

تعدیل اثر سبک سوژه با استفاده از تابع انتقال سبک در بازشناسی رفتار خودانگیخته چهره

امین محمدیان^۱، حسن آقائی نیا^۲ و فرزاد توحیدخواه^۳

چکیده

ظاهر بروز واحدهای کنشی و درد در چهره ممکن است به طور قابل توجهی برای افراد مختلف متفاوت باشد. بنابراین توزیع احتمال داده‌های آموزشی و داده فرد جدید در حالتیکه لازم است مدل یادگیرنده مستقل از داده شخص جدید آموزش دیده باشد، متفاوت است. در این کار با فرض اینکه سبک افراد در بروز حالات علت تنوعات بین فردی است، روش‌هایی برای افزایش توانایی تعمیم مدل بازشناسی جلوه‌های کنشی چهره ارائه شده است. این روش‌ها بر اساس مفهوم انتقال سبک یا تعدیل اثر سبک عمل می‌کنند. تفاوت روش پیشنهادی این است که در مقایسه با روش‌هایی که از نگاشت برای انتقال سبک استفاده می‌کنند، الگوریتم‌هایی را پیشنهاد داده است که در حین سادگی کارایی مناسب داشته باشند. آزمایش‌هایی بر روی پایگاه داده UNBC-McMaster انجام شده تا نتایج روش‌های پیشنهادی با کارهای قبل مقایسه شوند. نتایج نشان می‌دهد که این رویکرد می‌تواند در حین داشتن سرعت مناسب، به طور موثر نرخ شناسایی درد و واحدهای کنشی را در چهره بهبود دهد. بطوریکه بهترین متوسط نرخ بازشناسی واحدهای کنشی ۹۶,۸۴ (با معیار AUC) بوده و همان روش در شرایط داده تطبیق کم و داده تطبیق مناسب به ترتیب دارای نرخ بازشناسی درد ۸۷,۳۰ و ۹۳,۲۶ (با معیار AUC) بوده است.

کلیدواژه‌ها

بینایی ماشین، بازشناسی ژست و رفتار چهره، رفتار خودانگیخته چهره، تابع انتقال سبک

این تفاوت افراد یکی از دلایل اختلاف بین توزیع داده آموزش و آزمون در سامانه‌های بازشناسی مستقل از فرد است که به عنوان یک فرض پایه در نظر گرفته می‌شود [۱]. علاوه بر این تعداد نمونه‌های آموزش و تنوع آنها نقش مهمی در یادگیری و قدرت تعمیم‌دهی روش بازشناسی ایفا می‌کند. همچنین در رابطه با توزیع داده ناشی از بروز جلوه‌های بصیری چهره در حالت خودانگیخته^۱ نیز نمی‌توان فرض یکسان بودن توزیع داده آموزش و آزمون را در نظر گرفت [۲]. به‌عنوان مثال در رابطه با جلوه‌های بصیری چهره ناشی از درد، تنوع بروز واحدهای کنشی^۲ مرتبط با درد در افراد مختلف قابل مشاهده است [۳]. در این حالت یادگیری، کارایی یک روش مستقل از فرد دچار مشکل خواهد شد و این درحالی است که در واقعیت جلوه‌های کنشی چهره خودانگیخته می‌باشند. رابطه کارایی سامانه‌های تحلیل اطلاعات چهره با مسئله وابسته به فرد بودن

۱ مقدمه

افراد مختلف در نوشتن، صحبت کردن و بروز دادن جلوه‌های هیجانی از خود دست‌نویستار، گفتار و جلوه‌های متفاوتی را تولید خواهند کرد.

این مقاله در مهرماه ۱۳۹۹ ارسال، در دی‌ماه بازنگری و در بهمن‌ماه همان سال پذیرفته شد.

این پژوهش با پشتیبانی مالی پژوهشکده پردازش هوشمند علائم، بر اساس قرارداد شماره ۹۱/۱۱۳۰/ت/۹۱ به تاریخ ۹۱/۰۷/۳۰ انجام شده است.

^۱ پژوهشگاه توسعه فناوری‌های پیشرفته و دانشکده مهندسی پزشکی، دانشگاه صنعتی امیرکبیر، تهران، ایران

رایانامه: a.mohammadian@aut.ac.ir

^۲ دانشکده مهندسی برق، دانشگاه صنعتی امیرکبیر، تهران، ایران

رایانامه: aghaeini@aut.ac.ir

^۳ دانشکده مهندسی پزشکی، دانشگاه صنعتی امیرکبیر، تهران، ایران

رایانامه: towhidkhah@aut.ac.ir

نویسنده مسئول: امین محمدیان

DOR: 20.1001.1.23831197.1400.8.4.5.5

¹ Spontaneous

² Action Units

۱ کارهای مرتبط

در کنار روش‌های تطبیق مدل، نگاشت تطبیق ویژگی‌ها از جمله روش‌های مطرحی است که به جای نزدیک کردن مدل به فضای فرد جدید ویژگی‌های فرد هدف را به مدل نزدیک می‌کند. به‌عنوان مثال در روش $fMLLR^4$ ، ماتریس انتقال به بردارهای ویژگی داده فرد جدید اعمال می‌شود و آنها را به مدل‌های مستقل از گوینده، نزدیک می‌کند. مزیت اصلی $fMLLR$ این است که ماتریس انتقال محاسبه می‌شود و برای تبدیل به کار گرفته شود [۱۴، ۱۵].

همچنین در کار دیگری جهت تطبیق به حوزه^۵ جدید بصورت بدون سرپرستی روش یادگیری لغت نامه تفکیک‌پذیر خود را به فضای هدف تطبیق داده است تا نسبت به تغییرات حوزه مقاوم باشد [۱۶]. علاوه بر این از روش‌های تطبیق به حوزه جدید مبتنی بر یادگیری عمیق برای بازشناسی جلوه‌های هیجانی چهره استفاده شده است [۱۷].

برای مدل‌سازی جهت کشف زیرفضای تغییرات، مدل‌های خطی چون تحلیل مولفه اساسی (PCA^۶) مورد استفاده قرار گرفته‌اند. در توسعه تحلیل زیرفضا، مدل‌های مبتنی بر چند عامل متعامد چون روش دوخطی و تحلیل تنسور چندخطی پیشنهاد شده‌اند. در کاری که روش دوخطی پیشنهاد کرده چارچوب تفکیک سبک از محتوا به شکل مناسبی فرموله شده است [۱۸]. در کار آنها اطلاعات چهره به دو نوع اطلاعات چرخش سر و اطلاعات افراد به‌عنوان سبک و محتوا تجزیه شده است. در توسعه این مدل‌ها استفاده از تابع پایه غیرخطی جهت رابطه متغیر محتوا با مشاهدات مورد توجه قرار گرفته است [۱۹]. همچنین مدل‌های پویای فرایند گوسی (GPDMS^۷) جهت یادگیری حرکات انسانی و تحلیل سری‌های زمانی غیرخطی پیشنهاد شده است [۲۰]. در کار دیگری یک شبکه عصبی تفکیک‌کننده خمینه^۸ جهت تفکیک زیرخمینه‌های غیرخطی اطلاعات هویتی از اطلاعات هیجانی پیشنهاد شده است [۲۱].

دسته‌ای از روش‌ها صرفاً به انتقال سبک بین دو تصویر چهره افراد پرداخته‌اند و لزوماً دغدغه آنها تعدیل اثر سبک نیست. بطوریکه در کاری از ظرفیت شبکه‌های عصبی کانولوشنی عمیق برای این انتقال در زیر بخش‌های تصویر استفاده شده است [۲۲]. یا از یادگیری تخصصی برای ارائه یک انتقال سبک سریع استفاده کرده است که برای فضاهای چند حوزه‌ای مناسب است [۲۳].

از ماشین انتقال انتخابی برای بهبود تشخیص جلوه‌های بصری پرت با استفاده از طبقه‌بند تک طبقه استفاده کرده است [۲۴]. در حوزه بازشناسی جلوه‌های بصری از توالی تصویر، در کار دیگری به تقویت مدل یادگیری در مواجهه با تنوعات ناشی از سبک فرد جدید پرداخته است. در این کار به تولید توالی مجازی برای فرد جدید پرداخته و از آن برای تقویت یادگیری استفاده کرده است [۲۵]. در رویکرد دیگری یک چارچوب جهت انتقال سبک پیشنهاد شده است. در این چارچوب با تعریف مجموعه نقاط منبع به‌عنوان داده مخصوص هر فرد و مجموعه

نشان داده شده است [۲، ۴]. در کار دیگری کاهش نرخ بازشناسی جلوه‌های هیجانی چهره در نتیجه استفاده از روش‌های مستقل از فرد نشان داده شده است [۵، ۶].

عوامل تأثیرگذار بر ایجاد تنوعات در بروز جلوه‌های چهره شامل دو دسته درونی و بیرونی هستند [۷]. منبع تغییرات درونی را می‌توان عواملی چون هویت، جنسیت، سن و نژاد دانست. هر یک از این موارد چالش‌هایی را برای توسعه سامانه‌های بازشناسی در محیط‌های واقعی ایجاد کرده‌اند [۸]. در دسته تنوعات درونی، اثر فرد در بروز تنوعات مؤثر است. این اثر به‌صورت بین‌فردی و درون‌فردی عامل ایجاد تنوعات است. درون‌فردی یعنی اینکه بروز یک جلوه در دو بار متوالی توسط یک فرد دارای اختلاف است و اثر بین‌فردی یعنی دو فرد مختلف در بروز یک جلوه یا واحد کنشی با یکدیگر اختلاف دارند. اگر مجموعه‌ای از مشخصات که مستقل از نوع کنش چهره بوده و وابسته به فرد برزدهنده است را با عنوان سبک^۳ بروز تعریف شود. افراد دارای سبک بیان، سبک نوشتن و سبک بروز حالات مخصوص به خود هستند. می‌توان فرض کرد که برای بروز یک دست خط یا جلوه هیجانی توسط یک فرد، دو عامل سبک و محتوا با یکدیگر تعامل می‌کنند تا نمونه‌ای تولید شود [۹]. سبک متفاوت افراد در بروز این محتوا خود عاملی جهت ایجاد تنوع در بروز آن نمونه است. به همین دلیل اگر بتوان اثر سبک را تعدیل کرد می‌توان پیش‌بینی کرد که محتوا بهتر شناسایی خواهد شد.

