



دانشکده مهندسی کامپیوتر

تشخیص ناهنجاری با استفاده از یادگیری عمیق

گزارش سمینار کارشناسی ارشد
در رشته مهندسی کامپیوتر-گرایش هوش مصنوعی و رباتیک

نام دانشجو:
علی نادری پاریزی

استاد راهنما:
دکتر محسن سریانی

۱۴۰۱ آذر ماه

لَهُ مُلْكُ الْأَنْعَمِ الْأَنْجَوِي

چکیده

تشخیص ناهنجاری مسئله مهمی است که در زمینه‌های تحقیقاتی گوناگون مورد مطالعه قرار می‌گیرد و کاربردهای بسیار زیادی دارد. یک نیاز مرسوم در حوزه تجزیه و تحلیل داده‌های دنیای واقعی، پی بردن به این است که بدانیم کدام نمونه‌ها از نقطه‌نظر تشابه رفتار و ظاهر با اکثریت نمونه‌های موجود بسیار متفاوت هستند. این تفاوت می‌تواند به دلیل خطای اندازه‌گیری در هنگام جمع آوری داده‌ها باشد. گاهی اوقات این تفاوت می‌تواند نشان دهنده وجود پدیده‌ای ناشناخته باشد که در پشت‌پرده جامعه آماری مورد مطالعه در حال رخ دادن است و ما از آن بی‌خبر هستیم.

در علم داده اصطلاح ناهنجاری به داده‌ای تعلق می‌گیرد که از نقطه‌نظر یک معیار تشابه تعریف شده، میزان تشابه آن با سایر دادگان موجود بسیار کم باشد. برای مثال اگر عکس رادیولوژی فردی که بیماری ریوی دارد را با عکس‌های رادیولوژی گرفته شده از ریه افراد سالم مقایسه کنیم متوجه تفاوت این عکس با سایر عکس‌ها خواهیم شد. این عدم تشابه در دادگان، مشخص می‌کند که فرد دچار بیماری ریوی است. درواقع پزشکان با مشاهده این عدم شباهت‌ها به وجود بیماری بی می‌برند. عمل مقایسه دادگان می‌تواند به وسیله کامپیوتر نیز انجام شود که موضوع این سمینار است.

در این سمینار تلاش شده روش‌های مبتنی بر یادگیری عمیق برای تشخیص ناهنجاری را بررسی کنیم. از آنجا که کاربرد این موضوع در حوزه‌های مختلف بسیار وسیع است و مقالات بسیاری در رابطه با کاربردی‌های مختلف به چاپ رسیده است، سعی کردیم حوزه سمینار را محدود کرده و ضمن معرفی انواع کاربردهای مسئله تشخیص ناهنجاری، به بررسی روش‌هایی پردازیم که در رابطه با کاربرد پردازش تصویر و بینایی کامپیوتر هستند. با توجه به تعدد مقالات در سال‌های اخیر و وجود مقالات جامع در این حوزه، بیشتر مقالات جدید که در سال‌های اخیر منتشر شده‌اند را بررسی می‌کنیم و برای باقی روش‌ها به ارجاع دهی به مقالات دیگر اکتفا خواهیم کرد.

واژه‌های کلیدی: تشخیص ناهنجاری، پردازش تصویر، شبکه‌های عمیق

فهرست مطالب

۱	۱	مقدمه
۲	۱.۱	تشخیص ناهنجاری
۳	۲.۱	جنبهای مختلف تشخیص ناهنجاری
۳	۲.۱.۱	کاربردهای تشخیص ناهنجاری
۴	۱.۳.۱	امنیت سیستم و تشخیص نفوذ
۴	۲.۳.۱	تشخیص جعل اسناد و کلاهبرداری
۴	۲.۳.۱.۱	سلامت و پزشکی
۴	۴.۳.۱	سامانه‌های هوشمند و اینترنت اشیا
۴	۵.۳.۱	ناظارت ویدیویی و سیستم‌های امنیتی
۵	۶.۳.۱	خودروهای خودران
۵	۴.۱	چالش‌های تشخیص ناهنجاری
۶	۱.۴.۱	چالش‌های عمومی تشخیص ناهنجاری
۶	۲.۴.۱	چالش‌های تشخیص ناهنجاری که می‌توان با بکارگیری روش‌های عمیق به سراغ آنها رفت
۸	۵.۱	ساختار کلی روش‌های تشخیص ناهنجاری
۹	۶.۱	ساختار گزارش
۱۰	۲	مروری بر کارهای انجام شده برای تشخیص ناهنجاری
۱۰	۱.۲	مروری بر روش‌های سنتی
۱۱	۱.۱.۲	روش‌های مبتنی بر ردیابندی
۱۳	۲.۱.۲	روش‌های مبتنی بر معیار فاصله
۱۳	۳.۱.۲	روش‌های مبتنی بر مدل آماری

۱۳	استفاده از یادگیری عمیق برای تشخیص ناهنجاری	۲.۲
۱۴	استفاده از یادگیری عمیق برای یادگیری بازنمایی دادگان	۱.۲.۲
۱۴	خود کدگذار	۲.۲.۲
۱۸	مدل‌های مولد	۲.۲.۲
۲۱	مجموعه دادگان موجود برای تشخیص ناهنجاری	۳.۲
۲۳	کارهای آینده	۳
۲۳	تشخیص ناهنجاری با نظارت ضعیف	۱.۳
۲۴	موضوعات کاربردی جدید مرتبط با تشخیص ناهنجاری	۲.۳
۲۴	موضوع پیشنهادی برای پایان نامه	۳.۳
۲۶	مراجع	

فهرست شکل‌ها

۱	مثالی از تفاوت دادگان ناهنجار و نوین	۱.۱
۲	مثالی ساده از نقاط ناهنجار در میان مجموعه داده‌ای در فضای دوبعدی (نقاط O_1, O_2, O_3 نقاط ناهنجار را نشان می‌دهند) [۱]	۲.۱
۳	ناهنجاری نقطه‌ای و مجموعه‌ای [۲]	۳.۱
۵	ناهنجاری در کاربرد نظارت ویدیو [۳]	۴.۱
۵	مثال‌هایی از ناهنجاری در تصاویر: به ترتیب از سمت چپ، توده سرطان سینه، مین زیردریابی، نقص رنگ‌آمیزی کاشی تولید شده در کارخانه، نمونه نقص موجود در چرخ خودرو. [۴]	۵.۱
۱۲	ماشین بردار پشتیبان یک کلاسه	۱.۲
۱۲	بردار پشتیبان توصیفگر داده [۵]	۲.۲
۱۳	نمایش کلی روش جداساز پرت محلی [۶]	۳.۲
۱۴	بردار پشتیبان توصیفگر داده عمیق [۵]	۴.۲
۱۵	مدل خودکدگذار	۵.۲
۱۷	مدل خودکدگذار حذف نویز [۷]	۶.۲
۱۹	مدل خود رمز کننده [۷] variational	۷.۲
۲۰	مدل پیشنهادی برای ترکیب ویژگی‌های بصری و متنی برای تشخیص اخبار جعلی [۸]	۸.۲
۲۱	شبکه مولد رقابتی	۹.۲
۲۱	نمایش نحوه آموزش مدل F-AnoGan [۹]	۱۰.۲

فهرست جدول‌ها

۱۱	۱.۲ دسته‌بندی روش‌های سنتی
۲۲	۲.۲ مجموعه دادگان در دسترس برای تشخیص ناهنجاری
۲۲	۳.۲ الگوریتم‌های عمیق مورد استفاده در تشخیص ناهنجاری

فصل ۱

مقدمه

تشخیص ناهنجاری^۱ مسئله مهمی است که در زمینه‌های تحقیقاتی گوناگون مورد مطالعه قرار می‌گیرد و کاربردهای بسیار زیادی دارد. یک نیاز مرسوم در حوزه تجزیه و تحلیل داده‌های دنیای واقعی، پی بردن به این موضوع است که بدانیم کدام نمونه‌ها از نقطه نظر تشابه رفتار و ظاهر با اکثریت نمونه‌های موجود بسیار متفاوت هستند. این تفاوت می‌تواند به دلیل خطای اندازه‌گیری در هنگام جمع آوری داده‌ها باشد. گاهی اوقات این تفاوت می‌تواند نشان دهنده وجود پدیده‌ای ناشناخته باشد که در پشت‌پرده جامعه آماری مورد مطالعه در حال رخ دادن است و ما از آن بی‌خبر هستیم.



شکل ۱.۱: مثالی از تفاوت دادگان ناهنجار و نوین

در کنار ناهنجاری‌ها، دادگان دیگری نیز وجود دارند که با دادگان عادی متفاوت‌اند اما این تفاوت به اندازه‌ی کافی زیاد نیست. به این دادگان اصطلاحاً دادگان نوین^۲ گفته می‌شود. دادگان نوین درواقع دادگانی هستند که در دسته دادگان عادی قرار می‌گیرند اما چون هنوز کشف نشده‌اند به نظر می‌رسد که با دادگان عادی تفاوت داشته باشند. برای مثال، اکثر ببرهای دیده شده و شناخته شده به رنگ نارنجی و با خطوط راه راه سیاه هستند و دیدن ببر سفید برای ما تعجب آور خواهد بود. اما همه به خوبی می‌دانیم که ببر سفید درواقع یک ببر است که فقط رنگ آن غیرعادی است و نباید آن را در دسته جدایی

¹Anomaly detection

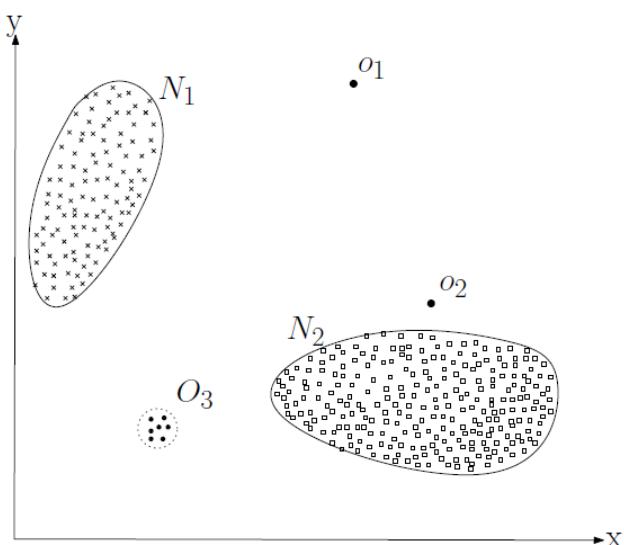
²Novelties

از حیوانات قرار داد.

در ادامه این فصل پس از تعریف ناهنجاری در دادگان، به بیان کاربردهای این بحث در حوزه‌های مختلف می‌پردازیم. سپس یک تعریف معیار که مرتبط با حوزه مورد نظر ما که همان پردازش تصویر است ارائه می‌دهیم. پس از تعریف حوزه مورد مطالعه و بررسی اهمیت موضوع، به توضیح ساختار کلی گزارش این سمینار خواهیم پرداخت.

۱.۱ تشخیص ناهنجاری

تشخیص ناهنجاری که با عنوان تشخیص دادگان پرت^۳ نیز شناخته می‌شود، به عملیاتی گفته می‌شود که طی آن به آشکارسازی نمونه‌هایی از مجموعه دادگان می‌پردازد که تفاوت زیادی با اکثریت دادگان موجود دارد. در واقع، اینجا تفاوت به معنی متفاوت بودن مشخصات و ویژگی‌های این نمونه‌ها با الگوی معمول موجود در مجموعه دادگان است. این مسئله یک موضوع فعال تحقیق در دهه‌های اخیر بوده است. مطالعه برای تشخیص نقاط خارج از دامنه در حوزه آمار به قرن ۱۹ میلادی بر می‌گردد که یکی از مقالات معروف آن مربوط به سال ۱۸۸۷ میلادی است [۱۰]. کاربردهای تشخیص ناهنجاری بسیار وسیع است و در حوزه‌های گوناگونی مورد استفاده قرار می‌گیرد.



شکل ۲.۱: مثالی ساده از نقاط ناهنجار در میان مجموعه داده‌ای در فضای دوبعدی (نقاط O_1, O_2, O_3 نقاط ناهنجار را نشان می‌دهند) [۱]

ناهنجاری‌ها انواع مختلفی دارند که بسته به کاربرد و مفاهیم مختلف تعریف می‌شوند. به طور کلی می‌توان برای ناهنجاری‌ها سه نوع مختلف درنظر گرفت که عبارت اند از ناهنجاری نقطه‌ای^۴، ناهنجاری مفهومی^۵، ناهنجاری مجموعه‌ای^۶. اکثر کارهای انجام شده در متون علمی در مورد ناهنجاری نقطه‌ای بحث شده است. در این گونه ناهنجاری دادگان به صورت نقاطی در فضا درنظر گرفته می‌شوند و دادگان ناهنجار، نقاطی در فضای مورد نظر هستند که با دیگر دادگان فاصله دارند و

³Outlier detection

⁴Point anomaly

⁵Contextual anomalies

⁶Collective anomalies

رفتاری تصادفی از خود نشان می‌دهند که اغلب تفسیر خاصی ندارند. برای مثال مبلغ بسیار بالای تراکنش در یک رستوران تراکنشی غیر عادی به حساب می‌آید که با در نظر گرفتن آن در فضای بازنمایی دادگان، نقطه شباهتی به دیگر دادگان نخواهد داشت. دسته دوم، ناهنجاری‌های مفهومی هستند که در این دسته مفهوم داده در یک مکان و یا زمان مختلف می‌تواند به صورت ناهنجاری درنظر گرفته شود. برای مثال عبور وسیله نقلیه در خیابان یک امر طبیعی است اما تردد وسایل نقلیه در مسیر عابرین پیاده یک پدیده غیرعادی است. نوع سوم ناهنجاری‌ها که اصطلاحاً ناهنجاری مجموعه‌ای گفته می‌شود، مفهوم ناهنجاری را در یک سلسله از رویدادها دنبال می‌کند در حالی که هر رویداد یک داده کاملاً عادی است. برای مثال در دنباله تراکنش‌های یک کارت اعتباری وجود چندین تراکنش یکسان با فواصل زمانی بسیار، کم مشکوک است.

