

آشکارسازی گل در ویدیو مسابقات فوتبال با استفاده از سیستم استنتاج فازی

محمدحسین سیگاری، حمید سلطانیانزاده و حمیدرضا پوررضا

چکیده

رویداد گل، مهمترین رویداد مسابقه فوتبال است و آشکارسازی آن در سیستم‌های خلاصه‌سازی، جستجو و بازیابی ویدیو مسابقات فوتبال کاربرد زیادی دارد. در این مقاله روشی جدید برای آشکارسازی گل در ویدیوی بخش تلویزیونی مسابقات فوتبال ارائه می‌گردد. روش پیشنهادی با پردازش داده‌های صوتی و تصویری در سطح پایین و سطح میانی، ویژگی‌های مناسب را استخراج کرده و از آنها برای آشکارسازی گل که یک مفهوم سطح بالاست، استفاده می‌کند. در پردازش‌های سطح پایین، ابتدا انرژی صوتی و هیستوگرام سه‌بعدی در فضای RGB استخراج شده و مرز بین شات‌ها تعیین می‌گردد. سپس در پردازش‌های سطح میانی، نوع نما برای هر شات تعیین می‌شود. همچنین لوگوی مسابقات در ویدیو آشکارسازی شده و بر اساس آن، بخش‌های بخش مجدد مشخص می‌شود. در مرحله آخر، پس از بخش‌بندی ویدیو به واحدهای معنایی، ویژگی‌های استخراج شده از پردازش‌های سطح پایین و سطح میانی، با استفاده از یک سیستم استنتاج فازی مورد بررسی قرار گرفته تا بخش‌هایی از ویدیو که شامل رویداد گل هستند، آشکارسازی شوند. نوآوری عمده این روش، به کارگیری دانش زمینه و قوانین اکتشافی در قالب سیستم استنتاج فازی است. این نوآوری علاوه بر ارائه یک راهکار برای توصیف ساده‌تر قوانین اکتشافی توسط فرد خبره و اعمال آن به سیستم، از مزایای مدل‌سازی فازی و استنتاج فازی نیز بهره‌مند است. آزمایش‌های انجام شده بر روی ۱۲ ویدیو مربوط به مسابقات جام جهانی ۲۰۱۰ آفریقای جنوبی نشان می‌دهد نرخ دقت و نرخ بازخوانی روش پیشنهادی به ترتیب ۹۰٫۹٪ و ۹۰٫۹٪ می‌باشد و نسبت به روش‌های دیگر عملکرد بهتری دارد.

کلید واژه‌ها

آشکارسازی گل؛ فوتبال؛ سیستم استنتاج فازی؛ ویدیو بخش تلویزیونی.

۱ مقدمه

تلویزیونی و حتی شخصی مورد توجه افراد زیادی قرار گرفته است. از این رو، تحقیقات زیادی در زمینه تحلیل خودکار ویدیو مسابقات فوتبال به منظور خلاصه‌سازی، آشکارسازی رویدادهای مهم، استخراج آمار، تحلیل استراتژیک و تاکتیکی مسابقه انجام شده است. از جمله این تحقیقات می‌توان به آشکارسازی موقعیت‌های حساس [۱]، آشکارسازی گل [۲] و [۳]، ردیابی بازیکنان [۴]، [۵] و توپ [۶] و خلاصه‌سازی ویدیو مسابقات فوتبال [۷]، [۸] اشاره نمود. در این میان، آشکارسازی گل، یکی از جذاب‌ترین تحقیقات است که نتایج آن می‌تواند در سیستم‌های خلاصه‌سازی، سیستم همیار داور و تحلیل تاکتیکی مسابقه کاربردهای فراوانی داشته باشد.

ورزش فوتبال یکی از پرطرفدارترین ورزش‌ها در دنیاست که علاقمندان زیادی دارد. در سال‌های اخیر، تحلیل فنی و آماری خودکار ویدیو مسابقات فوتبال، از دیدگاه داور، مربی‌گری،

این مقاله در مردادماه ۱۳۹۱ دریافت، در خردادماه ۱۳۹۲ بازنگری و در خردادماه ۱۳۹۲ پذیرفته شد.

محمدحسین سیگاری، دانشگاه تهران، دانشکده مهندسی برق و کامپیوتر
hoseyn_sigari@ieee.org

حمید سلطانیانزاده، دانشگاه تهران، دانشکده مهندسی برق و کامپیوتر؛ و آزمایشگاه تحلیل تصویر، بخش رادیولوژی، بیمارستان هنری فورد، دیترویت، میشیگان، ایالات متحده آمریکا
hszadeh@ut.ac.ir

حمیدرضا پوررضا، دانشگاه فردوسی مشهد، گروه مهندسی کامپیوتر
hpourreza@um.ac.ir

تصاویر می‌توانند مفاهیم مختلفی را در سطوح معنایی متفاوت دربرداشته باشند. از طرفی نمی‌توان مستقیماً بر اساس اطلاعات خام ویدیویی، رویدادها و مفاهیم ویدیو که در سطح بالا قرار دارند را استخراج کرد. بنابراین لازم است استخراج اطلاعات معنایی در سطوح مختلف انجام گیرد.

ماهیت پردازش‌ها در هر سطح متفاوت بوده و در نتیجه اطلاعات استخراج شده از آنها، در سطوح معنایی متفاوتی خواهد بود. شکاف معنایی، اهمیت بیشتری نسبت به شکاف حسگری دارد و در بسیاری از تحقیقات، شکاف معنایی بیشتر مورد توجه قرار گرفته است. براساس یکدسته بندی کلی، پردازش‌های انجام شده برای تحلیل معنایی ویدیو را می‌توان در سه سطح معنایی طبقه‌بندی کرد: (۱) پردازش‌های سطح پایین، (۲) پردازش‌های سطح میانی و (۳) پردازش‌های سطح بالا. هرچند مرز دقیقی میان این سطوح وجود ندارد، اما دسته‌بندی آنها باعث بهبود درک و حل مسئله می‌شود.

پردازش‌هایی مانند آشکارسازی مرز بین شات‌ها ([۱۲]، [۱۳]) و تشخیص رنگ چمن و رنگ لباس بازیکنان ([۱۴]) از جمله پردازش‌های سطح پایین در سیستم‌های آشکارسازی گل در ویدیو مسابقات فوتبال محسوب می‌شود. تشخیص نوع نما (نمای دور، نمای متوسط، نمای بسته و نمای خارج) ([۲]، [۱۵])، آشکارسازی و ردیابی توپ ([۱۶]، [۱۸])، بازیکنان ([۵]، [۱۵]، [۱۹]) و داور ([۱۵]، [۲۰]) و همچنین تشخیص پخش مجدد ([۲]، [۲۰])، از جمله پردازش‌های سطح میانی می‌باشند که معمولاً بر اساس ویژگی‌های سطح پایین استخراج شده و دربردارنده مفاهیم سطح بالاتری هستند. برای آشکارسازی گل لازم است پردازش‌های سطح بالا بر روی ویژگی‌های سطح میانی انجام شود تا وقوع گل در ویدیو تشخیص داده شود. از جمله این تحقیقات می‌توان به استفاده از قوانین اکتشافی ([۲]، [۱۵]، [۲۱])، درخت تصمیم [۱۷] و شبکه بیز پویا^۴ [۲۰] برای آشکارسازی گل اشاره نمود.

در این مقاله، با تأکید بر ارائه راهکاری برای کاهش شکاف معنایی، یک روش کارآمد برای آشکارسازی گل در ویدیو پخش تلویزیونی مسابقات فوتبال ارائه می‌گردد. در روش پیشنهادی، پس از بخش‌بندی ویدیو، با استخراج ویژگی‌های تصویری و صوتی از هر بخش، امکان وقوع گل در هر بخش با استفاده از یک سیستم استنتاج فازی، بررسی می‌گردد. سپس با اعمال یک حد آستانه، وقوع گل در ویدیو آشکارسازی می‌شود.

این مقاله در پنج بخش آماده شده است. در این بخش مقدمه‌ای بر موضوع و تعریف مسئله ارائه شد. در بخش دوم، با مروری کوتاه بر تحقیقات انجام شده در زمینه آشکارسازی گل در ویدیو پخش تلویزیونی مسابقات فوتبال، مزایا و معایب آنها مورد بررسی قرار می‌گیرد. در بخش سوم، روش پیشنهادی و

تحقیقات انجام شده در زمینه آشکارسازی گل را می‌توان به دو دسته کلی تقسیم‌بندی نمود. در دسته اول، آشکارسازی گل با استفاده از دوربین‌های دلخواه (قابل کنترل) که در مکان‌های مشخص و با زاویه معینی قرار گرفته‌اند، انجام می‌شود. از جمله این تحقیقات می‌توان به [۳] اشاره کرد که با استفاده از چهار دوربین در اطراف دروازه‌ها، به صورت بلادرنگ عبور توپ از خط دروازه را آشکارسازی نموده و از آن می‌توان به عنوان یک سیستم همیار داور استفاده نمود. مهمترین چالش‌ها در این دسته از تحقیقات، پردازش بلادرنگ تصاویر، رفع انسداد و هم‌پوشانی اشیا با استفاده از تصاویر چند دوربین و دستیابی به دقت‌های بالا (نزدیک به ۱۰۰٪) به منظور داشتن یک سیستم قابل اعتماد است.

دسته دوم تحقیقات که حجم بیشتری را به خود اختصاص داده‌اند، به آشکارسازی گل در ویدیو پخش تلویزیونی می‌پردازند. مهمترین چالش در این دسته از تحقیقات این است که تصاویر ویدیویی غیرقابل کنترل بوده و در نتیجه، تحلیل ویدیو معمولاً دقت و کارایی خوبی ندارد. این دسته از تحقیقات معمولاً برای تحلیل، خلاصه‌سازی و نمایه‌سازی آرشیوهای ویدیویی شبکه‌های تلویزیونی یا شخصی کاربرد دارد [۱۱].

به طور کلی، سیستم‌های آشکارسازی گل در ویدیو مسابقات فوتبال همانند سایر سیستم‌های تحلیل معنایی ویدیو همواره با دوچالش اساسی روبرو هستند: (۱) شکاف حسگری^۱ و (۲) شکاف معنایی^۲. به شکاف میان موجودیت شی در عالم واقعیت و توصیف آن در تصویر، شکاف حسگری می‌گویند. به عبارت دیگر، شکاف حسگری به مشکلات ابزارهای تصویربرداری و نقش آنها در تحلیل تصویر می‌پردازد. از جمله مهمترین مشکلات در زمینه شکاف حسگری، قابلیت کنترل دوربین، نرخ تصویربرداری و کیفیت وضوح تصویر است. در [۳]، با استفاده از چهار دوربین خاص با نرخ تصویربرداری ۲۰۰ قاب در ثانیه^۳، و نصب آنها در مکان‌های دلخواه، تا حدود زیادی، شکاف حسگری مربوط به تشخیص عبور توپ از خط دروازه رفع شده است. چرا که در روش ارائه شده اولاً تعداد و مکان دوربین‌ها طوری است که انسداد و هم‌پوشانی بین اشیا قابل تشخیص و رفع است، ثانیاً به دلیل نرخ تصویربرداری بالا، تصویر توپی که با سرعت بالا به سمت دروازه حرکت می‌کند، دچار کشیدگی و اعوجاج نمی‌شود.

