

نهان کاوی کور ویدئو با رویکرد یادگیری شبه‌ناظر برای الگوریتم‌های نهان‌نگاری ویدئوی مبتنی بر بردارهای حرکت

جواد مرتضوی مهریزی، مرتضی خادمی و هادی صدوقی یزدی

شناسایی حضور پیام سری در درون یک ویدئوی میزبان به صورت یک مسئله طبقه‌بندی مدل می‌شود. الگوریتم‌های نهان‌کاوی، ویدئوهای مشکوک را به عنوان ورودی دریافت کرده و آنها را تحت عنوان ویدئوی میزبان (پاک) یا ویدئوی نهان‌نگاری شده (آلوده) طبقه‌بندی می‌کنند. بنابراین یک الگوریتم نهان‌کاوی کور ویدئو شامل دو مرحله اصلی استخراج ویژگی و طراحی طبقه‌بند است [۱]. در هنگام جاسازی پیام سری در درون یک ویدئوی میزبان برخی از مشخصه‌های آماری ویدئو میزبان تغییر خواهد کرد به نحوی که باعث کاهش کیفیت تصویر و همبستگی بین فریم‌های مجاور می‌شود. بنابراین در فرایند نهان‌کاوی کور ویدئو ویژگی‌هایی استخراج می‌شود که مقدار آنها در زمان حضور پیام سری نسبت به زمانی که پیام سری وجود ندارد متفاوت باشد. در گام بعد این ویژگی‌ها برای طراحی طبقه‌بند مورد استفاده قرار می‌گیرند.

بررسی مقالاتی که تاکنون در زمینه نهان‌کاوی کور ویدئو ارائه شده است نشان می‌دهد در اکثر موارد برای طراحی طبقه‌بند از الگوریتم‌های یادگیری باناظر استفاده شده است. طبقه‌بند SVM پرکاربردترین الگوریتم یادگیری باناظر در فرایند نهان‌کاوی کور ویدئو است [۲] تا [۱۰]. یکی از ویژگی‌های مهم این روش، پایه ریاضی قوی آن است. در این روش، ابرصفحه طبقه‌بند با حل یک مسئله بهینه‌سازی با تابع محدب به دست می‌آید. یکی دیگر از الگوریتم‌های باناظر و پرکاربرد در فرایند نهان‌کاوی کور ویدئو، روش طبقه‌بندی k نزدیک‌ترین همسایه (KNN) است [۴] و [۱۱] تا [۱۳]. الگوریتم‌های یادگیری باناظر اگرچه دارای دقت و تفسیرپذیری بالایی هستند ولی در فرایند نهان‌کاوی کور ویدئو مشکلاتی دارند. اول این که در این دسته از روش‌های طبقه‌بندی تمام داده‌ها آموزشی دارای برچسب هستند و داده‌های بدون برچسب از فرایند آموزش کنار گذاشته می‌شوند. با توجه به تنوع بالای الگوریتم‌های نهان‌نگاری و وجود الگوریتم‌های نهان‌نگاری‌ای که کد آنها در دسترس نیست، فراهم‌آوردن داده‌های برچسب‌دار نیازمند هزینه، زمان و کار انسانی زیادی می‌باشد. بنابراین در مرحله آموزش تنها می‌توان از تعداد محدودی از الگوریتم‌های نهان‌نگاری که کد آنها در دسترس است استفاده کرد و نمی‌توان با اطمینان درباره کارایی طبقه‌بند برای الگوریتم‌های نهان‌نگاری‌ای که کد آنها در دسترس نیست و در مرحله آموزش در نظر گرفته نشده‌اند صحبت کرد. دومین مشکل الگوریتم‌های یادگیری باناظر رایج (مانند SVM و KNN) این است که آنها برون‌خطی هستند و هزینه آموزش مجدد آنها بالا است. برای آموزش مجدد یک الگوریتم طبقه‌بندی برون‌خطی به ازای یک مجموعه آموزشی جدید، بایستی یک مجموعه آموزشی بزرگ شامل داده‌های آموزشی قدیمی و جدید تشکیل شود و طبقه‌بند به وسیله آن آموزش داده شود. به عبارت دیگر نمی‌توان

چکیده: اکثر الگوریتم‌هایی که تا این زمان در زمینه نهان‌کاوی کور ویدئو ارائه شده‌اند منطبق بر یادگیری باناظر بوده و هزینه تولید داده‌های برچسب‌دار در آنها بالا می‌باشد. تحت این شرایط در فرایند نهان‌کاوی کور ویدئو تنها می‌توان از تعداد محدودی از الگوریتم‌های نهان‌نگاری که کد آنها در دسترس است برای آموزش طبقه‌بند استفاده کرد. بنابراین نمی‌توان درباره کارایی نهان‌کاوی برای شناسایی آن دسته از الگوریتم‌های نهان‌نگاری ویدئویی که کد آنها در دسترس نیست، مطمئن بود. همچنین الگوریتم‌های رایج عموماً برون‌خط می‌باشند و بنابراین آموزش مجدد سیستم زمان‌بر بوده و نمی‌توان سیستم را به صورت برخط بروز کرد. برای حل این مشکلات یک روش جدید نهان‌کاوی کور ویدئو با رویکرد یادگیری شبه‌ناظر در این مقاله ارائه شده است. در روش پیشنهادی با توجه به رفع محدودیت برچسب‌دار بودن داده‌های آموزشی، کارایی طبقه‌بند برای آن دسته از الگوریتم‌های نهان‌نگاری که کد آنها در دسترس نیست بهبود می‌یابد. همچنین ثابت می‌شود که روش پیشنهادی نسبت به روش‌های متداول در فرایند نهان‌کاوی کور ویدئو پیچیدگی زمانی کمتری داشته و یک روش برخط بهینه است. نتایج شبیه‌سازی بر روی پایگاه داده استاندارد نشان می‌دهد که روش پیشنهادی علاوه بر مزایای فوق، دارای دقت آشکارسازی قابل مقایسه با روش‌های متداول روز است.

کلیدواژه: شبکه خودسازمان‌ده پویای شبه‌ناظر، نهان‌کاوی کور ویدئو، نهان‌نگاری ویدئو، یادگیری شبه‌ناظر.

۱- مقدمه

امروزه توسعه اینترنت و پیشرفت سیستم‌های مخابراتی و نرم‌افزارهای فشرده‌سازی ویدئو باعث شده که اطلاعات ویدئو به راحتی و با هزینه‌ای کم بر روی شبکه انتقال داده شوند. با توجه به این نکته، نهان‌نگاری ویدئو^۱ به عنوان یکی از تکنیک‌های پنهان‌سازی اطلاعات بسیار مورد توجه قرار گرفته است. استفاده از ویدئو به عنوان میزبان در فرایند نهان‌نگاری باعث افزایش ظرفیت و کاهش نرخ جاسازی پیام می‌شود. بنابراین شناسایی حضور پیام سری در یک ویدئوی نهان‌نگاری شده کار آسانی نخواهد بود.

با گسترش الگوریتم‌های نهان‌نگاری ویدئو الگوریتم‌های نهان‌کاوی کور ویدئو^۲ نیز برای مقابله با آنها توسعه یافته‌اند. در نهان‌کاوی کور ویدئو،

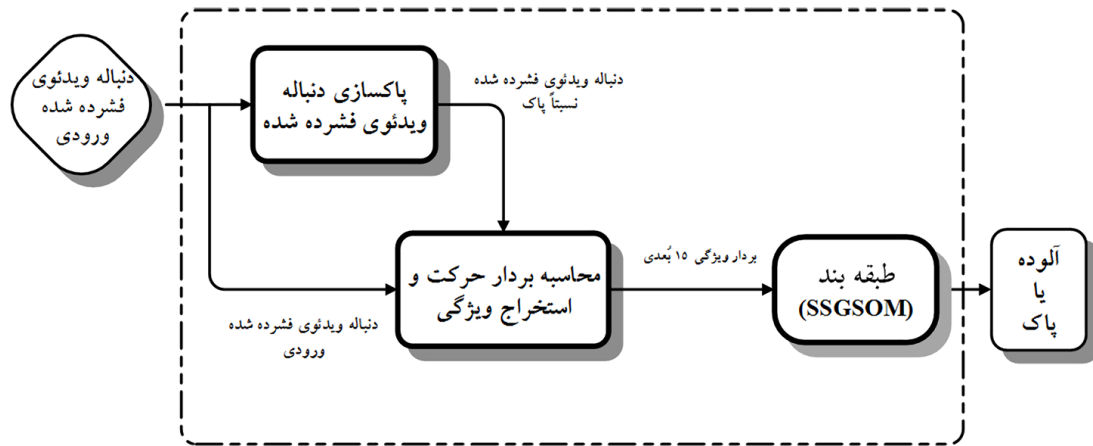
این مقاله در تاریخ ۶ بهمن ماه ۱۳۹۲ دریافت و در تاریخ ۱۹ آبان ماه ۱۳۹۳ بازنگری شد.

جواد مرتضوی مهریزی، دانشکده مهندسی، گروه برق- مخابرات، دانشگاه فردوسی مشهد، (email: javadmortazavi2005@yahoo.com).

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

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

1. Video Steganography
2. Video Blind Steganalysis



شکل ۱: بلوک دیاگرام نهان کاوی کور ویدئوی پیشنهادی.

متداول در نهان کاوی کور ویدئو پیچیدگی زمانی پایینی دارد و کاملاً برخط است و می‌تواند سیستم را به ازای یک داده آموزشی جدید بدون هزینه اضافی بروز کند. در نتیجه هزینه آموزش مجدد سیستم به شدت کاهش می‌یابد.

در قسمت ۲ این مقاله روش پیشنهادی در زمینه نهان کاوی کور ویدئو ارائه خواهد شد. در قسمت ۳ نحوه آموزش شبکه خودسازمان‌ده پویای شبه‌ناظر در فرایند نهان کاوی کور ویدئو ارائه می‌شود. در قسمت ۴ ضمن بیان دو اصل، ثابت می‌شود روش پیشنهادی نسبت به روش‌های مرسوم، پیچیدگی زمانی کمتری دارد. در قسمت ۵ نتایج شبیه‌سازی و در قسمت ۶ نتیجه‌گیری بیان می‌شود.

۲- نهان کاوی کور ویدئوی پیشنهادی

بلوک دیاگرام روش پیشنهادی در شکل ۱ نشان داده شده است. در این پژوهش هدف ارائه یک روش نهان کاوی کور ویدئو با رویکرد یادگیری شبه‌ناظر برای الگوریتم‌های نهان‌نگاری ویدئوی مبتنی بر بردارهای حرکت می‌باشد. دلیل انتخاب این دسته از الگوریتم‌های نهان‌نگاری ویدئو این است که تولید داده برچسب‌دار در آنها مشکل بوده و این می‌تواند مزایای استفاده از الگوریتم‌های یادگیری شبه‌ناظر را بهتر نشان دهد. شبیه‌سازی این دسته از الگوریتم‌های نهان‌نگاری ویدئو نیازمند آگاهی کامل از الگوریتم‌های فشرده‌سازی ویدئو (مانند سری MPEG) است. همچنین چون در این دسته از الگوریتم‌های نهان‌نگاری مشخصات آماری ضرایب فریم‌های ویدئو به صورت غیر مستقیم تحت تأثیر جاسازی پیام سری قرار می‌گیرند، کار شناسایی حضور پیام سری نسبت به سایر روش‌های نهان‌نگاری ویدئو مشکل‌تر است.

