می‌شود، تعریف شده است. رگرسیون غیرپارامتری به تکنیکی گفته می‌شود که باعث می‌شود تابع رگرسیون در مجموعه مشخصی از توابع قرار بگیرد که ممکن است نامتناهی باشد.

در ادامه مقاله، در بخش دوم، به مرور مختصری بر کارهای انجام شده در زمینه تشخیص و مکان‌یابی نقاط راهنما می‌پردازیم. در بخش سوم، روش پیشنهادی معرفی می‌شود. در بخش چهارم، نتایج به دست آمده از آزمایش ارائه شده و کارآیی روش پیشنهادی مورد ارزیابی قرار می‌گیرد. در بخش پنجم نیز، نتیجه‌گیری کلی از تحقیق انجام شده ارائه می‌شود.

۲- کارهای مرتبط

در بینایی کامپیوتر، اطلاعات چهره به صورت خودکار استخراج می‌شود. مکان‌یابی نقاط کلیدی چهره، معمولاً یک گام کلیدی است و بسیاری از روش‌های تجزیه و تحلیل چهره برای تشخیص دقیق این نقاط راهنما ساخته شده‌اند. برای مثال، تشخیص حالت چهره [۶] و الگوریتم‌های تخمین حالت سر [۷] ممکن است به شدت بر اطلاعات شکل چهره ارائه شده توسط موقعیت‌های نقاط راهنما تکیه کنند. نقاط راهنمای چهره در اطراف چشم‌ها می‌تواند حدس اولیه از مکان مرکز قرینه چشم را برای تشخیص چشم و ردیابی خیرگی چشم [۸] تهیه کند. برای بازشناسی چهره، مکان نقاط راهنما در تصویر دو بعدی با مدل سه‌بعدی سر با هدف بهبود دقت تشخیص [۹] ترکیب شده‌اند. اطلاعات چهره به دست آمده از طریق مکان‌یابی نقاط راهنمای چهره می‌تواند اطلاعات مهمی برای تعامل انسان و کامپیوتر، سرگرمی، نظارت امنیتی و برنامه‌های پزشکی فراهم کند. روش‌های مکان‌یابی نقاط راهنما، بسته به اینکه چگونه ظاهر شکل و الگوهای شکل چهره را مدل می‌کنند، به سه دسته عمده تقسیم می‌شوند: روش‌های جامع، مدل محلی محدود شده (CLM)^۱، روش‌های مبتنی بر رگرسیون. مدل ظاهر فعال (AAM)^۲ [۱۰] توسط تیلور و کوئتس ادواردز در سال ۱۹۹۸ و کوئتس و همکارانش در سال ۲۰۰۱ معرفی شد. این یک مدل آماری است که در آن تصاویر چهره با تعداد کمی از ضرایب، کنترل هر دو ظاهر چهره و تغییرات شکل را مناسب می‌کند. روش‌های محلی محدود (CLM) [۱۱، ۱۲] مکان نقاط راهنما را بر اساس الگوهای شکل چهره سراسری و اطلاعات ظاهر محلی مستقل در اطراف هر نقطه راهنما استخراج می‌کنند که بدست آوردن آن آسان‌تر و در برابر روشنایی و انسداد در مقایسه با ظاهر کلی قوی‌تر است.

روش‌های مبتنی بر رگرسیون، به طور مستقیم نگاشت از ظاهر تصویر را برای مکان‌های نقاط راهنما یاد می‌گیرد. متفاوت از روش‌های جامع و روش‌های مدل محلی محدود، آن‌ها معمولاً به طور صریح هیچ مدل شکل چهره سراسری را ایجاد نمی‌کنند. در عوض، محدودیت‌های شکل چهره ممکن است به صورت ضمنی تعبیه شوند. به طور کلی، روش‌های مبتنی بر رگرسیون را می‌توان به روش‌های رگرسیون مستقیم، روش‌های رگرسیون آبخاری و روش‌های رگرسیون مبتنی بر یادگیری عمیق طبقه بندی نمود. در کار دانتونه و همکاران [۱۳]، از جنگل‌های رگرسیون مشروط برای یادگیری نگاشت از وصله پنجره‌های نمونه برداری شده تصادفی در ناحیه چهره برای به روز رسانی‌های شکل چهره، استفاده شده است. رویکردهای سراسری پیشرو که توسط سان و همکاران [۱۴] و ژانگ و همکاران [۱۵] ارائه شده است، همگی از روش‌های یادگیری عمیق برای یادگیری نگاشت استفاده می‌کنند. در کار کاظمی و سالیوان [۱۶]، مجموعه‌ای از درختان رگرسیون به عنوان مدل‌های رگرسیون برای تنظیم چهره مورد استفاده قرار گرفته‌اند.

^۱ Constrained Local Model (CLM)

^۲ Active Appearance Model

با اصلاح تابع هدف، الگوریتم می‌تواند از تصاویر آموزشی با موقعیت‌های نقاط راهنمای چهره تا حدودی برچسب‌دار استفاده کند. در میان روش‌های مختلف رگرسیون آبشاری، روش کاهش تحت نظارت (SDM^1) از ژيانگ و دلانور عملکرد قابل ملاحظه‌ای بدست آورده است. در کار اولیه که توسط وو و همکاران ایشان ارائه شده است، مدل ما شین بولتزمن عمیق که یک مدل عمیق احتمالاتی است، برای گرفتن تغییرات شکل چهره به دلیل حالات و قیافه‌ها برای تشخیص و ردیابی نقاط راهنمای چهره مورد استفاده قرار گرفته است. اخیراً، مدل‌های شبکه‌های عصبی کانولوشنال (CNN^2) به مدل‌های یادگیری عمیق غالب برای تشخیص نقاط راهنمای چهره تبدیل شده است و بیشتر آن‌ها از چارچوب رگرسیون مستقیم سراسری یا چارچوب رگرسیون آبشاری استفاده می‌کنند. این روش‌ها را می‌توان به روش‌های یادگیری خالص و روش‌های ترکیبی طبقه‌بندی کرد. یک الگوریتم سریع و دقیق برای محاسبه حالت دوبعدی اشیاء در تصاویر به نام رگرسیون حالت آبشاری (CPR^3) ارائه شده است. در این کار که توسط پیوتر دلار و همکاران [۱۷] صورت گرفته است، CPR به تدریج یک حدس اولیه مشخص را به طور آزادانه تصحیح می‌کند که هر تصحیح توسط رگرسیون متفاوت انجام می‌شود. هر رگرسیون اندازه‌گیری تصویر ساده را انجام می‌دهد که وابسته به خروجی‌های رگرسیون قبلی است. کل سیستم به طور خودکار از نمونه‌های آموزشی یادداشت شده انسان یاد می‌گیرد. همه روش‌های معرفی شده، از مقاردهی اولیه تصادفی برای مکان یابی نقاط راهنمای چهره استفاده می‌کنند که به طور کلی برای همه تصاویر مناسب نیست و ممکن است در روند تشخیص نقاط راهنمای چهره دچار اشتباه شده و منجر به مکان یابی نادرست نقاط راهنما گردد.