May-22	1:14 pm	FOOD	Monaco Café	\$1,127.80	→ Point Anomaly
May-22	2:14 pm	WINE	Wine Bistro	\$28.00	
...					
Jun-14	2:14 pm	MISC	Mobil Mart	\$75.00	
Jun-14	2:05 pm	MISC	Mobil Mart	\$75.00	
Jun-15	2:06 pm	MISC	Mobil Mart	\$75.00	
Jun-15	11:49 pm	MISC	Mobil Mart	\$75.00	
May-28	6:14 pm	WINE	Acton shop	\$31.00	
May-29	8:39 pm	FOOD	Crossroads	\$128.00	
Jun-16	11:14 am	MISC	Mobil Mart	\$75.00	
Jun-16	11:49 am	MISC	Mobil Mart	\$75.00	

شکل ۳.۱: ناهنجاری نقطه‌ای و مجموعه‌ای [۲]

۲.۱ جنبه‌های مختلف تشخیص ناهنجاری

مسئله تشخیص ناهنجاری را از جنبه‌های مختلفی می‌توان مورد بررسی قرار داد. برای مثال می‌توان روش‌های موجود را بر اساس ماهیت دادگان موجود مورد بررسی قرار داد و با توجه به نوع داده، انواع روش‌ها را دسته‌بندی کرد. برای نمونه می‌توان ماهیت دادگان را به دو دسته کلی، دنباله‌ای^۷ (مانند صدا، موسیقی، فیلم، متن و ...) غیر دنباله‌ای (مانند عکس، علائم بیماری و ...) تقسیم کرد. و یا بر اساس تعداد ویژگی‌های داده ورودی به دو دسته ابعاد پایین و ابعاد بالا تقسیم کرد. همچنین می‌توان روش‌های تشخیص ناهنجاری‌ها را از دید در دسترس بودن برچسب دادگان نیز بررسی کرد. اما باید توجه داشت که پدیده‌های ناهنجار اصولاً کم اتفاق می‌افتد و تعداد آنها در دادگان موجود کم است. با این حال می‌توان روش‌های تشخیص ناهنجاری را بر اساس در دسترس بودن برچسب دادگان به سه دسته باناظر، با نظارت ضعیف و همچنین بدون ناظر تقسیم کرد.

۳.۱ کاربردهای تشخیص ناهنجاری

برای درک اهمیت و کاربرد مسئله تشخیص ناهنجاری می‌توان به حجم مقالات چاپ شده در این حوزه و دامنه وسیع موضوعات تحقیقاتی اشاره کرد که حول این موضوع انجام شده و یا در حال انجام است. در این قسمت برخی از کاربردهای مسئله تشخیص ناهنجاری را به تفصیل حوزه‌های کاربردی مختلف می‌آوریم.

⁷Streaming data

۱.۳.۱ امنیت سیستم و تشخیص نفوذ

تشخیص نفوذ در کاربرد امنیت سایبری که عمل تشخیص و اطلاع پیدا کردن از دسترسی‌های غیر مجاز به شبکه و یا سامانه‌های رایانه‌ای است می‌تواند یکی از کاربردهای مسئله تشخیص ناهنجاری باشد. در اینگونه مسائل با بررسی گزارش‌های سیستم در طول زمان به عنوان داده ورودی، به بررسی این قضیه می‌پردازند. همانطور که مشخص است، نوع ناهنجاری در این جا میتواند از دو نوع دنباله‌ای و یا مفهومی باشد.

۲.۳.۱ تشخیص جعل اسناد و کلاهبرداری

تشخیص مدارک جعلی در حوزه‌های مختلف مانند هویتی، بانکی، بیمه، کارت اعتباری و غیره بسیار کارآمد است. در اینگونه کاربردها نیز مدارک از جنبه‌های مختلفی با یکدیگر مقایسه می‌شوند تا مدارک جعلی از مدارک حقیقی تشخیص داده شوند. برای مثال، در جعل تراکنش‌های بانکی، میتوان با بررسی تاریخچه تراکنش‌ها، به عنوان داده ورودی، به یافتن تراکنش‌های غیر مجاز و جعلی پرداخت.

۳.۳.۱ سلامت و پزشکی

بررسی گزارش‌های پزشکی یک حوزه بسیار فعال در علم کامپیوتر و مهندسی پزشکی بوده است. مقایسه و بررسی این گزارش‌ها از دید مسئله تشخیص ناهنجاری نیز بسیار مورد مطالعه قرار گرفته و کاربردهای فراوانی دارد. برای مثال در بررسی تصاویر پزشکی می‌توان از دید مسئله تشخیص ناهنجاری به یافتن بیماری‌ها و نواقص بیمار و علت بیماری پرداخت. همچنین بررسی گزارش علائم بیمار مانند ضربان قلب، سیگنال‌های مغز، فشار خون و غیره توسط دستگاه‌های پزشکی با هدف آگاهی از شرایط بحرانی و کنترل شرایط بیمار بسیار مناسب است. در این نوع کاربردها دادگان به صورت دنباله‌ای از رویدادها به عنوان داده ورودی مورد بررسی قرار می‌گیرند تا در صورت بروز علائم و شرایط حیاتی غیر طبیعی از پیش‌آمدن اتفاقات ناگوار جلوگیری کنند.

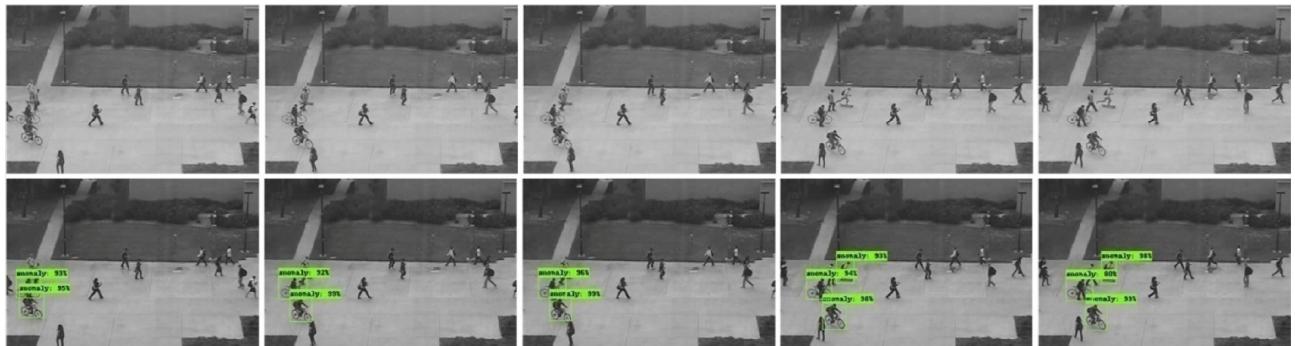
۴.۳.۱ سامانه‌های هوشمند و اینترنت اشیا

در سیستم‌های خانه هوشمند، سامانه‌های خودکار و اینترنت اشیا معمولاً بسیاری از حسگرهای دستگاهها با استفاده از شبکه‌هایی به هم متصل شده‌اند که برای بررسی وضعیت کلی سیستم و اطمینان از کارکرد صحیح سیستم می‌توان رویدادهای سامانه را در طول زمان مورد بررسی و ارزیابی قرار داد. کاربرد مسئله تشخیص ناهنجاری در اینجا بررسی گزارش‌های سامانه در طی زمان برای پی‌بردن به اتفاق افتادن شرایط نامتعادل و خطاهای سامانه است.

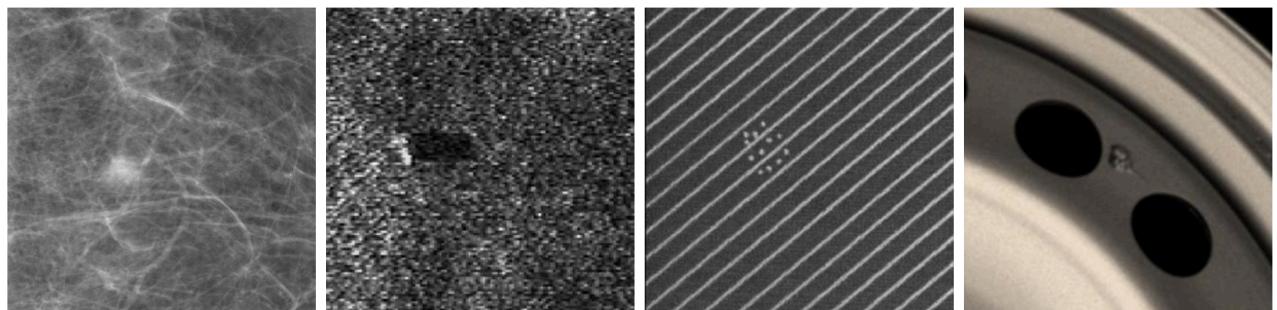
۵.۳.۱ نظارت ویدیویی و سیستم‌های امنیتی

دوربین‌های امنیتی در بسیاری از مکان‌ها برای بالابردن امنیت و همچنین نظارت بر افراد و وضعیت کلی مکان مورد استفاده قرار می‌گیرند اما بررسی و نظارت بر فیلم‌های ضبط شده توسط این دوربین‌ها کار بسیار دشوار و وقت‌گیری است که، در مقیاس وسیع این امر نزدیک به غیر ممکن می‌شود. برای مثال نظارت کارآمد دوربین‌های موجود در سطح شهر تهران برای کنترل ترافیک کار بسیار دشواری است و در صورتی که بخواهیم این کار را با استفاده از منابع انسانی انجام دهیم وقت و منابع بسیاری را می‌طلبد. یکی از کاربردهای مسئله تشخیص ناهنجاری در این حوزه بررسی ویدیوها و تلاش برای یافتن پدیده‌های غیر عادی است. برای مثال تشخیص ناهنجاری در تشخیص عبور غیرمجاز وسایل نقلیه، تشخیص تخلف‌های رانندگی، بررسی

امنیت مکان‌های عمومی، وضعیت خط تولید کارخانه برای یافتن کالاهای معیوب و کاربردهای دیگری از این قبیل بسیار مورد استفاده قرار می‌گیرد.



شکل ۴.۱: ناهنجاری در کاربرد ناظارت ویدیو [۲]



شکل ۵.۱: مثال‌هایی از ناهنجاری در تصاویر: به ترتیب از سمت چپ، توده سرطان سینه، مین زیردریایی، نقص رنگ‌آمیزی کاشی تولید شده در کارخانه، نمونه نقص موجود در چرخ خودرو. [۲]

۶.۳.۱ خودروهای خودران

یکی دیگر از حوزه‌های بسیار پرطرفدار در سال‌های اخیر ساخت خودروهای خودران و رانندگی خودکار وسایل نقلیه مختلف است. در این گونه سیستم‌ها نیز می‌توان با بررسی وضعیت حسگرها و دوربین‌های نصب شده بر روی وسیله نقلیه به بررسی خطرات احتمالی و شرایط غیرعادی مسیر در حال عبور پرداخت. با توجه به اینکه شرایط غیر عادی در رانندگی که منجر به تصادف و خطر شود به ندرت اتفاق می‌افتد و همچنان این شرایط می‌تواند به صورت‌ها و شکل‌ها مختلف روی دهد، استفاده از روش‌های تشخیص ناهنجاری در این حوزه کاربردی مورد توجه پژوهشگران قرار گرفته است.

۴.۱ چالش‌های تشخیص ناهنجاری

با توجه به ماهیت منحصر به فرد ناهنجاری‌ها، روش‌های تشخیص ناهنجاری، با چالش‌های اساسی و عمدتی روبرو هستند که برخی از آنها هنوز به صورت قابل قبولی حل نشده و تلاش برای حل آنها هنوز یک حوزه پژوهشی فعال است. در این بخش پیچیدگی‌های ذاتی و چالش‌های عمدتی در مسئله تشخیص ناهنجاری را شرح می‌دهیم.

۱.۴.۱ چالش‌های عمومی تشخیص ناهنجاری

بر خلاف سایر بحث‌های یادگیری ماشین که به یافتن پدیده‌ها و الگوهای مشخص می‌پردازند، روش‌های تشخیص ناهنجاری به دنبال یافتن الگوهایی غیرقابل پیش‌بینی، نامفهوم و کمیاب هستند که این باعث می‌شود پیچیدگی‌های منحصر به فرد و عمومی در روش‌های تشخیص ناهنجاری وجود داشته باشد.

۱. مجھول بودن^۸ : ناهنجاری‌ها با بسیاری از مجھولات مرتبط هستند، به عنوان مثال، نمونه‌هایی با رفتارهای ناگهانی ناشناخته، ساختارها و توزیع‌های داده ناشناخته. آنها تا زمانی که واقعاً رخ ندهند ناشناخته می‌مانند، مانند حملات تروریستی، کلاهبرداری‌ها و نفوذگاهی شبکه.

۲. ناهمگن^۹ بودن دسته‌های ناهنجار : ناهنجاری‌ها نامنظم هستند، درواقع، یک دسته از ناهنجاری‌ها ممکن است ویژگی‌های غیرطبیعی کاملاً متفاوتی با دسته دیگر از ناهنجاری‌ها داشته باشد. به عنوان مثال، در نظرارت تصویری، رویدادهای غیرعادی سرقت و یا تصادفات رانندگی از نظر بصری بسیار متفاوت هستند.

۳. کمیاب بودن و عدم توازن دادگان ناهنجار و عادی: برخلاف نمونه‌های عادی که اغلب بخش بزرگی از داده‌ها را تشکیل می‌دهند، ناهنجاری‌ها معمولاً نادر هستند. بنابراین، جمع‌آوری مقدار زیادی از نمونه‌های غیرعادی برچسب‌گذاری شده، اگر نگوییم غیرممکن، ولی دشوار است. این منجر به در دسترس نبودن داده‌های برچسب‌گذاری شده در مقیاس‌های بزرگ در اکثر کاربردها می‌شود. باید توجه داشت که رده‌بندی نادرست ناهنجاری‌ها معمولاً بسیار پرهزینه تراز نمونه‌های عادی است.

۴. گوناگونی انواع ناهنجاری: به صورت کلی ناهنجاری‌ها دارای سه دسته کلی هستند که در بخش قبل آنها را معرفی کردیم. با این وجود، یکی از چالش‌های عمومی تشخیص ناهنجاری، گوناگونی انواع مختلف آن خواهد بود.

۲.۴.۱ چالش‌های تشخیص ناهنجاری که می‌توان با بکارگیری روش‌های عمیق به سراغ آنها رفت

ماهیت پیچیده ناهنجاری باعث به وجود آمدن چالش‌های بسیاری در تشخیص آن شده است. برخی از چالش‌ها مانند مقیاس پذیری با توجه به اندازه دادگان، در سال‌های اخیر مورد توجه قرار گرفته است. اما چالش‌های اساسی و حل نشده دیگری برای تشخیص ناهنجاری وجود دارند که یادگیری عمیق می‌تواند در حل آنها بسیار کمک کننده باشد. از جمله‌ای این چالش‌ها می‌توان به موارد زیر اشاره کرد:

۱. نرخ پایین یادآوری در روش‌های تشخیص ناهنجاری: از آنجایی که ناهنجاری‌ها بسیار نادر و ناهمگن هستند، شناسایی همه ناهنجاری‌ها دشوار است. بسیاری از نمونه‌های عادی به اشتباه به عنوان ناهنجاری گزارش می‌شوند در حالی که ناهنجاری‌های واقعی و در عین حال پیچیده، نادیده گرفته می‌شوند. اگرچه تعداد زیادی از روش‌های تشخیص ناهنجاری در طول سال‌ها معرفی شده‌اند، روش‌های پیشرفت‌های فعلی، به‌ویژه روش‌های بدون نظرارت (به عنوان مثال [۱۱]، هنوز دارای نرخ مثبت اشتباه^{۱۰} بالایی در مجموعه داده‌های دنیای واقعی هستند [۱۲]). چگونگی کاهش نرخ مثبت اشتباه و افزایش نرخ یادآوری تشخیص، یکی از چالش‌های مهم و در عین حال دشوار است و با توجه به هزینه قابل توجه عدم شناسایی ناهنجاری‌ها در کاربردهای مختلف، از اهمیت ویژه‌ای برخوردار است.

