

دانشگاه صنعتي امیرکبیر  
(پلی تکنیک تهران)

دانشكده مهندسی کامپیوتر و فناوری اطلاعات

گزارش پروژه­ی درس طراحی کم­توان

مروری بر روش­های محاسبات تقریبی

نگارنده

سید سالار هاشمی طاهری

استاد درس

دکتر محمود ممتازپور

مرداد 1398



صفحه فرم ارزیابی و تصویب پایان نامه- فرم تأیید اعضاء كميته دفاع

در این صفحه (هر سه مقطع تحصيلي) بايد فرم ارزيابي یا تایید و تصویب پایان­نامه/رساله موسوم به فرم کمیته دفاع براي ارشد و دكترا و فرم تصويب براي كارشناسي، موجود در پرونده آموزشی را قرار دهند.

|  |  |  |
| --- | --- | --- |
| Untitled | **به نام خدا** |  |
| **تعهدنامه اصالت اثر** | **تاريخ:** |
|  |

اينجانب متعهد مي‌شوم كه مطالب مندرج در اين پايان نامه حاصل كار پژوهشي اينجانب تحت نظارت و راهنمايي اساتيد دانشگاه صنعتي اميركبير بوده و به دستاوردهاي ديگران كه در اين پژوهش از آنها استفاده شده است مطابق مقررات و روال متعارف ارجاع و در فهرست منابع و مآخذ ذكر گرديده است. اين پایان نامه قبلاً براي احراز هيچ مدرك هم‌سطح يا بالاتر ارائه نگرديده است.

در صورت اثبات تخلف در هر زمان، مدرك تحصيلي صادر شده توسط دانشگاه از درجه اعتبار ساقط بوده و دانشگاه حق پيگيري قانوني خواهد داشت.

كليه نتايج و حقوق حاصل از اين پایان نامه متعلق به دانشگاه صنعتي اميركبير مي‌باشد. هرگونه استفاده از نتايج علمي و عملي، واگذاري اطلاعات به ديگران يا چاپ و تكثير، نسخه‌برداري، ترجمه و اقتباس از اين پایان نامه بدون موافقت كتبي دانشگاه صنعتي اميركبير ممنوع است.   
نقل مطالب با ذكر مآخذ بلامانع است.

در صفحه تعهدنامه اصالت اثر، در قسمت بالا سمت چپ، تاریخ دفاع خود را جایگزین تاریخ نوشته شده کنید.

همچنین در صفحه تعهدنامه اصالت اثر، در خط اول، نام و نام خانوادگی خود را به صورت کامل با نام و نام خانوادگی نمونه، جایگزین کنید. در انتهای متن تعهد، در قسمت امضا نیز باید نام و نام خانوادگی کامل خود را وارد نماید.

تقدیم

به...تشکر

چكيده

نیاز به محاسبات و کارایی بیشتر در بسترهای مدرن، احساس نیاز به توسعه­ی روش­های جدیدی را جهت رویارویی با مشکلات ناشی از محدودیت انرژی و تاخیر را به وجود آورده­است. در این راستا یکی از روش­های نویدبخش ارائه شده، که به طور وسیعی در سال­های اخیر مورد مطالعه قرار گرفته است، محاسبات تقریبی می­باشد. این راهکار با هزینه­ی اضافه شدن مقداری تقریب در خروجی­، بهره­ی انرژی و سرعت را برای کاربر فراهم می­سازد. کاربردهای نوظهور در حوزه­هایی مثل پردازش تصویر، داده­کاوری و رباتیک شامل بخش­هایی هستند که در برابر خطا مقاوم هستند و اجرای تقریبی بخش­های مذکور باعث می­شود تا مصالحه­ای بین کیفیت و کارایی در اختیار طراحان قرار گیرد. اگرچه محاسبات تقریبی انرژی و تاخیر را کاهش می­دهد اما تضمینی در خصوص برآورده­سازی نیازهای کاربر در خصوص کیفیت خروجی را ارایه نمی­دهد. از این رو، مدیریت کیفیت خروجی در به کارگیری محاسبات غیردقیق چالشی اساسی است که طراحان باید به آن­ پاسخ مناسبی بدهند.

واژگان کلیدی:

محاسبات تقریبی (Approximate Computing)، ارزیابی کارایی (Performance evaluation).

|  |  |
| --- | --- |
| فهرست مطالب | صفحه |

[فصل اول 1](#_Toc102151015)

[1-1- ساختار گزارش 3](#_Toc102151016)

[فصل دوم 5](#_Toc102151017)

[1-2- طبقه‌بندی روش‌های تشخیص ناهنجاری از دیدگاه در دسترس بودن برچسب داده 6](#_Toc102151018)

[1-1-2- تشخیص ناهنجاری با نظارت 7](#_Toc102151019)

[2-1-2- تشخیص ناهنجاری نیمه‌نظارتی 7](#_Toc102151020)

[3-1-2- تشخیص ناهنجاری بدون نظارت 7](#_Toc102151021)

[2-2- طبقه‌بندی روش‌های تشخیص ناهنجاری از نظر رویکرد حل مسئله 8](#_Toc102151022)

[1-2-2- روش‌های آماری 9](#_Toc102151023)

[1-1-2-2- روش‌های پارامتری 9](#_Toc102151024)

[2-1-2-2- روش‌های غیرپارامتری 10](#_Toc102151025)

[2-2-2- روش‌های یادگیری ماشین 10](#_Toc102151026)

[1-2-2-2- دسته‌بندی 10](#_Toc102151027)

[2-2-2-2- نزدیک‌ترین همسایه 12](#_Toc102151028)

[3-2-2-2- خوشه‌بندی 12](#_Toc102151029)

[3-2- دسته‌بندی بر اساس نحوه تشخیص ناهنجاری 13](#_Toc102151030)

[1-3-2- بر اساس فاصله 13](#_Toc102151031)

[2-3-2- دسته‌بندی تک‌کلاسی 13](#_Toc102151032)

[3-3-2- بر اساس بازسازی 13](#_Toc102151033)

[4-2- معیار‌های ارزیابی روش‌های تشخیص ناهنجاری 14](#_Toc102151034)

[1-4-2- نرخ تشخیص 15](#_Toc102151035)

[2-4-2- دقت 15](#_Toc102151036)

[3-4-2- کارایی 15](#_Toc102151037)

[4-4-2- مقیاس‌پذیری 16](#_Toc102151038)

[5-2- شبکه‌های مولد تخاصمی و تشخیص ناهنجاری 16](#_Toc102151039)

[1-5-2- شبکه‌های مولد تخاصمی 17](#_Toc102151040)

[1-1-5-2- تحلیل نظری شبکه مولد تخاصمی 20](#_Toc102151041)

[2-1-5-2- مزایا و معایب 22](#_Toc102151042)

[1-3-2- نرخ خطا 23](#_Toc102151043)

[2-3-2- اندازهی خطا 24](#_Toc102151044)

[3-3-2- میانگین نسبی اندازهی خطا 24](#_Toc102151045)

[4-3-2- بیشینهی خطا 25](#_Toc102151046)

[5-3-2- میانگین مربع خطا 25](#_Toc102151047)

[6-3-2- نسبت سیگنال به نویز 25](#_Toc102151048)

[7-3-2- مونتکارلو 26](#_Toc102151049)

[8-3-2- استفاده از روشهای یادگیری ماشین 26](#_Toc102151050)

[4-2- سطوح پیادهسازی محاسبات تقریبی 27](#_Toc102151051)

[1-4-2- پیادهسازی در سطح سیستم 27](#_Toc102151052)

[1-1-4-2- روزنهدار کردن حلقه 28](#_Toc102151053)

[2-1-4-2- پراندن وظایف و دسترسی به حافظه 28](#_Toc102151054)

[3-1-4-2- استفاده از تعدادی نسخهی غیردقیق یک برنامه 29](#_Toc102151055)

[4-1-4-2- استفاده از شتابدهندههای برپایهی شبکهی عصبی 29](#_Toc102151056)

[2-4-2- پیادهسازی در سطح معماری 30](#_Toc102151057)

[1-2-4-2- تقریب از طریق مقیاس کردن دقت 30](#_Toc102151058)

[2-2-4-2- تقریب مقدار بارگیری 30](#_Toc102151059)

[3-2-4-2- استفاده از سختافزار غیردقیق یا خطادار 31](#_Toc102151060)

[4-2-4-2- بهخاطرسپاری 32](#_Toc102151061)

[3-4-2- پیادهسازی در سطح مدار 32](#_Toc102151062)

[1-3-4-2- کاهش ولتاژ کاری 32](#_Toc102151063)

[2-3-4-2- استفاده از سلولهای حافظهتقریبی 33](#_Toc102151064)

[5-2- بررسی مقالات ارجاعدادهشده 33](#_Toc102151065)

[1-5-2- هیستوگرام تاریخ چاپ مقالات 33](#_Toc102151066)

[2-5-2- هیستوگرام محل چاپ مقالات 34](#_Toc102151067)

[فصل سوم روش پیشنهادی 36](#_Toc102151068)

[1-3- مدل ALAD 38](#_Toc102151069)

[2-3- مدل R-CALAD 38](#_Toc102151070)

[فصل چهارم جمع‌بندي و نتيجه‌گيري جمع‌بندي و نتيجه‌گيري 43](#_Toc102151071)

[منابع و مراجع 46](#_Toc102151072)

|  |  |
| --- | --- |
| فهرست شکل­ها | صفحه |

[شكل ‏1‌-2- تعداد مقالات چاپ شده در حوزه­ی محاسبات تقریبی بر حسب سال چاپ.](#_Toc276969410) 22

[شكل ‏1‌-‌‌2- تعداد مقالات چاپ شده در حوزه­ی محاسبات تقریبی در مجلات و همایش­ها](#_Toc276969410) 23

# فصل اول

**مقدمه**هنگام تجزیه و تحلیل دادگان موجود در دنیای واقعی، شناسایی نمونه‌های غیرمشابه با سایر نمونه‌ها امری ضروری به نظر می‌رسد. چنین نمونه‌هایی با عنوان ناهنجاری شناخته می‌شوند و از عملیات شناسایی چنین نمونه‌هایی با عنوان مسئله تشخیص ناهنجاری یاد می‌شود. این مسئله یک بخش حائز اهمیت از زمینه تحقیقاتی داده‌کاوی است چرا که شامل کشف الگو‌های جذاب و نادر در داده‌هاست. این مسئله به طور گسترده در آمار و یادگیری ماشین مورد مطالعه قرار گرفته است و با مترادف‌هایی مانند تشخیص داده پرت، شناسایی نوآوری، تشخیص انحراف و استخراج استثنا نیز یاد می‌شود. تعریف رسمی و مورد قبول این مسئله به صورت زیر است:

"یك ناهنجاری مشاهده‌ای است که به ميزاني از سایر مشاهدات منحرف  
مي‌شود که ظن‌هایي را برای این که توسط مكانيسم متفاوتي توليد شده باشد، ایجاد مي‌کند."

ناهنجاری‌ها جزو پارامترهای مهم هر مجموعه داده‌ای‌ در نظر گرفته مي‌شوند و در دامنه وسیعی از کاربردها تاثیرگذار هستند. به عنوان مثال، الگوی غير معمول ترافيك در یك شبكه کامپیوتری ميتواند به معنای هك شدن رایانه و انتقال داده‌ها به مقصدهای غيرمجاز باشد. رفتار غير عادی در معاملاتی که توسط کارت‌های اعتباری انجام می‌شوند مي‌تواند نشانگر فعاليت‌های اقتصادی با هدف کلاهبرداری باشد. و یا یك ناهنجاری در تصویر MRI ممكن است وجود تومور بدخيم را نشان دهد. تشخيص ناهنجاری به طور گسترده در زمینه‌های کاربردی گوناگونی مانند: پزشكي، بهداشت عمومي، تشخيص کلاهبرداری، تشخيص نفوذ، پردازش تصویر، آسيب‌های صنعتي، شبكه‌های حسگر، رفتار روبات‌ها و داده‌های نجومي بكار گرفته شده است.

با هدف تشخیص نمونه‌های ناهنجار موجود در دادگان دنیای واقعی، تاکنون روش‌های متنوعی مورد استفاده قرار گرفته‌اند. به طور کلی می‌توان روش‌های تشخیص ناهنجاری را به دوسته روش‌های آماری و روش‌های مبتنی بر یادگیری ماشین اشاره کرد. روش‌های آماری اگرچه در برخی از موارد کارایی مناسبی دارند اما عملکرد صحیح و مناسب آن‌ها در گرو صحت پیش‌فرض‌های استفاده شده در همین روش‌هاست و در صورتی که پیش‌فرض‌های اولیه در مورد توزیع داده اشتباه باشد نتایج نهایی ناامید کننده خواهد بود[1]. مزیت اصلی روش‌های مبتنی بر یادگیری ماشین استفاده از تجربه‌های گذشته به منظور انجام پیش‌بینی‌های صحیح در آینده است. این روش‌ها تنها با مشاهده نمونه‌های گذشته به طراحی مدل می‌پردازند و پیش‌فرض خاصی نسبت به توزیع داده ندارند. این دسته از الگوریتم‌ها نیازمند تعداد مناسبی از داده‌ها هستند تا بتوانند مدل پیشنهادی خود را آموزش دهند. از جمله مهم‌ترین مدل‌های موجود در این دسته می‌توان به شبکه‌های عصبی اشاره کرد که سابقه طولانی در زمینه تشخیص ناهنجاری دارد. به عنوان مثال شبکه‌های عصبی رمزگذار[[1]](#footnote-1) و خودرمزگذار[[2]](#footnote-2) مدلی برای بازسازی داده‌های عادی آموزشی آموزش داده می‌شود و نمونه‌های با خطای بازسازی بالا به عنوان نمونه ناهنجار در نظر گرفته می‌شوند[2].

در سال 2017 از شبکه‌های عصبی مولد تخاصمی[[3]](#footnote-3) برای تشخیص ناهنجاری در زمینه تصاویر پزشکی (تصاویر شبکیه) و در مقایسه با سایر روش‌ها به موفقیت قابل توجهی دست یافت[3]. نتایج درخشان شبکه‌های مولد تخاصمی در عرصه پردازش تصویر و استخراج ویژگی سبب محبوبیت آن در زمینه‌های کاربردی مختلف شده است. به طور خاص این دسته از شبکه‌ها به عنوان یک چهارچوب قدرتمند برای مدل‌سازی مجموعه داده‌های پیچیده با ابعاد بالا شناخته می‌شوند. یک از اصلی‌ترین چالش‌های موجود در استفاده از شبکه‌های مولد تخاصمی مقابله با پیچیدگی‌های استنتاج است[4]. در سال‌های اخیر تلاش‌های گسترده‌ای انجام شده است تا با استفاده از شبکه‌های عصبی خود رمزگذار در ساختار شبکه‌های عصبی مولد تخاصمی از پیچیدگی‌های استنتاج کاسته شده و بر چالش‌های موجود غلبه کنند. علیرغم این تلاش‌ها همچنان ضعف‌هایی در روند یادگیری بلوک‌های موجود در ساختار شبکه‌های مولد تخاصمی موجود است و از تمامی ظرفیت موجود به منظور دریافت اطلاعات و آموزش هر چه بهتر مدل استفاده نمی‌شود.

## 1-1- ساختار گزارش

در فصل بعدی ابتدا به دسته‌بندی روش‌های تشخیص ناهنجاری از دیدگاه‌های مختلف می‌پردازیم، معیار‌های ارزیابی مدل‌های تشخیص ناهنجاری را معرفی می‌کنیم و در قسمت انتهایی فصل روش‌های تشخیص ناهنجاری مبتنی بر شبکه‌های مولد تخاصمی را بررسی و مرور می‌کنیم. در این قسمت سعی می‌شود تا ضمن دسته‌بندی روش‌‌های تشخیص ناهنجاری، مروری گذرا بر روش‌های به نسبت قدیمی‌تر نیز انجام شود. در ادامه به طور دقیق‌تر زنجیره‌ای از کارها مورد بحث قرار خواهد گرفت که در طول این زنجیره نقاط ضعف و کمبود‌های مدل پیشنهادی بر طرف می‌شود. فصل سوم به معرفی مدل ‌پیشنهادی و روش نوین تشخیص ناهنجاری اختصاص خواهد داشت. در این فصل الگوریتم پیشنهادی، که بر اساس حل یک مسئله بهینه‌سازی با در نظر گرفتن توزیع توام[[4]](#footnote-4) پارامترهای موجود در ساختار شبکه‌ عصبی مولد تخاصمی طراحی شده است، به شناسایی نمونه‌های ناهنجار در فضای نهفته[[5]](#footnote-5) و فضای ورودی می‌پردازد. در فصل چهارم عملکرد مدل روی دادگان‌های مختلف آزمایش می‌شود. در این فصل با دو نوع مختلف از دادگان روبرو خواهیم بود، دادگانی جدولی[[6]](#footnote-6) که شامل دادگان KDD و ARRHYTHMIA و دادگان تصاویر که شامل CIFAR-10 و SVHN[[7]](#footnote-7) است. در فصل پنجم و نهایی این گزارش پیشنهاد‌های موثر به منظور اصلاح و بهبود احتمالی روند آموزش شبکه‌ مولد تخاصمی ارائه می‌شود و در گام آخر جمع‌بندی مطالب ارائه شده و نتیجه‌گیری نهایی صورت خواهد پذیرفت.

