

طبقه‌بند خودسازمانده هندسی مبتنی بر یادگیری فعال برای نهان‌کاوی در محیط ویدئو با صرف حداقل برچسب

علی محی‌الدینی شاهره‌آبادی‌پور، مرتضی خادمی و هادی صدوقی یزدی

می‌شود و الگوریتمی مشخص، در سیگنال میزبان مخفی نموده و برای گیرنده ارسال می‌کند. گیرنده نیز با استفاده از کلید پیغام را آشکار می‌نماید. در مقابل، نهان‌کاوی فرایندی است که در آن فرد نهان‌کاو با بررسی خط ارتباطی به دنبال آشکارسازی وجود اطلاعات در این ارتباط است. الگوریتم‌هایی که در نهان‌کاوی مورد استفاده قرار می‌گیرد در یک دسته‌بندی کلی به دو دسته نهان‌کاوی کور و خاص تقسیم می‌شود. نهان‌کاوی کور به دنبال کشف وجود اطلاعات موجود در یک سیگنال مشکوک می‌باشد، با فرض این که از الگوریتم پنهان‌کننده، اطلاعی در دست نباشد (مانند [۳] تا [۶]) اما در نهان‌کاوی خاص (مانند [۷]) الگوریتم پنهان‌کننده اطلاعات مشخص است.

نهان‌کاوی‌ای که در این تحقیق مورد استفاده قرار می‌گیرد از نوع کور است و در صورت تشخیص وجود اطلاعات حداکثر ارتباط دو طرف قطع می‌شود. فایل میزبان مورد استفاده ویدئو است، در حالت کلی اطلاعات را می‌توان مانند [۲] در ماکروبلوک‌های (مثل ضرایب DCT^1 یا DWT^2) هر فریم پنهان نمود یا مانند [۸] در بردارهای حرکت که در اینجا فرض می‌شود اطلاعات در بردارهای حرکت در حین فشرده‌سازی توسط استانداردهای موجود پنهان شده است.

سیستم نهان‌کاوی: سیستمی که برای نهان‌کاوی مورد استفاده قرار می‌گیرد دارای سه بخش می‌باشد که در شکل ۲ دیده می‌شود. در ادامه این بخش به توضیح اجزای آن پرداخته می‌شود.

تخمین ویدئوی میزبان: در ابتدا با شیوه‌های متفاوت فایل ویدئوی عاری از اطلاعات استخراج می‌شود. در تحقیقات صورت‌پذیرفته برای تخمین ویدئوی میزبان روش‌های متفاوتی ارائه شده است.

در [۴] Kundur, Budhia و Zourntos اطلاعات پنهان‌شده را به صورت نویز مدل نمودند، در این حالت برای تخمین سیگنال میزبان از عمل میانگین‌گیری استفاده شده است. این عمل سبب می‌شود اطلاعاتی که وارد سیگنال شده و در هر فریم شبیه نویز عمل می‌نماید حذف شده و فریم اصلی بازسازی شود.

در [۶] که توسط Pankajakshan و Ho ارائه شد، از پیش‌بینی زمانی برای تخمین فریم میزبان استفاده شده است. این مقاله در ادامه کار انجام‌شده Xuan و همکاران در [۹] و Shi و سایرین در [۱۰] بوده است. پس از آن که Xuan و همکاران در [۹] یک نهان‌کاو برای تصویر ارائه دادند Shi و همکارانش نشان دادند با استفاده از پیش‌بینی مکانی می‌توان تداخل بردار ویژگی‌ها را کاهش داد. Pankajakshan و Ho برای تعمیم‌دادن این نهان‌کاو به حوزه ویدئو علاوه بر استفاده از پیش‌بینی مکانی تصویر (تخمین فریم میزبان) و تفاضل آن از تصویر اصلی که سبب ایجاد کاهش تداخل بردار ویژگی شده، از همبستگی زمانی بین فریم‌های

چکیده: طبقه‌بند یکی از سه بلوک تشکیل‌دهنده یک نهان‌کاو ویدئو است که برای آموزش نیازمند برچسب می‌باشد. در نهان‌کاوی کور به دلیل عدم دسترسی به الگوریتم‌های نهان‌نگاری تهیه برچسب مشکل است. در این مقاله از طبقه‌بند خودسازمانده پویای شبه‌ناظر برای رسیدن به حداقل برچسب استفاده شده و بدین منظور مفهومی به نام افزونگی هندسی گره‌های لایه زیرین شبکه خودسازمانده پویای شبه‌ناظر به کار گرفته شده است. نشان داده شده که این افزونگی منجر به ایجاد الگوهای تکراری برای شبکه خواهد شد، پس حذف چنین گره‌هایی بالمانع است. اثبات شده به دلیل وجود تناظر یک به یک بین گره‌ها و برچسب‌ها کاهش گره‌ها منجر به کاهش تعداد برچسب لازم می‌شود. نکته اساسی این که لازمه وجود افزونگی هندسی در میان تعدادی گره که مفهومی انتزاعی است، تشکیل دسته توسط آنهاست و بنابراین مبنای الگوریتم پیشنهادی شناسایی دسته‌ها و ادغام اعضای آنهاست. طبقه‌بند به دست آمده بر این مبنای طبقه‌بند خودسازمانده هندسی نام نهاده شده و اثبات می‌شود که این طبقه‌بند می‌تواند به مقدار بهینه حداقل برچسب دست یابد. نتایج شبیه‌سازی نشان‌دهنده برتری چشم‌گیر طبقه‌بند نسبت به الگوریتم‌های پیشین است.

کلیدواژه: طبقه‌بند خودسازمانده پویای شبه‌ناظر، نهان‌کاوی کور ویدئو، یادگیری شبه‌ناظر، یادگیری فعال.

۱- مقدمه

با پیشرفت علم روش‌های جدیدی برای مخفی کردن اطلاعات به صورت نهان‌نگاری برای انتقال از نقطه‌ای به نقطه دیگر به وجود آمده است. برای نهان‌نگاری اطلاعات نیاز به یک سیگنال میزبان است تا بتوان اطلاعات را در آن جاسازی کرد. اغلب در فضای مجازی از تصویر، صوت یا ویدئو استفاده می‌شود.

امنیت و ظرفیت دو عامل مهم در نهان‌نگاری به حساب می‌آیند و بر همین اساس ویدئو به دلیل داشتن حجم زیاد به عنوان یک واسط مهم به حساب می‌آید [۱] چرا که این توانایی را دارد تا بتوان حجم بالایی از اطلاعات را در آن پنهان نمود (ظرفیت زیاد). برای مثال در [۲] یک ویدئو به عنوان اطلاعات در یک ویدئوی دیگر پنهان شده است.

در شکل ۱ ساختار کلی سیستم نهان‌نگاری و نهان‌کاوی مشاهده می‌شود. سیستم نهان‌نگاری را می‌توان این گونه توضیح داد: فرستنده پیغام را از طریق یک کلید که بین فرستنده و گیرنده به اشتراک گذارده

این مقاله در تاریخ ۲۰ شهریور ماه ۱۳۹۵ دریافت و در تاریخ ۲۰ دی ماه ۱۳۹۶ بازنگری شد.

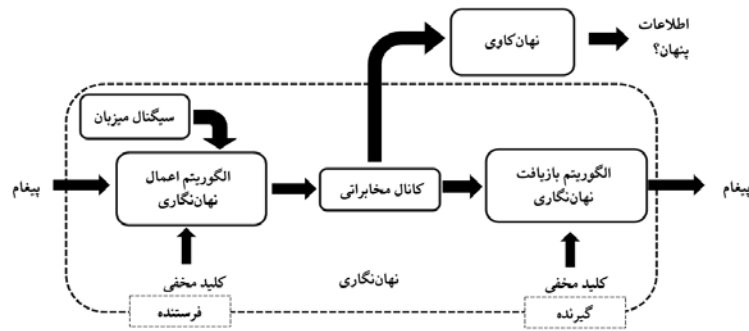
علی محی‌الدینی شاهره‌آبادی‌پور، دانشکده مهندسی، دانشگاه شهید باهنر، کرمان، (email: alimohiaddiny@gmail.com).

مرتضی خادمی، دانشکده مهندسی، دانشگاه فردوسی مشهد، مشهد، (email: khademi@um.ac.ir).

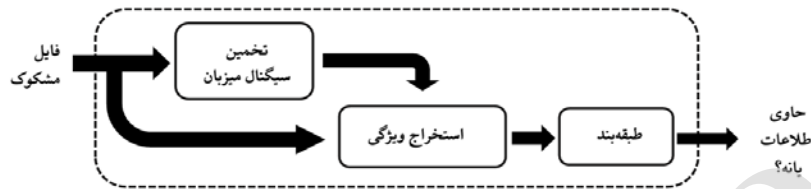
هادی صدوقی یزدی، دانشکده مهندسی، دانشگاه فردوسی مشهد، مشهد، (email: h-sadoghi@um.ac.ir).

1. Descrete Cosine Transform

2. Descrete Wavelet Transform



شکل ۱: ساختار کلی سیستم پنهان‌نگاری و پنهان‌کاری.



شکل ۲: ساختار کلی سیستم پنهان‌کاری.

واقع این تساوی به عنوان یک ویژگی مورد استفاده قرار گرفته است). بر این اساس در [۱۳] برای رسیدن به دقت بالاتر هم از همبستگی زمانی بردارهای حرکت و هم همبستگی مکانی آنها استفاده شده است، بر خلاف [۱۲] که تنها از همبستگی زمانی بردارهای حرکت بهره برده است. به همین دو دلیل بیان شده (استفاده از ویژگی‌های آمارگان مرتبه بالا و در نظر گرفتن همبستگی زمانی و مکانی بردارهای حرکت توأمان) این مرجع نسبت به [۱۲] و [۱۴] به دقت بالاتری دست یافته است. لازم به توضیح است در [۱۳] برای پنهان‌نگاری به جهت محاسبه‌ی دقت از الگوریتم پیشنهادی Aly در [۱۶] استفاده شده است. این الگوریتم بر خلاف الگوریتم‌های قبلی پنهان‌نگاری مثل [۱۷] و [۱۸] که برای پنهان کردن اطلاعات معیار اندازه را در نظر گرفته‌اند معیار خطا را برای پنهان کردن اطلاعات در نظر گرفته است. در واقع تا پیش از این فرض بر این بوده که هرچه اندازه بردار حرکت بزرگ‌تر باشد، میزان اعوجاج بالاتر است و به همان نسبت همبستگی بین بلوک‌های تصویر کاهش خواهد یافت، بنابراین پنهان کردن اطلاعات در چنین بردار حرکتی قابل شناسایی نبود. اما Aly نشان داد که گاهی بردارهای حرکت بزرگ میزان اعوجاج تولیدی کمتری نسبت به بردارهای حرکت کوچک داشته و معیار میزان خطای پیش‌بینی قابلیت اطمینان بیشتری در مقابل کشف شدن خواهد داشت. از این رو کشف وجود اطلاعات پنهان شده در سیگنالی که با روش به کار گرفته شده در [۱۶] پنهان شده باشد سخت‌تر است. نقطه قوت الگوریتم معرفی شده در [۱۳] نسبت به [۱۲] و [۱۴] در همینجاست.

در [۱۱] و [۱۲] با توجه به اختلاف بردارهای حرکت میزبان و بردارهای حرکت موجود سه نوع ویژگی استخراج شده است. ویژگی نوع اول مربوط به اختلاف اندازه بردارهای حرکت میزبان و موجود، ویژگی نوع دوم به اختلاف میزان خطایی که این دو نوع بردار تولید می‌کنند، مربوط است و ویژگی نوع سوم ترکیبی از دو ویژگی قبلی است. بر این مبنا یک بردار ویژگی ۱۵ بعدی استخراج شده که ۵ بعد اول مربوط به ویژگی نوع اول و به همین ترتیب ۵ بعد دوم و سوم متعلق به ویژگی نوع دو و سه بوده است. در این تحقیق نیز از همین ویژگی‌ها برای پنهان‌کاری استفاده شده است.

طبقه‌بند: پس از آن که ویژگی‌های مطلوب استخراج شد فایلی که از اختلاف ویدئوی موجود و تخمین ویدئوی میزبان به دست می‌آید به یک طبقه‌بند داده می‌شود. طبقه‌بند که پیش از این بر اساس ویژگی‌های

همسایه نیز که می‌تواند موجب تخمین فریم بعدی و در نتیجه کاهش بیشتر تداخل بردار ویژگی‌ها شود استفاده کردند.