⁸Unknownness

⁹Heterogeneous

¹⁰False positive

۲. تشخیص ناهنجاری در ابعاد بالا و با وجود دادگان نه لزوماً مستقل: ناهنجاری‌ها اغلب ویژگی‌های غیرعادی آشکاری را در فضایی با ابعاد پایین نشان می‌دهند، اما در فضایی با ابعاد بالا پنهان و غیرقابل تشخیص می‌شوند. تشخیص ناهنجاری در ابعاد بالا یک چالش قدیمی برای این مسئله بوده است. تشخیص ناهنجاری در فضایی با ابعاد پایین که با استفاده از زیرمجموعه کوچکی از ویژگی‌های اصلی، یا ویژگی‌های جدید ساخته شده صورت می‌گیرد، راه حلی ساده برای این چالش در نظر گرفته می‌شود. به عنوان مثال، در روش‌های مبتنی بر زیرفضا [۱۳، ۱۴، ۱۵، ۱۶]^{۱۱} و روش‌های مبتنی بر انتخاب ویژگی [۱۷، ۱۸، ۱۹] از این ایده استفاده شده است. با این حال، شناسایی و در نظر گرفتن برهمکنش‌ها (مثلاً مرتبه بالا، غیرخطی و ناهمگن) [۲۰] ممکن است در داده‌هایی با ابعاد بالا ضروری باشد، اما همچنان یک چالش بزرگ برای تشخیص ناهنجاری است. علاوه بر این، چگونه می‌توان تضمین کرد که فضای ویژگی جدید اطلاعات مناسب را برای روش‌های تشخیصی خاص حفظ می‌کند.

۳. یادگیری ناهنجاری‌ها با داده حداقل^{۱۱}: به دلیل دشواری و هزینه بالای جمع‌آوری داده‌های ناهنجاری برچسب‌گذاری شده در مقیاس بزرگ، تشخیص ناهنجاری به صورت کاملاً ناظارت شده، اغلب غیرعملی است. زیرا در دسترس بودن داده‌های آموزشی برچسب‌گذاری شده با دسته‌های عادی و غیرعادی را می‌طلبید. در دهه گذشته، تمرکز اغلب پژوهش‌های مرتبط، بر روی رویکردهای بدون ناظر بوده است که به هیچ داده آموزشی برچسب‌گذاری شده‌ای نیاز ندارد. با این حال، روش‌های بدون ناظر هیچ گونه آگاهی قبلی از ناهنجاری‌های واقعی ندارند. آنها به شدت بر فرض خود در مورد توزیع ناهنجاری‌ها تکیه می‌کنند. از سوی دیگر، جمع‌آوری داده‌های عادی برچسب‌گذاری شده و استفاده از تعداد کمی از داده‌های ناهنجار برچسب‌گذاری شده دشوار نیست. در عمل، اغلب پیشنهاد می‌شود که تا حد امکان از چنین داده‌های برچسب‌گذاری شده که به آسانی در دسترس هستند استفاده شود [۲۱]. بنابراین، استفاده از این داده‌های برچسب‌گذاری شده برای یادگیری بازنمایی‌های توصیف از نرمال و نابهنجار برای تشخیص دقیق ناهنجاری بسیار مهم است. تشخیص ناهنجاری به صورت نیمه ناظارت شده، که مجموعه‌ای از داده‌های آموزش عادی و برچسب‌گذاری شده را استفاده می‌کند، یک حوزه تحقیقاتی است که به این موضوع می‌پردازد. حوزه تحقیقاتی دیگر، تشخیص ناهنجاری با ناظارت ضعیف است که فرض می‌کند برچسب‌های مختلفی برای دسته‌های ناهنجار وجود دارند اما این برچسب‌ها جزئی یا ناقص (بخشی از برچسب موجود است)، نادریق (برچسب‌ها ممکن است کلی باشند) و یا نادرست هستند (برخی از برچسب‌ها ممکن است نادرست باشند). چالش اصلی این است که چگونه توزیع دادگان عادی و ناهنجار را با استفاده از تعداد کمی نمونه ناهنجار برچسب‌گذاری شده یاد بگیریم. همچنین چگونه مدل‌های تشخیصی را یاد بگیریم که به با استفاده از نمونه‌های ناهنجاری شناخته شده دسته‌ها و انواع ناشناخته ناهنجاری را نیز تشخیص دهد.

۴. تشخیص ناهنجاری مقاوم در برابر نویز: بسیاری از روش‌های تشخیص ناهنجاری با ناظارت ضعیف فرض می‌کنند که داده‌های آموزشی برچسب‌گذاری شده تمیز هستند. این فرض می‌تواند در برابر نمونه‌هایی با نویز بالا که به اشتباه به عنوان برچسب کلاس مخالف برچسب‌گذاری شده‌اند آسیب‌پذیر باشد. در چنین مواردی، ممکن است به جای استفاده از روش‌های باناظر، از روش‌های بدون ناظر استفاده کنیم اما این روش‌ها از داده‌های برچسب‌گذاری شده واقعی استفاده نمی‌کنند. علاوه بر این، اغلب داده‌های بدون برچسب آلوده به ناهنجاری در مقیاس بزرگ وجود دارد. مدل‌های مقاوم در برابر نویز می‌توانند از این داده‌های بدون برچسب برای تشخیص دقیق‌تر استفاده کنند. بنابراین،

¹¹Data efficient

نویز در اینجا می‌تواند داده‌های دارای برچسب اشتباہ و یا ناهنجاری‌های بدون برچسب باشد. چالش اصلی این است که میزان نویز می‌تواند به طور قابل توجهی با مجموعه داده‌ها متفاوت باشد و نمونه‌های نویزی ممکن است به طور نامنظم در فضای داده توزیع شده باشند.

۵. تشخیص ناهنجاری‌های پیچیده: بیشتر روش‌های موجود، برای تشخیص ناهنجاری‌های نقطه‌ای هستند که نمی‌توانند برای ناهنجاری شرطی و ناهنجاری گروهی استفاده شوند زیرا رفتارهای کاملاً متفاوتی با ناهنجاری‌های نقطه‌ای از خود نشان می‌دهند. یکی از چالش‌های اصلی در اینجا گنجاندن مفهوم ناهنجاری‌های مجموعه‌ای در مدل‌های ناهنجاری است. همچنین، روش‌های کنونی عمدتاً بر تشخیص ناهنجاری‌ها از یک منبع داده تمرکز می‌کنند، در حالی که بسیاری از کاربردها نیاز به تشخیص ناهنجاری‌ها با منابع داده ناهمگن چندگانه، به عنوان مثال، داده‌های چند بعدی، نمودار، تصویر، متن و داده‌های صوتی دارند. یکی از چالش‌های اصلی این است که برخی از ناهنجاری‌ها را می‌توان تنها با در نظر گرفتن دو یا چند منبع داده شناسایی کرد.

۶. قابل تعریف بود ناهنجاری:

در بسیاری از حوزه‌های حیاتی و ایمنی، ممکن است موارد عمدتی وجود داشته باشند که اگر مدل‌های تشخیص ناهنجاری مستقیماً به عنوان مدل‌های جعبه سیاه استفاده شوند، این خود خطراتی را به همراه خواهد داشت. برای مثال، به دلیل وجود موارد نادر که به عنوان ناهنجاری گزارش شده‌اند، ممکن است منجر به سوگیری الگوریتمی احتمالی علیه گروه‌های اقلیت موجود در داده‌ها شود، مانند گروه‌هایی که کمتر در سیستم‌های کشف تقلب و کشف جرم ارائه شده‌اند. یک رویکرد موثر برای کاهش این مشکل، داشتن الگوریتم‌های توضیح ناهنجاری است که سرنخ‌های ساده‌ای در مورد اینکه چرا یک نمونه داده خاص به عنوان ناهنجاری شناسایی می‌شود، ارائه می‌دهد. این الگوریتم به متخصصین و عوامل انسانی اجازه می‌دهد حساسیت مدل را برای این دادگان بررسی کرده و عملکرد آن را تصحیح کنند. ارائه چنین توضیحی می‌تواند به اندازه دقیق تر تشخیص در برخی کاربردها مهم باشد. با این حال، اکثر مطالعات تشخیص ناهنجاری تنها بر دقت تشخیص تمرکز می‌کند و توانایی ارائه توضیح ناهنجاری‌های شناسایی شده را نادیده می‌گیرند. استخراج توضیح ناهنجاری از روش‌های تشخیص خاص هنوز یک چالش است که حد زیادی حل نشده باقیمانده است. توسعه مدل‌های تشخیص ناهنجاری ذاتاً قابل تفسیر نیز بسیار مهم است، اما همچنان یک چالش اصلی برای ایجاد تعادل بین تفسیرپذیری و اثربخشی مدل است.

۵.۱ ساختار کلی روش‌های تشخیص ناهنجاری

اگر بخواهیم روش‌های تشخیص ناهنجاری را به صورت عمومی توصیف کنیم، می‌توانیم بگوییم که این روش‌ها عموماً دارای سه مرحله اصلی هستند. مرحله اول یادگیری بازنمایی داده‌ها^{۱۲} است. در این مرحله نگاشتی از دادگان ورودی به فضای معین آموخته می‌شود. این نگاشت را می‌توان به صورت تابعی مانند زیر تعریف کرد.

$$f(\cdot; \theta) : x \rightarrow y \quad (1.1)$$

در مرحله دوم به تعریف یک معیار سنجش برای ارزیابی خروجی مرحله قبل پرداخته می‌شود. این معیار که به صورت یک تابع بیان می‌شود با دریافت خروجی مرحله قبلی عددی را به عنوان یک امتیاز برای سنجش میزان تعلق داده ورودی به

¹²Data representation

دسته ناهنجار اختصاص می‌دهد که به آن امتیاز ناهنجاری ^{۱۳} گوییم.

$$d(f(x); \eta) : f(x) \rightarrow d, \quad d \in \mathbb{R} \quad (2.1)$$

در آخر نیز با درنظر گرفتن یک مقدار آستانه δ ، به تصمیم‌گیری در مورد داده ورودی با توجه به امتیاز اختصاص داده شده در مرحله دوم پرداخته می‌شود.

$$g(d(f(x))) = \begin{cases} \text{anomaly} & d \geq \delta \\ \text{not anomaly} & d < \delta \end{cases} \quad (3.1)$$

با توجه به این تعریف، رویکردهای موجود می‌توانند انواع زیر را داشته باشند:

۱. غیر پارامتری: نیازی به یادگیری θ و η و δ نیست.

۲. یک مرحله‌ای: تنها یکی از مجموعه پارامترهای موجود θ یا η یادگرفته می‌شوند.

۳. دو مرحله‌ای: هر دو مجموعه پارامتر θ و η به صورت مستقل و جداگانه یادگرفته می‌شوند.

۴. ادغامی^{۱۴}: هر دو مجموعه پارامتر θ و η باهم یادگرفته می‌شوند.

در صورت عدم وجود برچسب‌های دادگان موجود، ناچار به استفاده از روش بدون ناظر هستیم که در آن هیچ گونه اطلاعاتی در مورد ماهیت دادگان در دسترس نیست. در این موقع معمولاً δ از پیش تعریف شده است و یا همراه با η یادگرفته می‌شود.

در حالتی که تنها بخشی از دادگان برچسب خورده باشند و باقی برچسب نخورده، می‌توان از رویکرد یادگیری با نظارت ضعیف استفاده کرد. در این مورد نیز مقدار آستانه می‌تواند با استفاده از تنظیم دقیق مدل بدست آید.

۶.۱ ساختار گزارش

در فصل اول این سمینار به معرفی حوزه سمینار و تعریف مسئله تشخیص ناهنجاری و کاربردهای آن در حوزه‌های مختلف پرداخته شد. فصل دوم با بررسی روش‌های سنتی و عمیق مورد استفاده در مقالات روز و معرفی کارهای مرتبط با این سمینار به بررسی جزئی از روش‌ها و مقالات موجود چاپ شده در سال‌های اخیر خواهد پرداخت. در نهایت، فصل آخر، مسائل باز و کارهای آینده این حوزه معرفی شده و یک موضوع پیشنهادی برای پژوهه نهایی مطرح می‌شود.

¹³Anomaly score

¹⁴Integrated

۲ فصل

مروری بر کارهای انجام شده برای تشخیص ناهنجاری

برای درک بهتر مقالات و پژوهش‌های انجام شده با موضوع تشخیص ناهنجاری، خوب است ابتدا مروری بر ساختارهای کلی موجود که در روش‌های تشخیص ناهنجاری استفاده می‌شوند داشته باشیم. اینگونه روشها پایه و اساس بسیاری از مقالات روز هستند و شناختن مدل و نحوه کارکرد مدل ما را در درک بهتر ایده نویسندگان مقالات و هدف از استفاده از این روش‌ها در کارهای انجام شده کمک می‌کند. مطالعه روش‌های پایه کمک می‌کند تا کارهای انجام شده اخیر بهتر درک شوند.

این فصل شامل دو بخش اصلی است که در بخش اول به معرفی روش‌های سنتی پرداخته می‌شود. هدف از آوردن روش‌های سنتی آشنایی پایه‌ای با مسئله تشخیص ناهنجاری است. همچنین این روش‌ها می‌توانند با ترکیب شدن با روش‌های عمیق، مدل‌های جدیدی را بسازند و همچنین ایده‌ای برای وجود آوردن مدل‌های عمیق دیگر باشند. در بخش دوم ساختارهای عمیق مورد استفاده در تشخیص ناهنجاری آورده شده‌اند که به همراه آنها مثال‌هایی از کارهای انجام شده جدید نیز بررسی می‌شوند.

