



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

اردیبهشت ماه ۱۴۰۱

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 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 پی 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 آنها که کاربرد این موضوع در حوزه‌های مختلف بسیار وسیع است و مقالات بسیار متعددی در رابطه با کاربردهای مختلف به چاپ رسیده، 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	مسئله تشخیص ناهنجاری
۳	۲.۱	جنبهای مختلف تشخیص ناهنجاری
۳	۲.۱.۱	کاربردهای مسئله 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 مسئله تشخیص ناهنجاری
۶	۱.۴.۱	چالش‌های عمومی مسئله تشخیص ناهنجاری
۶	۲.۴.۱	چالش‌های تشخیص ناهنجاری که میتوان با بکارگیری روش‌های عمیق به سراغ آنها رفت
۹	۵.۱	ساختار کلی روش‌های تشخیص ناهنجاری
۱۰	۶.۱	ساختار گزارش
۱۱	۲	مروری بر کارهای انجام شده برای مسئله تشخیص ناهنجاری
۱۱	۱.۲	مروری بر روش‌های سنتی
۱۲	۱.۱.۲	روش‌های مبتنی بر رده‌بندی
۱۴	۲.۱.۲	روش‌های مبتنی بر معیار فاصله
۱۴	۳.۱.۲	روش‌های مبتنی بر مدل آماری

۱۴	استفاده از یادگیری عمیق برای مسئله تشخیص ناهنجاری	۲.۲
۱۵	استفاده از یادگیری عمیق برای استخراج ویژگی	۱.۲.۲
۱۵	خود رمزگذار	۲.۲.۲
۱۹	مدل‌های مولد	۲.۲.۲
۲۰	شبکه‌های مولد رقابتی (GAN)	۴.۲.۲
۲۲	مدل‌های جریانی	۵.۲.۲
۲۲	مجموعه دادگان موجود برای تشخیص ناهنجاری	۳.۲
۲۴	کارهای آینده	۳
۲۴	بررسی سیگنال‌های ناهنجاری ناظارتی	۱.۳
۲۴	تشخیص ناهنجاری با ناظارت ضعیف	۲.۳
۲۵	موضوعات کاربردی جدید مرتبط با مسئله تشخیص ناهنجاری	۳.۳
۲۶	موضوع پیشنهادی برای پایان نامه	۴.۳
۲۷	مراجع	
۲۷	کتاب‌نامه	

# فهرست تصاویر

۱	مثالی از تفاوت دادگان ناهنجار و نوین	۱.۱
۲	مثالی ساده از نقاط ناهنجار در میاد مجموعه دادهای در فضای دوبعدی (نقاط $O_1, O_2, O_3$ نقاط ناهنجاری را نشان می‌دهند) [۱۰]	۲.۱
۳	ناهنجاری نقطه‌ای و دنباله‌ای [۸]	۳.۱
۵	ناهنجاری در کاربرد نظارت ویدیو [۲۰]	۴.۱
۵	مثال‌هایی از ناهنجاری در تصاویر [۱۲]	۵.۱
۱۳	ماشین بردار پشتیبان یک کلاسه	۱.۲
۱۳	بردار پشتیبان توصیفگر داده عمیق [۳۶]	۲.۲
۱۴	نمایش کلی روش عامل پرت محلی [۵]	۳.۲
۱۶	بردار پشتیبان توصیفگر داده عمیق [۳۶]	۴.۲
۱۶	مدل خود رمز کننده	۵.۲
۱۸	مدل خود رمز کننده حذف نویز	۶.۲
۲۰	مدل خود رمز کننده variational	۷.۲
۲۱	شبکه مولد رقابتی	۸.۲
۲۲	نمایش نحوه آموزش مدل F-AnoGan [۴۰]	۹.۲

# فهرست جداول

۱۲	.....	۱.۲ دسته‌بندی روش‌های سنتی
۱۵	.....	۲.۲ الگوریتم‌های عمیق مورد استفاده در تشخیص ناهنجاری
۲۳	.....	۳.۲ مجموعه دادگان در دسترس برای تشخیص ناهنجاری
۲۳	.....	۴.۲ الگوریتم‌های عمیق مورد استفاده در تشخیص ناهنجاری

# فصل ۱

## مقدمه

تشخیص ناهنجاری<sup>۱</sup> 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 این 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 می‌توانند نشان دهنده وجود پدیده‌ای ناشناخته باشد که در پشت پرده 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 اندازی کافی زیاد نیست. به این دادگان اصطلاحاً دادگان نوین<sup>۲</sup> 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

<sup>1</sup>Anomaly detection

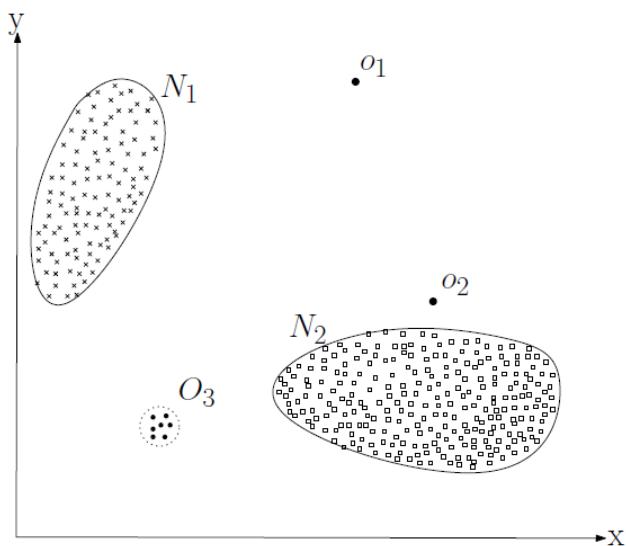
<sup>2</sup>Novelties

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 مسئله تشخیص ناهنجاری

تشخیص ناهنجاری که با عنوان تشخیص دادگان خارج از محدوده<sup>۳</sup> 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. 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 [۱۶]. کاربردهای تشخیص ناهنجاری بسیار وسیع است و در حوزه‌های گوناگونی مورد استفاده قرار می‌گیرد.



شکل ۲.۱: مثالی ساده از نقاط ناهنجار در میاد مجموعه دادهای در فضای دوبعدی (نقاط  $O_1, O_2, O_3$  نقاط ناهنجاری را نشان می‌دهند) [۱۰]

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<sup>۴</sup>، ناهنجاری مفهومی<sup>۵</sup>، ناهنجاری مجموعه‌ای<sup>۶</sup>. 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

<sup>3</sup>Outlier detection

<sup>4</sup>Point anomaly

<sup>5</sup>Contextual anomalies

<sup>6</sup>Collective anomalies

رفتاری تصادفی از خود نشان می‌دهند که اغلب تفسیر خاصی ندارند. برای مثال مبلغ بسیار بالای تراکنش در یک رستوران یک تراکنش غیر عادی به حساب می‌آید که با در نظر گرفتن آن در فضای بازنمایی دادگان این نقطه شباهتی به دیگر دادگان نخواهد داشت. دسته دوم ناهمجارتی‌های مفهومی هستند که در این دسته مفهوم داده در یک مکان و یا زمان مختلف می‌تواند به صورت ناهمجارتی 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 ناهمجارتی‌ها که اصطلاحاً ناهمجارتی مجموعه‌ای گفته می‌شود، مفهوم ناهمجارتی از در یک سلسله از رویدادها در حالی که هر رویداد یک داده کاملاً عادی است. برای مثال در دنباله تراکنش‌های یک کارت اعتبار 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: ناهمجارتی نقطه‌ای و دنباله‌ای [۸]

## ۲.۱ جنبه‌های مختلف تشخیص ناهمجارتی

مسئله تشخیص ناهمجارتی را 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 (مانند صدا، موسیقی، فیلم، متن و ...) غیر دنباله‌ای (مانند عکس، مشخصات بیمار 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 ناهمجارتی‌ها را از دید در دسترس بود برچسب دادگان مورد استفاده بررسی کرد. اما باید توجه داشت که پدیده‌های ناهمجارتی اصولاً کم اتفاق می‌افتد و تعداد آنها در دادگان موجود کم است. با این حال می‌توان روش‌های تشخیص ناهمجارتی را از دید رویکرد بر اساس در دسترس بودن برچسب دادگان به سه دسته باناظر، با نظارت ضعیف و همچنین بدون ناظر تقسیم کرد.

## ۳.۱ کاربردهای مسئله تشخیص ناهمجارتی

برای درک اهمیت و کاربرد مسئله تشخیص ناهمجارتی 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 ناهمجارتی را به تفصیل حوزه‌های کاربردی مختلف می‌آوریم.

## 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 ظربان 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 حیاطی غیر طبیعی از پیش‌آمدن اتفاقات ناگوار جلوگیری کنند.

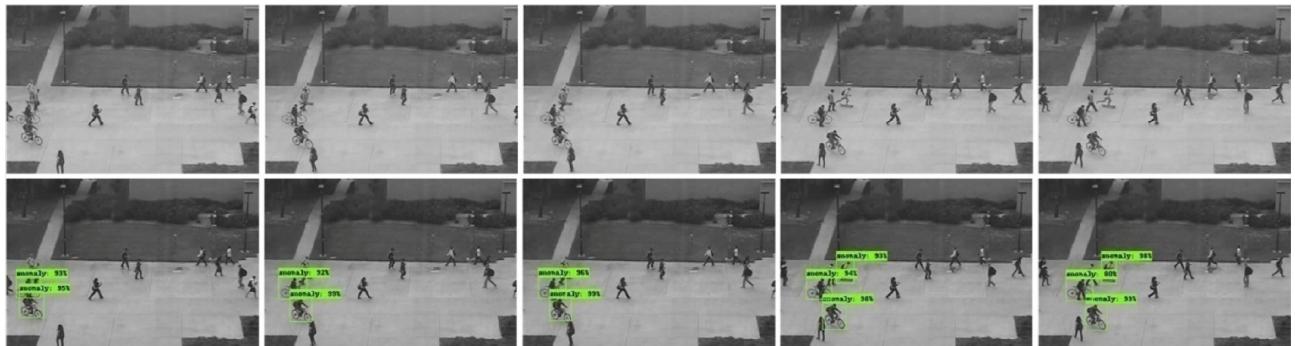
## ۴.۳.۱ سامانه‌های هوشمند و اینترنت اشیا

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.