# فصل دوم

**مروری بر کارهای پیشین**

در این فصل ابتدا به دسته‌بندی روش‌های مختلف تشخیص ناهنجاری و مرور روش‌های شاخص هر دسته پرداخته خواهد شد. در گام بعدی معیار‌های ارزیابی مدل‌های تشخیص ناهنجاری معرفی می‌شوند. در ادامه بر روی کارهایی که تاکنون در زمینه تشخیص ناهنجاری با استفاده از شبکه‌های مولد تخاصمی انجام گرفته‌اند مروری خواهیم داشت. در گام اول این قسمت، مقاله پایه با عنوان شبکه‌های مولد تخاصمی[[8]](#footnote-8) مورد بررسی قرار خواهد گرفت در ادامه، کار تشخیص ناهنجاری بدون نظارت با شبکه‌های عصبی تخاصمی به منظور راهنمایی عملیات اکتشاف نشانگر[[9]](#footnote-9) به اختصار AnoGan شرح داده خواهد شد و پس از آن f-AnoGan[[10]](#footnote-10) که در ادامه کار قبلی و توسط همان نویسندگان انجام شده است بررسی خواهد شد. با توجه به ضعف‌های موجود در ساختار f-AnoGan مقاله مکمل این مدل با نام استنتاج یادگرفته شده به روش تخاصمی[[11]](#footnote-11) به اختصار ALI مرور خواهد شد. . در گام بعدی مدل EGBAD[[12]](#footnote-12) که جزو اولین کار‌ها در زمینه تشخیص ناهنجاری که با الهام از مدل ALI خلق شده است مرور می‌شود. در مرحله بعدی با توجه تضمین نشدن شرط سازگاری حلقه[[13]](#footnote-13) در ALI مقاله آلیس[[14]](#footnote-14) مورد اشاره قرار می‌گیرد و در انتها تشخیص ناهنجاری یادگرفته شده به روش تخاصمی[[15]](#footnote-15) که در ادامه کارهای پیشین و همچنین مقاله پایه در این پروژه است به طور دقیق بررسی خواهد شد.

## 1-2- طبقه‌بندی روش‌های تشخیص ناهنجاری از دیدگاه در دسترس بودن برچسب داده

بیشتر روش‌های تشخیص ناهنجاری در مرحله آموزش خود نیاز دارند تا برچسب نمونه‌های مختلف در دسترس آن‌ها باشد تا بتوانند در خصوص طبیعی یا ناهنجار بودن یک نمونه در مرحله آزمایش تصمیم‌گیری کنند. فرایند تهیه و دستیابی به داده‌های دارای برچسب دقیق شامل طیف گسترده‌ای از عملیات‌های بسیار هزینه‌بر و دشوار است، از این‌رو تکنیک‌های تشخیص ناهنجاری را بر اساس میزان در دسترس بودن برچسب‌ها می‌توان به سه دسته: تشخیص ناهنجاری با نظارت، تشخیص ناهنجاری نیمه نظارتی و تشخیص ناهنجاری بدون نظارت تقسیم کرد.

### 1-1-2- تشخیص ناهنجاری با نظارت

در این دسته از روش‌ها هر دو الگوری رفتاری غیرطبیعی و طبیعی مدل ‌می‌شوند. در این مدل‌ها به داده‌های غیر طبیعی برچسب ناهنجاری و به داده‌های طبیعی برچسب عادی می‌زنند. در این رویکرد، برخی از مدل‌ها نمونه‌های ورودی را با نمونه‌های غیرعادی مقایسه می‌کنند و برخی دیگر نمونه‌ها را با نمونه‌های برچسب عادی مقایسه می‌کنند تا بر اساس آن در مورد ماهیت نمونه ورودی تصمیم‌گیری کنند.

### 2-1-2- تشخیص ناهنجاری نیمه‌نظارتی

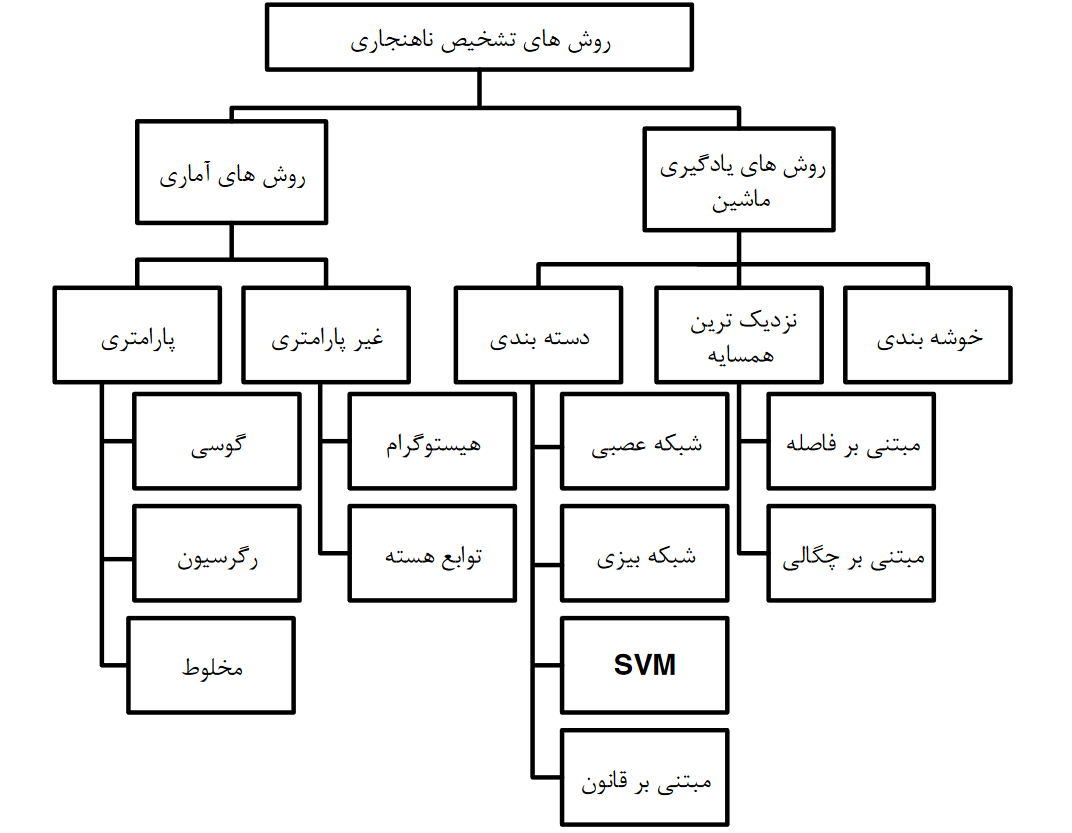
در تشخیص ناهنجاری نیمه‌نظارتی تنها الگوی رفتار طبیعی داده مدل می‌شود و به بیان دیگر تنها به برچسب‌های کلاس عادی نیاز داریم. از نظر کمی این دسته از روش‌ها کاربرد بیشتری نسبت به روش‌های تشخیص ناهنجاری نظارتی دارند.

### 3-1-2- تشخیص ناهنجاری بدون نظارت

اساس کار این دسته از روش‌ها همانند روش‌های خوشه‌بندی[[16]](#footnote-16) است و مدل کلاس داده‌های ناهنجاری را به صورت خودکار از سایر کلاس‌ها تمیز می‌دهد. این روش خوشه‌ای از داده‌ها با رفتار نزدیک به هم پیدا می‌کند و بدین ترتیب عملیات شناسایی ناهنجاری صورت می‌گیرد. اینگونه از مدل‌ها در بسیاری از تشخیص‌ها دچار مشکل می‌شوند چراکه ممکن است نمونه‌های ناهنجار خود باعث ایجاد خوشه‌هایی با الگوی مشابه داده‌های عادی شوند، به همین دلیل تکنیک‌های بدون نظارت در تولید نتایج دقیق کارآمد نیستند و اغلب دارای نرخ مثبت کاذب[[17]](#footnote-17) هستند.

## 2-2- طبقه‌بندی روش‌های تشخیص ناهنجاری از نظر رویکرد حل مسئله

روش‌های تشخیص ناهنجاری به طور کلی به دو دسته روش‌های آماری و روش‌های مبتنی بر یادگیری ماشین تقسیم می‌شوند. روش‌های آماری خود شامل دو دسته پارامتری و غیرپارامتری هستند و روش‌های یادگیری ماشین شامل خوشه‌بندی، نزدیک‌ترین همسایه و دسته‌بندی است. در شکل 2-1 دسته‌بندی این روش‌ها به طور دقیق‌تر به تصویر کشیده شده است. در ادامه هر یک از این روش‌ها بررسی و مرور خواهد شد.



1شکل 2-1: دسته‌بندی روش‌های تشخیص ناهنجاری.

### 1-2-2- روش‌های آماری

تشخیص ناهنجاری با روش‌های آماری به ترتیب بر اساس آمارگان‌های آماری مانند میانگین و انحراف از معیار، توزیع داده‌ها و توابع احتمال (برای ساختن نمایه‌های رفتاری) انجام می‌شود. در اینگونه از روش‌ها بر اساس آزمون‌های آماری هر نوع انحراف از رفتار عادی داده‌ها تشخیص داده می‌شود و داده مورد نظر به عنوان ناهنجاری در نظر گرفته می‌شود. به منظور توسعه مدل‌های اماری در تشخیص ناهنجاری از دو نوع تکنیک پارامتری و غیرپارامتری استفاده می‌شود. تکنیک‌های غیرپارامتری از اطلاعات زمینه‌ای داده‌ها استفاده نمی‌کنند، به بیان دیگر اطلاعی از توزیع داده ورودی ندارند در حالی که روش‌های پارامتری با استفاده از همین اطلاعات مدل را طراحی می‌کنند.

#### 1-1-2-2- روش‌های پارامتری

در روش‌های آماری فرض می‌شود داده‌های واقعی بر اساس پارامتر‌های مشخص از یک توزیع یا تابع خاص تولید می‌شوند، این دسته از روش‌ها خود به سه دسته کلی مدل رگرسیونی، مدل گاوسی و مدل مخلوط تقسیم می‌شوند.

در مدل رگرسیونی داده‌ها بر یک مدل رگرسیونی منطبق می‌شوند و باقی مانده[[18]](#footnote-18) هر داده که بر مدل منطبق نیست به عنوان معیار جهت تشخیص ناهنجاری به کار برده می‌شود.

در مدل گاوسی، فرض بر این است که داده‌ها به توزیع گاوسی تعلق دارند و پارامترهای مدل با استفاده از

استفاده از تخمین بیشینه درستنمایی[[19]](#footnote-19) تعیین می‌شوند. در این مدل‌ها از آزمون‌هایی نظیر آزمون کای-دو[[20]](#footnote-20) جهت شناسایی نمونه ناهنجار استفاده می‌شود.

مدل‌های مخلوط خود ترکیبی از سایر مدل‌های پارامتری هستند. چنین مدل‌هایی در برخی از کاربردها عملکرد بسیار موفقی از خود نشان داده‌اند. به عنوان مثال با به کارگیری یک مدل مخلوط از روش‌های پارامتری برای تشخیص ناهنجاری‌های شبکه، توانسته‌اند در طی زمان بسیار کوتاهی تمام ناهنجاری‌‌های موجود در شبکه که توسط سناریوهای مختلف ایجاد شده بودند را شناسایی کنند[1].

#### 2-1-2-2- روش‌های غیرپارامتری

در این روش از نمونه‌های عادی برای تولید مدل استفاده می‌شود و انحراف نمونه از مدل به عنوان امتیاز ناهنجاری در نظر گرفته می‌شود. این روش را می‌توان به دو دسته مدل‌های مبتنی بر هیستوگرام و مدل‌های مبتنی بر هسته تقسیم کرد.

در مدل‌های مبتنی بر هیستوگرام، هیستوگرام بر اساس تقریب از داده‌های عادی تولید می‌شود و برای اگر نمونه ورودی در محدوده‌های خاصی از هیستوگرام قرار گیرد به عنوان ناهنجاری شناخته می‌شود.

روش مدل‌سازی مبتنی بر هسته[[21]](#footnote-21) یک تابع تشابه بر اساس نمونه‌های موجود از داده استنباط می‌شود. در اینگونه از مدل‌ها در دسترس بودن نمونه‌های کافی به منظور بازنمایی کامل رفتار مجموعه داده ضروریست چراکه در غیر این صورت دقت مدل کاهش می‌یابد.

### 2-2-2- روش‌های یادگیری ماشین

روش‌های مبتنی بر یادگیری ماشین بر اساس تجربه حاصل از مشاهده نمونه‌های قدیمی و به کار گیری آن، ظرفیت تمایز میان رفتار‌های غیر­طبیعی و طبیعی داده تا حد مناسبی بهبود می‌بخشند. این طبقه از روش‌ها خود به سه گروه دسته‌بندی، نزدیک‌ترین همسایه و خوشه‌بندی تقسیم می‌شوند، در ادامه به بررسی هر یک از این دسته‌ها می‌پردازیم.

#### 1-2-2-2- دسته‌بندی

هدف اصلی از روش‌های مبتنی بر دسته‌بندی، اختصاص هر نمونه از داده به یکی از کلاس‌های از پیش‌ تعیین شده بر اساس ویژگی‌های آن نمونه است. از مزیت‌های این دسته از روش‌ها می‌توان به توانایی بالای آن‌ها در تمایز میان کلاس‌های مختلف داده در زمان آزمایش اشاره کرد. از روش‌های متداول که در این دسته جای دارند می‌توان به شبکه‌های بیزی، ماشین بردار پشتیبان[[22]](#footnote-22) ، برخی روش‌های مبتنی بر قانون و شبکه‌های عصبی اشاره کرد.

شبکه­های بیزی در واقع مدل‌های گرافیکی هستند که اتصالات میان نمونه­های مختلف را بر اساس محاسبه احتمال پیشین[[23]](#footnote-23) یک نمونه از داده به همراه دسته‌ای از پیش‌شروط مورد بررسی و ترجمه قرار می‌دهند. اساس کار این دسته از روش‌ها استفاده از یادگیری با نظارت است.

ماشین‌های بردار پشتیبان از دسته الگوریتم‌های یادگیری با نظارت هستند که در صورت استفاده از هسته نمونه‌ها را فضای با ابعاد بالاتر می‌برند و در فضای جدید نمونه‌ها را به دو کلاس تقسیم می‌کنند. استفاده از هسته زمانی توجیه پذیر است که نمونه‌ها در فضای با ابعاد پایین جداپذیر نباشند. این مدل به دلیل استفاده از یک مرز خطی به منظور جداسازی نمونه‌های غیرطبیعی و عادی به عنوان دسته‌بندی خطی شناخته می‌شود.

روش‌های مبتنی بر قانون بر اساس یک سری از قواعد رفتار و عملکرد نمونه‌های عادی را می‌آموزد، بنابراین اگر یک نمونه نتواند از این مجموعه قوانین پیروی کند به عنوان نمونه ناهنجار شناخته خواهد شد. از مطرح‌ترین روش‌هایی که در این دسته می‌گنجند می‌توان به درخت تصمیم[[24]](#footnote-24) اشاره کرد.

شبکه‌های عصبی رفتار سیستم عصبی انسان را تقلید می‌کنند و شامل مجموعه‌ای از فرایندهای بهم پیوسته هستند که به طور همزمان روی داده‌ عمل می‌کنند. در این دسته از روش‌ها از نمونه‌های عادی برای آموزش شبکه عصبی استفاده می‌شود. از نظر دسترسی به برچسب نمونه‌ها شبکه‌های عصبی را می‌توان مشترک بین دسته‌های یادگیری با نظارت و بدون نظارت در نظر گرفت. یکی از انواع این شبکه‌ها که در سال‌های اخیر به موفقیت چشم‌گیری دست یافته است، شبکه‌های مولد تخاصمی[[25]](#footnote-25) هستند. این دسته از شبکه‌ها با موفقیت بر روی داده‌های دنیای واقعی که دارای ابعاد بالا هستند اعمال شده‌اند و چهارچوب یادگیری خصمانه آن‌ها عملکرد مناسبی از خود بر جای گذاشته است. قابلیت این شبکه‌ها نشان‌دهنده ظرفیت آن‌ها برای استفاده در مسئله تشخیص ناهنجاری می‌باشد و به همین جهت به کارگیری شبکه‌های مولد تخاصمی در حوزه تشخیص ناهنجاری اخیرا مورد توجه و کاوش قرار گرفته است[5]. در تشخیص ناهنجاری به کمک شبکه‌های مولد تخاصمی با استفاده از فرایند آموزش تخاصمی رفتار عادی داده مدل می‌شود سپس با اندازه‌گیری امتیاز ناهنجاری روی نمونه‌های مختلف عمل شناسایی نمونه ناهنجار صورت می‌پذیرد. شبکه‌های مولد تخاصمی با کمک آموزش و نمونه‌گیری از مدل‌های مولد به نتایجی بسیار مناسبی در مقایسه با دیگر روش‌ها دست می‌یابند همچنین این مدل‌ها امکان آموزش داده‌های از دست رفته به کمک الگوریتم‌های یادگیری تقویتی[[26]](#footnote-26) را می‌دهد. بیان جزئیات بیشتر در خصوص این دسته از شبکه‌ها را به به بخش‌هایی که در آینده خواهیم داشت موکول می‌کنیم.