۱.۲ مروری بر روش‌های سنتی

اگر به یاد داشته باشید، در ابتدای فصل یک به این نکته اشاره شد که مسئله تشخیص ناهنجاری، یک موضوع فعال تحقیق در چند دهه اخیر است که یکی از مقالات معتبر چاپ شده آن مربوط به دهه ۱۹۶۰ میلادی می‌شود. از این رو، در طی این مدت بسیاری از روش‌ها برای یافتن دادگان خارج از محدوده معرفی و توسعه داده شده‌اند که از یادگیری عمیق استفاده نمی‌کنند. این روش‌ها به صورت عمده دادگان را مجموعه‌ای از نقاط در یک فضای چند بعدی فرض می‌کنند و تلاش آنها بر این است که نقاط خارج از محدوده را با توجه به ویژگی‌ها و مشخصات نقاط دیگر آشکار کنند. عمدتاً این اینگونه روش‌ها را می‌توان از نقطه‌نظر ایده اصلی به سه دسته کلی تقسیم کرد که عبارت‌اند از: استفاده از رده‌بندی، مبتنی بر معیار فاصله و استفاده از مدل‌های آماری (جدول ۱.۲). در ادامه به مرور کلی این روش‌ها خواهیم پرداخت. با توجه به اینکه تمرکز ما بر بررسی کامل این روش‌ها نیست پیشنهاد می‌شود برای آشنایی بیشتر با این‌گونه روش‌ها به مقاله چاندولا و همکاران مراجعه کنید [۱].

جدول ۱.۲: دسته‌بندی روش‌های سنتی

دسته‌بندی روش‌های سنتی در تشخیص ناهنجاری			
روش‌های شناخته شده	آنواع	خلاصه ایده	رویکرد
One-class SVM SVDD	یک کلاسه چند کلاسه	یادگیری یک مرز تفکیک میان دادگان عادی و ناهنجار	رده‌بندی
LOC ^۱ COF	فاصله تا نزدیک ترین همسایه	اقدام به تعریف یک معیار فاصله می‌کند تا دادگان عادی را از دادگان ناهنجار جدا کند	
K-means CBLOF	خوش‌بندی و سنجش فاصله تا نزدیک ترین خوش	مدل آماری	
PCA Isolation Forest	استفاده از تصویر سازی نقاط در فضایی با بعد کمتر		
Gaussian Mixture Model	روش‌های پارامتری	دادگان عادی در نواحی پر احتمال مدل آماری قرار می‌گیرند	مدل آماری
Kernel destiny estimator	روش‌های غیر پارامتری		

۱.۱.۲ روش‌های مبتنی بر رده‌بندی

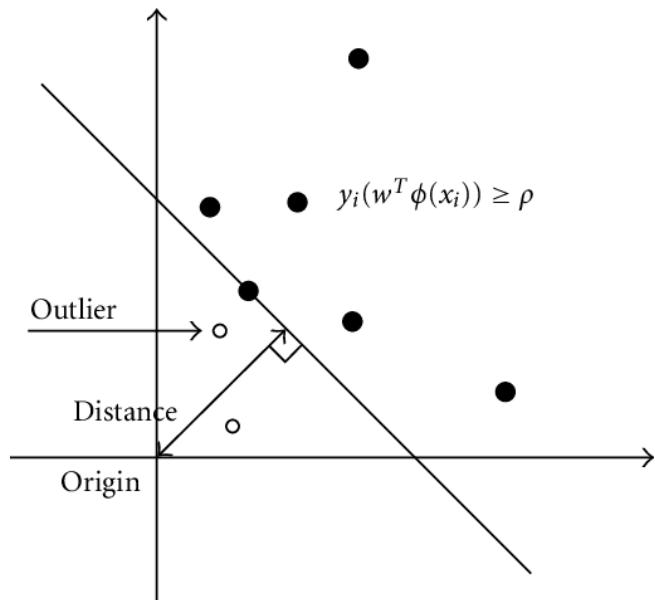
همانطور که در ابتدای این بخش گفته شد، یکی از ایده‌های کلی در روش‌های مورد استفاده برای تشخیص ناهنجاری استفاده از ایده رده‌بندی است. در اینگونه روش‌ها تلاش می‌شود یک مرز تفکیک میان دادگان عادی و دادگان ناهنجار رسم شود. اگر چنین مرزی وجود داشته باشد، می‌توان با استفاده از الگوریتم‌های رده‌بند موجود اقدام به یافتن این مرز و آشکارسازی داده‌های ناهنجار کرد. همانطور که مشخص است در اینگونه روش‌ها تنها یک دسته برای دادگان تعریف می‌شود که آن دسته دادگان عادی است. دیگر دادگانی که در این دسته قرار نمی‌گیرند به عنوان دادگان عادی در نظر گرفته می‌شوند. البته استفاده از رویکرد رده‌بندی چند کلاسه نیز در صورت وجود برچسب برای تمامی دادگان امکان پذیر است اما استفاده از این روش کمتر مرسوم است. یکی از معروف ترین روش‌های مورد استفاده دسته بند، ماشین بردار پشتیبان یک کلاسه^۲ است. در روش ماشین بردار پشتیبان که در یک روش معروف رده‌بندی است تلاش می‌شود دادگان دو دسته موجود توسط یک صفحه از یکدیگر جدا شوند. در الگوریتم بردار پشتیبان یک کلاسه سعی می‌شود صفحه جدا کننده را طوری مشخص کند تا دادگان معمول در یک طرف این صفحه و دادگان ناهنجار در سمت دیگر آن قرار گیرند. همچنین تلاش می‌شود صفحه مورد نظر تا حد امکان به نقاط داده عادی نزدیک باشد. پس از رسم این صفحه، دادگانی که به مبدأ مختصات نزدیک تر هستند در دسته ناهنجاری‌ها قرار می‌گیرند.^۳

در اینجا تابع نگاشتی که باید یاد گرفته شود همان تابع کرنل در ماشین بردار پشتیبان است و تابع امتیاز ناهنجاری نیز به صورت اندازه فاصله از مبدأ مختصات تعریف می‌شود. شکل ۱.۲ این روش را به تصویر کشیده است. توجه داشته باشید که در اینجا تنها یک دسته برای رده‌بندی تعریف می‌شود که آن دسته دادگان عادی است، پس نیازی به وجود برچسب برای تمامی دادگان نیست و این رویکرد به صورت کاملاً بدون ناظر خواهد بود.

نمونه دیگری از روش‌های مورد استفاده برای آشکارسازی ناهنجاری که از رویکرد رده‌بندی استفاده می‌کند، بردار پشتیبان توصیف‌گر داده^۴ است. در این روش سعی می‌شود کره‌ای با کوچک ترین اندازه شعاع ممکن حول دادگان موجود رسم شود.

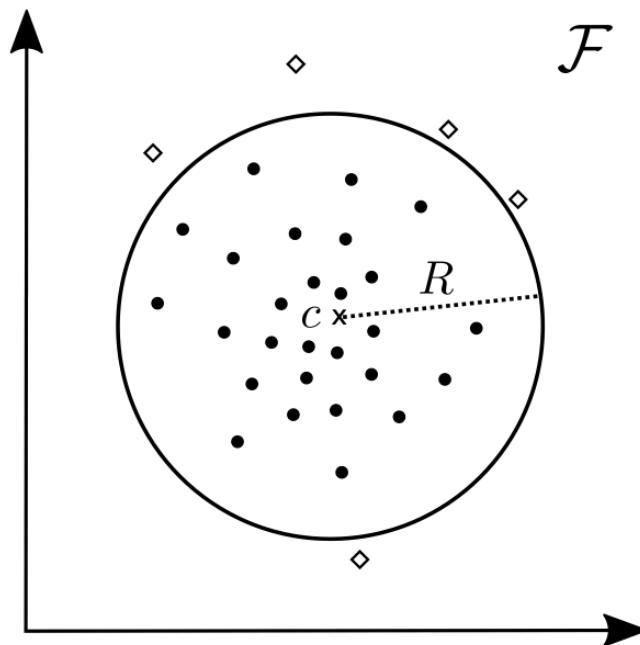
²One-class SVM

³Support Vector Data Description (SVDD)



شکل ۱.۲: ماشین بردار پشتیبان یک کلاسه

پس از رسم این کره، دادگانی که در خارج از آن قرار می‌گیرند به عنوان داده ناهمجارت شناخته خواهند شد [۵].

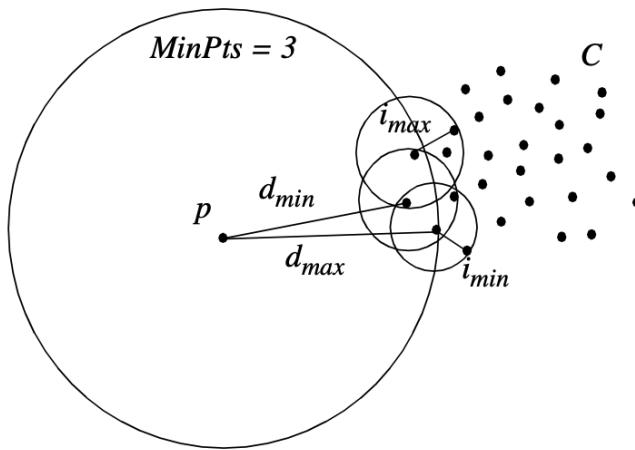


شکل ۲.۲: بردار پشتیبان توصیفگر داده [۵]

از جمله مزیت‌های این رویکرد، آموزش سریع، و دقیق بیشتر آن در موقعی است که دادگان برچسب خورده در اختیار هستند. و از معایب این روش در هنگام استفاده از رده‌بندی چند کلاسه می‌توان به نیاز برای چندین دسته داده عادی یاد کرد. همچنین این رویکردها نیاز به مشخص کردن ابرپارامتر برای مدل یادگیری دارند.

۲.۱.۲ روش‌های مبتنی بر معیار فاصله

اگر به دادگان موجود را به صورت نقاطی بازنمایی شده بر روی صفحه مختصات نگاه کنیم، می‌توانیم از معیار فاصله نقاط از یکدیگر به تصمیم‌گیری در مورد دادگان بپردازیم. در اینگونه رویکردها معمولاً اقدام به تعریف یک معیار فاصله می‌کنند تا دادگان عادی را از دادگان ناهنجار جدا کنند. یک نمونه روش معروف که در این دسته می‌گنجد روش معروف جداساز پرت محلی^۴ است. در این روش میانگین فاصله هر نقطه از همسایگان محلی محاسبه شده و اگر این میانگین از یک مقدار آستانه بیشتر باشد، داده به عنوان داده ناهنجار شناخته می‌شود. برای سادگی کار، میانگین فاصله نقطه تا تمام همسایگان را بر میانگین فاصله میان همسایگان نقطه محاسبه شده و مقدار آستانه برابر با عدد یک درنظر گرفته می‌شود [۶]. در استفاده از این روش نیز نیازی به وجود برچسب دادگان نیست همچنین این روش پارامتری برای یادگیری ندارد و در دسته روش‌های بدون پارامتر نیز قرار می‌گیرد. در واقع این گونه روش‌ها معمولاً به صورت بدون ناظر هستند.



شکل ۳.۲: نمایش کلی روش جداساز پرت محلی [۶]

۳.۱.۲ روش‌های مبتنی بر مدل آماری

ایده اصلی در این دسته از رویکردها بدین صورت است که، دادگان عادی همواره احتمال رخداد بالایی دارند، در نتیجه در نواحی از مدل آماری قرار می‌گیرند که احتمال وقوع آنها بیشتر است. برای مثال در روش مدل خطی پویا^۵ ابتدا دادگان از فضای ورودی به یک فضای از پیش تعیین شده نگاشت می‌شوند. سپس با استفاده از مدل بدست آمده سعی می‌شود مقدار ورودی را با توجه به دیگر دادگان موجود پیش‌بینی کند. در اینجا امتیاز ناهنجاری میزان تفاوت مقدار پیش‌بینی شده و مقدار حقیقی داده است. اگر مقدار اختلاف از یک مقدار آستانه از پیش تعیین شده، که با استفاده از آزمایش با دادگان برچسب خورده بدست آمده، بیشتر باشد، داده ورودی به دسته دادگان ناهنجار تعلق می‌گیرد.

۲.۲ استفاده از یادگیری عمیق برای تشخیص ناهنجاری

در بخش قبل، مروری مختصر و کلی بر روی روش‌های سنتی و برای درک بهتر مسئله تشخیص ناهنجار انجام شد. در این بخش به معرفی مدل‌های یادگیری عمیق پر استفاده در اینگونه مسائل پرداخته می‌شود که پایه و اساس خیلی از روش‌های

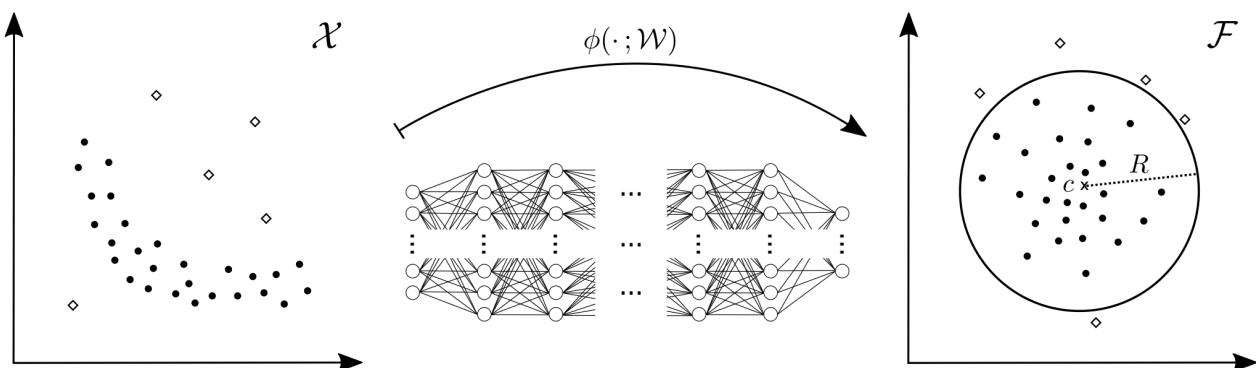
⁴Local Outlier Factor

⁵Dynamic linear model

ارائه شده در مقالات هستند و آشنایی با آنها به درک بهتر مطلب کمک بسیار زیادی خواهد کرد. پس از معرفی ساختار مورد بحث نمونه‌هایی از کارهای انجام شده که از آن استفاده می‌کنند را به اختصار معرفی خواهیم کرد.