همان‌طور که در شکل ۱ نشان داده شده است روش پیشنهادی شامل سه مرحله است. اولین مرحله، پاک‌سازی دنباله ویدئوی فشرده‌شده است. مرحله دوم از روش پیشنهادی، محاسبه بردار حرکت و استخراج ویژگی است که در این مرحله بردار ویژگی برای تعداد مشخصی از فریم‌ها تولید می‌شود. مرحله آخر، طراحی طبقه‌بند و قراردادن ویدئوی مشکوک در یکی از دو دسته پاک یا آلوده می‌باشد. در ادامه سه مرحله روش پیشنهادی با جزئیات توضیح داده می‌شود.

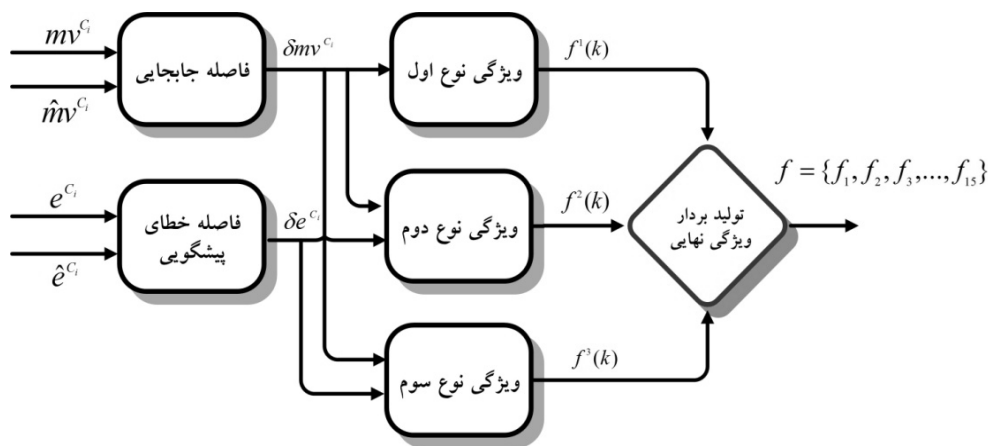
۲-۱ پاک‌سازی دنباله ویدئوی فشرده‌شده

در نهان کاوی کور ویدئو نوع الگوریتم نهان‌نگاری مشخص نمی‌باشد و هیچ‌گونه اطلاعاتی درباره ویدئوی نهان‌نگاری شده در اختیار نیست. با توجه به این محدودیت، برای آموزش یک طبقه‌بند از ویژگی‌های ویدئوی

سیستم را به صورت برخط و تنها به ازای داده آموزشی جدید بروز کرد. از طرفی با توجه به این که در فرایند نهان‌کاوی کور ویدئو تعداد داده‌های مجموعه آموزش در فرایند یادگیری بسیار زیاد است، این دسته از الگوریتم‌های طبقه‌بندی دارای هزینه محاسباتی سنگینی می‌باشند [۱۴]. الگوریتم‌های یادگیری بدون ناظر در مقایسه با الگوریتم‌های یادگیری باناظر اگرچه ارزان و سریع هستند، اما به علت قابلیت تفسیر و دقت پایین عملاً در فرایند نهان‌کاوی کور ویدئو کاربرد ندارند و به ندرت مورد استفاده قرار گرفته‌اند [۴].

در سال‌های اخیر تکنیک‌های یادگیری شبه‌ناظر پیشرفت قابل ملاحظه‌ای داشته است. یادگیری شبه‌ناظر مابین یادگیری باناظر و بدون ناظر قرار می‌گیرد و تلاش دارد با استفاده از تعداد کمی داده برچسب‌دار و تعداد زیادی داده بدون برچسب به دقتی نزدیک به یادگیری باناظر برسد. Zhang و همکاران [۱۵] یک روش نهان‌کاوی کور تصویر ثابت با استفاده از الگوریتم طبقه‌بندی Path-Based TSVM ارائه کرده‌اند. الگوریتم طبقه‌بندی Path-Based TSVM یک الگوریتم طبقه‌بندی شبه‌ناظر است که توأمأ از داده‌های برچسب‌دار و بدون برچسب در جهت تعیین مرز تصمیم‌گیری طبقه‌بند استفاده می‌کند. اگرچه استفاده از این روش طبقه‌بندی مشکل محدودیت برچسب‌دار بودن داده‌های آموزش را حل کرده است، ولی دارای مشکلاتی است که استفاده از آن را در فرایند نهان‌کاوی کور ویدئو و حتی نهان‌کاوی کور تصویر ثابت با مشکل روبه‌رو می‌کند. الگوریتم طبقه‌بندی Path-Based TSVM علاوه بر این که یک روش طبقه‌بندی برون‌خطی است و آموزش مجدد آن هزینه‌بر است، دارای تابع هدف محدب نبوده و حداقل کردن آن مشکل است. تا کنون روش‌های مختلفی برای حداقل کردن تابع هدف این طبقه‌بند ارائه شده است [۱۶] و [۱۷] ولی همه آنها به خصوص زمانی که تعداد اعضای مجموعه آموزشی زیاد است دارای پیچیدگی زمانی بسیار بالایی می‌باشند.

در زمینه نهان‌کاوی کور ویدئو تا کنون روشی با رویکرد یادگیری شبه‌ناظر ارائه نشده است. در این پژوهش یک روش جدید نهان‌کاوی کور ویدئو با رویکرد یادگیری شبه‌ناظر معرفی شده که با توجه به رفع محدودیت برچسب‌دار بودن داده‌های آموزشی و دسترسی آسان به داده‌های بدون برچسب، کارایی طبقه‌بند آن برای شناسایی الگوریتم‌های نهان‌نگاری‌ای که کد آنها در دسترس نیست بهبود یافته است. همچنین الگوریتم یادگیری شبه‌ناظر ارائه‌شده نسبت به الگوریتم‌های طبقه‌بندی



شکل ۲: مراحل استخراج ویژگی.

حد بالایی فاصله جابه‌جایی است

$$f^1(k) = \frac{|\{C_i | \delta mv^{C_i} = k\}|}{n}, \quad k = 1, 2, \dots, u \quad (3)$$

ویژگی نوع دوم: ویژگی نوع دوم، $f^2(k)$ ، نسبت δe^{C_i} را به ازای یک فاصله جابه‌جایی مشخص در کل n ماکروبلوک محاسبه می‌کند

$$f^2(k) = \frac{1}{\alpha} \sum_{\delta mv^{C_i} = k} \delta e^{C_i} \quad (4)$$

$$\alpha = \sum_{i=1}^n \delta e^{C_i}$$

ویژگی نوع سوم: ویژگی نوع سوم، $f^3(k)$ ، مشتقی از ویژگی نوع دوم است. همان‌طور که در (۵) نشان داده شده است ویژگی نوع سوم در واقع همان ویژگی نوع دوم است که در آن تغییرات فاصله جابه‌جایی نیز در نظر گرفته شده است

$$f^3(k) = \frac{1}{\beta} \sum_{\delta mv^{C_i} = k} (\|mv^{C_i} - \hat{mv}^{C_i}\| + 1) \times \delta e^{C_i} \quad (5)$$

$$\beta = \sum_{i=1}^n (\|mv^{C_i} - \hat{mv}^{C_i}\| + 1) \times \delta e^{C_i}$$

$$\|mv^{C_i} - \hat{mv}^{C_i}\| = \sqrt{(h^{C_i} - \hat{h}^{C_i})^2 + (v^{C_i} - \hat{v}^{C_i})^2}$$

با استفاده از سه نوع ویژگی تعریف‌شده در بالا، برای فریم‌های در نظر گرفته شده یک بردار ویژگی ۱۵ بعدی استخراج می‌شود. ۵ بُعد اول مربوط به ویژگی نوع ۱ بوده و ۵ بُعد دوم و سوم به ترتیب مربوط به ویژگی نوع ۲ و ۳ می‌باشند و به روشی مشابه (۶) تولید می‌شوند

$$f_k = \begin{cases} f^1(k) & , k = 1, 2, 3, 4 \\ \sum_{i=1}^n f^i(i) & , k = 5 \end{cases} \quad (6)$$

در این بخش حد بالای فاصله جابه‌جایی (u) برابر ۴ در نظر گرفته شد زیرا انتخاب مقادیر بزرگ‌تر از ۴ ضمن این که تأثیر چندانی در تمایز بین ویژگی‌های استخراجی از ویدئوی پاک و آلوده ندارد، باعث افزایش ابعاد بردار ویژگی و مدت زمان لازم برای آموزش طبقه‌بند می‌شود. تا این مرحله یک بردار ویژگی ۱۵ بعدی برای یک تعداد مشخص از فریم‌های دنباله ویدئوی ورودی تولید می‌شود. حال با استفاده از این بردار ویژگی، فریم‌های ورودی در یکی از دو دسته پاک یا آلوده طبقه‌بندی می‌شوند که در قسمت بعد درباره آن توضیح داده می‌شود.

پاک استفاده می‌شود. برای این منظور ابتدا با ابزاری تحت عنوان کالیبراسیون، ویدئوی پاک از روی ویدئوی مشکوک ورودی تخمین زده می‌شود.

روش‌های نهان‌نگاری ویدئوی مبتنی بر بردار حرکت در برابر فشرده‌سازی مجدد مقاوم نبوده و اطلاعات سری پنهان‌شده در بردارهای حرکت بعد از فشرده‌سازی مجدد از بین می‌روند [۸]. در نهان‌کاوی کور ویدئوی پیشنهادی از فشرده‌سازی مجدد به عنوان ابزاری برای پاک‌سازی دنباله ویدئوی فشرده‌شده استفاده می‌شود که شامل دو گام است. در گام اول، دنباله ویدئو فشرده‌شده با توجه به نوع کدگذار (۴-MPEG، H.264 و ...) و فشرده‌سازی می‌شود. سپس در گام دوم، دنباله ویدئوی خام با تنظیمات مشابه با دنباله ویدئوی فشرده‌شده اولیه، مجدداً فشرده‌سازی می‌شود.

۲-۲ محاسبه بردار حرکت و استخراج ویژگی

برای استخراج ویژگی با توجه به اختلاف بردارهای حرکت قبل و بعد از پاک‌سازی دنباله ویدئوی فشرده‌شده، یک سری ویژگی استخراج می‌شود. در شکل ۲ مراحل استخراج ویژگی نشان داده شده است. ابتدا دو عملگر δmv^{C_i} و δe^{C_i} برای هر ماکروبلوک C_i در فریم جاری با استفاده از (۱) و (۲) تعریف می‌شوند. در (۱) δmv^{C_i} فاصله جابه‌جایی نامیده می‌شود که در محاسبه آن $mv^{C_i} = (h^{C_i}, v^{C_i})$ و $\hat{mv}^{C_i} = (\hat{h}^{C_i}, \hat{v}^{C_i})$ به ترتیب نشان‌دهنده مؤلفه‌های افقی و عمودی بردار حرکت برای ماکروبلوک C_i قبل و بعد از پاک‌سازی دنباله ویدئوی فشرده‌شده است. همچنین در (۲) δe^{C_i} فاصله خطای پیشگویی نامیده می‌شود که در آن e^{C_i} و \hat{e}^{C_i} به ترتیب مقدار خطای پیشگویی برای ماکروبلوک C_i قبل و بعد از پاک‌سازی دنباله ویدئوی فشرده‌شده هستند و n تعداد کل ماکروبلوک‌ها است

$$\delta mv^{C_i} = |\hat{h}^{C_i} - h^{C_i}| + |\hat{v}^{C_i} - v^{C_i}|, \quad i = 1, 2, \dots, n \quad (1)$$

$$\delta e^{C_i} = |\hat{e}^{C_i} - e^{C_i}|, \quad i = 1, 2, \dots, n \quad (2)$$

برای استخراج ویژگی ابتدا توسط یک پنجره ثابت (به طول یک GOP یا GOV)، تعداد مشخصی از فریم‌ها شامل n ماکروبلوک در نظر گرفته می‌شود و با توجه به دو عملگر δmv^{C_i} و δe^{C_i} سه نوع ویژگی تولید می‌شود [۸].