#### 2-2-2-2- نزدیک‌ترین همسایه

روش نزدیک‌ترین همسایه مبتنی بر سنجش فاصله یا تراکم میان داده‌هاست به بیان دیگر امتیاز ناهنجاری مقدار همین فاصله است و بسته به مسئله و میزان در دسترس بودن برچسب‌ها این روش می‌تواند به عنوان روش یادگیری بدون نظارت و یا با نظارت به کار گرفته شود.

#### 3-2-2-2- خوشه‌بندی

روش مبتنی بر خوشه‌بندی از دسته روش‌های یادگیری بدون نظارت است که برای شناسایی مجموعه نمونه‌های شبیه به هم به کار برده می‌شود. ناهنجاری‌ها ممکن است تشکیل یک خوشه کوچک بدهند یا در هیچ خوشه‌ای جای نگیرند. این روش در مقایسه با روش‌های مبتنی بر فاصله از پیچیدگی محاسباتی کمتری برخوردار است و در عین حال نقطه ضعف این روش عملکرد نامناسب روی دادگان کوچک است چرا که مدل بینش مناسبی نسبت به داده ندارد و به عنوان مثال برای قسمتی از فضا که برای آن نمونه آموزشی نداریم همواره برچسب ناهنجاری در نظر می‌گیرد در صورتی که ممکن است در حضور تعداد داده کافی برچسب آن نمونه خاص برچسب عادی باشد.

## 3-2- دسته‌بندی بر اساس نحوه تشخیص ناهنجاری

در این قسمت بر اساس نحوه تشخیص ناهنجاری روش‌های موجود را به دسته بر اساس فاصله، دسته‌بندی تک‌کلاسی و بر اساس بازسازی[[27]](#footnote-27) تقسیم می‌کنیم.

### 1-3-2- بر اساس فاصله

از روش‌های مبتنی بر فاصله می‌توان به عنوان یکی از کلاس‌های اصلی روش‌های تشخیص ناهنجاری یاد کرد. در این روش‌ها با استفاده از محاسبه فاصله یک نمونه‌ خاص با نزدیک‌ترین همسایه‌ها و یا نزدیک‌ترین کلاستر، نمونه ناهنجار شناسایی می‌شود. بدیهی‌ست به کار بردن چنین روش‌هایی نیازمند طراحی یا انتخاب معیار فاصله مناسب است.

### 2-3-2- دسته‌بندی تک‌کلاسی

در رویکرد دسته‌بندی تک‌کلاسی تنها نمونه‌های عادی به یک دسته‌بند نظیر SVM آموزش داده می‌شوند، در واقع این مدل‌ها یک مرز تصمیم حول نمونه‌های عادی یاد می‌گیرند، در صورتی که نمونه ورودی داخل این مرز قرار گیرد به عنوان نمونه عادی و در غیر اینصورت به عنوان نمونه ناهنجار شناخته می‌شود.

### 3-3-2- بر اساس بازسازی

این دسته از الگوریتم‌ها بر اساس میزان خطای بازسازی به شناسایی نمونه‌های ناهنجار می‌پردازند. PCA[[28]](#footnote-28) و الگوریتم‌های مشتق از آن جزو همین دسته روش‌ها هستند. اخیرا اساس کار بیشتر کارهای پژوهشی و کاربردی در زمینه تشخیص ناهنجاری شبکه‌های عصبی هستند و به نظر می‌رسد این شبکه‌‌ها دارای سابقه طولانی در این زمینه هستند[6]. به عنوان مثال رویکردهای مبتنی بر خودکدگذار‌ و خودکدگذار متغیر[[29]](#footnote-29) روند بازسازی نمونه‌های عادی را فرامی‌گیرند و نمونه‌های با خطاری بازسازی زیاد را به عنوان ناهنجاری در نظر می‌گیرند. مدل‌های مبتنی بر انرژی و مدل‌های ترکیبی گوسی با خود رمزگذار عمیق[[30]](#footnote-30) به طور خاص در زمینه تشخیص ناهنجاری مورد تحقیق و پژوهش قرار گرفته‌اند. چنین روش‌هایی توزیع داده را با استفاده از خودرمزگذار یا روش‌های مشابه مدل می‌کنند و سپس بر اساس انرژی و یا ترکیب گوسی‌ها یک معیار آماری تشخیص ناهنجاری پدید می‌آورند. در سال‌های اخیر از شبکه‌های مولد تخاصمی به منظور تشخیص ناهنجاری استفاده شده است. در این مدل‌ها به هنگام آزمایش برای هر نمونه ورودی عمل استنتاج انجام می‌شود و با استفاده از انتشار خطای گرادیان نزولی[[31]](#footnote-31) در شبکه مولد، پارامترهای فضای نهفته بازیابی می‌شوند و با استفاده از این پارامترها می‌توان به شناسایی نمونه‌های ناهنجار پرداخت.

## 4-2- معیار‌های ارزیابی روش‌های تشخیص ناهنجاری

صرف نظر از رویکرد به کار گرفته شده، تشخیص ناهنجاری از مرحله یادگیری که در آن با استفاده از نمونه‌های آموزشی مدل آموزش داده می‌شود، آغاز می‌شود. پس از اتمام مرحله یادگیری، مدل دسته‌بندی نمونه‌هایی که تاکنون آن‌ها را مشاهده نکرده است را بر اساس معیارهای مورد نظر انجام می‌دهد. به منظور ارزیابی روش‌های مختلف تشخیص ناهنجاری معیار‌های استانداردی نظیر نرخ تشخیص[[32]](#footnote-32)، دقت[[33]](#footnote-33)، کارایی[[34]](#footnote-34) و مقیاس پذیری[[35]](#footnote-35) مورد استفاده قرار می‌گیرد.

### 1-4-2- نرخ تشخیص

نرخ تشخیص یکی از مقبول‌ترین معیارها برای تشخیص ناهنجاری است. نتیجه ارزیابی در میزان تشخیص ناهنجاری‌ها می‌تواند در چهار دسته مختلف گزارش شود که عبارتند از مثبت صحیح[[36]](#footnote-36)، منفی صحیح[[37]](#footnote-37)، مثبت کاذب و منفی کاذب[[38]](#footnote-38). معیار معمول مورد استفاده در میزان تشخیص مثبت صحیح و منفی صحیح است که در آن نسبت مواردی که توسط آشکارساز دسته‌بندی می‌شود تعریف می‌شود. مثبت کاذب معیار مهم دیگری است که به مواردی که به اشتباه به عنوان ناهنجاری دسته‌بندی می‌شوند اشاره دارد. نرخ منفی کاذب، نسبت موارد غیرطبیعی است که به طور نادرست به عنوان داده عادی دسته‌بندی شده‌اند.

### 2-4-2- دقت

این معیار به تمام موارد دسته‌بندی شده عادی و غیر عادی نسبت به تمامی داده‌‎‌ها اشاره دارد. معیار صحت[[39]](#footnote-39) نیز یک معیار کاربردی است که مشخص می‌کند چه تعداد از داده‌های ناهنجار که توسط مدل شناسایی شده‌اند به واقع داده ناهنجار هستند.

### 3-4-2- کارایی

برای محاسبه معیار کارایی از نمودار مشخصه عملکرد[[40]](#footnote-40) استفاده می‌شود. این مشخصه با ترسیم نرخ مثبت صحیح بر اساس نرخ مثبت کاذب به ازای مقادیر آستانه مختلف رسم می‌شود. از حد آستانه مورد نظر به عنوان نقطه برش در تعیین هویت نمونه به عنوان داده ناهنجار و یا غیر ناهنجار استفاده می‌شود.

### 4-4-2- مقیاس‌پذیری

این معیار به منظور تعیین میزان توانایی مدل در تشخیص ناهنجاری به هنگام افزایش اندازه مجموعه داده تعریف شده است. این معیار با هدف ایجاد اطمینان از توانایی مدل درکنترل تغییرات سریع حجم داده ایجاد شده است و برای محاسبه آن با توجه به جنس مجموعه داده مورد آزمایش روش‌های متفاوتی ارائه شده است[1].

## 5-2- شبکه‌های مولد تخاصمی و تشخیص ناهنجاری

هدف از مدل‌های یادگیری عمیق، کشف مدل‌های سلسله‌مراتبی قوی است. این مدل‌ها نشان‌دهنده توزیع احتمال انواع داده‌هایی است که در کاربردهای هوش مصنوعی نظیر تصاویر طبیعی، شکل موج صوتی حاوی گفتار به کار می‌رود. برجسته‌ترین موفقیت یادگیری عمیق در طراحی مدل‌های تمایزگر[[41]](#footnote-41) بوده است. این مدل‌ها قادرند تا ورودی با ابعاد بالا را دریافت کنند و کلاس هر یک از نمونه‌ها را با قدرت تشخیص خود با دقت بالا مشخص کنند. استفاده از الگوریتم‌های پس‌انتشار[[42]](#footnote-42) ، حذف تصادفی[[43]](#footnote-43) و واحدهای خطی تکه‌ای[[44]](#footnote-44) که دارای گرادیان با رفتار مناسب هستند عامل موفقیت چشم‌گیر یادگیری عمیق است.

استفاده کاربردی از الگوریتم‌هایی نظیر تخمین بیشینه درستنمایی و الگوریتم‌های مرتبط با آن همراه با چالش‌ها و دشواری‌های زیادی نظیر محاسبات احتمالاتی زیاد و خارج از کنترل است، وجود این چالش و همچنین سختی‌های موجود در استفاده از واحد‌های خطی تکه‌ای در حوزه مدل‌های مولد سبب شده است تا مدل‌های مولد عمیق کمتر مورد توجه قرار گیرند. شبکه‌های مولد تخاصمی بر این دست از چالش‌ها و دشواری‌ها فائق آمده و نقش مدل‌های مولد عمیق را پررنگ‌تر ساخته است.

### 1-5-2- شبکه‌های مولد تخاصمی

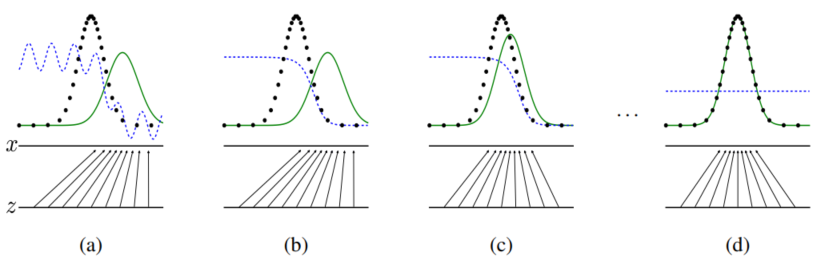
در شبكه‌های مولد تخاصمی، یک مدل مولد در برابر یك مدل تمایزگر قرار مي‌گير، مدل تمایزگر سعی می‌کند میان داده‌های واقعی و داده‌های تولیدی توسط شبکه مولد تمایز ایجاد کند. مدل مولد را مي‌توان مانند تيمي فرض کرد که سعي در توليد ارز جعلي با میزان شباهت بسیار بالا به ارز واقعی دارد و در طرف مقابل مدل تمایزگر مشابه پليس است که سعي در کشف ارز تقلبي دارد. رقابت در این بازی، هر دو تيم را به سمت بهبود روش های خود سوق مي‌دهد تا اینکه ارز تقلبي از ارز اصلي غيرقابل تشخيص باشد. این چارچوب مي‌تواند الگوریتم‌های آموزشي خاصي را برای انواع مختلف از مسائل و مدل‌ها بهينه‌سازی کند. بخش مولد با دریافت نویز تصادفي، از طریق پرسپترون چند لایه نمونه‌هایی با توزیع مشابه داده اصلی توليد مي‌کند و مدل تمایزگر با استفاده از مدل پرسپترونی چند لایه تلاش می‌کند تا میان نمونه‌‌های مختلف تمایز ایجاد کند. در این نوع از تعریف شبكه مي‌توان هر دو مدل را با استفاده از الگوریتم‌های پس‌انتشار و حذف تصادفي ایجاد کرد و برای نمونه‌گيری از مدل مولد تنها از الگوریتم انتشار رو به جلو[[45]](#footnote-45) استفاده کرد و در نتيجه به كارگيری هيچ الگوریتمی نظیر استنتاج تقریبي و یا زنجيره مارکوف ضروری نيست.

در شبكه مولد تخاصمی به طور همزمان دو مدل مولد و تمایزگر آموزش داده مي‌شود. مدل مولد توزیع داده را ضبط مي‌کند و مدل تمایزگر که احتمال این که نمونه از داده‌های توليد شده توسط باشد را تخمين مي‌زند. تابع هدف برای شبكه مولد به حداکثر رساندن احتمال اشتباه شبكه *D* است. این بستر منجر به یك بازی دو نفره مانند بازی‌های بیشینه-کمینه[[46]](#footnote-46) مي­شود. در فضای توابع دلخواه و یك راه حل منحصر به فرد وجود دارد و در این راه حل شبكه مولد توزیع داده‌های آموزشي را یاد گرفته است و شبكه تمایزگر احتمال را در همه جا برابر مقدار 2/1 نشان می‌دهد، به بیان دیگر میان داده‌های واقعی و داده تولید شده توسط شبکه مولد نمی‌تواند تمیز دهد.

مدل­سازی چارچوب تخاصمی به وسیله ایجاد یک مدل چند لایه پرسپترونی برای هر دو مدل مولد و تمایزگر انجام می‌شود. برای یادگيری توزیع مولد روی داده ، یك تابع نویز خالص را به عنوان ورودی تعریف مي­کنيم، سپس یك نگاشت از فضای نهفته به فضای داده را به عنوان نشان مي‌دهيم، در اینجا یك تابع مشتق‌پذیر است که توسط یك پرسپترون چند لایه با پارامترهای نمایش داده مي­شود. همچنين برای شبكه تمایزگر یك پرسپترون چند لایه لایه با یك خروجي اسكالر[[47]](#footnote-47) تعریف مي­کنيم. بيانگر احتمال این است که x از داده‌های اصلي به جای توزیع ناشی شده باشد. در این میان به شبكه آموزش داده مي‌شود تا احتمال تخصيص برچسب صحيح را به داده‌های واقعی و نمونه‌های توليدی از را به حداکثر برساند. به طور همزمان به شبكه آموزش داده مي‌شود تا تابع هدف را به حداقل برساند. به عبارت دیگر، شبكه‌های و بازی کمینه-بیشینه دو نفره زیر با تابع مطابق معادله 2-1 انجام می‌دهند.

|  |  |
| --- | --- |
| (1-2) |  |

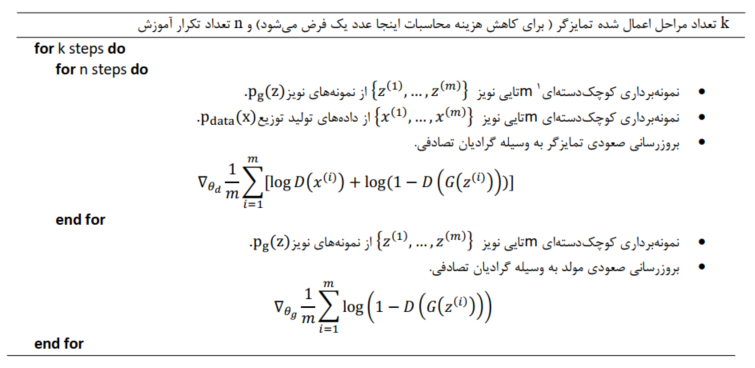
با تحليل نظری صورت گرفته بر روی شبكه‌های مولد تخاصمی، نشان داده شده است که این شبكه‌ها پتانسيل کافي برای بازیابي توزیع داده‌های اصلي را در قالب شبكه مولد دارند. در عمل معادله 2-1 ممکن است گرادیان کافی برای آموزش شبکه را فراهم نکند. در اوایل یادگیری، هنگامی که شبکه مولد ضعیف است، شبکه تمایزگر می‌تواند با اطمینان بالا نمونه‌های غیرواقعی را شناسایی کند چرا که نمونه‌های تولیدی با نمونه‌های آموزشی کاملا متفاوت هستند. در این حالت شباع می‌شود. در این حالت می‌توان به جای آموزش برای به حداقل رساندن تابع می‌توانیم را برای به حداکثر رساندن آموزش دهیم. این تابع در همان نقطه ثابت و قرار دارد اما گرادیان قوی‌تری در یادگیری فراهم می‌کند. رویكرد کلي شبکه‌های مولد تخاصمی در شكل 2-2 نشان داده شده است.