۱.۲.۲ استفاده از یادگیری عمیق برای یادگیری بازنمایی دادگان

یکی از ابتدایی ترین ایده‌هایی که در مورد استفاده از روش‌های سنتی موجود با توجه به معرفی و پیشرفت ساختارهای عمیق به ذهن می‌رسد، استفاده از این ساختارها به منظور استخراج ویژگی از دادگان با ابعاد بالا و نه لزوماً تفکیک پذیر خطی است. ساختارهای عمیق با توجه به قابلیت بالای یادگیری ترکیب‌های غیر خطی گوناگون، می‌توانند به عنوانتابع نگاشت دادگان در روش‌های سنتی استفاده شوند تا بتوانند بازنمایی بسیار بهتری از دادگان را برای انجام عملیات امتیازدهی و تشخیص ناهنجاری بدست آورند. مدل‌های عمیق بسیاری برای استخراج ویژگی‌ها در طول زمان برای انواع مختلف دادگان معرفی شده‌اند که می‌توانند برای این منظور استفاده بشوند (AlexNet ، VGG و ...). پس از بدست آمدن ویژگی‌ها در فضای جدید، یکتابع امتیاز دهی به دادگان اعمال می‌شود تا امتیاز ناهنجاری بدست آید و با توجه به آن عمل تشخیص ناهنجاری صورت گرفته شود. در این مورد تابع امتیاز ناهنجار می‌تواند کاملاً مستقل باشد و در فرآیند آموزش مدل عمیق برای استخراج ویژگی‌ها نقشی نداشته باشد. برای مثال در روش بردار پشتیبان توصیفگر داده که در فصل دوم معرفی شد می‌توان بجای تابع $f(\theta)$ که مسئول نگاشت دادگان به فضایی معین است، از یک شبکه عمیق مانند مدل پرسپترون چندلایه استفاده کرد. این مدل به دلیل توانایی یادگیری نگاشت غیر خطی دادگان می‌تواند بازنمایی بهتری از دادگان را برای مرحله دوم محاسبات، که همان عمل امتیاز دهی به نقاط است، بدست آورد. راف و همکاران با استفاده از این ایده، روش بردار پشتیبان توصیف‌گر داده عمیق را معرفی کردند که در مقایسه با روش‌های سنتی عملکرد بسیار بهتری را از خود نشان داده است [۵].



شکل ۴.۲: بردار پشتیبان توصیفگر داده عمیق [۵]

۲.۲.۲ خود کدگذار

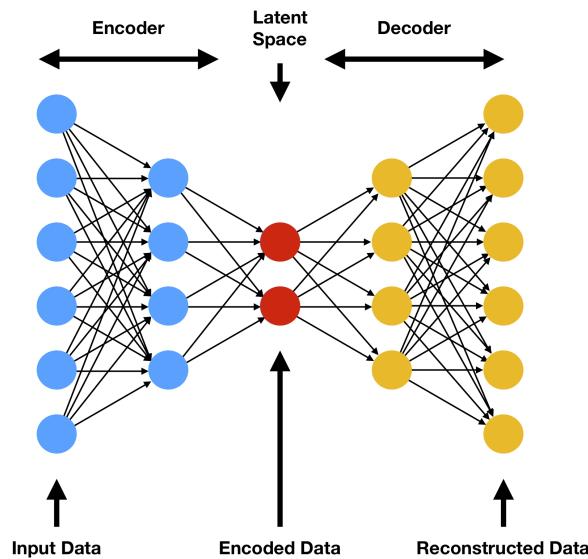
خود کدگذار^۶‌ها نوعی از شبکه‌های عصبی هستند که از روش پس انتشار^۷ برای یادگیری ویژگی‌های مفهومی استفاده می‌کنند. این شبکه‌ها به صورت دو مرحله‌ای اقدام به یادگیری می‌کنند که به ترتیب کدگذاری و کدگشایی نام دارند. در مرحله اول داده ورودی به شبکه کدگذار داده می‌شود و آن داده ورودی را به یک فضای با ابعاد پایین نگاشت می‌کند. به این فضای اصطلاح فضای باقی‌مانده^۸ یا فضای z می‌گویند. در مرحله دوم، بازنمایی بدست آمده وارد شبکه کدگشا شده تا داده از فضای

⁶AutoEncoder

⁷Backpropagation

⁸Latent space

باقی مانده دوباره به فضای ورودی باز گردانده شود. آنچه انتظار می‌رود آن است که خروجی مدل با ورودی مدل بسیار تشابه داشته باشد. در این صورت قسمت کدگذار توانسته بازنمایی خوبی از داده را در فضای باقی مانده ایجاد کند [۲۳].



شکل ۵.۲: مدل خودکدگذار

اگر بخواهیم کارکرد مدل شکل ۵.۲ را با فرمول ریاضی توصیف کنیم، با در نظر گرفتن داده X به عنوان ورودی مدل، کدگذار پس از دریافت این ورودی، آن را به فضای باقی مانده و به نقطه‌ای مانند z نگاشت می‌کند. اگر تابع کدگذار را f بنامیم معادله مرحله اول به صورت زیر خواهد بود.

$$f(X, \theta_1) : X \rightarrow z \quad (1.2)$$

که در اینجا ابعاد فضای z از ابعاد فضای ورودی X کمتر است. این بدان معنی است که در اینجا عمل کاهش ابعاد ورودی صورت گرفته است. اگر کدگشا را مانند تابعی درنظر بگیریم و آن را g بنامیم، این تابع با دریافت ورودی z ، اقدام به بازسازی داده ورودی می‌کند.

$$g(z, \theta_2) : z \rightarrow \hat{X} \quad (2.2)$$

در کاربردهای تشخیص ناهنجاری معمولاً در هنگام استفاده از این معماری، سعی می‌شود از تابع خطای مقایسه ورودی و خروجی مدل برای آموزش مدل استفاده کنند و در فرایند آموزش تنها از دادگان عادی استفاده شود. ایده اصلی در این گونه روش‌ها این است که با توجه به اینکه مدل تنها با دادگان عادی آموزش دیده است، دادگانی که توسط این مدل نتوانند به خوبی بازسازی شوند دارای ناهنجاری بوده‌اند. در واقع در اینجا تابع خطای همان تابع امتیاز ناهنجاری است به صورت زیر تعریف می‌شود.

$$L(X, g(f(x))) = d \quad (3.2)$$

پس از آموزش مدل مقدار آستانه d برای بدست آوردن بهترین نتیجه با آزمون و خطای روش‌های دیگر مانند استفاده از نمودار حساسیت و دقت تعیین می‌شود.

خودکدارها باید به تغییرات دادگان ورودی حساس باشند تا بتوانند با دقت مطلوب داده کد شده را بازسازی کنند. همچنین این حساسیت نباید به اندازه‌ای باشد که باعث شود مدل بجای یادگیری عملکرد مناسب، به بخار سپاری دادگان ورودی بپردازد و دچار بیشبرازش^۹ بشود. برای دستیابی به چنین توازنی، انواع مختلفی از خودکدارها معرفی شده‌اند که با افروzen یک مقدار تنظیم کننده^{۱۰} به تابع خطای اصلی معرفی شده، بدست می‌آیند.

$$L(X, g(f(X))) + \text{regularizer} \quad (4.2)$$

از ویژگی‌های تولید شده در لایه میانی توسط کدگذار نیز می‌توان به عنوان ویژگی‌های مناسب و با ابعاد پایین‌تر استفاده کرد و با بهره گیری از ایده روش‌های سنتی ماشین بردار پشتیبان، اقدام به رده‌بندی یک کلاسه برای تشخیص ناهنجاری‌ها کرد [۲۴].

خودکدار خلوت (SAE)

خودکدار خلوت^{۱۱} یکی از انواع خودکدارها است. ایده اصلی آن این است که، با توجه به اینکه ممکن است تعداد نورون‌های لایه مخفی به اندازه کافی نباشد، مفاهیم پیچیده شاید به خوبی توسط مدل یادگرفته نشوند. در نتیجه پیشنهاد می‌شود در لایه مخفی تعداد نورون‌های بیشتری قرار گیرند اما در تابع فعال سازی ترتیبی داده شود تا این نورون‌ها تاحد ممکن کم استفاده شوند و یا به اصطلاح، به صورت خلوت^{۱۲} فعال سازی شوند. برای دستیابی به چنین هدفی می‌توان از تنظیم کننده^{۱۳} در تابع خطای مدل استفاده کرد. نوع اول استفاده از تنظیم کننده نرم یک است^{۱۴} که معادله تابع خطای صورت زیر خواهد بود.

$$L(X, g(f(X))) + \lambda \sum_i^n |a^{(h)}| \quad (5.2)$$

با استفاده از این نوع تنظیم کننده، چون تابع نرم یک استفاده شده، در طی فرایند یادگیری سعی می‌شود وزن یال‌های متصل به نورون‌ها تا حد امکان صفر باشند و با صفر شدن این وزن‌ها، درواقع نورون‌های کمتری در فرایند محاسبه استفاده می‌شوند.

خودکدار حذف نویز

نوع دیگری از خودکدار که می‌توان از آن برای حذف نویز در داده استفاده کرد را به اصطلاح خودکدار حذف نویز^{۱۵} نام دارد. تفاوت این مدل با حالت کلی خودکدارها در فرایند آموزش مدل است. در این مدل داده ورودی ابتدا با استفاده از یک تولید کننده نویز، نویزی می‌شود. سپس به خودکدار داده می‌شود. شبکه باید بتواند نویز اضافه شده به دادگان را حذف کند. برای انجام این کار یک راه ساده، تعریف تابع خطای صورت مقایسه خروجی مدل و ورودی اصلی بدون نویز است. شبکه باید تلاش کند تا اختلاف تصویر باز سازی شده و تصویر اصلی را به حداقل برساند. پس از آموزش این مدل، شبکه قادر خواهد بود تا هرگونه ناهنجاری در داده که در اینجا همان نویز موجود در دادگان است را حذف کند. با مقایسه مقدار خروجی و ورودی مکان‌هایی که تفاوت زیادی بایکدیگر دارند به احتمال زیاد متعلق به دسته ناهنجاری هستند.

⁹Overfit

¹⁰Rgulizer

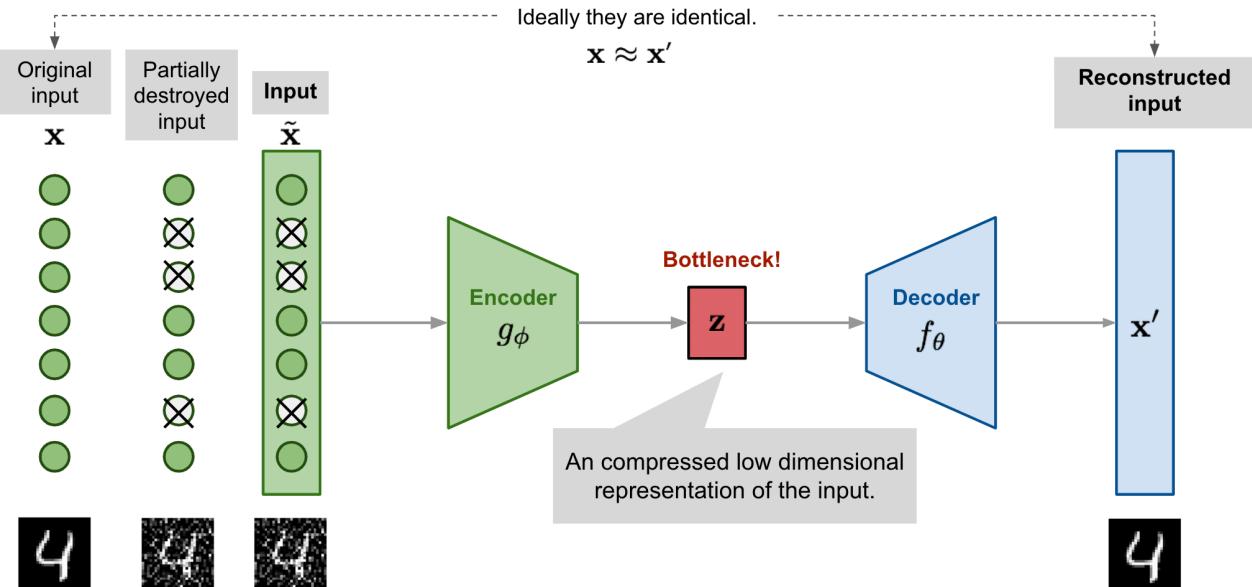
¹¹Sparse AutoEncoder (SAE)

¹²Sparse

¹³Regulizer

¹⁴L1-Rgulizer

¹⁵Denoising AutoEncoder



شکل ۶.۲: مدل خودکدگذار حذف نویز [۷].

RDA خودکدگذار

خودکدگذارهایی که تا کنون معرفی شدند، در مرحله آموزش مدل تنها از دادگان عادی و بدون ناهنجاری استفاده می‌کردند و دادگان ناهنجار تنها زمان آزمون استفاده می‌شدند. حال اگر بخواهیم دادگان ناهنجار را نیز در فرآیند آموزش مدل دخیل کنیم باید روش جدیدی را معرفی کنیم. خودکدگذار مقاوم^{۱۶} درواقع از ایده تجزیه و تحلیل مؤلفه بنیادی مقاوم^{۱۷} برگرفته شده است. در روش تجزیه و تحلیل مؤلفه بنیادی مقاوم، دادگان ورودی با استفاده از دو ماتریس مرتبه پایین^{۱۸} و خلوت^{۱۹} نمایش داده می‌شوند.

$$X = L + S \quad (6.2)$$

که در اینجا L نمایش داده ورودی در ابعاد پایین‌تر است و S قسمتی از دادگان است که نمی‌تواند توسط L به خوبی نمایش داده شود. این دوماتریس تحت شرط بهینه‌سازی و تابع هدف زیر آموزش داده می‌شوند^{۲۰}.

$$\|X - L - S\|_F^2 = 0 \quad (7.2)$$

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1 \quad (8.2)$$

این روش نیز سعی دارد دادگان ورودی را به استفاده از دو ماتریس نمایش دهد که ماتریس اول بازنمایی بدست آمده توسط خودکدگذار است و قسمت دوم نمایانگر ناهنجاری‌هایی است که نمی‌توانند توسط خودکدگذار به خوبی بازنمایی شوند.

$$X = L_D + S \quad (9.2)$$

¹⁶Robust Deep AutoEncoder

¹⁷Robust PCA

¹⁸Low rank

¹⁹Sparse

²⁰در اینجا $\|.\|_F$ نرم و $\|.\|_*$ جمع مقادیر یکتا (singular value) است.

اگر کدگذار و کدگشا را به عنوان دوتابع f و g در نظر بگیریم، معادل بهینه سازی مدل به صورت زیر خواهد بود.

$$\min_{\theta} \|L_D - G_{\theta}(F_{\theta}(L_D))\|_2 + \lambda \|S\|_1 \quad (10.2)$$

که شرایط زیر باید در فرایند بهینه سازی صدق کند:

$$X - L_D - S = 0 \quad (11.2)$$

فرایند امتیاز دهنده به ناهنجاری در این نوع خودکدگذار مشابه روش اصلی خواهد بود. در اینجا S در واقع همان ناهنجاری‌های موجود در دادگان هستند که پس از تکمیل فرایند آموزش می‌توانیم از آن استفاده کنیم. این روش در مقایسه با روش سنتی در کاربرد تشخیص ناهنجاری حدود ۷۰ درصد بهتر عمل کرده است [۲۵].