روش بعدی برای تخمین سیگنال میزبان مربوط به حوزه بردارهای حرکت می‌باشد. در [۱۱] و [۱۲] برای تخمین بردار حرکت میزبان از روش فشرده‌سازی مجدد فایل ویدئویی استفاده شده است. روش کار بدین صورت است که فایل ویدئویی فشرده شده، وافشرده می‌شود و سپس با همان الگوریتم سابق (در اینجا MPEG۲) مجدداً فشرده می‌شود. Cao, Feng و Zhao در [۱۰] با استفاده از روابط ریاضی ثابت کردند که می‌توان برای تخمین بردارهای حرکت میزبان از این روش استفاده کرد. در این تحقیق نیز برای تخمین بردارهای حرکت میزبان از همین روش استفاده می‌شود.

استخراج ویژگی: پس از آن که در بلوک اول موفق به تخمین ویدئوی میزبان شدیم، باید در مرحله دوم بر اساس اختلاف بین ویدئوی میزبان و موجود، ویژگی‌هایی را استخراج کنیم که اختلاف ویدئوی حاوی اطلاعات را با ویدئوی میزبان به خوبی نشان دهد. میزان موفقیت طبقه‌بند در شناسایی فایل حاوی اطلاعات از غیر آن به انتخاب ویژگی‌هایی بستگی دارد که بهتر بتواند اختلاف این دو نوع ویدئو را نشان دهد.

در [۴] با توجه به این که پنهان‌نگاری طیف گسترده دارای توزیع گوسی است، برای استخراج ویژگی از ویژگی‌های آنتروپی، کشیدگی و ۲۵امین صدک استفاده گردیده و میزان گوسی بودن سیگنال بر اساس این سه ویژگی سنجیده شده است.

در [۶] ابتدا تصویر به ۳ سطح وولت گسسته مکانی تبدیل و ویژگی‌ها از هر بخش استخراج شده است.

در [۱۳] Zhou, Duan, Wu, Deng برای پنهان‌کاری بردارهای حرکت از ویژگی‌های آمارگان مرتبه بالا نسبت به کار انجام شده در [۱۴] استفاده نمودند. با توجه به این نکته که استفاده بهینه از همبستگی موجود در سیگنال میزبان منجر به کشف وجود اطلاعات پنهان شده خواهد گردید، قابل ذکر است که در [۱۵] برای کشف وجود اطلاعات پنهان شده در بردارهای حرکت، از همبستگی بین بردارهای حرکت فریم‌های نوع P (که بر اساس فریم I یا P قبل خود به دست می‌آید) و B (که بر اساس فریم I یا P قبل و بعد خود به دست می‌آید) استفاده شده و به یک رابطه (تساوی) دست یافته است. این تساوی هنگام تزریق اطلاعات از بین رفته که بر اساس آن به وجود اطلاعات پی برده شده است (در

۲- طبقه‌بند خودسازمانده هندسی مبتنی بر یادگیری فعال

با توجه به این که طبقه‌بند خودسازمانده هندسی (GeSOM)^۴ از طریق بازطراحی SSGSOM به دست آمده است ابتدا به معرفی اجمالی SSGSOM پرداخته می‌شود، سپس ایده یادگیری فعال به طور مختصر معرفی خواهد شد تا پس از آن طرح ایده اصلی مقاله مطرح شود. شبکه خودسازمانده پویای شبه‌ناظر ارتقایافته شبکه GSOM^۵ است. شبکه GSOM که توسط Alahakoon و Halgamunge در سال ۲۰۰۰ ارائه شد کاملاً به صورت بدون ناظر است [۱۹]. Hsu و Halgamunge در سال ۲۰۰۸ نسخه شبه‌ناظر این شبکه را با عنوان شبکه خودسازمانده پویای شبه‌ناظر ارائه کردند [۲۰]. در این شبکه مطابق شکل ۳، طبقه‌بند دارای دو لایه است. هر گره در لایه GSOM (لایه زیرین) دارای تابع گوسی به صورت (۱) است

$$o_c = \exp \frac{-\|x(t) - w_c(t)\|^2}{\sigma_c^2} \quad (1)$$

در این رابطه $w_c(t)$ بردار وزن لایه GSOM و σ_c محدوده فعالیت گره شماره c است که به صورت میانگین فاصله‌های هر گره با کلیه گره‌هایی که در همسایگی مستقیم آن قرار دارد، تعریف می‌شود. این توابع گوسی از طریق وزن‌های $w_{c1}, w_{c2}, \dots, w_{cm}$ به m گره خروجی که هر یک نشان‌دهنده یک کلاس هستند متصل شده‌اند. در شبکه خودسازمانده پویای شبه‌ناظر، در حین عملیات یادگیری شبکه خودسازمانده پویا (که به صورت بدون ناظر انجام می‌شود)، عملیات یافتن وزن‌های بهینه که لایه GSOM را به لایه خروجی متصل می‌نمایند با یک روند شبه‌ناظر به صورت (۲) به تکرار λ بار آموزش داده می‌شود

$$w_{cm}(t+1) = w_{cm}(t) + \eta(\xi_m - PS_m)o_c \quad (2)$$

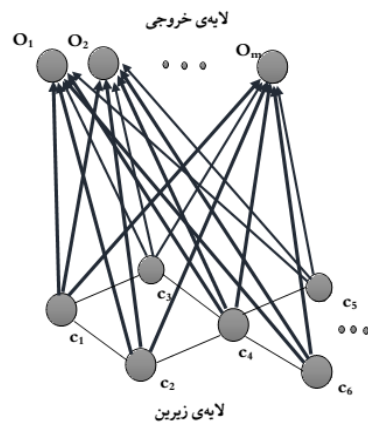
در این رابطه پارامتر λ یک عدد صحیح کوچک‌تر از ۵ است و η نرخ یادگیری است که معمولاً برابر ۰/۱ در نظر گرفته می‌شود. همچنین مقدار ξ_m با استفاده از (۳) تعریف می‌شود. همان طور که مشاهده می‌شود برای آموزش وزن اتصالات بین لایه‌ای و انتخاب وزن‌های بهینه از برچسب داده‌ها استفاده می‌شود

$$\xi_m = \begin{cases} 1 & \text{پیش‌بینی صحیح برای گره برنده یا پیش‌بینی اشتباه گره} \\ 0 & \text{پیش‌بینی صحیح برای گره برنده پیش‌بینی یا اشتباه گره برنده} \\ 0 & \text{برچسب گم شد} \end{cases} \quad (3)$$

در نهایت برای محاسبه برچسب ورودی ابتدا میزان امتیاز هر گره در لایه خروجی شبکه خودسازمانده پویای شبه‌ناظر با استفاده از (۴) محاسبه می‌شود

$$PS_m = \sum_{c=1}^{N_{unit}} w_{cm} o_c \quad (4)$$

که در آن N_{unit} تعداد گره‌ها در لایه پایین است. بعد از آن که میزان امتیاز تمام گره‌ها در لایه خروجی بر اساس این رابطه محاسبه شد، برچسب داده ورودی برابر با آن دسته یا گره خروجی است که میزان امتیاز آن از سایر گره‌ها بیشتر باشد.



شکل ۳: شبکه خودسازمانده پویای شبه‌ناظر.

استخراج‌شده و با استفاده از داده برچسب‌دار فرایند آموزش را طی کرده است، در مورد حاوی اطلاعات بودن یا نبودن فایل موجود تصمیم‌گیری می‌کند. طبقه‌بندها را در یک دسته‌بندی کلی می‌توان به سه دسته باناظر، شبه‌ناظر و بدون ناظر دسته‌بندی کرد. به طور معمول از طبقه‌بندهای با ناظر استفاده می‌شود.

در [۴] برای طبقه‌بندی از بین شبکه عصبی، SVM^۱ و kNN^۲ به دلیل سادگی بیشتر از kNN استفاده شده است. در [۷]، [۹] و [۱۰] برای شبیه‌سازی از طبقه‌بند kNN استفاده شده است. برای هر سه طبقه‌بند k برابر یک در نظر گرفته شده است. در [۱۳] از طبقه‌بند SVM برای جداسازی فایل حاوی اطلاعات از غیر آن استفاده شده است. در تمامی کارهای بیان‌شده از روش‌های یادگیری با ناظر برای طبقه‌بندی استفاده شده است. مشکل اصلی روش یادگیری با ناظر با توجه به عدم وجود کد الگوریتم‌های نهان‌نگاری در نهان‌کاوی کور، نیاز به صرف وقت و هزینه انسانی برای تهیه نمونه‌های برچسب‌دار برای آموزش طبقه‌بند می‌باشد.

در [۱۱] از طبقه‌بند خودسازمانده پویای شبه‌ناظر (SSGSOM)^۳ استفاده شده است. ویژگی شبه‌ناظر بودن این طبقه‌بند امکان استفاده کمتر از داده‌های برچسب‌دار را فراهم کرده است، در واقع یادگیری شبه‌ناظر علاوه بر استفاده از اطلاعات موجود در برچسب داده‌ها، از اطلاعات خام داده‌ها به علت حضور در فضای ویژگی نیز استفاده می‌کند (داده‌های بدون برچسب). بنابراین این اطلاعات می‌تواند منجر به کاهش تعداد برچسب مورد نیاز شود اما در این طبقه‌بند (SSGSOM) به کار رفته در [۱۱]) هنگامی که نسبت داده‌های برچسب‌دار نسبت به کل داده‌ها در فضای ویژگی کاهش می‌یابد، دقت نهان‌کاوی شبه‌ناظر به سرعت افت می‌کند (نقطه ضعف).

در تحقیق پیش رو هدف، طراحی نوع جدیدی از طبقه‌بند شبه‌ناظر SSGSOM است که با کمترین میزان داده برچسب‌دار بتواند به دقت نهایی (دقتی که با فرض برچسب‌دار بودن تمامی داده‌ها می‌توان بدان دست یافت) دست یابد. قبل از بیان مطالب مربوط به ایده تحقیق، ذکر این نکته ضروری است که با توجه به این که ایده تحقیق مربوط به بلوک سوم (طبقه‌بند) شکل ۲ می‌باشد مطالبی که از این پس ارائه خواهد شد همگی مربوط به همین بلوک (طبقه‌بند) خواهد بود.

1. Support Vector Machine
2. Kth Nearest Neighbor
3. Semi Supervised Growin Self Organizing Map

4. Geometrical Self Organizing Map
5. Growin Self Organizing Map

۱-۲ یادگیری فعال

$$\Delta w_c = \eta(\xi_m - PS_m) o_c \quad (5)$$

که در آن تغییر الگوی وزن‌های بین لایه‌ای که در چند مرحله صورت می‌پذیرد به دو عامل (۱) سطح فعالیت گره‌ها (O_c) و (۲) الگویی که شبکه از طریق آن آموزش می‌بیند (ξ_m) وابسته است. توجه شود η مقداری ثابت و PS_m بر اساس (۴) متأثر از O_c است. به این نکته مهم نیز توجه شود که برای آموزش یک سطر از وزن‌های بین لایه‌ای که منتسب به یک گره مشخص در لایه زیرین است از نزدیک‌ترین داده ورودی به آن گره با برچسب مربوط استفاده می‌شود. برای گره‌هایی که در لایه زیرین دارای سطح فعالیت و الگوی یکسانی باشند (منتسب کردن الگوی ورودی به گره‌ها به علت وجود تناظر یک به یک بین داده‌های برچسب‌دار و گره‌ها است که در ادامه در قضیه ۱ اثبات خواهد شد)، وزن‌های بین لایه‌ای به شکل الگوی تکراری آموزش خواهند یافت.

توجه شود آنچه بیان شد عبارت دقیقی نیست و عبارت دقیق‌تر این است که گفته شود هنگامی که یک داده با برچسب مربوط به شبکه اعمال می‌شود، بردار سطح فعالیت گره‌های مربوط به آن داده به صورت $O_1 = (o_1, o_2, \dots, o_c)$ به دست می‌آید. حال اگر داده دیگری با برچسب مربوط به خود به شبکه اعمال شود که برای این داده بردار $O_p = (o'_1, o'_2, \dots, o'_c)$ باشد به شرط آن که دو برچسب داده و همچنین بردارهای O_c مربوط برابر باشد در این صورت دو گره‌ی که در نزدیک‌ترین فاصله به دو داده هستند دارای افزونگی هندسی فضای ویژگی هستند که منجر به آموزش الگوی تکراری برای وزن‌های بین لایه‌ای خواهد شد. البته در ادامه نیز به جز در موارد لازم همان بیان غیر دقیق که فهم آن ساده‌تر است به کار برده می‌شود.