2شکل 2-2: رویکرد کلی شبکه‌های مولد تخاصمی[5].

همانطور که در شكل 2-2 مشاهده‌ می‌کنید توزیع تمایزگر (خط آبي شكسته) به‌روز‌رسانی می‌شود تا بتواند نمونه‌های توزیع داده‌های اصلي (خط مشكي نقطه‌چي) از داده‌های توليدشده توسط توزیع مولد )خط سبز پيوسته) تمیز دهد. خط افقي پایين بیانگر فضای نهفته است که متغیر z با توزیع یکنواخت از آن نمونه‌برداری شده است. خط افقی بالا بخشی از فضای داده واقعی x است. فلش‌های رو به بالا نشان می‌دهد که چگونه تابع ، z را به طور غیریکنواخت به x نگاشت می‌کند. به مرور زمان در مناطق چگال‌تر منقبض می‌شود و در مناطق با چگالی کمتر باز می‌شود. در قسمت (a) شکل 2-2 دو بلوک تخاصمی نزدیک همگرایی هستند،یعنی به نزدیک شده است و همچنین دسته‌بند تا حدی دقیق می‌باشد. در قسمت (b) در حلقه داخلی الگوریتم، آموزش می‌بیند تا بتواند نمونه‌های غیرواقعی را تشخیص دهد و به  همگرا شود. در (c) پس از به روزرسانی گرادیان ناشی از ، را به گونه‌ای هدایت می‌کند که به سمت مناطقی مایل شود که به توزیع داده واقعی نزدیک‌تر شود. در قسمت آخر (d) پس از انجام چند مرحله از آموزش اگر و ظرفیت کافی را داشته‌ باشند به نقطه تعادلی می‌رسند که در آن . و تمایزگر دیگر قادر نیست میان توزیع داده‌های واقعی و غیرواقعی تفاوتی قائل شود.

به منظور پیاده‌سازی این شبکه‌ها از روش عددی مبتنی بر تکرار استفاده می‌شود. تکمیل بهینه‌سازی شبکه در حلقه داخلی مرحله آموزش همراه با چالش‌هایی نظیر هزینه محاسباتی زیاد است و همچنین روی دادگان کوچک منجر به بیش‌برازش[[48]](#footnote-48) است. حال به جای اینکه در هر تکرار هر دو شبکه و بهینه شوند، به ازای k مرحله بهینه کردن یک مرحله بهینه می‌شود. با این کار تا زمانی که به اندازه کافی آهسته تغییر کند در نزدیکی نقطه بهینه باقی خواهد ماند. شبه کد الگوریتم شبکه‌های مولد تخاصمی در الگوریتم 2-1 آورده شده است.



3الگوریتم 2-1: آموزش گرادیان نزولی کوچک دسته‌ای[[49]](#footnote-49) شبکه‌های مولد تخاصمی.

#### 1-1-5-2- تحلیل نظری شبکه مولد تخاصمی

شبکه مولد به طور ضمنی یک تابع توزیع احتمال را به عنوان توزیع نمونه‌های تعریف کرده است(توجه داشته باشید z). در صورتی که ظرفیت مناسب و زمان کافی برای آموزش در اختیار الگوریتم 2-1 قرار گیرد در نهایت این الگوریتم توزیع داده ورودی را خواهد یافت. نتایج حاصل از شبکه‌های عصبی مولد تخاصمی بر اساس تنظیمات غیرپارامتری به دست آمده است. در این بخش نشان می‌دهیم که در بازی کمینه-بیشینه بین دو شبکه یک بهینه عمومی است. برای این منظور ابتدا تمایزگر بهینه را برای هر مولد در می‌گیریم.

**قضیه 2-1-** برای هر تابع مولد ثابت، تابع تمایزگر بهینه عبارت است از:

|  |  |
| --- | --- |
| (2-2) |  |

**اثبات:** معیار آموزش برای تمایزگر، با توجه به هر مولد ، به حداکثر رساندن مقدار می­باشد، پس داریم:

|  |  |
| --- | --- |
| (3-2) |  |

از طرفی می­دانیم برای هر،تابع دربازه در بیشینه است. هم­چنین می­دانیم تمایزگر نیاز به تعریف بیرون از مرز ندارد، پس در نتیجه نقطه بهینه برای به حداکثر رساندن می­باشد.

می­توان هدف از آموزش شبکه را به حداکثر رساندن لگاریتم درست­نمایی[[50]](#footnote-50) احتمال تعبیر کرد، که بیانگر آن است که هرجا از توزیع باشد و هر جا از توزیع باشد است. با این تعریف بازی کمینه-بیشینه در معادله ‏3‑2 را می­توان به­صورت زیر، بازنویسی کرد:

|  |  |
| --- | --- |
| (4-2) |  |

**قضیه 2-2-** کمینه سراسری تنها در حالتی قابل محاسبه است که اگر و تنها اگر و مقدار برابر با باشد.

**اثبات:** طبق قضیه 2-1، می­دانیم هنگامی­که باشد ، می­شود. در ادامه نشان دادیم که حال برای اثبات این­که بهترین مقدار این مقدار است، این عبارت را از تعریف کم می­کنیم. طبق همگرایی جنسن-شانون[[51]](#footnote-51) داریم:

|  |  |
| --- | --- |
| (5-2) |  |

می­دانیم تابع همگرایی جنسن-شانون بین دو توزیع همواره نامنفی است و صفر است اگر و تنها اگر دو توزیع برابر باشند. پس ما نشان دادیم مقدار بهینه برابر با است و نقطه بهینه هنگام برابری دو توزیع رخ می­دهد. بدین ترتیب اثبات این قضیه نیز به پایان رسید.

#### 2-1-5-2- مزایا و معایب

شبکه­های مولد تخاصمی نسبت به مدل­های قبلی دارای مزایا و معایبی می­باشد. این معایب در درجه اول این است که نمایش صریح وجود ندارد و باید در حین آموزش به خوبی با هماهنگ شود به طور خاص، نباید بدون به‌روزرسانی خیلی زیاد آموزش داده شود که در آن مقدار مقادیر زیادی z به همان مقدار x بدست می آید تا از تنوع کافی برای مدل سازی برخوردار باشد، در هنگام یادگیری نیازی به استنباط نیست و می توان طیف گسترده­ای از توابع را در مدل گنجانید. مزایای ذکر شده در بالا در درجه اول محاسباتی است. مدل­های تخاصمی همچنین ممکن است برخی از مزیت­های آماری را از شبکه مولد به دست آورند که مستقیماً با نمونه داده‌ها به روز نمی­شوند، اما فقط با گرادیان­هایی که از طریق تمایزگر جریان می­یابند، بروزرسانی می­شود.

این بدان معنی است که اجزای ورودی مستقیماً در پارامترهای مولد کپی نمی­شوند. یکی دیگر از مزیت­های شبکه‌های تخاصمی این است که آنها می­توانند توزیع­های بسیار تیز و حتی تخریب کننده را نشان دهند، در حالی که روش­های مبتنی بر زنجیره‌های مارکوف نیاز دارند که توزیع تا حدی مبهم باشد تا زنجیرها بتوانند میان حالت­ها مخلوط شوند.

### 2-5-2- مدل ANOGAN

مدل ANOGAN به منظور مدل‌سازی در زمینه پزشکی و به طور خاص مدل‌سازی وضعیت سلامت موضعی آناتومی طراحی شده است. این مدل از دسته الگوریتم‌های مولد و بدون‌نظارت است. با به وجود آمدن این مدل توانایی شبکه‌های مولد تخاصمی در ایجاد یک مدل با قدرت بازنمایی بالا در تشریح وضعیت آناتومی ثابت شد. لازم به ذکر است مدل پیشنهادی شبکه تمایزگر و شبکه مولد را به طور همزمان آموزش می‌دهد و با استفاده از هر دو شبکه مشخص می‌کند داده ورودی از جنس داده‌های آموزشی است و یا باید به عنوان ناهنجاری دسته‌بندی شود.

به منظور تشخیص ناهنجاری مدل مورد نظر بازنمایی نمونه‌های متنوع آناتومیکی طبیعی را می‌آموزد. در این کار به جای استفاده از بهینه­سازی تابع هزینه واحد، از تعادل نش[[52]](#footnote-52) میان هزینه‌ها استفاده شده است که سبب افزایش قدرت بازنمایی و رشد نرخ منفی صحیح[[53]](#footnote-53) به اختصار TNR مدل تولیدی، بهبود روند نگاشت ویژگی[[54]](#footnote-54) و همچنین دستیابی به دقت بالا در طبقه­بندی داده­های واقعی از داده­های غیرواقعی می­شود. در ادامه چگونگی طراحی مدل و نحوه شناسایی وضعیت و ظواهری که در داده‌های آموزش دیده نشده‌اند شرح داده خواهد شد.

#### 1-2-5-2- یادگیری بدون نظارت تنوع آناتومیکی طبیعی

مجموعه­ای از تصاویر پزشکی است که هر نمونه آن نمایانگر نمونه‌هایی از آناتومی‌های سالم است و با نمایش داده می­شود که ، در اینجا است یعنی اندازه یک تصویر برابر است. از هر تصویر ، تکه­ تصویر دو بعدی ، با ابعاد بطور تصادفی از موقعیت­های مختلف نمونه­گیری می­کنیم که منجر به داده­های ، می شود. در طول آموزش، فقط را در اختیار داریم و برای یادگیری توزیع حاشیه­ای، که نشان دهنده تنوع تصاویر آموزش است، از یک روش بدون نظارت استفاده می­شود. برای آزمایش، داریم ، که تصاویر مشاهد نشده با ابعاد استخراج شده از داده است و آرایه­ای از برچسب­های حقیقی مبتنی بر تصویر باینری با است. این برچسب­ها فقط در طول آزمایش استفاده می­شوند، تا کارایی روش تشخیص ناهنجاری ارزیابی شود.

شبکه مولد توزیع را روی داده از طریق نگاشت نمونه‌های z توسط تابع آموزش می­بیند؛ در واقع بردارهای تک بعدی با توزیع یکنواخت از فضای نهفته نمونه­برداری می­‌شوند و به فضای دو بعدی تصویر که در آن تصاویر آناتومی سالم وجود دارند نگاشت می‌شوند. در این تنظیمات، معماری شبکه مولد معادل یک رمزگذار پیچشی[[55]](#footnote-55) که از پشته­های پیچشی استفاده می­کند، در نظر گرفته می­شود. تمایزگر یک CNN استاندارد است که یک تصویر دو بعدی را به یک مقدار نگاشت می­کند. مقدار خروجی تمایزگر بیانگر احتمال این است که ورودی تمایزگر، از فضای تصاویر واقعی یعنی فضای نمونه‌های آموزشی نمونه­برداری شده باشد و یا توسط شبکه مولد تولید شده باشد. و به طور همزمان از طریق بازی کمینه-بیشینه با تابع و معادله ‏6‑2 بهینه­سازی می­شوند[3].

|  |  |
| --- | --- |
| (6-2) |  |

در این بازی شبکه تمایزگر آموزش می­بیند که احتمال اختصاص نمونه­های واقعی را بیشینه و نمونه­های تولیدی از با برچسب جعلی را کمینه کند. هم­چنین شبکه مولد آموزش می­بیند همزمان با حداقل کردن که معادل با حداکثرکردن است، شبکه تمایزگر را فریب دهد. به طور کلی در طول آموزش تخاصمی، مولد در تولید تصاویر واقع بینانه و تمایزگر در شناسایی صحیح تصاویر واقعی و تولید شده بهبود می­یابد.

#### 2-2-5-2- نگاشت تصاویر جدید به فضای نهفته

وقتی آموزش تخاصمی به پایان رسید، شبکه مولد یاد می­گیرد که را از فضای نهفته با نمایش z به تصویر واقعی (عادی) xنگاشت کند. شبکه­های GAN­ به­طور خودکار نگاشت معکوس را انجام نمی­دهد. فضای نهفته دارای گذار خطی است، بنابراین نمونه­گیری از دو نقطه نزدیک بهم در فضای نهفته، دو تصویر مشابه بصری نیز ایجاد می­کند.

با فرض اینکه تصویر را برای بررسی داریم، هدف این است که یک نقطه را در فضای پنهان پیدا کنیم که مطابق با تصویر باشد و از نظر بصری در آن نقطه شبیه به تصویر باشد و در توزیع حاشیه­ای قرار داشته باشد. میزان شباهت و بستگی به این دارد که چه تصویری از توزیع داده برای آموزش مولد استفاده می‌شود. برای پیدا کردن بهترین ، با نمونه­گیری تصادفی از توزیع فضای نهفته شروع می‌کنیم و آن را به شبکه مولد آموزش دیده، برای تولید تصویر اعمال می­کنیم. سپس بر اساس تصویر ایجاد شده یک تابع اتلاف تعریف می­کنیم، که گرادیان به روزرسانی ضرایب را فراهم می­کند و در نتیجه یک موقعیت بروز شده در فضای نهفته بدست می­آید. به عبارتی برای پیدا کردن شبیه­ترین تصویر ، نقطه در فضای نهفته در یک فرآیند تکراری از طریق با مراحل پس­انتشار بهینه می­شود.

تعریف تابع اتلاف برای نگاشت از تصاویر فضای نهفته شامل دو بخش است[7]، باقی­مانده اتلاف[[56]](#footnote-56) و باقی­مانده تمایز[[57]](#footnote-57). باقی­مانده اتلاف شباهت بصری بین تصویر تولید شده و تصویر مورد بررسی را تقویت می­کند. باقی­مانده تمایز، تصویر تولید شده را در حاشیه توزیع آموزش دیده قرار می­دهد. بنابراین، هر دو مؤلفه GAN آموزش دیده، تمایزگر و مولد ، برای یافتن ضرایب از طریق پس­انتشار مورد استفاده قرار می­گیرند.

باقی‌مانده اتلاف معیار عدم شباهت بصری بین تصویر مورد بررسی و تصویر تولید شده در فضای تصویر اندازه­گیری می­کند و به­صورت معادله ‏7‑2 تعریف می­شود.

|  |  |
| --- | --- |
|  | (7-2) |

با فرض یک مولد کامل و یک نگاشت کامل فضای نهفته، برای یک مورد بررسی ایده­آل، تصاویر و یکسان هستند. در این حالت باقی­مانده تلفات برابر با صفر است.

باقی‌مانده تمایز برای تأمین تصویر، باقی­مانده تمایز را بر اساس خروجی تمایزگر با اعمال تصویر تولید شده در تمایزگر محاسبه می­کنیم، در این­ رابطه آنتروپی متقاطع سیگموئید[[58]](#footnote-58)، که از تمایز اتلاف تصاویر واقعی حین آموزش متخاصم، با تابع لاجیت[[59]](#footnote-59) و تعریف می­شود.