گاهی اوقات ممکن است فرض ما بر استفاده از خودکدگذارها برای تشخیص ناهنجاری درست نباشد. با توجه به قابلیت یادگیری بالای خودرمزنگذارهای عمیق، در برخی کاربردها ممکن است امکان بازسازی ناهنجاری‌ها نیز همانند دادگان عادی وجود داشته باشد که برای این مورد باید چاره‌ای اندیشه‌یده شود [۲۶].

۳.۲.۲ مدل‌های مولد خودکدگذار VAE

مشکل خودکدگذارهایی که تا کنون معرفی کردیم در این است که، نگاشت دادگان به فضای باقیمانده به صورت قطعی صورت می‌گیرد. در واقع، هر نقطه از فضای ورودی به یک نقطه معین از فضای باقیمانده نگاشته می‌شود. از طرف دیگر اگر یک نقطه را به صورت تصادفی در فضای باقیمانده، مانند \hat{z} را در نظر بگیریم، نمی‌توان به طور قطع گفت که این نقطه به کدام دسته از نقاط تعلق خواهد گرفت. در واقع خودکدگذارهایی که تا کنون مطالعه کردیم به خوبی دادگان ورودی را به فضایی با ابعاد دیگر نگاشت می‌کردند اما در هیچ یک از این روش‌ها اختیاری برای کنترل روند و نحوه این نگاشت نداشتند. انواع مختلف این خودکدگذارها نیز بسته به نیاز، نگاشتهای گوناگونی را در اختیار ما قرار می‌دادند تا مناسب کاربرد انتخاب شده باشد. برای اینکه در طی فرایند یادگیری بر روی نحوه نگاشت دادگان به فضای باقیمانده کنترل داشته باشیم، نوع دیگری از خودکدگذارها تحت عنوان خودکدگذار VAE^{۲۱} معرفی شده است [۲۷]. در این روش بجای یادگیری نگاشت گسسته و قطعی دادگان به فضای باقیمانده، سعی می‌شود توزیع دادگان در فضای باقیمانده یاد گرفته شود. در این صورت دادگان در این فضا توزیع پیوسته‌ای خواهد داشت و عمل درونیابی نقاط در این فضا کار راحت‌تری خواهد بود.

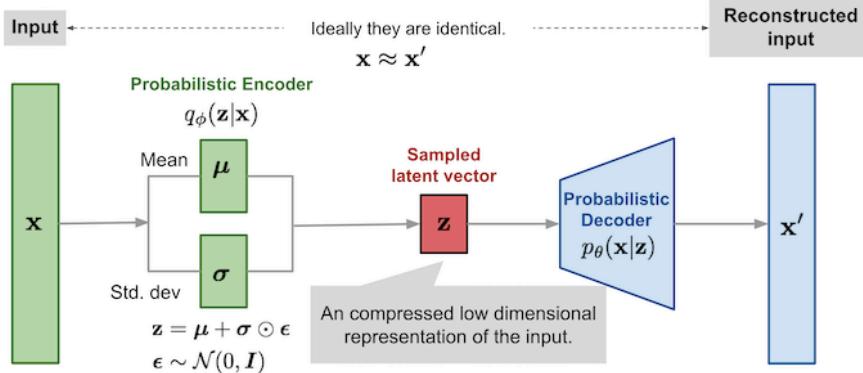
در این روش بجای تلاش برای کمینه کردن اختلاف ورودی مدل با خروجی بازسازی شده توسط مدل، سعی می‌شود تا احتمال درستنمایی نهایی^{۲۲} حداکثر شود. معادله تابع بهینه سازی این مدل به صورت زیر تعریف می‌شود.

$$\log(p(x)) \geq \log(p(x)) - KL(q_{\phi}(z|x) || p(z)) \quad (12.2)$$

$$\log(p(x)) - KL(q_{\phi}(z|x) || p(z)) = E_{z \sim q_{\phi}(x)} \log P_{\phi}(x|z) - KL(q_{\phi}(z|x) || p(z)) \quad (13.2)$$

²¹Variational AutoEncoder

²²Marginal likelihood



شکل ۷.۲: مدل خود رمز کننده [۷]. variational

$$\text{maximize } E_{z \sim q_\phi(x)} \log P_\phi(x|z) - KL(q_\phi(z|x) || p(z)) \quad (14.2)$$

در معادله (۱۴.۲) قسمت اول برای حداکثر کردن احتمال داده باز سازی شده است. قسمت دوم که در واقع می‌توان آن را به عنوان تنظیم کننده معادله در نظر گرفت، تلاش می‌کند تا توزیع دادگان در فضای بازنمایی \mathbb{Z} بسیار مشابه توزیع دادگان ورودی باشد. بنابراین بازنمایی دادگان در فضای باقیمانده بر خلاف مدل پایه به صورت غیر قطعی^{۲۳} خواهد بود. همچنانین دادگان بازسازی شده و امتیاز ناهنجاری برای دادگان نیز غیر قطعی و به صورت مقادیر احتمال خواهد بود. یک مثال خوب از کاربرد این گونه خودکدگذار برای تشخیص ناهنجاری، استفاده از خودکدگذار متغیر برای ترکیب ویژگی‌های بصری (تصویر) و ویژگی‌های متنی برای تشخیص اخبار جعلی بوده است [۸]. شکل ۸.۲ نحوه عملکرد این مدل را نشان می‌دهد.

شبکه‌های مولد رقابتی (GAN)

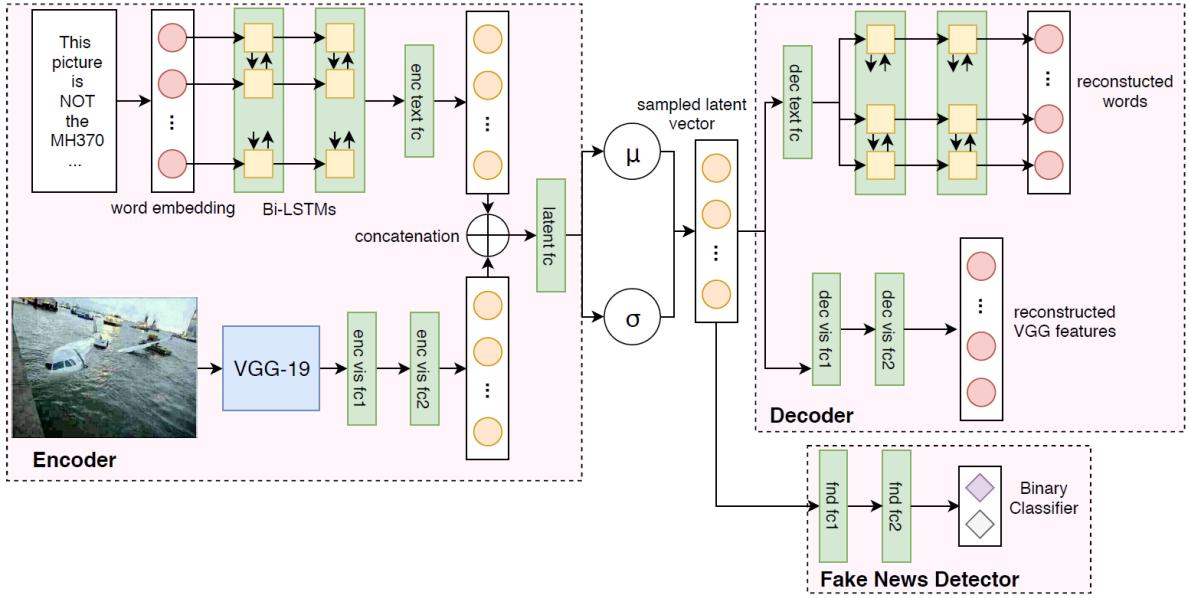
شبکه‌های مولد رقابتی^{۲۴} از دو قسمت اصلی تشکیل شده‌اند که به صورت رقابتی با یکدیگر آموزش می‌بینند. هر یک از این دو قسمت، سعی دارند عملکرد طرف مقابل را با بالا بردن کیفیت کار خود به چالش کشند. بخش اول این مدل که مولد^{۲۵} نام دارد، مسئولیت تولید داده مصنوعی را بر عهده دارد. این قسمت با گرفتن یک بردار ورودی از فضای باقیمانده، داده‌ای مصنوعی را تولید می‌کند. خروجی این قسمت به همراه یک نمونه از دادگان آموزش برای مقایسه و داوری، جهت تشخیص مصنوعی و یا حقیقی بودن، به تصمیم گیرنده^{۲۶} وارد می‌شوند. تصمیم گیرنده باید بتواند به داده حقیقی که از دادگان آموزش دریافت کرده است برچسب حقیقی و به داده تولید شده توسط بخش مولد برچسب مصنوعی بودن را اختصاص دهد. آموزش این مدل‌ها به صورت نوبتی صورت می‌گیرد و با شروع از بخش دوم، وزن‌ها در بخش دیگر ثابت می‌مانند و پس از چند مرحله که عملکرد این قسمت بهبود یافته، تغییر وزن‌ها در آن بخش متوقف شده و وزن‌های بخش دیگر آموزش می‌بینند. پس از بهبود عملکرد قسمت بعدی این چرخه ادامه پیدا می‌کند.تابع خطای مورد استفاده در مدل‌های مولد پایه به صورت زیر است.

²³Stochastic

²⁴Generative Adversarial Networks

²⁵Generator

²⁶Discriminator



شکل ۸.۲: مدل پیشنهادی برای ترکیب ویژگی‌های بصری و متنی برای تشخیص اخبار جعلی [۸].

$$\min_G \max_D V(D, G) = E_{X \sim p_{data}(x)}[\log D(x)] + E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (15.2)$$

برای استفاده از این مدل در تشخیص ناهنجاری، استفاده از تابع خطای تصمیم گیرنده به عنوان تابع امتیاز ناهنجاری می‌تواند مفید باشد. در اینصورت، تابع تصمیم گیرنده $D(X)$ وظیفه نگاشت دادگان به فضای تشخیص ناهنجاری را بر عهده دارد و تابع خطای این قسمت از مدل که به صورت زیر تعریف می‌شود به عنوان تابع امتیاز ناهنجاری بکار خواهد رفت.

$$d(x) = \log(1 - D(X)) \quad (16.2)$$

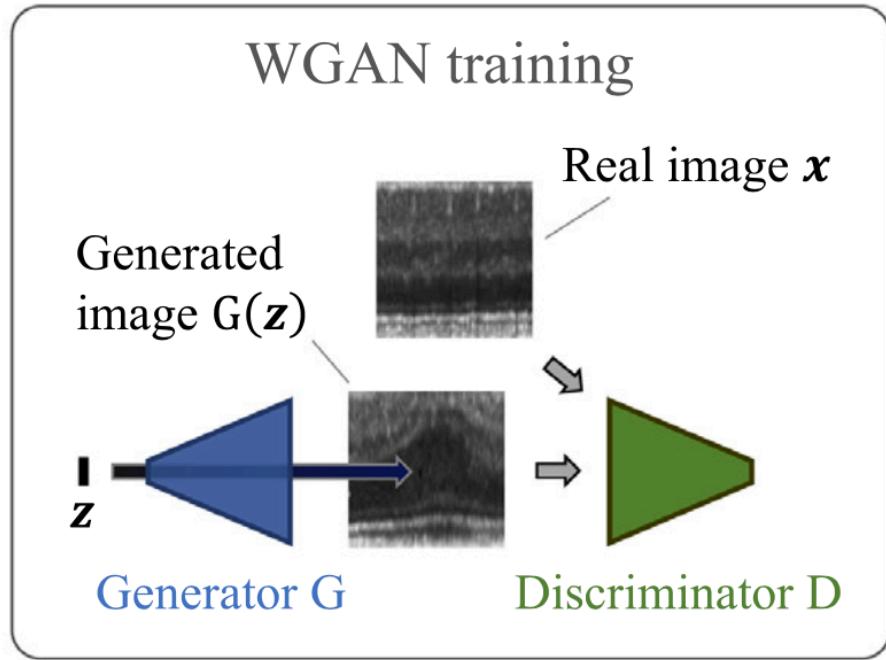
میزان آستانه تصمیم گیری δ نیز می‌تواند با استفاده از آزمون و خطا و یا با استفاده از منحنی حساسیت و دقت تعیین گردد.

برای اینکه بتوانیم از شبکه مولد نیز در این مسئله کمک بگیریم، می‌توانیم در فرایند آموزش دادگان، بجای انتخاب تصادفی یک نقطه از فضای z به عنوان ورودی شبکه مولد، با استفاده از یک رمز کننده دادگان ورودی را ابتدا به رمز کننده بدھیم تا بازنمایی دادگان در فضای باقیمانده بدست آید و سپس این داده را به عنوان ورودی به شبکه مولد بدھیم تا با این بازنمایی اقدام به تولید داده مصنوعی کند. چیزی که در اینجا توقع داریم این است که داده تولید شده توسط تابع مولد، بسیار شبیه به داده ورودی رمز کننده باشد. در این صورت تابع بهینه سازی مدل به صورت زیر خواهد بود.

$$\min_{\theta} \|G(E(X, \theta)) - X\| + \lambda \log(1 - D(G(E(X, \theta)))) \quad (17.2)$$

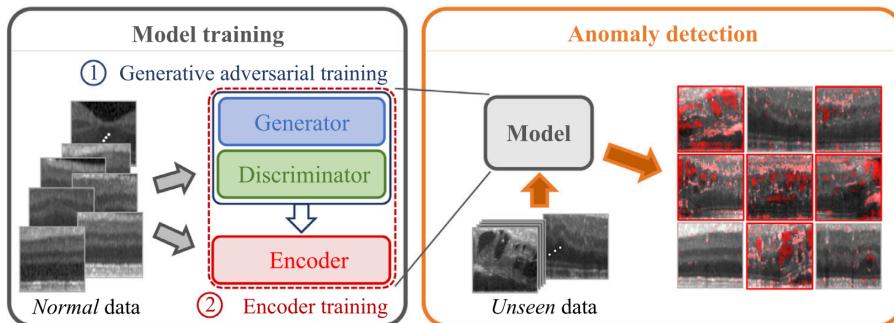
در این معادله پارامتر λ یک ابر پارامتر مدل است که به صورت دستی تعیین می‌شود. این روش در سال ۲۰۱۷ توسط چلگ و همکاران تحت عنوان AnoGan معرفی شد که برای آشکار سازی ناهنجاری‌ها در تصاویر گرفته شده از قرنیه چشم

²⁷OCT



شکل ۹.۲: شبکه مولد رقابتی

برای تشخیص بیماری و آسیب‌دیدگی مورد استفاده قرار گرفته است [۲۸]. مدل بهبود یافته این روش که با نام f-AnoGan توسط همین پژوهشگران معرفی شده است، با جای‌گذاری روند نگاشت دادگان به فضای باقی‌مانده با یک کدگذار از پیش آموزش دیده، سرعت محاسبات مدل را بهبود داده است [۹].