ویژگی نوع اول: ویژگی نوع اول، $f^1(k)$ ، احتمال یک فاصله جابه‌جایی مشخص را برای یک ماکروبلوک محاسبه می‌کند. در (۳) u

همان طور که در شکل ۴ نشان داده شده است در روش پیشنهادی به منظور آموزش شبکه خودسازمان‌ده پویای شبه‌ناظر در هر تکرار الگوریتم تنها یک داده آموزشی شامل یک بردار ویژگی با ۱۵ بُعد به همراه اطلاعات مربوط به برچسب آن وارد شبکه می‌شود. داده‌های آموزشی از نظر برچسب به دو گروه تقسیم می‌شوند. یک گروه داده‌های آموزش برچسب‌دار هستند که خود به دو دسته آلوده و پاک تقسیم می‌شوند و گروه دیگر، داده‌های بدون برچسب می‌باشند که نمی‌توان برچسب آنها را تعیین کرد و در اصطلاح بیان می‌شود که برچسب آنها گم شده است.

در ابتدای فرایند آموزش، لایه GSOM در شبکه خودسازمان‌ده پویای شبه‌ناظر (SSGSOM) به صورت یک شبکه مربع‌شکل با چهار گره تعریف می‌گردد. به هر کدام از این گره‌ها یک بردار وزن به ابعاد داده آموزشی و به صورت کاملاً تصادفی اختصاص داده می‌شود. در ادامه روند یادگیری لایه GSOM بایستی این توانایی را داشته باشد بر اساس چگونگی توزیع داده‌ها و به صورت خودکار، در کنار هر کدام از گره‌های لایه GSOM، گره جدیدی ایجاد کند. این امر باعث افزایش دانش موجود در لایه GSOM می‌شود. این دانش ممکن است مربوط به الگوریتم‌های پنهان‌نگاری‌ای باشد که کد آنها در دسترس نیست و امکان شبیه‌سازی آن برای تولید داده‌های برچسب‌دار وجود ندارد. به همین منظور در این شبکه با تعیین پارامتری به نام ضریب گسترش (SF)، میزان حساسیت شبکه به پراکندگی داده‌ها مشخص می‌شود. ضریب گسترش مقداری بین صفر و یک دارد و هرچه این پارامتر بزرگ‌تر باشد شبکه افزایش اندازه بیشتری خواهد داشت. با توجه به مقدار ضریب گسترش سطح آستانه GT به صورت زیر تعریف می‌شود

$$GT = -M \times \ln(SF) \quad (7)$$

که در آن M ابعاد یا تعداد ویژگی‌های داده ورودی می‌باشد. GT ماکسیمم خطای ممکن برای یک گره را مشخص می‌کند و هرگاه خطای یک گره از این سطح آستانه بیشتر شود، آن گره وارد فاز رشد می‌شود. با توجه به این که فرایند پنهان‌نگاری کور ویدیو به صورت یک مسئله طبقه‌بندی دوکلاسه مدل می‌شود، برای طراحی طبقه‌بند در پنهان‌کاو پیشنهادی، تعداد گره‌های لایه خروجی برای شبکه خودسازمان‌ده پویای شبه‌ناظر برابر ۲ در نظر گرفته می‌شود. با توجه به توضیحات بالا، عملیاتی که در هر تکرار از الگوریتم شبکه خودسازمان‌ده پویای شبه‌ناظر (SSGSOM) انجام می‌شود به شرح زیر است:

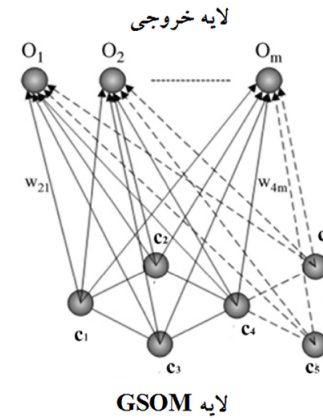
در گام اول از الگوریتم طبقه‌بندی SSGSOM برای به دست آوردن گره برنده (BMU)، فاصله اقلیدسی بین ورودی با تمام گره‌های موجود در لایه GSOM (که تعداد آنها برابر N_{units} است) با استفاده از (۸) محاسبه و ذخیره می‌شود. در این رابطه $x(t)$ و $w_k(t)$ به ترتیب نشان‌دهنده ورودی و وزن گره k ام در لحظه t می‌باشند و $d_k(t)$ برابر با فاصله اقلیدسی بین ورودی و گره k ام است

$$d_k(t) = \|x(t) - w_k(t)\|, \quad k = 1, 2, \dots, N_{units} \quad (8)$$

در گام دوم، گره برنده (BMU) که در واقع شبیه‌ترین گره موجود در شبکه به ورودی می‌باشد با استفاده از (۹) تعیین می‌شود

$$BMU(t) = \min_k(d_k(t)), \quad k = 1, 2, \dots, N_{units} \quad (9)$$

در گام سوم وزن گره برنده و گره‌های همسایه آن با استفاده از (۱۰) به روز می‌شود. در این رابطه $h_k(t)$ تابع همسایگی است که مطابق (۱۱) تعریف می‌شود و با گذشت زمان کاهش می‌یابد



شکل ۳: شبکه خودسازمان‌ده پویای شبه‌ناظر [۱۸].

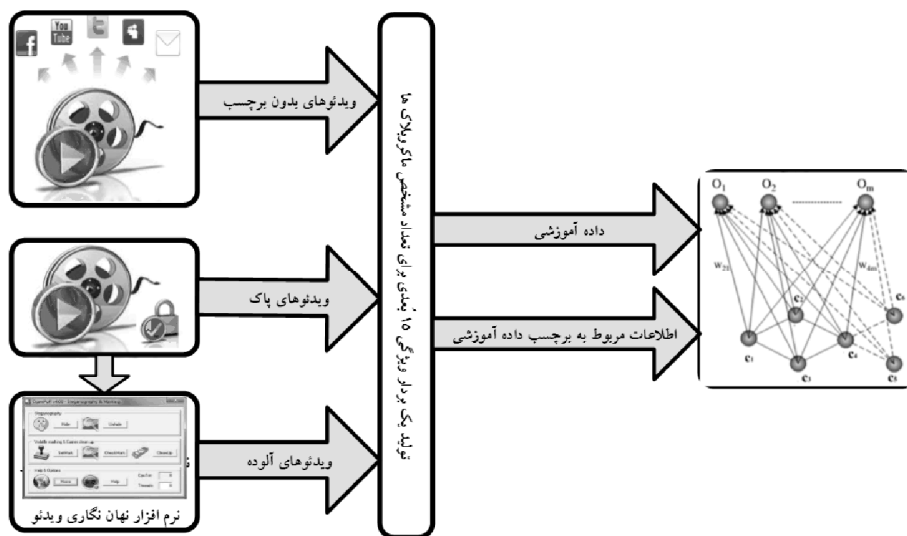
۳-۲ طبقه‌بندی

آخرین مرحله از روش پیشنهادی طراحی طبقه‌بند است که هدف آن قراردادن ویدئوی مشکوک در یکی از دو دسته پاک یا آلوده می‌باشد. در روش پیشنهادی برای اولین بار در پنهان‌کاو کور ویدیو از شبکه خودسازمان‌ده پویای شبه‌ناظر^۱ در طراحی طبقه‌بند استفاده شده است [۱۸]. همان طور که در شکل ۳ نشان آمده است شبکه خودسازمان‌ده پویای شبه‌ناظر دارای دو لایه است، در این شبکه یک لایه یادگیری شبه‌ناظر (لایه خروجی) به بالای یک شبکه خودسازمان‌ده پویا [۱۹] (لایه GSOM) اضافه شده است. در این شبکه هر گره c در لایه GSOM دارای یک تابع گوسی است که با وزن‌های $w_{c1}, w_{c2}, \dots, w_{cm}$ به m گره خروجی که هر یک نشان‌دهنده یک کلاس می‌باشند، متصل شده است. محدوده فعالیت یا واریانس هر یک از این توابع گوسی به صورت میانگین فاصله‌های هر گره با کلیه گره‌هایی که در همسایگی مستقیم آن قرار دارند، تعریف می‌شود. 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 صورت می‌پذیرد. شبکه خودسازمان‌ده پویای شبه‌ناظر یک الگوریتم یادگیری ذاتاً برخط است و می‌تواند داده آموزشی جدید را بدون نیاز به استفاده از داده‌های آموزشی قبلی به سیستم اعمال کرده و سیستم را به روز کند. همچنین می‌تواند به طور هم‌زمان با هر سه نوع روش یادگیری باناظر، شبه‌ناظر و بدون ناظر عمل کند و تمام اطلاعات موجود در مجموعه آموزشی را در جهت بهبود عملکرد سیستم مورد استفاده قرار دهد.

۳- آموزش شبکه خودسازمان‌ده پویای شبه‌ناظر در فرایند پنهان‌کاو کور ویدیو

برای تولید مجموعه آموزش در روش پیشنهادی از سه دسته ویدیو استفاده شده است. دسته اول ویدئوهای بدون برچسب هستند که با حجم بالا از طریق اینترنت و سایت‌های اشتراک ویدیو قابل دریافت هستند. جمع‌آوری این دسته از ویدئوها هزینه پائینی دارد که شامل ویدئوهای آلوده و پاک است. دسته دوم ویدئوهای پاک هستند که در این دسته از ویدئوها هیچ گونه پیام سری پنهان نشده و دسته سوم ویدئوهای آلوده هستند که با استفاده از ویدئوهای پاک و تعدادی محدودی از الگوریتم‌های پنهان‌نگاری ویدیو که کد آنها در دسترس است تولید می‌شوند.

