

# بازشناسی اعمال در ورزش کشتی با استفاده از هیستوگرام نقاط گراف حاصل از اسکلت سایه‌نما

محسن سریانی<sup>۱</sup>، دانشیار؛ علی متقی<sup>۲</sup>، دانشجوی دکتری؛ حمید سیفی<sup>۳</sup> استادیار

۱- دانشکده مهندسی کامپیوتر- دانشگاه علم و صنعت ایران- تهران- ایران، soryani@iust.ac.ir

۲- دانشکده مهندسی کامپیوتر- دانشگاه علم و صنعت ایران- تهران- ایران، mottaqi@iust.ac.ir

۳- گروه تربیت‌بدنی- دانشگاه علم و صنعت ایران- تهران- ایران، seifi@iust.ac.ir

**چکیده:** بازشناسی اعمال و رفتار انسان کاربردهای فراوانی در حوزه پردازش تصویر دارد و سال‌هاست که پژوهشگران و علاقه‌مندان در این حوزه کار می‌کنند. یکی از خلأهای تحقیقاتی موجود در این زمینه بازشناسی اعمال ورزشی دوفره است. در این پژوهش اعمال ورزشکاران در کشتی آزاد انتخاب شده و با تهیه یک مجموعه داده از فنون استفاده شده در کشتی، الگوریتمی برای شناسایی این فنون توسعه داده شده و آزمایش‌هایی پیاده‌سازی گردیده است. برای استخراج ویژگی از گراف به دست آمده از اسکلت افراد استفاده شده و با تهیه هیستوگرام دوبعدی قطبی از نقاط گراف در هر فریم و ترکیب آن‌ها در کل نمونه ویدیو به روش‌های مختلف، بردار ویژگی نهایی به دست می‌آید. برای دسته‌بندی از دو شیوه KNN و SVM به‌طور مستقل استفاده شده و نتایج به دست آمده از ۳ دسته ترکیب متفاوت ویژگی‌های مکانی برای هر یک از دو دسته‌بند، مورد مقایسه و تحلیل قرار گرفته است. بالاترین دقت شناسایی برای دسته‌بند KNN حدود ۹۰ درصد به دست آمده است.

**کلمات کلیدی:** بازشناسی عمل، کشتی آزاد، سایه‌نما، ویژگی‌های اسکلتی، گراف، دسته‌بندی، ماشین بردار پشتیبان، k-نزدیک‌ترین همسایه

## Action recognition in free style wrestling using histogram of graph vertices from silhouette skeletons

M. Soryani<sup>1</sup>, Associate Professor; A. Mottaqi<sup>2</sup>, PhD. Student; H. Seifi<sup>3</sup>, Assistant Professor

1- Faculty of Computer Engineering, Iran University of Science and Technology, Tehran, Iran, Email: soryani@iust.ac.ir

2- Faculty of Computer Engineering, Iran University of Science and Technology, Tehran, Iran, Email: mottaqi@iust.ac.ir

3- Department of Physical Education University of Science and Technology, Tehran, Iran, Email: seifi@iust.ac.ir

**Abstract:** Human Action and behavior recognition have many applications in computer vision and researchers have been working on this area for many years. Two-player sport action recognition is one of the research gaps in this scope. In this research, free style wrestling actions have been considered and by providing a dataset, an algorithm was developed to recognize such actions and different experiments were implemented. The free graph produced from player's skeletons is used for feature extraction. In each frame, a feature vector is built using 2-dimensional polar histogram of the graph points and by different combination of these vectors the final feature vector is produced for a video sample. Two classifiers; SVM and KNN were used independently to classify the actions based on different feature vector combinations. The highest score for action recognition is around 90% when KNN is used.

**Keywords:** Action recognition, wrestling, silhouette, skeleton features, graph, classification, Support Vector Machine, K-Nearest Neighbor.

تاریخ ارسال مقاله: ۱۳۹۶/۰۳/۰۴

تاریخ اصلاح مقاله: ۱۳۹۷/۰۳/۱۱

تاریخ پذیرش مقاله: ۱۳۹۷/۰۵/۱۷

نام نویسنده مسئول: محسن سریانی

نشانی نویسنده مسئول: ایران - تهران - خیابان هنگام - دانشگاه علم و صنعت - دانشکده مهندسی کامپیوتر.

## ۱- مقدمه

هیستوگرام و بردار ویژگی در هر فریم، این بردارها برای ویدیوی نمونه ورودی در بعد زمان ترکیب شده تا بردار ویژگی نهایی به دست آید. در تولید بردارهای اولیه و ترکیب آن‌ها، از چند روش با پارامترهای متفاوت استفاده شده و برای دسته‌بندی، دو دسته‌بند ماشین بردار پشتیبان (SVM) و ک-نزدیک‌ترین همسایه (KNN) به‌طور مستقل به کار گرفته شده است.

در ادامه این مقاله، در بخش ۲ به کارهای مرتبط پرداخته می‌شود. در بخش ۳ روش کار پیشنهادی ارائه می‌گردد. بخش ۴ آزمایش‌ها و تحلیل نتایج به دست آمده را دربر دارد و بخش نتیجه‌گیری پایان‌بخش مقاله است.

## ۲- کارهای مرتبط

پژوهش‌های زیادی با رویکردهای گوناگون در زمینه بازشناسی اعمال انسان انجام گرفته است. دسته‌ای از پژوهش‌ها اعمال ساده انفرادی مانند راه رفتن، دویدن و افتادن را در نظر گرفته‌اند [۱، ۲]. در مواردی نیز رفتارها و رویدادهای پیچیده مانند تعویض چرخ اتومبیل و یا مراسم جشن را بررسی کرده‌اند [۳، ۴]. همچنین در تحقیقات انجام شده در این زمینه، پژوهشگران از ویژگی‌های گوناگونی مانند  $SIFT$ ،  $HOG$ ،  $HOF$  برای دسته‌بندی اعمال و رویدادهای ویدیو بهره گرفته‌اند. ویژگی‌های اسکلتی یکی از انواع ویژگی‌هاست که کم‌وبیش مورد توجه قرار گرفته است. در سال ۱۹۷۵، جانسون با آزمایشی نشان داد که انسان‌ها می‌توانند اعمال را با شواهد بسیار کمی تشخیص دهند. برای این کار از شخصی با چراغ‌های متصل به مفاصل اصلی بدن که در اتاق تاریک قدم می‌زد فیلمی تهیه کرد [۵]. در کار دیگری Fujiyoshi و همکاران، پیشنهاد کردند که از اسکلت ستاره‌ای استخراج شده از سایه‌نما برای تحلیل حرکت استفاده شود [۶]. در [۷] اسکلت بدن انسان با ۱۰ نقطه مفصلی شامل سر، گردن، آرنج و مچ دودست، زانو و مچ دویا مدل‌سازی شده است. زاویه این ۱۰ نقطه نسبت به مرکز بدن بعلاوه نسبت طول به پهنای اسکلت به‌عنوان ویژگی استفاده شده است. در این کار روش دسته‌بندی شبکه بیز برای سه عمل دویدن، راه رفتن و پریدن مورد استفاده قرار گرفته است. Yu و همکاران در [۸] از نقاط انتهایی بدن برای تشخیص عمل بالا رفتن از حفاظ توری استفاده کردند. Zia و همکاران از زاویه مفاصل بدن که از تصاویر RGB استخراج شده بود، برای دسته‌بندی هشت عمل ساده استفاده کردند [۹].

امروزه یکی از کاربردهای رایج بنیایی ماشین، تشخیص و بازشناسی اعمال و رفتار انسان در ویدیو است. کاربردهای این حوزه شامل سامانه‌های نظارتی، تحلیل ویدیو، رباتیک و سامانه‌های تعاملی انسان و کامپیوتر می‌شود. پژوهش‌هایی که در این زمینه‌ها مورد توجه قرار گرفته بسیار گسترده هستند و از تشخیص حالت انسان گرفته تا بازشناسی اعمالی مانند افتادن و رفتارهای پیچیده‌ای مانند تعویض چرخ خودرو را دربر می‌گیرند. یکی از خلأهای تحقیقاتی که در این مورد وجود دارد، اعمال ورزشی دونفره مانند کشتی، کاراته و جودو است.