شکل ۹.۰.۲: نمایش نحوه آموزش مدل [۹] F-AnoGan

روش دیگری که از رویکرد شبکه‌های مولد استفاده می‌کند، GANomaly نام دارد. این روش با هدف تشخیص اشیاء ممنوعه در تصاویر گرفته شده با اشعه ایکس در خطوط کنترل وسایل توسعه داده شده است که عملکرد بهتری نسبت به روش AnoGan داشته است [۲۹].

۳.۲ مجموعه دادگان موجود برای تشخیص ناهنجاری

یکی از چالش‌های بر سر راه تولید و توسعه روش‌های تشخیص ناهنجاری عدم دسترسی به مجموعه دادگان دنیای واقعی و دربرگیرنده ناهنجاری‌های حقیقی است. در بسیاری از پژوهش‌های صورت گرفته، پژوهشگران برای ارزیابی و قابل مقایسه شدن کار خود با سایرین از مجموعه دادگان موجود و مورد استفاده در دسته‌بندی استفاده کرده اند [۲۹، ۲۰، ۳۱، ۳۰، ۵، ۳۲، ۳۳، ۲۵].

در این صورت امکان دارد کارایی روش معرفی شده در کاربردهای دنیای واقعی به خوبی مشخص نشود. از این رو در جدول ۲.۲ تعدادی مجموعه داده که دارای ناهنجاری‌های واقعی هستند جمع آوری شده تا بتواند مرجع خوبی برای استفاده در ارزیابی روش‌های معرفی شونده در کارهای آینده و پژوهش‌های دیگر باشد (برای مشاهده لیست کامل و در حال به روز رسانی مجموعه‌های داده موجود برای تشخیص ناهنجاری می‌توانید به آدرس <https://git.io/JTs93> مراجعه کنید) [۳۴].

جدول ۲.۲: مجموعه دادگان در دسترس برای تشخیص ناهنجاری

نام	حوزه کاربرد	اندازه	ابعاد	درصد ناهنجاری	جنس دادگان	مقالات مرجع
UCF-Crime	نظرات ویدیویی	۸.۱۳ میلیون فریم ویدیو	۱۳ دسته ناهنجار	ناشناخته	ویدیو	[۳۵]
HyperKvasir	تشخیص بیماری	درحال جمع آوری	درحال ناهنجار	ناشناخته	تصویر و ویدیو	[۳۷، ۳۶]
KDD Cup 99	تشخیص نفوذ	۶۵۷۴۹۷-۴۰۹۱	۷۷٪ درصد	۴۱	جدول	[۳۱، ۳۸]
UNSW-NB15	تشخیص نفوذ	۲۵۷۶۷۳	۹/۱۷ درصد	۴۹	جدول	[۱۲]
Webspam	آشکارسازی هرزتامه	۳۵۰۰۰۰	۳۶۶۱ درصد	۱۶/۶ میلیون	متن و جدول	[۳۹]
MVTec AD	آشکارسازی نقص در محصول	۵۳۵۴	۳۵۲۶ درصد	نامعلوم	تصویر	[۴۰]
ShanghaiTech Campus	نظرات ویدیویی	۳۱۷۳۹۸ فریم	۳۵۲۶ درصد	نامعلوم	ویدیو	[۴۱]

جدول ۳.۲: الگوریتم‌های عمیق مورد استفاده در تشخیص ناهنجاری

نام روش	مقاله مرجع	نحوه آموزش	رویکرد	معماری	تعداد لایه‌ها	نوع داده
RDA	[۲۵]	نیمه ناظارت شده	بازسازی	خودکدگزار	۳	ویدیو
AnoGAN	[۲۸]	نیمه ناظارت شده	بازسازی	مدل مولد	۴	تصویر
f-AnoGan	[۹]	نیمه ناظارت شده	بازسازی	شبکه کانولوشنی و خود رمزگذار	۴	تصویر
LSA	[۴۲]	نیمه ناظارت شده	پیش‌بینی	شبکه کانولوشنی	۷-۴	ویدیو
FFP	[۴۳]	نیمه ناظارت شده	پیش‌بینی	شبکه کانولوشنی	۱۰	ویدیو
EBGAN	[۳۳]	نیمه ناظارت شده	مدل مولد	شبکه کانولوشنی و پرسپترون چند لایه	۴-۳	تصویر
GT	[۳۰]	نیمه ناظارت شده	رده بندی	شبکه کانولوشنی	۱۶-۱۰	تصویر
E3Outlier	[۴۴]	نیمه ناظارت شده	رده بندی	شبکه کانولوشنی	۱۰	تصویر
REPEN	[۲۵]	بدون ناظر	معیار فاصله	پرسپترون چند لایه	۱	جدول
RDP	[۴۵]	بدون ناظر	معیار فاصله	پرسپترون چند لایه	۱	جدول
AE-OSVM	[۲۴]	بدون ناظر	رده بندی	خود رمزگذار و شبکه کانولوشنی	۵-۲	تصویر و جدول
Deep SVDD	[۵]	نیمه ناظارت شده	رده بندی	شبکه کانولوشنی	۴-۳	تصویر
Deep SAD	[۴۶]	نیمه ناظارت شده	رده بندی	شبکه کانولوشنی و پرسپترون چند لایه	۴-۳	تصویر و جدول
DAGMM	[۴۷]	بدون ناظر	خوشه بندی	خودکدگزار و پرسپترون چند لایه	۶-۴	جدول
MIL	[۴۸]	ناظارت ضعیف	شبکه کانولوشنی سه بعدی و پرسپترون چند لایه	شبکه کانولوشنی سه بعدی و پرسپترون چند لایه	۱۸ یا ۳۴-۳	ویدیو
DevNet	[۱۲]	ناظارت ضعیف	پرسپترون چند لایه	شبکه کانولوشنی سه بعدی و پرسپترون چند لایه	۴-۲	جدول
ALOCC	[۴۹]	نیمه ناظارت شده	خود رمزگذار و شبکه کانولوشنی	خود رمزگذار و شبکه کانولوشنی	۵	تصویر
OCAN	[۵۰]	نیمه ناظارت شده	LSTM-AE	استفاده از تابع امتیاز ناهنجاری	۴	داده دنباله‌ای
FenceGAN	[۳]	نیمه ناظارت شده	شبکه کانولوشنی و پرسپترون چند لایه	شبکه کانولوشنی و پرسپترون چند لایه	۵-۴	تصویر و جدول
OCGAN	[۵۱]	نیمه ناظارت شده	شبکه کانولوشنی	شبکه کانولوشنی	۳	تصویر

فصل ۳

کارهای آینده

در این سمینار ما با تعریف ناهنجاری و مسئله تشخیص ناهنجاری آشنا شدیم، نمونه‌هایی از کاربردهای وسیع این حوزه را معرفی کردیم که نشان دهنده اهمیت این موضوع بود. در ادامه چالش‌های موجود را شرح دادیم و نحوه عملکرد اینگونه روش‌ها را به صورت یک مدل عمومی ریاضی بیان کردیم. در فصل دوم نیز به مرور کارهای انجام شده پرداختیم تا با نمونه‌هایی از روش‌های موجود و ایده‌های اصلی این حوزه بیشتر آشنا شویم. در این فصل به معرفی موضوعات باز و کارهای قابل انجام خواهیم پرداخت و موضوعاتی را معرفی خواهیم کرد که در آینده می‌توانند مورد بررسی قرار بگیرد و روش‌های موجود هنوز در این موارد کاستی‌هایی را داشته‌اند.

۱.۳ تشخیص ناهنجاری با نظارت ضعیف

تشخیص ناهنجاری عمیق با نظارت ضعیف [۳۵، ۵۲] تلاش می‌کند تا از شبکه‌های عمیق استفاده کند تا مدل‌های مناسب تشخیص ناهنجاری را با استفاده از سیگنال‌های نظارتی ضعیف بیاموزد. به عنوان مثال می‌توان به استفاده از دادگانی اشاره کرد که به صورت ناقص، غیر دقیق و یا غلط برچسب گذاری شده‌اند. برچسب دادگان، دانش بسیار مهمی را درباره ناهنجاری‌ها دربر دارد که می‌تواند عامل مهمی باشد که توسط آن بتوان نرخ یادآوری را بهبود داد [۴۸، ۳۵، ۵۲، ۵۳]. یک امکان جالب توجه در کارهای آینده استفاده از دادگان ناهنجار با برچسب دقیق برای بالابردن دقت روش‌های موجود است و معمولاً چنین نمونه‌هایی در کاربردهای واقعی در دسترس خواهند بود، برای مثال استفاده از نمونه‌هایی که توسط سیستم‌های جلوگیری از کلاهبرداری و یا متخصصان آن زمینه مشخص خواهند شد می‌تواند در این کاربردها مفید واقع شود. با این حال، از آنجایی که ناهنجاری‌ها می‌توانند بسیار ناهمگن باشند، ناهنجاری‌های ناشناخته و یا جدیدی می‌توانند وجود داشته باشند که فراتر از گستره نمونه‌های ناهنجاری داده شده قرار دارند. بنابراین، یک جهت مهم، تشخیص ناهنجاری ناشناخته است که در آن هدف ما ساختن مدل‌های تشخیصی است که از ناهنجاری‌های برچسب‌گذاری شده و محدود، به ناهنجاری‌های ناشناخته تعمیم می‌یابند. برخی از مطالعات اخیر [۵۴، ۵۲، ۱۲، ۵۵] نشان می‌دهند که مدل‌های عمیق قادر به یادگیری ناهنجاری‌هایی هستند که فراتر از محدوده نمونه‌های ناهنجاری ارائه شده است. بررسی بیشتر میزان تعمیم پذیری و توسعه مدل‌ها برای بهبود عملکرد و دقت مدل‌های موجود بسیار مهم است.

۲.۳ موضوعات کاربردی جدید مرتبط با تشخیص ناهنجاری

برخی از کاربردها و مسائل تحقیقاتی در حال ظهور و جالب توجه وجود دارد که می‌توانند فرسته‌های مهمی برای گسترش روش‌های تشخیصی عمیق را به وجود آورند. مورد اول، مسئله تشخیص نقاط خارج از دامنه^۱ است که در آن سعی می‌شود دادگانی را که با توزیع عمومی سایر دادگان متفاوت هستند تشخیص داده شوند^[۵۶، ۵۷]. این ایده بسیار نوینی است تا با استفاده از یادگیری ماشین پدیده‌های ناشناخته و جدید محیط مورد بررسی را کشف کرده و از آن‌ها بهره‌مند شد. این موضوع خود یک مسئله تشخیص ناهنجاری نیز به حساب می‌آید اما، در این مسئله انتظار داریم برچسب دادگان برای دسته‌های عادی در دسترس باشند و سعی می‌شود دقت تشخیص برای دسته‌های دادگان عادی حفظ شود.

مورد دوم، موضوع یادگیری کنجکاوانه^۲ است^[۵۹، ۶۰، ۶۱] که هدف آن یادگیری یک تابع پاداش جایزه^۳ در یادگیری تقویتی با پاداش‌های پراکنده است. از دید کاربرد، معمولاً روش‌های یادگیری تقویتی در محیط‌هایی با پاداش‌های پراکنده، اغلب با مشکلاتی موافق خواهند شد و عملکرد صحیحی ندارند. در یادگیری کنجکاوانه، سعی می‌شود مشکل پراکنده‌ی پادash‌های محیط را با یک پاداش اضافی علاوه بر پادash‌های پراکنده اولیه از محیط برطرف شود. پاداش جازه معمولاً بر اساس جدید بودن حالت و یا نادر بودن حالت تعریف می‌شود. برای مثال اگر عامل حالات نادر را کشف کند، پاداش بسیار بالایی را دریافت خواهد کرد. حال آنکه این حالات کمیاب، مفهومی مشابه با ناهنجاری دارند. این ایده به ذهن می‌رسد که می‌توان از روش‌های تشخیص ناهنجاری برای مشخص کردن این حالات خاص کمک گرفت و یا از رویکردهای یادگیری کنجکاوی برای تشخیص ناهنجاری‌ها کمک گرفت، مانند کار انجام شده توسط وانگ و همکاران^[۶۲].

اکثر مدل‌های عمیق و غیر عمیق برای تشخیص ناهنجاری فرض می‌کنند که دادگان غیرعادی، نمونه‌هایی مستقل و به طور یکسان توزیع شده است (I.I.D). و این در حالی است که ناهنجاری‌ها در کاربردهای واقعی ممکن است به صورت مستقل و یکسانی توزیع شده نباشند. به عنوان مثال، وجود ناهنجاری در علائم بیماری و به صورت همزمان در تشخیص زودهنگام بیماری‌ها به طور متقابل اثر می‌گذارد و خطای تشخیص را تقویت می‌کند. که این خود مستلزم یادگیری ناهنجاری‌های غیر مستقل (None-I.I.D) است^[۶۳]. این نیاز در رویه‌های پیچیده بسیار مهم است، به عنوان مثال، در جایی که ناهنجاری‌ها فقط دارای انحرافات ظریف هستند، اگر این ویژگی‌های غیرعادی، غیر مستقل در نظر گرفته نشود، در فضای داده پنهان خواهند شد. در نهایت، کاربردهای جالب دیگر شامل تشخیص نمونه‌های متخاصم^[۶۴، ۶۵]، سیستم ضد جعل در سیستم‌های بیومتریک^[۶۶، ۶۷]، و یا تشخیص زودهنگام رویدادهای نادر فاجعه آمیز در این عرصه جای خواهند گرفت.

۳.۳ موضوع پیشنهادی برای پایان نامه

با توجه به کاربرد وسیع مسئله تشخیص ناهنجاری در حوزه پزشکی و مشکل کمبود دادگان برچسب خورده به دلیل چالش‌هایی که در این حوزه وجود دارد، استفاده از روش‌های تشخیص ناهنجاری در این حوزه بسیار مناسب است. در پردازش تصاویر پزشکی به دلیل منحصر به فرد بودن بافت بدن اشخاص مختلف و همچنین نادر بودن بیماری‌ها در میان افراد می‌توان از دید مسئله تشخیص ناهنجاری به وجود توده‌های سرطانی در تصاویر پزشکی نگاه کرد و به بررسی مسئله از این دید پرداخت که

¹Out Of Distribution Detection(ODD)

²curiosity learning

³Bonus reward function

می‌تواند به عنوان پیشنهادی برای پروژه پایانی بررسی شود.