1. Semi-Supervised Growing Self-Organizing Maps



شکل ۴: مراحل آموزش شبکه خودسازمان‌ده پویای شبه‌ناظر.

$$PS_m = \sum_{C=1}^{N_{units}} w_{cm} O_C \quad (15)$$

در گام پنجم وزن اتصالات بین لایه GSOM و لایه خروجی (w_{cm}) با استفاده از (۱۶) به تکرار λ بار آموزش داده می‌شود. در این رابطه λ یک عدد صحیح است و نرخ یادگیری η عددی بین صفر و یک می‌باشد. همچنین ξ_m با استفاده از (۱۷) تعریف می‌شود. همان‌طور که مشاهده می‌شود برای آموزش وزن اتصالات بین لایه‌ای و انتخاب وزن‌های بهینه از برچسب داده‌ها استفاده می‌شود

$$w_{cm}(t+1) = w_{cm}(t) + \eta(\xi_m - PS_m)O_C \quad (16)$$

$$\xi_m = \begin{cases} 1 & \text{has lable and correct prediction} \\ 0 & \text{has label and wrong prediction or miss label} \end{cases} \quad (17)$$

در گام ششم بایستی خطای گره برنده با استفاده از (۱۸) به روز شود. در این رابطه $E_{BMU}(t)$ نشان‌دهنده خطای گره برنده و ΔE_{BMU} مقدار تغییر خطای گره برنده در تکرار t ام است که با توجه به (۱۹) تعریف می‌شود. در (۱۹) M ابعاد یا تعداد ویژگی‌های داده ورودی و $w_{BMU}(t)$ وزن گره برنده در تکرار t ام است

$$E_{BMU}(t+1) = E_{BMU}(t) + \Delta E_{BMU} \quad (18)$$

$$\Delta E_{BMU} = \begin{cases} 0 & \text{case 1: has label and correct prediction} \\ \frac{\sqrt{M}}{2} & \text{case 2: has label and} \\ & \text{wrong prediction} \\ \|x(t) - w_{BMU}(t)\| & \text{case 3: miss label} \end{cases} \quad (19)$$

در گام هفتم از شبکه خودسازمان‌ده پویای شبه‌ناظر، لایه GSOM وارد فاز رشد می‌شود. در این مرحله اگر خطای جمع‌شونده گره برنده (رابطه (۱۸)) از سطح آستانه GT بزرگ‌تر شود فاز رشد برای آن گره اعمال می‌شود که دو حالت مختلف دارد. حالت اول آن است که گره برنده یک گره مرزی نباشد و در این صورت با توجه به (۲۰) خطای گره برنده بر روی گره‌هایی که در همسایگی آن قرار دارند توزیع می‌شود. در این رابطه $E_{neighbours}(t)$ نشان‌دهنده خطای گره همسایه در تکرار t ام است و $N_{neighbours}$ برابر با تعداد همسایه‌های گره برنده می‌باشد

$$E_{neighbours}(t+1) = E_{neighbours}(t) + \frac{1}{N_{neighbours}} \times \frac{E_{BMU}(t)}{2} \quad (20)$$

$$w_k(t+1) = w_k(t) + h_k(t)(x(t) - w_k(t)) \quad (10)$$

$k = 1, 2, \dots, N_{units}$

$$h_k(t) = \alpha(t) \left[\exp \frac{-\|x(t) - w_k(t)\|^2}{\sigma_k^2} \right] \quad (11)$$

که در (۱۱) $\alpha(t)$ نرخ یادگیری است. این نرخ در طول فرایند یادگیری کاهش می‌یابد تا وزن‌ها همگرا شوند. برای آن که این رفتار شبیه‌سازی شود مدل‌های متعددی وجود دارد که یکی از مدل‌های پرستفاده تابع تضعیف‌نمایی است که به صورت (۱۲) بیان می‌شود

$$\alpha(t) = \alpha \times e^{-\frac{t}{\gamma}} \quad (12)$$

که در آن α مقدار ابتدایی نرخ یادگیری بوده و عددی بین صفر و یک است. همچنین γ پارامتری است که مشخص می‌کند نرخ یادگیری با چه سرعتی افت نماید.

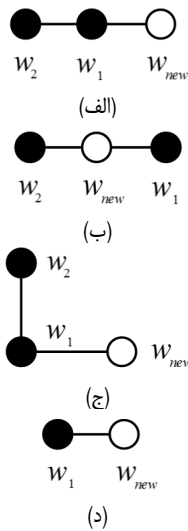
در گام چهارم بایستی برچسب داده ورودی $x(t)$ تعیین شود. برای این منظور با استفاده از (۱۵) میزان امتیاز هر گره در لایه خروجی شبکه SSGSOM محاسبه می‌شود. برای به دست آوردن میزان امتیاز هر گره خروجی (PS_m) ابتدا باید سطح فعالیت هر گره (O_C) طبق رابطه زیر در لایه GSOM محاسبه و ذخیره شود

$$O_C = \exp \frac{-\|x(t) - w_C(t)\|^2}{\sigma_C^2} \quad (13)$$

که در آن، $w_C(t)$ بردار وزن و σ_C محدوده فعالیت گره شماره C است که با استفاده از (۱۴) تعریف می‌شود

$$\sigma_C = \frac{\sum_{k \in N_{NC}} \|w_k(t) - w_C(t)\|^2}{\text{Count}(N_{NC})} \quad (14)$$

که در آن، N_{NC} برابر با شماره گره همسایه‌های مستقیم گره شماره C می‌باشد. در این صورت میزان امتیاز هر گره در لایه خروجی شبکه خودسازمان‌ده پویای شبه‌ناظر با استفاده از رابطه زیر محاسبه می‌شود. بعد از آن که میزان امتیاز تمام گره‌ها در لایه خروجی محاسبه شد برچسب داده ورودی برابر با آن دسته یا گره خروجی است که میزان امتیاز آن از گره دیگر بیشتر باشد



شکل ۶: وزن دهی اولیه گره‌های جدید.

$$w_{new} = \frac{r_1 + r_2}{2} \quad (24)$$

۴- پیچیدگی زمانی

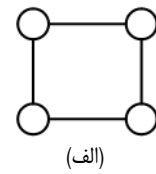
درباره کاربرد الگوریتم‌های طبقه‌بندی Path-Based TSVM و SVM در فرایند نهان‌کاوی کور در مقدمه توضیح داده شد. همان طور که گفته شد این الگوریتم‌های طبقه‌بندی در محیط نهان‌کاوی کور ویدئو (با توجه به تعداد زیاد داده‌های آموزشی) دارای پیچیدگی زمانی بالایی هستند و کاربرد آنها در دنیای واقعی مشکل است. مزیت دیگر شبکه خودسازمان‌ده پویای شبه‌ناظر نسبت به دو الگوریتم Path-Based TSVM و SVM این است که پیچیدگی زمانی آن در محیط نهان‌کاوی کور ویدئو پایین است. برای اثبات سریع‌تر بودن شبکه خودسازمان‌ده پویای شبه‌ناظر ابتدا پیچیدگی زمانی روش طبقه‌بندی SSGSOM محاسبه می‌شود.

اصل ۱: اگر M تعداد ویژگی‌های هر داده آموزشی، N تعداد اعضای مجموعه آموزشی و $N_{units}(t)$ برابر با تعداد گره‌های لایه GSOM در هر تکرار t باشد، در این صورت پیچیدگی زمانی شبکه خودسازمان‌ده پویای شبه‌ناظر برابر با $O(\sum_{t=1}^N N_{units}(t) \times (1 \cdot M + \lambda) + 4MN)$ است (برای اثبات به پیوست مراجعه شود).

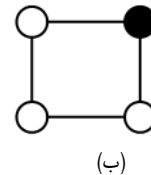
با توجه به پیچیدگی زمانی به دست آمده برای شبکه خودسازمان‌ده پویای شبه‌ناظر (SSGSOM) در اصل ۱، بایستی ثابت شود شبکه SSGSOM در محیط نهان‌کاوی کور ویدئو نسبت به دو روش طبقه‌بندی SVM و Path-Based TSVM سریع‌تر است. برای این منظور اصل ۲ بیان می‌شود.

اصل ۲: اگر تعداد اعضای مجموعه آموزشی از تعداد ویژگی‌های استخراج‌شده خیلی بزرگ‌تر باشد، در این صورت پیچیدگی زمانی شبکه خودسازمان‌ده پویای شبه‌ناظر از دو روش طبقه‌بندی Path-Based TSVM و SVM کمتر است (برای اثبات به پیوست مراجعه شود).

با توجه به شباهت بالا بین ویدئوهای آلوده و پاک برای آن که نهان‌کاوی دارای دقت شناسایی و تعمیم‌پذیری مناسب باشد معمولاً در مرحله استخراج ویژگی تعداد قابل توجهی ویژگی استخراج می‌شود. از طرف دیگر تعداد زیاد ویژگی‌ها نیازمند تعداد قابل ملاحظه‌ای داده آموزشی برای طراحی طبقه‌بند می‌باشد. در واقع با افزایش تعداد ویژگی‌ها پدیده‌ای با عنوان Curse Dimensionality مطرح می‌شود. طبق این پدیده در محیطی که تعداد ویژگی‌ها زیاد است تعداد نمونه‌های مجموعه

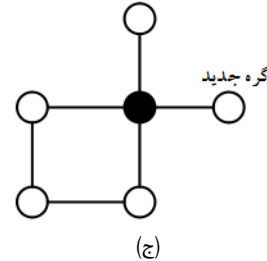


خطای گره زیاد (الف)



(ب)

گره جدید



(ج)

شکل ۵: تولید گره جدید از یک گره مرزی شبکه.

حالت دوم آن است که گره برنده یک گره مرزی باشد. در این صورت گره‌های جدید به شبکه اضافه شده و با استفاده از گره‌هایی که در همسایگی مستقیم آنها قرار دارند وزن‌دهی می‌شوند (شکل ۵). در مدل ارائه‌شده برای شبکه خودسازمان‌ده پویای شبه‌ناظر، هر گره دارای حداکثر چهار همسایه مستقیم است و یک گره مرزی است اگر حداقل یکی از گره‌هایی که در همسایگی مستقیم آن قرار دارد آزاد باشد. بنابراین در فاز رشد برای یک گره حداکثر ۳ گره جدید تولید می‌شود. گره جدید بعد از وزن‌دهی به تمام گره‌های لایه خروجی متصل می‌شود (مانند گره‌های c_5 و c_6 در شکل ۳). در شکل ۶ تمام حالت‌های ممکن بعد از اضافه شدن گره جدید نشان داده شده است. برای هر یک از حالت‌های نشان داده شده در شکل ۶ مقدار وزن گره جدید به صورت زیر محاسبه می‌شود. توجه شود در اینجا w_i نشان‌دهنده وزن گره i ام است.

حالت الف: گره جدید در یک طرف دو گره قدیمی قرار گرفته است. در این صورت وزن گره جدید به صورت زیر محاسبه می‌شود

$$w_{new} = \begin{cases} w_1 - (w_r - w_1) & \text{if } w_r > w_1 \\ w_1 + (w_r - w_1) & \text{if } w_1 > w_r \end{cases} \quad (21)$$

حالت ب: گره جدید بین دو گره قدیمی قرار گرفته است. در این صورت وزن گره جدید به صورت زیر محاسبه می‌شود

$$w_{new} = \frac{w_1 + w_r}{2} \quad (22)$$

حالت ج: گره جدید دارای یک گره در همسایگی مستقیم و یک گره در همسایگی غیر مستقیم خود است. در این صورت وزن گره جدید به صورت زیر محاسبه می‌شود

$$w_{new} = \begin{cases} w_1 - (w_r - w_1) & \text{if } w_r > w_1 \\ w_1 + (w_r - w_1) & \text{if } w_1 > w_r \end{cases} \quad (23)$$

حالت د: گره جدید فقط دارای یک گره در همسایگی مستقیم خود است. در این صورت وزن گره جدید با استفاده از (۲۴) محاسبه می‌شود. در این رابطه r_1 و r_2 به ترتیب مقدار حداکثر و حداقل بردار وزن مربوط به گره‌های موجود در شبکه می‌باشند

TN (منفی درست) و FN (منفی اشتباه) دارای تعاریفی مشابه برای ویدئوی پاک هستند.