اگر (۵) به صورت ماتریسی بیان شود خواهیم داشت

$$\begin{pmatrix} \Delta w_{11} & \dots & \Delta w_{1m} \\ \vdots & \ddots & \vdots \\ \Delta w_{c1} & \dots & \Delta w_{cm} \end{pmatrix} = \begin{pmatrix} k_1 o_1 & \dots & k_m o_1 \\ \vdots & \ddots & \vdots \\ k_1 o_c & \dots & k_m o_c \end{pmatrix} \quad (6)$$

که در آن

$$k_i = \eta(\xi_i - PS_i) \quad (7)$$

و $i = 1, \dots, m$ است. الگوهای تکراری به صورت سطرهای تکراری در ماتریس وزن‌های بین لایه‌ای (۶) خود را نشان می‌دهد. بدیهی است که هدف از اعمال الگوریتم LMS^F در ماتریس وزن‌های بین لایه‌ای برای آموزش درایه‌ها، یادگیری الگوهای متفاوت بوده و آموزش الگوی تکراری مشکلی را حل نخواهد کرد و بنابراین حذف آن نیز ایجاد مشکل نخواهد کرد. برای حذف هر سطر ماتریس وزن‌های بین لایه‌ای کافی است گره منتسب به آن حذف شود، بنابراین حذف گره مربوط به سطری از ماتریس که مساوی سطر دیگر است منجر به حذف افزونگی هندسی فضای ویژگی خواهد شد.

۲-۲-۲ رهیافت تعیین الگوریتم برای تشخیص گره‌های دارای افزونگی هندسی

با توجه به مطالب بیان‌شده، گره‌هایی که در دو مقدار ذکرشده سطح فعالیت و الگوی آموزش دارای مقدار یکسانی باشند، به گره‌هایی دارای افزونگی هندسی فضای ویژگی تعبیر می‌شوند که بدون رخداد خطا می‌توان آنها را حذف نمود. سؤال بسیار مهمی که باید بدان پاسخ داد این

در یادگیری فعال^۱ هدف رسیدن به بالاترین دقت در مسایل یادگیری ماشینی با صرف کمترین برچسب در فرایند آموزش است [۲۱]. در یادگیری فعال فرض می‌شود کارشناس بتواند از میان انبوه داده‌های بدون برچسب، بر اساس راهبردهای متفاوت داده‌ای را که می‌بایست برچسب آن انتخاب شود تعیین نماید. بر این مبنا برای بدست آوردن برچسب داده‌ها سناریوها و راهبردهای متفاوتی تدوین شده است. مثلاً در راهبرد نمونه‌برداری عدم قطعیت^۲ که عمومی‌ترین راهبرد یادگیری فعال است، کارشناس برچسب نمونه‌هایی از فضای ویژگی را درخواست می‌کند که قادر به پیش‌بینی آن نباشد و از درخواست برچسب نمونه‌هایی که می‌تواند بدون رخداد خطا پیش‌بینی نماید اجتناب می‌کند [۲۲]. کاربردهای یادگیری فعال چنان گسترده است که حتی تهیه لیستی از آن مشکل است اما به طور کلی بیشتر توجه به آن در حوزه دسته‌بندی داده‌هاست هرچند که در مواردی محدود مانند [۲۳] در کشف دسته‌های ناشناخته^۳ نیز مورد استفاده قرار گرفته است. در دسته‌بندی داده‌ها نیز اخیراً کاربردهای دسته‌بندی تصاویر سنجش از راه دور^۴ بیشتر مورد توجه قرار گرفته است [۲۴] تا [۲۷]. در این مقاله از ایده یادگیری فعال برای دسته‌بندی داده‌ها برای نهان‌کاوی ویدئو استفاده شده است. از دیدگاه طبقه‌بندی که برای دسته‌بندی مورد استفاده قرار می‌گیرد باید گفت یادگیری فعال توانایی کار با انواع طبقه‌بند برای بهینه‌کردن مصرف برچسب را داراست. برای نمونه در [۲۸] از ایده یادگیری فعال در کنار طبقه‌بند SVM استفاده شده است در حالی که در اینجا از ایده یادگیری فعال در کنار طبقه‌بند SSGSOM استفاده خواهیم کرد. به عبارت دقیق‌تر در این مقاله برای بازطراحی طبقه‌بند SSGSOM از یادگیری فعال استفاده می‌شود.

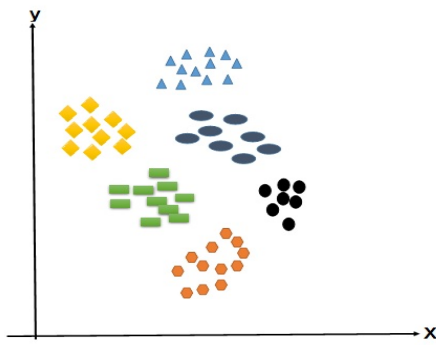
۲-۲-۲ افزونگی هندسی و مفهوم دسته

در این بخش ابتدا به بیان مشکل شبکه SSGSOM که منجر به استفاده به میزان زیاد از برچسب می‌شود می‌پردازیم و سپس ساختاری که منجر به رفع این مشکل می‌شود بیان خواهد شد.

۱-۲-۲ توصیف ریاضی افزونگی هندسی

در الگوریتم GSOM میزان پراکندگی گره‌ها توسط ضریب گسترش^۵ (SF) کنترل می‌شود. پس از انتخاب این ضریب توسط کاربر، الگوریتم مکان گره‌ها را در لایه پایین با توجه به تراکم داده‌ها به نحوی می‌یابد تا در مکان‌هایی از فضای ویژگی که داده‌ها تراکم بیشتری دارند به همان نسبت گره‌های بیشتری ایجاد شود. در نتیجه ساختاری از گره‌ها به دست می‌آید که تخلیص‌یافته داده اصلی با حفظ ساختار هندسی آن در فضای ویژگی است. این دقیقاً همان هدفی است که GSOM برای آن طراحی شده است. اما این نتیجه برای SSGSOM که در لایه زیرین خود از GSOM استفاده کرده مطلوب نیست. دلیل این عدم مطلوبیت را باید در مفهومی به نام افزونگی هندسی گره‌های لایه زیرین که در تعارض با هدف SSGSOM که استفاده حداقلی از داده برچسب‌دار است جستجو کرد. برای اثبات این مدعا باید از (۱) و (۲) کمک گرفت. رابطه (۲) را می‌توان به صورت (۵) بازنویسی نمود

1. Active Learning
2. Uncertainty Sampling
3. Unknown Class Discovery
4. Remote Sensing Image Classification
5. Spread Factor



شکل ۴: شش دسته داده در فضای ویژگی دوبعدی.

هستند چرا که الگوی خروجی یکسانی دارند.

در بین تعدادی از گره‌ها نیز لازمه این که تمامی گره‌ها دارای افزونگی هندسی باشند این است که هر گره با تمامی گره‌های دیگر در هر دو عامل مشترک باشند و این یعنی این که تمامی آنها در نزدیکی یکدیگر نسبت به سایر داده‌ها در فضای ویژگی قرار گرفته‌اند. داده‌هایی که نسبت به سایر داده‌ها در نزدیکی یکدیگر در فضای ویژگی مجتمع شده باشند را یک دسته می‌نامند. بنابراین تمامی گره‌هایی که تشکیل یک دسته می‌دهند دارای افزونگی هندسی هستند و کافی است برای حذف این افزونگی تمامی گره‌های دسته را در یک گره ادغام نمود. این حالت در شکل ۴ معادل تمامی گره‌های هر یک از دسته‌هاست.

در حالت چهارم برچسب دو گره یکسان ولی سطح فعالیتشان متفاوت است. در نگاه اول این طور به نظر می‌رسد که این حالت امکان وقوع ندارد چرا که لازمه یکسان بودن برچسب دو نمونه، قرابت در فضای ویژگی است که با متفاوت بودن سطح فعالیت که لازمه آن عدم قرابت در فضای ویژگیست قابل جمع نیست. اما این حالت یکی از ویژگی‌های جالب شبکه SSGSOM که دارای دو لایه است می‌باشد. برای فهم این حالت بار دیگر شکل ۴ را در نظر بگیرید. در این شکل اگر لازم شود داده‌های شبیه به دایره (دایره، بیضی و پنج‌ضلعی منتظم) که به صورت غیر خطی در بین سایر داده‌ها قرار گرفته‌اند جدا شوند کافی است به این سه دسته داده در لایه بالاتر برچسب‌های مشابه داده شود و با استفاده از طبقه‌بند فرایند آموزش طی شود. در این کاربرد این سه مجموعه داده دارای برچسب‌های یکسان بوده در حالی که سطح فعالیت آنها متفاوت است (توجه شود در این حالت هر چند برچسب‌ها یکسان است اما به دلیل یکی نبودن سطح فعالیت‌ها، گره‌های مربوط تشکیل دسته نمی‌دهند). آنچه بیان شد، توصیف مفهوم آشنایی به نام کرنل^۱ است که برای جداپذیر نمودن داده‌های غیر خطی از آن استفاده می‌شود. جهت جلوگیری از کشیدن بحث به مباحث حاشیه‌ای و خروج از اصل موضوع، از توضیح بیشتر در این رابطه خودداری می‌کنیم.

۲-۳ جمع‌بندی و استنتاج مفهوم دسته برای گره‌های دارای افزونگی هندسی

نتیجه بسیار مهمی که از دقت در چهار حالت ذکر شده در بخش قبل مشخص می‌شود این است که تنها در حالت سوم که گره‌ها دارای افزونگی هندسی هستند تشکیل دسته می‌دهند، بنابراین نکته طلایی برای شناسایی گره‌های دارای افزونگی هندسی و سپس ادغام آنها در یک گره، شناسایی دسته‌هایی است که گره‌ها تشکیل می‌دهند. بر اساس آنچه در این بخش بیان شد برای نشان دادن وجود افزونگی

جدول ۱: حالات مختلف برچسب و سطح فعالیت برای دو گره.

حالت	اول	دوم	سوم	چهارم
برچسب	متفاوت	یکسان	متفاوت	یکسان
سطح فعالیت	متفاوت	یکسان	متفاوت	متفاوت

است که آیا رهیافت حذف گره‌های دارای افزونگی هندسی در فضای ویژگی به روشی مشخص خواهد انجامید؟ و اگر آری، آن روش و الگوریتم چیست؟ برای پاسخ باید دو عاملی که سبب ایجاد افزونگی هندسی در فضای ویژگی شده‌اند را در نظر گرفت: برای هر دو گره دلخواه با احتساب کلیه حالت‌های ممکن رخداد این دو عامل (سطح فعالیت و برچسب) چهار حالت متصور است که در جدول ۱ نشان داده شده و در ادامه به بررسی آنها پرداخته می‌شود.

در حالت اول بر اساس جدول ۱ که دو گره هم دارای برچسب و هم سطح فعالیت متفاوتی هستند افزونگی هندسی وجود ندارد. برای تصور ذهنی این حالت از شکل ۴ استفاده شده است. در این شکل شش دسته داده متفاوت به اشکال دایره، بیضی، پنج‌ضلعی منتظم، مستطیل، لوزی و مثلث در یک فضای ویژگی دوبعدی تصور شده است. حالت اول را می‌توان در شکل ۴ تعبیر به دو گره از دسته‌های متفاوت نمود (برای مثال یک داده از دسته مستطیل با یک داده از دسته دایره). در حالت دوم دو گره دارای برچسب متفاوت اما سطح فعالیت یکسانی هستند. در این حالت نیز افزونگی هندسی بین دو گره وجود ندارد. این حالت انتزاعی بوده و در عمل امکان وقوع ندارد چرا که اگر دو گره دارای سطح فعالیت یکسانی باشند بردار O_e مربوط به هر دو داده ورودی مساوی است. در این حالت با توجه به (۴) حتماً برچسب خروجی یکسان خواهد بود و به عبارت دیگر هنگامی که به یک تابع ورودی‌های یکسان داده شود، نباید توقع داشت خروجی متفاوت دریافت نمود چرا که بر اساس (۴) نسبت برچسب به سطح فعالیت مثل خروجی به ورودی یک تابع است (البته دقت شود که برچسب مورد بحث PS_m است که مربوط به داده ورودی است و معادل نمودن آن با PS_m به این دلیل است که الگوریتم LMS وزن‌های بین لایه‌ای را در جهت یکسان نمودن این دو آموزش می‌دهد).