همچنین پیش‌بینی می‌شود تحلیل محتوا در داخل هر سبک بهبود بازشناسی محتوا را فراهم خواهد آورد. البته تفکیک محتوا از سبک مشکلاتی خواهد داشت. اینکه لزوماً مستقل از هم تغییر نمی‌کنند یا اینکه سبک و محتوا هم‌زمان قابلیت تغییر دارند و رابطه ترکیب آنها لزوماً خطی تعریف نشده و داده هم در بعضی موارد فاقد برچسب است.

در نگاه احتمالاتی به موضوع، تنوعات در داده آزمون باعث فاصله گرفتن توزیع نمونه‌های آزمون از آموزش خواهد شد. سازگاری فضای آزمون و آموزش عاملی برای افزایش قدرت تعمیم‌دهی روش یادگیرنده است [۱۰]. به‌طوری‌که در صورت وجود عدم سازگاری بین دو فضا حتی با کاهش خطای تقریب (خطای یادگیری داده آموزش) بر روی داده آموزش، نمی‌توان انتظار کاهش خطای تخمین (خطای مربوط به تشخیص صحیح داده آزمون) بر روی فضای آزمون را داشت. در نتیجه تعمیم‌دهی روش یادگیرنده کاهش خواهد یافت. اما در صورتی‌که اطلاعاتی پیرامون فضای آزمون در دسترس باشد، یک راهکار به‌کارگیری آن در فرایند یادگیری و ایجاد تمایل در آن است تا تعمیم‌دهی بر روی این حوزه جدید افزایش پیدا کند. به‌عنوان مثال می‌توان بر اساس دانش محدود از فرد جدید یا داده بدون برچسب از حوزه جدید نمونه‌های مجازی تولید و برای تقویت یادگیری سیستم بازشناسی استفاده کرد [۴، ۱۱-۱۳]. راهکار دیگر این است که توزیع فضای آزمون را به فضای آموزش نزدیک کرد تا سازگاری بین آنها برقرار گردد.

⁴ Feature Maximum Likelihood Linear Regression

⁵ Domain

⁶ Principal component analysis

⁷ Gaussian Process Dynamical Models

⁸ Manifold

³ Style

می‌شود. الگوریتم‌ها همه در قالب یک چارچوب پیشنهاد شده‌اند که تابع انتقال سبک^۹ (STF) نامگذاری شده است. جهت بررسی اثر الگوریتم‌های پیشنهادی نتایج بر روی تشخیص درد و واحدهای کنشی چهره به کار گرفته شده است و نتایج آن با روش‌های موجود مقایسه شده است.

۲ روش‌شناسی

در این کار، رویکرد انتقال سبک، برای تعدیل اثر سبک فرد جدید انتخاب شده است. همانطور که توضیح داده شد هدف اصلی انتقال سبک، ایجاد نوعی سازگاری بین فضای آزمون و آموزش است. با توجه به استفاده نسبتاً مشابه از یک چارچوب در بین روش‌های ارائه شده، ابتدا این چارچوب تعریف شده است، سپس روش‌های پیشنهادی ارائه شده بر اساس این چارچوب توضیح داده شده‌اند.

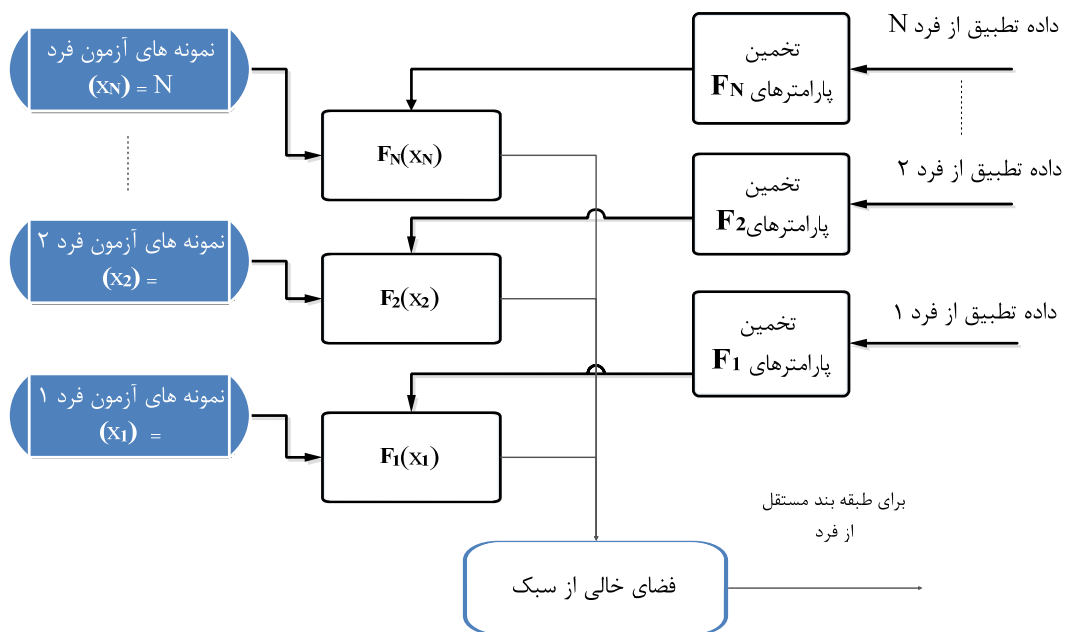
۲-۱ تعریف چارچوب انتقال سبک

همان‌طور که در شکل ۱ نشان داده شده است، با استفاده از داده تطبیق فرد n ، تابع $\{F_n | n = 1: N\}$ برای انتقال سبک هر فرد تنظیم می‌شود. که N تعداد کل افراد است. این تابع مخصوص هر فرد است که داده‌های جدید آن‌را به فضای خالی از سبک منتقل می‌کند تا بازنمایی محتوای نمونه جدید بهتر انجام شود. اگر مجموعه نمونه‌های آموزش با برچسب $\{x_i^{tr}, y_i^{tr}\}_{i=1}^{tr}$ موجود باشد، که $x_i^{tr} \in R^D$ نمونه i ام از نمونه‌های آموزش که به‌طور مستقل از توزیع احتمالی با چگالی $p_{tr}(x)$ به‌دست آمده، $y_i^{tr} \in \{1, \dots, K\}$ برچسب آموزش و K تعداد طبقات با توزیع احتمال شرطی با چگالی $p(y|x = x_i^{tr})$ است و علاوه بر نمونه‌های آموزش، مجموعه نقاط بدون برچسب $\{x_i^{te}\}_{i=1}^{te}$ موجود باشد که $x_i^{te} \in R^D$ نمونه i ام از نمونه‌های آزمون است که به‌طور مستقل از توزیع احتمالی با چگالی $p_{te}(x)$ به‌دست آمده است.

نقاط هدف به‌عنوان پارامترهایی از طبقه‌بند مستقل از فرد، توسط یک نگاشت نقاط منبع را به سمت نقاط هدف انتقال می‌دهد. در نتیجه این نگاشت سبک افراد مختلف به فضایی بدون سبک افکنده می‌شود. این کار ابتدا بر روی دست خط به کار گرفته شد [۲۶] اما در کار دیگری نتیجه به کارگیری آن برای بازنمایی جلوه‌های درد در چهره به کار گرفته شده است [۲۷]. در تکمیل کار انجام گرفته با فرض داده کافی برای تطبیق و جهت ارائه یک نگاشت انتقال که قادر باشد به‌طور کامل و مناسبی نمونه‌های جدید را منتقل کند نگاشت غیرخطی انتقال سبک توسعه داده شده و به کار گرفته شده است [۲۸]. در کار دیگر برای تعدیل اثر سبک از به کارگیری نگاشت انتقال سبک بر روی توالی از تصاویر جلوه‌های بصری افراد استفاده شده است [۲۹].

در این کار با فرض اینکه مشابه کارهای قبلی، سبک افراد در بروز حالات موضوع علت تنوعات بین فردی است پیشنهاداتی برای مواجهه با آن پیشنهاد شده است. قابل ذکر است که الگوریتم‌های تعدیل سبک پیشنهادی زیر دسته روش‌های انتقال یادگیری (نه روش‌های یادگیری یا ساخت مدل) و مرتبط با بحث تطبیق به فضای آزمون است. فضای الگوریتم پیشنهادی مربوط به مرحله به کارگیری مدل اصلی و نه ساخت مدل اصلی است.

