

# شناسایی حالت آنی با الگوی دودویی محلی یکنواخت در فضای چند رنگی و هیستوگرام جهت‌گیری گرادیان روی شش صفحه متقاطع

ویدا اسماعیلی<sup>۱</sup> محمود محصل فقهی<sup>۲</sup> سید امید شهیدی<sup>۳</sup>  
۱- دانشجوی دکتری- دانشکده مهندسی برق و کامپیوتر- دانشگاه تبریز- تبریز- ایران  
[v.esmaeili@tabrizu.ac.ir](mailto:v.esmaeili@tabrizu.ac.ir)  
۲- دانشیار- دانشکده مهندسی برق و کامپیوتر- دانشگاه تبریز- تبریز- ایران  
[mohasselfeghi@tabrizu.ac.ir](mailto:mohasselfeghi@tabrizu.ac.ir)  
۳- استادیار- گروه مهندسی برق - واحد قزوین- دانشگاه آزاد اسلامی- قزوین- ایران  
[shahdi@qiau.ac.ir](mailto:shahdi@qiau.ac.ir)

**چکیده:** احساسات درونی افراد مانند خوشحالی، ترس، ناراحتی، اضطراب و ... پس از کنترل و سرکوب شدن نیز به مدت کوتاه در چهره پدیدار می‌شوند. این حالت چهره‌ی کنترل‌ناپذیر، حالت آنی است و می‌تواند حقه و نیرنگ را نشان دهد. بنابراین، شناسایی آن در مراکز روان‌درمانی، قضایی و انتظامی بسیار کاربردی است. با این حال، لازم است حرکات سریع و ضعیف ماهیچه‌های صورت به-دقت شناسایی شوند. به همین منظور، ما در این مقاله علاوه بر بزرگنمایی حرکت، شش صفحه در امتداد بُعد زمان در نظر گرفتیم تا تغییرات را به‌خوبی شناسایی کنیم. همچنین، روش‌های الگوی دودویی محلی یکنواخت در فضای رنگی از شش صفحه متقاطع و هیستوگرام جهت‌گیری گرادیان روی شش صفحه متقاطع را پیشنهاد کردیم که نتایج مطلوبی در شناسایی حالت آنی تولید کردند. از آنجائیکه این روش‌ها در شرایط تغییرات نور و روشنایی بهتر عمل می‌کنند، می‌توانند در محیط کنترل نشده و واقعی مورد استفاده قرار گیرند. طبق نتایج آزمایش‌ها، بهترین دقت شناسایی حالت آنی با روش پیشنهادی روی سه پایگاه داده CASME I، CASME II و SMIC-NIR به ترتیب ۷۴/۱۲، ۸۶/۱۹ و ۶۶/۲۳ درصد است که نسبت به روش‌های مقالات پیشین بهبود یافته است.

**واژه‌های کلیدی:** شناسایی حالت آنی، الگوی دودویی محلی یکنواخت در فضای رنگی، هیستوگرام جهت‌گیری گرادیان.

نوع مقاله: پژوهشی

DOI: 10.52547/jiaeee.19.3.123

تاریخ ارسال مقاله: ۱۳۹۹/۱۱/۰۳

تاریخ پذیرش مشروط: ۱۴۰۰/۰۲/۱۱

تاریخ پذیرش مقاله: ۱۴۰۰/۰۳/۰۸

نام نویسنده‌ی مسئول: محمود محصل فقهی

نشانی نویسنده‌ی مسئول: ایران - تبریز - بلوار ۲۹ بهمن - دانشگاه تبریز - دانشکده‌ی مهندسی برق و کامپیوتر

## ۱- مقدمه

دارد [۱۰]. با این حال، آن‌ها ویژگی‌ها را در کل ترکیب نمی‌کنند [۱۳]. درحالی‌که، روش‌های کلی مانند DRMF، بدون نیاز به بخش-بندی و تقسیم نواحی، ویژگی‌ها را از کل تصویر استخراج می‌کنند [۱۴].

با وجود روش‌های متعددی که برای بررسی حالت آنی ارائه شده است، هنوز تحقیقات مربوط به آن اندک است. تحقیقات مربوط به شناسایی حالت آنی به دنبال پی بردن به حالتی از چهره هستند که نمایان شده است. حرکات نامحسوس ماهیچه‌های صورت که فقط در برخی مناطق رخ می‌دهند، گاهی موجب تشخیص اشتباه حالت چهره می‌شود. برای حل این مشکل، در برخی مقالات استفاده از روش بزرگنمایی حرکت اوپلر پیشنهاد شده است [۱۵، ۱].

با این حال، حرکات به حدی ریز هستند که در هنگام وجود تغییرات نور و روشنایی در تصویر، تشخیص آن‌ها چالش برانگیز می‌شود. بزرگنمایی حرکت نیز مصنوعات بیشتری را به تصویر اعمال می‌کند. بنابراین، باید روشی ارائه شود که نسبت به تغییرات روشنایی مقاوم باشد. در این صورت، می‌توان در محیط واقعی و خارج از آزمایشگاه کنترل شده، تصاویر را دریافت کرد و حالت آنی را به خوبی تشخیص داد. اعمال این روش در صفحات زمانی نیز موجب بهبود دقت شناسایی خواهد شد.

به همین منظور، در این مقاله، ابتدا حرکات ظریف ماهیچه‌های صورت در مناطق خاص دارای بالاترین فرکانس بزرگنمایی می‌شوند. این مناطق شامل اطلاعات برجسته و مهم حالت چهره می‌باشند. سپس، از دو روش توصیفگر ویژگی پیشنهادی نام‌های الگوی دودویی محلی یکنواخت در فضای چند رنگی روی شش صفحه متقاطع و هیستوگرام جهت‌گیری گرادیان روی شش صفحه متقاطع استفاده می‌شود. در واقع، ویژگی‌ها از شش صفحه زمانی از تصاویر متوالی شامل فریم اوج استخراج می‌شوند. زیرا این صفحات اطلاعات انتقال حرکت را دربر دارند. در نهایت، دسته‌بندی و تشخیص حالت چهره انجام می‌شود.

نوآوری این مقاله عبارت است از پیشنهاد دو روش توصیفگر ویژگی جدید که نسبت به تغییرات روشنایی مقاوم هستند. تاجایی‌که نویسندگان اطلاع دارند الگوی دودویی محلی یکنواخت<sup>۲</sup> در فضای چند رنگی در تحقیقات شناسایی حالت آنی مورد استفاده قرار نگرفته است. این روش و هیستوگرام جهت‌گیری گرادیان<sup>۴</sup> بر روی شش صفحه اعمال می‌گردد که در ادامه معرفی خواهند شد.