#### 3-2-5-2- تشخیص ناهنجاری­

در طی شناسایی ناهنجاری­ها در داده­ی جدید، ابتدا نمونه مورد بررسی جدید را به عنوان یک تصویر طبیعی یا غیر عادی ارزیابی می­کنیم. تابع اتلافی که برای نگاشت به فضای نهفته مورد استفاده قرار می­گیرد، در هر تکرار بروزرسانی می­شود و سازگاری تصاویر تولید شده با تصاویر را که در طول آموزش متخاصم مشاهده می­شود ارزیابی می­کنیم. بنابراین، یک نمره ناهنجاری که تناسب تصویر مورد جستجو را با مدل تصاویر عادی بیان می­کند، این معیار می­تواند مستقیماً از تابع اتلاف معادله ‏8‑2 بدست آید.

|  |  |
| --- | --- |
|  | (8-2) |

که در آن به ترتیب امتیاز باقی­مانده و امتیاز تمایزگر است که توسط باقی­مانده اتلاف و باقی­مانده تمایز در فضای نهفته تعریف می­شود. این مدل، نمره ناهنجاری بزرگی برای تصاویر غیر عادی بدست می آورد و یک نمره ناهنجاری کوچک بدین معنی است که این تصویر بسیار مشابه تصاویر قبلاً در طول آموزش دیده شده، می­باشد. برای تشخیص ناهنجاری مبتنی بر تصویر، از نمره ناهنجاری استفاده می­شود. علاوه بر این، از تصویر باقی­مانده برای شناسایی مناطق غیر عادی در یک تصویر استفاده می‌شود. برای مقایسه، علاوه بر این امتیاز ناهنجاری را مطابق معادله 2-9 تعریف می­شود، در اینجا همان امتیاز تمایزگر و امتیاز باقی­مانده می­باشد.

|  |  |
| --- | --- |
|  | (9-2) |

#### 4-2-5-2- مزایا و معایب­

علیرغم نتایج قابل قبول این مدل در تشخیص ناهنجاری در تصاویر پزشکی، همانطور که مشاهده کردیم، برای بررسی تصویر ، باید نقطه متناظر با آن در فضای نهفته پیدا شود به طوری که در آن نقطه از نظر بصری بیشتر شبیه به تصویر باشد. برای پیدا کردن بهترین ، باید با نمونه­گیری تصادفی از توزیع فضای نهفته شروع کنیم و آن را به شبکه مولد آموزش بدهیم و برای تولید تصویر اعمال می­کنیم. سپس بر اساس تصویر ایجاد شده می‌بایست یک تابع اتلاف تعریف کنیم، که گرادیان به روزرسانی ضرایب را فراهم کند و در نتیجه یک موقعیت بروز شده در فضای نهفته بدست آورد. به عبارتی برای پیدا کردن شبیه­ترین تصویر ، نقطه در فضای نهفته در یک فرآیند تکراری از طریق با مراحل پس­انتشار بهینه و پیدا می‌شود. این فرایند تکراری مبتنی بر تصادف است و هزینه و پیچیدگی محاسباتی زیادی به مدل تحمیل می‌کند. با توجه به وجود این مشکل می‌توان مکانیزمی طراحی کرد که مدل پیشنهادی فرایند نگاشت معکوس تصاویر در حین آموزش مدل فرابگیرد.

### 3-5-2- مدل f-AnoGan

مدل f-AnoGan در ادامه کار قبلی و توسط همان نویسندگان در سال 2019 ارائه شد. در مدلAnoGan از شبکه‌های عصبی عمیق کانولوشنی مولد تخاصمی[[60]](#footnote-60) به اختصار DCGAN برای آموزش بدون نظارت شبکه مولد و تمایزگر استفاده شده است. در مدل قبلی برای شناسایی نقطه متناسب با تصویر ورودی بر اساس الگوریتم پس‌انتشار از یک فرایند مبتنی بر تکرار استفاده می‌شد. به هنگام شناسایی ناهنجاری در کاربردهای دنیای واقعی این فرایند مبتنی بر تکراراز نظر پیچیدگی زمانی مشکلاتی را از نظر زمان ایجاد می‌کند.f-AnoGan فرایند مبتنی بر تکرار مورد نظر را با یادگیری یک نگاشت از فضای اصلی به فضای نهفته جایگزین می‌کند. علاوه بر این در ساختار مدل جدید از WGAN[[61]](#footnote-61) به جای DCGAN استفاده شده است.

چهارچوب ارائه شده در این کار شامل دو گام آموزشی روی تصاویر عادی است، در گام اول شبکه مولد تخاصمی آموزش می‌بیند و در گام بعدی بر اساس شبکه مولد آموزش دیده شده کدگذار آموزش می‌بیند. پس از آموزش قسمت‌های مختلف مدل بر اساس استنتاج برای هر تصویر یک امتیاز ناهنجاری محاسبه می‌شود. همانند کار قبلی شبکه مولد روی تصاویر عادی آموزش داده می‌شود و بازنمایی‌های تنوع آناتومی‌های سالم در فضای نهفته به دست می‌آید. علاوه بر این کدگذار نیز نگاشت تصاویر به فضای نهفته را آموزش می‌بیند. در ادامه در سه بخش به بررسی شبکه‌های تمایزگر و مولد، کدگذار و نحوه امتیاز‌دهی به داده‌ها به منظور تشخیص داده‌های ناهنجار می‌پردازیم.

#### 1-3-5-2- GAN یادگیری بدون نظارت تنوع آناتومیکی طبیعی

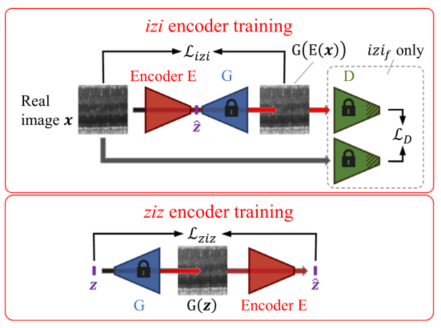
داده‌های آموزشی در این مدل به شکل نمایش داده می‌شود. در این نوع از نمایش و است. به بیان دیگر تعداد تصویر پزشکی داریم و از هر تصویر *به صورت تصادفی از مناطق مختلف تصویر تعداد نمونه با ابعاد انتخاب می‌شود. توجه شود*  *و*  و *.* برای ارزیابی مدل از داده تصویر با ابعاد  که از میان دادگان آزمایش انتخاب شده است، حاشیه نویسی پیکسل مربوطه به صورتو برچسب مربوطه استفاده می‌شود. این مجموعه داده آزمایش به صورت است و به طور همزمان شامل داده عادی و ناهنجار است.

برای یادگیری تنوع آناتومی موجود در تصاویر از WGAN استفاده می‌شود. این شبکه نگاشت غیرخطی از فضای نهفته *به فضای ورودی را یاد می‌گیرد. مانند مدل‌های دیگر در این مدل نیز شبکه مولد و تمایزگر به طور همزمان بهینه می‌شوند. در ابتدا از فضای با ابعاد d نمونه نویز انتخاب می‌کنیم. در طول فرایند آموزش شبکه مولد تلاش می‌کند تا توزیع شبکه مولد یعنی را تا حد امکان به توزیع داده ورودی یعنی نزدیک کند و در نتیجه بتواند شبکه تمایزگر را به گونه‌ای فریب دهد که نتواند تشخیص دهد از توزیع داده واقعی است و یا توسط شبکه مولد ایجاد شده است. در پایان فرایند آموزش شبکه مولد توزیع داده‌های آموزشی را یاد گرفته است و شبکه تمایزگر می‌تواند تخمین بزند داده تولیدی توسط مولد تا چه اندازه به توزیع داده واقعی نزدیک است.*

#### 2-3-5-2- یادگیری نگاشت سریع از فضای تصویر به فضای نهفته

شبکه مولد GAN را می‌توان به صورت *نمایش داد. این نمایش به توانایی نگاشت شبکه مولد از فضای نهفته به فضای داده ورودی اشاره دارد. در روند آموزشی* GAN *هیچ نگاشتی از فضای داده ورودی به فضای نهفته آموزش داده نمی‌شود. در مدل پیشنهادی این مقاله نگاشت معکوس به صورت* نمایش داده می‌شود و این نگاشت توسط یک کدگذار آموزش دیده می‌شود. آموزش این کدگذار با دو معماری مختلف قابل پیاده‌سازی است، روش اول z-image-z و روش دوم image-z-image نام دارد. روش اول که به صورت خلاصه ziz و روش دوم به طور خلاصه izi نامیده می‌شود. در هر دو معماری از خودکدگذار‌های کانولوشنی استفاده می‌شود. از کدگذار *E* برای نگاشت معکوس استفاده می‌شود و از شبکه مولد که در واقع یک WGAN با اوزان یادگرفته‌شده ثابت است به عنوان کدگشا استفاده می‌شود. تفاوت دو روش فوق در ترتیب استفاده از کدگذار و کدگشاست.

هنگام آموزش کدگذار تنها پارامترهای کدگذار بهینه می‌شوند و پارامترهای شبکه مولد ثابت هستند. معماری‌های مختلف جهت آموزش کدگذار در شکل 2-3 قابل مشاهد است.



4شکل 2-3: 1) آموزش کدگذار به روش izi: در این روش خطای *بر اساس باقی‌مانده[[62]](#footnote-62)* *از تفاوت تصاویر ورودی واقعی و تصویر بازسازی شده بهینه می‌شوند کد گذار آموزش می‌بیند. 2) آموزش کدگذار به روش : در این روش به صورت توام خطای*  که همان خطای باقی‌مانده میان تصویر ورودی واقعی و تصویر بازسازی شده است به همراه خطای *بر اساس خطای باقی‌مانده روی ویژگی‌های شبکه تمایزگر است بهینه می‌شود. 3) آموزش کدگذار به روش* : در این روش خطای باقی‌مانده میان نمونه‌های تصادفی و موقعیت‌های موجود در فضای نهفته بهینه می‌شود.

الگوریتم ziz با معکوس کردن ترتیب کدگذار و کدگشا در ساختار معمول یک خودکدگذار به وجود می‌آید. در هنگام آموزش یک نمونه از فضای نهفته انتخاب می‌شود و با استفاده از شبکه مولد که وزن‌های آن ثابت نگاه داشته شده است به فضای داده واقعی نگاشت می‌شود و کدگذار تلاش می‌کند تا معکوس این نگاشت به فضای نهفته را یاد بگیرد، بنابراین در این روش به هیچ تصویر واقعی یا به بیان دیگر به هیچ نمونه‌ای از فضای داده واقعی نیاز نیست. در واقع در معماری ziz ساختار یک کدگذار از فضای نهفته به فضای نهفته است و در همین حال نگاشت مورد نیاز از فضای نهفته به فضای واقعی داده ورودی ثابت در نظر گرفته شده است. تابع هدف این آموزش به صورت خطای MSE[[63]](#footnote-63) روی نمونه اولیه z و مقدار بازسازی شده آن توسط کدگذار تعریف شده است. تابع هدف این معماری را در زیر مشاهده می‌کنید:

|  |  |
| --- | --- |
| (10-2) |  |

d در معادله بیانگر ابعاد نمونه‌ها در فضای نهفته است.

در این روش کدگذار بر خلاف روش izi هیچ نمونه‌ای از فضای تصاویر واقعی نمی‌بیند و این مسئله می‌تواند بر آموزش صحیح کدگذار تاثیر منفی بگذارد.

در آموزش کدگشا به روش izi از ساختار کدگذار استاندارد استفاده می‌شود، بدین صورت که در ادامه کدگذار کدگشا (شبکه مولد) قرار خواهد گرفت. در فرایند آموزش ابتدا نگاشت معکوس از فضای داده واقعی به فضای نهفته توسط کدگذار انجام می‌شود و در ادامه نگاشت از فضای نهفته به فضای داده واقعی توسط کدگشا با ضرایب ثابت صورت می‌پذیرد. ساختار این روش به صورت از فضای واقعی به فضای واقعی است. تابع هدف این روش با استفاده از خطای MSE بدین شکل پیاده‌سازی می‌شود که خطای باقی‌مانده میان تصویر واقعی و تصویر خروجی مولد کمینه می‌شود، تابع هدف مورد نظر در ادامه آمده است:

|  |  |
| --- | --- |
| (11-2) |  |

در معادله 2-11 بیانگر جمع مربعات خطا در سطح پیکسل میان دو تصویر است. داده‌های آموزشی این روش داده‌های آموزش همان WGAN یعنی داده‌های عادی است. در این روش نمی‌توان به طور مستقیم میزان دقت کدگذار را در فضای نهفته اندازه گرفت و تنها می‌توان به صورت غیرمستقیم نگاشت مربوط به فضای نهفته را به فضای داده واقعی انتقال داده و در این فضا میزان دقت را اندازه گرفت به بیان دیگر میزان دقت به صورت تصویر-تصویر محاسبه می‌شود.

در روش دیگر به نام izifاز تمایزگر نیز استفاده می‌شود. در روش izi تابع هدف میزان شباهت در فضای تصویر را تحمیل می‌کند. هنگام نگاشت تصاویر جدید ممکن است با نمونه‌هایی روبرو شویم که در مرحله آموزش به صورت تنک از فضای نهفته متناظر آن‌ها نمونه گرفته باشیم وقتی نقطه متناظر را به فضای تصویر(فضای داده ورودی) می‌بریم با تصاویر تولیدی دیگر نمی‌توان تمایزگر را متقاعد کرد. در نتیجه این اتفاق تنها کمینه کردن تفاوت تصاویر در سطح پیکسل گاهی اوقات منجر به تولید تصاویر عادی می‌شوند که واقعی به نظر نمی‌رسند ولی هنوز خطای باقی‌مانده کمی حتی برای نمونه‌های ناهنجار دارند و این مورد سبب می‌شود تا دیگر خطای باقی مانده (خطای بازسازی) در فضای داده ورودی دیگر به عنوان معیار مناسب تشخیص ناهنجاری در نظر گرفته نشود.

نویسندگان مقاله دریافتند که باقی‌مانده که خود معیار مورد نظر ما برای تشخیص ناهنجاری است در فضای ویژگی توسط تمایزگر انباشته می‌شود و این عبارت حتما باید در تابع هدف مربوط به آموزش کدگذار گنجانده شود. بنابر این آمارگان تصاویر ورودی و تصاویر خروجی محاسبه می‌شود تا با استفاده از آن‌ها تصاویر خروجی شبکه مولد شبیه تصاویر ورودی بشود و بدین ترتیب روش izif پدید آید. تابع هدف این روش به شکل زیر است:

|  |  |
| --- | --- |
| (12-2) |  |

در معادله 2-12 ویژگی‌های شبکه تمایزگر که در واقع بردار ویژگی لایه‌های میانی این شبکه است با *و* *ابعاد این ویژگی با نماد نمایش داده می‌شود، همچنین* عامل وزن است. اوزان شبکه تمایزگر همان اوزانی است که در آموزش WGAN یاد گرفته شده‌اند و هنگام آموزش کدگشا ثابت در نظر گرفته شده‌اند. این مدل سبب می‌شود تا هم در فضای تصویر و هم در فضای نهفته کدگذار به جهت مناسبی حرکت کند.

#### 3-3-5-2- شناسایی ناهنجاری

در مرحله آزمایش میزان انحراف تصویر اصلی از تصویر بازسازی شده به منظور تشخیص ناهنجاری اندازه‌گیری می‌شود. تمامی موارد مورد نیاز برای بازسازی تصویر و تشخیص ناهنجاری در هنگام آموزش WGAN و کدگذار یادگرفته می‌شود. برای محاسبه امتیاز ناهنجاری مستقیما از تعریف خطای استفاده شده در آموزش کدگذار استفاده می‌شود. امتیاز نهایی که برای تشخیص ناهنجاری در مدل f-AnoGan استفاده می‌شود به صورت زیر است.

|  |  |
| --- | --- |
| (13-2) |  |

در اینجا و ، همچنین *عامل وزن است. عبارت مورد نظر برای نمونه‌های عادی دارای خطای کمی است و برای نمونه‌های ناهنجار دارای مقدار بزرگی است. از آنجایی که مدل تنها روی نمونه‌های عادی آموزش دیده است نمونه‌های بازسازی شده از نظر بصری شبیه تصویر ورودی هستند. توانایی بازسازی تصویر به طوری که شبیه تصویر ورودی باشد رابطه عکس دارد با درجه یا میزان تمایز ناهنجاری دارد. تصاویر عادی میزان انحراف کمی دارند در حالی‌که تصاویر ناهنجار که به بازسازی خود نگاشت می‌شوند میزان انحراف زیادی دارند. قدر مطلق خطای باقی‌مانده در سطح پیکسل به صورت زیر تعریف می‌شود.*

|  |  |
| --- | --- |
| (14-2) |  |

#### 4-3-5-2- مزایا و معایب

f-AnoGan مشکلات مدل قبلی (AnoGan) نظیر پیچیدگی زمانی بالا در هنگام اجرا رفع کرد. همچنین از شبکه مولد تخاصمی قوی‌تری در ساختار مدل خود استفاده کرد و به نتایج قبابل قبولی نیز دست یافت. در مدل قبلی از یک فرایند تصادفی و مبتنی بر تکرار برای نگاشت معکوس از فضای داده واقعی به فضای نهفته استفاده می‌شد که از نظر زمانی هزینه گزافی را به هنگام اجرا به مدل تحمیل می‌کرد. در مدل جدید f-AnoGan با استفاده از یک کدگشا پارامترهای نگاشت معکوس مورد نظر فراگرفته می‌شود. علی‌رغم این موفقیت این مدل همچنان از مشکلاتی نظیر عدم استفاده توام از هر دو فضای تصویر و نهفته برای آموزش کدگشا رنج می‌برد.