حالت سوم حالتی است که دو گره هم دارای سطح فعالیت یکسان و هم برچسب مشابه هستند. این همان حالتی است که در آن دو گره دارای افزونگی هندسی فضای ویژگی هستند. هنگامی که دو گره دارای برچسب یکسان باشند و در فضای ویژگی در نزدیکی یکدیگر قرار داشته باشند بر اساس (۱) سطح فعالیت یکسانی خواهند داشت. به عبارت دقیق‌تر در حالتی که دو گره در فضای ویژگی نسبت به سایر گره‌ها نزدیک به هم باشند دارای خواص مشترک بوده و بنابراین برچسب نزدیک‌ترین داده به هر کدام مشابه دیگری است و همچنین بر اساس (۱) بردار O_e دو داده یکسان خواهد بود (به دلیل سادگی از اثبات آن صرف نظر می‌شود). بنابراین شرط لازم برای این که دو گره دارای افزونگی هندسی فضای ویژگی باشند، قرابت آن دو در فضای ویژگی است. البته باید توجه نمود که منظور از یکسان بودن دو بردار O_e نزدیک بودن مقدار درایه‌های نظیر به نظیر آن دو به یکدیگر به حدی است که الگوی خروجی یکسانی به دست آید چرا که در لایه بالا انتخاب برچسب بر اساس راهبرد گره برنده به دست می‌آید، بنابراین الزامی برای دقیقاً یکی بودن آنها وجود ندارد. برای فهم این مطلب دو بردار $O_1 = (0.8, 0.1)$ و $O_2 = (0.6, 0.3)$ را در نظر بگیرید با فرض آن که ماتریس وزن‌های بین لایه‌ای به صورت $w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ باشد. با توجه به مطالب بیان شده این دو بردار یکسان

خواهد بود (اثبات این قضیه نیز در پیوست موجود است). در واقع قضیه ۲ بیان می‌دارد روشی که در بخش قبل برای از بین بردن افزونگی هندسی فضای ویژگی گره‌های لایه زیرین به دست آمد به حالت بهینه که همان استفاده حداقلی از برچسب‌ها است خواهد انجامید. این قضیه نه تنها تعداد حداقل گره‌ها در لایه پایین را مشخص می‌کند (حذف کامل افزونگی هندسی) بلکه مسیر رسیدن به این تعداد را نیز که پیش از این به دست آمد در درون خود دارد. برای حذف کامل افزونگی هندسی و رسیدن به تعداد حداقل گره کافی است داده‌های موجود در فضای ویژگی را دسته‌بندی نموده و گره‌ها را در مراکز این خوشه‌ها قرار دهیم و از نزدیک‌ترین داده موجود به هر گره برای فرایند آموزش استفاده نماییم. به عبارت دیگر با ادغام تمامی گره‌های موجود هر دسته در گرهی به نام مرکز دسته افزونگی هندسی گره‌ها از بین خواهد رفت و می‌توان به تعداد حداقل برچسب رسید.

اما در صورت قضیه ۲ فرض شده دسته‌ها به صورت خطی جداپذیر باشند و این فرض فضای مسئله را به شدت محدود کرده است. سؤال مهم این است که برای رفع این محدودیت چه باید کرد؟ به عبارت دیگر در حالت کلی و با فرض وجود N دسته (نه لزوماً جداپذیر خطی) حداقل تعداد گره لازم برای رسیدن به پاسخ با احتمال خطای صفر چه تعداد است؟

برای پاسخ به سؤال دو حالت متصور است. یک مورد این که بر اساس روابط موجود و در ساختار شبکه GeSOM قصد یافتن راه حلی را داشته باشیم. در این حالت بر اساس (۱) میزان شباهت هر داده با مرکز دسته خود با فاصله آن داده تا مرکز هر دسته رابطه عکس دارد. در دسته‌های غیر خطی به دلیل این که فاصله برخی داده‌ها از مرکز دسته خود بیشتر از فاصله تا سایر مراکز است، طبقه‌بند در تعیین برچسب دچار خطا می‌شود. برای تصحیح خطا دو راه حل وجود دارد: (۱) خطی‌سازی محلی، در این حالت برای داده‌هایی که غیر خطی می‌باشند دسته‌های جدیدی به صورتی در نظر گرفته می‌شود که دسته‌های جدید همگی خطی باشند. (۲) با توجه به این که الگوریتم طراحی شده باید قابلیت افزایش دقت پس از تخصیص برچسب در مرحله اول را دارا باشد می‌توان خطای ایجادشده در این مرحله را پذیرفت و سپس در مرحله بعد دقت را ارتقا داد. هر یک از دو روش بیان‌شده نیازمند صرف برچسب بیشتر است که به دلیل سادگی و حجم کم محاسبات روش دوم نسبت به روش اول، در این تحقیق از روش دوم استفاده می‌نماییم.

فرض دیگر این که بتوان در روابط موجود به ویژه (۳) تجدید نظر کرد. به دلیل این که این حالت از چهارچوب تحقیق خارج است از پیگیری آن صرف نظر نموده و به عنوان پیشنهاد در انتها برای تحقیقات آتی مطرح خواهد شد.

با توجه به دو قضیه مطرح‌شده برای دستیابی به الگوریتم GeSOM که می‌بایست حداقل تعداد برچسب برای فرایند آموزش شبکه را از طریق کاهش افزونگی هندسی به کار بگیرد، باید حداقل تعداد گره در لایه پایین ایجاد شود. حال اگر در لایه پایین N دسته داده متمایز وجود داشته باشد، می‌توان با انتخاب مراکز دسته‌ها و قراردادن گره‌ها در این مراکز، شبکه را ایجاد نمود. سپس نزدیک‌ترین نمونه به هر گره برای آموزش آن گره انتخاب شده و با توجه به برچسب آن فرایند آموزش وزن‌های بین لایه‌ای اجرا شود. اگر دسته‌ها به صورت خطی جداپذیر باشند طبقه‌بند با احتمال رخداد صفر برچسب تمامی داده‌ها را پیش‌بینی می‌نماید و در صورت جداپذیر خطی نبودن دسته‌ها، داده‌هایی که به مرکز دسته‌ای دیگر شباهت بیشتری داشته باشند طبقه‌بند برچسب را به اشتباه تخصیص

هندسی گره‌های موجود در یک دسته به هر گره یک داده برچسب‌دار اختصاص یافت. به عبارت ساده‌تر برای آن که کاهش تعداد گره‌ها منجر به کاهش تعداد برچسب مورد نیاز شود می‌بایست بین برچسب‌ها و گره‌ها تناظر یک به یک برقرار باشد. این گزاره که بین داده‌های برچسب‌دار و گره‌ها تناظری یک به یک برقرار است نیاز به اثبات دارد که در بخش بعد در قالب یک قضیه ارائه خواهد شد. همچنین پاسخ این سؤال بسیار مهم که آیا با حذف افزونگی هندسی گره‌های هر دسته، حالت بهینه (کمترین تعداد برچسب ممکن) به دست می‌آید در قالب قضیه‌ای دیگر جواب داده خواهد شد و سپس الگوریتم طبقه‌بند خواهد آمد.

قبل از بیان دو قضیه لازم است ساختار الگوریتم پیشنهادی بیان شود. GeSOM دارای دو لایه است که در لایه پایین یک سری گره^۱ قرار دارد. این گره‌ها دارای توابع گوسی بوده که (۱) آن را بیان می‌کند و به وسیله وزن‌های بین لایه‌ای به لایه بالا متصل می‌شوند. وزن‌های بین لایه‌ای به وسیله الگوریتم LMS فقط برای داده‌های برچسب‌دار آموزش می‌یابند (بر خلاف الگوریتم SSGSOM که هم از داده‌های برچسب‌دار و هم بدون برچسب برای آموزش وزن‌های بین لایه‌ای استفاده می‌کند). هدف یافتن مکان و تعداد گره‌ها در لایه پایین به نحوی است که حداقل برچسب لازم برای آموزش شبکه لازم باشد. به عبارت دقیق‌تر در این مقاله هدف این نیست که دقت نهایی شبکه SSGSOM مانند مقالات مشابهی همچون [۲۹] که هدفشان ارتقای دقت بوده است افزایش یابد بلکه هدف صرفاً رسیدن به همان دقت نهایی شبکه SSGSOM با صرف برچسب کمتر است.

به زبان ساده می‌توان گفت سخت‌افزار GeSOM مطابق شکل ۳ همانند SSGSOM بوده اما نحوه چینش گره‌ها در لایه پایین تغییر خواهد کرد تا با صرف برچسب کمتر به دقت نهایی دست یابیم.

۲-۳ مبانی ریاضی مورد نیاز

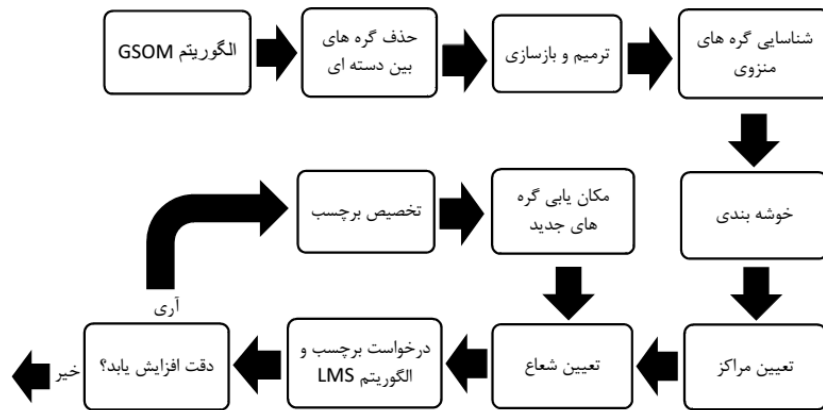
با توجه به توضیحات داده‌شده در بخش قبل و بر اساس معیار رسیدن به استفاده از حداقل تعداد برچسب برای فرایند آموزش شبکه، در این بخش الگوریتم پیشنهادی ارائه می‌شود.

قبل از ارائه الگوریتم به بیان دو قضیه که برای به حداقل رساندن افزونگی هندسی گره‌های لایه زیرین طرح شده است و نتایج آنها در الگوریتم استفاده می‌گردند پرداخته می‌شود.

قضیه ۱) در فرایند آموزش شبکه GeSOM برای رسیدن به حداقل برچسب لازم با فرض خطای صفر، باید به هر گره که در لایه پایین وجود دارد دقیقاً یک برچسب اختصاص داد (برای اثبات قضیه به پیوست مراجعه شود).

قضیه ۱ از تناظر یک به یک بین گره‌ها و برچسب‌ها خبر می‌دهد. طبق قضیه فوق برای رسیدن به حداقل برچسب در فرایند آموزش شبکه با فرض خطای صفر باید حداقل تعداد گره در لایه پایین ایجاد شود. حال این سؤال مطرح می‌گردد که حداقل تعداد گره که در لایه پایین امکان ایجاد شدن را داراست چه تعداد می‌باشد؟ برای پاسخ به این سؤال قضیه ۲ مطرح می‌شود.

قضیه ۲) فرض کنید در یک فضای M بعدی تعداد زیادی داده که در N کلاس مختلف چیده شده‌اند موجود هستند. با فرض این که دسته‌های مختلف به صورت خطی جداپذیر باشند، حداقل تعداد لازم گره برای ایجاد شبکه GeSOM با احتمال خطای صفر، برابر با تعداد دسته‌ها (N)



شکل ۵: ساختار الگوریتم ارائه شده برای طبقه بندی GeSOM.