بنابراین مسئله اصلی نحوه انتقال مناسب نمونه‌های آزمون نسبت به مدل اصلی آموزش دیده است. این رویکرد در کارهای گذشته هم مورد توجه قرار گرفته است. اما تفاوت روش پیشنهادی این است که در مقایسه با روش‌های که از نگاشت برای انتقال سبک استفاده می‌کنند، الگوریتم‌هایی را پیشنهاد داده است که در حین سادگی کارایی مناسب داشته باشند. در قالب تابعی این الگوریتم رابطه بین نمونه جدید و نمونه تطبیق موجود بررسی و انتقال مناسب برای آنرا پیشنهاد می‌کند. بطوریکه در این روش نحوه انتقال سبک به شکل مناسب‌تر و با حساسیت کمتری نسبت به حجم داده تطبیق تخمین زده و سپس انجام



شکل ۱ - روندنمای چارچوب تابع انتقال سبک

تابع F_N سبک فرد شماره N یعنی A^{SN} را به $A^{Speople}$ نزدیک می‌کند و برای محتوای c می‌شود (۷) و این باعث می‌شود تا تابع چگالی توزیع احتمال $p_{te}(x)$ به $p_{tr}(x)$ نزدیک شود.

$$A_N \mu_c^{te} \cong \mu_c^{tr} \quad (۷)$$

۲-۲ تابع انتقال سبک

در گسترش کار قبل [۲۷] که روشی پیشنهاد شد که به جای انتقال نمونه‌های فرد جدید توسط نگاشت، این کار توسط بردار انتقال سبک انجام می‌شد. هدف بردار تخمینی نزدیک کردن توزیع داده آزمون به توزیع فضای آموزش بود تا قدرت تعمیم‌دهی روش بازشناسی را افزایش دهد. اما در این کار در طراحی و ارائه روش‌های پیشنهادی موارد زیر در نظر گرفته شده است:

- مبحث انتقال سبک از نوع انتقال نمونه‌ها در داخل یک فضا است و لزوماً نیازمند نگاشت نیست. در صورتیکه بتوان تنها یک هدف برای انتقال مشخص کرد، تابع انتقال می‌تواند یک بردار انتقال باشد.
- روش بگونه‌ای باشد که در صورت در دسترس بودن حجم کم داده به عنوان داده تطبیق کارایی لازم را دارا باشد.
- این بهبود در شرایطی مزیت رقابتی خواهد داشت که امکان نگهداری داده آموزش نباشد و صرفاً مدل آموزش دیده و مشخصات آماری داده آموزش (میانگین) در دسترس است. بنابراین روش‌های تعدیل سبک کمک‌کننده خواهند بود.
- همچنین فرضیات روش پیشنهادی در رابطه با توزیع فضای آموزش و آزمون در تاثیر گذاری بر کارایی مدل مستقل هم حاکم باشند.

اما در دسته روش‌های مبتنی بر تابع انتقال سبک، مشابه بردار انتقال سبک به جای انتقال نمونه‌ها توسط نگاشت، از الگوریتمی مبتنی بر بردار انتقال سبک استفاده می‌شود. اما تفاوتش در این است که فرضیات جدید را در نظر می‌گیرد. فرض اول اینکه تنها یک بردار انتقال برای هر طبقه کافی است و فرض دوم بیشتر از یک بردار انتقال را برای هر طبقه در نظر می‌گیرد. بر اساس این فرضیات، در ادامه روش نزدیکترین بردار انتقال سبک (STNV) نسخه ۱ با فرض تک خوشه و نسخه ۲ با فرض چند خوشه برای هر طبقه پیشنهاد شده است. روندنمای الگوریتم روش STNV نسخه اول (STNV1) در شکل ۲ نشان داده شده است و جزئیات آن در جدول ۱ آورده شده است.

در مواردی که $p_{tr}(x) \neq p_{te}(x)$ ، توزیع نمونه‌ها بین دو فاز آموزش و آزمون متفاوت است، هدف اولیه، الگویی است که $p_{te}(x)$ را به $p_{tr}(x)$ نزدیک‌تر کند تا بتواند برچسب y^{te} دقیق‌تری را برای نمونه x^{te} پیش‌بینی کند. در بروز کنش‌های چهره تفاوت توزیع نمونه‌های دو فضا می‌تواند ناشی از سبک متفاوت افراد باشد.

یک بردار مشاهده با سبک (s) و محتوای (c) را می‌توان به صورت x^{sc} نوشت. مدل $x^{sc} = f(a^s; b^c; W)$ برای آن فرض می‌شود که در آن برداری است که سبک s نامتغیر با زمان را توصیف می‌کند، b^c برداری است برای توصیف محتوای c و W که مجموعه‌ای از پارامترهاست که بیانگر تعامل بین این دو عامل است. فرض می‌شود که f یک مدل دوخطی غیرمتقارن چون (۱) است که در آن w_{iju} با سبک تغییر می‌کند [۱۸].

$$x_{iu}^{sc}(t) = \sum_j a_{ju}^s b_{tj}^c \quad (۱)$$

که در آن $a_{ju}^s = \sum_i a_i^s w_{iju}^s$ و i, j, u به ترتیب بیانگر مؤلفه‌های سبک، محتوا و بردار مشاهدات است. فرمول در فرم ماتریسی به صورت (۲) می‌شود.

$$x^{sc}(t) = A^s b^c \quad (۲)$$

ماتریس $A^s \in R^{D \times J}$ بیانگر نگاشت خطی با سبک خاص از فضای محتوا به فضای مشاهده است. برای دو فرد l و m برای یک محتوای مشابه c، اگر $A^{sl} \neq A^{sm}$ سپس مشاهدات آنها نیز $x^{slc} \neq x^{smc}$ اگر $A^{Speople}$ را چون (۳) برای $N-1$ فرد آموزش تعریف شود و μ_c^{te} چون (۴) تعریف گردد.

$$A^{Speople} = \frac{1}{N-1} \sum_{n=1}^{N-1} A^{Sn} \quad (۳)$$

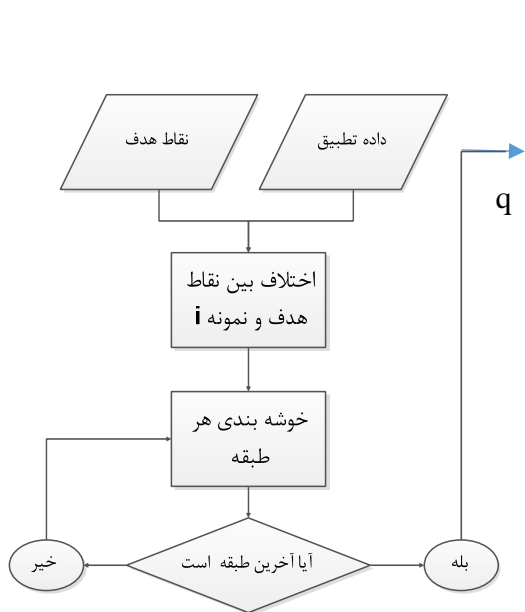
$$\mu_c^{te} = \frac{1}{O_{te}} \sum_{o=1}^{O_{te}} x_o^{te} = \frac{1}{O_{te}} \sum_{o=1}^{O_{te}} A^{SN} b^c \quad (۴)$$

O_{te} بیانگر شماره مشاهدات آزمون و N تعداد افراد آموزش است. برای فرد شماره N، اگر A^{SN} توسط تابع F_N به $A^{Speople}$ نزدیک‌تر شود، سپس $A_N A^{SN} \cong A^{Speople}$ تابع F_N همچون (۵) فاصله بین μ_c^{te} و μ_c^{tr} را کاهش می‌دهد.

$$\begin{aligned} A_N \mu_c^{te} &= A_N \frac{1}{O_{te}} \sum_{o=1}^{O_{te}} x_o^{te} = \frac{1}{O_{te}} \sum_{o=1}^{O_{te}} F_N (A^{SN} b^c) \\ &\cong \frac{1}{O_{te}} \sum_{o=1}^{O_{te}} A^{Speople} b^c = A^{Speople} b^c \end{aligned} \quad (۵)$$

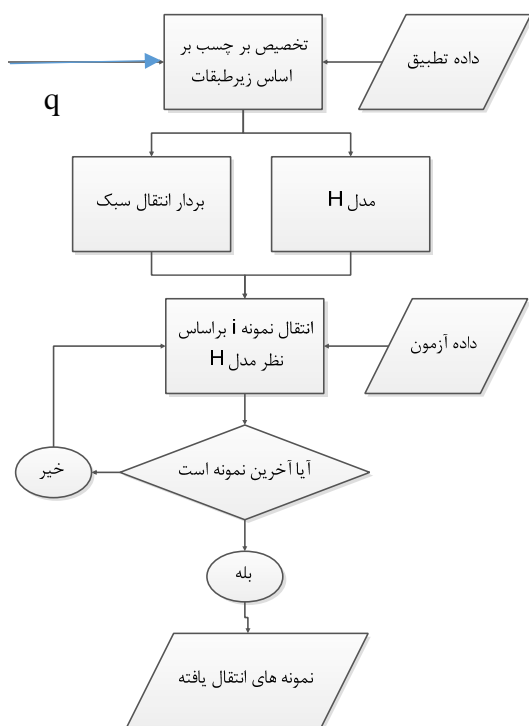
و μ_c^{tr} معادل (۶) می‌شود.

$$\mu_c^{tr} = \frac{1}{N-1} \sum_{n=1}^{N-1} A^{Sn} b^c = A^{Speople} b^c \quad (۶)$$



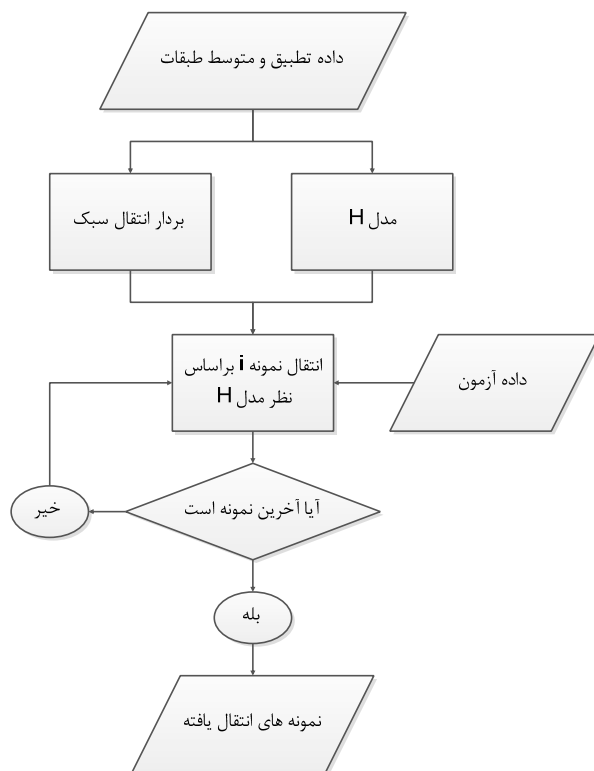
شکل ۳- روندنمای الگوریتم روش STNV2

ادامه روند در شکل ۴ نشان داده شده است.