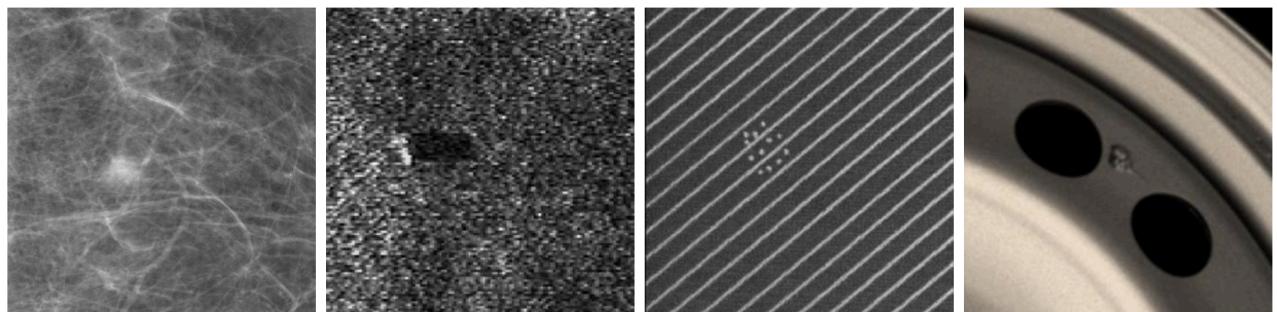
## ۵.۳.۱ نظارت ویدیویی و سیستم‌های امنیتی

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 ظبط شده توسط این دوربین‌ها کار بسیار دشوار و وقتی‌گیری 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 کار وقت و منابع بسیاری را طلب می‌کند. 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 نظارت ویدیو [۲۰]



به ترتیب از سمت چپ، توده سرطان سینه، مین زیردریایی، نقص رنگ‌آمیزی کاشی تولید شده در کارخانه، نمونه نقص موجود در چرخ خودرو.

شکل ۵.۱: مثال‌هایی از ناهنجاری در تصاویر [۱۲]

### ۶.۳.۱ خودروهای خودران

یکی دیگر از حوزه‌های بسیار پرطرفدار در سال‌های اخیر ساخت خودروهای خودران و رانندگی خودکار وسائل 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 درحال عبور پرداخت. با توجه به اینکه شرایط غیر عادی در رانندگی که منجر به تصادف و خطر شود به ندرت اتفاق می‌افتد و همچنین این شرایط می‌توانند به صورت‌ها و شکل‌ها مختلف روی دهنده، استفاده از روش‌های تشخیص ناهنجاری در این کاربردها بسیار مورد پسند پژوهشگران این حوزه قرار گرفته است.

## ۴.۱ چالش‌های مسئله تشخیص ناهنجاری

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

## ۱.۴.۱ چالش‌های عمومی مسئله تشخیص ناهنجاری

بر خلاف سایر بحث‌های یادگیری ماشین که به یافتن پدیده‌ها و الگوهای نامشخص 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.<sup>۷</sup>

۱. مجھول بودن<sup>۸</sup> : 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 جدید، کلاهبرداری‌ها و نفوذهای شبکه.

۲. ناهمگن<sup>۹</sup> بودن دسته‌های ناهنجار : ناهنجاری‌ها نامنظم هستند، و بنابراین، یک دسته از ناهنجاری‌ها ممکن است ویژگی‌های غیرطبیعی کاملاً متفاوتی از دسته دیگر از ناهنجاری‌ها نشان دهند. 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: ناهنجاری‌ها معمولاً نمونه‌های داده نادری هستند، در تضاد با نمونه‌های معمولی که اغلب بخش بزرگی از داده‌ها را تشکیل می‌دهند. بنابراین، جمع‌آوری مقدار زیادی از نمونه‌های غیرعادی برچسب‌گذاری شده، اگر نگوییم غیرممکن، دشوار است. این منجر به در دسترس نبودن داده‌های برچسب‌گذاری شده در مقیاس بزرگ در اکثر برنامه‌ها می‌شود. عدم تعادل طبقاتی نیز به این دلیل است که طبقه بندی نادرست ناهنجاری‌ها معمولاً بسیار پرهزینه‌تر از نمونه‌های عادی است.

۴. گوناگونی انواع ناهنجاری: سه نوع کاملاً متفاوت از ناهنجاری مورد بررسی قرار گرفته است [۲۸]. ناهنجاری‌های نقطه‌ای، نمونه‌های جداگانه‌ای هستند که در حالت عادی غیرعادی هستند. اکثر موارد فردی دیگر، به عنوان مثال، شاخص‌های سلامت غیرطبیعی یک بیمار. ناهنجاری‌های شرطی، که به آن ناهنجاری‌های زمینه‌ای نیز معروف است، به موارد غیرعادی فردی نیز اشاره دارد، اما در یک زمینه خاص، به عنوان مثال، نمونه‌های داده در زمینه خاص غیرعادی هستند، در غیر این صورت عادی هستند. زمینه‌ها می‌توانند در کاربردهای دنیای واقعی بسیار متفاوت باشند، به عنوان مثال، افت/افزایش ناگهانی دما در یک زمینه زمانی خاص، یا تراکنش‌های سریع کارت اعتباری در زمینه‌های فضایی غیرمعمول. ناهنجاری‌های گروهی، یا ناهنجاری‌های جمعی، زیرمجموعه‌ای از نمونه‌های داده‌ای هستند که در مجموع غیرعادی هستند. سایر نمونه‌های داده؛ اعضای منفرد ناهنجاری جمعی ممکن است ناهنجاری نباشند، به عنوان مثال، زیرگراف‌های فوق العاده متراکم که توسط حساب‌های جعلی در شبکه‌های اجتماعی تشکیل شده‌اند به عنوان یک مجموعه ناهنجاری هستند، اما گره‌های فردی در آن زیرگراف‌ها می‌توانند به اندازه حساب‌های واقعی عادی باشند.

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

ماهیت پیچیده مسئله موجود باعث به وجود آمدن چالش‌های بسیاری در تشخیص ناهنجاری شده است. برخی از چالش‌های مانند مقیاس پذیر با توجه به اندازه دادگان در سال‌های اخیر مورد توجه قرار گرفته است ام چالش‌های اساسی و حل نشده

<sup>7</sup>Unknownness

<sup>8</sup>Heterogeneous

دیگری برای تشخیص ناهنجاری وجود دارند که یادگیری عمیق میتواند در حل آنها بسیار کمک کننده باشد. از جمله‌ی 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 موارد 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 نظارت (به عنوان مثال، [۷۱، ۸۴، ۱۷۳]، هنوز اغلب دارای نتایج مثبت کاذب بالایی در مجموعه داده‌های دنیای واقعی هستند [۲۰، ۱۱۵]). چگونگی کاهش مثبت کاذب و افزایش نرخ یادآوری تشخیص یکی از چالش‌های مهم و در عین حال دشوار است، بهویژه برای هزینه قابل توجه عدم شناسایی ناهنجاری‌ها.

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 توجه می‌شوند. تشخیص ناهنجاری با ابعاد بالا یک مشکل طولانی مدت بوده است [۱۷۸]. انجام تشخیص ناهنجاری در فضای کم‌بعدی که توسط زیرمجموعه کوچکی از ویژگی‌های اصلی یا ویژگی‌های جدید ساخته شده است.

یک راه حل ساده است، به عنوان مثال، در روش‌های مبتنی بر زیرفضا [۱۲۳، ۸۴، ۷۷، ۷۰] و روش‌های مبتنی بر انتخاب ویژگی [۱۲، ۱۰۹، ۱۱۱]. با این حال، شناسایی برهمنکنش‌ها و جفت‌های ویژگی پیچیده (مثلاً مرتبه بالا، غیرخطی و ناهمگن) [۲۲] ممکن است در داده‌های با ابعاد بالا ضروری باشد، اما همچنان یک چالش بزرگ برای تشخیص ناهنجاری است. علاوه بر این، چگونگی تضمین فضای ویژگی جدید که اطلاعات مناسب را برای روش‌های تشخیص خاص حفظ می‌کند، برای تشخیص ناهنجاری دقیق پایین‌دستی حیاتی است، اما به دلیل ناشناخته‌های ذکر شده و ناهمگونی‌های ناهنجاری‌ها، چالش برانگیز است. همچنین، شناسایی ناهنجاری‌ها از نمونه‌هایی که ممکن است به یکدیگر وابسته باشند، مانند روابط زمانی، مکانی، مبتنی بر نمودار و سایر روابط واستنگی متقابل، چالش برانگیز است [۴، ۲۲، ۵۳].