گره برنده (BMU) به سمت گره ورودی جابه جا می شود گره های همسایه را نیز با خود جابه جا می نمود. میزان جابه جایی برخی از این گره ها به حدی بود که از فضای ویژگی دسته ها به طور کلی خارج شده که در این حالت توسط مرحله دوم شناسایی و حذف شدند اما برای برخی از داده ها میزان جابه جایی از فضای ویژگی به حدی نیست که به عنوان گره بین دسته ای محسوب شوند. این نوع گره ها را نه می توان حذف کرد زیرا بر خلاف گره های بین دسته ای نماینده تعدادی داده هستند و حذفشان به ساختار شبکه آسیب می رساند و نه می توان آنها را به حال خود رها کرد بدون نیاز به پراتنر- زیرا به دلیل انحراف از فضای ویژگی دسته مربوط به خود و نزدیک شدن به سایر دسته ها خطای طبقه بندی kNN را زیاد می کنند. بر این اساس در مرحله ترمیم و بازسازی از عملگر میانگین گیر برای حل مشکلات فوق استفاده می شود. عملگر میانگین گیر به دلیل خاصیت ذاتی خود سبب خواهد شد گره ها از مقادیر بیشینه و کمینه خود (در اینجا مرز دسته ها) فاصله گرفته و به سمت مرکز هر دسته (دسته ای که گره، نماینده آنها می باشد) جابه جا شوند که سبب افزایش نسبت فاصله بین دسته ای به فاصله درون دسته ای خواهد شد.

در مرحله بعد نوبت به شناسایی گره های منزوی می رسد. این نوع گره ها در درون دسته با توجه به فاصله زیادتری که نسبت به سایر گره ها دارند می توان تشخیص داد. برای شناسایی این گره ها کافی است فاصله هر گره تا نزدیک ترین گره در همان دسته را به عنوان یک متغیر تصادفی در نظر گرفت. با توجه به این که فاصله هر گره از نزدیک ترین گره در همان دسته مستقل از سایر فاصله ها است، بر طبق قضیه حد مرکزی هنگامی که تعداد این گره ها زیاد و زیادتر شود این فاصله دارای توزیع گوسی خواهد شد. حال می توان با یافتن میانگین و واریانس این فاصله ها، گره منزوی را این گونه شناسایی کرد: گره هایی که حداقل فاصله آنها (فاصله تا نزدیک ترین گره) از میانگین فاصله ها به علاوه انحراف معیار (جذر واریانس) فاصله ها بیشتر باشد. برای این که این فاصله زیاد برای خوشه بندی kNN ایجاد مشکل نکند این گره ها به طور موقت حذف شده و پس از اعمال خوشه بندی، دوباره اضافه می شوند. توجه شود به دلیل این که این نمونه ها نماینده تعداد زیادی داده هستند حذف کلی آنها (مانند گره های بین دسته ای) سبب تخریب ساختار شبکه خواهد شد.

در مرحله بعد با استفاده از گراف kNN خوشه بندی انجام می پذیرد. این مرحله از نوع یادگیری بدون ناظر بوده و حتی یک برچسب هم مصرف نمی شود.

خواهد داد. در این حالت طبق آنچه بیان شد باید با ایجاد دسته جدید برای این داده ها و صرف برچسب بیشتر از ایجاد خطا جلوگیری کرد. قبل از ارائه الگوریتم GeSOM، بیان این مطلب ضروری است که داده هایی که در فضای ویژگی قابل دسترسی هستند به صورت خام بوده و نه تنها دسته بندی شده نیستند بلکه حتی تعداد دسته ها (N) نیز مشخص نیست. بنابراین الگوریتم GeSOM باید داده ها را دسته بندی نموده و با توجه به دسته های ایجاد شده و مراکز دسته ها تعداد و محل قرار گرفتن هر گره را تعیین نماید. اما در دسته بندی داده ها امکان دارد داده های متمایز در یک دسته قرار بگیرند و بنابراین خطا مقداری غیر صفر خواهد داشت. پس الگوریتم باید به نحوی طراحی شود تا بتواند در مراحل بعد خطا را یافته و با برطرف کردن آن دقت سیستم را افزایش دهد که مکانیزم آن در ادامه توضیح داده خواهد شد.

۲-۲ الگوریتم GeSOM مبتنی بر یادگیری فعال برای به حداقل رساندن افزونگی هندسی

نمودار گردش عملیات الگوریتم GeSOM در شکل ۵ ارائه شده است. در ادامه به شرح تک تک مراحل می پردازیم. ابتدا الگوریتم GSOM بر روی داده ها اعمال می شود. با این کار دو هدف دنبال می شود:

(۱) کم کردن حجم داده ها که برای کم کردن حجم پردازش های آتی مفید است و

(۲) دستیابی به ساختار داده ها. اگر داده ها در N دسته متمایز دسته بندی شده باشند با اعمال GSOM می توان گره ها را به روی N دسته پخش نمود.

در مرحله بعد، هدف اصلی خوشه بندی داده ها است که برای آن از خوشه بندی بدون ناظر kNN استفاده خواهد شد. هر چند این خوشه بندی به دلیل قابلیت بالا برای خوشه بندی های غیر خطی به ویژه در مواردی که تعداد خوشه ها نیز مشخص نیست (مانند اینجا) بسیار مفید است اما به علت وجود داده در بین دسته های متمایز که سبب رخداد خطا در خوشه بندی خواهد شد، لازم است قبل از خوشه بندی نسبت به رفع این نقیصه اقدام نمود که این مهم در طی سه مرحله انجام می پذیرد.

بنابراین در مرحله دوم باید ابتدا به حذف گره های بین دسته ای پرداخت. هدف از این کار هر چه بیشتر جدا کردن خوشه ها نسبت به یکدیگر است. در هنگام پخش گره ها بر اثر اعمال الگوریتم GSOM بر روی دسته ها، تعدادی گره نیز در بین دسته ها ایجاد می شود که در این مرحله حذف می شوند.

در مرحله اول در طی اعمال الگوریتم GSOM به داده ها، زمانی که

محل رخداد خطا مشخص شود. با توجه به (۱) میزان شباهت هر داده با مرکز دسته (داده برچسب‌دار) با فاصله دو داده نسبت عکس دارد. بر این اساس هر چه ابعاد یک دسته بزرگ‌تر باشد، شباهت داده‌های مرزی با مرکز دسته کاهش بیشتری یافته و احتمال رخداد خطا برای این داده‌ها افزایش می‌یابد.

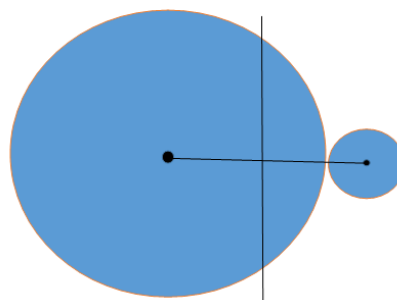
بر اساس مطالب بیان‌شده برای تخصیص برچسب جدید می‌بایست نسبت ابعاد دسته‌ها در نظر گرفته شود و به دسته‌های بزرگ‌تر برچسب بیشتری اختصاص یابد تا بتوان احتمال خطا را بیشتر و بیشتر کاهش داد. بر اساس قضیه ۱ برای هر برچسب می‌بایست یک گره در لایه پایین وجود داشته باشد. در این مرحله باید مکان گره‌های جدید در فضای هندسی مسأله مشخص شود. با توجه به این که هدف از ایجاد گره‌های جدید کاهش خطاست (افزایش دقت) بر اساس مطالب بیان‌شده که به حداقل رسیدن خطا را منوط به کاهش هرچه بیشتر ابعاد دسته‌ها می‌داند، الگوریتمی که برای مکان‌یابی گره‌های جدید طراحی شده گره‌ها را به صورتی در فضای هندسی دسته‌ها قرار می‌دهد تا دسته‌های جدید ایجادشده دارای حداقل ابعاد ممکن باشد.

پس از تعیین شعاع برای تمامی دسته‌ها، هم برای گره‌های جدید و گره‌های سابق (توجه شود هرچند مکان گره‌های سابق تغییر نکرده اما به دلیل ایجاد دسته‌های جدید که منجر به کاهش ابعاد دسته‌های سابق شده نیاز به تعیین شعاع جدید است) و اعمال الگوریتم LMS می‌توان انتظار داشت دقت افزایش یافته باشد. روند افزایش دقت بنا بر نظر کاربر تا برچسب‌دار کردن تمامی گره‌های ایجادشده در اثر اعمال GSOM که پس از پالایش در مراحل دو تا چهار همچنان باقی مانده‌اند می‌تواند ادامه یابد.

۳- نتایج شبیه‌سازی

همان طور که پیش از این توضیح داده شد با توجه به شکل ۲ نهان کاو پیش‌نهادی دارای سه بخش مجزا بوده که ایده این تحقیق مربوط به بلوک سوم است. برای شبیه‌سازی و آزمایش یک طبقه‌بند می‌بایست پایگاه داده به صورت مجموعه‌ای از داده و برچسب‌های مربوط موجود باشد. با توجه به این که طبقه‌بند پیشنهادی در نهان کاوی کور ویدئو مورد استفاده قرار گرفته است، پایگاه داده از این مجموعه انتخاب شده است. این پایگاه داده در [۱۱] ایجاد شده و با توجه به یکسان بودن تمامی شرایط و همچنین برای مقایسه نتایج با نتایج به دست آمده در [۱۱] از همین پایگاه داده استفاده می‌شود. این پایگاه داده شامل ۱۹ دنباله ویدئو با نوع نمایش YUV و فرمت نمونه‌برداری ۴:۲:۰ است که برای آزمایشات مورد استفاده قرار گرفته است. دنباله‌های استاندارد ویدئو از سایت <http://media.xiph.org/video/derf/> گرفته شده‌اند و مشخصات آنها در جدول ۲ آمده و همان طور که در جدول مشخص است دنباله‌های ویدئو از نظر مؤلفه حرکت دارای تنوع مناسب می‌باشند. از آنجایی که دنباله‌های ویدئو دارای تعداد فریم‌های متفاوت از ۹۰ تا ۱۰۶۵ می‌باشند همه دنباله ویدئو بدون همپوشانی به دنباله‌های ۵۰ فریمی تقسیم می‌شوند و در مجموع تعداد کل دنباله‌ها به بیش از ۱۱۹ عدد می‌رسد.

البته همان طور که پیش از این بیان شد، طبقه‌بند پیشنهادی محدود به نهان کاوی ویدئو نبوده و قابل استفاده برای سایر پایگاه‌های داده نیز می‌باشد. به همین جهت قبل از بیان نتایج برای پایگاه داده ویدئوهای ارائه‌شده در جدول ۲، نتایج شبیه‌سازی برای سه پایگاه داده twomoon، جدانپذیر خطی و جدانپذیر خطی که در شکل ۷ نشان داده شده‌اند آورده می‌شود. نتایج سه بار آزمایش حاصل از اعمال طبقه‌بند به سه مجموعه



شکل ۶: نمایش دو دسته داده کروی که شعاع یکی چند برابر دیگری است، به همراه عمود منصف خط واصل بین مراکز.

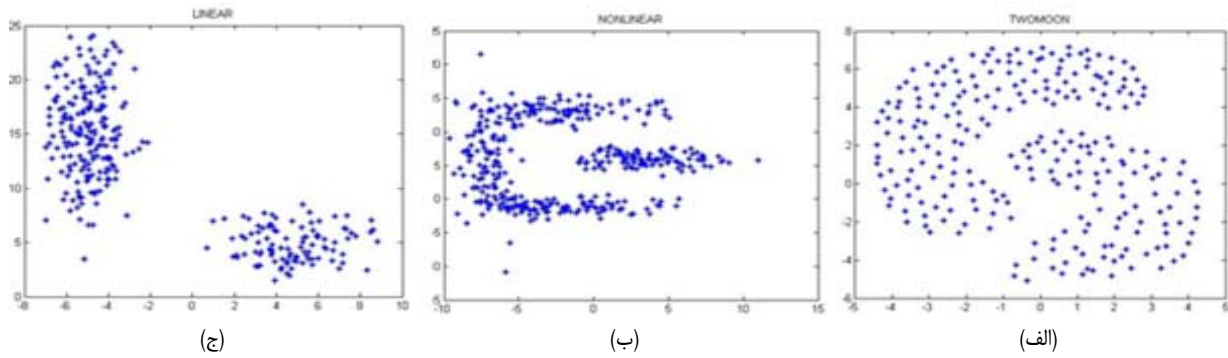
با توجه به مشخص شدن دسته‌ها در مرحله قبل، هدف در این مرحله تعیین مراکز دسته‌ها از طریق میانگین‌گیری بر روی اعضای دسته‌ها است. برای ساخت لایه زیرین GeSOM در مرکز هر دسته یک گره قرار داده می‌شود. توجه شود فقط گره‌هایی که در مراکز دسته‌ها ایجاد شده‌اند در شبکه مورد استفاده قرار می‌گیرند.