شکل ۴- روندنمای الگوریتم روش STNV2 (ادامه)

STNV2 مشابه روش STNV1 نیز روی نتایج نمونه‌هایی که انتقال داده برای آنها انجام نشده، اثری ندارد. اختلاف در این است که برای هر طبقه بیشتر از یک بردار انتقال سبک در نظر گرفته می‌شود. متناسب با توزیع نمونه تطبیق و تعداد آن، تعداد بردار مناسب متفاوت خواهد بود. در قسمت تعیین این بردارها، از قابلیت خوشه‌بندی فضای توزیع نمونه‌های تطبیق استفاده شده است. برای اینکه پراکندگی داده آزمون در نظر گرفته شود، بردار هدفی در نظر گرفته شده و فاصله همه نقاط داده تطبیق با نقطه هدف معادلشان محاسبه شده است. برای هر طبقه این فاصله‌ها خوشه‌بندی شده‌اند. تعداد طبقات داده تطبیق از طریق خوشه‌بندی به تعداد بیشتر (تعداد طبقات \times تعداد خوشه هر طبقه)



شکل ۲- روندنمای الگوریتم روش STNV1

مشابه روش STM، این روش نیز روی نتایج نمونه‌هایی که انتقال داده برای آنها انجام نشده است، اثری ندارد، چراکه تغییری در طبقه‌بند ایجاد نمی‌کند. اما این روش جهت کاهش اثر سبک، برای سبک نفر جدید، بردارهای انتقالی را فرض می‌کند که سعی دارد تا اثر سبک را در نمونه‌های جدید کاهش دهد. اهداف مورد نظر این انتقال همان میانگین طبقات می‌باشد که جزو نواحی با احتمال بالای توزیع داده آموزش هستند.

جدول ۱- الگوریتم روش STNV1

ورودی:	داده تطبیق برچسب‌دار $\{\mathbf{x}_i^{ad}, \mathbf{y}_i^{ad}\}_{i=1}^{l_{ad}}$
	داده آزمون بدون برچسب $\{\mathbf{x}_i^{ts}, \mathbf{y}_i^{ts}\}_{i=1}^{l_{ts}}$
	طبقه‌بند مستقل $G(\mathbf{x}, k)$
	بردار انتقال سبک به صورت (۶) تعریف می‌شود.
	$S_k = \boldsymbol{\mu}_k^{tr} - \boldsymbol{\mu}_k \quad (6)$
	$\boldsymbol{\mu}_k$ بردار میانگین طبقه k ام از داده تطبیق و $\boldsymbol{\mu}_k^{tr}$ بردار میانگین طبقه k ام از داده آموزش است.
	مدل H توسط نمونه‌های $\{\mathbf{x}_i^{ad}, \mathbf{y}_i^{ad}\}_{i=1}^{l_{ad}}$ آموزش می‌بیند.
	for $l = 1: l_{ts}$
	$\mathbf{y}_l^{ts} = H(\mathbf{x}_l^{ts})$
	$\hat{\mathbf{x}}_l^{ts} = \mathbf{x}_l^{ts} + S_{\mathbf{y}_l^{ts}}$
	End
	خروجی: $\hat{\mathbf{x}}_l^{ts}$
	مرحله آزمون:
	$\mathbf{y}_l^{final} = \underset{k=1:K}{\operatorname{argmin}} G(\hat{\mathbf{x}}_l^{ts}, k)$

روندنمای الگوریتم روش STNV نسخه دوم (STNV2) در شکل ۳ نشان داده شده است و جزئیات آن در جدول ۲ نشان داده شده است.

تخصیص نزدیکترین بردار انتقال سبک هم نزدیک است. از آنجا که داده تطبیق معمولاً حجم کمی دارد، اجرای این طبقه‌بند نیازمند حافظه زیاد نیست.

$$T_N = A^{S_{people}} - A_0^{SN} \quad (11)$$

اگر همه نمونه‌های فرد جدید توسط T_N در فضا منتقل شوند آنگاه μ_c^{te} منتقل شده می‌شود.

$$\mu_c^{te} = A_0^{SN} + T_N + \frac{1}{I_{te}} \sum_{i=1}^{I_{te}} \frac{1}{O_{te}} \sum_{o=1}^{O_{te}} (A_t^{SN}) b^i$$

$$= A^{S_{people}} \quad (12)$$

$$+ \frac{1}{I_{te}} \sum_{i=1}^{I_{te}} \frac{1}{O_{te}} \sum_{o=1}^{O_{te}} (A_t^{SN}) b^i$$

و μ^{tr} معادل ۱۳ می‌شود.

$$\mu^{tr} = \frac{1}{I_{tr}} \sum_{i=1}^{I_{tr}} \frac{1}{N_i - 1} \sum_{n=1}^{N_i - 1} (A_0^{SN} + A_t^{SN} b^i)$$

$$= A^{S_{people}} \quad (13)$$

$$+ \frac{1}{I_{te}} \sum_{i=1}^{I_{te}} \frac{1}{N_i - 1} \sum_{n=1}^{N_i - 1} (A_t^{SN}) b^i$$

N_i تعداد نمونه برای هر طبقه از محتوا و I_{tr} تعداد طبقات محتواست. قسمت متغیر با محتوا تغییر نکرده اما بخش نامتغیر با زمان، توسط بردار انتقال T_N سبک فرد شماره N یعنی A^{SN} به $A^{S_{people}}$ نزدیک می‌کند. مستقل از محتوای موجود در نمونه‌های فرد N ، بردار μ^{te} را به μ^{tr} نزدیکتر می‌کند. این باعث می‌شود تا تابع چگالی توزیع احتمال $p_{te}(x)$ به $p_{tr}(x)$ نزدیکتر شود.

۳ آزمایشات

۳-۱ دادگان خودانگیخته UNBC-McMaster

این دادگان توسط همکاری دانشگاه مک‌مستر و بریتیش کلمبیای شمالی تهیه شده است [۳]. دادگان برای کمی‌سازی درد توسط واحدهای کنشی مرتبط با آن تهیه و مورد استفاده قرار گرفته است و قابل استفاده برای بازشناسی شدت جلوه‌های درد در چهره و واحدهای کنشی مرتبط با آن است. از افراد درحین حرکت دادن شانه‌ای که درد داشته و نداشته و در شرایط فعال و غیرفعال خود فرد فیلم‌برداری انجام شده است. در شرایط فعال بودن فرد، فرد خودش شانه خود را می‌چرخاند. در شرایط غیرفعال، یک فیزیوتراپیست مسئول حرکت بوده است. در آزمایشات این مقاله تنها از داده مربوط به شرایط فعال (۲۴ فرد) استفاده شده است که مشخصات آن در جدول ۳ آورده شده است. زاویه دید دوربین در آغاز از روبرو است. برای ارزیابی یکی بودن نظرات افراد نمره‌دهنده تعدادی قاب به صورت تصادفی انتخاب شده است و به صورت مستقل نمره‌دهی شده که درصد هماهنگی آنها ۹۵٪ بوده است.

افزایش پیدا کرده است. طبقه‌بندی که با نمونه‌های تجزیه شده آموزش دیده است طبقه احتمالی نمونه جدید را تشخیص می‌دهد. براساس طبقه تخمین زده شده بردار انتقال معادل آن برای انتقال آن نمونه استفاده می‌شود.

جدول ۲- الگوریتم روش STNV2

ورودی : داده تطبیق برجسب‌دار $\{x_i^{ad}, y_i^{ad}\}_{i=1}^{I_{ad}}$
داده آزمون بدون برجسب $\{x_i^{ts}, y_i^{ts}\}_{i=1}^{I_{ts}}$
طبقه‌بند مستقل $G(x, k)$
مجموعه نقاط هدف که به صورت (V) محاسبه می‌شود [۲۶].

$$P_{k_i} = \mu_{k_i} + (x_i - \mu_{k_i}) \times \min\left\{1, \frac{\rho}{d(x_i, k_i)}\right\}$$

$$1 < k_i < K \quad (V)$$

x_i ، i امین نقطه از مجموعه منبع، μ_{k_i} بردار متوسط طبقه k_i ، K تعداد طبقات و $d(x_i, k_i)$ فاصله مالهالانوبیس^{۱۱} از طبقه k_i است که به صورت (۸) تعریف شده است.

$$d(x_i, k_i) = \sqrt{(x_i - \mu_{k_i})^T \Sigma_{k_i}^{-1} (x_i - \mu_{k_i})} \quad (8)$$

پارامتر $\rho \geq 0$ محدود کننده فاصله گرفتن نقطه افکنده شده از متوسط طبقه است: $d(P_{k_i}, k_i) < \rho$

for $k = 1$ to K do

اگر $X_k^{ad} = \{x_{k1}^{ad}, \dots, x_{kn_k}^{ad}\}$ نمونه‌های طبقه k باشند.
مجموعه $E = \{e_{ik}\}$ را می‌سازیم. که $e(x_{ki}^{ad}, P_{k_i})$ بردار اختلاف بین نقاط هدف P_{k_i} و نمونه i به صورت (۹) است.

$$e(x_{ki}^{ad}, P_{k_i}) = x_{ki}^{ad} - P_{k_i} \quad (9)$$

از الگوریتم K-means برای تقسیم‌بندی فضای هر طبقه به C زیرطبقه $\{C_1, \dots, C_C\}$ استفاده می‌شود.