۳. یادگیری ناهنجاری‌ها به صورت به صرفه<sup>۹</sup>: به دلیل دشواری و هزینه جمع‌آوری داده‌های ناهنجاری برچسب‌گذاری شده در مقیاس بزرگ، تشخیص ناهنجاری کاملاً نظارت شده اغلب غیرعملی است زیرا در دسترس بودن داده‌های آموزشی برچسب‌گذاری شده با کلاس‌های عادی و غیرعادی را فرض می‌کند. در دهه گذشته، تلاش‌های تحقیقاتی عمدۀ بر روی تشخیص ناهنجاری بدون نظارت متمرکز شده است که به هیچ داده آموزشی برچسب‌گذاری شده ای 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 برخی داده‌های ناهنجاری برچسب‌گذاری شده اغلب دشوار نیست. در عمل، اغلب پیشنهاد می‌شود که تا حد امکان از چنین داده‌های برچسب‌گذاری شده به آسانی در دسترس استفاده شود [۲]. بنابراین، استفاده از این داده‌های برچسب‌گذاری شده برای یادگیری بازنمایی‌های بیانی از نرمال/نابهنجاری برای تشخیص دقیق ناهنجاری بسیار مهم است. تشخیص ناهنجاری نیمه نظارت شده، که مجموعه‌ای از داده‌های آموزشی عادی برچسب‌گذاری شده را فرض می‌کند، یک جهت تحقیقاتی است که به این مشکل اختصاص داده شده است. خط تحقیقاتی دیگر، تشخیص ناهنجاری با نظارت ضعیف است که

<sup>9</sup>Data efficient

فرض می‌کند چند برچسب برای کلاس‌های ناهنجاری داریم، اما برچسب‌های کلاس جزئی/ناقص هستند (یعنی کل مجموعه کلاس ناهنجاری را در بر نمی‌گیرند)، نادقيق (یعنی برچسب‌های درشت دانه) یا نادرست است (یعنی برخی از برچسب‌های داده شده ممکن است نادرست باشند). دو چالش اصلی این است که چگونه نمایش‌های عادی/ناهنجاری بیانی را با مقدار کمی از داده‌های ناهنجاری برچسب‌گذاری شده یاد بگیریم و چگونه مدل‌های تشخیصی را یاد بگیریم که به ناهنجاری‌های جدید آشکار شده توسط داده‌های ناهنجاری برچسب‌گذاری شده تعییم داده می‌شوند.

۴. تشخیص ناهنجاری مقاوم در برابر نویز: بسیاری از روش‌های تشخیص ناهنجاری با نظارت ضعیف فرض می‌کنند که داده‌های آموزشی برچسب‌گذاری شده تمیز هستند، که می‌تواند در برابر نمونه‌های پر سر و صدایی که به اشتباه به عنوان برچسب کلاس مخالف برچسب‌گذاری شده‌اند آسیب‌پذیر باشد. در چنین مواردی، ممکن است به جای آن از روش‌های بدون نظارت استفاده کنیم، اما این روش از داده‌های برچسب‌گذاری شده واقعی استفاده نمی‌کند. 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، نویز در اینجا می‌تواند داده‌های دارای برچسب اشتباه یا ناهنجاری‌های بدون برچسب باشد. چالش اصلی این است که میزان نویزها می‌تواند به طور قابل توجهی با مجموعه داده‌ها متفاوت باشد و نمونه‌های نویز ممکن است به طور نامنظم در فضای داده توزیع شوند.

۵. تشخیص ناهنجاری‌های پیچیده: بیشتر روش‌های موجود برای ناهنجاری‌های نقطه‌ای هستند که نمی‌توانند برای ناهنجاری شرطی و ناهنجاری گروهی استفاده شوند زیرا رفتارهای کاملاً متفاوتی با ناهنجاری‌های نقطه از خود نشان می‌دهند. یکی از چالش‌های اصلی در اینجا گنجاندن مفهوم ناهنجاری‌های شرطی/گروهی در معیارها/مدل‌های ناهنجاری است. همچنین، روش‌های کنونی عمدتاً بر تشخیص ناهنجاری‌ها از منابع داده تکی تمرکز می‌کنند، در حالی که بسیاری از برنامه‌ها 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. توضیح ناهنجاری: در بسیاری از حوزه‌های حیاتی اینمی، ممکن است موارد عمده‌ای وجود داشته باشد اگر مدل‌های تشخیص ناهنجاری مستقیماً به عنوان مدل‌های جعبه سیاه استفاده شوند، خطراتی را به همراه خواهد داشت. برای مثال، موارد نادر داده‌ای که به عنوان ناهنجاری گزارش شده‌اند، ممکن است منجر به سوگیری الگوریتمی احتمالی علیه گروه‌های اقلیت ارائه شده در داده‌ها شود، مانند گروه‌هایی که کمتر در سیستم‌های کشف تقلب و کشف جرم ارائه شده‌اند. یک رویکرد موثر برای کاهش این نوع ریسک، داشتن الگوریتم‌های توضیح ناهنجاری است که سرنخ‌های ساده‌ای در مورد اینکه چرا یک نمونه داده خاص به عنوان ناهنجاری شناسایی می‌شود، ارائه می‌دهد. سپس کارشناسان انسانی می‌توانند تعصب را بررسی کرده و تصحیح کنند. ارائه چنین توضیحی می‌تواند به اندازه دقت تشخیص در برخی برنامه‌ها مهم باشد. با این حال، اکثر مطالعات تشخیص ناهنجاری تنها بر دقت تشخیص تمرکز می‌کنند و توانایی ارائه توضیح ناهنجاری‌های شناسایی شده را نادیده می‌گیرند. استخراج توضیح ناهنجاری از روش‌های تشخیص خاص هنوز یک مشکل تا حد زیادی حل نشده است، بهویژه برای مدل‌های پیچیده. 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<sup>۱۰</sup> است. در این مرحله نگاشتی از دادگان ورودی به فضای معین آموخته می‌شود. این نگاشت را می‌توان به صورت تابعی مانند زیر تعریف کرد.

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

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 دسته ناهنجار اختصاص می‌دهد که به آن امتیاز ناهنجاری<sup>۱۱</sup> 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  $\delta$ ، 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  $\theta$  و  $\eta$  و  $\delta$  نیست.

۲. یک مرحله‌ای: تنها یکی از مجموعه پارامترهای موجود  $\theta$  یا  $\eta$  یادگرفته می‌شوند.

۳. دو مرحله‌ای: هر دو مجموعه پارامتر  $\theta$  و  $\eta$  به صورت مستقل و جداگانه یادگرفته می‌شوند.

۴. ادغامی<sup>۱۲</sup>: هر دو مجموعه پارامتر  $\theta$  و  $\eta$  باهم یادگرفته می‌شوند.

در صورت عدم وجود برچسب‌های دادگان موجود، ناچار به استفاده از روش بدون ناظر هستیم که در آن از هیچ گونه اطلاعاتی در مورد ماهیت دادگان در دسترس نیست. در این موضع معمولاً  $\delta$  از پیش تعریف شده است و یا همراه با  $\eta$  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.

<sup>10</sup>Data representation

<sup>11</sup>Anomaly score

<sup>12</sup>Integrated

## 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. در فصل دوم به معرفی پیشنبازها و الگوریتم‌های پایه و پر استفاده در روش‌های تشخیص ناهنجاری پرداخته می‌شود. فصل سوم نیز با بررسی روش‌های 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، در فصل چهارم، مسائل باز و کارهای آینده این حوزه معرفی شده و چند نمونه پیشنهاد برای پژوهش‌هایی مطرح خواهد شد.

## ۲ فصل

### مروری بر کارهای انجام شده برای مسئله 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 نویسنده‌گان مقالات و هدف از استفاده از این روش‌ها در کارهای انجام شده کمک می‌کند. مطالعه روش‌های پایه کمک میکند تا کارهای انجام شده اخیر را که در فصل بعد مورد بررسی قرار گرفته‌اند بهتر درک بشوند.

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 و همچنین این روش‌ها میتوانند با ترکیب شدن با روش‌های عمیق، مدل‌های جدیدی را بسازند و همچنین ایده به وجود آوردن مدل‌های عمیق دیگر باشند. در بخش دوم روش‌های عمیق مورد استفاده در مسئله تشخیص ناهنجاری آورده شده‌اند که مطالعه این روش‌ها برای درک مطالب فصل بعدی ضروری است.

#### 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 فعل 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 برای 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<sup>۱</sup>. 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 [۱۰].

<sup>۱</sup> ر.ک. جدول ۱.۲

## 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 <sup>2</sup> 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

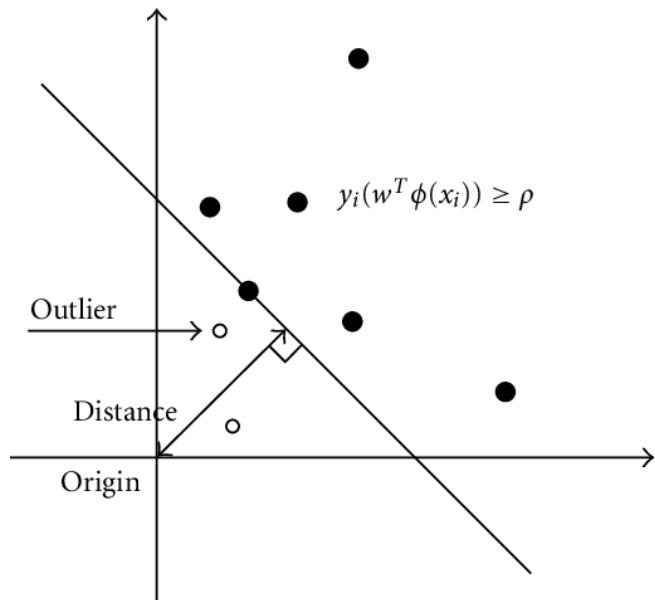
همانطور که در ابتدای این بخش گفته شد، یکی از ایده‌های کلی در روش‌های مورد استفاده برای تشخیص ناهنجاری استفاده از ایده رده‌بندی است. در اینگونه روش‌ها تلاش می‌شود یک مرز تفکیک میان دادگان عادی و دادگان ناهنجار رسم شود. اگر چنین مرزی وجود داشته باشد، می‌توان با استفاده از الگوریتم‌های رده‌بند موجود اقدام 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<sup>۳</sup> 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.<sup>۴</sup>

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<sup>۴</sup> است. در این روش سعی می‌شود کره‌ای با کوچک ترین اندازه شعاع ممکن حول دادگان موجود رسم شود.