شکاف معنایی، به مسئله عدم تطابق میان مفاهیم استخراج شده توسط یک سیستم تحلیل معنایی ویدیو و مفاهیم مقابل درک توسط انسان می‌پردازد. به عبارت دیگر، تفسیر معانی از اطلاعات تصویری در سیستم کامپیوتری و انسان متفاوت است، که این موضوع طراحی سیستم‌های تحلیل معنایی ویدیو را مشکل می‌کند.

¹Sensory Gap

²Semantic Gap

³Frame Per Second (FPS)

⁴Dynamic Bayesian Network (DBN)

زمانی می‌تواند کارایی خوبی داشته باشد که تعداد گل‌ها در بازی مشخص باشد.

در (Kolekar et al. 2009) [۱۵]، یک ساختار سلسله مراتبی برای آشکارسازی گل و موقعیت‌های حمله به سمت دروازه ارائه شده است. این ساختار شبه درختی، بر اساس قوانین اکتشافی طراحی شده است. در سطوح ابتدایی و میانی، پردازش‌هایی شامل استخراج انرژی صوتی، تشخیص نوع نما و آشکارسازی بخش‌های پخش مجدد و در سطوح بالا، پردازش‌هایی مانند آشکارسازی داور و دروازه‌بان و همچنین آشکارسازی گل و موقعیت‌های خطرناک انجام می‌شود. هرچند این روش ساده و نسبتاً کارآمد به نظر می‌رسد، اما در آن از قوانین اکتشافی بسیاری استفاده شده است که انعطاف‌پذیری و قابلیت استفاده مجدد از آن را برای ویدیوهای دیگر با مشکل مواجه می‌کند.

در (Tjondronegoro et al. 2010) [۲۱]، با استفاده از قوانین اکتشافی، سیستمی برای آشکارسازی رویداد در ویدیوی مسابقات ورزشی ارائه شده است. در این روش، ابتدا نوع نما تشخیص داده شده و بخش‌های پخش مجدد آشکارسازی می‌شوند. سپس بر اساس ترکیب این دو ویژگی، ویدیو به بخش‌های بازی-توقف^۴ بخش‌بندی می‌گردد. سپس بر اساس یک دسته از ویژگی‌ها شامل نسبت بخش بازی و توقف، میزان نمایش نمای بسته، نمای دروازه و نسبت بخش پخش مجدد، تعدادی از بخش‌های ویدیو به عنوان بخش برجسته انتخاب می‌شوند تا پردازش‌های بعدی بر روی آنها صورت گیرد. در مرحله آشکارسازی رویداد، هر یک از بخش‌های برجسته توسط قوانین اکتشافی مورد ارزیابی قرار می‌گیرند تا نوع رویداد در آن بخش تعیین گردد. بر اساس ادعای نویسندگان، مهمترین مزیت این روش این است که با تغییر قوانین اکتشافی می‌توان از این ساختار برای آشکارسازی رویدادهای مهم در ویدیو مسابقات فوتبال، بسکتبال و فوتبال استرالیایی استفاده نمود.

در ([۹]، [۱۰])، سیستم‌های مشابهی برای آشکارسازی و بازیابی رویدادهای مهم ویدیوهای فوتبال شامل گل، پنالتی، کرنر و ضربه آزاد ارائه شده است. در این تحقیقات، از سیستم استنتاج فازی به عنوان یک مرحله پیش‌پردازشی برای حذف شات‌های کم اهمیت از ویدیو استفاده شده است. این سیستم استنتاج فازی بر اساس نوع نما و میزان بخش‌های حرکت آهسته در هر شات، میزان اهمیت آن را تخمین زده و سپس شات‌های کم اهمیت را از ویدیو حذف می‌کند. سپس برای هر یک از شات‌های مهم، یک دسته ویژگی استخراج می‌گردد. این ویژگی‌ها توسط ماشین بردار پشتیبان به پنج کلاس شامل گل، پنالتی، کرنر، ضربه آزاد و سایر رویدادها تقسیم‌بندی می‌گردد. اگرچه راهکار پیشنهادی بر اساس ماشین بردار پشتیبان می‌باشد، اما در این مقاله هیچ توضیحی در مورد نوع تابع هسته و چگونگی استفاده از ماشین بردار پشتیبان

جزئیات مربوط به هر یک از قسمت‌های آن بیان می‌گردد. نتایج آزمایش‌ها و بحث در مورد آنها در بخش چهارم ارائه خواهد شد. آخرین بخش مربوط به نتیجه‌گیری و پیشنهاداتی برای ادامه پژوهش‌ها در این زمینه می‌باشد.

۲. مروری بر کارهای گذشته

در این بخش، تعدادی از مهمترین تحقیقات انجام شده در زمینه آشکارسازی گل در ویدیوی پخش تلویزیونی مسابقات فوتبال را مرور می‌کنیم. در اکثر این تحقیقات، آشکارسازی گل بر اساس قوانین اکتشافی و وابسته به دانش زمینه است. اما در معدودی از آنها، از روش‌های یادگیری ماشین در کنار قوانین اکتشافی^۱ و دانش زمینه استفاده شده که در این حالت معمولاً روش ارائه شده عمومیت بیشتری دارد و می‌تواند برای آشکارسازی رویدادهای مختلف مورد استفاده قرار گیرد، اما معمولاً دقت پایین‌تری نسبت به روش‌های مبتنی بر قوانین اکتشافی دارند.

در [۱۸]، برای آشکارسازی گل، ابتدا ویدیو در سطح شات بخش‌بندی می‌شود. سپس یک سری ویژگی‌های تصویری شامل هیستوگرام، میزان تحرک بین قاب‌ها، رنگ پس‌زمینه و یک سری ویژگی‌های صوتی شامل شدت انرژی صوتی و حجم صدا از هر شات استخراج می‌شوند. در مرحله آشکارسازی گل، از درخت تصمیم استفاده شده است. مهمترین ایراد در این تحقیق، استفاده مستقیم از ویژگی‌های سطح پایین برای آشکارسازی مفهوم سطح بالای گل است. به عبارت دیگر، عدم استفاده از ویژگی‌های سطح میانی برای ایجاد ارتباط معنایی میان ویژگی‌های سطح پایین و مفهوم سطح بالا، باعث عدم کارایی این روش می‌شود.

در (Huang et al. 2006) [۲۰]، برای آشکارسازی گل و سایر رویدادهای مهم بازی، یک روش مبتنی بر شبکه بیز پویا ارائه شده است. در این روش، پس از آشکارسازی یک سری ویژگی‌های سطح میانی شامل داور، لوگو، دروازه، خطوط اطراف زمین و تشخیص نوع نما و حرکت دوربین، این ویژگی‌ها توسط یک شبکه بیز پویا بررسی شده و رویدادهای گل، پنالتی، کارت، کرنر و ضربه آزاد در ویدیو آشکارسازی می‌شود. این روش بر خلاف سایر روش‌ها، مبتنی بر شات نیست و ویدیو را قاب^۲ به قاب مورد بررسی قرار می‌دهد. روش ارائه شده در این مقاله، تغییر محتوای قاب‌ها را در قاب‌های متوالی ارزیابی می‌کند و از این نظر، حجم محاسباتی بسیار زیادی دارد.

در [۲]، آشکارسازی گل بر اساس ترکیب تعدادی قانون اکتشافی و الگوریتم خوشه‌بندی فازی^۳ انجام شده است. ویژگی‌های استخراج شده برای آشکارسازی گل، تشخیص نوع نما و آشکارسازی بخش‌های پخش مجدد است. روش پیشنهادی

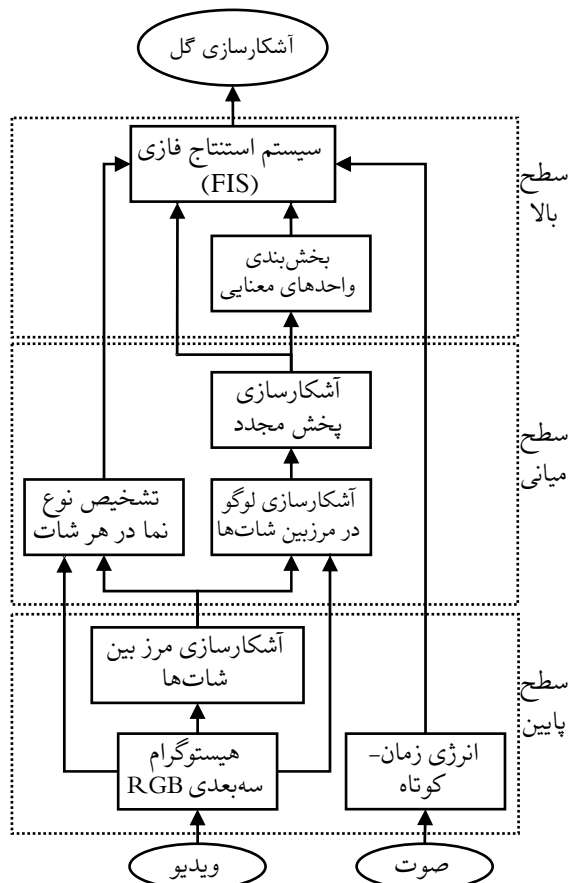
^۱Heuristic

^۲Frame

^۳Fuzzy C-Means (FCM)

^۴Play-Break

معنایی مورد ارزیابی قرار می‌گیرد. ورودی‌های سیستم استنتاج فازی عبارتند از: (۱) مدت زمان پخش مجدد، (۲) مدت زمان پخش نمای بسته و نمای خارج و (۳) میانگین انرژی صوتی. سیستم استنتاج فازی با استفاده از روش استنتاج مددانی، برای هر واحد معنایی یک عدد در بازه $[0,1]$ به عنوان «امکان وقوع گل» تولید می‌کند. هر چه این عدد بیشتر باشد، به معنای این است که با امکان بیشتری رویداد گل در واحد معنایی مورد نظر اتفاق افتاده است. بنابراین با اعمال یک حدآستانه، می‌توان وقوع گل در هر واحد معنایی را آشکارسازی نمود. در ادامه، هر یک از بخش‌های روش پیشنهادی با جزئیات بیشتری مورد بررسی قرار می‌گیرد.



شکل ۱ بلوک دیاگرام روش پیشنهادی برای آشکارسازی گل.