طبق آخرین اطلاعات به دست آمده هنوز کار جدی در این زمینه و به‌ویژه در مورد کشتی انجام نگرفته است. براین اساس بازشناسی اعمال و فنون کشتی آزاد به‌عنوان اولین پژوهش در این زمینه انتخاب و برای انجام آزمایش‌ها نیز از مجموعه داده‌ای که برای این پژوهش تهیه گردید، استفاده شده است. شناسایی خودکار فنون کشتی می‌تواند در کاربردهایی مثل خلاصه‌سازی ویدیو جهت پخش تلویزیونی، آموزش کشتی و کمک به داوری مورد استفاده قرار گیرد.

در زمینه بازشناسی اعمال رویکردهای متعددی وجود دارد و این رویکردها طیف وسیعی از ویژگی‌ها، روش‌های استخراج ویژگی و شیوه‌های دسته‌بندی را دربر می‌گیرند. یکی از انواع ویژگی‌ها که مورد توجه قرار گرفته، ویژگی‌های استخراج شده از اسکلت بدن انسان است. در این پژوهش شناسایی حرکات انسان بر اساس گراف به دست آمده از نقاط اتصال اسکلت سایه‌نما انجام می‌گیرد. به این ترتیب که ابتدا اسکلت سایه‌نمای پیش‌زمینه، به دست می‌آید و سپس اسکلت به دست آمده تبدیل به گراف می‌گردد و از این گراف برای استخراج ویژگی استفاده می‌شود. هیچ محدودیتی در مورد ساختار سایه‌نما و به تبع آن اسکلت و گراف حاصل وجود ندارد، یعنی اینکه ساختار گراف‌ها از نظر تعداد رأس‌ها، تعداد یال‌ها و اتصالات می‌تواند از فریمی به فریم دیگر تغییر کند. بنابراین برخلاف بسیاری از کارهای مشابه که مفاصل و اعضاء بدن را برای ایجاد گراف استفاده می‌کنند، نیازی به ردگیری و مکان‌یابی دقیق اعضای بدن نیست.

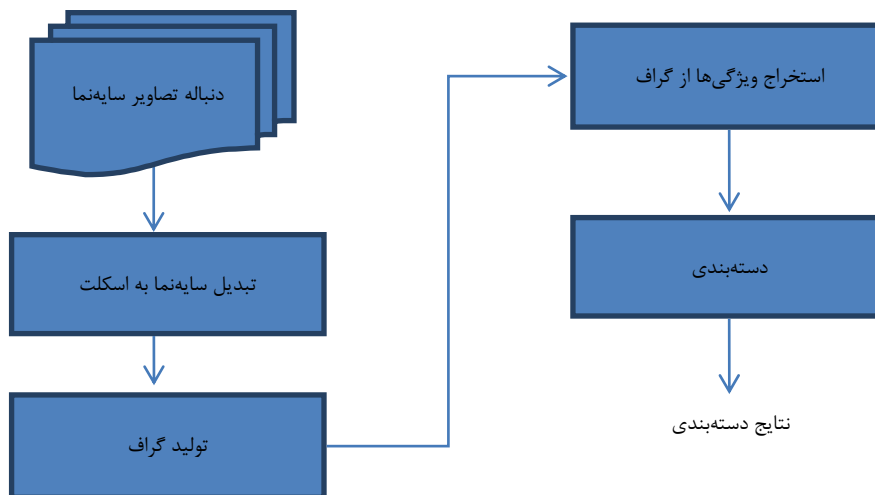
ویژگی‌های خام اولیه بر اساس مکان رأس‌های گراف مربوطه به دست می‌آیند. از آنجاکه ساختار گراف متغیر است، برای تولید ویژگی‌های با طول یکسان از هیستوگرام نقاط استفاده می‌شود. پس از به دست آوردن

### ۳- روش کار

دو مرحله اصلی در بازشناسی اعمال، تولید ویژگی‌های و دسته‌بندی است. در این پژوهش از دو الگوریتم دسته‌بندی موجود ماشین بردار پشتیبان و ک-نزدیک‌ترین همسایه استفاده می‌کنیم و تمرکز کار را روی تولید ویژگی‌ها و اجرای آزمایش‌ها روی داده‌های چالشی کشتی قرار می‌دهیم. مراحل اصلی روش پیشنهادی در شکل ۱ و چهار نمونه از تصاویر کشتی تهیه‌شده از مجموعه داده در شکل ۲ نشان داده شده‌است. داده‌های ورودی، دنباله ویدیوهای کشتی هستند که پیش‌زمینه آن‌ها جدا شده‌است. هر دنباله ورودی متناظر با یک عمل یا فن انتخاب‌شده در کشتی آزاد است.

ویژگی‌های موردنیاز از اسکلت‌های تولیدشده از سایه‌نما استخراج می‌گردند. ابتدا در هر فریم از دنباله، اسکلت سایه‌نما تولید و گراف متناظر با آن به‌دست می‌آید. در مرحله بعد هیستوگرام رأس‌های گراف در هر فریم به‌دست‌آمده و سپس هیستوگرام‌های فریم‌های دنباله در بعد زمان به‌صورت نرمال شده باهم ترکیب می‌شوند و بردار ویژگی نهایی تولید می‌گردد. در پایان با استفاده از دو روش دسته‌بندی مستقل، دسته مربوط به فن مورد بررسی تعیین می‌گردد. در دسته‌بندی به روش KNN، هر نمونه از داده‌ها با کلیه نمونه‌های دیگر مقایسه می‌گردد و از بین k نزدیک‌ترین همسایه نمونه جاری، با رأی اکثریت کلاس نمونه انتخاب می‌شود.

Xia و همکاران در [۱۰] روشی را برای بازشناسی اعمال با هیستوگرام سه‌بعدی مکان مفاصل ارائه کردند. آن‌ها برای به‌دست آوردن مکان مفاصل در هر فریم، دوربین عمق کینکت<sup>۱</sup> را به‌کار بردند. در این پژوهش برای نمایش حالت از ۱۲ مفصل استفاده می‌گردد و با استفاده از تابع وزن گاوسی، نقاط مفصلی واقع در بین‌های مجاور نیز در محاسبه فراوانی هر بین در نظر گرفته می‌شود تا روش ارائه‌شده مقاومت بیشتری در برابر تغییرات جزئی حالت داشته باشد. در [۱۱] اعمال انسان به‌صورت منحنی‌هایی در جبر لی<sup>۲</sup> نمایش داده می‌شود. آن‌ها از سه مجموعه‌داده تهیه‌شده با دوربین‌های عمق کینکت که موقعیت مفصل‌های مشخصی از اسکلت انسان را در فضای سه‌بعدی دارا هستند، استفاده کردند. در این روش با نگاشت موقعیت نسبی اعضا در هر اسکلت به یک نقطه از فضای جبر لی، منحنی‌های اعمال به‌دست می‌آیند. در کار دیگری Oflin و همکاران نیز از مفاصل اسکلت‌های سه‌بعدی برای استخراج ویژگی از سه مجموعه داده آماده استفاده کردند [۱۲]. آن‌ها با تقسیم هر دنباله اسکلت متناظر با اعمال به چند قطعه زمانی، سری زمانی مفاصل را در هر قطعه به‌دست آورده و بر اساس واریانس سری‌ها، مفاصل با بیشترین اطلاعات را برای تولید ویژگی‌های نهایی مشخص و استفاده کردند. امینیان و همکاران برای بازشناسی اعمال انفرادی در ویدیوهای چنددیدی، از گراف وضعیت بدن<sup>۳</sup> برای استخراج ویژگی استفاده کردن [۱۳]. آن‌ها iang و همکاران در [۱۴] طرحی را برای بازشناسی اعمال بر پایه اسکلت ارائه کردند. در این کار با استفاده از آنترپی مکان مفاصل، سهم هر مفصل در هر عمل را تحلیل کرده و توزیع مکانی دوبه‌دوی مفاصل مؤثر را تحت عنوان بافتار اسکلت<sup>۴</sup> به‌دست آوردند.



شکل ۱: مراحل اصلی روش پیشنهادی

رأس‌های این گراف هستند که متناظر با نقاط انتهایی و اتصالات اسکلت می‌باشند و بین هر دو رأسی از گراف که نقاط متناظر آن‌ها در اسکلت به همدیگر متصل هستند یک یال در نظر گرفته می‌شود. مجموعه یال‌ها با  $E \in V^2$  نمایش داده می‌شود. در بیان حالت و تشخیص عملی که بازیکنان انجام می‌دهند، علاوه بر اطلاعات ساختاری اسکلت، اطلاعات مکانی اعضا و مفاصل نیز مهم هستند. بنابراین در کنار ساختار گراف متناظر با اسکلت، مختصات مکانی رأس‌های گراف در دستگاه مختصات دکارتی در مجموعه  $V$  قرار می‌گیرند.