۳- روش پیشنهادی

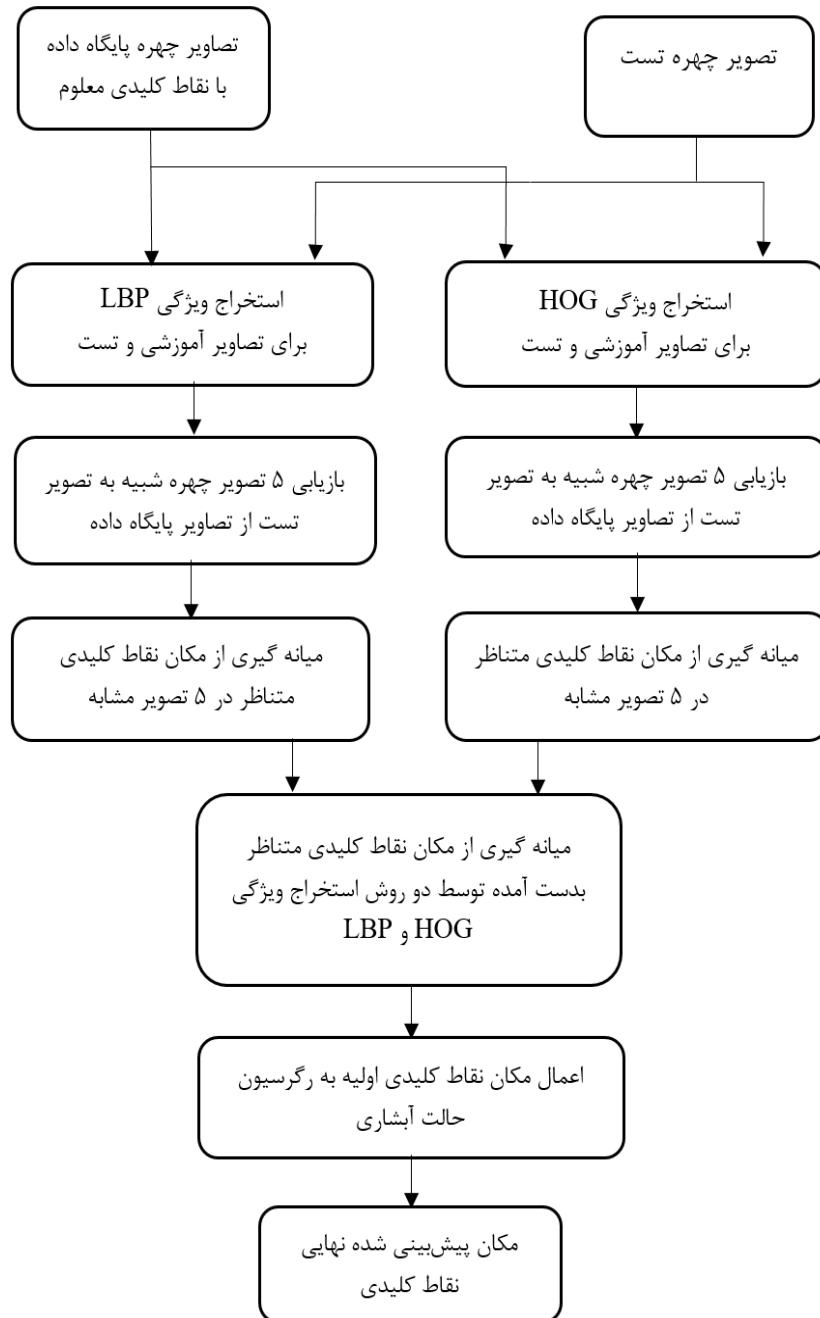
برای پیدا کردن مکان دقیق نقاط کلیدی چهره، لازم است که از یک مکان اولیه (شکل اولیه) برای آنها استفاده شود. روش‌های موجود برای این امر، عمدتاً از یک روش تصادفی استفاده می‌کنند که در بعضی شرایط نوری و حالات چهره ممکن است منجر به نتایج نادرست شود. از این رو، یک انتخاب اولیه خوب و نزدیک به واقعیت کمک موثری در همگرایی نهایی مکان نقاط خواهد بود. با الهام از این موضوع، در این مقاله، بدنبال ارائه ایده‌ای برای انتخاب بهتر مکان اولیه نقاط کلیدی خواهیم بود. ایده اصلی روش پیشنهادی بر این اساس استوار خواهد بود که نقاط کلیدی اولیه، از نقاط کلیدی تصاویر چهره موجود در پایگاه داده انتخاب شود. برای نیل به این هدف، لازم است تصویر چهره تست را آنالیز نموده و با استخراج ویژگی‌های مفید از آن، به بازیابی تصاویر چهره موجود در پایگاه داده پردازیم. تصاویر بازیابی شده، از نظر شرایط نوری، حالت چهره، زاویه چهره و ... به تصویر تست نزدیک بوده و نقاط کلیدی آنها یک انتخاب مناسب و نزدیک به واقعیت خواهد بود. با این انتخاب، از مقاردهی اولیه قوی برای رگرسیون حالت آبشاری با شکل‌های اولیه همبسته حالت و بافت برای تصویر تست استفاده شده است. با بررسی همبستگی الگوی باینری محلی و هیستوگرام گردایان جهت دار بین چهره تست و چهره‌های آموزشی، چهره‌های آموزشی که بیشترین همبستگی را با چهره تست دارند، به عنوان مقاردهی اولیه همبسته بافت انتخاب می‌شوند تا از انتخاب تصادفی مقاردهی اولیه شکل‌های چهره که برای رسیدن به همگرایی از رگرسیون با مراحل زیاد استفاده می‌کنند، جلوگیری شود. برای ساختن مقاردهی اولیه قوی‌تر برای حالت‌های مختلف، از حالت ناهنجار چهره تست مطابق با پنج تصویر که دارای نقاط راهنمای مختلف هستند با