۳-۱ هیستوگرام سه بعدی RGB

هیستوگرام یک ابزار آماری برای محاسبه فراوانی نمونه‌ها در بازه‌های مختلف است. متداول‌ترین روش محاسبه هیستوگرام در تصاویر RGB، محاسبه هیستوگرام ۱۶ یا ۳۲ ستونی برای هر یک از مولفه‌های فضای RGB است. به عنوان مثال، با فرض محاسبه هیستوگرام ۱۶ ستونی، برای هر مولفه R، G و B، بردار هیستوگرام ۱۶ ستونی محاسبه می‌شود که در نهایت با کنار هم قرار دادن این بردارها، یک بردار ۴۸ عضوی به عنوان هیستوگرام کل تصویر بدست می‌آید. با توجه به این که رنگ در فضای RGB وابسته به مقدار تمام مولفه‌ها است و در ضمن در این روش، هیستوگرام هر مولفه جداگانه محاسبه می‌گردد، این روش محاسبه هیستوگرام، توصیف خوبی از رنگ‌های درون تصویر ارائه نمی‌کند.

برای طبقه‌بندی یک مسئله چندکلاسی ازائه نشده است. ضمن اینکه تعداد نمونه‌های بسیار کم (بین ۵ تا ۹ نمونه) از هر کلاس، استفاده از روش‌های مبتنی بر یادگیری مانند ماشین بردار پشتیبان را دشوار می‌نماید.

همان طور که ملاحظه گردید، پایه و اساس تمام روش‌های مورد مطالعه، قوانین اکتشافی و دانش زمینه است که توسط طراح در سیستم به کار گرفته می‌شود. در اغلب تحقیقات، دانش زمینه در قالب قوانین اکتشافی و به شکل غیرمدون و غیرمنسجم مورد استفاده قرار گرفته است. در برخی از تحقیقات، علاوه بر قوانین اکتشافی، از روش‌های یادگیری ماشین نیز استفاده شده است. اگرچه استفاده از روش‌های یادگیری ماشین باعث افزایش عمومیت‌پذیری سیستم می‌شود، اما این روش‌ها نیاز به تعداد کافی و متنوع نمونه‌های آموزشی دارد. با توجه تنوع زیاد نحوه وقوع رویداد گل در بازی فوتبال، تهیه نمونه‌های آموزشی کافی و متنوع از مشکلات این دسته از روش‌ها محسوب می‌گردد. در این مقاله، با استفاده از سیستم استنتاج فازی، دانش زمینه و قوانین اکتشافی به طور مدون و منسجم برای آشکارسازی گل مورد استفاده قرار گرفته است. در این حالت، علاوه بر این که دانش زمینه به شکلی ساده و قابل فهم در قالب مجموعه‌ای از قوانین فازی ارائه می‌شود، از مدل استنتاج فازی به عنوان یک ابزار قدرتمند برای تصمیم‌گیری در مورد امکان وقوع گل استفاده می‌گردد. اگرچه در تحقیقات گذشته از سیستم استنتاج فازی به عنوان یک فرایند پیش‌پردازش برای تحلیل ویدیو فوتبال استفاده شده است، اما در این مقاله از سیستم استنتاج فازی برای تصمیم‌گیری در مورد امکان وقوع گل بهره می‌بریم. ضمن این که برخلاف بسیاری از تحقیقات گذشته، آشکارسازی گل در یک شات صورت نمی‌پذیرد، بلکه پس از بخش بندی ویدیو به واحدهای معنایی (شامل چندین شات)، وقوع گل در یک واحد معنایی آشکارسازی می‌گردد.

۳ روش پیشنهادی

روش پیشنهادی شامل سه مرحله اصلی می‌باشد: (۱) پردازش‌های سطح پایین، (۲) پردازش‌های سطح میانی و (۳) پردازش‌های سطح بالا. بلوک دیاگرام روش پیشنهادی در شکل ۱ نشان داده شده است.

در مرحله اول، پردازش‌های سطح پایین به منظور استخراج هیستوگرام سه بعدی RGB، آشکارسازی مرز بین شات‌ها و محاسبه انرژی صوتی انجام می‌شود. در مرحله بعد، بر اساس ویژگی‌های سطح پایین، نوع نما (نمای دور، نمای متوسط، نمای بسته و نمای خارج) برای هر شات مشخص شده و لوگو در ویدیو آشکارسازی می‌شود. همچنین بر اساس زمان نمایش لوگو در ویدیو، بخش‌های پخش مجدد در ویدیو آشکارسازی می‌گردد.

با آشکارسازی بخش‌های پخش مجدد، ویدیو در سطح بالا به واحدهای معنایی بخش بندی می‌گردد. سپس با استفاده از یک سیستم استنتاج فازی، ویژگی‌های استخراج شده از هر واحد

در روش پیشنهادی، پس از محاسبه هیستوگرام سه بعدی هر قاب، برای آشکارسازی مرز بین شات‌ها، تفاضل هیستوگرام قاب‌های متوالی محاسبه می‌شود. در صورتی که مجموع قدرمطلق تفاضل هیستوگرام قاب‌های متوالی بیشتر از حد آستانه T_{shot} شود، مرز میان دو قاب، به عنوان مرز بین شات‌ها تلقی می‌شود. حد آستانه T_{shot} به صورت تطبیقی و از رابطه زیر تعیین می‌گردد:

$$T_{shot} = average(HD_i | HD_i) AHD \quad (۱)$$

در رابطه فوق، HD_i مجموع قدرمطلق تفاضل هیستوگرام سه بعدی قاب i و $i-1$ و AHD میانگین تمام HD_i ها است. به این ترتیب، T_{shot} برابر با میانگین تمام HD_i هایی است که مقدار آنها بیشتر از مقدار میانگین باشد.

با توجه به اینکه در ویدیو ممکن است گذار تدریجی میان شات‌ها وجود داشته باشد، این حد آستانه تا حد ممکن پایین در نظر گرفته شده تا گذارهای تدریجی نیز آشکارسازی گردند. اگرچه در برخی موارد ممکن است جابجایی سریع اشیاء یا حرکت سریع دوربین در نمای بسته و نمای متوسط، به عنوان مرز بین شات‌ها آشکارسازی شود، اما با انتخاب این حد آستانه تقریباً می‌توان مطمئن بود که تمام گذارها آشکارسازی می‌شوند. با توجه به توضیحاتی که در ادامه ارائه می‌گردد، آشکارسازی اشتباه مواردی که واقعاً گذار بین شات‌ها نیست، باعث کاهش کارایی سیستم نمی‌شود، اما عدم آشکارسازی مرز بین شات‌ها ممکن است باعث بروز اختلال در روش پیشنهادی گردد. به این ترتیب با پذیرش خطای آشکارسازی اشتباه مرز میان شات‌ها، از بروز خطا در عدم آشکارسازی مرز میان شات‌ها جلوگیری می‌کنیم.

۳-۳ انرژی صوتی

یکی از ویژگی‌های سطح پایین که می‌تواند در آشکارسازی گل و سایر رویدادهای مهم مسابقه فوتبال بسیار موثر باشد، انرژی صوتی است. در تصاویر پخش تلویزیونی، پس از به ثمر رسیدن گل، هیجان و تشویق تماشاچیان و گوینده تلویزیونی به طور قابل توجهی افزایش می‌یابد. بنابراین، با محاسبه انرژی صوتی بخش‌های ویدیو می‌توان تا حدود زیادی به میزان اهمیت آن بخش از ویدیو پی برد. در روش پیشنهادی، میزان انرژی صوتی در هر قاب از ویدیو محاسبه می‌شود. به عبارت دیگر، انرژی داده‌های صوتی در پنجره‌های غیرهمپوشان به طول ۴۰ میلی ثانیه محاسبه می‌شوند.

۳-۴ آشکارسازی نوع نما

در حالت کلی، نما در ویدیو مسابقات فوتبال به چهار گروه تقسیم‌بندی می‌شود: (۱) نمای دور، (۲) نمای متوسط، (۳) نمای بسته و (۴) نمای خارج. نوع نما در هر شات تقریباً ثابت است. بنابراین می‌توان برای هر شات یک نوع نما تعیین کرد.

در حالت کلی، معمولاً نمای دور یا نمای متوسط از بازی پخش می‌شود. اما پس از وقوع گل یا یک رویداد برجسته در

روش بهتر برای توصیف رنگ‌های تصویر، استفاده از هیستوگرام سه بعدی RGB است. در این روش، ابتدا هر مولفه از فضای RGB به n سطح تقسیم‌بندی شده و در نتیجه، کل فضای RGB به n^3 مکعب کوانتیزه می‌شود. هر مکعب نماینده یک رنگ کوانتیزه شده در فضای RGB است. سپس هیستوگرام تصویر برای هر یک از این رنگ‌های کوانتیزه شده محاسبه می‌گردد. این روش محاسبه هیستوگرام، توصیف بسیار بهتری از رنگ‌های تصویر ارائه می‌نماید، اما مهمترین عیب آن، حجم محاسبات بسیار زیاد نسبت به روش قبل است. در روش پیشنهادی، $n=6$ فرض شده و هیستوگرام سه بعدی بدست آمده شامل ۲۱۶ عضو می‌باشد.

شایان ذکر است که استفاده از هیستوگرام رنگ در فضاهای رنگی دیگر مانند HSV و YCbCr نیز مشکلاتی مشابه هیستوگرام رنگ در فضای RGB را دارد. به عنوان مثال، اگرچه هیستوگرام مولفه H در فضای HSV می‌تواند توصیف بسیار خوبی از فضای رنگی تصویر ارائه دهد که در بسیاری از موارد بهتر از هیستوگرام RGB است، اما به تنهایی نمی‌تواند توصیف مناسبی از رنگ‌های سطح خاکستری داشته باشد. بنابراین برای توصیف کامل رنگ‌ها در فضای HSV نیز باید از تمام مولفه‌های آن استفاده نمود. در این حالت، کوانتیزه کردن سه بعدی فضای HSV پیچیده‌تر از کوانتیزه کردن سه بعدی فضای RGB خواهد بود. ضمن این که تبدیل تصویر از فضای رنگی RGB به HSV بار محاسباتی مضاعف به سیستم اعمال می‌کند.

۳-۲ آشکارسازی مرز بین شات‌ها

شات مجموعه‌ای از قاب‌های متوالی است که به صورت پیوسته از یک دوربین گرفته شده‌اند، به طوریکه مشخصات دوربین ثابت است یا تغییرات کندی دارد. با آشکارسازی مرز بین شات‌ها، بخش‌بندی ویدیو در سطح پایین انجام می‌شود. بنابراین آشکارسازی مرز بین شات‌ها به عنوان یک پردازش سطح پایین می‌تواند در پردازش‌های سطح بالاتر مفید واقع شود.

تغییر شات معمولاً به صورت ناگهانی و در یک قاب انجام می‌شود که به آن گذار ناگهانی یا برش^۱ می‌گویند، اما ممکن است به لحاظ زیبایی‌شناسی، تغییر شات به صورت تدریجی و در طی چند قاب انجام شود. به تغییرات تدریجی شات، گذار تدریجی می‌گویند که انواع مختلفی دارد [۲۲]. معمولاً آشکارسازی مرز بین شات‌ها در حالتی که گذار ناگهانی رخ دهد، با دقت زیاد انجام می‌شود، اما آشکارسازی گذار تدریجی کمی مشکل است. روش‌های متعددی از جمله [۱۲]، [۱۳]، [۲۳]، برای آشکارسازی مرز بین شات‌ها ارائه شده که تقریباً در تمام آنها، تاکید بر آشکارسازی گذارهای تدریجی میان شات‌هاست. اما اغلب این روش‌ها حجم محاسبات زیادی دارند و استفاده از آنها در سیستم پیشنهادی ضروری نیست.