### ۳-۲-۱- استخراج ویژگی

مبنای ویژگی‌های استخراج‌شده، گراف‌های به‌دست‌آمده از اسکلت در فریم‌های متوالی هستند. ابتدا با محاسبه هیستوگرام نقاط در نواحی افراز شده در فریم  $t$ ، بردار اولیه ویژگی در زمان  $t$  به‌طور مستقل به‌دست می‌آید که ویژگی‌های مکانی فریم  $t$  را داراست. سپس این بردارها به‌گونه‌ای باهم ترکیب می‌شوند که هم اطلاعات زمانی به‌صورت نسبی حفظ شود و هم یک بردار ویژگی با طول یکسان برای تمام نمونه‌ها به‌دست آید.

### ۳-۲-۱-۱- استخراج ویژگی‌های مکانی

پس از ایجاد اسکلت پیش‌زمینه و تبدیل آن به گراف، ابتدا دستگاه

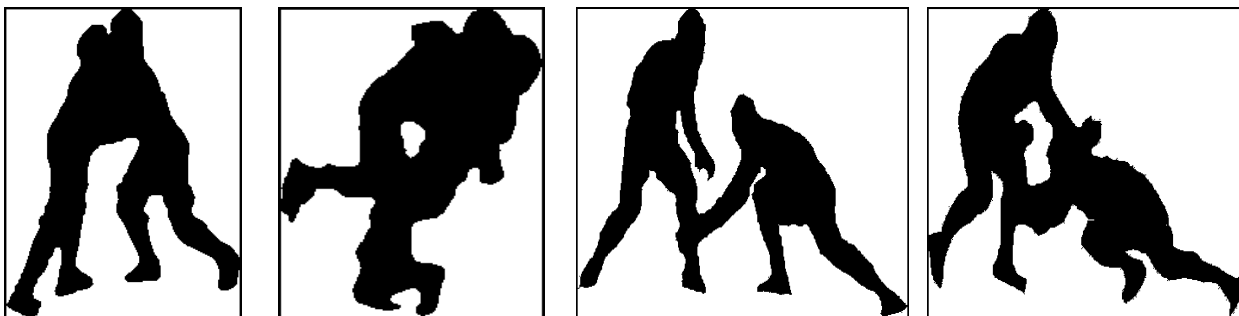
SVM یک دسته‌بند قوی و پرکاربرد است که با تغییر فضای ویژگی‌ها سعی در جداسازی خطی کلاس‌ها دارد.

### ۳-۱- ایجاد اسکلت و تولید گراف

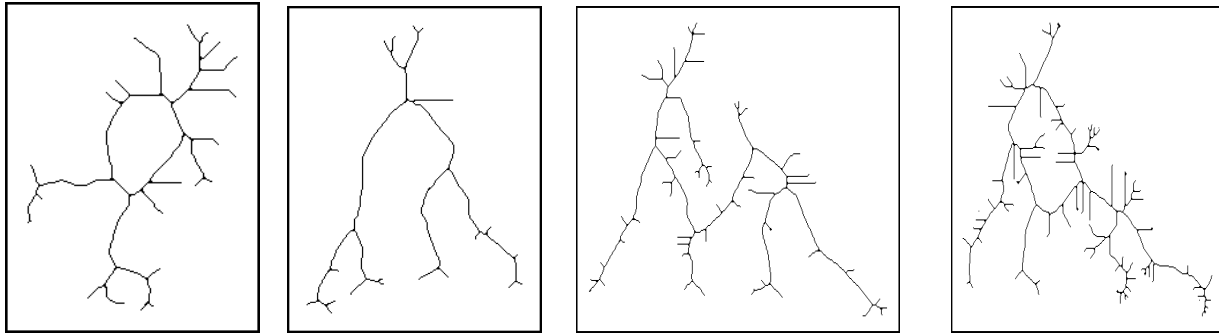
پس از تولید سایه‌نما، اسکلت پیش‌زمینه با بکارگیری عملیات مورفولوژی روی سایه‌نما استخراج می‌شود. در شکل ۳ چند نمونه از سایه‌نماها و در شکل ۴ اسکلت‌های متناظر با آن‌ها نشان داده شده است. در ویدیوهای مورد استفاده به دلیل پیچیده بودن اعمال کشتی، تفکیک اعضا کشتی گیران از یک دید و با دقت قابل قبول عملاً غیرممکن می‌باشد. با تولید اسکلت و تبدیل آن به گراف آزاد، یعنی گرافی که لزوماً متناظر با اعضا و مفاصل بدن نیست، با این چالش مقابله شد. اسکلت به‌دست‌آمده به‌گونه‌ای است که تا حد زیادی زائده‌های ریز موجود در سایه‌نما در تولید آن بی‌تأثیر هستند. با توجه به اینکه این زائده‌ها، چه نویز باشند و چه بخشی از بدن انسان، معمولاً تأثیری در قدرت بازشناسی اعمال ندارند و حتی ممکن است منجر به افزونگی داده‌ها و تغییرات داخل کلاسی ناخواسته شوند، این موضوع می‌تواند باعث افزایش کارایی سیستم شود. در اسکلت حاصل از سایه‌نما اطلاعات مفیدی از اعضا و مفاصل بدن وجود دارد. برای استخراج این اطلاعات اسکلت به‌دست‌آمده در مرحله قبل با استفاده از روش ارائه شده در [۱۵] و [۱۶] و با کمک پیاده‌سازی انجام‌شده در [۱۷] تبدیل به گراف  $G(V,E)$  با  $n$  رأس می‌گردد.  $V \in R^{n*2}$



شکل ۲: نمونه‌هایی از تصاویر کشتی مجموعه داده



شکل ۳: چهار نمونه سایه‌نمای تولیدشده از فریم‌های شکل ۲



شکل ۴: اسکلت‌های به‌دست‌آمده از سایه‌نماهای شکل ۳

در رابطه (۳) نماد  $1^T$  یک بردار  $n$  تایی با مؤلفه‌های برابر با ۱ می‌باشد. اکنون در فریم زمان  $t$ ، گراف  $G_t(V_t, E_t)$  موجود است که از اسکلت دو بازیکن به‌دست‌آمده است. با توجه به ساختار بدن انسان، حرکات تعاملی افراد و استفاده از گراف آزاد تعداد رأس‌های گراف ( $n$ )، در فریم‌ها متفاوت خواهد بود. در آزمایش‌هایی که انجام گرفت نیز به‌روشنی این واقعیت دیده‌شده و تعداد رأس‌ها از ۱۵ تا ۳۵ تغییر می‌کند. این تغییر به دلایلی از جمله جابجایی دو فرد و همچنین اعضای آن‌ها نسبت به همدیگر، چرخش بازیکنان و انسداد آلت. اگر از همه این نقاط در کلیه فریم‌ها استفاده شود، دو مشکل بروز می‌کند. اولاً ابعاد بردار ویژگی در نمونه‌های گوناگون تغییر می‌کند؛ ثانیاً طول این بردارها بزرگ خواهد بود. برای رفع هم‌زمان این دو مشکل هیستوگرام دوبعدی قطبی آله کار گرفته شد.

در هیستوگرام دوبعدی قطبی، افراز ناحیه به چند بخش بر اساس مختصات قطبی در آن ناحیه انجام می‌شود. برای این کار ناحیه موردنظر به‌صورت شعاعی و زاویه‌ای بخش‌بندی می‌گردد. بخش‌های به‌دست‌آمده از تقاطع خطوط، به‌عنوان بین یا نواحی افراز در نظر گرفته‌شده و مجموع نقاط در هریک از آن‌ها به‌دست می‌آید. پس از این مرحله، مختصات رئوس گراف در دستگاه مختصات قطبی محاسبه می‌شود. اگر  $V_D = \{(x_i, y_i)\}$  مطابق رابطه (۳) مجموعه نقاط در مختصات دکارتی و  $V_P = \{(r_i, \theta_i)\}$  مجموعه نقاط در دستگاه مختصات قطبی باشد که در آن‌ها  $1 \leq i \leq n$  و عدد کل رأس‌هاست، مختصات قطبی رأس‌ها با استفاده از رابطه (۴) به‌دست می‌آید.

$$r_i = \sqrt{x_i^2 + y_i^2} \quad \theta_i = \arctg \frac{y_i}{x_i} \quad (4)$$

با توجه به اینکه مکان مبدأ مختصات منطبق بر نقطه پایین سمت چپ سایه‌نمای پیش‌زمینه است، تمامی نقاط گراف اسکلت در ناحیه اول دستگاه مختصات قرار می‌گیرند.