اگر $\{x_{kci}^{adn} = x_{ki}^{adn}\}_{i=1}^{I_{adn}}$ and $\{y_{kci}^{adn} = M(C_c, y_{ki}^{adn})\}_{i=1, c=1}^{I_{adn}, C}$ آنجا y_{kci}^{adn} برجسب جدید نمونه i در زیر طبقه C از طبقه k است که از ترکیب برجسب طبقه و زیرطبقه، توسط M تخصیص داده می‌شود. در این حالت بردار انتقال سبک به صورت (۱۰) تعریف می‌شود.

$$S_{kc} = P_k - \mu_{kc} \quad (10)$$

μ_{kc} اشاره دارد به بردار میانگین مربوط به زیرطبقه C از طبقه k داده تطبیق

End for

آموزش می‌بیند. $\{x_i^{adn}, y_i^{adn}\}_{i=1}^{I_{adn}}$ توسط نمونه‌های H مدل
for $l = 1: I_{ts}$
 $y_l^{ts} = H(x_l^{ts})$
 $\hat{x}_l^{ts} = x_l^{ts} + S_{y_l^{ts}}$
End

خروجی:

مرحله آزمون:

$$y_l^{final} = \underset{k=1}^K \operatorname{argmin} G(\hat{x}_l^{ts}, k)$$

آنجا که G طبقه‌بند مستقل از فرد است و H در این کار طبقه‌بند K نزدیکترین همسایه (KNN^{12}) در نظر گرفته شده است. این طبقه‌بند بر اساس نزدیکترین نقاط، طبقه‌ای برای نمونه جدید پیش‌بینی می‌کند. این طبقه‌بند به دلیل سادگی و توانایی کار با داده کم انتخاب شده است. دلیل دیگر نگاه محلی آن به طبقه‌بندی است که به هدف ما برای

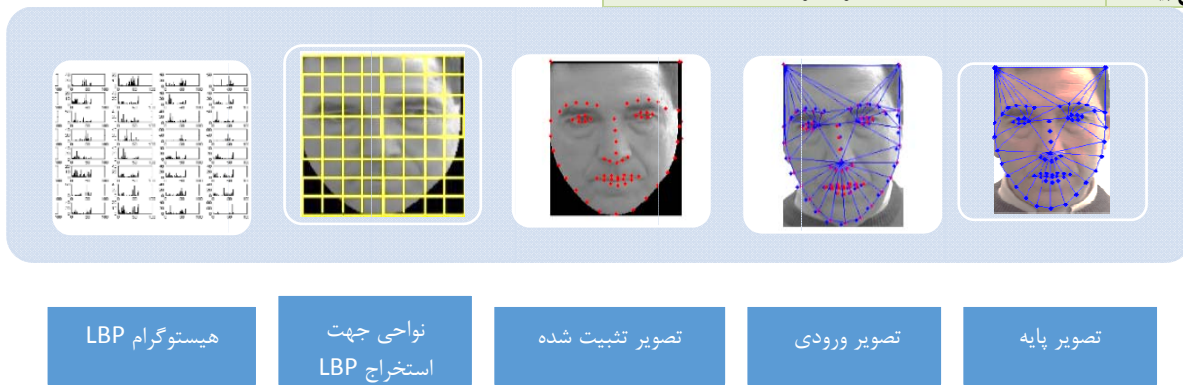
¹¹ Mahalanobis

¹² K-Nearest Neighbors (KNN)

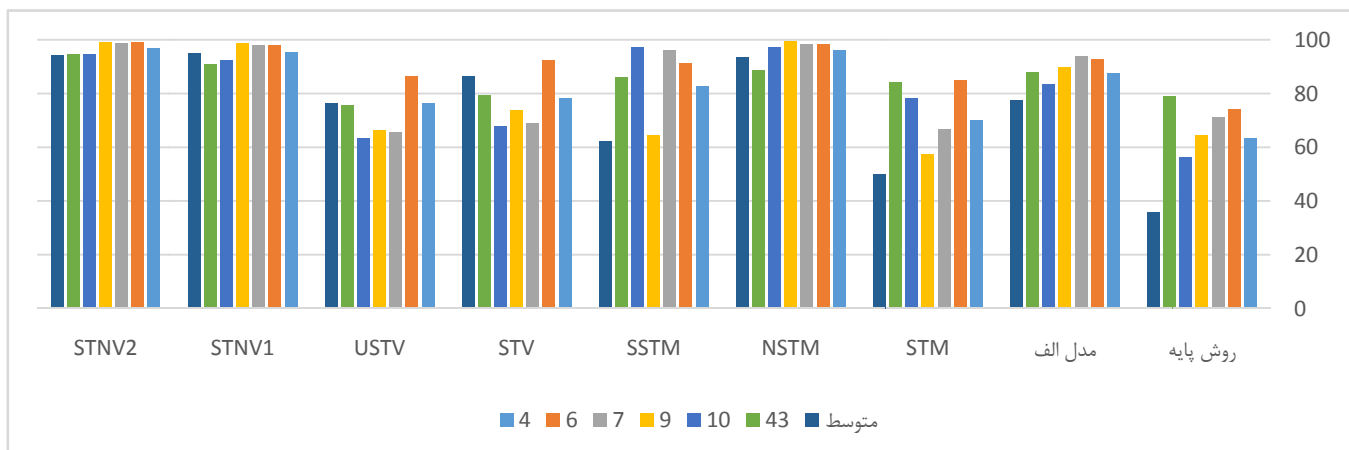
جدول ۳- مشخصات دادگان UNBC-McMaster

مشخصه	توضیح
تعداد افراد	۲۵ (مرد و زن)
تعداد عکس/ویدئو	نزدیک به ۴۸ هزار قاب

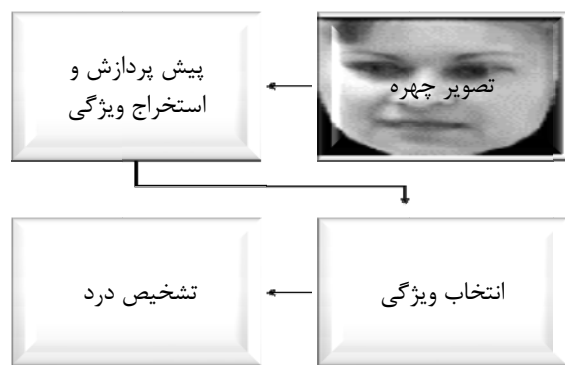
حالات	قاب های متوالی	استاتیک/ویدئو
روشنایی	تک چهره	تک یا چند چهره
متعلقات	رنگی	رنگی/خاکستری
نحوه برجسب زنی	مختلف	وضعیت سر
داخل اتاق	بیان جلوه های ناشی از درد	بیان حالات چهره
وضعیت ۶۶ نقطه مشخصه چهره برای هر قاب تصویر	حالات خودبخودی	نحوه توضیح بیان
قاب به قاب و مبتنی بر FACS		



شکل ۵ - روندنمای استخراج ویژگی از تصاویر چهره



شکل ۶ - AUC تشخیص واحدهای کنشی توسط روش های مختلف



نمونه ای از تصاویر دادگان UNBC-McMaster در شکل ۷ نشان داده شده است.



شکل ۷ - نمونه ای از تصاویر دادگان UNBC-McMaster [۳].

شکل ۸ - روندنمای تشخیص درد در تصاویر چهره

از آنجاکه در این دادگان به دلیل شرایط واقعی ثبت تصاویر، تغییرات سر قابل توجه است، استفاده از روش تثبیت مناسب، کمک قابل توجهی خواهد کرد. به همین دلیل تعدادی از نقاط مشخصه چهره که در نتیجه بروز جلوه های مختلف تغییرات کمتری دارند، در نظر گرفته شده اند و با ایجاد مثلث هایی در چهره

۲-۳ بازشناسی درد

مطابق روند شکل ۸ تشخیص درد در نظر گرفته شده است.

- توسط روش PAW^۱ چهره همه افراد به یک تصویر واحد افکنده شده است شکل ۵.
- در این قسمت از اطلاعات هندسی و بافت چهره جهت تشخیص درد استفاده شد که قبلاً نشان داده شده است که هر دو این اطلاعات برای تشخیص درد مفید می‌باشند [۳۰]. جهت توصیف اطلاعات بافت الگوهای دودوئی محلی و به‌عنوان اطلاعات هندسی موقعیت نقاط مشخصه، در آزمایشات به‌کار گرفته شده است. تصویر چهره به نواحی کوچکتر از ۸ تقسیم شد و برای هر ناحیه $LBP_{u2, 8; 1}$ (دارای بعد ۵۹) محاسبه شده است. $u2$ بیانگر استفاده از الگوهای یکنواخت، ۸ بیانگر نقاط ۸ گانه در اطراف دایره‌ای با شعاع یک است. از همه تصاویر الگوهای دودوئی محلی استخراج و ۵۰۰ ویژگی برتر با رتبه‌بندی تکی انتخاب شده است.
- در حالت بدون سرپرستی (USTV) از اطلاعات تطبیق بدون دانستن برجسب [۲۷].
- به‌کارگیری دو نسخه پیشنهادی از روش نزدیکترین بردار انتقال سبک جهت بهبود طبقه‌بند پایه (STNV1 و STNV2) نتیجه تشخیص ۶ واحد کنشی با معیار AUC برای روش‌های ارائه‌شده در شکل ۶ آورده شده است. اگرچه که از داده تطبیق استفاده شده ولی نتایج هم‌زمان بر روی داده تطبیق و آزمون گزارش شده است تا تعداد نمونه‌های هر فرد برای هر واحد کنشی متفاوت نباشد.
- جهت بررسی اثر تعداد داده تطبیق بر روی روش‌های مختلف نتایج تشخیص مستقیم درد توسط روش‌های مختلف به‌ازای ۱۲، ۲۰، ۴۰، ۶۰ نمونه و نصف نمونه‌های موجود برای فرد هدف در جدول ۴ آورده شده است.