مراجع

- [1] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Comput. Surv.*, vol.41, jul 2009.
- [2] R. Chalapathy and S. Chawla, “Deep learning for anomaly detection: A survey,” 01 2019.
- [3] B. Murugan, M. Elhoseny, K. Shankar, and J. Uthayakumar, “Region-based scalable smart system for anomaly detection in pedestrian walkways,” *Comput. Electr. Eng.*, vol.75, pp.146–160, may 2019.
- [4] T. Ehret, A. Davy, J.-M. Morel, and M. Delbracio, “Image anomalies: A review and synthesis of detection methods,” *Journal of Mathematical Imaging and Vision*, vol.61, no.5, pp.710–743, 2019.
- [5] L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, E. Müller, and M. Kloft, “Deep one-class classification,” in *Proceedings of the 35th International Conference on Machine Learning* (J. Dy and A. Krause, eds.), vol.80 of *Proceedings of Machine Learning Research*, pp.4393–4402, PMLR, 10–15 Jul 2018.
- [6] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, “Lof: Identifying density-based local outliers,” in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, SIGMOD ’00, (New York, NY, USA), pp.93–104, Association for Computing Machinery, 2000.
- [7] “From autoencoder to beta-vae,” <https://lilianweng.github.io/lil-log/2018/08/12/from-autoencoder-to-beta-vae.html>. Accessed: 2018-08-12.
- [8] D. Khattar, J. S. Goud, M. Gupta, and V. Varma, “Mvae: Multimodal variational autoencoder for fake news detection,” in *The World Wide Web Conference*, WWW ’19, (New York, NY, USA), pp.2915–2921, Association for Computing Machinery, 2019.
- [9] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, “f-anogan: Fast unsupervised anomaly detection with generative adversarial networks,” *Medical Image Analysis*, vol.54, pp.30–44, 2019.
- [10] F. E. Grubbs, “Procedures for detecting outlying observations in samples,” *Technometrics*, vol.11, pp.1–21, 1969.
- [11] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, “Lof: identifying density-based local outliers,” in *ACM SIGMOD Conference*, 2000.
- [12] G. Pang, C. Shen, and A. van den Hengel, “Deep anomaly detection with deviation networks,” in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery data mining*, pp.353–362, 2019.

- [13] F. Keller, E. Müller, and K. Böhm, “Hics: High contrast subspaces for density-based outlier ranking,” *2012 IEEE 28th International Conference on Data Engineering*, pp.1037–1048, 2012.
- [14] “Feature bagging for outlier detection,” pp.157–166, 2005.
- [15] F. Liu, K. Ting, and Z.-H. Zhou, “Isolation-based anomaly detection,” *ACM Transactions on Knowledge Discovery from Data*, vol.6, no.1, pp.1 – 39, 2012.
- [16] T. Pevný, “Loda: Lightweight on-line detector of anomalies,” *Machine Learning*, vol.102, pp.275–304, 2016.
- [17] G. Pang, L. Cao, L. Chen, D. Lian, and H. Liu, “Sparse modeling-based sequential ensemble learning for effective outlier detection in high-dimensional numeric data,” in *AAAI Conference on Artificial Intelligence*, 2018.
- [18] G. Pang, L. Cao, L. Chen, and H. Liu, “Learning homophily couplings from non-iid data for joint feature selection and noise-resilient outlier detection,” in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, pp.2585–2591, 2017.
- [19] F. Azmandian, A. Yilmazer, J. G. Dy, J. A. Aslam, and D. R. Kaeli, “Gpu-accelerated feature selection for outlier detection using the local kernel density ratio,” *2012 IEEE 12th International Conference on Data Mining*, pp.51–60, 2012.
- [20] L. Cao, “Coupling learning of complex interactions,” *Information Processing and Management*, vol.51, no.2, pp.167–186, 2015.
- [21] C. C. Aggarwal. *Outlier Analysis*. Springer, 2017.
- [22] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, “Support vector method for novelty detection,” in *Proceedings of the 12th International Conference on Neural Information Processing Systems*, NIPS’99, (Cambridge, MA, USA), pp.582–588, MIT Press, 1999.
- [23] M. Bhuvaneshwari, E. G. M. Kanaga, J. Anitha, K. Raimond, and S. T. George, “Chapter 7 - a comprehensive review on deep learning techniques for a bci-based communication system,” in *Demystifying Big Data, Machine Learning, and Deep Learning for Healthcare Analytics* (P. N, S. Kautish, and S.-L. Peng, eds.), pp.131–157, Academic Press, 2021.
- [24] M.-N. Nguyen and N. Vien, “Scalable and interpretable one-class svms with deep learning and random fourier features,” 04 2018.
- [25] C. Zhou and R. C. Paffenroth, “Anomaly detection with robust deep autoencoders,” in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’17, (New York, NY, USA), pp.665–674, Association for Computing Machinery, 2017.
- [26] D. Gong, L. Liu, V. Le, B. Saha, M. R. Mansour, S. Venkatesh, and A. v. d. Hengel, “Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection,” 2019.
- [27] D. P. Kingma and M. Welling, “An introduction to variational autoencoders,” *Foundations and Trends® in Machine Learning*, vol.12, no.4, pp.307–392, 2019.

- [28] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, “Unsupervised anomaly detection with generative adversarial networks to guide marker discovery,” in *Information Processing in Medical Imaging* (M. Niethammer, M. Styner, S. Aylward, H. Zhu, I. Oguz, P.-T. Yap, and D. Shen, eds.), (Cham), pp.146–157, Springer International Publishing, 2017.
- [29] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, “Ganomaly: Semi-supervised anomaly detection via adversarial training,” in *Asian Conference on Computer Vision*, pp.622–637, Springer, 2018.
- [30] I. Golan and R. El-Yaniv, “Deep anomaly detection using geometric transformations,” in *Advances in Neural Information Processing Systems* (S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, eds.), vol.31, Curran Associates, Inc., 2018.
- [31] C. P. Ngo, A. A. Winarto, C. K. L. Kou, S. Park, F. Akram, and H. K. Lee, “Fence gan: Towards better anomaly detection,” 2019.
- [32] S. Wang, Y. Zeng, X. Liu, E. Zhu, J. Yin, C. Xu, and M. Kloft, “Effective end-to-end unsupervised outlier detection via inlier priority of discriminative network,” in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox, and R. Garnett, eds.), pp.5960–5973, Curran Associates, Inc., 2019.
- [33] H. Zenati, C. S. Foo, B. Lecouat, G. Manek, and V. R. Chandrasekhar, “Efficient gan-based anomaly detection,” 2018.
- [34] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, “Deep learning for anomaly detection: A review,” *ACM Computing Surveys (CSUR)*, vol.54, no.2, pp.1–38, 2021.
- [35] Y. Tian, G. Pang, Y. Chen, R. Singh, J. W. Verjans, and G. Carneiro, “Weakly-supervised video anomaly detection with robust temporal feature magnitude learning,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021.
- [36] H. Borgli, V. Thambawita, P. H. Smedsrud, S. Hicks, D. Jha, S. L. Eskeland, K. R. Randel, K. Pogorelov, M. Lux, D. T. D. Nguyen, D. Johansen, C. Griwodz, H. K. Stensland, E. Garcia-Ceja, P. T. Schmidt, H. L. Hammer, M. A. Riegler, P. Halvorsen, and T. de Lange, “HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy,” *Scientific Data*, vol.7, no.1, p.283, 2020.
- [37] G. Pang, C. Ding, C. Shen, and A. v. d. Hengel, “Explainable deep few-shot anomaly detection with deviation networks,” *arXiv preprint arXiv:2108.00462*, 2021.
- [38] S. Hawkins, H. He, G. J. Williams, and R. A. Baxter, “Outlier detection using replicator neural networks,” in *International Conference on Data Warehousing and Knowledge Discovery*, 2002.
- [39] G. Pang, L. Cao, L. Chen, and H. Liu, “Learning representations of ultrahigh-dimensional data for random distance-based outlier detection,” *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018.
- [40] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, “Mvtex ad — a comprehensive real-world dataset for unsupervised anomaly detection,” *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.9584–9592, 2019.

- [41] W. Liu, W. Luo, D. Lian, and S. Gao, “Future frame prediction for anomaly detection - a new baseline,” *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.6536–6545, 2017.
- [42] D. Abati, A. Porrello, S. Calderara, and R. Cucchiara, “Latent Space Autoregression for Novelty Detection,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision and Pattern Recognition*, 2019.
- [43] W. Liu, W. Luo, D. Lian, and S. Gao, “Future frame prediction for anomaly detection - a new baseline,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.6536–6545, 2018.
- [44] G. Pang, L. Cao, L. Chen, and H. Liu, “Learning representations of ultrahigh-dimensional data for random distance-based outlier detection,” in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’18, (New York, NY, USA), pp.2041–2050, Association for Computing Machinery, 2018.
- [45] H. Wang, G. Pang, C. Shen, and C. Ma, “Unsupervised representation learning by predicting random distances,” in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, IJCAI’20, 2021.
- [46] L. Ruff, R. A. Vandermeulen, N. Görnitz, A. Binder, E. Müller, K.-R. Müller, and M. Kloft, “Deep semi-supervised anomaly detection,” in *International Conference on Learning Representations*, 2020.
- [47] B. Zong, Q. Song, M. R. Min, W. Cheng, C. Lumezanu, D. ki Cho, and H. Chen, “Deep autoencoding gaussian mixture model for unsupervised anomaly detection,” in *ICLR*, 2018.
- [48] W. Sultani, C. Chen, and M. Shah, “Real-world anomaly detection in surveillance videos,” in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [49] M. Sabokrou, M. Khalooei, M. Fathy, and E. Adeli, “Adversarially learned one-class classifier for novelty detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.3379–3388, 2018.
- [50] P. Zheng, S. Yuan, X. Wu, J. Li, and A. Lu, “One-class adversarial nets for fraud detection,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol.33, 03 2018.
- [51] P. Perera, R. Nallapati, and B. Xiang, “Organ: One-class novelty detection using gans with constrained latent representations,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.2893–2901, 2019.
- [52] G. Pang, C. Shen, H. Jin, and A. v. d. Hengel, “Deep weakly-supervised anomaly detection,” 2019.
- [53] G. Pang, L. Chen, L. Cao, and H. Liu, “Learning representations of ultrahigh-dimensional data for random distance-based outlier detection,” in *KDD 2018 - Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp.2041–2050, Association for Computing Machinery, July

2018. Funding Information: This work is partially supported by the ARC Discovery Grant DP180100966. Publisher Copyright: © 2018 Association for Computing Machinery.; 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2018 ; Conference date: 19-08-2018 Through 23-08-2018.
- [54] G. Pang, A. van den Hengel, C. Shen, and L. Cao, “Toward deep supervised anomaly detection,” in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, ACM, aug 2021.
 - [55] L. Ruff, R. A. Vandermeulen, N. Görnitz, A. Binder, E. Müller, K.-R. Müller, and M. Kloft, “Deep semi-supervised anomaly detection,” 2019.
 - [56] D. Hendrycks and K. Gimpel, “A baseline for detecting misclassified and out-of-distribution examples in neural networks,” 2016.
 - [57] K. Lee, K. Lee, H. Lee, and J. Shin, “A simple unified framework for detecting out-of-distribution samples and adversarial attacks,” in *Advances in Neural Information Processing Systems* (S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, eds.), vol.31, Curran Associates, Inc., 2018.
 - [58] J. Ren, P. J. Liu, E. Fertig, J. Snoek, R. Poplin, M. A. DePristo, J. V. Dillon, and B. Lakshminarayanan, “Likelihood ratios for out-of-distribution detection,” 2019.
 - [59] Y. Burda, H. Edwards, A. Storkey, and O. Klimov, “Exploration by random network distillation,” 2018.
 - [60] Y. Burda, H. Edwards, D. Pathak, A. Storkey, T. Darrell, and A. A. Efros, “Large-scale study of curiosity-driven learning,” 2018.
 - [61] D. Pathak, P. Agrawal, A. A. Efros, and T. Darrell, “Curiosity-driven exploration by self-supervised prediction,” 2017.
 - [62] H. Wang, G. Pang, C. Shen, and C. Ma, “Unsupervised representation learning by predicting random distances,” in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20* (C. Bessiere, ed.), pp.2950–2956, International Joint Conferences on Artificial Intelligence Organization, 7 2020. Main track.
 - [63] G. Pang, “Non-iid outlier detection with coupled outlier factors,” 2019.
 - [64] K. Grosse, P. Manoharan, N. Papernot, M. Backes, and P. McDaniel, “On the (statistical) detection of adversarial examples,” 2017.
 - [65] A. Paudice, L. Muñoz-González, A. Gyorgy, and E. C. Lupu, “Detection of adversarial training examples in poisoning attacks through anomaly detection,” 2018.
 - [66] S. Fatemifar, M. Awais, A. Akbari, and J. Kittler, “A stacking ensemble for anomaly based client-specific face spoofing detection,” 2020.
 - [67] D. Pérez-Cabo, D. Jiménez-Cabello, A. Costa-Pazo, and R. J. López-Sastre, “Deep anomaly detection for generalized face anti-spoofing,” *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp.1591–1600, 2019.

Abstract

Detection of abnormalities is important, which is studied in various research fields and has many applications. A common need in the field of real-world analysis is to find out which examples are very different from the majority of existing examples in terms of similarity of behavior and appearance. This difference could be due to measurement errors during data collection. Sometimes this difference can indicate the existence of unknown phenomena that are happening behind the scenes of the statistical population of the study cases and we are unaware of it.

In data science, the term anomaly belongs to data, from the point of view of a defined similarity criterion, its similarity with other existing data is very low. For example, if we compare the radiology photo of a person with lung disease with the radiology photos taken from the lungs of healthy people, we will notice the difference between this photo and other photos. This dissimilarity in the data indicates that the person has a lung disease. In fact, doctors find out the existence of disease by observing these dissimilarities. The act of comparing data can also be done by computer, which is the subject of this seminar.

In this seminar, we tried to examine methods based on deep learning for abnormality detection. Since the application of this topic is very wide in various fields and many articles have been published in relation to various applications, we tried to limit the scope of the seminar and while introducing the various applications of the problem of anomaly detection, examine the methods related to the application Image processing and computer vision. Considering the number of articles in recent years and the existence of comprehensive articles in this field, we will review most of the new articles that have been published in recent years, and for the rest of the methods, we will limit ourselves to referring to other articles.



Department of computer engineering

Deep learning for anomaly detection

MSc. Seminar

Computer engineering - Artificial intelligence and
robotics

Ali Naderi Parizi

Supervisor:

Dr. Mohsen Soryani

November 2022