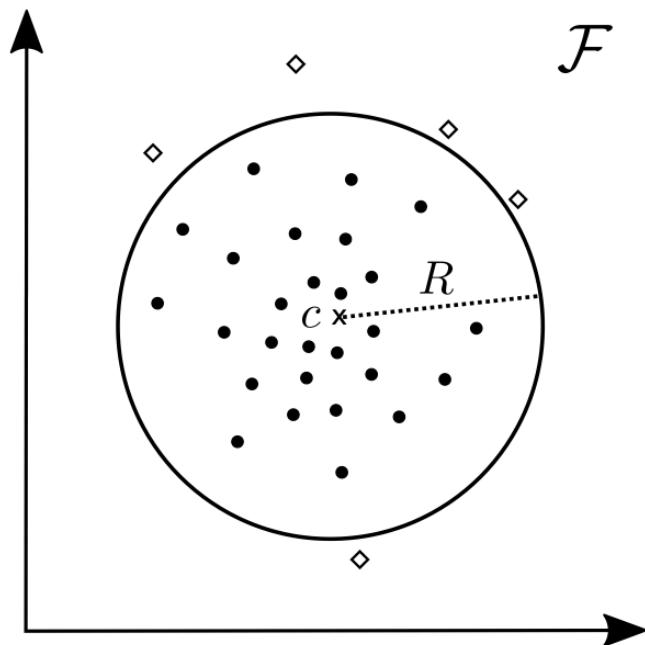
<sup>3</sup>One-class SVM

<sup>4</sup>Support 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 [۳۶].

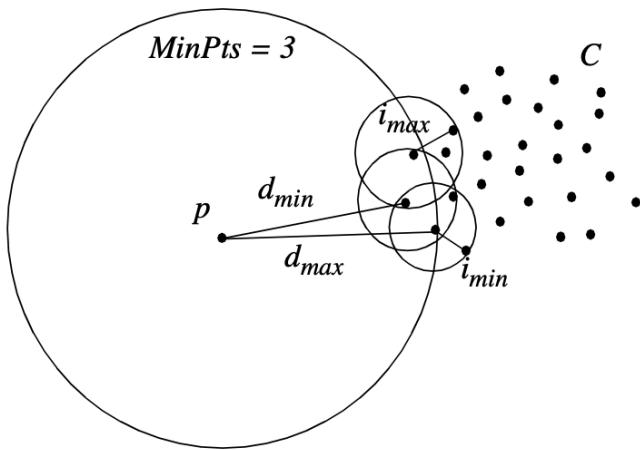


شکل ۲.۲: بردار پشتیبان توصیفگر داده عمیق [۳۶]

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 تعیین 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 عامل پرت محلی<sup>۵</sup> است. در این روش میانگین فاصله هر نقطه از همسایگان محلی محاسبه شده و اگر این میانگین از یک مقدار آستانه محلی<sup>۶</sup> 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 [۵]. 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 عامل پرت محلی [۵]

## 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 مدل آماری قرار می‌گیرند که احتمال وقوع آنها بیشتر است. برای مثال در روش مدل خطی پویا<sup>۶</sup> ابتدا دادگان را از فضای ورودی به یک فضای از پیش تعیین شده نگاشت می‌کنیم. سپس با استفاده از مدل بدست آمده سعی در پیش‌بینی مقدار ورودی با توجه به دیگر دادگان موجود می‌کنیم. 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، به دسته دادگان ناهنجار تعلق می‌گیرد.

## ۲.۲ استفاده از یادگیری عمیق برای مسئله 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 مسائله پرداخته می‌شود که پایه و اساس خیلی از روش‌های

<sup>۵</sup>Local Outlier Factor

<sup>۶</sup>Dynamic liner 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. جدول ۲.۲ لیستی از روش‌های مورد بحث در این بخش را جمع آوری کرده است.

جدول ۲.۲: الگوریتم‌های عمیق مورد استفاده در تشخیص ناهنجاری

نام مدل	مقاله مرجع
AE	عدم نیاز به برچسب دادگان
SAE	درنظر گرفتن ابعاد بالاتر در لایه میانی برای در نظر گرفتن ویژگی‌های بیشتری از دادگان
DAE	کاهش حساسیت مدل به نویز موجود در دادگان
RDA	استفاده از دادگان ناهنجار در طی فرآیند یادگیری
VAE	نگاشت دادگان به یک توزیع احتمال بجای یادگیری نگاشت بدون پایه و اساس در لایه میانی
DCAE	
DTS	
GAN	

## ۱.۲.۲ استفاده از یادگیری عمیق برای استخراج ویژگی

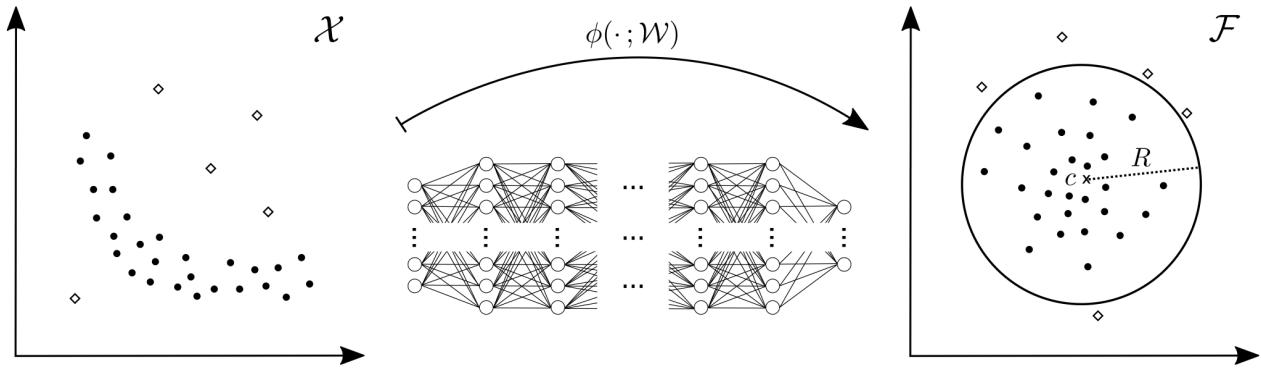
یکی از ابتدایی ترین ایده‌هایی که در مورد استفاده از روش‌های سنتی موجود با توجه به معرفی و پیشرفت ساختارهای عمیق به ذهن می‌رسد، استفاده از این ساختارها به منظور استخراج ویژگی از دادگان با ابعاد بالا و نه لزوماً تفکیک پذیر خطی است. ساختارهای عمیق با توجه به قابلیت بالای یادگیری ترکیب‌های غیر خطی گوناگون، می‌توانند به عنوان تابع 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 شوند (AlexNet ، VGG و ...). پس از بدست آمدن ویژگی‌ها در فضای جدید، یک تابع امتیاز دهنده 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 فضای معین برای بدست آورد بازنمایی خوبی از دادگان 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 دهنده به نقاط است بدست آورد. راف و همکاران با استفاده از این ایده، روش بردار پشتیبان توصیف‌گر داده عمیق را معرفی کردن که در مقایسه با روش‌های سنتی عملکرد بسیار بهتری را از خود نشان داده است [۳۶].

## ۲.۲.۲ خود رمزگذار

خودرمز گشا<sup>۷</sup>‌ها نوعی از شبکه‌های عصبی هستند که از روش پس انتشار<sup>۸</sup> برای یادگیری ویژگی‌های مفهومی استفاده می‌کنند. این شبکه‌ها به صورت دو مرحله‌ای اقدام به یادگیری می‌کنند که به ترتیب رمزگاری و رمزگشایی نام دارند. در مرحله اول داده ورودی به شبکه رمز کننده داده می‌شود و رمز کننده داده ورودی را به یک فضا با ابعاد پایین نگاشت می‌کند. به این فضا

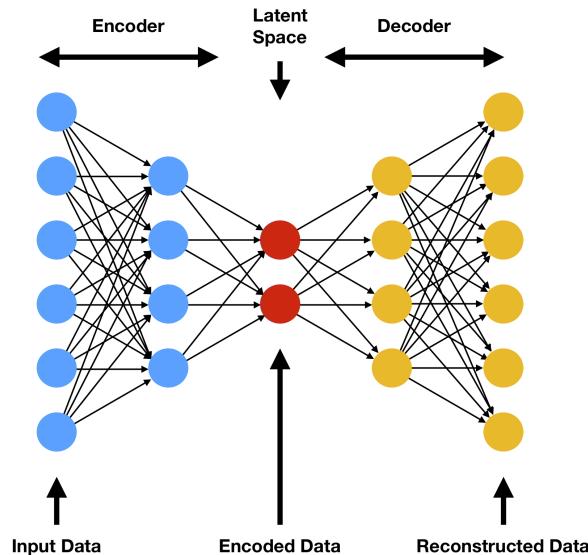
<sup>7</sup>AutoEncoder

<sup>8</sup>Backpropagation



شکل ۴.۲: بردار پشتیبان توصیفگر داده عمیق [۳۶]

به اصطلاح فضای باقی‌مانده<sup>۹</sup> یا فضای  $z$  می‌گویند. در مرحله دوم، بازنمایی بdst آمده وارد شبکه رمزگشا شده تا داده از فضای باقی‌مانده دوباره به فضای ورودی باز گردانده شود. آنچه که انتظار می‌رود آن است که خروجی مدل با آنچه در ورودی به مدل داده شده بسیار شبیه باشند. در این صورت قسمت رمز کنند توانسته بازنمایی خوبی از داده را در فضای باقی‌مانده ایجاد کند [۳].