در ادامه، در بخش ۲ روش پیشنهادی و پس از آن در بخش ۳ آزمایش‌ها و بحث و در بخش آخر، نتیجه‌گیری کلی از مقاله ارائه خواهد شد.

سعی در پنهان کردن یا تغییر احساسات واقعی در شرایط خاص (مثلاً برای حفظ موقعیت اجتماعی)، اغلب با بروز حالت آنی چهره<sup>۱</sup> همراه است. این حالت غیرارادی می‌تواند به‌عنوان سرخ دروغ یا فریب در گستره‌ی وسیعی از کاربردها از جمله در دادگاه‌ها، مراکز تحقیقاتی روانشناسی، هنگام بازجویی پلیس و... مورد توجه قرار گیرد [۲، ۱]. در حقیقت، حالت آنی یکی از حالت‌های جهانی چهره با مدت زمان کوتاه و شدت کم است [۳]. طول کشیدن متغیر از ۰.۴ تا ۰.۳ ثانیه و انقباض خفیف ماهیچه‌ها، شناسایی آن را برای چشم غیرمسلح و حتی بینایی ماشین دشوار ساخته است [۱].

اکمان [۴] در سال ۱۹۶۹ متوجه شد که تغییر حالتی در چهره برخی بیماران افسرده در مدت کمتر از ۰.۲ ثانیه مشاهده می‌شود که به‌سرعت به حالت قبلی بازمی‌گردد. مثلاً بیمار مضطرب برای فریب درمانگر، خود را شاد جلوه می‌دهد. اما در مدت کوتاهی، حالت اضطراب در چهره وی نمایان می‌شود. بطوریکه، اگر در آن لحظه پلک زده شود یا توجه کافی نشود، شاید نتوان به این حالت پی برد. در واقع، این حالت خارج از کنترل که به‌نام حالت آنی شناخته می‌شود، احساس واقعی و درونی شخص است.

پس از آن در سال ۲۰۰۲، اکمان ابزار آموزش حالت آنی [۵] را برای بهبود شناسایی طراحی کرد که عملکرد آن فقط ۴۷٪ بود [۶]. بنابراین، محققان استفاده از روش‌های شناسایی الگو را پیشنهاد کردند. روش‌های مورد استفاده در مقالات اخیر عبارتند از: روش‌های مبتنی بر هندسه درمقابل روش‌های مبتنی بر ظاهر و روش‌های محلی درمقابل کلی. در روش‌های مبتنی بر هندسه از شکل، اندازه، فاصله و موقعیت اجزای صورت بردار ویژگی استخراج می‌شود [۳]. DRMF<sup>۲</sup> یکی از این روش‌هاست که مدل شکل و حرکت اجزا را از مجموعه آموزشی می‌آموزد [۹، ۸، ۷]. این روش، محاسبات ساده‌ای دارد [۱۰، ۷]. با این حال، برخی نقاط ویژگی هنوز نمی‌توانند بدرستی با استفاده از آن یافت شوند [۱۱].

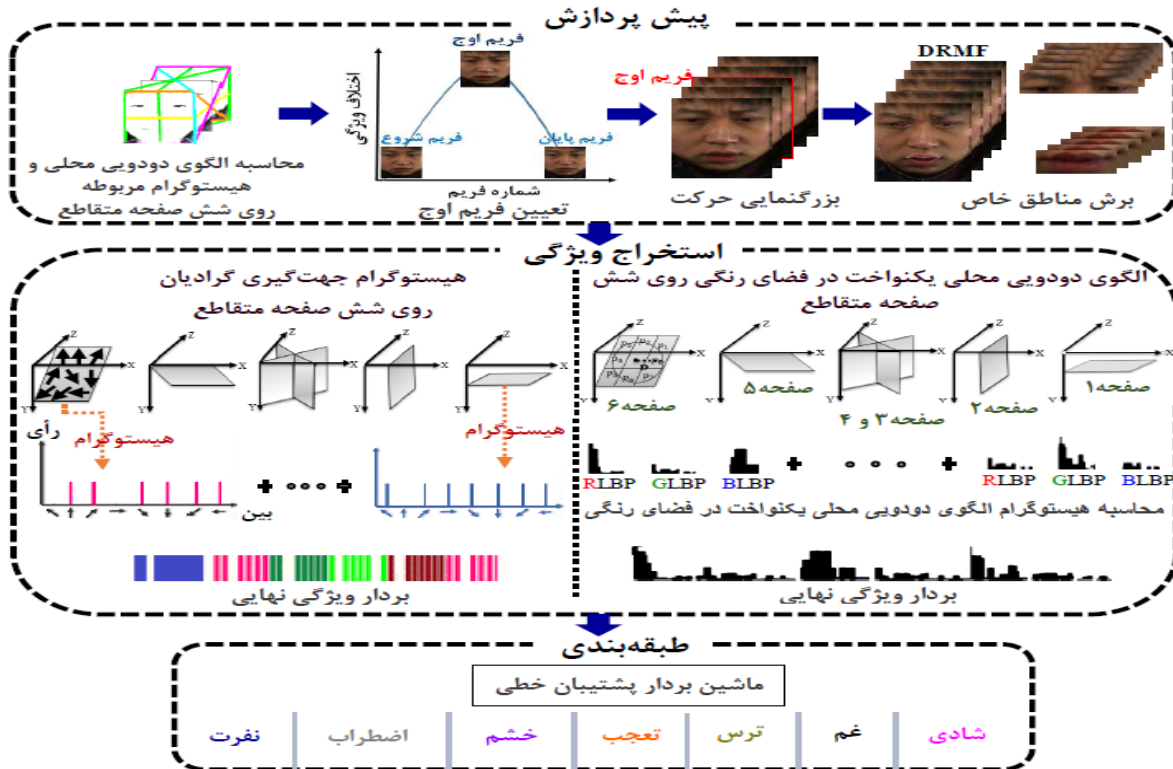
در مقابل، روش‌های مبتنی بر ظاهر، ویژگی‌ها را از بافت پوست مانند چین و چروک‌های روی چانه، پیشانی، اطراف چشم و خطوط اطراف دهان استخراج می‌کنند [۱۲]. روش‌های مبتنی بر الگوی دودویی محلی، نمونه‌ای از این روش‌ها هستند که بر روی کل تصویر چهره یا نواحی خاصی از آن اعمال می‌شوند. این روش‌ها در مقایسه با روش‌های قبلی، عملکرد بهتری در توصیف حرکات ظریف دارند [۱۲، ۳].