۵-۱ شرایط شبیه‌سازی

یک پایگاه داده شامل ۱۹ دنباله ویدئو با نوع نمایش YUV و فرمت نمونه‌برداری ۴:۲:۰ برای آزمایشات مورد استفاده قرار گرفت. در جدول ۲ مشخصات دنباله‌های ویدئوی استاندارد آورده شده است. همان‌طور که از جدول ۲ 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 هستند، همه دنباله‌های ویدئو بدون هم‌پوشانی به دنباله‌های ۵۰ فریمی تقسیم می‌شوند و در مجموع تعداد کل دنباله‌ها به بیش از ۱۱۹ عدد می‌رسد. دنباله‌های ویدئو از سایت <http://media.xiph.org/video/derf/> گرفته شده است.

برای پنهان‌سازی اطلاعات سری در درون دنباله‌های ویدئو از سه روش نهان‌نگاری مبتنی بر بردارهای حرکت [۲۰] تا [۲۲] استفاده گردید. در این مرحله بیت‌های پیام سری در داخل بردارهای حرکت جاسازی می‌شوند که برای پیاده‌سازی از استاندارد مشهور ۲-MPEG استفاده شد. همه دنباله‌های ویدئویی توسط $TM5$ (MPEG۲ Test Model ۵) [۲۳] با تنظیمات استاندارد فشرده شدند. اندازه هر بلاک 8×8 و اندازه هر ماکروبلاک نیز 16×16 می‌باشد. هر GOP شامل یک فریم نوع I، سه فریم نوع P و هشت فریم نوع B می‌باشد که به صورت پررودیک تکرار می‌شود (IBBPBBPBBPBB). توجه شود قدرت جاسازی به صورت متوسط بیت‌های جاسازی شده در یک فریم (bpf) در نظر گرفته شده است. پنهان‌سازی اطلاعات سری در دنباله‌های ویدئو برای تعداد بیت‌های متفاوت شامل ۱۰، ۲۰، ۳۰، ۴۰ و ۵۰ بیت در یک فریم (bpf) تکرار شده است.

تعداد ۹۶ عدد (حدود ۸۰ درصد) دنباله ویدئو به صورت کاملاً تصادفی از بین ۱۱۹ دنباله ویدئوی موجود در پایگاه داده برای مرحله آموزش طبقه‌بند انتخاب می‌شوند. از ۲۳ عدد دنباله ویدئوی باقی‌مانده در پایگاه داده برای تولید مجموعه آزمایش استفاده خواهد شد. برای تولید مجموعه آموزش ابتدا تمام ۹۶ دنباله ویدئو توسط $TM5$ با تنظیمات استاندارد فشرده‌سازی شده و عمل استخراج ویژگی برای تولید نمونه‌های پاک انجام می‌شود. سپس برای تولید نمونه‌های آلوده دوباره ۹۶ دنباله ویدئو به نسبت مساوی توسط سه روش نهان‌نگاری مبتنی بر بردارهای حرکت [۲۰] تا [۲۲] فشرده‌سازی می‌شوند. بنابراین مجموعه آموزش دارای ۱۹۲ دنباله ویدئو است که از همین مجموعه برای تولید داده‌های بدون برچسب استفاده شده است.

۵-۲ دقت نهان‌کاوی پیشنهادی

همان‌طور که توضیح داده شد یکی از اهداف پژوهش، حل مسئله نهان‌کاوی کور ویدئو با رویکرد یادگیری شبه‌ناظر است. در یادگیری شبه‌ناظر تعداد داده‌های برچسب‌دار بر روی دقت سیستم تأثیر مستقیم دارد. بنابراین در گام اول، تأثیر تعداد داده‌های برچسب‌دار در مجموعه آموزش بر روی دقت شناسایی نهان‌کاوی پیشنهادی بررسی می‌شود. برای این منظور تعداد داده‌های برچسب‌دار در مجموعه آموزش از ۱۰ تا ۹۰ درصد تعداد کل اعضای مجموعه آموزش تغییر داده می‌شود. سپس مجموعه آموزش و آزمایش به سیستم اعمال شده و دقت سیستم محاسبه می‌شود. در اینجا این روند ۱۰ بار تکرار می‌شود و میانگین دقت محاسبه‌شده در ۱۰ بار تکرار به عنوان خروجی ثبت می‌گردد. دو روش

جدول ۱: مقایسه بین تعداد ویژگی‌های استخراجی و تعداد اعضای مجموعه آموزشی در فرایند نهان‌کاوی کور ویدئو.

تعداد ویژگی‌های استخراج شده	تعداد اعضای مجموعه آموزشی	مرجع
۱۴	۲۳۵۲	[۵]
۲۷۴	۲۴۰۰۰	[۶]
۲	۱۳۶۴	[۷]
۱۲	۸۸۰	[۱۰]
۳۹	۳۰۰۰	[۱۱]
۳	۱۱۲۰	[۱۲]
۳	۱۶۸	[۱۳]

آموزشی به صورت نمایی با افزایش تعداد ویژگی‌ها رشد می‌کند [۱۴]. بنابراین اگرچه در فرایند نهان‌کاوی کور ویدئو تعداد ویژگی‌های استخراجی زیاد می‌باشد ولی برای طراحی یک طبقه‌بند قابل اطمینان باید تعداد اعضای مجموعه آموزشی خیلی بزرگ‌تر از تعداد ویژگی‌ها در نظر گرفته شود. در جدول ۱ تعداد ویژگی‌های استخراجی در مقایسه با تعداد اعضای مجموعه آموزشی برای پژوهش‌هایی که تا این زمان در زمینه نهان‌کاوی کور ویدئو انجام شده‌اند نشان داده شده است. همان‌طور که در جدول ۱ مشاهده می‌شود تعداد اعضای مجموعه آموزشی حداقل ۵۰ برابر تعداد ویژگی‌ها می‌باشد که ادعای فوق را تصدیق می‌کند. بنابراین با توجه به اصل ۲ و تعداد بسیار زیاد داده‌های آموزشی نسبت به تعداد ویژگی‌های استخراجی می‌توان گفت روش طبقه‌بندی SSGSOM از دو روش طبقه‌بندی SVM و Path-Based TSVM در محیط نهان‌کاوی کور ویدئو سریع‌تر است.

۵- نتایج شبیه‌سازی

در این قسمت تلاش می‌شود از دو جنبه مختلف عملکرد الگوریتم پیشنهادی، شبیه‌سازی و بررسی شود. ابتدا دقت نهان‌کاوی سیستم پیشنهادی نسبت به تغییر تعداد بیت جاسازی شده در هر فریم، تغییر تعداد داده‌های برچسب‌دار و شناسایی الگوریتم‌های نهان‌نگاری با کد غیر قابل دسترس محاسبه می‌شود. در مرحله دوم سرعت سیستم با هدف بررسی میزان محاسبات لازم برای مرحله آموزش طبقه‌بند محاسبه خواهد شد. در تمام آزمایش‌ها مقدار پارامتر ضریب گسترش برابر 0.35 ، مقدار اولیه نرخ یادگیری برابر یک، نرخ یادگیری برای آموزش وزن‌های بین لایه‌ای برابر 0.1 و مقدار پارامتر λ برابر ۳ در نظر گرفته شده است (مقادیر مذکور با هدف رسیدن به حداکثر دقت شناسایی و با انجام آزمایشات مختلف تعیین شده‌اند).

در فرایند نهان‌کاوی کور ویدئو یک سری معیار برای بررسی دقت شناسایی تعریف می‌شود. یکی از معروف‌ترین معیارهای بررسی دقت شناسایی در فرایند نهان‌کاوی کور ویدئو، دقت^۱ است که به صورت زیر تعریف می‌شود [۱۵]

$$A = \frac{TP + TN}{TP + FN + TN + FP} \quad (25)$$

در (۲۵)، TP (مثبت درست) حالتی است که در آن طبقه‌بندی کننده به درستی ویدئوی مشکوک را یک ویدئوی آلوده شناسایی کرده است در حالی که FP (مثبت اشتباه) به معنای آن است که طبقه‌بندی کننده به اشتباه ویدئوی مشکوک را یک ویدئوی آلوده شناسایی کند. همچنین

1. Accuracy

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: دقت شناسایی روش پیشنهادی در مقایسه با [۸].

الگوریتم نهان‌نگاری		[۲۰]	[۲۱]	[۲۲]
		TP	TN	TP
مرجع [۸]		۰٫۸۳۱	۰٫۸۱۳۲	۰٫۷۳۲۵
روش پیشنهادی		۰٫۸۳۹	۰٫۷۹۶۰	۰٫۷۳۴۲

سیستم اعمال شده و دقت شناسایی سیستم بر حسب معیارهای TP و TN در جدول گزارش شده است. در این آزمایش نیز متوسط تعداد بیت‌های جاسازی شده در هر فریم ۵۰ عدد است. همان‌طور که مشاهده می‌شود دقت روش پیشنهادی با روش ارائه‌شده در [۸] قابل مقایسه می‌باشد در حالی که هزینه تولید مجموعه آموزش در روش پیشنهادی به مراتب از روش ارائه‌شده در [۸] کمتر است.

شبکه خودسازمان‌ده پویای شبه‌ناظر باعث بهبود عملکرد نهان‌کاو نسبت به الگوریتم‌های یادگیری باناظر مانند SVM، برای شناسایی الگوریتم‌های نهان‌نگاری‌ای که کد آنها در دسترس نیست، می‌شود. برای بررسی درستی این مطلب، یک مجموعه آموزشی جدید تولید می‌شود. در مجموعه آموزشی جدید، تنها داده‌های آموزشی تولیدشده توسط دو الگوریتم از سه الگوریتم نهان‌نگاری شبیه‌سازی شده ([۲۰] تا [۲۲]) برچسب زده می‌شوند و داده‌های آموزشی مربوط به الگوریتم سوم به صورت بدون برچسب به مجموعه آموزش اضافه شده و فرض می‌شود کد الگوریتم نهان‌نگاری سوم در دسترس نیست. سپس دقت سیستم پیشنهادی برای شناسایی الگوریتم نهان‌نگاری با کد غیر قابل دسترس در مقایسه با [۸] بررسی می‌شود. در [۸] از الگوریتم یادگیری باناظر SVM به عنوان طبقه‌بند استفاده شده است. در هنگام آموزش نهان‌کاو، در روش پیشنهادی از کل مجموعه آموزش در طراحی طبقه‌بند استفاده می‌شود در حالی که برای آموزش SVM در [۸] بایستی داده‌های بدون برچسب از مجموعه آموزش کنار گذاشته شوند. نتایج در جدول ۴ نشان داده شده است و همان‌طور که مشاهده می‌شود روش پیشنهادی نسبت به [۸]

طبقه‌بندی SVM و Path-Based TSVM به عنوان مرجع مقایسه استفاده می‌شوند. SVM یک الگوریتم یادگیری باناظر قدرتمند است که به عنوان پرکاربردترین طبقه‌بند در زمینه نهان‌کاو کور ویدئو مطرح است. Path-Based TSVM نیز تنها الگوریتم یادگیری شبه‌ناظر است که در زمینه نهان‌کاو کور تصویر ثابت استفاده شده است. اشاره به این نکته الزامی است که در این آزمایش متوسط تعداد بیت‌های جاسازی شده در هر فریم ۵۰ عدد است. در شکل ۷ دقت شناسایی روش پیشنهادی در مقایسه با دو الگوریتم دیگر رسم شده است. همان‌طور که مشخص است زمانی که تعداد داده‌های برچسب‌دار از ۴۰ درصد تعداد کل داده‌های آموزشی بیشتر است دقت الگوریتم پیشنهادی، قابل مقایسه با دو الگوریتم دیگر است ولی زمانی که تعداد داده‌های برچسب‌دار بسیار کم است دقت سیستم کاهش می‌یابد که این به دلیل ماهیت برخط‌بودن سیستم می‌باشد. چون در الگوریتم‌های برون‌خطی مانند Path-Based TSVM یا SVM تمام داده‌ها برچسب‌دار و به طور هم‌زمان در اختیار هستند در نتیجه مرز تصمیم‌گیری با دقت بالاتری تعیین می‌شود، در حالی که به یک سیستم برخط در هر تکرار تنها یک داده آموزشی وارد می‌شود.