تمام اقداماتی که تا این مرحله انجام پذیرفت برای حذف افزونگی هندسی گره‌های لایه زیرین GeSOM بوده است. این هدف از طریق ادغام گره‌های هر دسته در مرکز آن دسته صورت پذیرفت تا بتوان به حداقل تعداد گره که معادل حداقل برچسب است دست یافت.

در این مرحله با توجه به (۱) برای گره‌های لایه پایین می‌بایست تعیین شعاع نمود. بر خلاف الگوریتم سابق که مقدار σ_c برای دسته‌های مجاور یکسان تعریف می‌شد، در الگوریتم پیشنهادی این مقدار برای هر گره با توجه به ابعاد دسته تعریف می‌شود و بنابراین برای دسته‌های بزرگ‌تر مقدار σ_c بیشتر از دسته‌های کوچک‌تر خواهد بود. به توجه شکل ۶ می‌توان به لزوم این کار پی برد. در این شکل دو دسته داده کروی که شعاع یکی چند برابر دیگری است و در کنار یکدیگر قرار گرفته‌اند، دیده می‌شود. داده‌هایی از دسته بزرگ‌تر که در نزدیکی مرکز دسته کوچک‌تر قرار گرفته‌اند و در شکل در سمت راست خط عمود منصف واصل بین دو مرکز قرار دارند، برای آن که به اشتباه برچسب دسته کوچک‌تر به آنها توسط شبکه تعلق نگیرد، می‌بایست شباهت بیشتری به مرکز دسته بزرگ‌تر نسبت به دسته کوچک‌تر داشته باشند. اما با توجه به (۱) اگر مقدار σ_c برای هر دو دسته یکسان در نظر گرفته شود، با توجه به فاصله کمتر این داده‌ها تا مرکز دسته کوچک‌تر نسبت به مرکز دسته خود، طبقه‌بند به اشتباه برچسب مرکز دسته کوچک را برای آنها در نظر خواهد گرفت. اما در حالت دوم که مقدار σ_c برای دسته بزرگ‌تر بیشتر از مقدار آن برای دسته کوچک‌تر در نظر گرفته می‌شود، میزان شباهت داده‌های مورد بحث به مرکز خود بیشتر از حالت قبل خواهد شد و طبقه‌بند در تعیین برچسب این داده‌ها دچار خطا نخواهد شد.

در مرحله درخواست برچسب و الگوریتم LMS با توجه به قضیه ۱ باید برای هر گره یک برچسب درخواست نمود (برچسب نزدیک‌ترین نمونه به مرکز هر دسته) و سپس با استفاده از الگوریتم LMS وزن‌های بین لایه‌های را آموزش داد. این مرحله در صورت لزوم چند بار تکرار می‌شود تا ضرایب بهینه به دست آید چرا که احتمال دارد الگوریتم در نقاط بهینه محلی گیر افتاده و ضرایب بر اساس مقدار بهینه تعیین نشوند.

پس از پایان یافتن مراحل آموزش طبقه‌بند، برای تعیین دسته داده‌های ورودی، طبقه‌بند می‌تواند مورد استفاده قرار گیرد اما به دلیل وجود احتمال رخداد خطا در طی فرایند آموزش به دو دلیل (۱) غیر خطی بودن داده‌ها در حالت کلی و (۲) احتمال رخداد اشتباه در مرحله خوشه‌بندی، می‌توان فرایند آموزش را برای بالابردن دقت ادامه داد. برای افزایش دقت می‌بایست



شکل ۷: سه پایگاه داده (الف) twomoon (ب) جدناپذیر خطی و (ج) جدناپذیر خطی.

جدول ۲: دنباله‌های ویدئوی استاندارد.

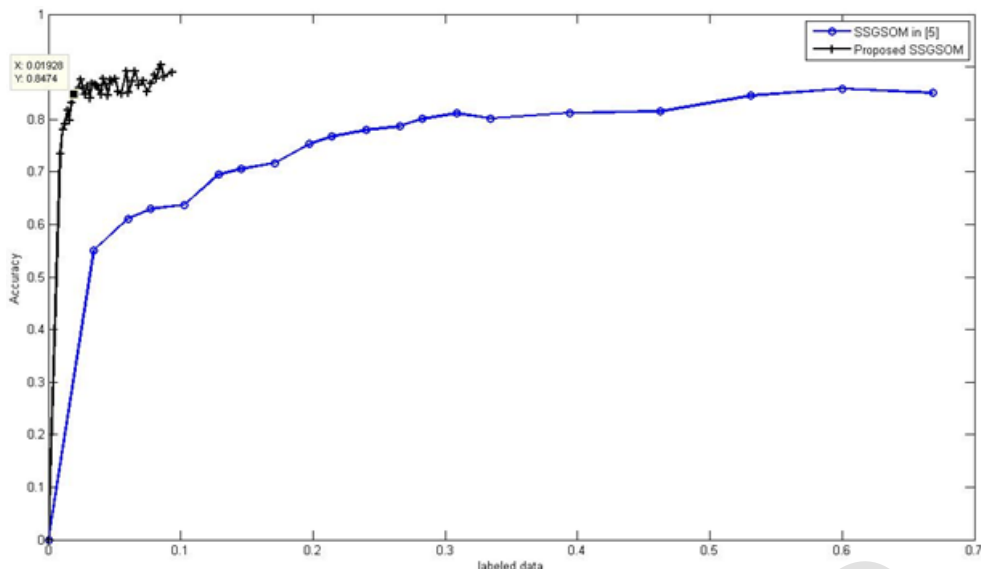
نام دنباله ویدئو	حرکت دوربین	تعداد فریم	اندازه فریم	حرکت اشیا
Container	نمای افق گرد	۳۰۰	۳۵۲×۲۸۸	انتقالی (کند)
Bus	نمای افق گرد	۳۰۰	۳۵۲×۲۸۸	انتقالی (سریع)
Coastguard	نمای افق گرد	۳۰۰	۳۵۲×۲۸۸	انتقالی
Flower	نمای افق گرد	۲۵۰	۳۵۲×۲۸۸	انتقالی
Football	ناچیز	۳۶۰	۳۵۲×۲۸۸	غیر انتقالی (سریع)
Mobile	نمای افق گرد	۳۰۰	۳۵۲×۲۸۸	انتقالی، چرخشی
Stefan	نمای افق گرد	۹۰	۳۵۲×۲۸۸	غیر انتقالی (سریع)
Tennis	نمای افق گرد، زوم	۱۵۰	۳۵۲×۲۸۸	غیر انتقالی
Carphone	ناچیز	۳۸۲	۱۷۶×۱۴۴	غیر انتقالی (سریع)
Foreman	چرخشی	۳۰۰	۳۵۲×۲۸۸	غیر انتقالی (سریع)
Hall monitor	ناچیز	۳۰۰	۳۵۲×۲۸۸	غیر انتقالی
Mother and daughter	ناچیز	۳۰۰	۳۵۲×۲۸۸	غیر انتقالی
News	ناچیز	۳۰۰	۳۵۲×۲۸۸	انتقالی، چرخشی
Paris	ناچیز	۱۰۶۵	۳۵۲×۲۸۸	غیر انتقالی، چرخشی (سریع)
Silent	ناچیز	۳۰۰	۳۵۲×۲۸۸	غیر انتقالی، چرخشی (سریع)
Suzie	ناچیز	۱۵۰	۱۷۶×۱۴۴	غیر انتقالی
Tempete	زوم	۲۶۰	۳۵۲×۲۸۸	غیر انتقالی
Waterfall	زوم	۲۶۰	۳۵۲×۲۸۸	انتقالی
Vector	نمای افق گرد	۳۰۰	۳۵۲×۲۸۸	غیر انتقالی (سریع)

جدول ۳: نتایج سه بار آزمایش حاصل از اعمال طبقه‌بند به سه مجموعه داده توصیف‌شده.

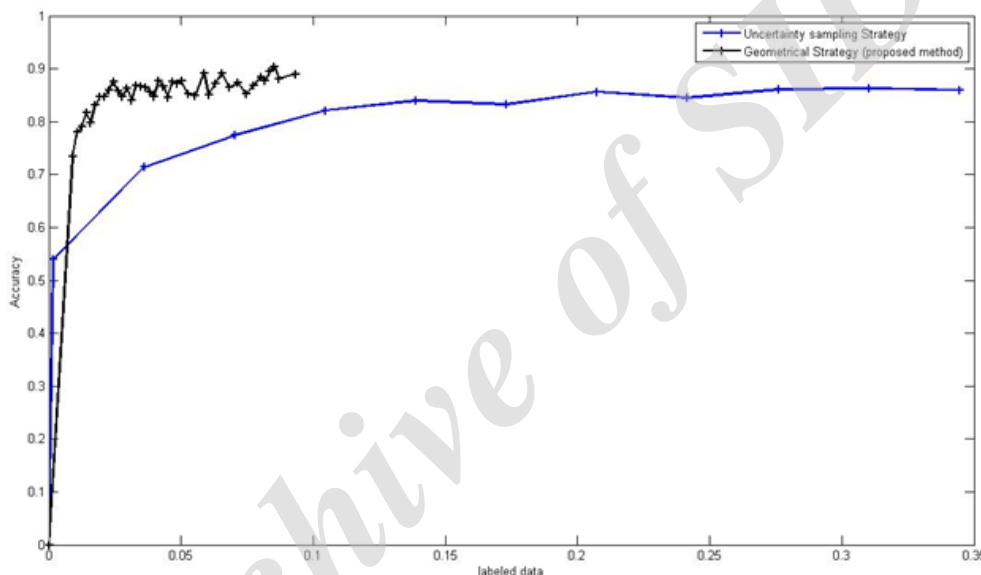
شماره آزمایش	مراحل تکرار	دسته شماره ۱ (Twomoon)		دسته شماره ۲ (جدناپذیر خطی)		دسته شماره ۳ (جدناپذیر خطی)	
		تعداد برچسب	دقت (%)	تعداد برچسب	دقت (%)	تعداد برچسب	دقت (%)
۱	مرحله اول	۲	۷۴	۲	۱۰۰	۵	۷۳
	مرحله دوم	۳	۷۸	۳	۱۰۰	۷	۸۱
	مرحله سوم	۵	۹۲	۴	۱۰۰	۹	۹۲
۲	مرحله اول	۲	۶۲	۲	۱۰۰	۶	۹۱
	مرحله دوم	۳	۶۴	۳	۹۸	۹	۹۶
	مرحله سوم	۴	۸۸	۵	۹۵	۱۳	۹۶
۳	مرحله اول	۲	۴۱	۲	۱۰۰	۷	۸۱
	مرحله دوم	۳	۸۸	۳	۱۰۰	۱۰	۷۷
	مرحله سوم	۴	۹۰	۵	۱۰۰	۱۲	۸۱

به دقت نهایی دست یابد. در مابقی حالات هر چند طبقه‌بند به مقدار بهینه دست نیافته است (به دلیل غیر خطی بودن داده‌ها) اما میزان برچسب مورد استفاده طبقه‌بند بسیار ناچیز است (این مهم در هنگام مقایسه میزان برچسب استفاده‌شده GeSOM با SSGSOM در ادامه مشخص می‌شود).

سه مجموعه داده توصیف‌شده در طی سه مرحله در جدول ۳ بر حسب میزان برچسب مصرف‌شده و دقت به دست آمده نمایش داده شده است. همان طور که مشهود است در مراحل دوم و سوم با مصرف برچسب بیشتر دقت طبقه‌بند افزایش یافته است. برای داده جدناپذیر خطی طبقه‌بند توانسته مطابق پیش‌بینی قضیه دو با حداقل برچسب ممکن (در اینجا دو)



شکل ۸: مقایسه دقت نهان‌کاوی پیشنهادی و SSGSOM استفاده‌شده در [۱۱] برای پایگاه داده ویدئوهای ارائه‌شده.



شکل ۹: دقت تشخیص استراتژی هندسی و مقایسه با استراتژی عدم قطعیت در یادگیری فعال برای پایگاه داده ویدئوهای ارائه‌شده.