مختصات جدید بر اساس مستطیل محیطی سایه‌نمای متناظر با گراف جاری تعریف می‌شود. علت استفاده از سایه‌نما به‌جای اسکلت این است که سایه‌نما دقیقاً منطبق بر پیش‌زمینه تصویر مربوطه می‌باشد، بنابراین انتظار می‌رود مختصات گراف مقاومت بیشتری در برابر تغییرات جزئی اسکلت داشته باشد. این دستگاه منطبق بر دستگاه مختصات دکارتی رایج است که مبدأ آن گوشه پایین سمت چپ مستطیل تعریف‌شده است. دستگاه مختصات جدید نسبت به مختصات صفحه تصویر دارای یک جابجایی و یک دوران حول محور  $y$  هاست. ابتدا با استفاده از رابطه ۱ جابجایی انجام می‌شود:

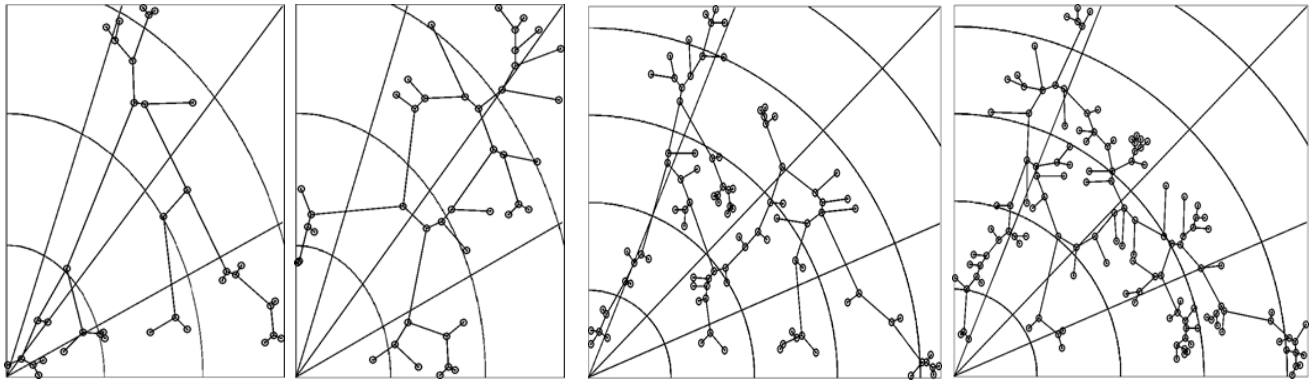
$$(x', y') = (x, y) - (x_0, y_0) \quad (1)$$

در این رابطه  $(x_0, y_0)$  مبدأ دستگاه مختصات جدید در دستگاه مختصات صفحه تصویر،  $(x, y)$  مختصات نقطه مورد نظر و عضوی از مجموعه  $n$  تایی  $V$  در صفحه تصویر و  $(x', y')$  مختصات آن نقطه در دستگاه جدید است. پس از جابجایی، دوران حول محور  $x$  با منفی کردن مؤلفه عمودی ( $y$ ) به‌دست می‌آید:

$$(x'', y) = (x', -y') = (x - x_0, y_0 - y) \quad (2)$$

اگر  $V_D \in N^{2 \times 2}$  مختصات نقاط گراف در صفحه تصویر و  $V_I \in N^{2 \times 2}$  مختصات نقاط گراف در مختصات دکارتی باشد، عملیات تبدیل گفته‌شده با نمایش ماتریسی رابطه (۳) انجام می‌گیرد:

$$V_D = [V_I \quad 1^T] \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ -x_0 & y_0 \end{bmatrix} \quad (3)$$



شکل ۵: افزاز ناحیه پیش‌زمینه در مختصات قطبی به صورت شعاعی و زاویه‌ای و گراف‌های متناظر با شکل ۴

$$n_b = n_a * n_c \quad (۷)$$

$$b_i = [r_i/d_r] * n_\theta + [\theta_i/d_\theta] + 1 \quad (۸)$$

$$V_t = [V_t^1, V_t^2, \dots, V_t^{n_b}] \quad (۹)$$

در رابطه (۹)، تعداد  $V_t^k$  های واقع در بردار B است که در آن  $b_{k=i}$  در آن می‌باشد.  $1 \leq k \leq n$  و  $1 \leq i \leq m_b$

### ۲-۲-۳- ترکیب ویژگی فریم‌ها در بعد زمان

پس از به‌دست آوردن بردار ویژگی در هر فریم، با ترکیب آن‌ها بردار ویژگی نهایی به‌دست می‌آید. با توجه به روش‌های استفاده‌شده برای دسته‌بندی، طول بردارهای ویژگی در کلیه نمونه‌ها باید برابر و مستقل از طول نمونه‌ها باشد. فرض کنید تمام بردارهای اولیه  $V_t$  که از فریم‌های نمونه به‌دست آمده‌اند، پشت سر هم قرار می‌گیرند تا بردار تجمیع شده  $A$  مطابق با رابطه (۱۰) به‌دست آید. در این رابطه T تعداد فریم در دنباله است و طول نمونه را در بعد زمان مشخص می‌کند. برای ترکیب بردارهای اولیه، روی دنباله فریم‌های هر نمونه و در بعد زمان، تعداد  $n_w$  پنجره با طول L تعریف می‌شود، به‌این‌ترتیب که پنجره‌ها پشت سر هم و بدون همپوشانی بوده و هر پنجره L فریم پشت سرهم را دارا باشد. داخل هر پنجره، بردارهای اولیه باهم جمع می‌شوند و بردارهای حاصل از پنجره‌ها با استفاده از روش الحاق باهم ترکیب می‌گردند تا بردار ویژگی نهایی  $F$  به طول  $n_b * n_w$  مطابق رابطه (۱۱) یا (۱۲) به‌دست آید:

$$F = [1 \leq t \leq N_w \left( \sum_{t=W_i}^{W_{i+1}-1} V_t \right)] \quad (۱۱)$$

$$w_i = (i - 1) * L + 1 \quad \text{که در آن:}$$

بنابراین این ناحیه را به  $n_a$  بازه زاویه‌ای یکسان و  $n_c$  دایره متوالی با تفاضل شعاعی ثابت و به مرکز مبدأ مختصات تقسیم می‌کنیم و فراوانی رئوس گراف در هر ناحیه را محاسبه می‌کنیم. بزرگ‌ترین شعاع افزاز برابر با فاصله نرمال شده گوشه بالای سمت راست مستطیل محیطی از مبدأ مختصات بوده و R نامیده می‌شود (رابطه (۵)). اندازه تفاضل زاویه‌ای هر بین (ناحیه)  $d_\theta$  و تفاضل شعاعی دایره‌های متوالی یعنی  $d_r$  از رابطه (۶) به‌دست می‌آیند:

$$R = \sqrt{x_{\max}^2 + y_{\max}^2} \quad (۵)$$

$$d_r = R/n_c d_\theta = (\frac{\pi}{2})/n_a \quad (۶)$$

تعداد کل نواحی به‌دست‌آمده از افزاز یعنی  $m_b$  نیز از رابطه (۷) به‌دست می‌آید. به‌این‌ترتیب با شمارش تعداد رأس‌های موجود در هر ناحیه، یک بردار  $m_b$  تایی از اعداد ایجاد می‌شود که ویژگی‌های مکانی فریم جاری را تشکیل می‌دهد. نام این بردار را  $V_t$  انتخاب می‌کنیم. برای ایجاد این بردار، ابتدا با محاسبه ناحیه هر رأس از گراف با استفاده از رابطه (۸)، بردار  $B = [b_1, b_2, \dots, b_n]^T$  متناظر با  $V_t$  ایجاد می‌شود که  $b_i$  شماره

$$A = [V_1, V_2, \dots, V_T] \quad (۱۰)$$

$$= [v_1^1, v_1^2, \dots, v_1^{n_b}, v_2^1, v_2^2, \dots, v_2^{n_b}, \dots, v_T^{n_b-1}, v_T^{n_b}]$$