### 4-5-2- مدل ALI

این شبکه در سال 2017 در کنفرانس ICLR معرفی شد[4]. این شبکه­ها با هدف یادگیری نگاشت معکوس از دامنه ورودی­ها به دامنه توزیع تعریف شد. در این شبکه، علاوه بر شبکه مولد که در معماری اصلی نیز تعریف شده بود، یک رمزگذار[[64]](#footnote-64) E نیز وجود دارد که از دامنه داده­های ورودی به دامنه ویژگی­ها می­برد. بدین ترتیب خروجی بخش مولد یک دوتایی[[65]](#footnote-65) است؛ که یکی از دامنه ویژگی­ها و دیگری از دامنه داده­های ورودی است. این مدل به طور همزمان شبکه مولد و شبکه استنتاج را با استفاده از یک فرآیند تخاصمی به کار می­برند. شبکه مولد، نمونه­ها را از یک فضای نهفته آماری به فضای داده­ها نگاشت می­کند و شبکه استنتاج نمونه­های آموزش را از فضای داده به فضای متغیرهای نهفته نگاشت می­کند. به این صورت یک بازی خصمانه بین دو شبکه انجام می­شود. در این جا شبکه متمایزگر باید یاد بگیرد تا تفاوت بین جفت ورودی فضای نهفته و فضای داده را تشخیص دهد. شبکه تمایزگر در این­جا علاوه بر تفکیک در فضای داده، در فضای ویژگی نیز تفکیک می­کند. به این معنا که تشخیص می­دهد دوتایی وارد­شده، داده واقعی است با ویژگی تولیدشده توسط E و یا داده جعلی است که توسط و با ویژگی­های درست شده­است. در تصویر زیر چارچوب کلی این الگوریتم، به نمایش درآمده است:



5شکل 2-3: معماری شبکه ALI.

**دو تابع توزیع احتمال روی و در نظر بگیرید:**

* **تابع توزیع تعریف شده برای رمزگذار**  *E*
* **تابع توزیع تعریف شده برای رمزگشا**

**این دو توزیع، توابع توزیع حاشیه­ای دارند که برای ما آشناست: توزیع حاشیه­ای رمزگذار تابع توزیع داده­های اصلی است و توزیع حاشیه­ای رمزگشا معمولا به عنوان یک تابع توزیع ساده مانند تابع توزیع استاندارد در نظر بگیریم. بدین ترتیب روند تولید معکوس می­باشد.**

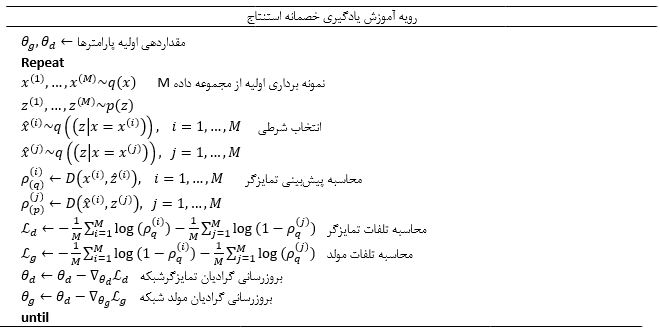
**هدف اصلی شبکه** ALI **مطابقت این دو توزیع است. اگر این شرط محقق شود، ما اطمینان حاصل می­کنیم که تمام توزیع­های حاشیه­ای و توزیع­های شرطی مطابقت دارد. برای دستیابی به این توابع توزیع، یک بازی متخاصمانه صورت می گیرد. جفت از دو توزیع در نظر گرفته می‌شود و یک شبکه تمایزگر می‌آموزد تا بین این دو خروجی، تمایز قائل شود؛ در حالی که دو شبکه رمزگشا و رمزگذار می­آموزند تا این شبکه را فریب دهند. در نهایت تابع هدفی که این بازی بر اساس آن صورت می­گیرد به صورت زیر است:**

|  |  |
| --- | --- |
| (10-2) |  |

**ویژگی جالب رویکردهای خصمانه این است که آنها نیازی به محاسبه تابع چگالی شرطی ندارند. آنها فقط نیاز دارند که به نحوی نمونه برداری شوند که این امکان را به وجود آورد که بتواند از پس انتشار گرادیان استفاده کند. در مورد شبکه** ALI **، این بدان معنی است که گرادیان ها باید از شبکه تمایزگر به شبکه­های رمزگذار و رمزگذار انتشار یابند.**

**به طور دقیق­تر شبکه تمایزگر آموزش می بیند که بین نمونه­هایی که از رمزگذار و نمونه هایی که از رمزگشا تولید می­شود، تمایز بگذارد. شبکه مولد و شبکه رمزگذار نیز می­آموزند که شبکه تمایزگر را فریب دهند؛ یعنی جفت تولید کنند که غیر قابل تشخیص باشد.**

**در الگوریتم 2‑2 شبکه** ALI **توصیف شده است. اثبات می­شود که با فرض یک تمایزگر بهینه، شبکه مولد، واگرایی جنسن-شانون را بین به حداقل می‌رساند.**



6الگوریتم 2-2: رویه آموزش یادگیری خصمانه استنتاج.

#### 1-4-5-2- مقایسه مدل‌های ALI و GAN­

شبکه ALI شباهت زیادی به شبکه GAN دارد ، اما دو تفاوت اساسی با آن دارد:

1- بخش مولد دارای دو مؤلفه است: بخش رمزگذار، که نمونه­های داده را به فضای z نگاشت می­کند و بخش رمزگشایی که نمونه­ها را از (منبع منبع نویز) به فضای ورودی نگاشت می­کند.

2- بخش تمایزگر به منظور تمایز بین جفت ، آموزش دیده می­شود.

#### 2-4-5-2- رویکرد‌های جایگزین برای استنتاج در GAN

روش استفاده شده در مدل ALI تنها راه استنتاج در شبکه‌های عصبی مولد تخاصمی نیست. راه دیگر برای انجام این کار استفاده از شبکه استنتاج جلورو[[66]](#footnote-66) در ساختارGAN است. در مدل InfoGAN[[67]](#footnote-67) [8] با کمینه کردن اطلاعات متقابل[[68]](#footnote-68) میان مجموعه **از فضای نهفته و**  **به وسیله توزیع کمکی**

نگاشت معکوس را یاد می‌گیرد. InfoGan نیاز دارد تا تابع احتمال پسین[[69]](#footnote-69) قابل تخمین و ارزیابی باشد. در مدل ALI تنها نیاز است که بتوان از شبکه استنتاج نمونه گرفت تا بدین وسیله تابع پیچیده توزیع پسین را بازنمایی کرد. عمل انجام شده در این کار مشابه این است که یک کدگذار برای بازسازی آموزش دهیم. به عنوان کثال پیدا کردن کدگذار به طوری که . نمونه که در جمله قبل بدان اشاره شد از نظر رویه‌ای شبیه به InfoGAN اما در این روش از یک شبکه مولد با ضرایب ثابت و همچنین تابع توزیع پسین گوسی با واریانس قطری ثابت استفاده شده است.

روند آموزش را می‌توان به دو فاز تقسیم کرد. در فاز اول شبکه مولد تخاصمی به صورت معمول آموزش می‌بیند. در فاز دوم کدگشا ثابت در نظر گرفته می‌شود و کدگذار به روش مدل ALI آموزش داده می‌شود. در این روش کدگذار و کدگشا در هنگام آموزش هیچ تعاملی با هم ندارند و در واقع کدگذار بر اساس هر چه کدگشا آموخته است آموزش می‌بیند. مشخص است اگر کدگذار و کدگشا با هم تعامل داشته باشند روند مدل‌سازی داده بهبود خواهد یافت.

#### 3-4-5-2- مزایا و معایب

**اگرچه وجود مکانیزم مناسب جهت نگاشت معکوس از فضای داده واقعی به فضای نهفته سبب موفقیت و عملکرد قابل قبول مدل** ALI **شده است اما در این مدل هیچ سازوکاری جهت کنترل میزان شباهت تصویر اصلی و تصویر بازسازی سده در آن تعبیه نشده است و به همین جهت در برخی موارد هیچ شباهتی میان داده اصلی و داده بازسازی شده وجود ندارد. به بیان دیگر پس از یافتن نقطه متناسب با داده مورد نظر در فضای داده واقعی در فضای نهفته، نقطه مورد نظر را به شبکه مولد می‌دهیم و خروجی حاصل را با ورودی ابتدایی مقایسه می‌کنیم. انتظار می‌رود داده ورودی و خروجی در این چرخه شباهت زیادی داشته باشند ولی همانطور که گفته شد در برخی از نمونه‌ها شباهتی میان این دو تصویر وجود ندارد.**

### 5-5-2- مدل EGBAD

همانطور که گفته شد شبکه‌های عصبی تخاصمی قادرند توزیع‌‌های پیچیده دنیای واقعی با ابعاد بالا را مدل کنند و همین امر سبب می‌شود تا بتوان از این شبکه‌ها در زمینه تشخیص ناهنجاری نیز کرد. با توجه به تعداد کارهای انگشت شمار در این زمینه، مدل EGBAD را می‌توان از اولین کارها در تشخیص ناهنجاری با استفاده از شبکه‌های عصبی تخاصمی به شمار آورد.

مدل ارائه شده بر اساس شبکه عصبی تخاصمی دوطرفه به اختصار BiGAN[[70]](#footnote-70) بنا نهاده شده است. وظیفه نگاشت معکوس از فضای داده ورودی به فضای نهفته نیز بر عهده کدگذار است. کدگذار، شبکه مولد و تمایزگر در اینجا به طور همزمان آموزش می‌بینند و وجود بلوک کدگذار سبب کاهش هزینه‌های محاسباتی در گام آزمایش می‌شود. بر خلاف ساختار استاندارد GAN که در آن تمایزگر تنها تصویر واقعی و تصویر تولیدی شبکه مولد را ورودی می‌گیرد، بازنمایی این تصاویر در فضای نهفته هم به عنوان ورودی به شبکه تمایزگر داده می‌شود.

استراتژی مورد استفاده در گام آموزش مدل مشابه کار [4] است. همانطور که در قسمت قبل بررسی شد در این استراتژی آموزشی تاکید بر آن است که شبکه مولد و کدگذار به طور توامان آموزش داده شوند. در هنگام آموزش معادله زیر بهینه می‌شود.

|  |  |
| --- | --- |
| (11-2) |  |

**در معادله 2-11**  بیانگر تابع توزیع داده ورودی است، بیانگر توزیع نمونه‌ها در فضای نهفته است، و به ترتیب بیانگر تابع توزیع کدگذار و کدگشا هستند.

پس از آموزش مدل نوبت به تعریف معیاری می‌رسد که به وسیله آن بتوان نمونه‌های ناهنجار را تشخیص داد. بیان ریاضی این معیار در ادامه آمده است.

|  |  |
| --- | --- |
| (12-2) |  |

**در معادله 2-12**  و را می‌توان به دو روش تعریف کرد. در روش اول تابع خطای آنتروپی[[71]](#footnote-71) متقابل استفاده می‌شود. ورودی این تابع در واقع خروجی تمایزگر است و خود شامل خروجی کدگذار و نمونه متناظر با آن است و به صورت تعریف می‌شود. در اینجا خروجی تمایزگر میزان اطمینان آن در قبال واقعی بودن نمونه ورودی است. در روش دوم تعریف از خطای تطبیق ویژگی [[72]](#footnote-72) استفاده می‌شود. تابع امتیاز این روش به صورت تعریف می‌شود. در واقع لاجیت‌های تمایزگر است و بیان می‌کند ویژگی‌های تصویر بازسازی شده تا چه اندازه شبیه ویژگی‌های تصویر واقعی هستند. توجه شود هر چه مقدار امتیاز محاسبه شده بالاتر باشد نمونه مورد نظر با احتمال بیشتری نمونه ناهنجار است.

#### 1-5-5-2- مزایا و معایب

***مدل مورد بحث در این قسمت ثابت کرد می‌توان با تعریف تابع امتیاز مناسب و استفاده از ساختارهای تخاصمی به روز ناهنجاری را داده‌های دنیای واقعی شناسایی کرد. در این مدل هیچ نظارتی بر میزان تشابه داده ورودی و داده بازسازی وجود ندارد در حالی که انتظار می‌رود داده ورودی و داده بازسازی شده توسط شبکه مولد برای نمونه‌های هنجار یکسان باشد. در ادامه به بررسی این مشکل و روش ارائه شده برای حل آن پرداخته می‌شود.***

### 6-5-2- مدل ALICE

در حالت استاندارد شبکه GAN تنها نگاشت یک طرفه از فضای نهفته به فضای داده بدست می­آورد، یعنی فاقد مکانیسم معکوس (از فضای داده به فضای نهفته) است و این امر مانع می­شود که این شبکه ها قادر به استنباط باشند. توانایی محاسبه تابع توزیع متغیر نهفته شرطی ممکن است برای تفسیر داده­ها و برای برنامه­های پایین دستی (به عنوان مثال، طبقه­بندی متغیر نهفته) مهم باشد.

تلاش­های زیادی برای یادگیری همزمان یک مدل دو طرفه کارآمد برای تولید نمونه­هایی با کیفیت بالا برای هر دو فضای نهفته و داده صورت گرفته است. در میان این طرح­ها، یکی از طرح­ها که به موفقیت چشم­گیری دست یافته است، شبکه یادگیر استنباط خصمانه ALI است[4]. دراین مدل در یک چارچوب شبکه مولد متخاصم، شبکه تمایزگر می­آموزد تا تفاوت بین دو توزیع توام را تشخیص دهد.

با این که شبکه ALI یک رویکرد جالب و خلاقانه است، اما یک ایراد اساسی دارد؛ این که بازسازی­های صورت گرفته از داده­ها در بعضی موارد حتی به داده های اصلی شبیه هم نیستند. دلیل این امر این است که شبکه ALI تنها به دنبال مطابقت دو توزیع توامان است، اما همبستگی بین دو متغیر تصادفی شرطی در هر یک از این توابع مشخص و اعمال نمی­شود. در نتیجه حاصل، راه­حل­هایی می­شود که هدف ALI را برآورده سازند اما در بازسازی داده­های مشاهده شده با مشکل روبرو هستند. این شبکه هم­چنین مشکلاتی در کشف رابطه صحیح جفت­ها در زمان تغییر دامنه دارد.

#### 1-6-5-2- یادگیری تخاصمی با اندازه‌گیری اطلاعات

به یاد داریم که تابع هدف در شبکه ALI به صورت معادله ‏11‑2 بود:

|  |  |
| --- | --- |
| (13-2) |  |

نقطه تعادل این معادله هنگامی است که  **باشد. ارتباط بین متغیرهای تصادفی توسط** ALI **محدود و مقید نشده است. در نتیجه، این امکان وجود دارد که توزیع همسان برای یک کاربرد خاص نامطلوب باشد. در واقع بسیاری از کاربردها به ثبات چرخه و وجود یک نگاشت معنی­دار دو طرفه بین دامنه ها احتیاج دارند.**

**جهت مقابله با مشکل توزیع­های نامطلوب اما برابر، بر روی راه­حل­های شبکه** ALI **باید محدویتی بر روی توزیع های و اعمال شود. این کار با کنترل "عدم قطعیت" بین جفت متغیرهای تصادفی، یعنی با استفاده از آنتروپی­های شرطی انجام می­شود.**

#### 2-6-5-2- آنتروپی شرطی[[73]](#footnote-73)

**آنتروپی شرطی یک معیار نظریه اطلاعاتی است که عدم قطعیت متغیر تصادفی را هنگام مقید شدن بر روی با کمک توزیع توامان تعیین می­کند:**

|  |  |  |  |
| --- | --- | --- | --- |
| |  |  | | --- | --- | | (14-2) |  | |  |

**عدم قطعیت متغیر مقید بر روی متغیر با مرتبط است. در حقیقت، اگر**  **باشد در این صورت به طور قطعی به** Z **وابسته می­باشد. به کمک کنترل میزان عدم قطعیت**  *و***، می توان راه حل های** ALI **را در توزیع های توامانی که نگاشت آن ها منجر به نتایج بهتری می شود، محدود کرد. در نهایت با افزودن یک عامل تنظیم کننده آنتروپی شرطی، به تابع هدف زیر دست می یابیم:**

|  |  |
| --- | --- |
| **(2 -15)** |  |

وابسته به متغیرهای تصادفی توزیع های توامان است. در حالت ایده آل، پس از شناسایی تمام نقاط تعادل تابع هدف ALI، می­توان با محاسبه آنتروپی شرطی آن­ها، راه­حل مطلوب را انتخاب کنیم. با این حال، در عمل این راه غیرقابل استفاده است، زیرا ما از قبل به نقاط تعادل دسترسی نداریم. در ادامه یک راه­حل برای محاسبه آنتروپی شرطی ارائه می­شود.

#### 3-6-5-2- فرایند یادگیری

در نبود تابع توزیع احتمال صریح که برای محاسبه آنتروپی شرطی مورد نیاز است، می­توان حدود آنتروپی شرطی را با استفاده از معیار ثبات چرخه[[74]](#footnote-74) محدود کرد. در این­جا برای بازسازی به طریق زیر عمل می شود:

|  |  |
| --- | --- |
| **(2 -16)** |  |

به کمک روال تولید بالا، تلاش می شود تا با احتمال بالایی شبیه اصلی باشد. اثبات می­شود که به کمک این روال تولید ها، حد بالای آنتروپی شرطی می­باشد.