شکل ۵.۲: مدل خود رمز کننده

اگر بخواهیم کار کرد مدل شکل ۵.۲ را با فرمول ریاضی توصیف کنیم، با در نظر داده  $X$  به عنوان ورودی مدل، رمز کنند با گرفتن این ورودی، آن را به فضای باقی‌مانده و به نقطه  $z$  نگاشت می‌کند. اگر تابع رمز کننده 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 رمزگشا را مانند تابعی درنظر بگیریم و آن را  $g$  بنامیم، این تابع با دریافت ورودی  $z$ ، اقدام به بازسازی

<sup>9</sup>Latent space

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  $\delta$  برای بدست آوردن بهترین نتیجه با آزمون و خطای و یا روش‌های دیگر مانند استفاده از نمودار حساسیت و دقت تعیین می‌شود.

خودرمز کننده‌ها باید به تغییرات دادگان ورودی حساس باشند تا بتوانند با دقت مطلوب داده رمز شده را بازسازی کنند. همچنین این حساسیت نباید به اندازه‌ای باشد که باعث بشود مدل بجای یادگیری عملکرد مناسب، به بخارط سپاری دادگان ورودی بپردازد و دچار بیش‌برازش<sup>۱۰</sup> بشود. برای دستیابی به چنین توازنی، انواع مختلفی از خودرمز کننده‌ها معرفی شده‌اند که با افزودن یک مقدار تنظیم کننده<sup>۱۱</sup> به تابع خطای اصلی معرفی شده، بدست می‌آیند.

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

### خود رمزگذار SAE

خود رمز کننده SAE<sup>۱۲</sup> یکی از انواع خودرمز کننده‌ها است. ایده اصلی این گونه رمز کننده‌ها این است که، با توجه به اینکه تعداد نورون‌ها لایه مخفی به اندازه کافی زیاد نباشند شاید نتوانند به خوبی مفاهیم پیچیده را یاد بگیرند. در نتیجه پیشنهاد می‌شود در لایه مخفی تعداد نورون‌های بیشتری قرار گیرند اما از تابع فعال سازی ترتیبی داده شود تا این نورون‌ها تا حد ممکن کم استفاده شوند و یا به اصطلاح، به صورت خلوت<sup>۱۳</sup> فعال سازی آنها صورت بگیرید. برای دستیابی به چنین هدفی می‌توان از تنظیم کننده<sup>۱۴</sup> در تابع خطای مدل استفاده کرد. نوع اول استفاده از تنظیم کننده نرم یک است<sup>۱۵</sup> که معادله تابع خطای به 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 تاجای امکان صفر باشد و با صفر شدن این وزن‌ها، در واقع نورون‌های کمتری در فرایند محاسبه استفاده می‌شوند.

<sup>10</sup>Overfit

<sup>11</sup>Rgulizer

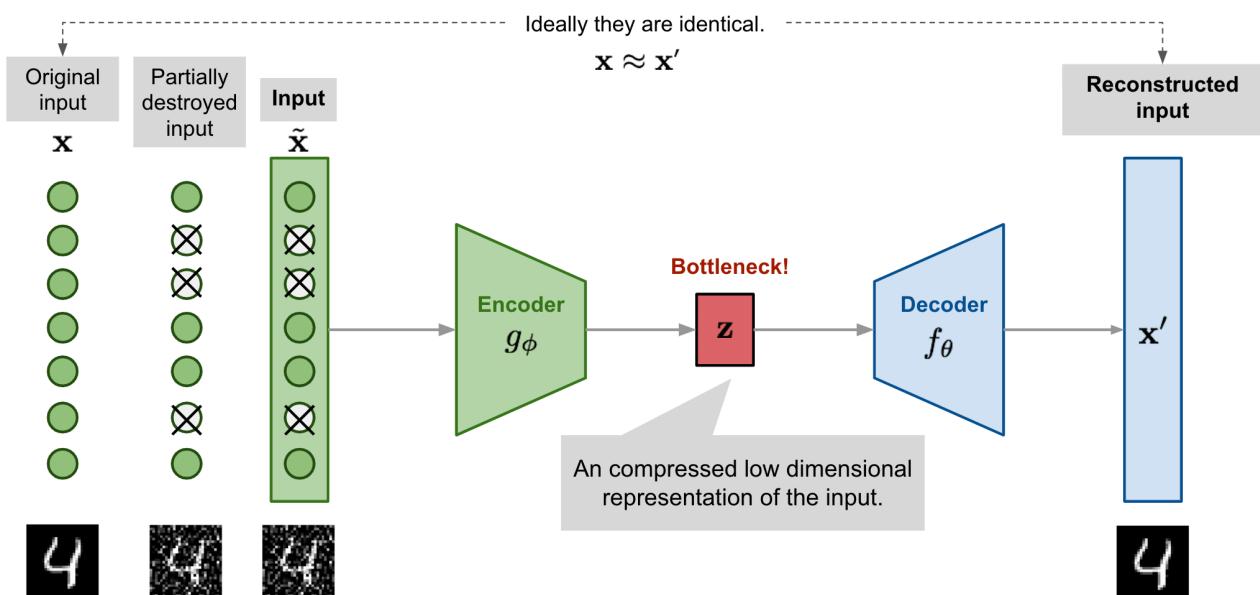
<sup>12</sup>Sparse AutoEncoder (SAE)

<sup>13</sup>Sparse

<sup>14</sup>Regulizer

<sup>15</sup>L1-Rgulizer

نوع دیگری از خود رمز کننده که می‌توان از آن برای حذف نویز در داده استفاده کرد را به اصطلاح خودرمز کننده حذف نویز<sup>۱۶</sup> نام دارد. تفاوت این مدل با حالت کلی خود رمز کننده‌ها در فرایند آموزش مدل است. در این مدل داده ورودی ابتدا با استفاده از یک تولید کننده نویز، نویزی می‌شود. سپس به مدل خودرمز کننده داده می‌شود. شبکه نهایی باید بتواند نویز اضافه شده با تصاویر را حذف کند. برای انجام این کار یک راه ساده، تعریفتابع خطابه صورت مقایسه خروجی مدل و ورودی اصلی بدون نویز است. شبکه باید تلاش کند تا اختلاف تصویر بازسازی 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. سپس با مقایسه مقدار خروجی و ورودی مکان‌هایی که تفاوت زیادی بایکدیگر دارند به احتمال تعلق به دسته ناهنجاری در آنها زیاد است.



شکل ۶.۲: مدل خود رمز کننده حذف نویز

### خودرمز کننده RDA

خودرمز کننده‌هایی که تا کنون معرفی شدند، در مرحله آموزش مدل تنها از دادگان عادی و بدون ناهنجاری استفاده می‌کردند و دادگان ناهنجار تنها زمان آزمون مدل استفاده می‌شدند. حال اگر بخواهیم دادگان ناهنجار را نیز در فرآیند آموزش مدل دخیل کنیم باید روش جدیدی را معرفی کنیم. خودرمز کننده مقاوم<sup>۱۷</sup> درواقع از ایده تجزیه و تحلیل مؤلفه بنیادی مقاوم<sup>۱۸</sup> برگرفته شده است. در روش تجزیه و تحلیل مؤلفه بنیادی مقاوم، دادگان ورودی با استفاده از دو ماتریس مرتبه پایین<sup>۱۹</sup> و خلوت<sup>۲۰</sup> نمایش داده می‌شوند.

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

<sup>16</sup>Denoising Auto Encoder

<sup>17</sup>Robust Deep AutoEncoder

<sup>18</sup>Robust PCA

<sup>19</sup>Low rank

<sup>20</sup>Sparse

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<sup>۲۱</sup>.

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

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

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

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

اگر رمزکننده و رمزگشا را به عنوان دو تابع  $f$  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 دهی به ناهنجاری در این نوع خودرمزکننده 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 [۵۱].

### ۳.۲.۲ مدل‌های مولد

#### خودرمزکننده VAE

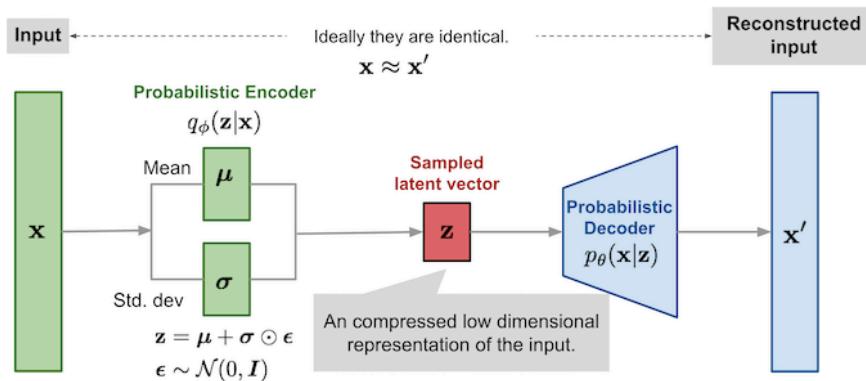
مشکل که خودرمزکننده‌هایی که تا کنون معرفی کردیم در این است که نگاشت دادگان به فضای باقی‌مانده به صورت قطعی صورت می‌گیرد. یعنی هر نقطه از فضای ورودی به یک مقطعه معین از فضای باقی‌مانده نگاشته می‌شود. از طرف دیگر اگر یک نقطه را به صورت تصادفی در فضای باقی‌مانده، مانند  $\hat{z}$  را در نظر بگیریم، نمی‌توان به طور قطعی گفت که این نقطه به کدام دسته از نقاط تعلق خواهد گرفت. در واقع خودرمزکننده‌هایی که تا کنون مطالعه کردیم به خوبی دادگان ورودی را به فضایی با ابعاد دیگر نگاشت می‌کردند اما در هیچ یک از این روش‌ها ما اختیاری برای کنترل روند و نحوه این نگاشت نداشتیم. انواع مختلف این رمزکننده‌ها نیز بسته به نیاز، نگاشتهای گوناگونی را در اختیار ما قرار می‌دادند تا مناسب کاربرد انتخاب شده باشند. برای اینکه در طی فرایند یادگیری ما بر روی نحوه نگاشت دادگان به فضای باقی‌مانده کنترل داشته باشیم، نوع دیگری از خودرمزکننده‌ها تحت عنوان خودرمزکننده VAE<sup>۲۲</sup> معرفی شده است [؟]. در این روش بجای یادگیری نگاشت گسسته و قطعی دادگان به فضای باقی‌مانده، سعی می‌شود توزیع دادگان در فضای باقی‌مانده یادگرفته شود که

<sup>۲۱</sup> در اینجا  $\|\cdot\|_F$  نرم جمع مقدار یکتا (singular value) است.