عدم قطعیت در شکل ۹ به نمایش درآمده است. راهبرد عدم قطعیت هر چند نسبت به حالت بدون استفاده به نتایج بهتری دست یافته است (رسیدن به دقت نهایی با استفاده از ۲۰ درصد برچسب مورد استفاده نسبت به ۶۰ درصد گذشته) اما تا رسیدن به مقدار ۲ درصد راهبرد هندسی اختلاف چشم‌گیری وجود دارد. برای توضیح تفوق راهبرد هندسی بر راهبرد عدم قطعیت علاوه بر استفاده بهینه از افزونگی هندسی موجود در گره‌های لایه زیرین این نکته نیز قابل ذکر است که فرض اولیه راهبرد عدم قطعیت برای موفقیت، جداپذیری خطی داده‌ها است. در حالت کلی داده‌ها این گونه نیستند و برای همین در لایه زیرین از مفهوم کرنل با استفاده از توابع گوسی استفاده شده تا داده‌ها به فضایی منتقل شوند که در آنجا جداپذیری خطی باشد و در نتیجه بتوان با انتخاب نمونه‌های مرزی برای برچسب‌زنی توسط راهبرد عدم قطعیت به نتیجه رسید. اما با این وجود نیز داده‌ها جداپذیر خطی نخواهند شد زیرا همان طور که پیشتر توضیح داده شد مبنای افزایش گره با تابع گوسی در لایه زیرین، افزایش رزولوشن شبکه متناسب با حجم داده‌ها است بدین معنی که در مکان‌هایی که داده‌ها تراکم بیشتری دارند شبکه گره‌های بیشتری ایجاد می‌کند. در این صورت چون هدف از ایجاد گره خطی‌سازی در محیطی بالاتر نیست

اکنون نوبت مقایسه دقت نهان‌کاوی پیشنهادی با کار انجام‌گرفته در [۱۱] بر اساس پایگاه داده ویدیوی ارائه‌شده در جدول ۲ می‌باشد. شکل ۸ نتیجه این مقایسه را نشان می‌دهد. محور افقی نسبت تعداد برچسب استفاده‌شده بر حسب تعداد کل داده می‌باشد و محور عمودی دقت نهان‌کاوی را نشان می‌دهد. نهان‌کاوی مورد استفاده قرار گرفته در [۱۱] (طبقه‌بند SSGSOM) با مصرف حدود ۶۰ درصد داده‌ی برچسب‌دار به دقت نهایی که حدود ۸۵٪ می‌باشد دست یافته است. در حالی که نهان‌کاوی پیشنهادی (طبقه‌بند GeSOM) توانسته با صرف تنها حدود ۲٪ داده برچسب‌دار به همان دقت دست یابد.

با توجه به این که برای رسیدن به حداقل برچسب از ایده یادگیری فعال استفاده نمودیم از این منظر می‌توان مقایسه‌ای بین راهبرد پرس و جوی هندسی (با توجه به این که برای رسیدن به حداقل برچسب مورد استفاده فضای هندسی بردار ویژگی مورد تحلیل قرار داده شد و سعی شد افزونگی هندسی حذف شود، این راهبرد، هندسی نامیده شد) مطرح‌شده در این تحقیق با پرس و جوهای موجود ترتیب داد. برای این مقایسه راهبرد پرس و جوی عدم قطعیت که عمومی‌ترین راهبرد می‌باشد انتخاب و شبیه‌سازی شده است. نتایج مقایسه راهبرد هندسی در مقایسه با راهبرد

اثبات) برای اثبات کافی است نشان داده شود اولاً هر بردار دلخواه مانند A را می‌توان بر اساس ترکیب خطی m بردار مستقل دیگر (بردارهای پایه) بیان نمود و ثانیاً هیچ یک از بردارهای پایه را نمی‌توان از ترکیب خطی سایرین به دست آورد

$$A = (a_1, a_2, \dots, a_m) \quad (\text{پ-۱})$$

اگر بردارهای پایه به صورت زیر تعریف شود

$$b_1 = (1, 0, \dots, 0) \quad (\text{پ-۲})$$

$$b_2 = (0, 1, \dots, 0) \quad (\text{پ-۳})$$

و به همین ترتیب

$$b_m = (0, 0, \dots, 1) \quad (\text{پ-۴})$$

آن گاه می‌توان بردار A را این گونه بیان کرد

$$A = a_1 b_1 + a_2 b_2 + \dots + a_m b_m \quad (\text{پ-۵})$$

بنابراین اثبات قسمت اول قضیه کامل شد. اما در مورد قسمت دوم می‌توان این گونه استدلال نمود: هر یک از بردارهای پایه در مؤلفه‌ای دارای مقدار یک است که سایر بردارها همگی دارای مقدار صفر هستند و هیچ گاه ترکیب خطی چند صفر، یک نخواهد شد. بدین ترتیب اثبات قضیه کامل می‌شود.

با توجه به (پ-۶) که به صورت زیر است

$$w_{cm}(t+1) = w_{cm}(t) + \eta(\xi_m - PS_m) o_c \quad (\text{پ-۶})$$

و برای آموزش وزن‌های بین لایه‌ای از آن استفاده می‌شود خواهیم داشت

$$\Delta w_c(t+1) = k_m o_c \quad (\text{پ-۷})$$

که می‌توان این رابطه را به صورت ماتریسی به شکل (پ-۸) نوشت

$$\begin{pmatrix} \Delta w_{c1} & \dots & \Delta w_{cm} \\ \vdots & \ddots & \vdots \\ \Delta w_{c1} & \dots & \Delta w_{cm} \end{pmatrix} = \begin{pmatrix} k_1 o_1 & \dots & k_m o_1 \\ \vdots & \ddots & \vdots \\ k_1 o_c & \dots & k_m o_c \end{pmatrix} \quad (\text{پ-۸})$$

حال با توجه به شکل ۳ مشخص است بردار O_c دارای c بعد است که هر بعد آن یک گره در لایه پایین شکل مذکور می‌باشد. بر اساس قضیه مطرح شده برای بردار O_c حداکثر c (تعداد گره‌ها) بردار مستقل وجود دارد. این بردارهای مستقل برای سادگی به صورت زیر تعریف می‌شود

$$o_1 = (1, 0, \dots, 0) \quad (\text{پ-۹})$$

$$o_2 = (0, 1, \dots, 0) \quad (\text{پ-۱۰})$$

و به همین ترتیب

$$o_c = (0, 0, \dots, 1) \quad (\text{پ-۱۱})$$

حال هنگامی که این بردارهای مستقل به فرایند آموزش وزن‌ها طبق (پ-۸) سپرده شوند هر بردار یکی از سطرها را بر اساس برچسب داده ورودی آموزش خواهند داد. بنابراین تمامی درایه‌ها آموزش می‌یابند. حال اگر الگوی دیگری به شبکه اعمال شود، وزنی برای آموزش آن وجود نخواهد داشت و شبکه ناچار خواهد بود با تغییر در مقدار یکی از وزن‌های آموزش یافته فرایند آموزش را برای آن نمونه طی کند و این پاک شدن وزن سابق به معنی از بین رفتن آموزش الگوی سابق است. بنابراین در این حالت نمی‌توان بیشتر از c برچسب از l برچسب موجود را به شبکه آموزش داد. پس تعداد برچسب‌ها نمی‌تواند از تعداد گره‌ها بیشتر باشد.

لزوماً منتج به جدایی‌پذیری خطی نخواهد شد و در نتیجه راهبرد عدم قطعیت به نتیجه مطلوب دست نمی‌یابد (که در عمل نیز همین گونه شد). این در حالی است که در راهبرد هندسی برای خطی‌سازی داده‌ها در محیط با توابع گوسی از ویژگی عمودبودن بردارها استفاده شده است. در واقع گره‌های با تابع گوسی در لایه زیرین به نحوی چیده شده‌اند تا تصویر ناشی از الگوهای متفاوت در فضای توابع گوسی عمود بر یکدیگر شود و هیچ دو بردار متفاوتی که عمود بر یکدیگر نباشند یافت نشود. در ضمن برای مصرف بهینه برچسب از حداقل ممکن توابع گوسی استفاده شده است (برای اثبات این گزاره باید دقت داشت که لازمه از بین رفتن افزودنی هندسی گره‌ها در لایه زیرین عمودبودن بردارهای ساخته شده از توابع گوسی آنهاست).

۴- جمع‌بندی

در این مقاله طبقه‌بند GeSOM مبتنی بر یادگیری فعال جهت استفاده حداقلی از داده برچسب‌دار معرفی و کاربرد آن در نهان‌نگاری ویدئو نشان داده شد. برای دستیابی به حداقل برچسب از افزودنی هندسی موجود در فضای ویژگی لایه پایین استفاده شد. بر این مینا برای به حداقل رساندن افزودنی هندسی که مبتنی بر دو قضیه ارائه شده منجر به حداقل رسیدن برچسب مورد استفاده در فرایند آموزش شبکه می‌شود دسته‌هایی که گره‌ها تشکیل می‌دهند شناسایی و اعضایشان در مرکز دسته ادغام شد. نتایج حاصل از شبیه‌سازی که در بخش ۳ ارائه گردید (شکل ۸) نشان از برتری چشم‌گیر طبقه‌بند GeSOM نسبت به طبقه‌بند SSGSOM دارد.

اما برای بهبود عملکرد الگوریتم در مورد داده‌های غیر خطی سه پیشنهاد برای تحقیقات آتی ارائه می‌شود:

۱) توسعه تعریف σ_c در (۱) از حالت تک‌بعدی به چندبعدی با هدف خطی‌سازی داده‌های غیر خطی

۲) بازتعریف (۱) بر اساس زنجیره شباهت

۳) استفاده از نسخه شبه‌ناظر kNN به جای GeSOM

۵- سپاس‌گزاری

در انتها وظیفه خود می‌دانیم از جناب مهندس جواد مرتضوی به دلیل در اختیار گذاشتن تمامی پایگاه داده‌های خود که با زحمت بسیار آنها را تهیه نموده بود تشکر نماییم.

پیوست

در پیوست به اثبات دو قضیه که در بخش ۲-۳ ارائه شد پرداخته می‌شود. برای درک شهودی بهتر شکل ۳ را در نظر بگیرید.

قضیه ۱) در فرایند آموزش شبکه GeSOM برای رسیدن به حداقل برچسب لازم باید برای هر گره که در لایه پایین وجود دارد، دقیقاً یک برچسب اختصاص داد.

اثبات) برای اثبات از برهان خلف استفاده می‌شود. دو حالت برای فرض خلف ممکن است: یکی این که فرض شود برای یکی از گره‌ها یا چند گره بیش از یک برچسب اختصاص یابد یعنی تعداد برچسب‌ها (l) از تعداد گره‌ها (c) بیشتر باشد. در این حالت اثبات خواهد شد که شبکه ظرفیت آموزش حداکثر c الگوی متمایز را خواهد داشت و برای $l-c$ الگوی باقیمانده فرایند آموزش امکان‌پذیر نیست. برای اثبات این مطلب قضیه پ-۱ مطرح می‌شود:

قضیه پ-۱) در یک فضای m بعدی حداکثر m بردار مستقل وجود دارد.