^۱Cut

۳-۵ آشکارسازی لوگو

تقریباً در تمام ویدیوهای مسابقات فوتبال، در ابتدا و انتهای بخش‌های پخش مجدد و حرکت آهسته، لوگوی مسابقات به شکل یک گذار پخش می‌شود. بنابراین با آشکارسازی گذارهای تدریجی و سپس آشکارسازی لوگو می‌توان بخش‌های پخش مجدد در ویدیو را تشخیص داد. قاب‌هایی از یک نمونه گذار لوگو در شکل ۳ نشان داده شده است. همان‌طور که مشاهده می‌گردد، لوگو در بسیاری از قاب‌های میانی این گذار به خوبی مشخص است.

در روش پیشنهادی برای آشکارسازی لوگو، از درخت تصمیم استفاده شده است. در این روش، پس از تعیین مرز بین شات‌ها، گذار بین شات‌ها مشخص می‌گردد. سپس هیستوگرام سه‌بعدی قاب‌هایی که در گذار بین شات‌ها قرار دارند، توسط یک درخت تصمیم مورد ارزیابی قرار می‌گیرند تا وجود لوگو در آن قاب‌ها تعیین شود (شکل ۴). با توجه به این که لوگو در چند قاب از گذار به خوبی ظاهر می‌شود، می‌توان با ترکیب نتایج حاصل از آشکارسازی لوگو در قاب‌های متوالی یک گذار، خطای عدم آشکارسازی لوگو را به حداقل رساند. بنابراین در مرحله آموزش درخت تصمیم، سعی می‌شود تا نمونه‌های آموزشی منفی بیشتر از نمونه‌های آموزشی مثبت باشد تا نرخ خطای آشکار سازی اشتباه نسبت به نرخ خطای عدم آشکار سازی کاهش یابد. چرا که در بخش بعد خواهیم دید که عدم آشکارسازی لوگو، تاثیرات منفی کمتری نسبت به آشکارسازی اشتباه آن خواهد داشت.



(الف)



(ب)



(ت)



(پ)



(ج)



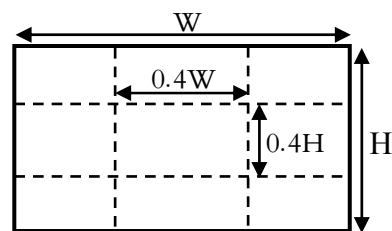
(ث)

شکل ۳ قاب‌هایی از یک نمونه گذار لوگو.

مسابقه، معمولاً نمای بسته بازیکنی که موجب بروز آن رویداد شده پخش می‌شود و ممکن است نمایی از تماشاچیان و هواداران آن تیم نیز نمایش داده شود. بنابراین تشخیص نوع نمای بسته و نمای خارج در میان شات‌ها می‌تواند کمک موثری در تشخیص وقوع گل داشته باشد. در روش پیشنهادی، نوع نما به دو دسته تقسیم‌بندی می‌شود: (۱) نمای دور و نمای متوسط و (۲) نمای بسته و نمای خارج. علت این تقسیم‌بندی در این است که برای آشکارسازی گل، وقوع نمای بسته و نمای خارج به یک میزان مورد توجه است. ضمن این که تقسیم‌بندی نوع نما به دو دسته، ساده‌تر و سریع‌تر از تقسیم‌بندی نما به چهار دسته می‌باشد.

برای تشخیص نوع نما، ابتدا رنگ چمن بر اساس ۱۰۰ قاب تصادفی از ویدیو تعیین می‌گردد، به طوری که رنگ انتخاب شده، بیشترین فراوانی را در میانگین هیستوگرام سه‌بعدی این ۱۰۰ قاب داشته باشد. سپس، نسبت پیکسل‌های چمن به کل پیکسل‌های قاب محاسبه شده و با نماد R_{Grass} نشان داده می‌شود. در نمای خارج از زمین، R_{Grass} تقریباً صفر است. بنابراین با اعمال حدآستانه $R_{Grass} < 0.05$ می‌توان نمای خارج از زمین را تشخیص داد.

در نمای بسته، تصویر بزرگی از چهره در مرکز قاب نمایش داده می‌شود. بنابراین، میزان پیکسل‌های مربوط به رنگ پوست در مرکز قاب زیاد است. برای این منظور، ابتدا قاب مطابق شکل ۲، به ۹ بخش تقسیم می‌شود به طوری که بخش مرکزی، بزرگتر از سایر بخش‌ها است. نسبت پیکسل‌های چمن به کل پیکسل‌ها در بخش مرکزی با نماد R_{Grass}^{Center} و نسبت پیکسل‌های پوست به کل پیکسل‌ها در بخش مرکزی با نماد R_{Skin}^{Center} نمایش داده می‌شود. اگر $R_{Grass}^{Center} \leq 0.5$ و $R_{Skin}^{Center} \geq 0.1$ برقرار باشد، نوع نمای قاب، نمای بسته تشخیص داده خواهد شد. قاب‌هایی که از نوع نمای خارج یا نمای بسته تعیین نشوند، نمای دور یا متوسط هستند. هرچند در روش پیشنهادی، فقط نمای خارج و نمای بسته مورد توجه قرار می‌گیرد.



شکل ۲ تقسیم‌بندی قاب به ۹ بخش برای تعیین نمای بسته.

برای تشخیص نوع نمای شات، سه قاب از شات انتخاب می‌شود. اگر شات در بازه زمانی $[t_1, t_2]$ واقع شده باشد، سه قاب مورد نظر به ترتیب در لحظات $t_1 + 0.1(t_2 - t_1)$ ، $t_1 + 0.5(t_2 - t_1)$ و $t_1 + 0.9(t_2 - t_1)$ انتخاب می‌شود. سپس نوع نما در هر سه قاب تعیین شده و با استفاده از روش رای اکثریت، نوع نمای شات تعیین می‌گردد.

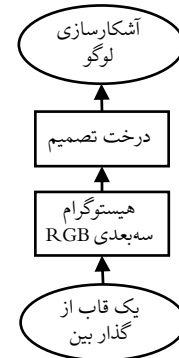
همان نقطه انتهای پخش مجدد است. اما برای تعیین نقطه ابتدای آن، باید به مدت زمان پخش مجدد توجه نمود. اگر مدت زمان یک پخش مجدد L_{Replay} باشد، مدت زمان واحد معنایی متناظر آن دو برابر L_{Replay} است. با توجه به اینکه در این حالت ممکن است نقطه ابتدای واحد معنایی منطبق بر آغاز یک شات نباشد، بنابراین نقطه ابتدای واحد معنایی برابر با نقطه ابتدای شات متناظر قرار می‌گیرد، طوری که مدت زمان واحد معنایی کمتر از دو برابر L_{Replay} نشود.

۳-۸ سیستم استنتاج فازی

در روش پیشنهادی، از سیستم استنتاج فازی برای امکان‌سنجی وقوع گل در یک واحد معنایی استفاده می‌شود. ورودی‌های سیستم استنتاج فازی عبارتند از: (۱) مدت زمان پخش مجدد، (۲) نسبت مدت زمان پخش نمای بسته و نمای خارج به کل مدت زمان واحد معنایی و (۳) میانگین انرژی صوتی در واحد معنایی قبل از پخش مجدد. این سیستم با استفاده از روش استنتاج ممدانی، برای هر واحد معنایی یک عدد در بازه $[0,1]$ به عنوان «امکان وقوع گل» تولید می‌کند. هر چه این عدد بیشتر باشد، به معنای این است که با امکان بیشتری رویداد گل در واحد معنایی مورد نظر اتفاق افتاده است.

همانند سایر سیستم‌های استنتاج فازی، برای هر یک از ورودی‌ها یک تابع عضویت فازی تعریف می‌شود تا ورودی‌های قطعی را به متغیرهای فازی تبدیل کند. تابع عضویت ورودی‌های این سیستم به ترتیب در شکل‌های ۵، ۶ و ۷ نمایش داده شده‌اند و برای خروجی سیستم استنتاج فازی، تابع عضویت مطابق شکل ۸ تعریف گردید. پایگاه قوانین در این سیستم شامل قوانین مندرج در جدول ۱ است.

در بسیاری از سیستم‌های استنتاج فازی، برای تعریف توابع عضویت متغیرهای ورودی از شکل مثلثی و دوزنقه‌ای استفاده می‌شود. مهمترین علت استفاده از این شکل توابع، سادگی آن است. ضمن این که اغلب در تعریف توابع عضویت برای متغیرهای ورودی سعی می‌شود تا مجموع درجه عضویت تمام مقادیر در طول دامنه آن متغیر، یک باشد. بنابراین بر خلاف توابعی مانند توابع گوسی، با استفاده از توابع مثلثی و دوزنقه‌ای می‌توان شرط فوق را ارضا کرد. به همین دلیل برای تعریف تابع عضویت متغیرهای ورودی از شکل مثلثی و دوزنقه‌ای استفاده شد. در مورد توابع عضویت خروجی، از توابع منفرد^۱ استفاده گردید. چرا که برای متغیرهای خروجی لازم نیست تا درجه عضویت تمام مقادیر در طول دامنه، یک باشد. بنابراین ساده‌ترین تابع تابع منفرد برای تعریف توابع عضویت خروجی استفاده شد. هر چند به دلیل این که از روش مرکز ثقل برای غیرفازی‌سازی استفاده می‌شود، استفاده از توابع عضویت با شکل مثلثی برای متغیر خروجی نیز



شکل ۴ ساختار روش پیشنهادی برای آشکارسازی لوگو.

۳-۶ آشکارسازی بخش‌های پخش مجدد

بخش‌های پخش مجدد در ویدیو فوتبال از اهمیت بالایی برخوردار هستند. چرا که این بخش‌ها معمولاً بعد از رویدادهای مهم نمایش داده می‌شوند و رویداد مهم اخیر را از زوایای متفاوت و با حرکت آهسته نشان می‌دهند. معمولاً مدت زمان پخش مجدد با میزان اهمیت رویداد رخ داده رابطه مستقیم دارد. یعنی هر چه رویداد رخ داده مهم‌تر باشد، مدت زمان پخش مجدد افزایش می‌یابد. بنابراین می‌توان با آشکارسازی بخش‌های پخش مجدد، میزان اهمیت رویدادهای ویدیو را تخمین زد.