¹ Supervised Descent Method

² Convolutional Neural Network

³ Cascaded Pose Regression

توجه به شبکه‌های کانولوشن آبخاری، استفاده شده است. سپس شکل‌های اولیه همبسته حالت با شکل چهره‌های میانگین و حالت چهره تست ناهنجار، ساخته می‌شود. در نهایت، شکل‌های اولیه همبسته حالت و بافت با یکدیگر ترکیب شده و به عنوان مقاردهی اولیه قوی در کار انجام شده مورد استفاده قرار می‌گیرد. روش به کار رفته باعث بهبود در مقاردهی اولیه شکل‌های چهره‌های تست می‌گردد و به کاهش تعداد مراحل رگرسیون حالت آبخاری و به افزایش سرعت آشکار سازی نقاط راهنمای چهره منجر می‌شود. بلوک دیاگرام روش پیشنهادی در شکل ۱ نشان داده شده است.



شکل ۱. بلوک دیاگرام روش پیشنهادی برای استخراج نقاط کلیدی چهره

الگوریتم ۱ مراحل انجام روش پیشنهادی را برای استخراج نقاط کلیدی چهره نشان می‌دهد.

الگوریتم ۱: الگوریتم مقداردهی اولیه بر اساس همبستگی بافت

ورودی: تصویر تست، تصاویر آموزشی، نقاط کلیدی چهره‌های آموزشی

خروجی: نقاط کلیدی اولیه برای تصویر تست

- (۱) محاسبه هیستوگرام الگوی باینری محلی برای تمام تصاویر
- (۲) بدست آوردن ماتریس هیستوگرام برای تصویر تست و تصاویر آموزشی
- (۳) محاسبه ضریب و فاصله همبستگی تصویر تست و تصاویر آموزشی
- (۴) انتخاب ۵ تصویر با همبستگی زیاد برای تصویر تست
- (۵) استفاده از میانه نقاط کلیدی ۵ تصویر همبسته با تصویر تست به عنوان مقداردهی اولیه

در روش انجام گرفته، از ویژگی‌های الگوی باینری محلی و همچنین هیستوگرام گرادیان جهت‌دار برای بدست آوردن تصاویر شبیه به تصویر تست مورد نظر استفاده شده است. از آنجا که انسداد و تغییرات حالت، ظاهر چهره را تغییر می‌دهد و توصیفگر بافت جزئیات ظاهر محلی را می‌گیرد، می‌توان یک تصویر اولیه همبسته بافت را انتخاب کرد تا اطلاعات انسداد از چهره تست را به جای شکل اولیه تصادفی در نظر بگیرد.

۳-۱- ویژگی الگوی باینری محلی (LBP)

یک روش ابتکاری بر اساس آنالیز همبستگی بافت بین چهره تست و چهره‌های آموزشی پیشنهاد می‌شود و نقاط کلیدی چهره‌های آموزشی که بیشترین همبستگی را با چهره تست دارند، به عنوان مقداردهی اولیه به جای استفاده از یک نمونه تصادفی انتخاب می‌شوند. ابتدا از اپراتور LBP برای آنالیز بافت استفاده می‌شود تا ویژگی‌های تصاویر بدست آید که تصاویر را با آستانه زنی همسایگی 3×3 هر پیکسل با مقدار مرکز، برچسب گذاری می‌کند. از هیستوگرام برچسب‌ها می‌توان به عنوان توصیف کننده بافت استفاده کرد. به این صورت که تصویر به قسمت‌های یکسان تقسیم شده و برای هر قسمت و تمامی پیکسل‌های موجود با توجه به اختلاف پیکسل‌های همسایه با پیکسل مرکزی یک ماتریس ایجاد می‌گردد. اگر عدد پیکسل همسایه بزرگتر از عدد مرکز باشد عدد ۱ و اگر کوچکتر باشد عدد صفر جای گذاری می‌گردد. سپس به صورت پاد ساعت گرد اعداد هر ماتریس که به صورت باینری می‌باشد، خوانده می‌شود. در ادامه عدد دهدهی مربوط به عدد باینری بدست می‌آید. برای همه پیکسل‌ها این کار انجام می‌گردد تا در نهایت همه اعداد دهدهی که از m ناحیه بدست آمده‌اند، یک ماتریس هیستوگرام $m \times n$ را برای هر تصویر بدست آورند. از آنجا که کار انجام شده با هدف انتخاب شکل‌های اولیه مناسب برای رگرسیون حالت آبخاری انجام شده است و باید چند نمونه از مناسب‌ترین شکل‌ها با چهره تست از چهره‌های آموزشی انتخاب شود، به روشی نیاز هست تا ارتباط همبستگی بین چهره تست و چهره‌های آموزشی را ارزیابی کند. در روش ارائه شده، ضریب همبستگی پیرسون برای اندازه‌گیری همبستگی بین چهره تست و چهره‌های آموزشی انتخاب شده است. ضریب همبستگی پیرسون بین ماتریس هیستوگرام چهره تست استخراج شده از روش‌های استخراج ویژگی TM و هر ماتریس هیستوگرام چهره TRM^γ به صورت رابطه (۱) محاسبه می‌شود.