روش‌های محلی تنها در مناطق خاصی، ویژگی‌ها را استخراج می‌کنند [۱۳]. انتخاب نواحی خاص دارای بیشترین جزئیات حرکات ماهیچه‌ها مانند دهان، چشم و ابرو، علاوه بر در اختیار قرار دادن اطلاعات مهم، موجب حذف مناطق غیرضروری می‌گردد. در نتیجه، استخراج ویژگی از این نواحی نسبت به کل تصویر نیاز به زمان کمتری

## ۲- روش پیشنهادی برای شناسایی حالت آنی چهره

سپس، روش پیشنهادی در سه مرحله توضیح داده خواهد شد. در مرحله اول، پیش پردازش ارائه می گردد. سپس، روش های استخراج ویژگی پیشنهادی به تفصیل شرح داده می شوند. پس از آن، روش طبقه بندی حالت های چهره بیان خواهد شد. چارچوب کلی روش پیشنهادی در شکل (۱) نشان داده شده است.

در این بخش، روشی برای شناسایی حالت آنی پیشنهاد می گردد. به همین منظور، ابتدا پژوهش های مرتبط پیشین مرور خواهد شد و



شکل (۱): چارچوب کلی روش پیشنهادی برای شناسایی حالت آنی چهره

### ۱-۲- مروری بر پژوهش های مرتبط پیشین

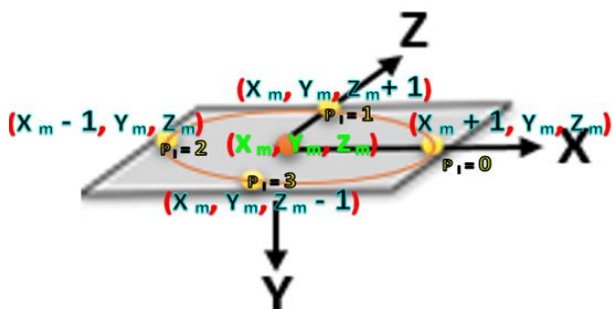
الگوی دودویی محلی روی سه صفحه متعامد<sup>۵</sup> [۱۶]، با مقایسه مقادیر پیکسل های همسایه و پیکسل میانی در هر صفحه، تغییرات بافت پویا را در تصاویر خاکستری متوالی رمزگذاری می کند. یکی از صفحات این روش تمام پیکسل های یک تصویر را شامل می شود. صفحه دوم، پیکسل های متناظر در صفحات عمودی تصاویر متوالی و صفحه سوم، پیکسل های متناظر در صفحات افقی این تصاویر را دارا هستند. پیکسل میانی در این سه صفحه مشترک است (برای مطالعه بیشتر به [۱۶] مراجعه شود).

با این حال، پیکسل هایی در نقاط تقاطع صفحات وجود دارند که دو مرتبه محاسبه می شوند. مثلاً یکبار در صفحه اول و یکبار در صفحه دوم با پیکسل میانی مقایسه می شوند. برای رفع این مشکل، از الگوی دودویی محلی با شش نقطه تقاطع<sup>۶</sup> که فقط یکبار هر پیکسل همسایه را در نظر می گرفت، استفاده شد [۱۷]. این روش سریع تر از الگوی دودویی محلی روی سه صفحه متعامد بود اما عملکرد شناسایی حالت آنی را بهبود چشمگیری نداد.

لی و همکاران [۱] علاوه بر الگوی دودویی محلی، هیستوگرام گرادینان های جهت دار<sup>۷</sup> و هیستوگرام جهت گیری گرادینان تصویر را روی همان سه صفحه محاسبه کردند (برای توضیحات بیشتر به [۱] مراجعه شود). بهترین نتیجه برای شناسایی حالت آنی با استفاده از روش هیستوگرام جهت گیری گرادینان تصویر روی سه صفحه متعامد حاصل شد.

اخیراً، اسماعیلی و همکاران [۱۵] چهار صفحه قطری را معرفی کردند. آن ها با ترکیب روش شار نوری و الگوی دودویی محلی در این صفحات، بردار ویژگی را استخراج کردند و پس از تفکیک و دسته بندی حالت های چهره با استفاده از ماشین بردار پشتیبان خطی، توانستند در ۷۹،۶٪ و ۷۲،۳۸٪ از پایگاه داده های CASME I و CASME II حالت آنی چهره را شناسایی کنند.

با این حال، هنوز هم می توان با ارائه روشی مناسب تر، دقت شناسایی را بهبود بخشید. روشی که در شرایط واقعی که در تصاویر تغییرات نور وجود دارد نیز عملکرد مطلوبی داشته باشد. اغلب روش های استفاده شده در مقالات پیشین مانند روش شار نوری و الگوی دودویی محلی روی سه صفحه متعامد یا شش نقطه تقاطع، نسبت به



شکل (۲): مختصات نقاط همسایه پیکسل میانی در صفحه‌ی یک (مختصات ۴ نقطه‌ی همسایه با رنگ آبی مشخص شده است)

کد الگوی دودویی محلی روی هر شش صفحه محاسبه می‌شود (برای محاسبه کد الگوی دودویی محلی به [۱۶] مراجعه کنید). توجه داشته باشید که این صفحات در پیکسل میانی یکدیگر را قطع کرده‌اند. پس از محاسبه هیستوگرام از شش صفحه، آن‌ها را به هم متصل می‌کنیم تا هیستوگرام کلی بدست آید. در نهایت، با مقایسه هیستوگرام هر فریم با هیستوگرام فریم شروع با استفاده از روش مجموع مربعات تفاضل، فریم اوج مشخص می‌شود. این فریم، فریمی است که بیشترین تفاوت را با فریم نشان‌دهنده حالت خنثی چهره دارد.