ناحیه متعلق به نقطه  $i$  از گراف را نشان می‌دهد. اکنون با محاسبه هیستوگرام شماره نقاط واقع در بردار B، بردار  $V_t$  از رابطه (۹) به‌دست می‌آید. در این رابطه  $b_i$  بخش افزاز شده  $i$  ام و  $V_t$  برابر با دنباله هیستوگرام‌های این بخش‌ها در فریم جاری  $t$  هستند:

به بیان دیگر می‌توان نوشت:

$$F = \left[ \sum_{t=1}^L V_t, \sum_{t=L+1}^{2*L} V_t, \dots, \sum_{t=(N_w-1)*L+1}^{L*N_w} V_t \right]^T \quad (12)$$

هریک از ویدیوهای کشتی، حرکات و فنون موردنظر به صورت دستی جدا شدند به شکلی که در هر نمونه ویدیو تنها یک عمل وجود داشته باشد. طول نمونه‌ها برحسب فریم متفاوت و از ۲۲ تا ۷۴ متغیر است. در تولید این مجموعه داده و انتخاب نمونه‌ها تا حد امکان محدودیتی در مدت زمان اجرای عمل، به عبارتی تعداد فریم‌های هر نمونه اعمال نگردید؛ در عین حال سعی شده است حرکات اضافی و نامرتبط در نمونه وجود نداشته باشد و طول نمونه‌ها خیلی زیاد نباشد. به دلیل اینکه این نمونه‌ها از منابع گوناگون تهیه شده‌اند، از نظر کیفیت تصویر با هم تفاوت دارند. همچنین در تهیه مجموعه داده و انتخاب و جداسازی فنون از راهنمایی اساتید و کارشناسان کشتی استفاده شده است.

#### جدول ۱: مشخصات دنباله ویدیوهای مجموعه داده

#### ۲-۴- تولید ویژگی

طبق آنچه در بخش ۳ گفته شد، اسکلت افراد در هر فریم از نمونه ایجاد شده و گراف متناظر با آن به دست می‌آید. در تولید بردار ویژگی نهایی، چند پارامتر و چند روش ترکیب به کار رفته و با تغییر آن‌ها، چند گروه ویژگی به دست می‌آید. برای تولید ویژگی‌های مکانی مربوط به هر فریم پارامتر  $m_b$  یعنی تعداد بین‌ها در افراز گراف، از مجموعه مقادیر  $\{12, 16, 20, 24, 30, 32, 36, 40, 48, 56\}$  انتخاب گردید. هر کدام از این مقادیر از حاصل ضرب تعداد شیار در تعداد قطاع به دست می‌آیند که در شکل ۵ نمونه‌هایی از آن‌ها نشان داده شده است. در بعد زمان با استفاده از سه روش ترکیب ویژگی‌های مکانی، سه نوع بردار ویژگی مکانی-زمانی به دست می‌آید. سطرها و ستون‌های جدول ۲ بر اساس مقادیر  $m_b$  و سه روش ترکیب ویژگی‌های مکانی اولیه سازمان‌دهی شده‌اند. طول بردارها و به دنبال آن زمان دسته‌بندی، ارتباط مستقیم با این پارامترها دارد. همچنین تغییر این پارامترها روی دقت دسته‌بندی تأثیر خواهد داشت. در ترکیب ویژگی‌های مکانی در بعد زمان، از سه روش استفاده شد که در ادامه می‌آیند.

$\begin{matrix} \hat{m}_b \\ \hat{m}_t \end{matrix}$	$\begin{matrix} \hat{m}_b \\ \hat{m}_t \end{matrix}$	$\begin{matrix} \hat{m}_b \\ \hat{m}_t \end{matrix}$	$\begin{matrix} \hat{m}_b \\ \hat{m}_t \end{matrix}$
۳۳	۴۱ تا ۳۳	۱۳	فن کمر
۳۰	۴۳ تا ۲۴	۱۵	زیر یک‌خم
۲۷	۳۳ تا ۲۲	۱۵	زیر دوخم
۶۰	۷۴ تا ۲۰	۱۴	بارانداز
۴۶	۷۱ تا ۳۶	۱۲	پیچ پیچک
۳۹	۷۴ تا ۲۰	۶۹	مجموع

در روابط (۱۱) و (۱۲)،  $w_i$  شماره اولین فریم در پنجره  $i$  ام و  $F$  بردار ویژگی نهایی برای نمونه جاری است که در دسته‌بندی از آن استفاده می‌شود.

برخی مزایای شیوه استفاده شده برای تولید ویژگی عبارت‌اند از: (۱) استقلال روش استخراج ویژگی از توپولوژی گراف به دست آمده؛ به این معنی که محدودیتی در مورد ساختار اسکلت و گراف متناظر با آن وجود ندارد. (۲) عدم وابستگی طول بردار ویژگی به تعداد فریم‌های دنباله؛ با تنظیم پارامترها، طول بردار نهایی ویژگی برای همه نمونه‌ها یکسان خواهد بود. (۳) نامتغیر بودن نسبت به مقیاس؛ بردار ویژگی به ابعاد ناحیه پیش‌زمینه تصویر وابسته نیست. (۴) به هر اندازه‌ای که نیاز باشد اطلاعات زمانی اعمال استخراج و استفاده می‌شود. با افزایش تعداد پنجره‌ها یعنی  $m_w$  اطلاعات زمانی کمتری را از دست می‌دهیم. (۵) ابعاد ویژگی‌ها قابل کنترل بوده و از بزرگی بی‌رویه ابعاد جلوگیری می‌شود. ابعاد ویژگی‌های با دو پارامتر  $m_b$  و  $m_w$  رابطه مستقیم دارد. (۶) مقاومت بخش دسته‌بندی در برابر تغییرات درون کلاسی داده‌ها افزایش می‌یابد. با وجود این مزایا یک ایراد این روش آن است که مسلماً مقداری از اطلاعات زمانی اعمال از دست می‌رود. به‌ویژه اگر تعداد پنجره‌ها کم باشد این عامل می‌تواند اثر منفی روی عملکرد سامانه داشته باشد.

#### ۴- آزمایش‌ها و نتایج

روش پیشنهادی روی مجموعه داده تهیه شده از کشتی آزاد آزمایش گردید. ویژگی‌های نهایی برای دسته‌بندی با استفاده از روش ارائه شده در بخش ۳ و از گراف‌های حاصل از اسکلت استخراج شدند. آزمایش‌ها با تغییر پارامترهای مؤثر در تولید ویژگی‌ها انجام شده و بردارهای ویژگی گوناگونی به دست آمد. برای هر حالت از پارامترها، ویژگی‌های تولید شده به دو دسته  $k$ -نزدیک‌ترین همسایه و ماشین بردار پشتیبان داده شده و خروجی‌ها به دست آمدند.

#### ۱-۴- مجموعه داده

در این پژوهش نمونه‌هایی از فنون کشتی آزاد از منابع گوناگون تهیه گردید. این نمونه‌ها شامل فنون زیر یک‌خم، زیر دوخم، فن کمر، پیچ پیچک و بارانداز هستند که در جدول ۱ مشخصات آن‌ها آمده است. تقریباً نیمی از کل نمونه‌ها از کشتی‌های واقعی ملی یا بین‌المللی و بقیه با اجرای کشتی توسط سه گروه دونفره و با دوربین ثابت تهیه شدند. از

**روش ۱:** در این روش طبق آنچه در بخش ۳-۲-۲ گفته شد فریم‌های دنباله در ۸ پنجره گروه‌بندی شدند ( $m_w=8$ ). از آنجا که طول دنباله‌ها باهم متفاوت است، طول پنجره‌ها (پارامتر  $L$ )، به عبارتی تعداد فریم در هر پنجره از دنباله‌ای به دنباله دیگر متفاوت خواهد بود.

**روش ۲:** در این روش با افزودن فریم‌های خالی به انتهای هر دنباله یکسان‌سازی طول نمونه‌ها انجام می‌شود و پس از آن مانند روش ۱، فریم‌ها در ۸ پنجره زمانی گروه‌بندی می‌شوند. در این شیوه طول پنجره‌ها در تمام دنباله‌ها یکسان خواهد بود.