همان طور که گفته شد، در ابتدا و انتهای بخش‌های پخش مجدد تصویری از لوگوی مسابقات به صورت گذار تدریجی پخش می‌شود. معمولاً مدت زمان این گذار حدود یک ثانیه است و در طول چند قاب، تصویر لوگو ثابت است یا به کندی حرکت می‌کند. بنابراین، بهترین زمان برای آشکارسازی لوگو در همین قاب‌هایی است که تحرک لوگو کم است. اما همان‌طور که گفته شد، آشکارسازی لوگو همراه با خطاست که این خطا می‌تواند آشکارسازی لحظات ابتدا/انتهای پخش مجدد را با مشکل جدی مواجه کند. برای مقابله با این خطا، پس از آشکارسازی لوگو در قاب‌های مربوط به گذار، از یک پنجره لغزان به طول ۵ قاب استفاده می‌شود. به این ترتیب که اگر حداقل در ۳ قاب از ۵ قاب این پنجره، لوگو آشکارسازی شده باشد، مرکز این پنجره به عنوان نقطه ابتدا/انتهای پخش مجدد در نظر گرفته می‌شود. به این ترتیب، اگر در ۲ قاب از ۵ قاب، به اشتباه لوگو آشکارسازی نشود، باز هم این روش می‌تواند نقاط ابتدا/انتهای پخش مجدد را به درستی تعیین کند.

۳-۷ بخش‌بندی ویدیو و تعیین واحدهای معنایی

قبلاً اشاره شد که با آشکارسازی مرز بین شات‌ها، ویدیو در سطح پایین بخش‌بندی می‌شود، و با آشکارسازی بخش‌های پخش مجدد، ویدیو در سطح بالا قابل بخش‌بندی است. بخش‌های ویدیو در سطح بالا با عنوان واحد معنایی نام‌گذاری می‌شود. برای تعیین نقاط ابتدا/انتهای هر واحد معنایی، از نقاط ابتدا/انتهای بخش‌های پخش مجدد کمک می‌گیریم. در واقع، هر واحد معنایی، یک پخش مجدد را شامل می‌شود و نقطه انتهای هر واحد معنایی

¹Singleton

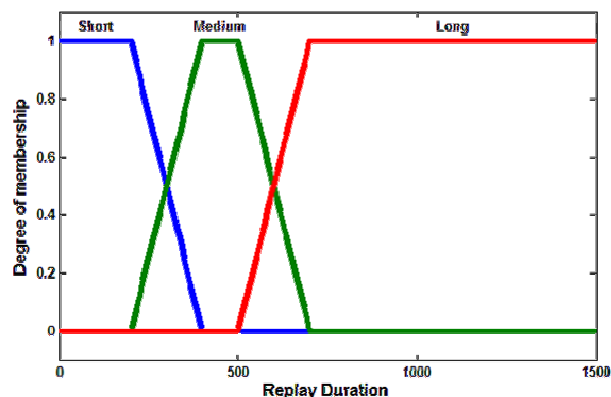
جدول ۱ پایگاه قوانین سیستم استنتاج فازی.

Input 1	Input 2	Input 3	Output
Short	-	-	Very Low
Medium	Low	-	Very Low
Medium	Medium	Low	Low
Medium	Medium	Medium	Low
Medium	Medium	High	Medium
Medium	High	Low	Low
Medium	High	Medium	Low
Medium	High	High	Medium
Long	Low	Low	Low
Long	Low	Medium	Low
Long	Low	High	Medium
Long	Medium	Low	Medium
Long	Medium	Medium	Medium
Long	Medium	High	High
Long	High	Low	High
Long	High	Medium	High
Long	High	High	Very High

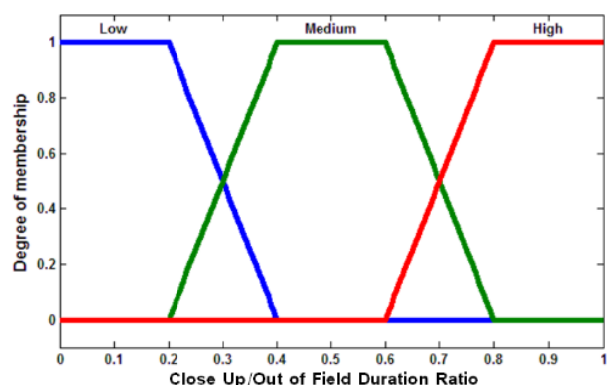
تغییری در نتایج ایجاد نمی‌کند.

معمولا تعداد توابع عضویت برای هر متغیر بر اساس نظر فرد خبره تعیین می‌شود. اما نکته مهم این است که افزایش تعداد توابع عضویت برای هر متغیر، منجر به افزایش تعداد قوانین می‌شود. چرا که باید برای هر مقدار از متغیرهای ورودی، حداقل یک قانون در پایگاه قوانین وجود داشته باشد. افزایش تعداد قوانین، ضمن این که طراحی سیستم را پیچیده‌تر می‌کند، از نظر محاسباتی نیز حجم محاسبات بیشتری خواهد داشت. اگر تعداد متغیرهای ورودی یک سیستم n و تعداد توابع عضویت برای هر متغیر m باشد، حداکثر تعداد قوانین لازم برای پوشش کل فضای ورودی m^n خواهد شد. با توجه به آنچه گفته شد، برای هر یک از متغیرهای ورودی سه تابع عضویت تعیین گردید. همچنین برای متغیر خروجی (امکان وقوع گل) ۵ تابع عضویت تعریف شد.

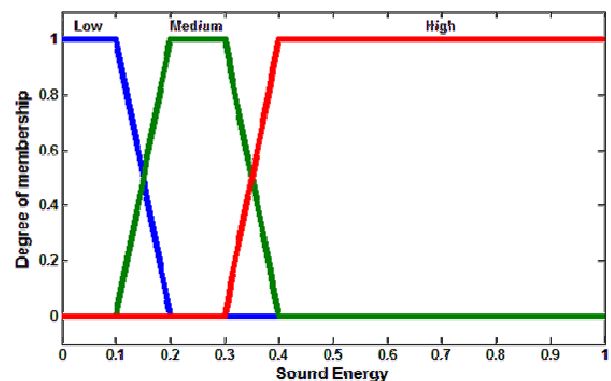
سیستم استنتاج فازی در روش پیشنهادی از روش ممدانی برای استنتاج از پایگاه قوانین فازی کمک می‌گیرد و خروجی فازی، بر اساس روش مرکز ثقل به خروجی غیرفازی تبدیل می‌گردد. روش استنتاج ممدانی که به روش Min-Max نیز مشهور است، پرکاربردترین روش استنتاج فازی می‌باشد. این روش در عین سادگی، روشی بسیار موثر و کارآمد در استنتاج فازی محسوب می‌گردد. در این روش، استلزام^۱ (بررسی میزان تطابق ورودی‌ها با قوانین فازی) بر اساس عملگر کمینه‌گیری و تجمیع^۲ (ترکیب نتایج



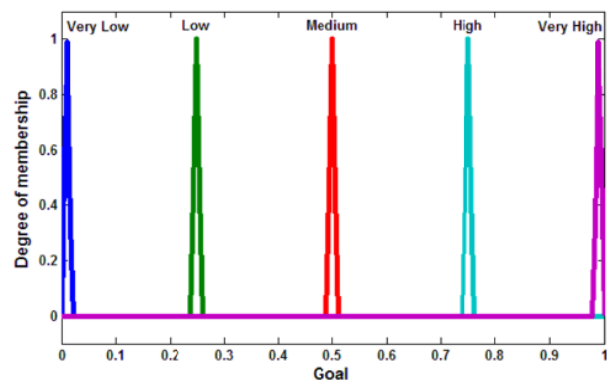
شکل ۵ تابع عضویت ورودی اول سیستم استنتاج فازی: مدت زمان پخش مجدد.



شکل ۶ تابع عضویت ورودی دوم سیستم استنتاج فازی: نسبت مدت زمان پخش نمای بسته و نمای خارج به کل مدت زمان واحد معنایی.



شکل ۷ تابع عضویت ورودی سوم سیستم استنتاج فازی: میانگین انرژی صوتی در واحد معنایی قبل از پخش مجدد.



شکل ۸ تابع عضویت خروجی سیستم استنتاج فازی: امکان وقوع گل.

¹Implication

²Aggregation

ابتدا/انتهای پخش مجدد کم‌رنگ می‌گردد. با این وجود، از میان ۳۳۴ پخش مجدد که در تمام ویدیوهای پایگاه داده وجود داشت، فقط ۵ مورد آشکارسازی نگردید و خطایی در آشکارسازی اشتباه رخ نداد. به این ترتیب، نرخ خطای آشکارسازی اشتباه صفر و نرخ خطای عدم آشکارسازی ۱,۵٪ می‌باشد.

جدول ۲ ویدیوهای مورد استفاده در پایگاه داده.

شماره ویدیو	توضیحات	مدت زمان (دقیقه)
۱	آلمان-آرژانتین (نیمه اول)	۵۸:۴۰
۲	آلمان-آرژانتین (نیمه دوم)	۴۹:۵۶
۳	آلمان-انگلیس (نیمه اول)	۵۸:۱۸
۴	آلمان-انگلیس (نیمه دوم)	۵۰:۵۱
۵	آلمان-اسپانیا (نیمه اول)	۵۷:۰۶
۶	آلمان-اسپانیا (نیمه دوم)	۵۱:۱۲
۷	آلمان-اروگوئه (نیمه اول)	۵۳:۴۲
۸	آلمان-اروگوئه (نیمه دوم)	۵۰:۰۴
۹	یونان-آرژانتین (نیمه اول)	۵۵:۲۸
۱۰	یونان-آرژانتین (نیمه دوم)	۵۰:۴۹
۱۱	اسلواکی-ایتالیا (نیمه اول)	۵۴:۵۵
۱۲	اسلواکی-ایتالیا (نیمه دوم)	۵۴:۱۳

۴-۲ ارزیابی تشخیص نوع نما

برای ارزیابی دقت روش پیشنهادی در تشخیص نوع نما، از ویدیو شماره ۱ و ۲ استفاده گردید. همان‌طور که گفته شد، روش پیشنهادی نوع نما را در دو دسته کلی طبقه‌بندی می‌نماید که ماتریس اغتشاش حاصل از آزمایش‌ها در جدول ۳ نشان داده شده است. با توجه به جدول ۳، نرخ متوسط تشخیص نوع نما ۹۲,۲٪ می‌باشد.

جدول ۳ نتایج ارزیابی روش پیشنهادی برای تشخیص نوع نما.

	نمای بسته / نمای خارج	نمای متوسط / نمای دور	نرخ دقت
نمای بسته / نمای خارج	۲۹۹	۲۷	۹۱,۷٪
نمای متوسط / نمای دور	۴۶	۵۶۳	۹۲,۴٪

علت اصلی بروز خطا در تشخیص نوع نما، تداخل میان (نمای متوسط-نمای بسته) و (نمای متوسط-نمای خارج) است. در حالت کلی نوع نما به چهار دسته (نمای دور، نمای متوسط، نمای بسته و نمای خارج) تقسیم‌بندی می‌شود. در این حالت، بیشترین تداخل بین کلاسی در میان نمونه‌های (نمای دور-نمای متوسط)، (نمای متوسط-نمای بسته)، (نمای متوسط-نمای خارج) و (نمای بسته-نمای خارج) بروز پیدا می‌کند. با توجه به این که در روش پیشنهادی، دو دسته نمای دور و نمای متوسط در هم ادغام شده‌اند، تداخل میان این دو کلاس خطایی ایجاد

حاصل از تمام قوانین) بر اساس عملگر بیشینه‌گیری انجام می‌گیرد. به همین دلیل، روش ممدانی به روش Min-Max مشهور است.