فریم اوج و دو فریم قبل و بعد آن وارد مرحله بزرگنمایی حرکت می‌شوند تا شناسایی حالت چهره به درستی انجام گیرد. در این مرحله، ورودی با استفاده از هرم لاپلاسین به زیربخش‌هایی با باند فرکانسی مختلف تبدیل می‌شود. با عبور هر باند فرکانسی از فیلتر پایین‌گذر و ضرب در ضریب تقویت  $\lambda$ ، حرکات ظریف بزرگنمایی می‌شوند. در نهایت، بازسازی و تولید خروجی با اضافه کردن سیگنال اصلی به آن انجام می‌شود (برای مطالعه جزئیات بیشتر به [۱۹] مراجعه شود).

از آنجائیکه چشم و ابرو و دهان مناطق پر اهمیتی در شناسایی حالت آبی هستند [۱۰]، ما این مناطق را جدا می‌کنیم. ابتدا با استفاده از آشکارساز لندمارک<sup>۸</sup> DRMF، ۶۶ نقطه صورت نمایان می‌گردد و با استفاده از مختصات این نقاط مناطق مورد نظر برش داده می‌شوند. البته ما ده پیکسل در چهار جهت اضافه می‌کنیم تا بتوانیم جزئیات بیشتری از حالت چهره را رمزگذاری کنیم.

### ۲-۳- استخراج ویژگی

در این مرحله، ویژگی‌های مطلوب از دو ناحیه‌ی برش داده شده استخراج می‌شوند. ما از دو روش پیشنهادی برای استخراج ویژگی استفاده می‌کنیم.

### ۲-۳-۱- الگوی دودویی محلی یکنواخت در فضای چند رنگی روی شش صفحه متقاطع

پنج تصویر از مناطق مورد نظر که خروجی مرحله قبل هستند به‌عنوان ورودی دریافت می‌شوند. تصاویر ورودی، رنگی هستند و ما از هر سه کانال رنگی R، G و B استفاده می‌کنیم. در حقیقت، الگوی دودویی

تغییرات روشنایی و سایه حساس هستند و می‌بایست بر روی تصاویر استاندارد که در شرایط کنترل‌شده‌ی آزمایشگاه جمع‌آوری شده‌اند، اعمال شوند. همچنین، استفاده از صفحات مناسب که اطلاعات مفیدی را برای استخراج ویژگی دریافت می‌کنند، مهم است. بنابراین در ادامه، روش پیشنهادی را بیان می‌کنیم.

### ۲-۲- پیش‌پردازش

در این مرحله، ابتدا فریم اوج شناسایی می‌شود. این فریم، حالت چهره را بهتر از سایر فریم‌ها نمایش می‌دهد. زیرا فریم اوج، فریمی است که بیشترین چین و چروک ایجاد شده در پوست ناشی از حرکات ماهیچه‌های صورت را نشان می‌دهد. به عبارت دیگر، حرکات ماهیچه‌های صورت در حالت آبی به‌قدری ظریف است که در اغلب فریم‌ها و تصاویر متوالی مشخص نیست و تصاویر تقریباً حالت خنثی چهره را نشان می‌دهند. اما فریم اوج که بیشترین حرکات ماهیچه‌های صورت در آن نمایان است، می‌تواند حالتی از چهره که نمایان شده است (مثلاً حالت شادی یا اخم در چهره و یا ...) را نشان دهد.

فریم‌های متوالی حالت آبی را پشت سر هم بصورت حجم سه بعدی قرار می‌دهیم. این تصاویر هم‌تراز شده‌اند (برای مطالعه بیشتر به [۱۸، ۱] مراجعه شود). راستای افقی را محور x، راستای عمودی را y و راستای امتداد فریم‌ها را z می‌گیریم. z بُعد زمان است. فرض کنید که  $P_l \in 0, \dots, N_{neig} - 1$  همسایه محلی اطراف پیکسل میانی در یک فریم و D، فاصله پیکسل میانی تا نقطه همسایه باشد.  $N_{neig}$  تعداد نقاط همسایگی پیکسل میانی است. نقاط همسایه روی دایره‌ای قرار دارند که مرکز آن پیکسل میانی است.

در این صورت، مختصات هر نقطه همسایه اطراف پیکسل میانی در صفحات یک، دو، سه (صفحه  $45^\circ$ )، چهار (صفحه  $135^\circ$ )، پنج و شش از روابط (۱) تا (۶) قابل محاسبه است. شش صفحه‌ی معرفی شده در شکل (۱) نشان داده شده است. مختصات نقاط همسایه پیکسل میانی در صفحه‌ی یک که طبق رابطه (۱) محاسبه شده است، در شکل (۲) قابل مشاهده است. در شکل (۲)، فرض شده است که ۴ نقطه همسایه وجود دارد ( $N_{neig} = 4$ ) و فاصله پیکسل میانی تا نقطه همسایه ۱ است ( $D=1$ ).

$$(x_m + D_x \cos(\frac{2\pi P_l}{N_{neig}}), y_m, z_m + D_z \sin(\frac{2\pi P_l}{N_{neig}})) \quad (1)$$

$$(x_m, y_m - D_y \sin(\frac{2\pi P_l}{N_{neig}}), z_m + D_z \cos(\frac{2\pi P_l}{N_{neig}})) \quad (2)$$

$$(x_m + D_x \sin(\frac{2\pi P_l}{N_{neig}}), y_m - D_y \sin(\frac{2\pi P_l}{N_{neig}}), z_m - D_z \cos(\frac{2\pi P_l}{N_{neig}})) \quad (3)$$

$$(x_m - D_x \sin(\frac{2\pi P_l}{N_{neig}}), y_m - D_y \sin(\frac{2\pi P_l}{N_{neig}}), z_m + D_z \cos(\frac{2\pi P_l}{N_{neig}})) \quad (4)$$

$$(x_m + D_x \cos(\frac{2\pi P_l}{N_{neig}}), y_m - D_y \sin(\frac{2\pi P_l}{N_{neig}}), z_m - D_z \cos(\frac{2\pi P_l}{N_{neig}})) \quad (5)$$

$$(x_m - D_x \cos(\frac{2\pi P_l}{N_{neig}}), y_m - D_y \sin(\frac{2\pi P_l}{N_{neig}}), z_m - D_z \cos(\frac{2\pi P_l}{N_{neig}})) \quad (6)$$

## ۲-۴- طبقه‌بندی

ماشین بردار پشتیبان روشی مناسب برای طبقه‌بندی است که در مقالات متعدد مورد استفاده قرار گرفته است [۱۵, ۱]. این روش، نتایج نسبتاً مطلوبی را تولید کرده است [۱۵, ۱]. ما از روش ماشین بردار پشتیبان خطی استفاده می‌کنیم.