$$\rho_\gamma = \frac{cov(TM, TRM^\gamma)}{\rho(TM), \rho(TRM^\gamma)} = \frac{E[(TM - E(TM))(TRM^\gamma - E(TRM^\gamma))]}{\rho(TM), \rho(TRM^\gamma)} \quad (1)$$

$$\gamma = 1, \dots, \Gamma$$

که در آن $cov(.,.)$ کوواریانس، $\sigma(.,.)$ انحراف استاندارد و Γ تعداد کل چهره‌های آموزشی می‌باشد. با توجه به هر ماتریس

هیستوگرام که به اندازه $m \times n$ است، ρ_γ از طریق رابطه (۲) محاسبه می گردد.

$$\rho_\gamma = \frac{\sum_{i=1}^m \sum_{j=1}^n (TM_{ij} - \overline{TM})(TRM_{ij}^\gamma - \overline{TRM}^\gamma)}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n (TM_{ij} - \overline{TM})^2} \sqrt{\sum_{i=1}^m \sum_{j=1}^n (TRM_{ij}^\gamma - \overline{TRM}^\gamma)^2}} \quad (2)$$

که \overline{TRM}^γ و \overline{TM} به ترتیب مقدار میانگین از ماتریس های TRM^γ و TM هستند. سپس می توان از ضریب همبستگی ρ_γ برای محاسبه فاصله همبستگی d_γ مطابق رابطه (۳) استفاده کرد:

$$d_\gamma = 1 - \rho_\gamma, \gamma = 1, \dots, \Gamma \quad (3)$$

d_γ کوچکتر نشان می دهد که چهره آموزشی با چهره تست همبستگی بیشتری دارد. ۵ چهره که بیشترین همبستگی را بین Γ چهره آموزشی دارند، انتخاب می شوند و شکل های آن ها به عنوان شکل های اولیه برای مقاردهی اولیه چهره تست انتخاب می گردد.

۲-۳- هیستوگرام گرادیان جهت دار (HOG)

همانطور که قبلا گفته شد، از یک روش استخراج ویژگی چهره دیگر به نام هیستوگرام گرادیان های جهت دار (HOG) برای بدست آوردن شکل اولیه مناسب با تصویر تست استفاده شده است. استدلال اساسی در توصیف گر هیستوگرام گرادیان های جهت دار این است که شکل و ظاهر شیء محلی در یک تصویر می تواند با توزیع گرادیان های شدت یا جهت های لبه توصیف گردد. برای بدست آوردن ماتریس هیستوگرام به این صورت عمل می شود که تصویر تست و تصاویر آموزشی به نواحی بهم متصل کوچک به نام سلول تقسیم می گردند و برای پیکسل های موجود در هر سلول، هیستوگرام جهت های گرادیان محاسبه می شود. در ادامه با استفاده از ویژگی های بدست آمده توسط هیستوگرام گرادیان های جهت دار برای تصویر تست و تصاویر آموزشی، پنج تصویر که با توجه به ویژگی ها شباهت زیادی با تصویر تست دارند، برای مقاردهی اولیه انتخاب می شوند؛ به این صورت که از شکل های این پنج تصویر میانه گرفته می شود و شکل بدست آمده به عنوان مقاردهی اولیه برای تشخیص و تنظیم نقاط راهنمای چهره مورد استفاده قرار می گیرد.

۳-۳- آموزش نقاط راهنما با روش رگرسیون آبخاری قوی (RCPR)

با توجه به مقاردهی اولیه بهتر برای نقاط راهنما که با استفاده از استخراج ویژگی تصاویر و نیز همبستگی ماتریس ویژگی بین تصویر تست و تصاویر آموزشی انجام شده است، نقاط راهنما با رگرسیون حالت آبخاری آموزش داده می شوند تا به شکل نهایی برسند و سپس در مرحله تست با توجه به آموزش صورت گرفته، تنظیم نقاط راهنمای تصویر تست انجام می گردد. برای آموزش نقاط راهنمای اولیه با استفاده از رگرسیون حالت آبخاری (CPR) به صورت الگوریتم ۲ عمل می شود.

الگوریتم ۲. الگوریتم رگرسیون حالت آبخاری

ورودی: تصویر I ، مقاردهی های اولیه S_0 ، رگرسورهای $R^{1 \dots T}$ ، ویژگی های شاخص شکل $h^{1 \dots T}$.

```

For  $t=1:T$  do
     $x^t = h^t(S^{t-1}, I)$ 
     $\delta S = R^t(x^t)$ 
     $S^t = S^{t-1} + \delta S$ 
End
    
```


در این روش پیشنهاد شده است که انسداد به طور مستقیم در حین یادگیری به منظور بهبود تخمین شکل ترکیب گردد. این روش به نشانه گذاری داده های مرجع برای انسداد در مجموعه آموزش نیاز دارد. این اطلاعات می تواند با هزینه جزئی در طی روش نشانه گذاری اضافه گردد و یک پرچم به هر نقطه راهنما که از لحاظ دیداری خود را رمزگذاری می کند اضافه گردد. این کار در شکل (۲) نمایش داده شده است.

1	4	7
32%	24%	30%
2	5	8
15%	23%	17%
3	6	9
13%	23%	14%

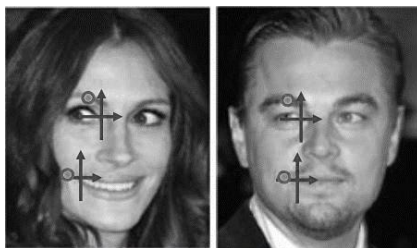
(ب)



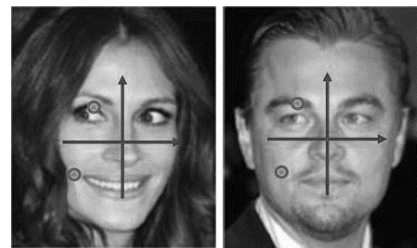
(الف)

شکل ۲. آشکارسازی نقاط راهنما با استفاده از رگرسیون حالت آبخاری قوی. (الف) نمونه‌ای از آشکارسازی نقاط راهنما، (ب) ناحیه بندی تصویر چهره برای تشخیص ناحیه‌هایی که دارای انسداد می‌باشند [۱۸].