جدول ۴- AUC تشخیص درد برای تعداد نمونه تطبیق مختلف

روش/تعداد نمونه	۱۲	۲۰	۴۰	۶۰	نصف
مدل الف	۷۵,۱۰	۷۵,۰۳	۷۴,۷۲	۷۴,۱۸	۸۱,۷۷
STM	۷۷,۹۲ ^۱	۷۹,۸۲ ^۱	۸۱,۶۴ ^۱	۸۱,۰۱ ^۱	۸۱,۵۸ ^۱
NSTM	۸۷,۱۵ ^{۱,۲}	۹۰,۱۹ ^۱	۸۹,۴۱ ^۱	۸۹,۶۰ ^۱	۹۵,۱۵ ^{۱,۲}
SSTM	۸۴,۵۷ ^{۱,۲}	۸۲,۱۳ ^۱	۸۲,۵۴ ^۱	۸۲,۹۴ ^۱	۸۴,۸۲ ^{۱,۲}
STNV1	۸۷,۱۶ ^{۱,۲}	۸۹,۵۰ ^{۱,۲}	۸۹,۶۷ ^{۱,۲}	۸۹,۶۲ ^{۱,۲}	۹۱,۷۳ ^{۱,۲}
STNV2	۸۷,۳۰ ^{۱,۲}	۸۹,۲۲ ^{۱,۲}	۸۹,۶۲ ^{۱,۲}	۹۰,۰۴ ^{۱,۲}	۹۳,۲۶ ^{۱,۲}
روش پایه			۷۵,۱۵		

"نصف" بیانگر استفاده از نصف داده موجود برای هر فرد است.

۱- معنی‌داری نسبت به روش پایه ۲- معنی‌داری نسبت به STM

"مدل الف" در شرایط داده کم عقب‌تر از روش STM قرار داشته و با افزایش داده (نصف داده فرد هدف) "مدل الف" و روش NSTM بر روش STM پیشی گرفته‌اند. روش STNV2 در بعضی از شرایط بر NSTM برتری داشته و با معیار AUC، روش NSTM در شرایط داده مناسب به عنوان برترین روش عمل کرده است. اما دو روش STNV1 و STNV2 به‌صورت معنی‌داری ($p < 0.05$) در شرایط مختلف داده نسبت به روش STM بهبود داشته‌اند.

کارایی چهار روش STM، STNV1، STNV2 و NSTM که از جهت روایی با یکدیگر در شرایط مختلف رقابت می‌کردند از جهت سرعت هم مورد بررسی قرار گرفت. در شکل ۹ زمان محاسبه (ثانیه) هر الگوریتم جداگانه گزارش شده است.

۳-۳ بازشناسی واحدهای کنشی

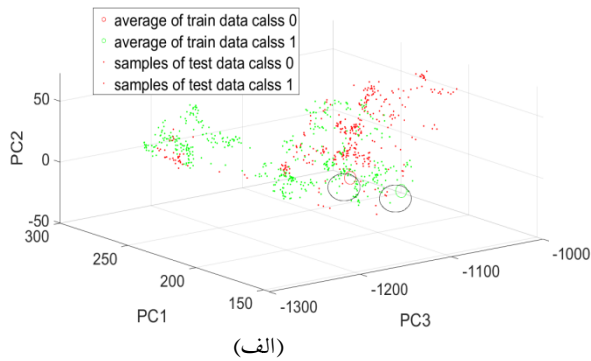
از آنجاکه عامل سبک به‌عنوان یک عامل درونی مؤثر بر کارایی تشخیص واحدهای کنشی است. پیش‌بینی می‌شود که اگر بتوان از طریق تعدیل اثر تغییرات بین سبکی در افراد آموزش و آزمون بهبودی در نرخ تشخیص ایجاد کرد. در روش پایه، ۳۷۷۶ ویژگی LBP با تعداد بین ۵۹ و ۴۴ ویژگی مربوط به نقاط مشخصه چهره از هر قاب استخراج شده است. رتبه‌بندی ویژگی‌ها در رابطه با قدرت تفکیک‌پذیری بین طبقات انجام شده و تعداد ۲۰۰ ویژگی برتر در هر مرحله بر اساس داده آموزش انتخاب شده‌اند. روش ارزیابی متقابل LOSO بر روی افراد مختلف جهت ارزیابی به‌کار گرفته شده است.

۴ نتایج آزمایشات

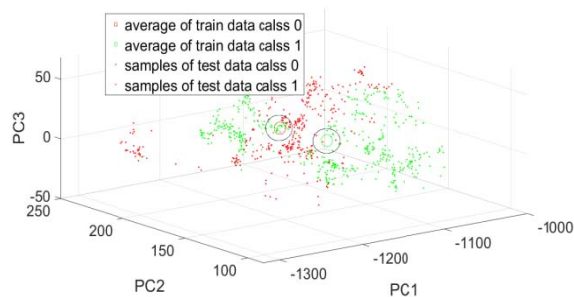
- در این قسمت روش‌های مختلفی به شرح زیر، جهت مقایسه پیاده‌سازی شده‌اند و نتایج آنها گزارش شده است:
- طبقه‌بند بیزی که توسط داده آموزش، آموزش دیده باشد (طبقه‌بند پایه).
 - طبقه‌بند بیزی که توسط داده آموزش و داده فرد هدف (داده تطبیق)، آموزش دیده است (مدل الف).
 - به‌کارگیری روش انتقال سبک خطی جهت بهبود طبقه‌بند پایه (STM) [۲۷].
 - به‌کارگیری روش انتقال سبک غیرخطی جهت طبقه‌بند روش پایه (NSTM) [۲۸].
 - به‌کارگیری روش انتقال سبک انتخابی جهت طبقه‌بند روش پایه (SSTM) [۲۸].
 - بردار انتقال سبک در حالت نیمه‌سرپرستی (STV) از اطلاعات تصاویر خنثی افراد [۲۷].

¹ Piecewise Affine Warp

نمونه‌های فرد جدید مقدار متوسط طبقه ۰ و طبقه ۱ نمونه‌های آموزش بصورت دایره وار نمایش داده شده است.



(الف)



(ب)

شکل ۱۱ - توزیع نمونه‌های یک فرد قبل (الف) و بعد از انتقال (ب) توسط الگوریتم STNV1

همانطور که نتایج شکل ۱۱ نشان می‌دهد وضعیت توزیع نمونه‌ها برای دو طبقه ۰ و ۱ درد بعد از انتقال نمونه‌ها توسط الگوریتم نسبت به دو متوسط طبقات ۰ و ۱ داده آموزش به وضعیت مناسبی تغییر وضعیت داده است.

۴-۱ مقایسه با نتایج کارهای دیگران

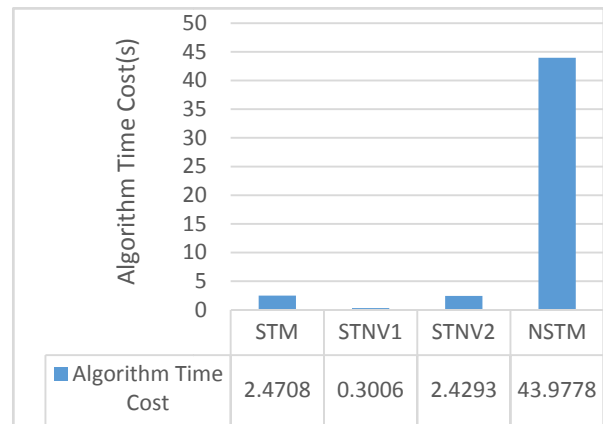
این روش نسبت به روش STM برای تعدادی از واحدهای کنشی نیز بهبود داشته است. نتایج به دست آمده با نتایج کارهای مشابه در جدول ۵ مقایسه شده است [۲, ۳].

جدول ۵- مقایسه نتایج AUCs تشخیص واحدهای کنشی بر روی

دادگان UNBC-McMaster

روش/واحد کنشی	۴	۶	۷	۹	۱۰	۴۳
[۳]	۵۷,۱۰	۸۵,۴	۸۰,۴	۸۵,۳	۸۹,۲	۸۷,۵
[۲]	-	۷۹,۲	۶۳,۴	-	۷۷,۲	۸۷,۵
[۲]	-	۹۰,۷۰	۹۲,۱۰	-	۹۳,۰۰	۹۲,۲
STNV2	۹۴,۳۰	۹۴,۷۷	۹۴,۹۰	۹۹,۲۰	۹۸,۶۹	۹۹,۱۶

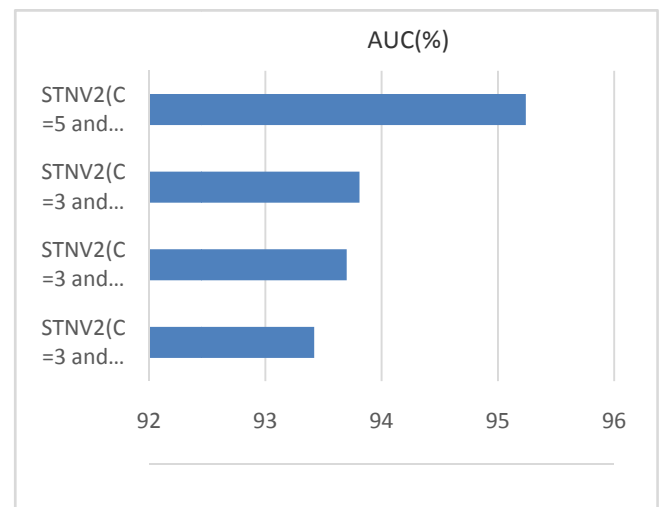
نتایج نشان می‌دهد که روش STNV2 نسبت به روش‌های موجود بر روی این دادگان مناسب‌تر عمل کرده است. همچنین نرخ تشخیص آن برای هر واحد کنشی بیانگر مؤثر بودن آن در تشخیص واحدها بوده است، خصوصاً واحدهایی چون ۴ و ۷ که فعالیت کوچکتری داشته و سخت‌تر تشخیص داده می‌شوند. در



شکل ۹ - هزینه زمان محاسبه روش‌های STM، STNV1، STNV2 و NSTM جهت بهبود تشخیص درد

همانطور که از نتایج مشخص است دو روش STNV1 و STNV2 در مقایسه با NSTM از سرعت بالای محاسباتی برخوردار هستند. در مجموع با در نظر گرفتن معیار سرعت و رویایی (AUC) این دور روش (STNV2 و STNV1) برتر از NSTM عمل می‌کنند.