در جدول ۳ دقت شناسایی روش پیشنهادی در مقایسه با روش ارائه‌شده در [۸] برای روش‌های نهان‌نگاری مطرح در [۲۰] تا [۲۲] آورده شده است. در این آزمایش تعداد داده‌های برچسب‌دار در مجموعه آموزش برای روش پیشنهادی برابر با ۵۰ درصد تعداد کل اعضای مجموعه آموزش در نظر گرفته شد در حالی که در روش ارائه‌شده در [۸] تمام داده‌های آموزشی برچسب‌دار می‌باشند. مجموعه آموزش و آزمایش به

جدول ۵: مدت زمان لازم برای آموزش اولیه سیستم نهان کاوی کور ویدئو (بر حسب ثانیه).

T SVM	SVM	SSGSOM	N
۲,۲۳	۰,۹۴	۰,۷۲	۱۰۰
۲۲,۴۶	۲,۱۱	۳,۱۹	۳۰۰
۹۱,۸۵	۱۶,۲۱	۶,۷۶	۵۰۰
۲۷۵,۴۹	۲۷,۵۵	۱۱,۲۳	۷۰۰
۵۱۷,۳۲	۴۱,۹۵	۲۲,۵۴	۹۰۰
۹۶۰,۲	۵۴,۸۵	۳۲,۲۱	۱۱۰
۱۲۲۰,۳	۹۸,۲۳	۴۱,۲۴	۱۳۰۰

جدول ۶: مدت زمان لازم برای آموزش مجدد سیستم نهان کاوی کور ویدئو (بر حسب ثانیه).

T SVM	SVM	SSGSOM	N
۱۴۰۰,۲۳	۱۳۰,۴	۵,۰۸	۴۰۰

مجدد سیستم است. در هر دو حالت مدت زمان لازم برای آموزش سیستم با استفاده از یک دستگاه رایانه با مشخصات زیر محاسبه گردید:

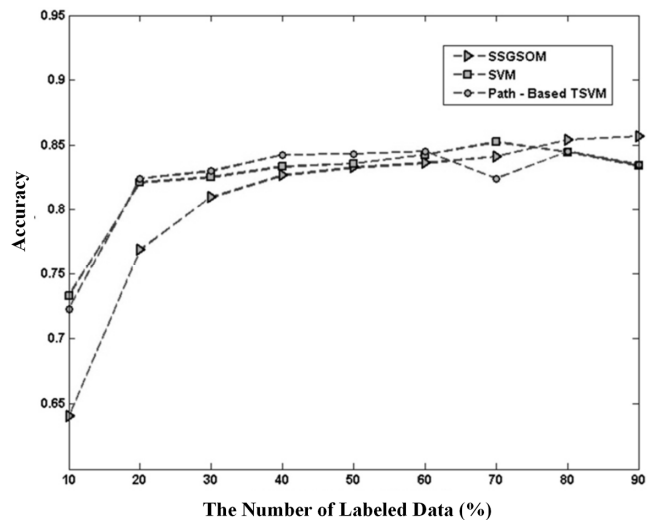
- OS Name : Microsoft Windows v Home Premium (Version ۶.۱.۷۶۰۱ Service Pack ۱ Build ۷۶۰۱)
- System Type : x۶۴ – based PC
- Processor : Intel (R) Core (TM) i۷ CPU Q ۷۲۰ @ ۱,۶۰ GHz, ۱۶۰۰ Mhz, ۴ Core (s), ۸ Logical Processor (s)
- Installed Physical Memory (RAM) : ۶,۰۰ GB

در جدول ۵ مدت زمان لازم برای آموزش اولیه روش پیشنهادی در مقایسه با SVM و Path-Based TSVM نسبت به تغییر تعداد داده‌های آموزشی (N) نشان داده شده است. همان طور که از نتایج جدول مشخص است هنگامی که تعداد اعضای مجموعه آموزشی زیاد می‌شود مدت زمان لازم برای آموزش طبقه‌بند در روش پیشنهادی به طور قابل ملاحظه‌ای از دو روش رقیب کمتر است.

در جدول ۶ مدت زمان لازم برای آموزش مجدد سیستم نشان داده شده است. ابتدا سیستم نهان‌کاوی با یک مجموعه آموزشی نسبتاً بزرگ به تعداد ۱۰۰۰ نمونه آموزش داده می‌شود. حال سیستم باید به ازای یک مجموعه آموزشی جدید به تعداد ۴۰۰ نمونه به روز شود. برای آموزش مجدد سیستم، با توجه به ماهیت برون‌خطی بودن الگوریتم‌های SVM و Path-Based TSVM بایستی ابتدا یک مجموعه آموزشی جدید به تعداد ۱۴۰۰ نمونه تشکیل شود و مرحله آموزش با این مجموعه بزرگ تکرار شود که مسلماً بسیار هزینه‌بر خواهد بود. شبکه خودسازمان‌ده پویای شبه‌ناظر یک الگوریتم طبقه‌بندی ذاتاً برخط و بهینه است (برای اثبات به پیوست مراجعه شود) و به همین دلیل می‌توان مجموعه آموزشی جدید را مستقیماً به سیستم اعمال کرد که این یک مزیت بسیار مناسب در فرایند نهان‌کاوی کور ویدئو محسوب می‌شود.

۶- نتیجه‌گیری

در این مقاله یک روش جدید در زمینه نهان‌کاوی کور ویدئو با رویکرد یادگیری شبه‌ناظر ارائه گردید. در روش پیشنهادی هدف، حمله به الگوریتم‌های نهان‌نگاری ویدئوی مبتنی بر بردارهای حرکت می‌باشد که شامل سه مرحله است. در مرحله اول با استفاده از فشرده‌سازی مجدد، دنباله ویدئوی فشرده‌شده ورودی پاک‌سازی می‌شود، سپس در مرحله دوم با توجه به اختلاف بردارهای حرکت میزبان قبل و بعد از فشرده‌سازی



شکل ۷: دقت شناسایی نهان‌کاوی پیشنهادی بر حسب تعداد داده‌های برچسب‌دار.

جدول ۴: دقت شناسایی برای الگوریتم‌های نهان‌نگاری‌ای که کد آنها در دسترس نیست.

روش پیشنهادی	دقت شناسایی	کد
۰,۸۲۳	۰,۸۵۴	[۱۸]
۰,۷۷۷	۰,۸۱۲	[۱۹]
۰,۶۶۸	۰,۷۲۷	[۲۰]

برای شناسایی الگوریتم نهان‌نگاری با کد غیر قابل دسترس بهتر عمل می‌کند. برای توضیح می‌توان گفت در لایه GSOM داده‌ها به صورت بدون ناظر خوشه‌بندی شده و داده‌های مشابه در یک خوشه قرار داده می‌شوند. از طرفی لایه GSOM در شبکه خودسازمان‌ده پویای شبه‌ناظر می‌تواند توسعه پیدا کند. هر داده بدون برچسب که دارای ویژگی‌هایی متفاوت باشد و تا کنون در آموزش شبکه در نظر گرفته نشده است باعث رشد شبکه می‌شود و به عبارت دیگر باعث گسترش دانش لایه GSOM می‌شود در حالی که هر داده برچسب‌دار باعث اصلاح لایه خروجی در شبکه خودسازمان‌ده پویای شبه‌ناظر می‌شود. پس اگر در فرایند نهان‌کاوی کور ویدئو حجم بالایی از داده‌های بدون برچسب به شبکه خودسازمان‌ده پویای شبه‌ناظر اعمال شود، این شبکه با استفاده از دانش لایه GSOM و اطلاعات جزئی لایه خروجی برای الگوریتم‌های نهان‌نگاری‌ای که کد آنها در دسترس نیست، بهتر عمل می‌کند.

۵-۳ سرعت نهان‌کاوی پیشنهادی

همان طور که در قسمت ۴ بیان شد می‌توان انتظار داشت در محیط نهان‌کاوی کور ویدئو، روش طبقه‌بندی SSGSOM از دو روش طبقه‌بندی معمول در زمینه نهان‌کاوی کور یعنی SVM و Path-Based TSVM سریع‌تر باشد. در این قسمت هدف این است که درستی این مطلب از طریق شبیه‌سازی کامپیوتری هم بررسی شود. در ادامه مدت زمان لازم برای آموزش شبکه خودسازمان‌ده پویای شبه‌ناظر با دو الگوریتم طبقه‌بندی پرکاربرد در زمینه نهان‌کاوی کور یعنی SVM و Path-Based TSVM مقایسه می‌شود. برای این منظور مدت زمان لازم برای آموزش سیستم در دو حالت محاسبه می‌شود. در حالت اول فرض بر این است که سیستم برای اولین بار آموزش داده می‌شود و مدت زمان لازم برای آموزش الگوریتم نسبت به تغییر تعداد داده‌های آموزشی محاسبه می‌شود. ولی در حالت دوم فرض بر این است که سیستم قبلاً آموزش داده شده و حال بعد از دسترسی به یک مجموعه داده آموزشی جدید بایستی به روز شود و به عبارت دیگر در حالت دوم هدف محاسبه مدت زمان لازم برای آموزش

همسایه‌های مستقیم گره شماره C می‌باشد که حداکثر مقدار آن برابر ۴ می‌باشد. بنابراین پیچیدگی زمانی مورد نیاز برای محاسبه سطح فعالیت برای هر گره در بدترین حالت ممکن برابر با $O(4M)$ است. در این صورت پیچیدگی زمانی مورد نیاز برای محاسبه میزان امتیاز هر گره در لایه خروجی برابر با $O(4N_{units}M)$ است. این روند باید برای تمام گره‌ها در لایه خروجی تکرار شود پس پیچیدگی زمانی کل برای مرحله تولید برچسب داده ورودی، در یک مسئله طبقه‌بندی دوکلاسه برابر $O(8N_{units}M)$ است.

در گام پنجم از الگوریتم طبقه‌بندی SSGSOM وزن اتصالات بین لایه‌ای با استفاده از (۱۶) به روز می‌شود. در این مرحله وزن هر اتصال به تکرار λ (برابر ۳) بار آموزش می‌بیند. به دلیل این که PS_m و O_C در گام قبل محاسبه و ذخیره شده‌اند و پیچیدگی زمانی ξ_m از مرتبه $O(1)$ است می‌توان گفت پیچیدگی زمانی این مرحله برابر با $O(N_{units}N\lambda)$ می‌باشد که در اینجا N تعداد گره در لایه خروجی است. با توجه به این که $N\lambda$ یک مقدار ثابت و برابر ۶ است می‌توان گفت پیچیدگی زمانی این مرحله از مرتبه $O(6N_{units})$ است.