**روش ۳:** در این روش مانند روش ۲ یکسان‌سازی طول نمونه‌ها انجام می‌گیرد اما گروه‌بندی زمانی را انجام نمی‌دهیم؛ به عبارتی طول پنجره‌های زمانی ۱ در نظر گرفته می‌شود. طول بردارهای ویژگی در روش‌های ۱ و ۲ برای تعداد بین یکسان برابر است ولی در روش ۳ طول بردارها بزرگ‌تر و در واقع ۸ برابر طول بردارهای متناظر در روش ۱ و ۲ هستند.

برای دسته‌بندی اعمال در هر ۳ روش ترکیب ویژگی‌ها، از دو دسته‌بند SVM با هسته خطی و با روش اعتبارسنجی تقاطعی nFold با  $n=10$  و دسته‌بند KNN با  $k=5$  استفاده شده است.

#### ۱-۱-۴- تأثیر مشخصات ویژگی‌ها

در آزمایش‌های انجام‌شده سه روش ترکیب ویژگی و ۱۰ مقدار متفاوت برای تعداد بین‌ها استفاده شده که سطرها و ستون‌های اصلی جدول ۲ را تشکیل می‌دهند. برای بررسی تأثیر ویژگی‌ها روی دقت دسته‌بندی، از میانگین‌های درج‌شده در این جدول و نمودارهای شکل ۶ استفاده می‌کنیم. مطابق این جدول بالاترین دقت برای هر دو دسته‌بند با استفاده از ویژگی‌های روش ۲ به دست آمده است. دقت میانگین SVM با این ویژگی‌ها برابر  $71/9$  و برای KNN برابر  $85/2$  و دقت میانگین کل برابر  $78/5$  می‌باشد. از نظر دقت، پس از ویژگی‌های روش ۲، ویژگی‌های روش ۱ با دقت میانگین کل  $73/1$  در رتبه بعدی و ویژگی‌های روش ۳ با میانگین دقت  $62/5$  در رتبه آخر قرار گرفته‌اند. بنابراین استخراج ویژگی با یکسان‌سازی طول نمونه‌ها و گروه‌بندی فریم‌ها در بعد زمان بهترین نتیجه را دارد. علت می‌تواند این باشد که اولاً با یکسان‌سازی، مؤلفه‌های دو بردار ویژگی با موقعیت زمانی تقریباً یکسان باهم مقایسه می‌شوند؛ بعلاوه با گروه‌بندی فریم‌ها مقاومت دسته‌بند در برابر تغییرات داخل کلاسی افزایش می‌یابد.

از طرف دیگر با توجه به جدول ۲ و نمودارهای شکل ۶، به تأثیر پارامتر افزایش  $m_b$  روی دقت پی می‌بریم. شکل ۶-د نشان می‌دهد که تعداد بین‌های استفاده‌شده با میانگین دقت دسته‌بندی به روش KNN رابطه

**تأثیر مشخصات ویژگی‌ها**

نتایج به دست آمده با سه روش ترکیب ویژگی در جدول ۲ آمده است. ستون‌های این جدول به ترتیب از راست به چپ عبارت‌اند از: تعداد بین‌ها در هر فریم، دقت دسته‌بندی به صورت جدا برای سه روش ترکیب ویژگی ۱، ۲ و ۳، دقت میانگین دسته‌بندی به روش SVM و KNN و ستون آخر دقت میانگین کل. برای هر روش ترکیب ویژگی نیز دقت به روش SVM، KNN و میانگین آن‌ها نوشته شده است. مشاهده می‌شود که بالاترین

#### ۱-۴- تحلیل نتایج

**جدول ۲: دقت دسته‌بندی در حالت‌های مختلف و میانگین دقت‌ها**

تعداد بین‌ها	روش ترکیب ویژگی ۱			روش ترکیب ویژگی ۲			روش ترکیب ویژگی ۳			Total Ave
	Ave1	KNN	SVM	Ave2	KNN	SVM	Ave3	KNN	SVM	
۱۲	۷۳/۲	۷۵/۴	۱۷/۰	۷۷/۵	۸۲/۶	۷۲/۵	۷۴/۷	<u>۷۸/۳</u>	۷۱/۰	۷۵/۱۴
۱۶	<u>۷۶/۸</u>	۷۵/۴	<u>۷۸/۳</u>	<u>۸۳/۳</u>	۸۸/۴	<u>۷۸/۳</u>	<u>۷۵/۴</u>	<u>۷۵/۴</u>	۷۵/۴	<u>۷۸/۵۳</u>
۲۰	<u>۷۶/۸</u>	۷۶/۸	<u>۷۶/۸</u>	۸۲/۷	<u>۸۹/۹</u>	۷۵/۴	۶۷/۴	۶۵/۲	۶۹/۶	۷۵/۶۲
۲۴	۶۸/۸	۷۲/۵	۶۵/۲	۷۸/۳	۸۵/۵	۷۱/۰	۶۳/۱	۶۵/۲	۶۲/۳	۷۰/۲۹
۳۰	<u>۷۸/۳</u>	۷۱/۰	۷۱/۰	۷۹/۸	۸۹/۹	۶۹/۶	۶۵/۲	۵۶/۵	۷۳/۹	۷۳/۲
۳۲	۷۶/۸	۷۳/۹	۷۶/۸	۸۰/۵	۸۷/۰	۷۳/۹	۶۵/۳	۶۰/۹	۶۹/۶	۷۳/۶۸
۳۶	۶۹/۶	۷۱/۰	۶۹/۶	۷۹/۷	۸۸/۴	۷۱/۰	۶۰/۱	۵۳/۶	۶۶/۷	۷۰/۰۵
۴۰	۷۳/۹	۷۵/۴	۷۳/۹	۷۸/۳	۸۴/۱	۷۲/۵	۵۸/۷	۴۷/۸	۶۹/۶	۷۰/۵۵
۴۸	۷۳/۹	۷۳/۹	۷۳/۹	۸۰/۵	۸۴/۱	۷۶/۸	۵۸/۷	۴۶/۴	۷۱/۰	۷۱/۰۲
۵۶	۶۳/۸	۷۲/۵	۶۳/۸	۷۶/۸	۶۵/۲	۶۸/۲	۵۰/۸	۴۳/۵	۷۱/۰	۶۳/۳
Ave	۷۳/۱	۷۴/۵	۷۱/۸	۷۸/۵	<u>۸۵/۲</u>	۷۱/۹	۶۸/۰	۵۸/۹	۶۸/۰	۷۱/۷۱



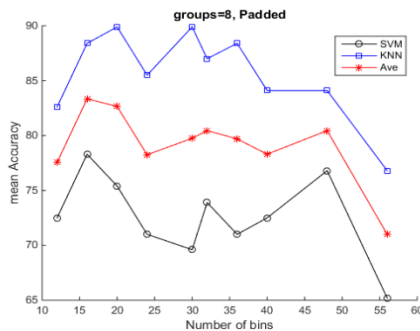
هرچند تعداد بین‌ها روی دقت دسته‌بندی به روش SVM نیز تأثیر دارد، اما رابطه روشی نمی‌توان بین آن‌ها پیدا کرد.

### ۲-۱-۴- مقایسه دسته‌بندها

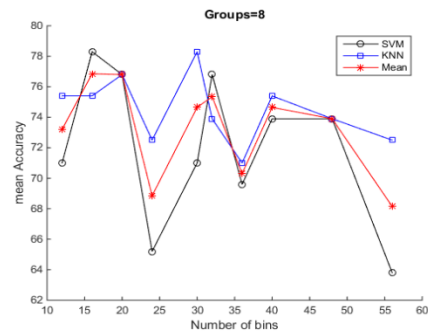
برای مقایسه دقت خروجی‌های KNN و SVM از داده‌های جدول ۲ و نمودارهای شکل ۶ استفاده می‌کنیم. در این جدول میانگین دقت حاصل از به‌کارگیری روش‌های ترکیب ویژگی ۱، ۲ و ۳ برای هر کدام از دسته‌بندها جداگانه نوشته شده است. مشاهده می‌شود که KNN با دقتی برابر ۸۹/۹ درصد بیشترین دقت را دارد؛ در حالی که بیشترین دقت SVM برابر ۷۸/۳ درصد و با اختلاف قابل توجه ۱۱/۶ درصد است. با مقایسه میانگین دقت‌ها در سه روش ترکیب ویژگی، مشاهده می‌شود که در هر سه روش، میانگین دقت KNN از SVM بیشتر است. همچنین میانگین کل دقت SVM برابر با ۷۰/۵ و برای KNN برابر با ۷۲/۸ است. با مراجعه به شکل ۶-ب نیز مشاهده می‌شود که برای تمامی مقادیر  $m_b$  در ویژگی‌های روش ۲، دقت KNN بالاتر از SVM قرار گرفته است. نتیجه دیگری که از شکل ۶-د می‌توان گرفت این است که تغییرات دقت

معکوس دارد. این ارتباط در روش ترکیب ۳ نیز در شکل ۶-ج به روشنی دیده می‌شود، اما در روش ترکیب ۲ کمتر و در روش ترکیب ۱، این ارتباط بسیار ناچیز است.