همچنین جهت بررسی اثر پارامترهای تعداد خوشه (C) و تعداد نمونه همسایگی (K) در الگوریتم STNV2 نتایج تشخیص درد برای مقادیر متفاوتی از آنها در شکل ۱۰ نشان داده شده است.



شکل ۱۰ - اثر تعداد خوشه (Cluster) و تعداد نمونه همسایگی (K) در الگوریتم STNV2

با بررسی آماری که بین خروجی الگوریتم STNV2 توسط پارامترهای (تعداد خوشه و نمونه همسایگی) مختلف برای بهبود بازشناسی درد انجام گرفت، نتایج متفاوت بود اما نتایج آنها اختلاف معنی‌داری با حالت اولیه از جهت آزمون فرضیه نداشت.

جهت تحلیل بهتر کارکرد الگوریتم STNV1، بررسی وضعیت انتقال نمونه‌های آزمون در فضا بررسی گردید. برای این منظور نتایج وضعیت نمونه‌ها در فضا بعد از کاهش به سه بعد اول تحلیل مولفه‌های اصلی (PCA) قبل و بعد از استفاده از الگوریتم STNV1 در شکل ۱۱ نمایش داده شده است. در کنار نمایش

این است که مدل در مرحله استفاده باید داده آموزش را به همراه داشته باشد. در حالیکه در الگوریتم پیشنهادی، در حین اینکه از ظرفیت KNN برای بهبود بازشناسی استفاده کرده است اما نیازمند نگهداری نمونه‌های آموزش برای مرحله آزمون نیست. همانطور که در توضیح الگوریتم STNV1 و STNV2 در جداول ۱ و ۲ آورده شده است. از اطلاعات فضای آموزش تنها میانگین طبقات نگهداری شده است. پس نمونه‌های آموزش در مرحله تعدیل اثر سبک نیاز نمی‌شوند. این یک تفاوت اساسی بین به کارگیری قابلیت نزدیکترین همسایگی در روش پیشنهادی با روش طبقه‌بندی KNN است.

بنابراین الگوریتم‌های پیشنهادی در شرایطی امکان به کارگیری و در حین حال مزیت رقابتی دارند که **اولا** امکان نگهداری داده آموزش نبوده و تنها مدل آموزش دیده و مشخصات آماری (میانگین طبقات) داده آموزش در دسترس باشند. در ضمن داده‌ای از فرد جدید به عنوان داده تطبیق در اختیار داشته باشیم. **دوما** فرضیاتی که در رابطه با توزیع فضای نمونه‌ها و نحوه تصمیم‌گیری مدل مستقل در نظر گرفته شده است بر شرایط حاکم باشد. در غیر این صورت استفاده از این روش‌ها نباید نتایج مناسبی به همراه داشته باشد.

سوما فرض روش پیشنهادی این است که در قسمت مدل اصلی و مستقل، معمولا روشی استفاده شود که مانند KNN زیر مجموعه روش‌های کاهل نیست پس نیازمند نگهداری داده آموزش در مرحله آزمون نباشد بلکه مدل بر اساس داده آموزش ساخته می‌شود و در مرحله آزمون به کار گرفته می‌شود. حال وظیفه الگوریتم‌های پیشنهادی ارتقاء کارایی مدل مستقل آموزش دیده است.

نکته دیگر اینکه روش‌های انتقال سبک که زیر مجموعه روش‌های انتقال یادگیری قرار می‌گیرند. اساسا مربوط به مرحله یادگیری نمی‌شوند. بلکه به عنوان روش‌های پس یادگیری مورد استفاده قرار می‌گیرند. بر همین اساس این روش نباید با روش‌های یادگیری مدل (چون نوروفازی یا نزدیکترین همسایه) مورد مقایسه قرار گیرد.

در ارتقاء تابع انتقال سبک، می‌توان اثر **فاصله** سنج‌های مختلف در انتخاب نزدیکترین بردار را مورد بررسی قرار داد. همچنین می‌توان اثر مدل طبقه‌بندی مستقل (به خصوص طبقه‌بندهایی که مجاورت مبنای تصمیم‌گیری آنهاست) را در رابطه با الگوریتم پیشنهادی بیشتر مورد بررسی قرار داد. اگرچه که استفاده از KNN برای به کارگیری داده تطبیق با توجه به ویژگی آن (سادگی و توانایی کار با داده کم) شروع خوبی است. اما ارتقاء الگوریتم پیشنهادی از طریق جایگزینی برای KNN در بخش مدل یادگیری می‌تواند در کارهای دیگر مورد توجه قرار بگیرد.

همچنین در این کار فرضیات در نظر گرفته شده همه با فرض جمع‌پذیری خطی رابطه سبک و محتوا بود. در نظر گرفتن فرضیات دیگر چون تغییرات همزمان محتوا و سبک و استفاده از

این بخش نتایج روش STNV با دیگر نتایج گزارش شده بر روی دادگان UNBC-McMaster با معیارهایی چون F1، AUC، Hit و RR مقایسه شده است (جدول ۶).

جدول ۶- مقایسه نتیجه پیشنهادی با نتایج دیگران در تشخیص درد

مرجع	RR	Hit	AUC	F1	روش ارزیابی
[۳۱]	-	۸۱,۲	-	-	LOSO
[۳۲]	-	-	۸۳,۹	-	LOSO
[۲]	-	-	۸۹,۵	-	LOSO
[۳۰]	۹۶,۹	-	-	-	۱۰ بار ۱۰ دسته
[۳۳]	-	-	-	۵۷	LOSO
STNV2	-	-	۹۳,۲۶	-	LOSO

F1=F1 measure, AUC=area under the curve,
RR=recognition rate and Hit rate= true positive rate

این روش نسبت به روش STM برای تشخیص درد به صورت معنی‌داری بهبود داشته است. پیش‌بینی می‌شود که رویکرد روش STNV2 قابل استفاده برای دیگر مسائل یادگیری و بازشناسی نیز خواهد بود.

۵ جمع‌بندی و پیشنهادات

در این کار با هدف استفاده مناسب از داده تطبیق، تابع انتقال سبک پیشنهاد شد. نتایج آنها بر روی تشخیص واحدهای کنشی و تشخیص درد گزارش گردید. این دسته از روش‌ها در شرایط وجود اختلاف بین توزیع فضای آزمون و آموزش مطرح هستند، زیرا این اختلاف قدرت تعمیم‌دهی روش یادگیرنده را کم خواهد کرد. اما روش‌های پیشنهادی با استفاده از چارچوب انتقال سبک و داشتن اطلاعاتی از فرد جدید سعی در نزدیک کردن توزیع فضای آزمون به آموزش می‌کنند. در این رویکرد در راستای کاهش **خطای تخمین** بدون تغییر **خطای تقریب**، تنها با استفاده از انتقال نمونه‌های آزمون، خطای تخمین را کاهش داده و تعمیم‌دهی به حوزه جدید را افزایش می‌دهد. تشخیص درست شرایط مسائلی که سبک افراد در آنها نقش مهمی در کاهش نرخ بازشناسی ایفا خواهد کرد و استفاده مناسب از تابعی که قابلیت مناسب انتخاب جابجایی نمونه‌ها در فضا را دارا باشد عامل موفقیت خواهد بود.

در هسته الگوریتم‌های پیشنهادی از روش KNN استفاده شده است که این ابهام را بوجود می‌آورد که اگر نمونه‌های داده تطبیق را به داده آموزش اضافه کرده و از روش KNN برای طبقه‌بندی نمونه‌های جدید استفاده می‌شد نتیجه بهتری بدست می‌آمد. برای توضیح این مسئله لازم است مطالبی توضیح داده شود. اول اینکه روش KNN که زیر مجموعه روش‌های **کاهل**^۱ در مقایسه با روش‌های **مشاق**^۲ قرار می‌گیرد دارای یک عیب بزرگی است و آن