شکل S به طور سنتی به عنوان یک سری از مکان های P بخش $S_p = [x_p, y_p]$, $p \in 1 \dots P$ نشان داده می‌شود. همچنین شامل بخشی از دید $S_p = [x_p, y_p, v_p]$ است، که $v_p \in \{0,1\}$ می‌باشد. پیشنهاد شده است که این بعد سوم همزمان با موقعیت‌های بخش یاد گرفته شود. این اطلاعات می تواند برای دستیابی بهتر به انسدادها در هنگام تخمین شکل مفید باشد و رگرسیون به خوبی عمل کند. با توجه به انتخاب ویژگی‌های محلی به جای ویژگی‌های سراسری در تصویر چهره همانطور که در شکل (۳) نشان داده شده است، چنین شاخص سازی در برابر تغییرات بزرگ ثابت نیست و الگوریتم را قوی می‌کند و باعث بهبود عملکرد سیستم شده و در نتیجه به ویژگی‌های مفیدتر برای آموزش دست پیدا می‌کند. بعلاوه، این امکان به وجود می‌آید تا از ویژگی‌های کاندیدای مفیدتری که در اطراف نقاط راهنمای برجسته توزیع شده است، نمونه برداری شود. به عنوان مثال، یک ویژگی تفاوت پیکسل خوب می‌تواند به درستی تشخیص دهد که مرکز چشم تاریک تر از نوک بینی است و یا دو مرکز چشم مشابه هستند که به استخراج ویژگی‌های مفیدتر کمک بسیاری می‌کند.



(ب)



(الف)

شکل ۳. ویژگی‌های شاخص شکل برای نقاط راهنمای چهره. (الف) پیکسل‌های شاخص شکل که به صورت سراسری استخراج شده‌اند، (ب) پیکسل‌های شاخص شکل استخراج شده به صورت محلی [۱۹].

همانطور که در شکل (۳) مشاهده می‌شود انتخاب ویژگی نقاط راهنما به صورت سراسری از کل تصویر چهره برای بدست آوردن ویژگی‌های نقاط راهنما استفاده می‌کند که ممکن است باعث بدست آوردن ویژگی‌های غیر مفید شده و در نتیجه باعث افزایش خطا در رگرسیون گردد. با استخراج ویژگی به صورت محلی، می‌توان ویژگی‌های نقاط راهنما را دقیق‌تر استخراج کرد و

ویژگی‌های دارای اهمیت زیاد را بدست آورد که در رگرسیون برای دستیابی به جواب نهایی بهتر، کمک بسیاری خواهد کرد. در این مطالعه، از استخراج ویژگی محلی و استخراج ویژگی سراسری استفاده شده است به این صورت که برای هر کدام از نقاط راهنما یک هم‌سایگی در نظر گرفته می‌شود و اختلاف پیکسل بین نقطه راهنمای مذکور و هم‌سایگی‌ها بدست آمده و با توجه به این تفاوت، ویژگی‌هایی از نقطه راهنما استخراج می‌گردد که ابتدا برای آموزش رگرسیون مورد استفاده قرار می‌گیرد و سپس برای تخمین شکل نهایی تصویر تست بهره برداری می‌شود.

۴- آزمایش‌ها

در این بخش به بررسی نتایج حاصل از مقداردهی اولیه نقاط راهنما و همچنین نتیجه نهایی بدست آمده از رگرسیون حالت آبشاری قوی پرداخته خواهد شد و نقاط قوت و ضعف کار انجام شده مورد بررسی قرار خواهد گرفت.

۴-۱- مجموعه داده‌ها

مجموعه داده‌های یادداشت شده در بینایی کامپیوتر بسیار مهم است. بنابراین، تعدادی از مجموعه داده‌ها حاوی چهره‌هایی با موقعیت‌های مختلف چهره، حالت‌های چهره، روشنایی و انسداد در گذشته جمع‌آوری شده‌اند. با این حال، اکثر این تصاویر شامل تصاویر در شرایط نامشخص نیستند. از این رو، به تازگی تعدادی از مجموعه داده‌های حاوی چهره در شرایط غیرقابل پیش‌بینی، "حالت آزاد یا طبیعی" جمع‌آوری شده است. مشهورترین مجموعه داده‌های چهره طبیعی عبارتند از: HELEN، LFPW، 300-W، AFW، AFLW و COFW. در ادامه یک مرور کلی از مجموعه داده‌های فوق ارائه می‌شود و در مورد تغییرات مختلف و نقاط راهنمای موجود در آن‌ها اطلاعاتی داده می‌شود.

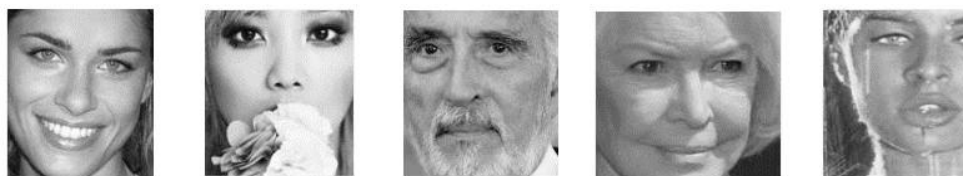
LFPW (۲۹ نقطه راهنما): مجموعه داده LFPW از ۱۱۳۲ تصویر برای آموزش و ۳۰۰ تصویر برای تست ساخته شده است. مجموعه داده LFPW آدرس‌هایی را که به تصاویر پیوند دارند، فراهم می‌کند و برخی آدرس‌ها خراب می‌باشند. فقط می‌توان ۷۷۸ تصویر آموزشی و ۲۱۶ تصویر تست را استفاده کرد.

300-W (۶۸ نقطه راهنما، ۵۱ نقطه راهنما): 300-W به دلیل تغییرات زیاد در زاویه، حالت، روشنایی، پس‌زمینه، انسداد و کیفیت تصویر بسیار چالش برانگیز است. این مجموعه داده از مجموعه داده‌های محبوب موجود، از جمله HELEN، AFW، LFPW و مجموعه داده جدید IBUG به وجود آمده است.