ماشین بردار پشتیبان خطی دارای کرنل خطی است. این روش یک روش یادگیری با نظارت است. بنابراین، میبایست ورودی و خروجی به آن داده شود. برای انتخاب پارامتر، اعتبار سنجی متقابل بر روی داده‌های آموزش انجام می‌شود.

## ۳- آزمایش‌ها و بحث

در این بخش ابتدا پایگاه داده‌های مورد استفاده که تصاویر حالت آنی از آن انتخاب شده‌اند، بیان می‌شود. سپس، جزئیات اجرای آزمایش‌ها و نتایج حاصل از آن‌ها شرح داده می‌شود. در نهایت، بحث عملکرد روش پیشنهادی و مقایسه آن با روش‌های مرتبط ارائه خواهد شد.

### ۳-۱- پایگاه داده‌های مورد استفاده

ما تصاویر ورودی را از پایگاه داده‌های CASME I [۲۱] و CASME II [۲۲] استفاده می‌کنیم. زیرا آن‌ها دارای برچسب فریم اوج هستند و بنابراین، می‌توانیم درصد تعیین دقیق این فریم را مشخص کنیم.

این پایگاه داده‌ها خودبه‌خودی هستند. به این معنا که سعی شده است حالت‌های آنی در شرایط واقعی اتفاق افتد و ضبط شود. به عبارت دیگر، برای جمع‌آوری این پایگاه داده‌ها از شرکت‌کنندگان خواسته شده است تا با تماشای فیلم‌های دلخراش یا بسیار خنده‌دار و ... سعی کنند احساسات حقیقی خود را کنترل و پنهان کنند و حالت خنثی چهره خود را حفظ کنند.

CASME II و CASME I به‌ترتیب دارای ۱۹۵ و ۲۴۷ نمونه حالت آنی هستند که رزولوشن تصاویر آن‌ها به‌ترتیب ۷۲۰ × ۱۲۸۰ و ۲۴۰ × ۲۸۰ پیکسل است. تصاویر پایگاه داده CASME II با نرخ فریم بالا (۲۰۰ فریم در ثانیه) ثبت شده‌اند. درحالی‌که، نرخ فریم در CASME I شصت فریم در ثانیه است. در آن‌ها تصویر حالت خنثی چهره در ابتدا و انتهای هر نمونه وجود دارد و تصاویر آن‌ها بدون نویز است.

برخلاف پایگاه داده‌های فوق، تصاویر پایگاه داده SMIC-NIR [۱] RGB نیستند زیرا تصاویر آن با دوربین نزدیک به مادون قرمز ضبط و ثبت شده‌اند. این پایگاه داده دارای ۷۱ نمونه حالت آنی است. رزولوشن تصاویر آن ۶۴۰ × ۴۸۰ پیکسل است. ما از این پایگاه داده فقط برای روش هیستوگرام جهت‌گیری گرادیدان روی شش صفحه متقاطع استفاده می‌کنیم.

محلی یکنواخت بصورت جداگانه روی هر کانال رنگی روی هر صفحه از شش صفحه‌ی معرفی شده محاسبه می‌شود. سپس، نتایج حاصل از سه کانال رنگی به هم می‌پیوندند. با الحاق نتایج کانال‌ها، الگوی دودویی محلی یکنواخت در فضای چند رنگی بدست می‌آید. در واقع، در الگوی دودویی محلی چند کاناله برای هر کانال رنگی یک هیستوگرام نتایج جداگانه محاسبه شده و ویژگی‌ها همه به یک میزان تأثیر می‌گذارند.

در نهایت، هیستوگرام‌های محاسبه شده از شش صفحه نیز به هم ملحق می‌شوند. معادلات در زیر آورده شده است (روابط (۷) تا (۹)). در این معادلات، ChLBP الگوی دودویی محلی سه کانال R, G و B است و  $pv$  نشان‌دهنده مقدار پیکسل در کانال قرمز و سبز و آبی است.  $S$  (تابع علامت) هنگامی صفر است که  $v$  منفی باشد، در غیر اینصورت یک است [۲۰].

$$U(ChLBP) = |S(pv_{r_1} - pv_m) - S(pv_0 - pv_m)| + \sum_{r_1}^{N_{neg}-1} |S(pv_{N_{neg}} - pv_m) - S(pv_{r_1} - pv_m)| \quad (7)$$

$$ChLBP = \begin{cases} \sum_{r_1}^{N_{neg}-1} S(pv_{N_{neg}} - pv_m), & \text{if } U(ChLBP) \leq 2 \\ N_{neg} + 1, & \text{otherwise} \end{cases} \quad (8)$$

$$Multi - Color ULBP Hist = \begin{bmatrix} RLBP_1 & GLBP_1 & BLBP_1 \\ RLBP_2 & GLBP_2 & BLBP_2 \\ RLBP_3 & GLBP_3 & BLBP_3 \\ RLBP_4 & GLBP_4 & BLBP_4 \\ RLBP_5 & GLBP_5 & BLBP_5 \\ RLBP_6 & GLBP_6 & BLBP_6 \end{bmatrix} \quad (9)$$

### ۲-۳-۲- هیستوگرام جهت‌گیری گرادیدان روی شش

#### صفحه متقاطع

دومین روش استخراج‌کننده ویژگی پیشنهادی، هیستوگرام جهت‌گیری گرادیدان روی شش صفحه متقاطع است. در این روش، با حرکت ماسک‌های سوپل  $[-1 \ 0 \ 1]^T$  و  $[-1 \ 0 \ 1]$  در راستای افقی و عمودی در هر صفحه از صفحات معرفی شده، گرادیدان محاسبه می‌شود. جهت گرادیدان پیکسل‌ها در هر صفحه با استفاده از رابطه (۱۰) و هیستوگرام جهت‌گیری گرادیدان با استفاده از رابطه (۱۱) قابل محاسبه است.

$$\theta = \arctan(\nabla P), \quad -\pi < \theta < \pi \quad (10)$$

$$Hist = \sum \delta(K(\theta), B) \quad (11)$$

که در آن  $\delta(K(\theta), B)$  دلتای کرونکر است و هنگامی ۱ می‌شود که  $K(\theta) = B$  باشد.  $B$  تعداد بین‌های هیستوگرام است. تابع کمی‌سازی  $\theta$  نگاشت  $\{1, \dots, q\} \rightarrow [-\pi, \pi]$  است که  $q$  سطح کمی‌سازی  $\theta$  را نشان می‌دهد.