^۱ Lazy

^۲ Eager

- Proceedings of the 2004 IEEE Computer Society Conference on, 2004.
- [10] V. Vapnik, *The nature of statistical learning theory*. Springer, 2000.
- [11] M. Amin, A. Hassan, and T. Farzad, "Incorporating prior knowledge from the new person into recognition of facial expression," *Signal, Image and Video Processing*, vol. -, no. -, pp. -, 2014.
- [12] S. Schwarcz, A. Gorban, X. Gibert, and D.-S. Lee, "Adapting Style and Content for Attended Text Sequence Recognition," in *The IEEE Winter Conference on Applications of Computer Vision*, 2020, pp. 1597-1606.
- [13] B. Bozorgtabar, M. S. Rad, H. K. Ekenel, and J.-P. Thiran, "Using photorealistic face synthesis and domain adaptation to improve facial expression analysis," in *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019: IEEE, pp. 1-8.
- [14] ک. جعفری. "تطبیق سریع گوینده برای سیستم بازشناخت گفتار." کارشناسی ارشد، دانشکده مهندسی پزشکی صنعتی امیرکبیر، ۱۳۸۸.
- [15] R. Hariharan and O. Viikki, "An integrated study of speaker normalisation and HMM adaptation for noise robust speaker-independent speech recognition," *Speech Communication*, vol. 37, no. 3-4, pp. 349-361, 2002, doi: 10.1016/s0167-6393(01)00039-5.
- [16] S. Wu, G. Gao, Z. Li, F. Wu, and X.-Y. Jing, "Unsupervised visual domain adaptation via discriminative dictionary evolution," *Pattern Anal Applic*, pp. 1-11, 2020.
- [17] N. Kalischek, P. Thiam, P. Bellmann, and F. Schwenker, "Deep Domain Adaptation for Facial Expression Analysis," in *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, 3-6 Sept. 2019 2019, pp. 317-323, doi: 10.1109/ACIIW.2019.8925055.
- [18] J. B. Tenenbaum and W. T. Freeman, "Separating Style and Content with Bilinear Models," *Neural Computation*, vol. 1283, pp. 1247-1283, 2000. [Online]. Available: file:///G:/PHDsubject/new articles/style/dynamic/seperation manifold/2000 - Separating style and content with bilinear models. - Tenenbaum, Freeman.pdf
- [19] A. Elgammal and C.-S. Lee, "Homeomorphic Manifold Analysis (HMA): Generalized separation of style and content on manifolds," *Image and Vision Computing*, vol. 31, no. 4, pp. 291-310, 2013, doi: <http://dx.doi.org/10.1016/j.imavis.2012.12.003>.
- [20] J. M. Wang, D. J. Fleet, and A. Hertzmann, "Multifactor Gaussian process models for style-content separation," presented at the Proceedings of the 24th international conference on Machine learning, 2007.
- [21] S. Seyyedsalehi and S. Seyyedsalehi, "Simultaneous Learning of Nonlinear Manifolds Based on the Bottleneck Neural Network," (in English), *Neural*
- روش‌های جداسازی خمینه^۱ می‌تواند از رویکردهایی باشد که در کارهای آتی به آن توجه شود. علاوه بر این استفاده از قابلیت مدل‌های مولد چون شبکه‌های مولد تخصصی (GAN^۲) برای غنی‌سازی اطلاعات کم از افراد جدید می‌تواند الهام بخش سبک فرد جدید باشد.
- ## ۶ سپاسگزاری
- جا دارد تا از حمایت‌های مسئولین وقت پژوهشکده پردازش هوشمند علائم (پردازش داده) در پیشبرد کارهای مرتبط با این مقاله تشکر و قدردانی شود.
- ## مراجع
- [1] X.-Y. Zhang and C.-L. Liu, "Style transfer matrix learning for writer adaptation," presented at the Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 20-25 June 2011, 2011.
- [2] J. Chen, X. Liu, P. Tu, and A. Aragonés, "Learning person-specific models for facial expression and action unit recognition," *Pattern Recognition Letters*, vol. 34, no. 15, pp. 1964-1970, 2013, doi: <http://dx.doi.org/10.1016/j.patrec.2013.02.002>.
- [3] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. Matthews, "Painful data: The UNBC-McMaster shoulder pain expression archive database," in *Automatic Face & Gesture Recognition and Workshops (FG 2011)*, 2011 IEEE International Conference on, 2011: IEEE, pp. 57-64.
- [4] ا. محمدیان، ح. آ. نیا، و ف. توحیدخواه. "بازشناسی جلوه‌های هیجانی چهره مستقل از فرد مبتنی بر دانش اولیه از شخص جدید." فصلنامه مهندسی پزشکی زیستی، دوره ۶، شماره ۳، سال ۱۳۹۱.
- [5] A. Maronidis, D. Bolis, A. Tefas, and I. Pitas, "Improving subspace learning for facial expression recognition using person dependent and geometrically enriched training sets," *Neural Networks*, vol. 24, no. 8, pp. 814-823, 2011, doi: 10.1016/j.neunet.2011.05.015.
- [6] A. Mohammadian, H. Aghaeinia, and F. Towhidkhan, "Geometric and texture based facial expression recognition," presented at the in Proc. of 19th Iranian Conference on BioMedical Engineering(persian), 2012.
- [7] S. Gong, S. J. McKenna, and A. Psarrou, *Dynamic vision*. Imperial College Press, 2000.
- [8] G. Sandbach, S. Zafeiriou, M. Pantic, and D. Rueckert, "Recognition of 3D facial expression dynamics," *Image and Vision Computing*, vol. 30, no. 10, pp. 762-773, 2012. [Online]. Available: <http://doc.utwente.nl/84216/>.
- [9] A. Elgammal and C.-S. Lee, "Separating style and content on a nonlinear manifold," presented at the Computer Vision and Pattern Recognition, 2004. CVPR 2004.

¹ Manifold² Generative Adversarial Networks

pain expression archive database," *Image and Vision Computing*, vol. 30, no. 3, pp. 197-205, 2012.

- [33] Z. Hammal and J. F. Cohn, "Automatic detection of pain intensity," presented at the Proceedings of the 14th ACM international conference on Multimodal interaction, 2012.



امین محمدیان در مقطع کارشناسی مهندسی پزشکی در سال ۱۳۸۱ از دانشگاه صنعتی امیرکبیر فارغ التحصیل شده و کارشناسی ارشد خود را در گرایش بیوالکترونیک بخش مهندسی برق دانشگاه تربیت مدرس تهران در سال ۱۳۸۴ به اتمام رسانده است. وی فارغ التحصیل دکتری مهندسی پزشکی از دانشگاه صنعتی امیرکبیر است. ایشان هم‌اکنون استادیار پژوهشی پژوهشگاه توسعه فناوری‌های پیشرفته هستند. زمینه‌های تحقیقاتی ایشان عبارتند از: رایانش عاطفی، یادگیری ماشین و علم داده.



حسن آقایی‌نیا کارشناسی و کارشناسی ارشد خود را به ترتیب در سال‌های ۱۳۶۵ و ۱۳۶۷ در رشته مهندسی برق گرایش الکترونیک از دانشگاه صنعتی امیرکبیر اخذ کرده است. ایشان همچنین در رشته مهندسی برق گرایش الکترونیک در سال ۱۳۷۱ مدرک D.E.A و در سال ۱۳۷۵ دکترا خود را از دانشگاه والنسین فرانسه (UVHC) دریافت کردند. ایشان هم‌اکنون دانشیار دانشکده مهندسی برق دانشگاه صنعتی امیرکبیر هستند. ایشان همچنین عضو هیأت تحریریه مجله علمی و پژوهشی IJICTR هستند. زمینه‌های تحقیقاتی ایشان عبارتند از: سیستم‌های مخابرات دیجیتال و پردازش سیگنال با محوریت پردازش تصویر.



فرزاد توحیدخواه مدارک کارشناسی و کارشناسی ارشد خود را به ترتیب در سال‌های ۱۳۶۴ و ۱۳۶۸ در رشته مهندسی برق از دانشگاه صنعتی امیرکبیر اخذ کرده و سپس برای طی دوره دکتری عازم کشور کانادا شد. ایشان در سال ۱۳۷۵ مدرک دکترای خود را در رشته مهندسی پزشکی از دانشگاه ساسکاچوان دریافت کرد. ایشان هم‌اکنون استاد تمام و عضو هیأت علمی دانشکده مهندسی پزشکی دانشگاه صنعتی امیرکبیر بوده و به مدت ۱۱ سال ریاست این دانشکده را به عهده داشته است. ایشان همچنین سردبیر فصلنامه علمی-پژوهشی مهندسی پزشکی، عضو هیأت تحریریه مجله امیرکبیر دانشگاه صنعتی امیرکبیر و عضو کمیته مهندسی پزشکی شورای عالی برنامه‌ریزی وزارت علوم، تحقیقات و فناوری هستند. زمینه‌های تحقیقاتی ایشان عبارتند از: کنترل حرکات انسان، مدل‌سازی و کنترل سیستم‌های بیولوژیکی، و مهندسی سیستم‌های عصبی.

Processing Letters, vol. 40, no. 2, pp. 191-209, 2014/10/01 2014, doi: 10.1007/s11063-013-9322-9.

- [22] H. Zhao, J. Zheng, Y. Wang, X. Yuan, and Y. Li, "Portrait style transfer using deep convolutional neural networks and facial segmentation," *Computers & Electrical Engineering*, vol. 85, p. 106655, 2020/07/01/2020, doi: <https://doi.org/10.1016/j.compeleceng.2020.106655>.
- [23] Z. Xu, M. Wilber, C. Fang, A. Hertzmann, and H. Jin, "Adversarial training for fast arbitrary style transfer," *Computers & Graphics*, vol. 87, pp. 1-11, 2020/04/01/2020, doi: <https://doi.org/10.1016/j.cag.2020.01.001>.
- [24] H. Fujita, T. Matsukawa, and E. Suzuki, "One-class Selective Transfer Machine for Personalized Anomalous Facial Expression Detection," in *VISIGRAPP (5: VISAPP)*, 2018, pp. 274-283.
- [25] A. Mohammadian, H. Aghaeinia, and F. Towhidkhal, "Diverse videos synthesis using manifold-based parametric motion model for facial understanding," *IET Image Processing*, vol. 10, no. 4, pp. 253-260. [Online]. Available: <https://digital-library.theiet.org/content/journals/10.1049/iet-ipr.2014.0905>
- [26] X.-Y. Zhang and C.-L. Liu, "Writer adaptation with style transfer mapping," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 7, pp. 1773-1787, 2013.
- [27] ا. محمدیان، ح. آقایی‌نیا، و ف. توحیدخواه، "بهبود تشخیص جلوه‌های کنشی درد در چهره توسط انتقال سبک،" مجله ماشین بینایی و پردازش تصویر، دوره ۲، شماره ۱، سال ۱۳۹۴.
- [28] A. Mohammadian, H. Aghaeinia, F. Towhidkhal, and S. z. seyedsalehi, "Subject adaptation using selective style transfer mapping for detection of facial action units," *Expert Systems with Applications*, vol. 56, pp. 282-290, 2016/09/01/2016, doi: <https://doi.org/10.1016/j.eswa.2016.03.023>.
- [29] A. Mohammadian, H. Aghaeinia, and F. Towhidkhal, "Video-based facial expression recognition by removing the style variations," *IET Image Processing*, vol. 9, no. 7, pp. 596-603. [Online]. Available: <https://digital-library.theiet.org/content/journals/10.1049/iet-ipr.2013.0697>
- [30] R. A. Khan, A. Meyer, H. Konik, and S. Bouakaz, "Pain detection through shape and appearance features," in *Multimedia and Expo (ICME), 2013 IEEE International Conference on*, 2013: IEEE, pp. 1-6.
- [31] A. B. Ashraf et al., "The painful face—pain expression recognition using active appearance models," *Image and Vision Computing*, vol. 27, no. 12, pp. 1788-1796, 2009.
- [32] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, S. Chew, and I. Matthews, "Painful monitoring: Automatic pain monitoring using the UNBC-McMaster shoulder