۳-۹ آشکارسازی گل

پس از این که سیستم استنتاج فازی، امکان وقوع گل را در هر واحد معنایی تعیین نمود، با اعمال حد آستانه T_G می‌توان گل را آشکارسازی نمود. در روش پیشنهادی، $T_G = 0.5$ برای آشکارسازی گل استفاده شده است. به عبارت دیگر، اگر امکان وقوع گل در یک واحد معنایی بیش از نصف باشد، در آن واحد معنایی گل اتفاق افتاده است.

۴ نتایج آزمایش‌ها

برای ارزیابی روش پیشنهادی، یک پایگاه داده شامل ۱۲ ویدیو از مسابقات فوتبال جام جهانی ۲۰۱۰ آفریقای جنوبی تهیه شد که مدت زمان آن حدود ۱۱ ساعت می‌باشد. توضیحات بیشتر در مورد هر ویدیو در جدول ۲ ارائه شده است.

وضوح کیفیت ویدیوهای این پایگاه داده 1280×720 پیکسل و نرخ قاب آنها ۲۵ قاب در ثانیه است. با توجه به ابعاد بزرگ قاب‌های ویدیویی این پایگاه داده، قبل از انجام آزمایش‌ها، ابعاد ویدیو به 320×180 کاهش یافت تا حجم محاسبات کمتر شود. در ادامه، به ارائه نتایج آزمایش‌ها و بحث در مورد آنها خواهیم پرداخت.

۴-۱ ارزیابی آشکارسازی لوگو و پخش مجدد

همان‌طور که قبلاً گفته شد، آشکارسازی لوگو با استفاده از درخت تصمیم انجام می‌گیرد. با توجه به اینکه درخت تصمیم نیاز به نمونه‌های آموزشی دارد، در مرحله آموزش از ۱۸۶ تصویر لوگو و ۲۰۰۰ تصویر غیرلوگو استفاده شد که تعداد نمونه‌های آموزشی منفی بیشتر از نمونه‌های آموزشی مثبت است. این موضوع باعث می‌شود تا در نهایت، نرخ خطای آشکارسازی اشتباه کمتر از نرخ خطای عدم آشکارسازی باشد. با توجه به اینکه الگوریتم آشکارسازی پخش مجدد بر اساس آشکارسازی لوگو در چند قاب متوالی عمل می‌کند، بنابراین آشکار نکردن لوگو نسبت به آشکارسازی اشتباه آن، تاثیر کمتر بر کاهش کارایی سیستم دارد. از این رو، تعداد نمونه‌های مثبت و منفی برای آموزش درخت تصمیم به این شکل انتخاب گردید. ضمن اینکه، نمونه‌های غیرلوگو تنوع بیشتری دارند و لازم است تا درخت تصمیم با تعداد کافی و متنوع از این نمونه‌ها آموزش ببیند، اما نمونه‌های آموزشی لوگو شباهت بیشتری به هم داشته و تنوع آنها کم است.

در مرحله آموزش درخت تصمیم، تمام نمونه‌های آموزشی بدون هیچ گونه خطایی آموزش دیدند. در مرحله آشکارسازی، با توجه به اینکه آشکارسازی نقطه ابتدا/انتهای پخش مجدد بر اساس آشکارسازی لوگو در ۳ قاب از ۵ قاب متوالی انجام می‌شود، خطای آشکارسازی لوگو تا حد زیادی در آشکارسازی نقطه

گل اول بازی آلمان-انگلیس (نیمه اول) از مواردی است که آشکارسازی نشده است. در این مورد، هیچ شاتی با نمای بسته بعد از وقوع گل و قبل از پخش مجدد نمایش داده نشده است. به همین علت، سیستم میزان امکان وقوع گل در آن واحد معنایی را پایین تر از حد آستانه تعیین نموده است. نمونه‌هایی از قاب‌های مربوط به شات پس از وقوع گل که در آن هیجان بازیکنان نشان داده شده است، در شکل ۱۰ نشان داده شده است. همان طور که در این شکل مشخص است، شات یک نمای متوسط از بازیکنان را نشان می‌دهد.



شکل ۱۰ نمایش برخی از قاب‌های مربوط به گل اول بازی آلمان-انگلیس (نیمه اول) که سیستم نتوانسته گل را آشکارسازی نماید.

مورد دیگری که در آن گل آشکارسازی نشده، مربوط به اولین گل بازی آلمان-ارگوت (نیمه اول) است. در این مورد، مدت زمان پخش مجدد حدود ۱۸ ثانیه (معادل ۴۵۰ قاب) است. در این حالت، مدت زمان پخش مجدد کوتاه تلقی شده و امکان وقوع گل در آن واحد معنایی کم ارزیابی شده است.

یکی از مواردی که سیستم به اشتباه آن را گل آشکارسازی کرده است، مربوط به بازی یونان-آرژانتین (نیمه دوم) است. در دقیقه ۲۶ این ویدیو، ضمن وقوع یک موقعیت خطرناک روی دروازه یونان، سانحه‌ای برای دروازه‌بان تیم یونان رخ می‌دهد که باعث می‌شود کارگردان برای این موقعیت خطرناک و سانحه رخ داده، پخش مجدد طولانی در نظر بگیرد. ضمن این که چندین بار نمای بسته بازیکنان و دروازه‌بان نمایش داده شده و هیاهوی تماشاچیان زیاد است. قاب‌هایی از این صحنه در شکل ۱۱ نشان شده است.



شکل ۱۱ نمایش برخی از قاب‌های مربوط به موقعیت خطرناک و صحنه حادثه بازی یونان-آرژانتین (نیمه دوم) که سیستم به اشتباه آن را گل آشکارسازی کرد.

در بازی اسلواکی-ایتالیا (نیمه دوم) بعد از وقوع یک گل، درگیری میان بازیکنان دو تیم ایجاد می‌شود. این رویداد که در

نمی‌کند. این موضوع برای تداخل بین نمای بسته و نمای خارج نیز صدق می‌کند. به این ترتیب تداخل میان نمای متوسط با نمای بسته و نمای خارج، عامل اصلی بروز خطا در روش پیشنهادی برای تشخیص نوع نما می‌باشد. برخی از قاب‌هایی که تشخیص نما در آنها به اشتباه انجام شده است، در شکل ۹ نشان داده شده است.



شکل ۹ برخی از قاب‌ها که الگوریتم تشخیص نما برای آنها با خطا مواجه شده است.

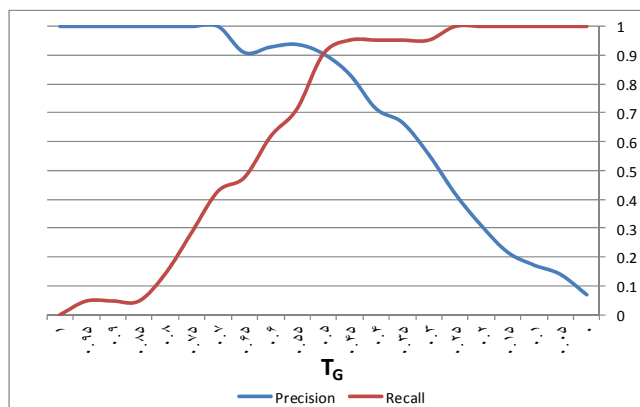
۴-۳ ارزیابی آشکارسازی گل

پس از این که ویدیو به واحدهای معنایی بخش‌بندی شد، امکان وقوع گل در هر واحد معنایی توسط سیستم استنتاج فازی در بازه [0,1] مشخص می‌شود. برای آشکارسازی گل، با اعمال حد آستانه $T_G = 0.5$ بر روی مقدار امکان وقوع گل، گل آشکارسازی می‌گردد. بر این اساس، نتایج آشکارسازی گل در ویدیوهای مختلف در جدول ۴ ارائه شده است.

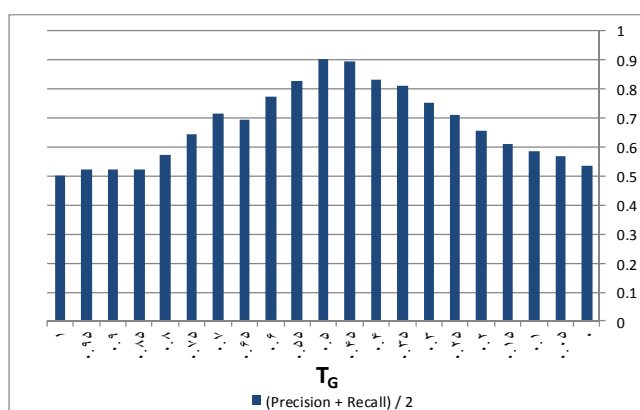
جدول ۴ نتایج آشکارسازی گل در ویدیوهای فوتبال.

شماره ویدیو	تعداد واقعی گل	تعداد عدم آشکارسازی گل	تعداد آشکارسازی اشتباه گل
۱	۱	۰	۰
۲	۳	۰	۰
۳	۳	۱	۰
۴	۲	۰	۰
۵	۰	۰	۰
۶	۱	۰	۰
۷	۲	۰	۰
۸	۳	۰	۰
۹	۰	۰	۱
۱۰	۲	۰	۱
۱۱	۱	۰	۰
۱۲	۴	۱	۰
جمع کل	۲۲	۲	۲

در مجموع، ۲۲ گل در این ویدیوها به ثمر رسیده است. بر اساس ارزیابی‌های انجام شده، با استفاده از $T_G = 0.5$ ، تمام گل‌ها به جز دو مورد آشکارسازی شده‌اند، و در دو مورد، آشکارسازی اشتباه گل صورت گرفته است. به این ترتیب، نرخ دقت و بازخوانی روش پیشنهادی ۹۰,۹٪ و ۹۰,۹٪ می‌باشد.



شکل ۱۳ نمودار تغییرات نرخ دقت و نرخ بازخوانی نسبت به تغییرات مقدار حد آستانه T_G .



شکل ۱۴ نمودار تغییرات میانگین نرخ دقت و نرخ بازخوانی نسبت به تغییرات مقدار حد آستانه T_G .

بیشتری انجام گیرد.

با مقایسه نتایج ارائه شده در جدول ۵ مشاهده می‌گردد که نتایج حاصل از روش پیشنهادی نسبت به سایر روش‌ها بهتر است. البته با توجه به این که روش ارائه شده در (Tjondronegoro et al. 2010)، یک ساختار کلی برای آشکارسازی رویدادهای مهم در مسابقات ورزشی مختلف ارائه می‌نماید، به نظر می‌رسد خطای بیشتر روش (Tjondronegoro et al. 2010) نسبت به سایر روش‌ها مربوط به استفاده از یک ساختار عمومی برای آشکارسازی گل در مسابقه فوتبال باشد.

جدول ۵ مقایسه نرخ دقت و بازخوانی روش پیشنهادی با سایر روش‌ها برای آشکارسازی گل در ویدیو مسابقات فوتبال.