COFW (۲۹ نقطه راهنما): مجموعه داده‌های چهره COFW برای ارائه چهره در شرایط دنیای واقعی طراحی شده است. از چهره‌هایی که با توجه به اختلاف در وضعیت ظاهری، حالت، استفاده از لوازم جانبی مانند عینک آفتابی و کلاه و تعامل با اشیاء (به عنوان مثال غذا، دست، میکروفن و غیره) تغییرات زیادی در شکل و انسداد داشته باشند، استفاده شده است. در نتیجه ۱۰۰۷ تصویر از چهره‌های بدست آمده از منابع مختلف است. همه تصاویر در آزمایشگاه با استفاده از ۲۹ نقطه راهنما مانند LFPW، برچسب‌گذاری شده‌اند. هر دو موقعیت نقاط راهنما و همچنین حالت انسداد / غیرمنسدد تصاویر مشخص شده است، COFW به طور متوسط بیش از ۲۳٪ انسداد دارد. برای افزایش تعداد تصاویر آموزشی، و از آنجایی که COFW دارای نقاط راهنما مانند LFPW است، برای آموزش از ۸۴۵ چهره LFPW و همچنین ۵۰۰ چهره COFW که در مجموع ۱۳۴۵ تصویر را شامل می‌شود، استفاده شده است و برای آزمایش ۵۰۷ چهره باقی مانده COFW مورد استفاده قرار گرفته است.

۴-۲- نتایج مقداردهی اولیه

برای مقداردهی اولیه با استفاده از استخراج ویژگی LBP و HOG، از تمام ۱۳۴۵ تصویر آموزشی و ۵۰۷ تصویر تست استخراج ویژگی انجام گرفته است. برای هر تصویر آموزشی و تست، ویژگی‌های بدست آمده از روش LBP به صورت یک ماتریس $m \times n$ می‌باشد که بسته به نوع تصویر ویژگی‌های متفاوتی بدست می‌آید. برای روش HOG نیز یک بردار ویژگی $1 \times n$ بدست می‌آید. در ادامه، برای تصویر تست مورد نظر، همبستگی ماتریس ویژگی بدست آمده با روش LBP و HOG را با تمامی ۱۳۴۵ تصویر آموزشی بدست آورده و به ترتیب ۵ تصویری که بیشترین همبستگی را با تصویر تست دارند، مشخص می‌کنیم. در شکل‌های (۴) و (۵) نمونه‌ای از تصاویر همبسته با تصویر تست که مورد استفاده قرار گرفته است، به ترتیب برای استخراج ویژگی HOG و LBP نشان داده شده است. به این ترتیب که ۵ تصویر همبسته تر انتخاب می‌شوند (شکل‌های مرجع آن‌ها بر روی تصویر تست نشان داده شده است). در ادامه با میانه‌گیری از این شکل‌ها، شکل نهایی که برای رگرسیون حالت آنبشاری قوی مورد استفاده قرار می‌گیرد، بدست می‌آید.



(الف)



(ب)

شکل ۴. الف) نتایج حاصل از تصاویر همبسته با تصویر تست از طریق استخراج ویژگی HOG که از چپ به راست مرتب شده است، ب) اعمال نقاط راهنمای تصاویر همبسته روی تصویر تست



(الف)



(ب)

شکل ۵. الف) نتایج حاصل از تصاویر همبسته با تصویر تست از طریق استخراج ویژگی LBP با تصویر تست که از چپ به راست مرتب شده است، ب) اعمال نقاط راهنمای تصاویر همبسته روی تصویر تست.

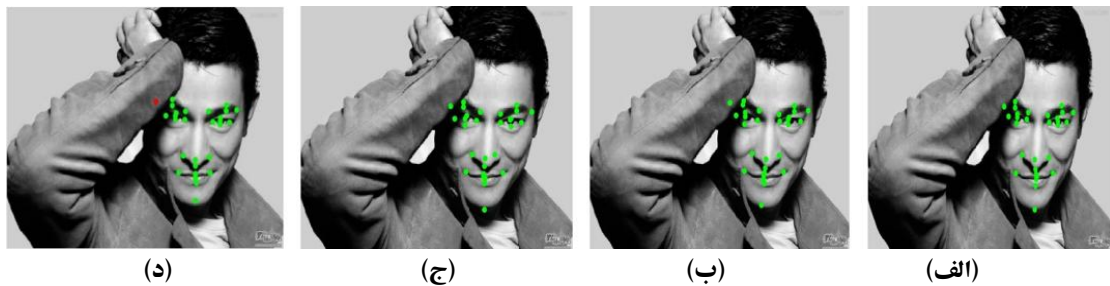
در ادامه از شکل‌های تصاویر همبسته با تصویر تست برای هر دو روش استخراج ویژگی HOG و LBP، میانه گرفته می‌شود تا یک شکل نهایی برای هر کدام بدست آید. سپس از شکل‌های بدست آمده از هر کدام از روش‌ها دوباره میانه گرفته می‌شود تا یک شکل نهایی همانطور که در شکل (۶) نشان داده شده است، برای تصویر تست بدست آید.



شکل ۶. نتیجه نهایی برای مقداردهی اولیه. (الف) شکل نهایی بدست آمده از روش استخراج ویژگی HOG، (ب) شکل نهایی بدست آمده از روش استخراج ویژگی LBP، (ج) شکل حاصل از میانه دو روش استخراج ویژگی

۴-۳- نتایج حاصل از رگرسیون حالت آبخاری قوی

پس از بدست آوردن شکل اولیه یا مقداردهی اولیه برای یک تصویر تست نمونه، با استفاده از روش ترکیب LBP و HOG، از رگرسیون حالت آبخاری قوی برای بدست آوردن تشخیص نهایی نقاط راهنما استفاده شده است. رگرسیون حالت آبخاری برای ۱۳۴۵ تصویر آموزشی، آموزش داده می‌شود و برای تصویر تست مورد استفاده قرار می‌گیرد که به نتیجه نهایی تشخیص نقاط راهنما منجر می‌شود. همانطور که در شکل (۷) نشان داده شده است، رگرسیون آبخاری برای دو روش مطرح شده مورد استفاده قرار می‌گیرد که به نتایج بهتری نسبت به مقداردهی اولیه تصادفی دست پیدا می‌کند.