در گام ششم بایستی خطای گره برنده با استفاده از (۱۸) به روز شود. پیچیدگی زمانی این مرحله برابر با $O(M)$ می‌باشد. گام هفتم شامل فاز رشد لایه پایینی شبکه (GSOM) می‌باشد. در حالت اول اگر گره برنده یک گره مرزی نباشد در این صورت با توجه به (۲۰) خطای گره برنده بر روی گره‌هایی که در همسایگی آن قرار دارند توزیع می‌شود. با توجه به مشخص بودن مقدار خطای گره می‌توان گفت پیچیدگی زمانی در این حالت برابر با $O(N_{neighbours})$ می‌باشد. در بدترین حالت ممکن خطا بر روی تمام گره‌های موجود در شبکه توزیع می‌شود پس پیچیدگی زمانی در بدترین حالت برابر با $O(N_{units} - 1)$ می‌باشد. حالت دوم آن است که گره برنده یک گره مرزی باشد در این حالت حداکثر ۳ گره جدید به شبکه اضافه می‌شود و با توجه به نحوه وزن‌دهی برای گره‌های جدید می‌توان گفت پیچیدگی زمانی برای وزن‌دهی به یک گره جدید از مرتبه $O(M)$ می‌باشد. بنابراین پیچیدگی زمانی فاز رشد در بدترین حالت ممکن برابر با $O(3M + N_{units} - 1)$ است.

برای محاسبه پیچیدگی زمانی کل الگوریتم SSGSOM کافی است که جمع پیچیدگی زمانی ۷ مرحله را به دست آورد. سپس مراحل بالا برای تعداد کل داده‌های مجموعه آموزشی تکرار می‌شود. در این صورت پیچیدگی زمانی الگوریتم SSGSOM به صورت (پ-۱) به دست می‌آید

$$O\left(\sum_{t=1}^N N_{units}(t) \times (10M + 8) + 4MN\right) \quad (\text{پ-۱})$$

پ-۲ اثبات اصل ۲

اصل ۲: اگر تعداد اعضای مجموعه آموزشی از تعداد ویژگی‌های استخراج‌شده، خیلی بزرگ‌تر باشد در این صورت پیچیدگی زمانی شبکه خودسازمان‌ده پویای شبه‌ناظر از دو روش طبقه‌بندی Path-Based TSVM و SVM کمتر است.

اثبات: اگر M تعداد ویژگی‌های هر داده آموزشی و N تعداد اعضای مجموعه آموزشی باشد در این صورت پیچیدگی زمانی الگوریتم Path-based TSVM برابر با $O(N^2 + N^2M + M^2)$ می‌باشد [۱۵] در حالی که پیچیدگی زمانی روش طبقه‌بندی SVM استاندارد با QP^1 برابر با $O(N^2)$ محاسبه شده است [۲۵].

مجدد، یک سری ویژگی استخراج می‌شود. در مرحله سوم با استفاده از ویژگی‌های استخراجی در مرحله دوم، یک طبقه‌بند آموزش داده می‌شود که هدف آن قراردادن ویدئوی مشکوک در یکی از دو دسته آلوده یا پاک است. در روش پیشنهادی برای اولین بار از شبکه خودسازمان‌ده پویای شبه‌ناظر برای طراحی طبقه‌بند استفاده شد. از مزایای روش پیشنهادی می‌توان به موارد زیر اشاره کرد:

- دقت شناسایی بالا و قابل مقایسه با الگوریتم‌های یادگیری باناظر
- کاهش هزینه تولید مجموعه آموزشی از نظر تولید داده‌های برچسب‌دار
- بهبود کارایی نهان‌کارو برای شناسایی الگوریتم‌های نهان‌نگاری‌ای که کد آنها در دسترس نیست
- کاهش پیچیدگی زمانی
- کاهش هزینه آموزش مجدد سیستم

پیوست

پ-۱ اثبات اصل ۱

اصل ۱: اگر M تعداد ویژگی‌های هر داده آموزشی، N تعداد اعضای مجموعه آموزشی و $N_{units}(t)$ برابر با تعداد گره‌های لایه GSOM در هر تکرار t باشد در این صورت پیچیدگی زمانی شبکه خودسازمان‌ده پویای شبه‌ناظر برابر با $O\left(\sum_{t=1}^N N_{units}(t) \times (10M + 8) + 4MN\right)$ می‌باشد.

اثبات: در اینجا هدف به دست آوردن پیچیدگی زمانی شبکه خودسازمان‌ده پویای شبه‌ناظر (SSGSOM) می‌باشد. فرض کنید مجموعه آموزشی دارای N عضو باشد. همچنین هر داده آموزشی دارای M ویژگی است. همان طور که قبلاً توضیح داده شد الگوریتم SSGSOM کاملاً برخط می‌باشد پس در هر تکرار الگوریتم یک داده ورودی با تعداد ویژگی M وارد شبکه می‌شود.

در گام اول از الگوریتم طبقه‌بندی SSGSOM برای به دست آوردن گره برنده، فاصله اقلیدسی بین ورودی با تمام گره‌های موجود در شبکه که تعداد آنها برابر N_{units} است با استفاده از (۸) محاسبه و ذخیره می‌شود. پیچیدگی زمانی این مرحله برابر با $O(MN_{units})$ است [۲۴].

در گام دوم الگوریتم گره برنده با استفاده از (۹) تعیین می‌شود. برای به دست آوردن گره برنده کافی است به تعداد N_{units} بار عمل مقایسه انجام شود. پیچیدگی زمانی این مرحله برابر با $O(N_{units})$ است.

در گام سوم از آنجایی که در گام اول فاصله اقلیدسی بین ورودی با تمام گره‌های شبکه SSGSOM محاسبه و ذخیره شده است، پیچیدگی زمانی محاسبه تابع همسایگی در (۱۱) از مرتبه $O(1)$ است اما بایستی مقدار خطای محاسبه‌شده برای هر گره به تمام ابعاد وزن گره اضافه شود که پیچیدگی زمانی این مرحله از مرتبه $O(M)$ است. این روند برای تمام گره‌های شبکه تکرار می‌شود و بنابراین پیچیدگی زمانی کل برای این مرحله برابر با $O(MN_{units})$ است.

در گام چهارم با استفاده از (۱۵) میزان امتیاز هر گره در لایه خروجی شبکه SSGSOM محاسبه می‌شود. ابتدا سطح فعالیت هر گره (O_C) محاسبه می‌شود که شامل دو مرحله است. اولین مرحله محاسبه فاصله اقلیدسی بین ورودی $x(t)$ و گره شماره C در لایه GSOM است که در گام اول محاسبه شده است. دومین مرحله محاسبه محدوده فعالیت گره شماره C (σ_C) با استفاده از (۱۴) است. بنابراین پیچیدگی زمانی مورد نیاز برای محاسبه سطح فعالیت گره شماره C برابر با $O(\text{Count}(N_{NC})M)$ است که در اینجا N_{NC} برابر با شماره گره

است. لایه GSOM که در آن داده‌ها به صورت بدون ناظر خوشه‌بندی می‌شوند و لایه خروجی که در آن دانش لایه پایین برای تعیین نوع کلاس هر خوشه استفاده می‌شود. آقای Bezdek و همکاران [۲۶] بیان داشته‌اند با فرض این که همه داده‌ها در دسترس باشد لایه GSOM در شبکه خودسازمان‌ده پویای شبه‌ناظر در واقع نوعی الگوریتم خوشه‌بندی Fuzzy C-means می‌باشد پس بهینه است، چون نهایتاً به آنچه الگوریتم خوشه‌بندی Fuzzy C-means همگرا می‌شود، میل می‌کند. در لایه خروجی شبکه خودسازمان‌ده پویای شبه‌ناظر هدف، به دست آوردن وزن‌های بهینه بین لایه GSOM و لایه خروجی است. یکی از روش‌هایی که برای حل این گونه مسایل استفاده می‌شود الگوریتم Least Mean Square Error است که در هر تکرار از وزن با استفاده از (پ-۷) به روز می‌شود

$$w(t+1) = w(t) + \frac{1}{\gamma} \mu(-\nabla E\{e^{\tau}\}) \quad (\text{پ-۷})$$

در (پ-۷) $w(t)$ وزن در تکرار t ام است، μ نرخ یادگیری نام دارد که مشخصات همگرایی الگوریتم را مشخص می‌کند و $E\{e^{\tau}\}$ میانگین مربعات خطای بین سیگنال مرجع و مقدار تخمین زده شده آن می‌باشد. در این صورت میزان اصلاح وزن در هر تکرار با استفاده از (پ-۸) محاسبه خواهد شد

$$\Delta w = \frac{1}{\gamma} \mu(-\nabla E\{e^{\tau}\}) \quad (\text{پ-۸})$$

در لایه خروجی شبکه خودسازمان‌ده پویای شبه‌ناظر میزان امتیاز یا فعالیت هر گره (PS_m) با استفاده از (۱۵) محاسبه می‌شود و هر گره‌ای که امتیاز بیشتری کسب نماید به عنوان کلاس داده مورد نظر شناسایی خواهد شد. ξ_m که در (۱۷) تعریف شده است مشخص‌کننده میزان فعالیت مطلوب برای هر گره m در لایه خروجی است. پس می‌توان گفت میانگین مربعات خطا در شبکه خودسازمان‌ده پویای شبه‌ناظر با استفاده از (پ-۹) تعریف می‌شود

$$\begin{cases} E\{e^{\tau}\} = e^{\tau} \\ e = \xi_m - PS_m \end{cases} \quad (\text{پ-۹})$$

حال منفی گرادیان میانگین مربعات خطای تعریف شده در (پ-۹) محاسبه می‌شود. در نتیجه میزان اصلاح وزن با استفاده از روش حداقل میانگین مربعات خطا برای شبکه خودسازمان‌ده پویای شبه‌ناظر مطابق (پ-۱۰) محاسبه می‌شود و همان طور که مشخص است با مقدار اصلاح وزن در شبکه SSGSOM (رابطه (۱۶)) برابر است. پس می‌توان نتیجه گرفت لایه دوم هم در شبکه خودسازمان‌ده پویای شبه‌ناظر بهینه است

$$\begin{aligned} \Delta w_{cm} &= \frac{1}{\gamma} \mu(-\nabla E\{e^{\tau}\}) = \frac{1}{\gamma} \mu(-\tau e \frac{\partial e}{\partial w_{cm}}) = \\ &= \mu(\xi_m - PS_m) \times (-O_c) = \eta(\xi_m - PS_m) O_c \end{aligned} \quad (\text{پ-۱۰})$$

مراجع

- [1] I. J. Cox, M. L. Miller, J. A. Bloom, J. Fridrich, and T. Kalker, *Digital Watermarking and Steganography*, Morgan Kaufmann Publishers, USA, 2008.
- [2] Y. Deng, Y. Wu, H. Duan, and L. Zhou, "Digital video steganalysis based on motion vector statistical characteristics," *Optik*, vol. 124, pp. 1705-1710, 2013.
- [3] X. Xu, J. Dong, W. Wang, and T. Tan, "Video steganalysis based on the constraints of motion vectors," in *Proc. 20th IEEE Int. Conf. on Image Processing, ICIP'13*, pp. 4422-4426, Melbourne, Australia, 15-18 Sep. 2013.