قابل ذکر است که از رأی ساده برای هر کانال (بین) هیستوگرام استفاده می‌شود و اندازه گرادیدان در رأی دادن به هر کانال بی‌تأثیر است. در نهایت، از نرمال‌سازی  $L1$  برای هر هیستوگرام استفاده می‌شود و سپس، هیستوگرام‌های شش صفحه به هم متصل می‌شوند.

### ۳-۲- آزمایش‌ها و نتایج

در این جا، جزئیات اجرای روش‌ها و نتایج آزمایش‌های شناسایی حالت آنی را بیان می‌کنیم. قابل ذکر است که شبیه‌سازی‌ها با نرم-افزار MATLAB و در سیستمی با پردازشگر Core i7، 3.5 GHZ و 8 RAM گیگا بایت انجام شده است.

پس از مشخص شدن فریم اوج، پنج تصویر شامل فریم اوج، دو فریم قبل از این فریم و دو فریم بعد از آن وارد مرحله بزرگنمایی حرکات ظریف و برش نواحی چشم‌ها و ابروها و دهان می‌شوند. انتخاب درست ضریب بزرگنمایی می‌تواند عملکرد شناسایی حالت آنی را بهبود بخشد (شکل ۴). ما حرکات را در ویدئوی حالت آنی متشکل از پنج فریم فوق با ضریب تقویت ۲، ۴، ۸، ۱۲، ۱۶ و ۲۰ بزرگنمایی می‌کنیم. سپس، ویدئو را به تصاویر پیاپی تبدیل می‌کنیم و برش مناطق موردنظر را اعمال می‌کنیم.

برای مشخص کردن فریم اوج، تصاویر متوالی خاکستری در هر نمونه یک آرایه سه بعدی در نظر گرفته می‌شوند. ویژگی‌ها از سه فریم متوالی استخراج می‌شود. بدین ترتیب که ابتدا از فریم‌های شماره ۱، ۲ و ۳ و سپس از فریم‌های ۲، ۳ و ۴ و به همین ترتیب تا آخرین فریم هر نمونه، بردار ویژگی استخراج می‌شود. بنابراین، فاصله زمانی را یک قرار دادیم. فاصله پیکسل میانی از پیکسل همسایه را در راستای محورهای x، y و z یک قرار دادیم. تعداد نقاط همسایه پیکسل میانی در هر صفحه می‌تواند ۴، ۸، ۱۶ و یا ۲۴ باشد که ما آن را ۸ انتخاب کردیم و کد الگوی دودویی محلی را در بلوک‌های ۳×۳ پیکسل محاسبه نمودیم. همچنین، برای محاسبه مقادیر پیکسل‌های همسایه که دقیقاً بر روی دایره قرار نداشتند از میانگین نزدیکترین پیکسل‌ها (دو پیکسل مجاور) استفاده کردیم.

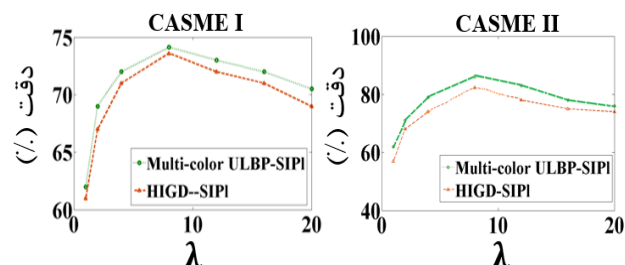
ویژگی‌ها از این مناطق استخراج می‌شوند. تصاویر RGB متوالی ناحیه چشم‌ها و ابروها و همچنین، تصاویر RGB دهان را آرایه سه بعدی در نظر می‌گیریم و شش صفحه معرفی شده را به آن‌ها اعمال می‌کنیم. با استفاده از روش الگوی دودویی محلی یکنواخت در فضای چند رنگی روی شش صفحه متقاطع پیشنهادی ویژگی‌ها را استخراج می‌کنیم. هیستوگرام‌های بدست آمده را بهم متصل می‌کنیم تا بردار ویژگی تولید شود. در این روش، ۸ پیکسل همسایه برای هر پیکسل میانی در نظر گرفتیم ( $N_{neig} = 8$ ) و فاصله پیکسل میانی از هر پیکسل را یک قرار دادیم. زیرا با انتخاب این پارامترها بهترین عملکرد حاصل شد.

در نهایت، با اتصال هیستوگرام‌های الگوی دودویی محلی بلوک‌ها در هر صفحه و پس از آن به هم پیوستن هیستوگرام‌های صفحات، هیستوگرام الگوی دودویی محلی از شش صفحه متقاطع به صورت ماتریسی با ابعاد ۶×۲۵۶ بدست آمد. هر سطر مربوط به یک صفحه‌ی روش پیشنهادی است.

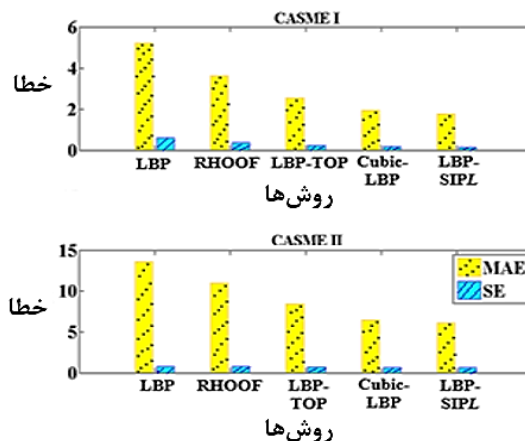
همچنین، ویژگی‌ها را با استفاده از دومین روش پیشنهادی یعنی هیستوگرام جهت‌گیری گرادیان روی شش صفحه متقاطع استخراج کردیم. در این روش، با استفاده از جهت گرادیان پیکسل‌ها در صفحات معرفی شده، هیستوگرام‌ها بدست می‌آیند که با اتصال آن‌ها بردار ویژگی استخراج می‌شود. تعداد بین برای هر هیستوگرام ۸ است ( $B = 8$ ).