نکته حائز اهمیت این است که می­توان عامل تنظیم آنتروپی را به تابع هدف شبکه ALI، بدون اعمال تغییرات اضافی دیگری، در روال آموزش این شبکه اضافه کرد. بدین ترتیب تابع بهینه­سازی برای شبکه ALICE به­صورت معادله ‏15‑2 خواهد بود.

|  |  |
| --- | --- |
| **(2 -17)** |  |

ویژگی پایداری چرخش در مقالات پیش از نیز وجود داشته است این ویژگی در این مقالات به کمک نرم درجه[[75]](#footnote-75) یک و دو و داده­های­ واقعی مانند تصاویر محاسبه ­شده ­است. وجود تابع اتلاف بر اساس نرم درجه 2 مبتنی بر پیکسل، سبب می­شود که نمونه­های خروجی این شبکه­ها تصاویر تاری باشند. به همین علت در این شبکه از یک شبکه تمایزگر که اختلاف بین ها و های بازسازی­شده را اندازه­گیری می­کند، استفاده­ شده ­است.

#### 4-6-5-2- مزایا و معایب

همانطور که بررسی شد در این کار با استفاده از کدگذار تلاش شد تا شرط پایداری در شبکه‌های عصبی مولد برقرار شود به بیان دیگر اگر تصویر ورودی را به کدگذار بدهیم و نقطه متناظر در فضای نهفته را بدست آوریم و سپس آن را به عنوان ورودی به شبکه مولد بدهیم انتظار داریم نتیجه نهایی شبیه به تصویر اولیه باشد. علی‌رغم عملکرد مناسب این مدل هنوز توزیع‌های توامی وجود دارد که از آن‌ها استفاده نشده است و همین امر سبب می‌شود تا از تمامی اطلاعات موجود استفاده نشود.

## 6-2- جمع‌بندی

در این فصل ابتدا به دسته‌بندی روش‌های مختلف تشخیص ناهنجاری پرداختیم، در خلال همین دسته‌بندی ‌ها برخی مدل‌های به نسبت قدیمی‌تر و سنتی بررسی شدند. در ادامه اهمیت جایگاه شبکه‌های عصبی تخاصمی روشن شد. در گام بعدی معیار‌های ارزیابی مدل‌های تشخیص ناهنجاری معرفی شده‌اند. در ادامه پس از بررسی مدل پایه شبکه عصبی تخاصمی زنجیره کارهایی که روی شبکه‌هایی عصبی تخاصمی به منظور بهبود صورت گرفته است بررسی شده است. در ادامه مدل AnoGan شرح داده شد و با توجه به مشکل آن در نگاشت معکوس از فضای داده ورودی به فضای نهفته مدل f-AnoGan که در ادامه کار قبلی است بررسی شد. با توجه به ضعف‌های موجود در ساختار f-AnoGan مقاله مکمل این مدل ALI مرور شد. . در گام بعدی مدل EGBAD که جزو اولین کار‌ها در زمینه تشخیص ناهنجاری که با الهام از مدل ALI خلق شده است مرور می‌شود. در مرحله بعدی با توجه تضمین نشدن شرط سازگاری حلقه در ALI مقاله آلیس بررسی شد.

# فصل سوم روش پیشنهادی

در فصول قبل به بررسی روش­های تشخیص ناهنجاری با رویکردهای مختلف پرداختیم. همان­طور که پیش­تر ذکر شد، با پیشرفت روزافزون زیرساخت­های محاسباتی و افزایش توان پردازشی، علاقه و توجه محققان حوزه هوش مصنوعی به سمت شبکه­های عصبی جلب شد. در میان این انواع شبکه­های عصبی ارائه ­شده در چند سال اخیر، شبکه های مولد تخاصمی به نتایج قابل توجه و درخشانی در کاربرد­های مختلف دست­ یافته­اند. علی­رغم نتایج قابل دفاع این نوع شبکه­ها در زمینه­های پردازش تصویر[3]، پردازش گفتار و پردازش متن [9]در حوزه تشخیص ناهنجاری آن طور که شایسته است بدان توجه نشده. در زمینه شناسایی نمونه‌های ناهنجار در دادگان دنیای واقعی به ندرت می‌توان الگوریتمی یافت که بر مبنای شبکه‌های مولد تخاصمی طراحی شده باشد؛ در حقیقت بیشتر الگوریتم‌های تخاصمی برای کاردبر‌های دیگر طراحی شده است و صرفا همان ساختار در فضای جدید تشخیص ناهنجاری مورد استفاده قرار گرفته است و با این حال همین مدل‌های عام منظوره توانسته به نتایج نسبتا معقولی دست یابد[6].

در این بخش روشی مبتنی بر شبکه‌های مولد تخاصمی که به صورت خاص با هدف تشخیص ناهنجاری پیشنهاد شده است. در طراحی این روش تمرکز اصلی بر روی ارائه روشی است که بتواند در کاربردهای دنیای واقعی نظیر تصاویر پزشکی مورد استفاده قرار بگیرد. مانند بسیاری از الگوریتم‌های مبتنی بر یادگیری، در این‌ جا دو مرحله اصلی آموزش و آزمایش وجود دارد. در قسمت آموزش همانند دیگر چهارچوب‌های تخاصمی به نوبت بخش مولد و بخش تمایزگر را آموزش می‌دهیم تا هر دو بخش در عین تناسب به نوبت به‌روزرسانی شود. همچنین یک مرحله پیش‌پردازش شامل نرمال‌سازی تصاویر به منظور افزایش دقت مدل نهایی انجام شده است. در همین مرحله نمونه‌ها ناهنجار را از تصاویر آموزشی حذف می‌کنیم و فقط از آن‌ها در مرحله آزمایش برای ارزیابی مدل پیشنهادی استفاده می‌کنیم. به بیان دیگر مدل تنها توزیع داده‌های عادی را می‌آموزد؛ روند آموزش مدل به حدی قدرتمند است که می‌تواند به خوبی فضای داده‌های ناهنجار را از داده‌های عادی تبیین کند.

در ادامه و در قسمت تشخیص ناهنجاری، یک ورودی بدون برچسب به ساختار شبکه وارد می‌شود و به کمک اختلاف بازسازی ارائه شده توسط شبکه برای آن ورودی در هر دو فضای داده و فضای نهان، امتیاز ناهنجاری برای هر ورودی محاسبه می‌شود. با توجه به این نکته که شبکه بر روی دادگان نرمال آموزش داده می‌شود انتظار می‌رود که برای داده‌های ناهنجار بازسازی ضعیف‌تری داشته باشد و بدین ترتیب این اختلاف بیشتر شود و در نهایت این داده‌ها امتیاز ناهنجار بیشتری بگیرند. بدین ترتیب با انتخاب یک حدآستانه[[76]](#footnote-76) و یا انتخاب یک نسبت معین از میان داده‌های با بیشترین امتیاز، داده‌های که ناهنجار هستند شناسایی می‌شوند.

در قسمت‌های بعد جزئیات دقیق‌تر روش پایه و همچنین جزئیات هر قسمت از الگوریتم پیشنهادی مورد بررسی قرار خواهد گرفت.

## 1-3- مدل ALAD

در این بخش، یک روش تشخیص ناهنجاری مبتنی بر شبکه مولد متخاصم را بررسی می­کنیم که در زمان آزمون بسیار کارآمد است. در این روش به طور هم­زمان یک شبکه رمزگذار را در حین آموزش فرا می­گیرد و بدین ترتیب استنتاج سریع­تر و کارآمدتر را در زمان آزمون امکان­پذیر می­کند. علاوه بر این در شبکه معرفی­شده، تکنیک­هایی که اخیرا برای بهبود بیشتر شبکه رمزگذار و تثبیت آموزش شبکه مولد متخاصم ترکیب­شده و نشان­ داده­ شده که این تکنیک­ها عملکرد و کارایی را در کاربرد تشخیص ناهنجاری بهبود می­بخشند. آزمایشات روی طیف وسیعی از داده­های جدولی و تصویری، کارایی و اثربخشی این رویکرد را در عمل نشان می­دهد[6].

شبکه­های GAN استاندارد از نمونه­گیری کارآمد پشتیبانی می­کنند و روش­های مختلفی وجود دارد که می­تواند آن­ها را برای تشخیص ناهنجاری تطبیق دهد. به عنوان مثال، برای یک نقطه داده ، می­توان از نمونه‌گیری استفاده کرد تا احتمال ناهنجار بودن را تخمین زند. تخمین دقیق احتمال به تعداد زیادی نمونه نیاز دارد و در نتیجه محاسبه احتمال، بار محاسباتی سنگینی دارد.

روش دیگر معکوس کردن[[77]](#footnote-77) شبکه مولد برای یافتن متغیرهای نهفته است که به معنای به حداقل رساندن خطای بازسازی با تابع هدف گرادیان نزولی تصادفی می باشد. این روش هم­چنین از نظر محاسباتی بسیار پرهزینه است زیرا هر محاسبه گرادیان نیاز به یک پس انتشار از طریق شبکه مولد دارد.

به واسطه بهره­وری محاسباتی بالا و قابلیت مدل­سازی داده­های ابعاد بالا، از شبکه­های مولد متخاصمی به همراه یک شبکه رمزگذار (که نمونه­ها را از فضای داده به فضای نهفته نگاشت می­کند) استفاده می­شود. نمایش نهفته هر نمونه از فضای داده در چنین مدل­هایی صرفاً با عبور از شبکه رمزگذار انجام می­شود. هم­چنین این مدل پیشرفت­های اخیر که برای بهبود شبکه رمزگذار صورت گرفته مانند افزودن یک شبکه تمایزگر برای بهبود سازگاری چرخه را شامل می­شود.

شبکه ALI توزیع توامان داده­ها را به همراه یک شبکه رمزگذار مدل می­کند. این مدل یک شبکه تمایزگر دارد که را به عنوان ورودی می­گیرد و بررسی می­کند که این جفت ورودی از کدام منبع – شبکه مولد و یا شبکه رمزگذار – تولید شده است.

|  |  |
| --- | --- |
| **(3 -1)** |  |

با این­که به لحاظ نظری توزیع توامان به یک نقطه میل می­کند، اما در عمل اغلب نتیجه یکسان نیست و لزوما به یک نقطه همگرا نمی­شوند و این پدیده سبب نقض پایداری چرخه می­شود. نبود پایداری چرخه به این معناست که باشد. این مشکل برای روش های تشخیص ناهنجاری مبتنی بر بازسازی چالش­های جدی ایجاد می­کند. برای حل این مشکل، چارچوب ALICE پیشنهاد می­کند که آنتروپی شرطی را به صورت زیر به روش تخاصمی برای سازگاری چرخه تقریب بزنیم.

|  |  |
| --- | --- |
| **(3 -2)** |  |

در این تعریفبه معنای توزیع توامان می­باشد.این مسئله شامل یکپارچه­سازی آنتروپی شرطی در شبکه رمزگذار و شبکه مولد است:

|  |  |
| --- | --- |
| **(3 -3)** |  |

که افزودن این عامل به مسئله معادل افزودن یک شبکه تمایزگر جدید به مسئله می­باشد. با افزودن این عامل، مسئله بهینه­سازی به صورت معادله ‏4-3 خواهد بود:

|  |  |
| --- | --- |
| **(3 -4)** |  |

### 1-1-3- تثبیت آموزش GAN بر پایه ALICE

برای تثبیت آموزش در مدل پایه ALICE، توزیع­های شرطی را با اضافه­کردن یک قید آنتروپی شرطی دیگر تنظیم می­کنیم و سپس عملیات نرمال­سازی طیفی را انجام می­دهیم.

توضیح دقیق­تر این­که، فضای نهفته شرطیرا با یک شبکه تمایزگر مخالف دیگر با نقطه تعادل مشترک تنظیم می­کنیم.

|  |  |
| --- | --- |
| **(3 -5)** |  |

با کنار هم قرار دادن تمامی این اجزا، شبکه ALAD تلاش می­کند تا نقطه تعادل این مسئله را آموزش ببیند.

|  |  |
| --- | --- |
| **(3 -6)** |  |

در نهایت شماتیک کلی شبکه طراحی­شده به صورت شکل ‏1‑3 خواهد بود .

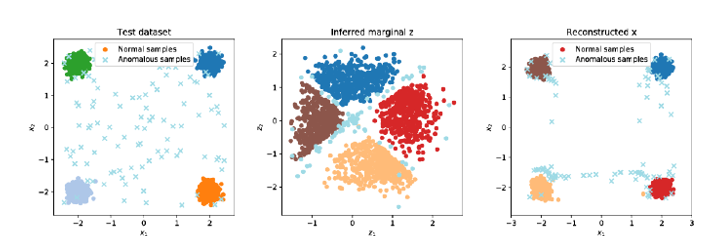


7شکل 3-1: شمای کلی شبکه ALAD.

اضافه کردن مرحله نرمال­سازی طیفی با انگیزه یادشده مقاله [10] می­باشد. در این مقاله نشان داده­شده که با افزودن قیود لیپسچیتز[[78]](#footnote-78) به تمایزگر شبکه GAN، فاز آموزش تثبیت خواهد شد. در عمل نشان داده شده که با تنظیم مجدد پارامترهای وزن، بهبود بسیار خوبی روی عملکرد شبکه خواهیم داشت. بدین صورت که بزرگ­ترین مقادیر ویژه ماتریس وزن را در شبکه تمایزگر ثابت نگه داریم. این روش از نظر محاسباتی کارآمد است و هم­چنین آموزش را تثبیت می­کند. با آزمایش­های صورت­گرفته نشان داده شد که افزودن این قیود، نه تنها برای شبکه تمایزگر، بلکه برای شبکه رمزگذار نیز سودمند است. قابل توجه است که مدل ALICE شامل این مرحله نمی باشد.

### 2-1-3- تشخیص ناهنجاری

شبکه ALAD یک روش تشخیص ناهنجاری مبتنی بر بازسازی است و بدین صورت عمل می­کند که فاصله نمونه از بازسازی را توسط شبکه GAN ارزیابی می­کند. نمونه­های عادی باید به طور دقیق بازسازی شوند در حالی که نمونه­های ناهنجار احتمالاً به طور ضعیف­تری بازسازی می­شوند. نحوه تشخیص ناهنجاری در شکل 3-2 نشان شده است.



8شکل 3-2: نمونه‌ای از خروجی شبکه ALAD یه همراه داده‌های ناهنجار. در این تصویر ضربدر‌ها نمونه‌های ناهنجار و دایره‌های رنگی نمونه‌های عادی هستند.

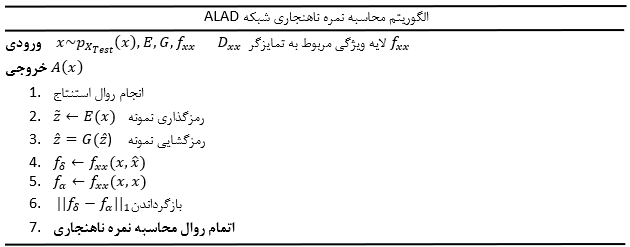
به منظور مدل کردن توزیع داده‌ها به طور موثر از شبکه‌های مولد تخاصمی استفاده می‌شود. تابع مولد برای یادگیری توزیع داده­های عادی استفاده می­شود، به طوری که

می­باشد. هم­چنین باید توزیع حاشیه­ای داده­ها را بیاموزیم تا بازنمایی­های نهفته منجر به بازسازی صحیح نمونه­های عادی شود. دو عنصر پایداری چرخه شرطی متقارن در این مدل، به اطمینان از این امر کمک می­کند.

مؤلفه کلیدی دیگر ALAD نمره ناهنجاری است که فاصله بین نمونه­های اصلی و بازسازی آن­ها را اندازه­گیری می­کند. انتخاب اولیه­ای که به ذهن می­رسد، فاصله اقلیدسی بین نمونه­های اصلی و بازسازی آن­ها در فضای داده است. اما، این معیار ممکن است معیار مطمئنی برای اندازه­گیری تشابه نباشد. به عنوان مثال، این معیار در مورد تصاویر می­تواند بسیار پرخطا باشد؛ زیرا تصاویر با ویژگی­های تصویری مشابه الزاماً از نظر فاصله اقلیدسی نزدیک به یک­دیگر نیستند. معیار تعریف­شده در این روش از فاصله بین نمونه­ها در فضای ویژگی­های تمایزگر محاسبه می­شود، که توسط لایه قبل از لاجیت تعریف شده است. از این ویژگی­ها هم­چنین به عنوان کدهای CNN یاد می­شود. به­طور دقیق­تر می­توان گفت با آموزش یک مدل برای داده­های عادی و محاسبه*E* ، *G* ، *Dxz* ، *Dxx* و *Dzz* یک تابع نمره­دهی را بر اساس خطای بازسازی نرم1 مطابق معادله 4-9 تعریف می­شود. در این تعریف تابع فعال­ساز[[79]](#footnote-79)های لایه قبل از لاجیت و یا همان کد CNN می­باشد. این نوع تعریف به ما این اطمینان را می­دهد که نمونه به درستی کدگذاری و بازسازی­شده و در نتیجه از توزیع داده واقعی می­باشد.

|  |  |
| --- | --- |
| **(3 -7)** |  |

با این تعریف، نمونه­ها با بیش­تر به احتمال بالاتری داده ناهنجار خواهند بود. در ادامه در الگوریتم ‏1‑3 روال محاسبه ارائه می­شود.