<sup>۲۲</sup>Variational AutoEncoder

در این صورت دادگان در این فضا توضیع 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<sup>۲۳</sup> حداکثر شود. معادل تابع بهینه سازی این مدل به صورت زیر تعریف می‌شود.



شکل ۷.۲: مدل خود رمز کننده variational

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

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

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

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 درنظر گرفت، تلاش می‌کند تا توزیع دادگان در فضای بازنمایی  $z$  بسیار مشابه توزیع دادگان ورودی باشند. بنابراین بازنمایی دادگان در فضای باقیمانده بر خلاف مدل پایه به صورت غیر قطعی<sup>۲۴</sup> خواهد بود. همچنین دادگان بازسازی شده و همچنین امتیاز ناهنجاری برای دادگان نیز غیر قطعی و به صورت احتمال خواهد بود.

#### ۴.۲.۲ شبکه‌های مولد رقابتی (GAN)

شبکه‌های مولد رقابتی<sup>۲۵</sup> 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 بکشند. بخش اول این مدل که مول

<sup>23</sup>Marginal likelihood

<sup>24</sup>Stochastic

<sup>25</sup>Generative Adversarial Networks

<sup>26</sup>Generator

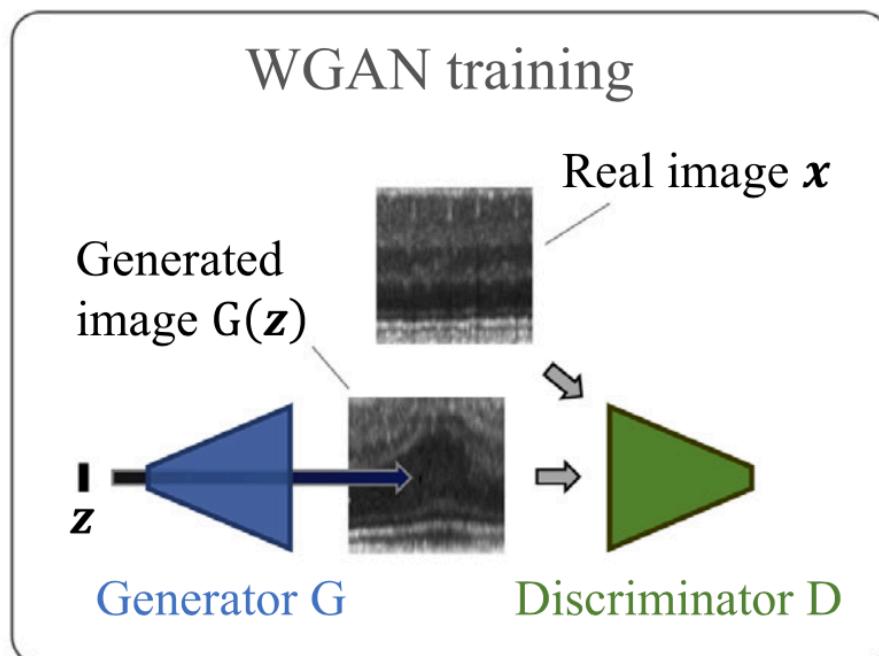
نام دارد، مسئولیت تولید داده مصنوعی را بر عهده دارد. این قسمت با گرفتن یک بردار ورودی از فضای باقیمانده، داده‌ای مصنوعی را تولید می‌کند. خروجی این قسمت به همراه یک نمونه از دادگان آموزش برای مقایسه و داوری جهت تشخیص مصنوعی و یا حقیقی بودن به قسمت دوم مدل که تصمیم گیرنده<sup>۲۷</sup> نام دارد وارد می‌شوند. بخش دوم 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 یافت، تغییر وزن‌ها در آن بخش متوقف شده و وزن‌های بخش دیگر آموزش می‌بینند و پس از بهبود عملکرد قسمت بعدی این چرخه ادامه پیدا می‌کند. تابع خطای مورد استفاده در مدل‌های مولد پایه به صورت زیر است.

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

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  $\delta$  نیز می‌تواند با استفاده از آزمون و خطا و یا با استفاده از منحنی حساسیت و دقت تعیین گردد.



شکل ۸.۲: شبکه مولد رقابتی

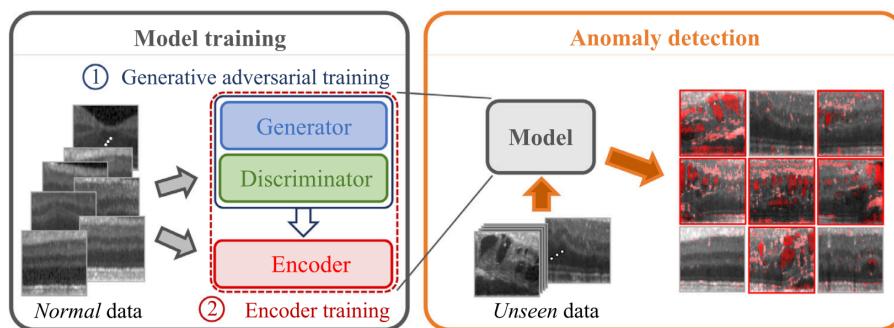
برای اینکه بتوانیم از شبکه مولد نیز در این مسئله کمک بگیریم، می‌توانیم در فرایند آموزش دادگان، بجای انتخاب تصادفی یک نقطه از فضای  $z$  به عنوان ورودی شبکه مولد، با استفاده از یک رمز کننده دادگان ورودی را ابتدا به رمز کننده

<sup>27</sup>Discriminator

بدهیم 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  $\lambda$  یک ابر پارامتر مدل است که به صورت دستی تعیین می شود. این روش در سال ۲۰۱۷ توسط چلگ و همکاران تحت عنوان AnoGan معرفی شد که برای آشکار سازی ناهنجاری های در تصاویر گرفته شده از قرنیه چشم<sup>۲۸</sup> برای تشخیص بیماری چشم مورد استفاده قرار گرفته است [۴۱]. 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 رمز کننده از پیش آموزش دیده، سرعت محاسبات مدل را بهبود دادند [۴۰].



[۴۰] F-AnoGan نمایش آموزش مدل

روش دیگری که از رویکرد شبکه های مولد استفاده می کند، 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 [۲].

## ۵.۲.۲ مدل های جریانی

مدلهای جریانی<sup>۲۹</sup> مشابه خود رمز کننده VAE سعی می کنند تا توزیع دادگان ورودی را بیاموزند. در این رویکرد، سعی می شود تا با استفاده از یک دنباله از توابع مبدل معکوس پذیر که دادگان را از فضای باقیمانده به فضای ورودی نگاشت می کنند، توزیع دادگان آموخته شود. یکی از ایراداتی که بر خود رمز کننده VAE در یادگیری توزیع دادگان وارد است، در نظر گرفتن توزیع مانند گوسی برای دادگان است.

## ۳.۲ مجموعه دادگان موجود برای تشخیص ناهنجاری

یکی از چالش های بر سر راه تولید و توسعه روش های تشخیص ناهنجاری عدم دسترسی به مجموعه دادگان دنیای واقعی و در برگیرنده ناهنجاری های حقیقی است. در بسیاری از پژوهش های صورت گرفته، پژوهشگران برای ارزیابی و قابل مقایسه شدن کار خود با سایرین از مجموعه دادگان موجود و مورد استفاده در دسته بندی استفاده کرده اند [۲، ۱۴، ۴۸، ۴۷، ۳۶، ۲۱، ۵۱].

<sup>28</sup>OCT

<sup>29</sup>Flow based models

در این صورت امکان دارد کارایی روش معرفی شده در کاربردهای دنیای واقعی به خوبی مشخص نشود. از این رو در جدول ۳.۲ 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 درحال بروز رسانی مجموعه‌های داده موجود برای تشخیص ناهنجاری می‌توانید به آدرس <https://git.io/JTs93> سر بزنید) [۲۷].