سپس، با استفاده از روش مجموع مربعات تفاضل، سطرهای هیستوگرام‌ها با سطرهای هیستوگرام اول مقایسه می‌شود. تا اختلاف ویژگی با فریم اول (حالت خنثی) محاسبه شود. بدین ترتیب، فریم دارای بیشترین تغییرات و حرکات در میان فریم‌های یک نمونه، فریم اوج اعلام می‌شود. خطا در تعیین فریم اوج در مقایسه با روش‌های استفاده شده در مقالات مرتبط در شکل (۳) مشاهده می‌گردد. در این شکل، نمودارهای زرد رنگ نشان‌دهنده میانگین خطای مطلق و نمودارهای آبی نشان‌دهنده خطای استاندارد هستند.

در پایان، بردارهای ویژگی به ماشین بردار پشتیبان خطی داده شدند تا آن‌ها را تفکیک کند و حالت چهره تشخیص داده شود. نتایج شناسایی حالت آنی با استفاده از روش‌های پیشنهادی در زیر آمده است. افزایش دقت شناسایی با روش‌های پیشنهادی در مقایسه با سایر روش‌ها در جدول ۱ و جدول ۲ مشهود است.



شکل (۴): دقت شناسایی حالت آنی با ضرایب بزرگنمایی  $\lambda$  مختلف



شکل (۳): خطا در تعیین فریم اوج

همانطور که در جداول ۱ و ۲ مشاهده شد روش‌های پیشنهادی درصد شناسایی بالاتری دارند. طبق آزمایش‌های انجام شده، این روش‌ها نسبت به تغییرات روشنایی مقاوم هستند. الگوی دودویی محلی یکنواخت در فضای چند رنگی روی شش صفحه متقاطع نسبت به هیستوگرام جهت‌گیری گرادبان روی شش صفحه متقاطع نتایج بهتری دارد اما نیاز به تصاویر رنگی دارد. در شرایطی که تصاویر RGB نباشند (مانند تصاویر پایگاه داده SMIC-NIR)، می‌توان از روش هیستوگرام جهت‌گیری گرادبان روی شش صفحه متقاطع که عملکرد مطلوبی در تغییرات نور دارد استفاده کرد. در واقع، این روش اثر روشنایی و کنتراست را بوسیله‌ی نادیده گرفتن مقدار مشتقات مرتبه اول کاهش می‌دهد. همچنین، جهت‌گیری گرادبان به روشنایی پیکسل بستگی ندارد.

#### ۴- نتیجه‌گیری

حالت آنی که با حرکت غیرارادی ماهیچه‌های صورت پدیدار می‌شود، نیت واقعی افراد را نشان می‌دهد. بنابراین می‌تواند برای فهمیدن دروغ در بازجویی‌ها و مذاکرات، جلوگیری از فریب، تشخیص قابل اطمینان‌تر در روانشناسی بالینی و ... مورد استفاده قرار گیرد. با این وجود، ارائه‌ی روشی برای نمایان کردن بهتر حرکات ظریف آن می‌تواند مفید باشد. به همین منظور، ما در این مقاله روش الگوی دودویی محلی یکنواخت در فضای رنگی از شش صفحه متقاطع و هیستوگرام جهت‌گیری گرادبان روی شش صفحه متقاطع را که می‌تواند تغییرات خفیف را در طول زمان از تعداد فریم اندک استخراج کنند، پیشنهاد کردیم. هر دو روش نسبت به تغییرات روشنایی مقاوم هستند. هیستوگرام جهت‌گیری گرادبان روی شش صفحه متقاطع می‌تواند بر روی تصاویر خاکستری و RGB اعمال شود و نسبت به رنگی بودن تصویر نیز حساس نیست. با توجه به آزمایش‌های انجام شده، روش‌های پیشنهادی عملکرد خوبی در شناسایی حالت آنی در مقایسه با روش‌های مرتبط داشتند. در آینده، این روش‌ها می‌توانند برای تشخیص سایر تغییرات و حرکات کوچک بکار روند.

#### مراجع

- [1] Li, X., et al, "Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods", IEEE transactions on affective computing, Vol. 9, No. 4, pp. 563-577, 2017.
- [2] Esmaeili, V., Shahdi, S.O., "Automatic micro-expression apex spotting using Cubic-LBP", Multimedia Tools and Applications, pp. 1-19, 2020.
- [3] Takalkar, M., Xu, M., Wu, Q., Chaczko, Z., "A survey: facial micro-expression recognition", Multimedia Tools and Applications, Vol. 77, No. 15, pp. 19301-19325, 2018.
- [4] Ekman, P., Friesen, W.V., "Nonverbal leakage and clues to deception", Psychiatry, Vol. 32, No. 1, pp. 88-106, 1969.

جدول (۱): نتایج شناسایی حالت آنی چهره در پایگاه داده SMIC-NIR

روش	دقت (%)
هیستوگرام جهت‌گیری گرادبان	۳۵/۲۱
هیستوگرام جهت‌گیری گرادبان روی سه صفحه متعامد	۳۵/۵۲
الگوی دودویی محلی روی سه صفحه متعامد	۶۴/۷۹
هیستوگرام جهت‌گیری گرادبان روی شش صفحه متقاطع (روش پیشنهادی)	۶۶/۲۳

جدول (۲): نتایج شناسایی حالت آنی چهره در دو پایگاه داده دیگر.

روش	CASME I دقت (%)	CASME II دقت (%)
هیستوگرام جهت‌گیری گرادبان روی سه صفحه متعامد	۶۴/۶۶	۵۵/۸۷
الگوی دودویی محلی از سه سطح متعامد	۶۴/۶۷	۵۵/۸۷
ترکیب شار نوری و الگوی باینری محلی روی صفحات قطری	۷۲/۳۸	۷۹/۶۰
هیستوگرام جهت‌گیری گرادبان روی شش صفحه متقاطع (روش پیشنهادی)	۷۳/۶۰	۸۳/۷۱
الگوی دودویی محلی یکنواخت در فضای چند رنگی روی شش صفحه متقاطع (روش پیشنهادی)	۷۴/۱۲	۸۶/۱۹

#### ۳-۳- بحث و مقایسه

با توجه به اینکه حالت آنی با حرکات مختصر و خفیف همراه است، در توالی فریم‌ها بهتر از یک فریم آشکار می‌گردد. روش‌های مبتنی بر الگوی دودویی محلی روی شش صفحه متقاطع، حداقل سه فریم متوالی را دریافت می‌کنند و تغییرات فریم‌های قبلی و بعدی را با فریم میانی می‌سنجند.