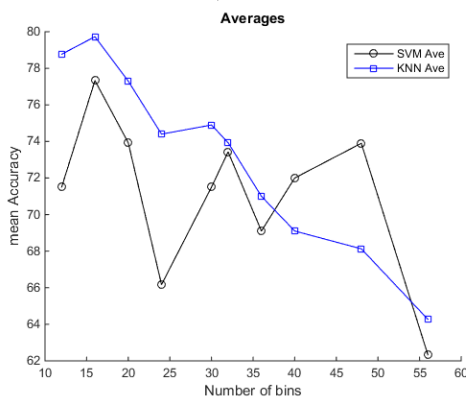
علت این که با افزایش تعداد بین‌ها دقت میانگین KNN در روش ترکیب ۳ کاهش می‌یابد، می‌تواند در ابعاد ویژگی‌ها و تعداد نقاط گراف نهفته باشد. در تولید ویژگی‌ها به روش ۳، یکسان‌سازی طول نمونه‌ها انجام شده و بدون گروه‌بندی فریم‌ها، ویژگی‌های نهایی استخراج می‌شوند؛ در نتیجه چگالی نقاط گراف در بین‌ها، به نسبت اندازه پنجره گروه‌بندی استفاده شده در روش ۱ و ۲ کاهش می‌یابد. بنابراین با افزایش تعداد بین‌ها بسیاری از ویژگی‌ها صفر و بقیه نیز اعداد صحیح کوچکی خواهند بود. نتیجه نهایی این است که با افزایش تعداد بین‌ها تغییرات داخل کلاسی افزایش و در نتیجه دقت تشخیص کاهش می‌یابد. در ویژگی‌های روش ۱ و ۲ که گروه‌بندی فریم‌ها انجام می‌گیرد، تغییرات نسبی ویژگی‌ها در نمونه‌های هم‌کلاس نسبت به تغییرات اسکلت و گراف کمتر می‌گردد؛ بنابراین تنها با افزایش یا کاهش بین‌ها دقت تفاوت زیادی نخواهد داشت.



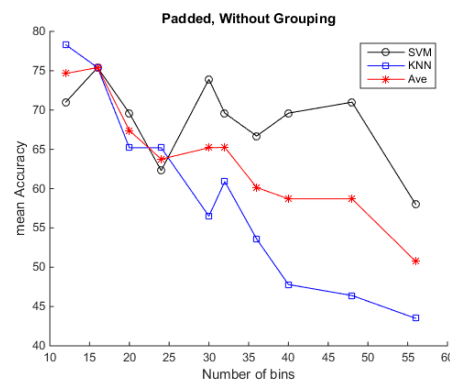
(ب)



(الف)



(د)



(ج)

شکل ۶: نمودار نتایج به‌دست آمده با روش‌های متفاوت ترکیب ویژگی‌های مکانی. الف: روش ترکیب ۱، ب: روش

ترکیب ۲، ج: روش ترکیب ۳، د: میانگین دقت هر دسته‌بند به سه روش

مناسب‌تر برای افزایش دقت استفاده کرد.

تشخیص KNN به‌طور متوسط بیشتر از SVM است؛ به‌عبارت‌دیگر SVM در برابر تغییر بین‌ها مقاوم‌تر است.

Predicted \ Actual	C1	C2	C3	C4	C5	دقت
C1	11	0	1	0	1	85
C2	2	5	7	0	1	33
C3	0	3	12	0	0	80
C4	1	0	0	13	0	93
C5	5	0	0	1	6	50

جدول (۴): ماتریس سردرگمی تولیدشده با KNN

### ۳-۱-۴- تفاوت دقت در کلاس‌ها

در جدول ۳ دقت تشخیص در تک‌تک کلاس‌ها برای روش ترکیب ویژگی ۲ و بر اساس دو الگوریتم دسته‌بندی به‌صورت جداگانه آمده است. ستون‌های این جدول به ترتیب از راست به چپ برای هر دسته‌بند عبارت‌اند از: تعداد بین‌ها در هر فریم، دقت هر کلاس (۵ مورد) و دقت میانگین همه کلاس‌ها است. در بیشتر حالات، دقت در کلاس ۴ که فن بارانداز کشتی است، بیشترین و در فنون یک‌خم و کمر کمترین است. همچنین سطر آخر جدول ۳ نشان می‌دهد که میانگین دقت این دو کلاس نیز از بقیه کمتر است. بر اساس دانش ما در مورد کشتی آزاد، اجرای بارانداز از نظر بصری تفاوت قابل‌توجهی با ۴ فن دیگر دارد، درحالی‌که فنون زیر یک‌خم و دوخم شباهت زیادی باهم دارند. دو فن کمر و پیچ‌پیچک نیز نسبت به بقیه شباهت بیشتری دارند. با مراجعه به ماتریس‌های سردرگمی جدول‌های ۴ و ۵، مشاهده می‌شود که بین دو فن یک‌خم و دوخم و همچنین بین دو فن کمر و پیچ‌پیچک بیشترین جابجایی نمونه، به‌ویژه در جدول (۴) وجود دارد. همان‌گونه که انتظار می‌رود بیشترین جابجایی‌ها مربوط به فنون مشابه است که انسان‌ها نیز آن‌ها را نزدیک به هم می‌دانند. نتایج به‌دست‌آمده از آزمایش‌ها این استدلال شهودی را تأیید کرده و از آن می‌توان در انتخاب ویژگی‌های

### ۵- نتیجه‌گیری

در این پژوهش بازشناسی اعمال ورزش کشتی به‌عنوان کاربرد جدیدی از پردازش ویدیو و با استفاده از ویژگی‌های استخراج‌شده از گراف اسکلت سایه‌نما موردبررسی قرار گرفت. ویژگی‌های مکانی فریم‌ها به ۳ روش متفاوت ترکیب‌شده و به‌طور جداگانه به دو دسته‌بند اعمال گردیدند. در انجام این تحقیق چالش‌های فراوانی از جمله نبود مجموعه داده و پیچیدگی ذاتی اجرای فنون کشتی وجود داشت. برای تهیه مجموعه داده، نمونه ویدیوهایی از منابع گوناگون تهیه‌شده و با انتخاب ۵ عمل (فن) از کشتی آزاد آزمایش‌ها انجام گرفت. برای استخراج ویژگی، اسکلت بازیکنان تبدیل به گراف شده و با افزودن گراف‌ها، هیستوگرام نقاط نواحی در هر فریم تولید گردید. سپس پنجره‌های زمانی روی فریم‌ها ایجادشده و با الحاق ویژگی‌های به‌دست‌آمده از هر پنجره، بردار ویژگی نهایی به‌دست آمد. با تغییر پارامترهای تعداد پنجره‌ها، تعداد نواحی افراز و چگونگی افراز نقاط گراف، بردارهای متفاوتی از ویژگی‌های به‌دست آمد. برای دسته‌بندی از دو الگوریتم ک-نزدیک‌ترین همسایه و ماشین بردار پشتیبان خطی استفاده شد. بهترین دقت