شکل ۷. شکل نهایی بدست آمده از روش رگرسیون حالت آبخاری با استفاده از (الف) استخراج ویژگی HOG، (ب) استخراج ویژگی LBP، (ج) ترکیب LBP-HOG، (د) شکل مرجع برای مقایسه با شکل‌های پیش‌بینی شده

۴-۴- اعتبارسنجی نتایج

در مطالعه صورت گرفته، به جای استفاده از شکل اولیه تصادفی برای رگرسیون حالت آبخاری قوی، از شکل اولیه استخراج شده با استفاده از ویژگی‌های تصاویر چهره استفاده شده است. با توجه به شکل اولیه استخراج شده از روش ترکیب LBP و HOG، خطای شکل اولیه نسبت به شکل نهایی برای تصاویر تست بدست می‌آید. برای این کار ابتدا فاصله اقلیدسی تمام نقاط راهنمای شکل اولیه نسبت به شکل مرجع محاسبه می‌شود. این کار برای تمامی تصاویر تست انجام می‌شود. سپس، از خطای تک تک نقاط راهنما برای هر تصویر میانگین گرفته می‌شود تا خطای کل برای هر تصویر بدست آید. در ادامه از خطای بدست آمده

برای همه تصاویر میانگین گیری می شود تا خطای نهایی برای مقداردهی اولیه با استفاده از روش های انجام شده بدست آید. جدول (۱) خطای بدست آمده از روش های HOG، LBP و ترکیب LBP-HOG را نشان می دهد. این خطای بدست آمده برای روش های استفاده شده در این تحقیق با روش جامع رگرسیون حالت آبخاری قوی که از مقداردهی اولیه تصادفی استفاده می کند، مقایسه شده است. مقداردهی اولیه تصادفی به این صورت مورد استفاده قرار گرفته است که یک شکل اولیه تصادفی از تصاویر آموزشی انتخاب می شود و برای شکل اولیه به هر تصویر تست که برای رگرسیون انتخاب شده است، اعمال می گردد.

جدول ۱. خطای بدست آمده برای مقداردهی اولیه تصاویر تست نسبت به شکل مرجع.

نام روش	[۱۸] RCPR	LBP	HOG	LBP+HOG
خطا	۳۶/۶۰	۱۲/۳۹	۱۰/۷۲	۱۰/۱۰

با توجه به جدول (۱) استخراج ویژگی با استفاده از ترکیب LBP-HOG نسبت به روش های دیگر خطای کمتری دارد و به نتایج بهتری برای مقداردهی اولیه یا شکل اولیه دست می یابد. روش استخراج ویژگی HOG نیز دارای خطای کمتری نسبت به سایر روش های پیشنهاد شده می باشد. با توجه به خطای بدست آمده برای مقداردهی اولیه یا شکل اولیه نسبت به شکل مرجع، می توان نشان داد که مقداردهی اولیه با استفاده از استخراج ویژگی و همبستگی بین تصاویر آموزشی و تصویر تست بسیار بهتر از مقداردهی اولیه تصادفی عمل می کند.

پس از بدست آوردن مقداردهی اولیه یا شکل اولیه برای تصاویر از رگرسیون حالت آبخاری قوی برای تشخیص و تنظیم این نقاط راهنما استفاده شده است. جدول (۲) خطای حاصل از تنظیم نقاط راهنمای چهره با استفاده از شکل اولیه بدست آمده از روش های استخراج ویژگی تصاویر را نشان می دهد و با کار انجام شده در [۱۸] مقایسه شده است. به این صورت که، فاصله اقلیدسی شکل اولیه نسبت به شکل مرجع برای تمام نقاط راهنما و تمام تصاویر تست بدست می آید. سپس برای هر تصویر از تمام خطاهای بدست آمده از نقاط راهنما میانگین گیری می شود. در ادامه، از خطای بدست آمده برای هر کدام از تصاویر، میانگین گرفته می شود که نشان دهنده خطای نهایی برای کل روش انجام شده می باشد. برای بدست آوردن درصد خطا برای روش های استفاده شده در این مطالعه به این صورت عمل می شود که خطای بدست آمده از فاصله اقلیدسی شکل اولیه و شکل مرجع، به فاصله اقلیدسی دو گوشه چشم تقسیم می شود و با ضرب کردن در ۱۰۰، درصد نهایی از خطای بدست آمده از روش های به کار گرفته شده بدست می آید.

جدول ۲. درصد خطای حاصل از نقاط راهنمای پیش بینی شده نسبت به نقاط راهنمای مرجع.

روش	[۲۱] ESR	[۲۰] SDM	[۱۸] RCPR	HOG	LBP	LBP+HOG
خطای میانگین نرمالیزه	۱۱/۲	۷/۷	۸/۵۰	۶/۲۷	۶/۷۲	۵/۸۷

نتایج بدست آمده نشان می دهد که روش استخراج ویژگی LBP-HOG بهترین عملکرد را برای تنظیم نقاط راهنمای چهره دارد. این درحالی است که RCPR نسبت به این روش دارای خطای بیشتری می باشد و تقریباً دارای خطای برابر با روش استخراج ویژگی HOG می باشد. جدول (۳) درصد خطای بدست آمده برای روش های به کار رفته را نشان می دهد.