شش صفحه معرفی شده تغییرات را با گذشت زمان در جهات مختلف کدگذاری می‌کنند. این صفحات، اکثر پیکسل‌های دارای اطلاعات مفید و تأثیرگذار را در بر می‌گیرند. کاهش خطای استاندارد و میانگین خطای مطلق این مورد را تأیید می‌کند. همانطور که در شکل (۳) مشاهده شد، تعیین فریم اوج با استفاده از الگوی دودویی محلی دارای خطای بیشتری است. اما وقتی این روش روی چند صفحه اعمال می‌شود، موجب کاهش خطا می‌شود. به طوری که، الگوی دودویی محلی روی شش صفحه متعامد<sup>۱۱</sup> کمترین خطا را در تعیین فریم اوج دارد. همچنین، با انتخاب مناطق خاص، اطلاعات زائد مربوط به ظاهر صورت حذف می‌شوند. بزرگنمایی حرکت نیز تأثیر به‌سزایی در عملکرد شناسایی حالت آنی دارد. اما باید توجه کرد که حرکات با ضریب مناسبی بزرگنمایی و تقویت شوند. زیرا همانطور که در شکل (۴) مشاهده شد، تقویت بیش از حد، دقت را کاهش می‌دهد. دلیل این امر آن است که در این صورت حرکات ناخواسته مانند حرکت سر و ... نیز تقویت می‌شوند.

- [19] Le Ngo, A.C., Oh, Y.H., Phan, R.C.W., See, J., "Eulerian emotion magnification for subtle expression recognition", In 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, pp. 1243-1247, 2016.
- [20] Shahdi, S.O., Abu-Bakar, S.A.R., "Multi-color ulbp with wavelet transform in invariant pose face recognition", In 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), pp. 52-57, 2011.
- [21] Yan, W.J., Wu, Q., Liu, Y.J., Wang, S.J., Fu, X., "CASME database: a dataset of spontaneous micro-expressions collected from neutralized faces", In 2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG), IEEE, pp. 1-7, 2013.
- [22] Yan, W.J., Li, X., Wang, S.J., Zhao, G., Liu, Y.J., Chen, Y.H., Fu, X., "CASME II: An improved spontaneous micro-expression database and the baseline evaluation", PloS one, Vol. 9, No. 1, p. e86041, 2014.
- [5] Ekman, P., "Microexpression training tool (METT)", San Francisco: University of California, 2002.
- [6] Yu, M., Guo, Z., Yu, Y., Wang, Y., Cen, S., "Spatiotemporal Feature Descriptor for Micro-Expression Recognition Using Local Cube Binary Pattern", IEEE Access, Vol. 7, pp. 159214-159225, 2019.
- [7] Merghani, W., Davison, A.K., Yap, M.H., "A review on facial micro-expressions analysis: datasets, features and metrics", Computer Vision and Pattern Recognition, arXiv preprint arXiv:1805.02397, 2018.
- [8] Asthana, A., Zafeiriou, S., Cheng, S., Pantic, M., "Robust discriminative response map fitting with constrained local models", In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3444-3451, 2013.
- [9] Asmara, R.A., Choirina, P., Rahmad, C., Setiawan, A., Rahutomo, F., Yusron, R.D.R., Rosiani, U.D., "Study of DRMF and ASM facial landmark point for micro expression recognition using KLT tracking point feature", In Journal of Physics: Conference Series, Vol. 1402, No. 7, p. 077039, IOP Publishing, 2019.
- [10] Liong, S.T., See, J., Wong, K., Le Ngo, A.C., Oh, Y.H., Phan, R., "Automatic apex frame spotting in micro-expression database", In 2015 3rd IAPR Asian conference on pattern recognition (ACPR), IEEE, pp. 665-669, 2015.
- [11] Oh, Y.H., See, J., Le Ngo, A.C., Phan, R.C.W., Baskaran, V.M., "A survey of automatic facial micro-expression analysis: databases, methods, and challenges", Frontiers in psychology, 9, p. 1128, 2018.
- [12] Guo, C., Liang, J., Zhan, G., Liu, Z., Pietikäinen, M. and Liu, L., "Extended Local Binary Patterns for Efficient and Robust Spontaneous Facial Micro-Expression Recognition", IEEE Access, 7, pp. 174517-174530, 2019.
- [13] Allaert, B., Bilasco, I.M., Djeraba, C., "Micro and macro facial expression recognition using advanced Local Motion Patterns", IEEE Transactions on Affective Computing, 2019.
- [14] Liong, S.T., See, J., Wong, K., Phan, R.C.W., "Less is more: Micro-expression recognition from video using apex frame", Signal Processing: Image Communication, Vol. 62, pp. 82-92, 2018.
- [15] Esmaili, V., Mohassel Fegghi, M., Shahdi, S.O., "Autonomous Apex Detection and Micro-Expression Recognition using Proposed Diagonal Planes", International Journal of Nonlinear Analysis and Applications, Vol. 11, No. Special Issue, pp. 483-497, 2020.
- [16] Zhao, G., Pietikainen, M., "Dynamic texture recognition using local binary patterns with an application to facial expressions", IEEE transactions on pattern analysis and machine intelligence, Vol. 29, No. 6, pp. 915-928, 2007.
- [17] Wang, Y., See, J., Phan, R.C.W., Oh, Y.H., "Lbp with six intersection points: Reducing redundant information in lbp-top for micro-expression recognition", In Asian conference on computer vision, pp. 525-537, Springer, Cham, 2014.

### زیر نویس ها

<sup>1</sup> Facial Micro-Expression

<sup>2</sup> Discriminative Response Map Fitting

<sup>3</sup> Uniform Local Binary Pattern (ULBP)

<sup>4</sup> Histogram of Gradient Direction (HiGD)

<sup>5</sup> LBP on Three Orthogonal Planes (LBP-TOP)

<sup>6</sup> LBP with Six Intersection Points

<sup>7</sup> Histograms of Oriented Gradients (HOG)

<sup>8</sup> Landmark

<sup>9</sup> Bins

<sup>10</sup> Spontaneous

<sup>11</sup> LBP on Six Intersection Planes (LBP-SIPL)

[۱۸] اخلاقی، سینا، حسن‌پور، حمید، ابوالقاسمی، حمید، "مدل‌سازی چهره با استفاده از میانگین‌گیری بر پایه دگر‌دیسی تصویر و تجزیه مرتبه پایین"، مجله مهندسی برق و الکترونیک ایران، سال پانزدهم- شماره دوم، ۶۷-۵۵، ۱۳۹۷.