نرخ بازخوانی	نرخ دقت	روش
۷۱,۴٪	۶۲,۵٪	روش ارائه شده در (Tjondronegoro et al. 2010)
۹۲,۰٪	۸۲,۳٪	روش ارائه شده در (Kolekar et al. 2009)
۹۸,۹٪	۷۸,۴٪	روش ارائه شده در (Huang et al. 2006)
۹۵,۲٪	۸۳,۳٪	روش پیشنهادی با $T_G = 0.45$
۹۰,۹٪	۹۰,۹٪	روش پیشنهادی با $T_G = 0.5$
۷۳,۴٪	۹۳,۷٪	روش پیشنهادی با $T_G = 0.55$

دقیقه ۳۶ این ویدیو حادث شده، آن قدر بحرانی می‌شود که برای دقایقی بازی متوقف می‌شود. از این رو کارگردان پخش مجدد طولانی از این اتفاق را نمایش داده و مکرراً نمای بسته از بازیکنان نشان می‌دهد. به همین دلیل سیستم به اشتباه این رویداد مهم را گل تلقی کرده است. قاب‌هایی از این رویداد در شکل ۱۲ نشان داده شده است.



شکل ۱۲ نمایش برخی از قاب‌های مربوط به صحنه حادثه و درگیری بازیکنان در بازی اسلواکی-ایتالیا (نیمه دوم) که سیستم به اشتباه آن را گل آشکارسازی کرده است.

در دو موردی که سیستم نتوانسته گل‌ها را آشکارسازی نماید، مدت زمان پخش مجدد و پخش نمای بسته کوتاه و میزان هیجان تماشاچیان کم بوده است. از این رو، سیستم امکان وقوع گل در این لحظات را کمتر از ۰,۵ تعیین نموده است. همچنین در مواردی که آشکارسازی گل به اشتباه انجام شده است، به دلایلی مانند حساس بودن صحنه و اهمیت رویدادی که اتفاق افتاده است، مدت زمان پخش مجدد و پخش نمای بسته بازیکنان طولانی بوده است. به این ترتیب، سیستم امکان وقوع گل در این لحظات را بیشتر از ۰,۵ تعیین کرده که این امر موجب بروز اشتباه شده است. برای بررسی بیشتر تاثیر مقدار حد آستانه T_G ، نمودار تغییرات نرخ دقت و نرخ بازخوانی برای مقادیر مختلف T_G در شکل ۱۳ رسم شده است. همچنین نمودار میانگین نرخ دقت و بازخوانی در شکل ۱۴ رسم شده است. برای دستیابی به بهترین حالت (زمانی که میانگین نرخ دقت و بازخوانی بیشینه است)، باید حد آستانه حدود ۰,۴۵ تا ۰,۵ انتخاب گردد. درحالت کلی، با توجه به کاربرد و بر اساس نمودارهای ارائه شده در شکل ۷ و ۸ می‌توان حد آستانه مناسب را برای سیستم تعیین نمود.

۴-۴ مقایسه با سایر روش‌ها

روش پیشنهادی با روش‌های ارائه شده در (Huang et al. 2006، Kolekar et al. 2009، Tjondronegoro et al. 2010) مورد مقایسه قرار گرفته است. نتایج این مقایسه در جدول ۵ ارائه شده است. در این مقایسه سعی شده تا نرخ دقت و بازخوانی به ازای مقادیر مختلف T_G ارائه شود تا مقایسه میان روش‌ها با سهولت

۴-۵ سرعت پردازش

الگوریتم پیشنهادی، با استفاده از نرم افزار Matlab R2011a پیاده سازی شده و بر روی یک کامپیوتر شخصی با پردازنده AMD 4200+ Dual Core (2.2 GHz)، حافظه 2 GB و سیستم عامل Microsoft Windows Seven 32 bit اجرا گردید. بر اساس نتایج بدست آمده، میانگین مدت زمان لازم برای پردازش هر قاب با ابعاد 180×320 حدود ۱۸۵ میلی ثانیه است. به این ترتیب، سرعت پردازش روش پیشنهادی در حدود ۵٫۵ قاب در ثانیه می باشد. شایان ذکر است که بیشتر حجم محاسبات مربوط به بخش محاسبه هیستوگرام سه بعدی است که بر اساس ارزیابی های انجام شده، به طور میانگین ۹۱٪ از حجم پردازش ها را به خود اختصاص می دهد.

۵ نتیجه گیری و پیشنهادات

در این مقاله، یک روش آشکارسازی گل در ویدئو پخش تلویزیونی مسابقات فوتبال ارائه گردید که پس از بخش بندی ویدئو در سطح شات و سطح معنایی، با استخراج ویژگی های سطح پایین و سطح میانی و به کمک سیستم استنتاج فازی، امکان وقوع گل در هر واحد معنایی را بررسی می کند.

سیستم پیشنهادی نیز همانند بسیاری از روش های دیگر، وابسته به دانش زمینه و قوانین اکتشافی است که توسط طراح سیستم در سیستم تعبیه می شود. تفاوت عمده روش پیشنهادی با سایر روش ها، استفاده از سیستم استنتاج فازی به عنوان یک ابزار قدرتمند برای کاربرد مدل فازی و بیان قوانین اکتشافی در قالب قوانین ساده و قابل فهم است. به عبارت دیگر، استفاده از سیستم استنتاج فازی در روش پیشنهادی، دو مزیت به همراه داشته است. اول این که دانش زمینه به شکلی ساده و قابل فهم در قالب پایگاه قوانین فازی ارائه می شود. دوم این که، با تکیه بر کارایی و قدرت سیستم استنتاج فازی، تصمیم گیری در مورد امکان وقوع گل در یک واحد معنایی انجام می گیرد. این روش، علاوه بر اینکه دقت بسیار خوبی دارد، از نظر محاسباتی نیز چندان پیچیده نیست.

در مقایسه ای که میان کارایی سیستم استنتاج فازی و روش های مبتنی بر آستانه گیری برای آشکارسازی گل انجام شد، مشخص گردید که نرخ دقت و بازخوانی سیستم استنتاج فازی حدود ۵٪ بیشتر از بهترین حالت از روش های آستانه گیری است. هرچند با تنظیم دقیق مقادیر حد آستانه می توان دقت روش آستانه گیری را به روش مبتنی بر سیستم استنتاج فازی نزدیک کرد، اما تنظیم دقیق چند حد آستانه کار ساده ای نیست، به ویژه اگر تعداد ورودی ها برای آستانه گیری افزایش یابد. به عبارت دیگر، دقت و بازخوانی روش های مبتنی بر آستانه گیری حساسیت زیادی به مقادیر حد آستانه تعیین شده دارد. اما در روش مبتنی بر استنتاج فازی، به دلیل استفاده از خواص مجموعه های فازی که می تواند ابهام در ورودی ها را مدل سازی کند، دقت و بازخوانی سیستم حساسیت کمتری به تعریف و تنظیم قوانین فازی و توابع عضویت دارد.

در روش پیشنهادی، از سیستم استنتاج فازی برای تعیین امکان وقوع گل در هر بخش از ویدئو استفاده می شود. سپس با اعمال حد آستانه، رویداد گل آشکارسازی می گردد. جهت تبیین دقیق تر کارایی سیستم استنتاج فازی در روش پیشنهادی، مقایسه ای میان سیستم استنتاج فازی و روش های مبتنی بر آستانه گیری انجام شد. در این مقایسه، ویژگی های استخراج شده از ویدئو کاملاً یکسان است و تفاوت روش ها در نحوه تصمیم گیری وقوع گل می باشد. برای انجام این مقایسه، مقادیر مختلف حد آستانه برای ورودی های مختلف مورد آزمایش قرار گرفت. سعی شد مقادیر حد آستانه مورد بررسی، با مقادیر مورد استفاده در ساخت توابع عضویت فازی در سیستم استنتاج فازی مشابهت داشته باشد. نتایج این آزمایش ها در جدول ۶ نشان داده شده است. در جدول ۶، حدود آستانه ای که منجر به بهترین نتیجه در روش های مبتنی بر آستانه گیری می شود، نیز ذکر شده است.

همانطور که در جدول ۶ ملاحظه می شود، نتایج بدست آمده از روش های مبتنی بر آستانه گیری، جز با تنظیم دقیق حدود آستانه، چندان مناسب نیست. بنابراین هرچند با استفاده از روش های مبتنی بر آستانه گیری و تنظیم دقیق حدود آستانه می توان متوسط نرخ دقت و بازخوانی را تا ۸۶٪ نیز افزایش داد، اما متوسط نرخ دقت و بازخوانی با استفاده از سیستم استنتاج فازی در بهترین حالت حدود ۵٪ بیشتر است.

جدول ۶ مقایسه نرخ دقت و بازخوانی روش مبتنی بر سیستم استنتاج فازی با روش های مبتنی بر آستانه گیری برای آشکارسازی وقوع گل.

روش تصمیم گیری	حد آستانه بر روی ورودی اول	حد آستانه بر روی ورودی دوم	حد آستانه بر روی ورودی سوم	نرخ دقت	نرخ بازخوانی	متوسط نرخ دقت و بازخوانی
مبتنی بر آستانه گیری	۳۰۰	-	-	۱۶٫۵٪	۱۰۰٪	۵۸٫۳٪
	۴۰۰	-	-	۲۴٫۷٪	۱۰۰٪	۶۲٫۴٪
	۵۰۰	-	-	۳۷٫۹٪	۱۰۰٪	۶۷٫۰٪
	۶۰۰	-	-	۶۶٫۷٪	۹۰٫۹٪	۷۸٫۸٪
	۷۰۰	-	-	۸۴٫۲٪	۷۲٫۷٪	۷۸٫۵٪
	۸۰۰	-	-	۹۱٫۷٪	۵۰٫۰٪	۷۰٫۸٪
	۹۰۰	-	-	۱۰۰٪	۳۶٫۴٪	۶۸٫۲٪
	۳۰۰	۰٫۳	-	۲۴٫۱٪	۹۰٫۹٪	۵۷٫۵٪
	۳۰۰	۰٫۷	-	۴۰٫۰٪	۲۷٫۳٪	۳۳٫۷٪
	۶۰۰	۰٫۳	-	۸۵٫۷٪	۸۱٫۸٪	۸۳٫۸٪
	۶۰۰	۰٫۷	-	۱۰۰٪	۲۲٫۷٪	۶۱٫۴٪
	۳۰۰	۰٫۳	۰٫۳۵	۳۳٫۹٪	۸۶٫۴٪	۶۰٫۲٪
	۳۰۰	۰٫۷	۰٫۳۵	۵۵٫۶٪	۲۲٫۷٪	۳۹٫۱٪
	۶۰۰	۰٫۳	۰٫۳۵	۸۹٫۵٪	۷۷٫۳٪	۸۳٫۴٪
	۶۰۰	۰٫۷	۰٫۳۵	۱۰۰٪	۱۸٫۲٪	۵۹٫۱٪
سیستم استنتاج فازی	۷۰۰	۰٫۴	۰٫۲۵	۱۰۰٪	۷۲٫۷٪	۸۶٫۴٪
	روش پیشنهادی با $T_G = 0.45$			۸۳٫۳٪	۹۵٫۲٪	۸۹٫۳٪
	روش پیشنهادی با $T_G = 0.5$			۹۰٫۹٪	۹۰٫۹٪	۹۰٫۹٪
	روش پیشنهادی با $T_G = 0.55$			۹۳٫۷٪	۷۳٫۴٪	۸۳٫۶٪