جدول ۳.۲: مجموعه دادگان در دسترس برای تشخیص ناهنجاری

لیست مجموعه دادگان در دسترس عموم و در برگیرنده ناهنجاری						
مقالات مرجع	جنس دادگان	درصد ناهنجاری	ابعاد	اندازه	حوزه کاربرد	نام
[۴۴]	ویدیو				نظرارت ویدیویی	UCF-Crime
[۲۶، ۴]	تصویر و ویدیو				تشخیص بیماری	HyperKvasir
[۴۴]	ویدیو				نظرارت ویدیویی	UCF-Crime
[۴۴]	ویدیو				نظرارت ویدیویی	UCF-Crime

جدول ۴.۲: الگوریتم‌های عمیق مورد استفاده در تشخیص ناهنجاری

روشهای عمیق معروف برای تشخیص ناهنجاری			
نام مدل	مقاله مرجع	ساختار عمیق مورد استفاده	لينك پیاده سازی
RDA	ژائو و همکاران [۵۱]	خود رمزنگار	<a href="https://git.io/JfYG5">https://git.io/JfYG5</a>
AnoGAN	چگل و همکاران [۴۱]	شبکه مولد به همراه خود رمزنگار	<a href="https://git.io/JfYG5">https://git.io/JfYG5</a>
f-AnoGan	چگل و همکاران [۴۰]	شبکه مولد به همراه خود رمزنگار	
LSA	آباتی و همکاران [۱]	خود رمزنگار	
ALAD	زناتی و همکاران [۴۹]	مدل مولد دوجهته	
GANomaly	آکای و همکاران [۲]	خود رمزنگار به همراه مدل مولد	
FFP	لیو و همکاران [۱۹]	مدل مولد	
EBGAN	زناتی و همکاران [۴۸]	مدل مولد	
GT	گولان و همکاران [۱۴]	استفاده از ساختارهای پایه عمیق	
E3Outlier	پانگ و همکاران [۲۴]	استفاده از ساختارهای پایه عمیق	
REPEN	ژائو و همکاران [۵۱]	استفاده از ساختارهای پایه عمیق	
RDP	وانگ و همکاران [۴۶]	استفاده از ساختارهای پایه عمیق	
AE-OSVM	نگوین و همکاران [۲۲]	ترکیب خود رمزنگار و روش سنتی	
OC-NN	چالاپاتی و همکاران [۹]	استفاده از ساختارهای پایه عمیق	
Deep SVDD	روف و همکاران [۳۶]	استفاده از ساختارهای پایه عمیق	
Deep SAD	روف و همکاران [۳۷]	ترکیب خود رمزنگار و روش سنتی	
DAGMM	زونگ و همکاران [۵۲]	استفاده از ساختارهای پایه عمیق	
MIL	سلطانی و همکاران [۴۳]	خود رمزنگار	
DevNet	پانگ و همکاران [۲۹]	خود رمزنگار	
ALOCC	سبکرو و همکاران [۳۹]	خود رمزنگار	
OCAN	ژنگ و همکاران [۵۰]	استفاده از ساختارهای پایه عمیق	
FenceGAN	ژائو و همکاران [۲۱]	شبکه مولد	
OCGAN	پرا و همکاران [۳۳]	شبکه مولد	
SPADE	کوهن و همکاران [۱۱]	ساختارهای پایه عمیق	

## 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 گیرد و روش‌های موجود هنوز در این موارد کاستی‌هایی را داشته‌اند.

### ۱.۳ بررسی سیگنال‌های ناهنجاری ناظارتی

در مسائل تشخیص ناهنجاری یکی از چالش‌های مهم، در دسترس نبودن دادگان برچسب خورده کافی و مناسب یادگیری است. با توجه به اینکه ناهنجاری‌ها ذاتاً به ندرت اتفاق می‌افتد و همچنین مفهوم ناهنجاری در کاربردهای گوناگون متفاوت است، استفاده از روش‌های با ناظر در این مورد مسئله‌ای چالش برانگیز است. یکی از روش‌های پرکاربرد در تشخیص ناهنجاری، یادگیری بازنمایی دادگان عادی است که در فصل سوم به برخی از آنها پرداختیم. در استفاده از این روش‌ها، این احتمال وجود دارد که دادگان آموزش ممکن است دارای خطای خطا باشند که این مورد ممکن است در فرایند یادگیری تاثیر بگذارد. همچنین در بسیاری از کاربردها مرز مشخص و دقیقی میان دادگان عادی و ناهنجار وجود ندارد. از این رو یافتن روش‌هایی که در مقابل خطای مقاوم باشند یکی از مسائل و چالش‌های بازی در این حوزه است. برای مقابله با چالش کمبود دادگان ناهنجار، ابداع روش‌های داده‌افزایی و بازنمایی دادگان که نیاز به تعداد کمی داده داشته باشند نیز یکی از مسائل باز است.

### ۲.۳ تشخیص ناهنجاری با ناظارت ضعیف

تشخیص ناهنجاری عمیق با ناظارت ضعیف [۲۸، ۴۴] تلاش می‌کند تا از شبکه‌های عمیق استفاده کند تا مدل‌های مناسب تشخیص ناهنجاری را با استفاده از سیگنال‌های ناظارتی 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 [۲۵، ۲۸، ۴۴، ۴۳]. یک امکان جالب توجه در کارهای آینده استفاده از تعدادی 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 مجموعه گسترده نمونه‌های ناهنجاری داده شده قرار دارند. بنابراین، یک جهت مهم در اینجا، تشخیص ناهنجاری ناشناخته است، که در آن هدف ما ساختن مدل‌های تشخیص است که از ناهنجاری‌های برچسب‌گذاری شده محدود به ناهنجاری‌های ناشناخته تعمیم می‌یابند. برخی از مطالعات اخیر [۳۰، ۲۸، ۲۹، ۳۸] نشان می‌دهند که مدل‌های تشخیص عمقی قادر به یادگیری ناهنجاری‌هایی هستند که فراتر از محدوده مثال‌های ناهنجاری ارائه شده است. درک و بررسی بیشتر میزان تعمیم پذیری و توسعه مدل‌هایی برای بهبود بیشتر عملکرد دقت بسیار مهم است.

### ۳.۳ موضوعات کاربردی جدید مرتبط با مسئله تشخیص ناهنجاری

برخی از کاربردها و مسائل تحقیقاتی در حال ظهور و جالب توجه وجود دارد که می‌توانند فرصت‌های مهمی برای گسترش روش‌های تشخیصی عمیق را به وجود آورند. اولی مسئله، مسئله تشخیص نقاط خارج از دامنه<sup>۱</sup> است که در آن سعی می‌شود دادگانی را که با توزیع عمومی سایر دادگان متفاوت هستند تشخیص داده شوند [۱۷، ۱۸، ۲۵]. 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 هستند و سعی می‌شود دقت تشخیص برای دسته‌های دادگان عادی حفظ شود.

موضوع یادگیری کنجکاوانه<sup>۲</sup> است [۷، ۶، ۳۱] که هدف آن یادگیری یک تابع پاداش جایزه<sup>۳</sup> در یادگیری تقویتی با پادash‌های پراکنده است. از دید کاربرد معمولاً روش‌های یادگیری تقویتی معمولاً در محیط‌هایی با پادash‌های پراکنده اغلب با مشکلاتی مواجه می‌شوند و عملکرد صحیحی ندارند. در یادگیری کنجکاوانه، سعی می‌شود مشکل پراکنده‌گی 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 ناهنجاری دارند. و این ایده به ذهن میرست 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 [۴۵].

اکثر مدل‌های عمیق و غیر عمیق برای تشخیص ناهنجاری فرض می‌کنند که دادگان غیرعادی، نمونه‌های داده‌ای مستقل و به طور یکسان توزیع شده است (I.I.D). و این در حالی است که ناهنجاری‌ها در کاربردهای واقعی ممکن است از 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 [۲۲]. این نیاز در سناریوهای 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 می‌شوند. در نهایت، کاربردهای جالب دیگر شامل تشخیص نمونه‌های متخاصم [۱۵، ۳۲]، سیستم ضد جعل در سیستم‌های بیومتریک [۱۲، ۳۴]، و یا تشخیص زودهنگام رویدادهای نادر فاجعه آمیز در این عرصه جای خواهند

<sup>1</sup>Out Of Distribution Detection(ODD)

<sup>2</sup>curiosity learning

<sup>3</sup>Bonus reward function

گرفت.

#### ۴.۳ 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 می‌تواند به عنوان پیشنهادی برای پژوهش پایانی بررسی شود.

## كتاب نامه

- [1] Abati, Davide, Porrello, Angelo, Calderara, Simone, and Cucchiara, Rita. Latent Space Autoregression for Novelty Detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision and Pattern Recognition*, 2019.
- [2] Akcay, Samet, Atapour-Abarghouei, Amir, and Breckon, Toby P. Ganomaly: Semi-supervised anomaly detection via adversarial training. In *Asian Conference on Computer Vision*, pages 622–637. Springer, 2018.
- [3] Bhuvaneshwari, M., Kanaga, E. Grace Mary, Anitha, J., Raimond, Kumudha, and George, S. Thomas. Chapter 7 - a comprehensive review on deep learning techniques for a bci-based communication system. In N, Pradeep, Kautish, Sandeep, and Peng, Sheng-Lung, editors, *Demystifying Big Data, Machine Learning, and Deep Learning for Healthcare Analytics*, pages 131–157. Academic Press, 2021.
- [4] Borgli, Hanna, Thambawita, Vajira, Smedsrud, Pia H, Hicks, Steven, Jha, Debesh, Eskeland, Sigrun L, Randel, Kristin Ranheim, Pogorelov, Konstantin, Lux, Mathias, Nguyen, Duc Tien Dang, Johansen, Dag, Griwodz, Carsten, Stensland, Håkon K, Garcia-Ceja, Enrique, Schmidt, Peter T, Hammer, Hugo L, Riegler, Michael A, Halvorsen, Pål, and de Lange, Thomas. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Scientific Data*, 7(1):283, 2020.
- [5] Breunig, Markus M., Kriegel, Hans-Peter, Ng, Raymond T., and Sander, Jörg. Lof: Identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, SIGMOD '00, pages 93–104, New York, NY, USA, 2000. Association for Computing Machinery.
- [6] Burda, Yuri, Edwards, Harri, Pathak, Deepak, Storkey, Amos, Darrell, Trevor, and Efros, Alexei A. Large-scale study of curiosity-driven learning, 2018.
- [7] Burda, Yuri, Edwards, Harrison, Storkey, Amos, and Klimov, Oleg. Exploration by random network distillation, 2018.
- [8] Chalapathy, Raghavendra and Chawla, Sanjay. Deep learning for anomaly detection: A survey. 01 2019.
- [9] Chalapathy, Raghavendra, Menon, Aditya Krishna, and Chawla, Sanjay. Anomaly detection using one-class neural networks. *arXiv preprint arXiv:1802.06360*, 2018.
- [10] Chandola, Varun, Banerjee, Arindam, and Kumar, Vipin. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3), jul 2009.
- [11] Cohen, Niv and Hoshen, Yedid. Sub-image anomaly detection with deep pyramid correspondences. *CoRR*, abs/2005.02357, 2020.