مراجع

- [1] N. Zarmehi and M. A. Akhaee, "Digital video steganalysis toward spread spectrum data hiding," *IET Image Processing*, vol. 10, no. 1, pp. 1-8, Dec. 2015.
- [2] A. Kolakalur, I. Kagalidis, and B. Vuksanovic, "Wavelet based color video steganography," *IACSIT International J. of Engineering and Technology*, vol. 8, no. 3, pp. 165-169, Jun. 2016.
- [3] H. Karimi, M. G. Shayesteh, and M. A. Akhaee, "Steganalysis of JPEG images using enhanced neighbouring joint density features," *IET Image Processing*, vol. 9, no. 7, pp. 545-552, Jul. 2015.
- [4] U. Budia, D. Kundur, and T. Zourntos, "Digital video steganalysis exploiting statistical visibility in the temporal domain," *IEEE Trans. on Information Forensics and Security*, vol. 1, no. 4, pp. 502-516, Dec. 2006.
- [5] J. S. Jainsky, D. Kundur, and D. R. Halverson, "Towards digital video steganalysis using asymptotic memoryless detection," in *Proc. of the 9th Int. Workshop on Multimedia and Security*, pp. 161-168, Dallas, TX, USA, 20-21 Sept. 2007.
- [6] V. Pankajakshan and A. T. S. Ho, "Improving video steganalysis using temporal correlation," in *Proc. of the 3rd Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing, IIH-MSP'07*, vol. 1, pp. 287-290, Kaohsiung, Taiwan, 26-28 Nov. 2007.
- [7] Y. Su, C. Zhang, L. Wang, and C. Zhang, "A new video steganalysis based on mode detection," in *Proc. of the Int. Conf. on Audio, Language and Image Processing*, pp. 1507-1510, Shanghai, China, 7-9 Jul. 2008.
- [8] K. Tasdemir, F. Kurugollu, and S. Sezer, "Spatio-temporal rich model-based video steganalysis on cross sections of motion vector planes," *IEEE Trans. on Image Processing*, vol. 25, no. 7, pp. 3316-3328, Jul. 2016.
- [9] G. Xuan, et al., "Steganalysis based on multiple features formed by statistical moments of wavelet characteristic functions," in *Proc. of the 7th Information Hiding Workshop*, Lecture Notes in Computer Science, vol. 3727, pp. 262-277, Jun. 2005.
- [10] Y. Q. Shi, et al., "Steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image, and neural network," in *Proc. of the IEEE Int. Conf. on Multimedia and Expo*, pp. 269-272, Amsterdam, Netherlands, 6-6 Jul. 2005.
- [11] ج. مرتضوی مهریزی، م. خادمی و ه. صدوقی یزدی، "نهان‌کاوی کور ویدئو با رویکرد یادگیری شبه‌ناظر برای الگوریتم‌های نهان‌نگاری ویدئوی مبتنی بر بردارهای حرکت،" *تشریح مهندسی برق و مهندسی کامپیوتر ایران، ب- مهندسی کامپیوتر*، سال ۱۳، شماره ۲، صص. ۱۰۰-۹۹، زمستان ۱۳۹۴.
- [12] Y. Cao, X. Zhao, and D. Feng, "Video steganalysis exploiting motion vector reversion-based features," *IEEE Trans. on Signal Processing*, vol. 19, no. 1, pp. 35-38, Nov. 2012.
- [13] Y. Deng, Y. Wu, H. Duan, and L. Zhou, "Digital video steganalysis based on motion vector statistical characteristics," *International Journal for Light and Electron Optics*, vol. 124, no. 14, Jul. pp. 1705-1710, 2013.
- [14] Y. S. Su, C. Zhang, and C. Zhang, "A video steganalytic algorithm against motion vector-based steganography," *Signal Process.*, vol. 91, no. 18, pp. 1901-1909, Aug. 2011.
- [15] X. Xu, J. Dong, W. Wang, and T. Tan, "Video steganalysis based on the constraints of motion vectors," in *Proc. IEEE Int. Conf. on Image Processing, ICIP'13*, pp. 4422-4426, Melbourne, VIC, Australia, 15-18 Sept. 2013.
- [16] H. A. Aly, "Data hiding in motion vectors of compressed video based on their associated prediction error," *IEEE Trans. on Information Forensics and Security*, vol. 6, no. 1, pp. 14-18, Nov. 2011.
- [17] C. Xu, X. Ping, and T. Zhang, "Steganography in compressed video stream," in *Proc. IEEE First Int. Conf. Innovative Computing, Information and Control, ICICIC'06*, vol. 1, pp. 269-272, Beijing, China, 30 Aug.-1 Sept. 2006.
- [18] D. Y. Fang and L. W. Chang, "Data hiding for digital video with phase of motion vector," in *Proc. IEEE Int. Symp. Circuits and Systems*, pp. 1422-1425, Island of Kos, Greece, 21-24 May 2006.
- [19] D. Alahakoon and K. Halgamunge, "Dynamic self-organizing maps with controlled growth for knowledge discovery," *IEEE Trans. on Neural Networks*, vol. 11, no. 3, pp. 601-614, May 2000.
- [20] A. Hsu and S. K. Halgamunge, "Class structure visualization with semi-supervised growing self-organizing maps," *Neurocomputing*, vol. 71, no. 16-18, pp. 3124-3130, Oct. 2008.

فرض خلف دیگر این که به یک یا چند گره هیچ برچسبی تخصیص داده نشود. در این حالت گرهی که هیچ برچسبی برای آموزش بدان اختصاص نیافته را در نظر می‌گیریم. با توجه به این که ضرایب بین لایه‌های فقط هنگامی وارد فرایند آموزش می‌شوند که یک داده برچسب‌دار وارد فرایند آموزش شود، بنابراین در این حالت وزن‌های بین لایه‌های مربوط به گره مورد بحث از مقادیر اولیه خود که به صورت تصادفی انتخاب شده‌اند تغییر نخواهند یافت و برچسب‌های تخصیص داده شده توسط الگوریتم به داده‌های مربوط به گره کاملاً تصادفی هستند. با توجه به مطالب گفته‌شده در این حالت جواب بهینه برای این گره و داده‌های مشابه به آن به دست نیامده و احتمال رخداد خطا وجود دارد، پس این فرض نیز قابل قبول نیست و اثبات قضیه کامل می‌شود.

قضیه ۲) فرض کنید در یک فضای M بعدی تعداد زیادی داده که در N کلاس مختلف چیده شده‌اند موجود هستند. با فرض این که دسته‌های مختلف به صورت خطی جداپذیر باشند، حداقل تعداد لازم گره برای ایجاد شبکه GeSOM که منجر به حداقل تعداد مورد نیاز برچسب بشود برابر با تعداد دسته‌ها (N) خواهد بود.

اثبات) برای اثبات قضیه کافی است نشان داده شود با احتساب فرض مسئله، با تعداد N گره می‌توان شبکه را بدون رخداد هیچ گونه خطایی آموزش داد و با تعداد $N-1$ یا هر مقدار دیگری که کوچک‌تر از N گره نمی‌توان شبکه را بدون خطا آموزش داد و احتمال رخداد خطا مقداری غیر صفر خواهد داشت.

ابتدا نشان داده می‌شود با N گره مستقل می‌توان شبکه را بدون رخداد هیچ گونه خطایی آموزش داد. هر یک از N گره را در مرکز یکی از دسته‌های موجود قرار داده و نزدیک‌ترین داده موجود به گره برای آموزش انتخاب می‌شود. با توجه به (۱) برای یافتن میزان شباهت هر داده به گرهی که فرایند آموزش با محوریت آن انجام می‌شود با فرض ثابت‌بودن مقدار σ ، پارامتر فاصله تا گره تعیین‌کننده است. بنابراین میزان شباهت هر داده با گره مورد نظر با فاصله آن داده از گره رابطه معکوس دارد. از طرف دیگر با توجه به فرض جداپذیری خطی دسته‌ها تمامی اعضای هر دسته دارای کمترین فاصله از مرکز همان دسته نسبت به سایر مراکز دسته‌ها هستند. بنابراین میزان شباهت هر داده به گرهی که متعلق به همان دسته است از سایر گره‌ها بیشتر می‌باشد. بر این اساس پس از طی فرایند آموزش شبکه، تمامی داده‌هایی که متعلق به یک دسته مشخص هستند، برچسب متعلق به مرکز همان دسته را بدون رخداد هیچ خطایی خواهند پذیرفت.

حال فرض می‌کنیم تعداد گره‌ها کمتر از تعداد دسته‌ها (N) باشد. در این حالت پس از تقسیم گره‌ها در بین دسته‌ها با هر روش ممکن، با توجه به مفهوم اصل لانه کبوتری (اصل لانه کبوتری می‌گوید اگر تعدادی کبوتر و تعدادی لانه وجود داشته باشد که کبوترها در آنها منزل گزینند اگر تعداد کبوترها از تعداد لانه‌ها بیشتر باشد حتماً لانه‌ای یافت می‌شود که در آن بیشتر از یک کبوتر وجود داشته باشد) حداقل یک (یا شاید چند) دسته پیدا خواهد شد که هیچ گرهی در آن قرار ندارد. بدیهی است که برای داده‌های مربوط به این دسته با توجه به نبود هیچ گونه برچسبی فرایند آموزش اجرا نشود و وزن‌های بهینه بین لایه‌های به دست نیاید و در همان مقادیر اولیه خود که به صورت تصادفی انتخاب شده است باقی بماند. در این حالت به طور قطع شبکه نمی‌تواند در مورد برچسب این دسته از داده‌ها با احتمال خطای صفر اظهار نظر کند. بدین ترتیب اثبات قضیه کامل می‌شود.

علی محی‌الدینی شاهم‌آبادی‌پور در سال‌های ۱۳۹۰ و ۱۳۹۳ به ترتیب مدارک کارشناسی و کارشناسی ارشد خود را در رشته مهندسی برق گرایش مخابرات از دانشگاه فردوسی مشهد دریافت کرد. وی از سال ۱۳۹۵ دانشجوی دکتری مهندسی برق گرایش مخابرات- سیستم دانشگاه باهنر کرمان شده است. زمینه تحقیقاتی وی هوش مصنوعی و هوش جمعی در حوزه‌ی توده رباتیکی است.

مرتضی خادمی تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد مهندسی برق به ترتیب در سال‌های ۱۳۶۴ و ۱۳۶۶ در دانشگاه صنعتی اصفهان به پایان رسانده است. نامبرده از سال ۱۳۶۶ الی ۱۳۷۰ به عنوان عضو هیات علمی (مربی) در دانشگاه فردوسی مشهد به کار مشغول بود. پس از آن به دوره دکترای مهندسی برق در دانشگاه ولونگونگ (استرالیا) وارد گردیده و در سال ۱۳۷۴ موفق به اخذ درجه دکترا در مهندسی برق از دانشگاه مذکور گردید. دکتر خادمی از سال ۱۳۷۴ مجدداً در دانشکده مهندسی دانشگاه فردوسی مشهد مشغول به فعالیت گردید و اینک نیز استاد این دانشکده است. زمینه‌های علمی مورد علاقه نامبرده شامل موضوعاتی مانند مخابرات ویدئویی، فشرده‌سازی ویدئو، پردازش تصویر، پردازش سیگنال‌های پزشکی و پنهان‌سازی اطلاعات در ویدئو می‌باشد.

هادی صدوقی یزدی هم‌اکنون استاد گروه مهندسی کامپیوتر در دانشگاه فردوسی مشهد می‌باشد. ایشان در سال ۱۳۷۳ مدرک کارشناسی مهندسی برق خود را از دانشگاه فردوسی مشهد و کارشناسی ارشد و دکترای مهندسی برق را به ترتیب در سال ۱۳۷۵ و ۱۳۸۴ از دانشگاه تربیت مدرس دریافت نمود. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: شناسایی آماری الگو، یادگیری ماشین، بینایی ماشین، پردازش سیگنال.

- [21] B. Settles, *Active Learning Literature Survey*, Computer Sciences Technical Report 1648 University of Wisconsin-Madison Updated on: Jan. 2010.
- [22] B. Settles, R. J. Brachman, W. W. Cohen, and T. Dietterich, *Active Learning, Synthesis Lectures on Artificial Intelligence and Machine Learning*, Morgan & Claypool Publishers, 2012.
- [23] H. Wu and S. Prasad, "Dirichlet process based active learning and discovery of unknown classes for hyperspectral image classification," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4882-4895, Aug. 2016.
- [24] Q. Shi, B. Du, and L. Zhang, "Spatial coherence based batch-mode activelearning for remote sensing image classification," *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2037-2050, Jul. 2015.
- [25] Z. Zhang, E. Pasolli, H. Lexie Yang, and M. M. Crawford, "Multimetric active learning for classification of remote sensing data," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 7, pp. 1007-1011, Jul. 2016.
- [26] L. Wan, K. Tang, M. Li, Y. Zhong, and A. K. Qin, "Collaborative active and semisupervised learning for hyperspectral remote sensing image classification," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2384-2396, May 2015.
- [27] E. Pasolli, H. L. Yang, and M. M. Crawford, "Active-metric learning for classification of remotely sensed hyperspectral images," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 1925-1939, Apr. 2016.
- [28] H. Lin and P. Li, "Circuit performance classification with active learning guided sampling for support vector machines," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 9, pp. 1467-1480, Sep. 2015.
- [29] A. Mehrizi and H. Sadoghi Yazdi, "Semi-supervised GSOM integrated with extreme learning machine intelligent data analysis," *Intelligent Data Analysis*, vol. 20, no. 5, pp. 1115-1132, Sept. 2016.

Archive of SID