- Image Understanding vol. 113, no. 5, pp. 622-632, 2009.
- [4] Figueroa, P. J., Leite, N. J. and Barros, R. M. L. "Tracking soccer players aiming their kinematical motion analysis", Computer Vision and Image Understanding, vol. 101, no. 2, pp. 122-135, 2006.
- [5] Khatoonabadi, H. S. and Rahmati, M. "Automatic soccer players tracking in goal scenes by camera motion elimination", Image and Vision Computing, vol. 27, pp. 469-479, 2009.
- [6] Miura, J., Shimawaki, T.; Sakiyama, T. and Shirai, Y. "Ball route estimation under heavy occlusion in broadcast soccer video", Computer Vision and Image understanding, vol. 113, no. 5, pp. 653-662, 2009.
- [7] Ekin, A., Tekalp, A. M. and Mehrotra, R. "Automatic soccer video analysis and summarization", IEEE Transactions on Image Processing, vol. 12, no. 7, pp. 796-807, 2003.
- [8] Lotfi, E. and Pourreza, H. R. "Summarizing soccer videos without detecting the events", International Journal of Computer Science and Network Security, vol. 6, no. 10, pp. 130-135, 2006.
- [9] Lotfi, E.; Akbarzadeh-T, M. R.; Pourreza, H. R.; Yaghoubi, M. and Dehghan, F. "A hybrid method for soccer video events retrieval using fuzzy systems", Joint International Fuzzy Systems Association World Congress and European Society of Fuzzy Logic and Technology Conference, Lisbon, Portugal, pp. 381-385, 2009.
- [10] Lotfi, E. and Yaghoubi, M. "Concept retrieval based on a combination of fractal coding, fuzzy rule based system and SVM", Fractals, vol. 19, no. 2, pp. 185-194, 2011.
- [11] Money, A. G. and Agius, H. "Video summarisation: a conceptual framework and survey of the state of the art", Journal of Visual Communication and Image Representation, vol. 19, no. 2, pp. 121-143, 2008.
- [12] Amiri, A. and Fathy, M. "Video shot boundary detection using QR-decomposition and gaussian transition detection", EURASIP Journal on Advances in Signal Processing, 2009.
- [۱۳] امینیان مدرس، امیرفرید؛ زینالپور تبریزی، زینب و سریان، محسن، «تشخیص مرز شات در ویدیو با استفاده از تئوری بیز و فیلتر همسایگی»، چهاردهمین کنفرانس ملی انجمن کامپیوتر ایران، تهران، ایران، ۱۳۸۷.
- [14] Wang, L., Zeng, B.; Lin, S.; Xu, G. and Shum, H.-Y. "Automatic extraction of semantic colors in sports video", International Conference on Acoustics, Speech, and Signal Processing, pp. 617-620, 2004.
- [15] Kolekar, M. H., Palaniappan, K.; Sengupta, S. and Seetharaman, G. "Semantic concept mining based on hierarchical event detection for soccer video indexing", Journal of Multimedia, vol. 4, no. 5, pp. 298-312, 2009.
- [16] Chen, H.-T.; Tien, M.-C.; Chen, Y.-W.; Tsai, W.-J. and Lee, S.-Y. "Physics-based ball tracking and 3D

عمده مشکل روش پیشنهادی، باقی ماندن بخش عمده‌ای از شکاف معنایی میان ویژگی‌های سطح پایین و مفاهیم سطح بالا است. به عبارت دیگر، اگر چه استفاده از ویژگی‌های سطح میانی مانند آشکارسازی پخش مجدد و تشخیص نوع نمای شات، تا حد زیادی می‌تواند در آشکارسازی گل موثر و مفید باشد، اما استفاده از این ویژگی‌ها کافی نیست. چرا که استفاده صرف از این ویژگی‌ها نمی‌تواند محتوای ویدیو را به درستی توصیف کند. همین عامل باعث بروز خطا در آشکارسازی اشتباه یا عدم آشکارسازی گل می‌شود. در ادامه پژوهش‌ها می‌توان با استخراج ویژگی‌های دیگر مانند آشکارسازی دروازه، دامنه ویژگی‌های استخراج شده را گسترش داد تا بتوان با دقت بهتری آشکارسازی گل را انجام داد. همچنین استفاده از سیستم استنتاج تطبیقی فازی-عصبی^۱ (ANFIS) می‌تواند با ارائه یک راهکار مبتنی بر یادگیری، نیاز سیستم را به استفاده از قوانین اکتشافی کمتر نماید. به عبارت دیگر، سیستم استنتاج تطبیقی فازی-عصبی می‌تواند قوانین فازی مناسب را بر اساس تعداد کافی از نمونه آموزشی استخراج نماید. برای کارهای آینده پیشنهاد می‌شود، سیستمی برای آشکارسازی سایر رویدادهای مهم فوتبال مانند پنالتی و کارت زرد یا قرمز طراحی شود. البته طراحی چنین سیستم‌هایی نیازمند استخراج ویژگی‌های دیگری از جمله آشکارسازی محوطه جریمه، آشکارسازی داور و تشخیص سوت داور است تا بتوان بر اساس این ویژگی‌ها، آشکارسازی رویدادهای مورد نظر را انجام داد.

قدردانی

نویسندگان مقاله از سرکار خانم محبوبه کهخائی جوان بابت نظرات ایشان در ویرایش ادبی مقاله سپاس‌گزاری می‌نمایند. همچنین لازم است از آقای وحید کیانی بابت همکاری‌شان در تهیه پایگاه داده تقدیر گردد.

مراجع

- [1] Ye, Q., Huang, Q.; Gao, W. and Jiang, S. "Exciting event detection in broadcast soccer video with mid-level description and incremental learning", ACM international conference on Multimedia, Singapore, pp. 455-458, 2005.
- [2] Han, Y., Liu, G. and Chollet, G. "Goal event detection in broadcast soccer videos by combining heuristic rules with unsupervised fuzzy C-means algorithm", International Conference on Control, Automation, Robotics and Vision, Hanoi, Vietnam, pp. 888-891, 2008.
- [3] D'Orazio, T., Leo, M.; Spagnolo, P.; Nitti, M. and Mosca, N. "A visual system for real time detection of goal events during soccer matches", Computer Vision and

¹Adaptive Neuro-Fuzzy Inference System (ANFIS)

محمدحسین سیگاری مدرک کارشناسی و کارشناسی ارشد خود را در رشته مهندسی کامپیوتر، به ترتیب در سال ۱۳۸۵ و ۱۳۸۷ از دانشگاه فردوسی مشهد و دانشگاه علم و صنعت ایران با کسب رتبه اول دریافت نمود.



هم اکنون ایشان دانشجوی مقطع دکتری تخصصی در رشته مهندسی کامپیوتر، گرایش هوش مصنوعی در دانشکده مهندسی برق و کامپیوتر دانشگاه تهران می باشد. علاقه مندی های علمی ایشان شامل پردازش تصویر، بینایی ماشین، شناسایی الگو، بیومتریک، سیستم های تحلیل معنایی ویدیو و کاربردهای پردازش تصویر در صنعت می باشد.

حمید سلطانیانزاده مدرک کارشناسی پیوسته خود را در رشته مهندسی برق گرایش الکترونیک در سال ۱۳۶۵ از دانشگاه تهران با کسب رتبه اول دریافت نمود. سپس ایشان مدارک کارشناسی ارشد و دکتری خود را در رشته مهندسی برق به ترتیب در گرایش های کنترل و پردازش سیگنال و بیوالکتریک در سال های ۱۳۶۹ و ۱۳۷۱ از دانشگاه میشیگان آمریکا دریافت نمود. هم اکنون ایشان استاد دانشکده مهندسی برق و کامپیوتر دانشگاه تهران و محقق ارشد گروه رادیولوژی بیمارستان هنری فورد آمریکا می باشد. علاقه مندی های علمی ایشان شامل تصویربرداری پزشکی، پردازش تصاویر و سیگنال های پزشکی، پردازش تصویر، بینایی ماشین و شناسایی الگو می باشد.



حمیدرضا پوررضا مدرک کارشناسی خود را در رشته مهندسی برق گرایش الکترونیک در سال ۱۳۶۸ از دانشگاه فردوسی مشهد کسب کرد. سپس ایشان مدارک کارشناسی ارشد و دکتری خود را به ترتیب در رشته مهندسی برق گرایش



الکترونیک و مهندسی کامپیوتر گرایش هوش مصنوعی در سال های ۱۳۷۲ و ۱۳۸۲ از دانشگاه امیرکبیر دریافت نمود. هم اکنون ایشان دانشیار گروه مهندسی کامپیوتر دانشکده مهندسی دانشگاه فردوسی مشهد می باشد. علاقه مندی های علمی ایشان پردازش تصویر، بینایی ماشین، شناسایی الگو و سیستم های حمل و نقل هوشمند می باشد.

trajectory reconstruction with applications to shooting location estimation in basketball video", Journal of Visual Communication & Image Representation, vol. 20, pp. 204-216, 2009.

- [17] Hosseinkhani, J., Soltanian-Zadeh, H.; Kamarei, M. and Staadt, O. "Ball detection with the aim of corner event detection in soccer video", International Symposium on Parallel and Distributed Processing with Applications Workshops Busan, South Korea, pp. 147-152, 2011.
- [18] Chen, S.-C.; Shyu, M.-L.; Chen, M. and Zhang, C. "A decision tree-based multimodal data mining framework for soccer goal detection", International Conference on Multimedia and Expo, Taipei, Taiwan, pp. 265-268, 2004.

[۱۹] کنگرلو، کاوه و کبیر، احسان ...، «آشکارسازی بازیکنان در تصاویر فوتبال»، چهارمین کنفرانس ماشین بینایی و پردازش تصویر ایران، مشهد، ایران، ۱۳۸۵.

- [20] Huang, C.-L., Shih, H.-C. and Chao, C.-Y. "Semantic analysis of soccer video using dynamic bayesian network", IEEE Transactions on Multimedia, vol. 8, no. 4, pp. 749-760, 2006.
- [21] Tjondronegoro, D. W. and Chen, Y.-P. P. "Knowledge-discounted event detection in sports video", IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, vol. 40, no. 5, pp. 1009-1024, 2010.

[۲۲] بوالی، امیر، مقدمه ای بر زیباشناسی و کاربردهای دراماتیک انواع انتقال های بصری، تهران، ایران، معاونت آموزش و پژوهش سازمان صدا و سیما، ۱۳۸۶.

- [23] Saoudi, A. and Essafi, H. "Spatio-temporal video slice edges analysis for shot transition detection and classification", World Academy of Science, Engineering and Technology, vol. 28, pp. 45-50, 2007.