- [12] Ehret, Thibaud, Davy, Axel, Morel, Jean-Michel, and Delbracio, Mauricio. Image anomalies: A review and synthesis of detection methods. *Journal of Mathematical Imaging and Vision*, 61(5):710–743, 2019.
- [13] Fatemifar, Soroush, Awais, Muhammad, Akbari, Ali, and Kittler, Josef. A stacking ensemble for anomaly based client-specific face spoofing detection, 2020.
- [14] Golan, Izhak and El-Yaniv, Ran. Deep anomaly detection using geometric transformations. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.
- [15] Grosse, Kathrin, Manoharan, Praveen, Papernot, Nicolas, Backes, Michael, and McDaniel, Patrick. On the (statistical) detection of adversarial examples, 2017.
- [16] Grubbs, Frank E. Procedures for detecting outlying observations in samples. *Technometrics*, 11:1–21, 1969.
- [17] Hendrycks, Dan and Gimpel, Kevin. A baseline for detecting misclassified and out-of-distribution examples in neural networks. 2016.
- [18] Lee, Kimin, Lee, Kibok, Lee, Honglak, and Shin, Jinwoo. A simple unified framework for detecting out-of-distribution samples and adversarial attacks. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.
- [19] Liu, Wen, Luo, Weixin, Lian, Dongze, and Gao, Shenghua. Future frame prediction for anomaly detection - a new baseline. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6536–6545, 2018.
- [20] Murugan, B.S., Elhoseny, Mohamed, Shankar, K., and Uthayakumar, J. Region-based scalable smart system for anomaly detection in pedestrian walkways. *Comput. Electr. Eng.*, 75(C):146–160, may 2019.
- [21] Ngo, Cuong Phuc, Winarto, Amadeus Aristo, Kou, Connie Khor Li, Park, Sojeong, Akram, Farhan, and Lee, Hwee Kuan. Fence gan: Towards better anomaly detection. 2019.
- [22] Nguyen, Minh-Nghia and Vien, Ngo. Scalable and interpretable one-class svms with deep learning and random fourier features. 04 2018.
- [23] Pang, Guansong. Non-iid outlier detection with coupled outlier factors. 2019.
- [24] Pang, Guansong, Cao, Longbing, Chen, Ling, and Liu, Huan. Learning representations of ultrahigh-dimensional data for random distance-based outlier detection. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery amp; Data Mining*, KDD ’18, pages 2041–2050, New York, NY, USA, 2018. Association for Computing Machinery.
- [25] Pang, Guansong, Chen, Ling, Cao, Longbing, and Liu, Huan. 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, pages 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.

- [26] Pang, Guansong, Ding, Choubo, Shen, Chunhua, and Hengel, Anton van den. Explainable deep few-shot anomaly detection with deviation networks. *arXiv preprint arXiv:2108.00462*, 2021.
- [27] Pang, Guansong, Shen, Chunhua, Cao, Longbing, and Hengel, Anton Van Den. Deep learning for anomaly detection: A review. *ACM Computing Surveys (CSUR)*, 54(2):1–38, 2021.
- [28] Pang, Guansong, Shen, Chunhua, Jin, Huidong, and Hengel, Anton van den. Deep weakly-supervised anomaly detection, 2019.
- [29] Pang, Guansong, Shen, Chunhua, and van den Hengel, Anton. Deep anomaly detection with deviation networks. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 353–362, 2019.
- [30] Pang, Guansong, van den Hengel, Anton, Shen, Chunhua, and Cao, Longbing. Toward deep supervised anomaly detection. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. ACM, aug 2021.
- [31] Pathak, Deepak, Agrawal, Pulkit, Efros, Alexei A., and Darrell, Trevor. Curiosity-driven exploration by self-supervised prediction, 2017.
- [32] Paudice, Andrea, Muñoz-González, Luis, Gyorgy, Andras, and Lupu, Emil C. Detection of adversarial training examples in poisoning attacks through anomaly detection, 2018.
- [33] Perera, Pramuditha, Nallapati, Ramesh, and Xiang, Bing. Oegan: One-class novelty detection using gans with constrained latent representations. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2893–2901, 2019.
- [34] Pérez-Cabo, Daniel, Jiménez-Cabello, David, Costa-Pazo, Artur, and López-Sastre, Roberto Javier. Deep anomaly detection for generalized face anti-spoofing. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1591–1600, 2019.
- [35] Ren, Jie, Liu, Peter J., Fertig, Emily, Snoek, Jasper, Poplin, Ryan, DePristo, Mark A., Dillon, Joshua V., and Lakshminarayanan, Balaji. Likelihood ratios for out-of-distribution detection, 2019.
- [36] Ruff, Lukas, Vandermeulen, Robert, Goernitz, Nico, Deecke, Lucas, Siddiqui, Shoaib Ahmed, Binder, Alexander, Müller, Emmanuel, and Kloft, Marius. Deep one-class classification. In Dy, Jennifer and Krause, Andreas, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 4393–4402. PMLR, 10–15 Jul 2018.
- [37] Ruff, Lukas, Vandermeulen, Robert A., Görnitz, Nico, Binder, Alexander, Müller, Emmanuel, Müller, Klaus-Robert, and Kloft, Marius. Deep semi-supervised anomaly detection. In *International Conference on Learning Representations*, 2020.

- [38] Ruff, Lukas, Vandermeulen, Robert A., Görnitz, Nico, Binder, Alexander, Müller, Emmanuel, Müller, Klaus-Robert, and Kloft, Marius. Deep semi-supervised anomaly detection, 2019.
- [39] Sabokrou, Mohammad, Khalooei, Mohammad, Fathy, Mahmood, and Adeli, Ehsan. Adversarially learned one-class classifier for novelty detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3379–3388, 2018.
- [40] Schlegl, Thomas, Seeböck, Philipp, Waldstein, Sebastian M., Langs, Georg, and Schmidt-Erfurth, Ursula. f-anogan: Fast unsupervised anomaly detection with generative adversarial networks. *Medical Image Analysis*, 54:30–44, 2019.
- [41] Schlegl, Thomas, Seeböck, Philipp, Waldstein, Sebastian M., Schmidt-Erfurth, Ursula, and Langs, Georg. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In Niethammer, Marc, Styner, Martin, Aylward, Stephen, Zhu, Hongtu, Oguz, Ipek, Yap, Pew-Thian, and Shen, Dinggang, editors, *Information Processing in Medical Imaging*, pages 146–157, Cham, 2017. Springer International Publishing.
- [42] Schölkopf, Bernhard, Williamson, Robert, Smola, Alex, Shawe-Taylor, John, and Platt, John. Support vector method for novelty detection. In *Proceedings of the 12th International Conference on Neural Information Processing Systems*, NIPS’99, pages 582–588, Cambridge, MA, USA, 1999. MIT Press.
- [43] Sultani, Waqas, Chen, Chen, and Shah, Mubarak. Real-world anomaly detection in surveillance videos. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [44] Tian, Yu, Pang, Guansong, Chen, Yuanhong, Singh, Rajvinder, Verjans, Johan W, and Carneiro, Gustavo. Weakly-supervised video anomaly detection with robust temporal feature magnitude learning. In *Proceedings of the IEEE/CVF international conference on computer vision*, 2021.
- [45] Wang, Hu, Pang, Guansong, Shen, Chunhua, and Ma, Congbo. Unsupervised representation learning by predicting random distances. In Bessiere, Christian, editor, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 2950–2956. International Joint Conferences on Artificial Intelligence Organization, 7 2020. Main track.
- [46] Wang, Hu, Pang, Guansong, Shen, Chunhua, and Ma, Congbo. Unsupervised representation learning by predicting random distances. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, IJCAI’20, 2021.
- [47] Wang, Siqi, Zeng, Yijie, Liu, Xinwang, Zhu, En, Yin, Jianping, Xu, Chuanfu, and Kloft, Marius. Effective end-to-end unsupervised outlier detection via inlier priority of discriminative network. In Wallach, H., Larochelle, H., Beygelzimer, A., d’Alché-Buc, F., Fox, E., and Garnett, R., editors, *Advances in Neural Information Processing Systems 32*, pages 5960–5973. Curran Associates, Inc., 2019.
- [48] Zenati, Houssam, Foo, Chuan Sheng, Lecouat, Bruno, Manek, Gaurav, and Chandrasekhar, Vijay Ramaseshan. Efficient gan-based anomaly detection. 2018.

- [49] Zenati, Houssam, Romain, Manon, Foo, Chuan Sheng, Lecouat, Bruno, and Chandrasekhar, Vijay R. Adversarially learned anomaly detection. *2018 IEEE International Conference on Data Mining (ICDM)*, pages 727–736, 2018.
- [50] Zheng, Panpan, Yuan, Shuhan, Wu, Xintao, Li, Jun, and Lu, Aidong. One-class adversarial nets for fraud detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33, 03 2018.
- [51] Zhou, Chong and Paffenroth, Randy C. Anomaly detection with robust deep autoencoders. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’17, pages 665–674, New York, NY, USA, 2017. Association for Computing Machinery.
- [52] Zong, Bo, Song, Qi, Min, Martin Renqiang, Cheng, Wei, Lumezanu, Cristian, ki Cho, Dae, and Chen, Haifeng. Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In *ICLR*, 2018.

# 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 error 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 a 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**

Master seminar report  
Computer engineering - Artificial intelligence and  
robotics

Student name:  
Ali Naderi Parizi

Professor:  
Dr. Mohsen Soryani

November 2022