جدول ۳. درصد خطای بدست آمده روش های استخراج ویژگی پس از رگرسیون حالت آبخاری قوی برای تنظیم نقاط راهنما

روش	[۲۱] ESR	[۲۰] SDM	[۱۸] RCPR	HOG	LBP	LBP+HOG
خطای میانگین نرمالیزه	۵/۷۰	۵/۸۲	۵/۶۲	۵/۸۴	۶/۲۷	۵/۳۷

روش‌های به کار رفته در این پژوهش با چندین روش پیشرفته در مجموعه داده COFW با استفاده از NME (خطای میانگین نرمالیزه) با استفاده از رابطه (۴) مقایسه شده است.

$$NME = \frac{1}{N} \sum_1^N \frac{\frac{1}{M} \sum_j^M |S_{i,j}^p - S_{i,j}^g|_2}{|gl_i - gr_i|_2} \quad (4)$$

که N تعداد کل تصاویر در مجموعه تست است. M تعداد نقاط راهنما در یک تصویر می‌باشد. $S_{i,j}^p$ نشان‌دهنده موقعیت پیش‌بینی شده از نقطه راهنمای i ام برای تصویر i ام می‌باشد. $S_{i,j}^g$ موقعیت مرجع از نقطه راهنمای i ام برای تصویر i ام را نشان می‌دهد و gl_i و gr_i موقعیت‌های مرجع از گوشه‌های چشم راست و چپ را نشان می‌دهند.

۵- نتیجه‌گیری

روش‌های انجام شده برای مقداردهی اولیه تصویر تست باعث شده است تا شکل اولیه که به رگرسیون اعمال می‌شود به شکل نهایی تصویر چهره بسیار نزدیک باشد و این کار باعث بهبود عملکرد رگرسیون حالت آبخاری قوی برای تنظیم نقاط راهنمای چهره می‌شود. در زمینه تشخیص نقاط راهنمای چهره با استفاده از رگرسیون از جدیدترین روش‌های موجود بهره گرفته شده است و نسبت به این روش‌ها، روش پیشنهاد شده دارای دقت بسیار بالاتری می‌باشد و نتایج بدست آمده نشان از برتری روش پیشنهادی نسبت به روش‌های موجود در زمینه مکان‌یابی نقاط راهنمای چهره با استفاده از رگرسیون دارد.

همه روش‌ها با استفاده از نرم‌افزار Matlab2018 پیاده‌سازی شده‌اند و بر روی رایانه ای با CPU ۳,۶۰ گیگاهرتز و سیستم عامل ۶۴ بیتی ویندوز ۱۰ اجرا شده‌اند. نتایج حاصل از آزمایشات نشان می‌دهد که روش استخراج ویژگی LBP-HOG که ترکیبی از دو روش LBP و HOG می‌باشد برای پیش‌بینی نقاط راهنمای چهره بهتر از روش‌های پیشرفته در این زمینه عمل می‌کند و نتایج بسیار خوبی را برای تنظیم نقاط چهره بدست می‌آورد. روش به کار رفته با استفاده از استخراج ویژگی HOG برای مقداردهی اولیه شکل چهره به خوبی عمل می‌کند و شکل اولیه مناسب‌تری را برای رگرسیون فراهم می‌آورد. با توجه به این که رگرسیون حالت آبخاری قوی (RCPR) و روش‌های رگرسیونی دیگر از مقداردهی اولیه یا شکل اولیه تصادفی برای نقاط راهنما استفاده می‌کنند، انتخاب شکل اولیه مناسب و نزدیک به شکل نهایی هزینه محاسبات را به شدت پایین می‌آورد و باعث بهبود عملکرد برای رسیدن به هدف نهایی می‌شود. خطای میانگین نرمالیزه برای روش‌های به کار رفته و روش رگرسیون حالت آبخاری قوی نشان می‌دهد که استفاده از شکل اولیه با کاهش خطا افزایش سرعت آموزش و تست و بهبود عملکرد می‌گردد.

مراجع

- [1] G. A. Baxes, *Digital image processing: principles and applications*. Wiley New York, 1994.
- [2] R. Klette, *Concise computer vision*. Springer, 2014.
- [3] Y. Wu and Q. Ji, "Facial landmark detection: A literature survey," *International Journal of Computer Vision*, vol. 127, no. 2, pp. 115-142, 2019.
- [4] E. Hjelms and B. K. Low, "Face detection: A survey," *Computer vision and image understanding*, vol. 83, no. 3, pp. 236-274, 2001.
- [5] N. Wang, X. Gao, D. Tao, H. Yang, and X. Li, "Facial feature point detection: A comprehensive survey," *Neurocomputing*, vol. 275, pp. 50-65, 2018.
- [6] M. Pantic and L. J. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 12, pp. 1424-1445, 2000.
- [7] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation in computer vision: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 4, pp. 607-626, 2008.

- [8] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 3, pp. 478-500, 2009.
- [9] Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1701-1708.
- [10] I. Matthews and S. Baker, "Active appearance models revisited," *International journal of computer vision*, vol. 60, no. 2, pp. 135-164, 2004.
- [11] D. Cristinacce and T. F. Cootes, "Feature detection and tracking with constrained local models," in *Bmvc*, 2006, vol. 1, no. 2, p. 3: Citeseer.
- [12] J. M. Saragih, S. Lucey, and J. F. Cohn, "Deformable model fitting by regularized landmark mean-shift," *International Journal of Computer Vision*, vol. 91, no. 2, pp. 200-215, 2011.
- [13] M. Dantone, J. Gall, G. Fanelli, and L. Van Gool, "Real-time facial feature detection using conditional regression forests," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2578-2585: IEEE.
- [14] Y. Sun, X. Wang, and X. Tang, "Deep convolutional network cascade for facial point detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 3476-3483.
- [15] Z. Zhang, P. Luo, C. C. Loy, and X. Tang, "Facial landmark detection by deep multi-task learning," in *European conference on computer vision*, 2014, pp. 94-108: Springer.
- [16] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1867-1874.
- [17] P. Dollár, P. Welinder, and P. Perona, "Cascaded pose regression," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 1078-1085: IEEE.
- [18] X. P. Burgos-Artizzu, P. Perona, and P. Dollár, "Robust face landmark estimation under occlusion," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1513-1520.
- [19] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast keypoint recognition using random ferns," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 3, pp. 448-461, 2009.
- [20] X. Xiong and F. De la Torre Frade. Supervised descent method and its applications to face alignment. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, May 2013
- [21] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):681-685, June 2001.