جدول ۳: نتایج دسته‌بندی با ویژگی‌های روش ۲ به تفکیک کلاس‌ها

Mean Acc	KNN					SVM					Bins	
	C5	C4	C3	C2	C1	Mean Acc	C5	C4	C3	C2		C1
۸۲/۶	۶۶/۷	۱۰۰	۹۳/۳	۶۰	۹۲/۳	۷۲/۵	۵۸/۳	۱۰۰	۸۰	۴۰	۸۴/۶	۱۲
۸۸/۴	۸۳/۳	۱۰۰	۱۰۰	۷۳/۳	۸۴/۶	<u>۷۸/۳</u>	۵۸/۳	۱۰۰	۹۳/۳	۵۳/۳	۸۴/۶	۱۶
<u>۸۹/۹</u>	۸۳/۳	۱۰۰	۹۳/۳	۸۰	۹۲/۳	۷۵/۴	۶۶/۷	۱۰۰	۸۶/۷	۴۰	۸۴/۶	۲۰
۸۵/۵	۷۵	۱۰۰	۹۳/۳	۶۶/۷	۹۲/۳	۷۱	۵۰	۱۰۰	۸۰	۴۰	۸۴/۶	۲۴
۸۷	۹۱/۷	۹۲/۹	۹۳/۳	۶۶/۷	۹۲/۳	۷۳/۹	۵۸/۳	۱۰۰	۸۰	۴۶/۷	۸۴/۶	۳۲
۸۴/۱	۶۶/۷	۹۲/۹	۱۰۰	۷۳/۳	۸۴/۶	۷۲/۵	۴۱/۷	۱۰۰	۸۶/۷	۴۶/۷	۸۴/۶	۴۰
<u>۸۹/۹</u>	۸۳/۳	۱۰۰	۹۳/۳	۸۰	۹۲/۳	۶۹/۶	۴۱/۷	۱۰۰	۸۰	۴۰	۸۴/۶	۳۰
۸۸/۴	۷۵	۹۲/۹	۱۰۰	۷۳/۳	۱۰۰	۷۱	۵۰	۱۰۰	۷۳/۳	۴۰	۹۲/۳	۳۶
۸۴/۱	۶۶/۷	۹۲/۹	۱۰۰	۶۶/۷	۹۲/۳	۷۶/۸	۵۰	۱۰۰	۹۳/۳	۴۶/۷	۹۲/۳	۴۸
۷۶/۸	۵۸/۳	۷۸/۶	۸۶/۷	۷۳/۳	۸۴/۶	۶۵/۲	۵۰	۸۵/۷	۶۶/۷	۴۶/۷	۷۶/۹	۵۶
۸۵/۲	۷۴/۴	۹۳/۹	<u>۹۵/۷</u>	۷۱	۹۰/۱	۷۱/۹	۵۰	<u>۹۹</u>	۸۱/۹	۴۲/۹	۸۴/۶	Class mean Acc

behavior recognition,” In *Advanced Video and Signal-Based Surveillance (AVSS), 2012 IEEE Ninth International Conference on*, pp. 136-142. IEEE, 2012.

- [4] J. K. Aggarwal and M. S. Ryoo, “Human activity analysis: A review,” *ACM Comput. Surv.*, vol. 43, no. 3, pp. 1-43, Apr. 2011.
- [5] G. Johansson, “Visual motion perception,” *Sci. Am.*, vol. 232, pp. 76-88, 1975.
- [6] H. Fujitayoshi, A. J. Lipton, and T. Kanade, “Real-time human motion analysis by image skeletonization,” *IEICE Trans. Inf. Syst.*, vol. 87, no. 1, pp. 113-120, 2004.
- [7] C. Zou and Z. Liu, “Behavior Classification Method Based on Skeleton Model from Video Images,” In *Computer Science and Information Technology, 2008. ICCSIT'08. International Conference on*, pp. 309-312. IEEE, 2008.
- [8] E. Yu and J. K. Aggarwal, “Human action recognition with extremities as semantic posture representation,” in *Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on*, pp. 1-8, 2009.
- [9] M. Z. Uddin, “Human Activity recognition using body joint-angle features and hidden markov model,” *ETRI J.*, vol. 33, no. 4, pp. 569-579, Aug. 2011.
- [10] L. Xia, C.-C. Chen, and J. K. Aggarwal, “View invariant human action recognition using histograms of 3d joints,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pp. 20-27, 2012.
- [11] R. Vemulapalli, F. Arrate, and R. Chellappa, “Human action recognition by representing 3d skeletons as points in a lie group,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 588-595, 2014.
- [12] F. Ofli, R. Chaudhry, G. Kurillo, R. Vidal, and R. Bajcsy, “Sequence of the most informative joints (SMIJ): A new representation for human skeletal action recognition,” *J. Vis. Commun. Image Represent.*, vol. 25, no. 1, pp. 24-38, Jan. 2014.

[۱۳] امیر فرید امینیان مدرس، بازشناسی و تحلیل فعالیت انسان در

تصاویر ویدیویی چند دیدی، رساله دکتری، دانشگاه علم و صنعت

ایران، ص. ۶۸-۵۳، خرداد ۱۳۹۳.

- [14] M. Jiang, J. Kong, G. Bebis, and H. Huo, “Informative joints based human action recognition using skeleton contexts,” *Signal Process. Image Commun.*, vol. 33, pp. 29-40, Apr. 2015.
- [15] X. Bai, L. J. Latecki, and W.-Y. Liu, “Skeleton pruning by contour partitioning with discrete curve evolution,” *Pattern Anal. Mach. Intell. IEEE Trans. On*, vol. 29, no. 3, pp. 449-462, 2007.
- [16] K. Liu, Y. S. Huang, and C. Y. Suen, “Identification of fork points on the skeletons of handwritten Chinese characters,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 10, pp. 1095-1100, 1999.
- [17] A. B. Iraola, “Skeletonization, skeleton pruning and simple skeleton graph construction example in Matlab,” 2009.

بازشناسی ۸۹/۹ درصد می‌باشد با دسته‌بند k-نزدیک‌ترین همسایه و

ویژگی‌ها

ای

تولیدشده

ده

جدول

۵:

ماتریس

Predicted \ Actual	C1	C2	C3	C4	C5	دقت
C1	11	1	0	0	1	85
C2	1	9	5	0	0	60
C3	0	1	14	0	0	93
C4	1	0	3	10	0	71
C5	3	0	0	1	8	67

س سردرگمی تولیدشده با SVM

با یکسان‌سازی طول نمونه‌ها و گروه‌بندی فریم‌ها به‌دست آمد که با توجه به پیچیدگی فنون و به‌عنوان اولین کار انجام‌شده دقت قابل قبولی به نظر می‌رسد. نتایج به‌دست‌آمده نشان می‌دهد که اگر تعداد بین‌ها خیلی زیاد شود، دقت بازشناسی کاهش می‌یابد. همچنین گروه‌بندی زمانی فریم‌ها باعث افزایش دقت می‌شود.

در ادامه کار تصمیم بر این است که مجموعه داده توسعه یابد، به‌گونه‌ای که علاوه بر افزایش تعداد فنون، تعداد نمونه‌ها افزایش یافته و اجرای فنون از دیدهای مختلف تحت پوشش قرار گیرد. همچنین می‌توان چگونگی افراز ناحیه پیش‌زمینه را در جهت افزایش کارایی تغییر داد.

## قدردانی

بدین‌وسیله از کارکنان اداره تربیت‌بدنی دانشگاه علم و صنعت ایران، کارشناسان و کشتی‌گیرانی که در اجرای کشتی و انتخاب و جداسازی فنون همکاری نموده‌اند، سپاسگزاری می‌شود.

## مراجع

- [1] G. Chéron, I. Laptev, and C. Schmid, “P-cnn: Pose-based cnn features for action recognition,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3218-3226, 2015.
- [2] I. Atmosukarto, N. Ahuja, and B. Ghanem, “Action recognition using discriminative structured trajectory Groups,” *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on*. pp. 899-906. IEEE, 2015.
- [3] S. Zaidenberg, B. Boulay, and F. Brémond, “A generic framework for video understanding applied to group

زیرنویس‌ها

Support Vector Machine<sup>1</sup>

---

K-Nearest Neighbor<sup>۳</sup>

Scale Invariant Feature Transform<sup>۴</sup>

Histogram of Oriented Gradient<sup>۴</sup>

Histogram of Optical Flow<sup>۵</sup>

Johansson<sup>۶</sup>

fence<sup>۷</sup>

Kinect Depth<sup>۸</sup>

Lie algebra<sup>۹</sup>

Body Posture Graph<sup>۱۰</sup>

Skeleton Context<sup>۱۱</sup>

Occlusion<sup>۱۲</sup>

Bipolar Histogram<sup>۱۳</sup>

Scale Invariant<sup>۱۴</sup>