یکی از مزیت‌های مهم شبکه خودسازمان‌ده پویای شبه‌ناظر این است که دارای یک فاز رشد قابل کنترل است. در ابتدای فرایند آموزش تعداد گره‌های موجود در شبکه اندک (معمولاً ۴) می‌باشد و در حین آموزش شبکه تعداد گره‌ها زیاد می‌شود، در حالی که میزان رشد گره‌ها با استفاده از ضریب گسترش (SF) قابل کنترل است. اگر در شبکه خودسازمان‌ده پویای شبه‌ناظر ضریب گسترش کوچک باشد در این صورت با توجه به این که تعداد گره‌های نهایی شبکه از تعداد کل داده‌های آموزشی خیلی کمتر است می‌توان گفت رابطه $\sum_{t=1}^N N_{units}(t) \ll N^{\tau}$ برقرار است. ولی اگر ضریب گسترش بزرگ و نزدیک یک باشد در این صورت باز هم تعداد نهایی گره‌های شبکه از تعداد اعضای مجموعه آموزشی بیشتر نخواهد بود و با توجه به رشد تدریجی تعداد گره‌های شبکه می‌توان گفت رابطه $\sum_{t=1}^N N_{units}(t) \leq N^{\tau}$ برقرار است. بنابراین می‌توان نوشت

$$\begin{aligned} \sum_{t=1}^N N_{units}(t) \times (1 \cdot M + \lambda) + \epsilon MN &\leq \\ N^{\tau} (1 \cdot M + \lambda) + \epsilon MN & \end{aligned} \quad (\text{پ-۲})$$

حال برای آن که ثابت شود شبکه SSGSOM از الگوریتم Path-Based TSMV سریع‌تر است کافی است رابطه زیر اثبات شود

$$N^{\tau} (1 \cdot M + \lambda) + \epsilon MN < N^{\tau} + N^{\tau} M + M^{\tau} \quad (\text{پ-۳})$$

اگر تعداد اعضای مجموعه آموزشی α برابر تعداد ویژگی‌های استخراج‌شده در نظر گرفته شود ($N = \alpha M$) در این صورت می‌توان (پ-۳) را به صورت (پ-۴) نوشت. حال بایستی در بدترین حالت ممکن مقادیری از α که به ازای آن (پ-۴) همواره مثبت است محاسبه شود. با توجه به این که تعداد ویژگی‌ها همواره بزرگ‌تر یا مساوی ۱ است بدترین حالت ممکن برای برقراری (پ-۴) آن است که تعداد ویژگی‌ها برابر یک در نظر گرفته شود. در این شرایط اگر $\alpha \geq 17,22$ باشد (پ-۴) همواره برقرار است

$$(\alpha^{\tau} - 9\alpha^{\tau})M - \lambda\alpha^{\tau} - \epsilon\alpha + 1 > 0 \quad (\text{پ-۴})$$

بنابراین می‌توان گفت بدون توجه به تعداد ویژگی‌های استخراج‌شده، اگر تعداد اعضای مجموعه آموزشی حداقل ۱۷,۲۲ برابر تعداد ویژگی‌های استخراجی باشد شبکه خودسازمان‌ده پویای شبه‌ناظر از الگوریتم Path-Based TSMV سریع‌تر است. به طور مشابه برای آن که ثابت شود شبکه SSGSOM از الگوریتم SVM استاندارد سریع‌تر است کافی است رابطه زیر اثبات شود

$$N^{\tau} (1 \cdot M + \lambda) + \epsilon MN < N^{\tau} \quad (\text{پ-۵})$$

اگر $N = \alpha M$ ($\alpha \gg 1$) در نظر گرفته شود در این صورت می‌توان (پ-۵) را به صورت (پ-۶) نوشت. با توجه به این که تعداد ویژگی‌ها همواره بزرگ‌تر یا مساوی ۱ است برای آن که (پ-۶) همواره مثبت باشد بدترین حالت ممکن آن است که تعداد ویژگی‌ها برابر یک باشد و در این شرایط اگر $\alpha \geq 18,22$ باشد (پ-۶) همواره برقرار است

$$(\alpha^{\tau} - 10\alpha^{\tau})M - \lambda\alpha^{\tau} - \epsilon\alpha > 0 \quad (\text{پ-۶})$$

بنابراین در بدترین شرایط ممکن و بدون توجه به تعداد ویژگی‌های استخراج‌شده، شبکه خودسازمان‌ده پویای شبه‌ناظر از الگوریتم SVM استاندارد سریع‌تر است اگر تعداد اعضای مجموعه آموزشی حداقل ۱۸,۲۲ برابر تعداد ویژگی‌های استخراجی باشد.

پ-۳ اثبات بهینه‌بودن شبکه SSGSOM

همان طور که بیان گردید شبکه SSGSOM از دو لایه تشکیل شده

- [21] 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, 21-24 May 2006.
- [22] 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.
- [23] <http://www.mpeg.org/MPEG/MSSG/tm5/>
- [24] T. Vollmer, T. Soule, and M. Manic, "A distance measure comparison to improve crowding in multi-model optimization," in *Proc. 3rd Int. Symp. on Resilient Control Systems, ISRCS'10*, pp. 31-36, 2010.
- [25] W. Tsang, T. Kwok, and P. Cheung, "Very large SVM training using core vector machines," *J. of Machine Learning Research*, vol. 6, no. 6, pp. 363-392, Apr. 2005.
- [26] J. C. Bezdek, E. C. K. Tsao, and N. R. Pal, "Fuzzy kohonen clustering networks," *Pattern Recognition*, vol. 27, no. 5, pp. 757-764, May 1994.
- [4] K. Kancherla and S. Mukkamala, "Video steganalysis using spatial and temporal redundancies," in *Proc. International Conf. on High Performance Computing & Simulation*, pp. 200-207, 21-24 Jun. 2009.
- [5] B. Liu, F. Liu, and P. Wang, "Inter-frame correlation based compressed video steganalysis," in *Proc. Congress on Image and Signal Processing, CISP'08*, pp. 42-46, 27-30 May. 2008.
- [6] K. Kancherla and S. Mukkamala, "Video steganalysis using motion estimation," in *Proc. of Int. Joint Conf. on Neural Networks*, pp. 1510-1515, Atlanta, GA, USA, 14-19 Jun. 2009.
- [7] V. Pankajakshan, G. Doerr, and P. Kumar, "Detection of motion-incoherent components in video streams," *IEEE Trans. on Information Forensics and Security*, vol. 4, no. 1, pp. 49-58, Mar. 2009.
- [8] 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, Jan. 2012.
- [9] Q. Liu, A. H. Sung, and M. Qiao, "Video steganalysis based on the expanded markov and joint distribution on the transform domains-detecting MSU stego video," in *Proc. 7th Int. Conf. on Machine Learning and Applications*, pp. 671-674, San Diego, CA, USA, 11-13 Dec. 2008.
- [10] S. Yuting, C. Zhang, and C. Zhang, "A video steganalytic algorithm against motion-vector-based steganography," *Signal Processing*, vol. 91, no. 8, pp. 1901-1909, Aug. 2011.
- [11] V. Pankajakshan and A. T. S. Ho, "Improving video steganalysis using temporal correlation," in *Proc. 3rd Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing*, vol. 1, pp. 287-290, 26-28 Nov. 2007.
- [12] U. Budhia, D. Kundur, and T. Zourntos, "Digital video steganalysis exploiting statistical visibility in the temporal domain," *IEEE Trans. on Information Forensics and Security*, vol. 92, no. 4, pp. 502-516, Dec. 2006.
- [13] K. Kashyap and P. K. Bora, "Spatial averaging based steganalysis scheme to detect antipodal watermarks," in *Proc. Int. Conf. Signal Processing and Communications, SPCOM'10*, 5 pp., 18-21 Jul. 2010.
- [14] H. G. Schaathun, *Machine Learning in Image Steganalysis*, John Wiley & Sons, United Kingdom, 2012.
- [15] X. Zhang and S. P. Zhong, "An improved path-based TSVM algorithm for blind steganalysis classification," *Artificial Intelligence and Computational Intelligence*, vol. 5855, pp. 453-462, 2009.
- [16] T. Joachims, "Transductive inference for text classification using support vector machines," in *Proc. of The 26th Int. Conf. on Machine Learning, ICML'99*, pp. 200-209, 1999.
- [17] O. Chapelle and A. Zien, "Semi-supervised classification by low density separation," in *Proc. of the 10th International Workshop on Artificial Intelligence and Statistics, AISTAT'05*, pp. 57-64, 6-8 Jan. 2005.
- [18] 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.
- [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] C. Xu, X. Ping, and T. Zhang, "Steganography in compressed video stream," in *Proc. IEEE Int. Conf. Innovative Computing, Information and Control*, vol. 1, pp. 269-272, Beijing, China, 30 Aug.-1 Sep. 2006.

جواد مرتضوی مهریزی در سال ۱۳۸۸ تحصیلات خود را در مقطع کارشناسی در گرایش مخابرات در دانشگاه آزاد اسلامی واحد مشهد به پایان رسانده است. سپس مدرک کارشناسی ارشد خود را در گرایش سیستم از دانشگاه فردوسی مشهد در سال ۱۳۹۲ دریافت نمود. زمینه تحقیقاتی مورد علاقه ایشان عبارتند از: مخابرات تصویری، مخابرات امن، پردازش سیگنال، شناسایی آماری الگو و رهگیری اهداف.

مرتضی خادمی متولد ۱۳۳۷، مدرک کارشناسی و کارشناسی ارشد خود را به ترتیب در سالهای ۱۳۶۴ و ۱۳۶۶ در گرایش‌های الکترونیک و مخابرات از دانشگاه صنعتی اصفهان اخذ نمود. نامبرده بلافاصله در دانشگاه فردوسی مشهد به عنوان عضو هیأت علمی استخدام گردیده و پس از چهار سال تدریس در سال ۱۳۷۰ برای ادامه تحصیل به کشور استرالیا عزیمت نموده و در سال ۱۳۷۴ مدرک دکتری را اخذ کرد. وی هم‌اکنون به عنوان استاد گروه برق دانشگاه فردوسی مشهد مشغول به کار می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان، مخابرات تصویری، فشرده سازی تصویر و صوت و پردازش سیگنال‌های پزشکی می‌باشد. وی تاکنون بیش از یکصد مقاله در کنفرانس‌ها و مجلات داخل و خارج کشور در زمینه‌های فوق به چاپ رسانیده است. اجرای چهار طرح پژوهشی برای صنعت و دانشگاه و ترجمه تنها کتاب پردازش تصویر به زبان فارسی از دیگر فعالیت‌های وی محسوب می‌گردد. ایشان در امور اجرایی نیز فعال بوده و علاوه بر عضویت در کمیته‌های مختلف مدیریتی در سطح دانشگاه و همکاری در برگزاری کنفرانس‌های مهندسی برق (ICEE2004) و کنفرانس پردازش تصویر و بینایی ماشین (IPMV2006)، در سمت‌های معاون دانشکده، معاون دانشگاه، مشاور رئیس دانشگاه و عضو شورای اسلامی شهر مقدس مشهد نیز فعالیت داشته است.

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 دکتری مهندسی برق را به ترتیب در سال‌های 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، داده‌کاوی و بهینه‌سازی.