9الگوریتم 3-1: شبه کد الگوریتم ALAD.

معیار استفاده شده در این­جا از ایده تطابق ویژگی­های از دست­رفته الهام گرفته شده است[11]. اما در این­جا به جای استفاده از ویژگی­های محاسبه شده درشبکه تمایزگر GAN استاندارد (که اختلاف را بین نمونه­های تولید شده و داده­های واقعی را محاسبه می­کند)، از ویژگی­های محاسبه­شده در شبکه تمایزگر استفاده می­شود. هم­چنین در این­جا به جای استفاده از این معیار در حین آموزش شبکه GAN، از این معیار در هنگام روال استنتاج بهره می­جوییم.

سوالی که در این­جا مطرح می­شود این است که : چرا نباید از خروجی تمایزگر به عنوان معیار فاصله استفاده کرد. پاسخ این سوال بدین صورت است که هدف از شبکه تمایزگر تمایز بین یک جفت نمونه واقعی و بازسازی آن می­باشد و شبکه رمزگذار و شبکه مولد داده­های واقعی و توزیع متغیر نهفته را کاملاً ضبط خواهند کرد. در این حالت قادر به تفکیک بین نمونه­های واقعی و نمونه­های بازسازی شده نخواهد بود و بدین ترتیب یک پیش بینی تصادفی را تولید می­کند که معیار ناهنجاری مناسبی نخواهد بود.

## 2-3- مدل RCGAN[[80]](#footnote-80)

تمرکز مقاله [12] بر ارائه ساختاری مبتنی بر شبکه‌های عصبی تخاصمی به منظور هر چه بهتر کردن تشخیص نمونه‌های ناهنجار به وسیله پوشش تمام فضای نهفته و فضای داده ورودی است. اساس این کار بر پایه تعریف تابع جریمه، بیان جدیدی از تابع هزینه و همچنین استفاده نوآورانه از تمایزگر است. عملکرد مناسب این مدل در خلال بررسی نتایج آن قابل مشاهده است.

### 1-2-3- شناسایی نمونه‌های تخاصمی از دیدگاه آماری

داده‌های هنجار مورد استفاده در مرحله آموزش به عنوان تابع توزیع احتمال با نام تعریف می‌شود. با توجه به این‌که در مرحله آموزش تنها به نمونه‌های تابع توزیع  دسترسی داریم این مدل نیز به عنوان آموزش بی‌نظارت دسته‌بندی می‌شود. هدف در این مدل آموزش تابع است، این تابع در واقع به هر نمونه یک امتیاز ناهنجاری اختصاص می‌دهد به گونه‌ای که برای نمونه‌های ناهنجار مقدار بیشتری اختصاص می‌دهد.

### 2-2-3- GAN

شبکه مولد تخاصمی توزیع داده ورودی را با نمونه گرفتن از توزیع فضای نهفته و با استفاده از تبدیل مدل می‌کند به بیان دیگر  *توزیع احتمال شرطی را تعریف می‌کند. توزیع مولد را می‌توان به صورت زیر تعریف کرد.*

|  |  |
| --- | --- |
| **(3 -8)** |  |

در چارچوب GAN تمایزگر وظیف دارد تا میان داده‌هایی که از توزیع می‌آیند و داده‌های مصنوعی که از توزیع نشئت می‌گیرد تمیز دهد. تمایزگر و مولد به هنگام تطابق () در حالت بهینه قرار دارند. کدگذار مدل ALI تلاش می‌کند تا توزیع توام خود را را با توزیع توام مولد تطابق دهد. تابع ا=هدف مدل ALI نیز با توجه به همین عبارات به صورت زیر تنظیم می‌شود.

|  |  |
| --- | --- |
| **(3 -9)** |  |

در معادله 3-9 بیانگر شبکه مولد است که از هر دو فضای نهفته () و فضای داده واقعی () ورودی می‌گیرد و خروجی آن احتمال تعلق **و**  **به توزیع**  است. همانطور که در فصل گذشته دیدیم به منظور ارضای شرط پایداری حلقه، به ساختار مدل اضافه می‌شود تا میان داده تولیدی شبکه مولد و داده واقعی شباهت ایجاد کند.

بیشتر مدل‌هایی که اخیرا به تشخیص ناهنجاری به وسیله شبکه‌های عصبی تخاصمی پرداخته‌اند مبتنی بر شبکه‌های عصبی دو طرفه هستند، اگرچه این مدل‌ها قادرند تا تصاویر هنجار را با امتیاز پایین ناهنجاری بازسازی کنند اما هیچ ضمانتی در اختصاص نمره بالای ناهنجاری به نمونه‌های ناهنجار وجود ندارد.

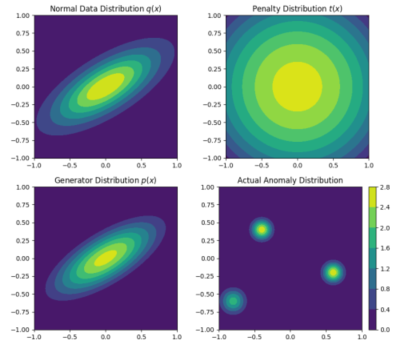
### 3-2-3- منظم‌سازی شبکه مولد و تمایزگر

برای رفع محدودیت‌های ذکر شده و قادر ساختن شبکه‌های عصبی تخاصمی برای تشخیص نمونه‌های ناهنجار از هنجار در این کار توزیع جریمه به گونه‌ای تعریف می‌شود که باشد، نمونه‌های تولیدی از این توزیع باید به عنوان داده تقلبی توسط تمایزگر شناخته شود. تابع هدف مدل پیشنهادی به صورت زیر است.

|  |  |
| --- | --- |
| **(3 -10)** |  |

تابع توزیع یک تابع تصادفی مانند توزیع گوسی انتخاب می‌شود. توزیع گوسی کاربرد‌های زیادی در زمینه‌ آموزش تخاصمی دارد. در انتهای آموزش شبکه‌های مولد و تمایزگر متمایل به توزیع داده‌های هنجار خواهد بود. در مدل پیشنهادی این مقاله نمونه که در آن  **نمونه ناهنجار است، نزدیک به توزیع داده هنجار خواهد، بنابراین بازسازی با خود فاصله خواهد داشت. این فاصله سبب می‌شود تا تشخیص نمونه ناهنجار آسان شود. همانطور که در شکل 3-3 مشخص است، نمونه‌های ناهنجار به گونه‌ای در مدل هدایت می‌شوند که بازسازی آن‌ها به سمت داده هنجار صفر متمایل شود. چنین رویکردی سبب می‌شود تا میان داده‌های ناهنجار و بازسازی آن‌ها تفاوت زیادی حاصل شود و در نتیجه تشخیص نمونه‌های ناهنجار گارانتی شود.**

**فرض کنید دارای پشتیبانی کمتر از مقدار در یک مکعب واحد است به ازای بعد‌های مختلف است، بنابراین داخل این ابعاد بزرگ و خارج از آن‌ها کوچک است. در مقابل دارای توزیع یکنواخت داخل یک کره واحد است. با این تنظیمات هنگامی که بزرگ است کوچک خواهد و هنگامی که بزرگ باشد کوچک خواهد بود و در این حالت پیش‌بینی‌های مدل بهبود خواهند یافت. مثال ساده دو بعدی موجود در شکل 3-3 را مشاهده کنید. در این شکل یک توزیع گوسی باریک است و تابع توزیع یکنواخت در است. توجه کنید از نمونه‌های تخاصمی مستقل است.**



10شکل 3-3: در این شکل نشان داده شده است که الگوریتم ارائه شده سبب می‌شود تا نقاطی که در آن‌ها کوچک است پوشش داده شود. این تنظیم مستقل از ارتباط میان و توزیع نمونه‌های ناهنجار انجام می‌شود.

### 4-2-3- پایداری چرخه

به منظور ارضای شرط پایداری چرخه تمایزگر به معماری مورد نظر اضافه شده است. تابع هدف چرخه پایداری به صورت زیر است.

|  |  |
| --- | --- |
| **(3 -11)** |  |

در معادله 3-11 بازسازی داده ورودی  **است. تابع هدف کامل ارائه شده در این کار به به شکل زیر است.**

|  |  |
| --- | --- |
| **(3 -12)** |  |

**پس از آموزش مدل نوبت به محاسبه امتیاز ناهنجاری می‌رسد، در مدل پشنهادی این کار تابع امتیاز پیشنهادی مطابق معادله زیر است.**

|  |  |
| --- | --- |
| **(3 -13)** |  |

امتیاز ناهنجاری بیانگر میزان کیفیت بازسازی  **است. الگوریتم ارائه شده در این کار مدل را مجبور به تولید خطای بزرگ برای نمونه‌های ناهنجار می‌کند در حالی که تابع هدف چرخه پایداری مدل مجبور به تولید بازسازی مناسب برای نمونه‌های هنجار می‌کند. این اختلاف امتیاز میان نمونه‌های هنجار و ناهنجار معیار مناسبی برای تشخیص نمونه‌های ناهنجار است.**

## 3-3- مدل RCALAD[[81]](#footnote-81)

بخش فعلی به معرفی مدل پیشنهادی اختصاص داده شده است. روش مورد بحث به منظور تشخیص ناهنجاری به دسته روش‌های مبتنی بر بازسازی تعلق دارد و از نظر دسترسی برچسب داده‌ها همانند سایر مدل‌های مبتنی بر شبکه‌های عصبی تخاصمی بررسی شده جزو دسته الگوریتم‌های بدون نظارت به حساب می‌آید. در این روش شبکه مولد موجود در ساختار GAN نگاشت از فضای نهان به فضای داده ورودی را فرا می‌گیرد، به بیان دیگر در این قسمت از شبکه عصبی تخاصمی توزیع داده هنجار ورودی مدل می‌شود. برای یادگیری نگاشت معکوس از فضای داده ورودی به فضای نهان از ساختار کدگذار-کدگشا استفاده شده است.

با افزودن تمایزگر به ساختار ارائه شده شرط چرخه پایداری در هر دو فضای نهفته و فضای ورودی در مدل پیشنهادی ارضا می‌شود اما تمایزگر مورد استفاده در کارهای قبلی، تنها به توزیع‌های منفرد داده‌ها در هر دو فضای نهفته و داده ورودی توجه می‌کردند و عملا برخی از جریان اطلاعات موجود در شبکه که متعلق به توزیع توامان متغیرهای فضای داده ورودی و نگاشت متناسب آن در فضای نهفته است، بلا استفاده باقی می‌ماند.

قابل توجه است که این روال و اثر بکاگیری این نوع تمایزگر توامان یک مرتبه پیش از این ثابت شده است؛ در واقع برای آموزش به هنگام افزودن شبکه کدگذار به ساختار اولیه GAN در هنگام معرفی مدل ALI، دو راه پیشرو بود. یک راه افزودن یک شبکه تمایزگر در کنار تمایزگر اولیه موجود، برای تمیز بین متغیرهای فضای پنهان بود و یک راه، تقویت شبکه تمایزگر اولیه و آموزش این شبکه به نحویست که توزیع توامان در هردو فضای نهان و واقعی را فرا بگیرند و بتوانند داده­های آموزشی را از داده­های تولیدشده توسط این مدل تشخیص دهد.

در مدل­های پیشین، جریان اطلاعات شامل یک فرایند دو مرحله­ایست، به این ترتیب که ابتدا از روی داده اولیه در فضای واقعی یک نگاشت توسط کدگذار به فضای نهان انجام می­شود و سپس از روی همین داده یک نگاشت معکوس به عنوان بازسازی به فضای داده اولیه توسط شبکه مولد انجام می­شود. سپس بار دیگر همین بازسازی به کدگذار فرستاده شده و در واقع بازسازی متغیر در فضای نهان بدست می­آید. در این مدل­ها متغیرها در هر کدام از فضاها با بازسازی آن­ها به صورت جداگانه مورد بررسی قرار می­گرفتند و تفکیک­شده تمیز داده می­شدند و بدین ترتیب ارتباط موجود بین این دو فضا نادیده گرفته می­شد.

بر مبنای استفاده از همین اطلاعات از دست رفته و سابقه بکارگیری توزیع­های توامان در این زمینه، در این جا مدل RCALAD پیشنهاد شده­است. در این­ ساختار در کنار چهارچوب اولیه شبکه­های مولد تخاصمی، یک شبکه تمایزگر توامان افزوده شده، تا با بکارگیری بیشترین اطلاعات موجود، مدل به جهت بهتری هدایت شده و سازگاری چرخه تا دو مرحله بررسی شود؛ یعنی برای آموزش مدل از اطلاعات هر دوفضا به صورت توامان استفاده شود و در نهایت مدل آموزش بهتری ببیند.

یک پیش­فرض اشتباه نیز در کارهای پیشین وجود داشت؛ این که اگر آموزش و بازسازی برای داده­های هنجار به خوبی انجام بگیرد، برای داده­های ناهنجار باید خراب شود و از داده اولیه ورودی متفاوت باشد. در صورتی که این فرض سهل­انگارانه است و هیچ استلزامی برای این قضیه در هیچ یک از مدل­های پیشین ارائه نشده است. در مدل ارائه شده با افزودن نمونه­گیری از کل فضا، این استلزام برای بازسازی در فضای هنجار ایجاد و مدل را به سمت فضای بازسازی داده­های هنجار متمایل کردیم[[82]](#footnote-82). در بخش­های آتی به بررسی جزئیات دقیق­تر این مدل می­پردازیم.

نتایج آزمایش­ها بر روی هر دو نوع داده تصویر و جدولی بیانگر کارایی و اثربخشی روش پیشنهادی می­باشد و نمایانگر سازگاری نتایج تئوری و عملی بدست­آمده برای این مسئله می­باشد.

### 1-3-3- معماری شبکه

در این ساختار همانند کارهای پیشین با هدف افزایش کارایی محاسباتی، یک کدگذار توام با شبکه مولد در ساختار کلی شبکه عصبی تخاصمی آموزش داده می‌شود. نگاشت معکوس از فضای داده ورودی به فضای نهفته به سادگی با تعبیه کدگذار در معماری پیشنهادی محاسبه می‌شود. این کدگذار تنها از فضای داده ورودی نمونه می‌گیرد و به طور تقریبی بازنمایی متناسب با آن را در فضای نهفته را تولید می‌کند. در اینجا برای آموزش هم‌زمان هر دو شبکه مولد و کدگذار از یک شبکه تمایزگر توامان با نام *Dxz* استفاده شده است. این تمایزگر بررسی می‌کند که جفت متغیر ورودی متعلق به توزیع داده ورودی و نقطه متناظر با آن در فضای نهفته است و یا توسط شبکه مولد و نمونه‌گیری از فضای نهفته z تولید شده است.

به منظور ارضای شرط پایداری حلقه از تمایزگر *Dxx* افزوده شده است، این تمایزگر به صورت توام نمونه داده ورودی و نمونه بازسازی شده متناظر آن را به عنوان ورودی اول و دوتایی به عنوان ورودی دوم دریافت می‌کند و برای تقویت شرط پایداری حلقه در فضای نهفته تمایزگر *Dzz* به این مجموعه اضافه شده است. این تمایزگر شرط پایداری حلقه را در خلال روند تولید نمونه بازسازی شده چک می‌کند. ورودی این تمایزگر نمونه ورودی شبکه مولد از فضای نهفته و خروجی کدگذار به عنوان ورودی اول و دوتایی است.

در این مدل یک شبکه تمایزگر جامع با نام *Dxxzz*  به ساختار فوق افزوده شده است. در این شبکه تلاش شده تا از تمامی اطلاعات موجود در مدل استفاده شود یعنی در کنار بررسی هر دو متغیر و بازسازی آن‌ها در همان فضا، توزیع توامان چهارتایی آن‌ها ( با بررسی جریان اطلاعات در دو گام ) در روند تشخیص نمونه ناهنجار به کار گرفته شود. این شبکه وظیفه تمییز بین نمونه‌های چهارتایی و را دارد به بیان دیگر تلاش می‌کند تا و بازسازی ارائه شده توسط شبکه و همینطور و بازسازی خروجی شبکه مولد توسط کدگذار تا حد امکان به یکدیگر نزدیک کند. هدف از تعبیه این تمایزگر در این ساختار حل مشکل چرخه پایداری توامان[[83]](#footnote-83) یا به اختصار JCC می‌باشد. تعریف دقیق مسئله JCC در ادامه بررسی می‌شود.

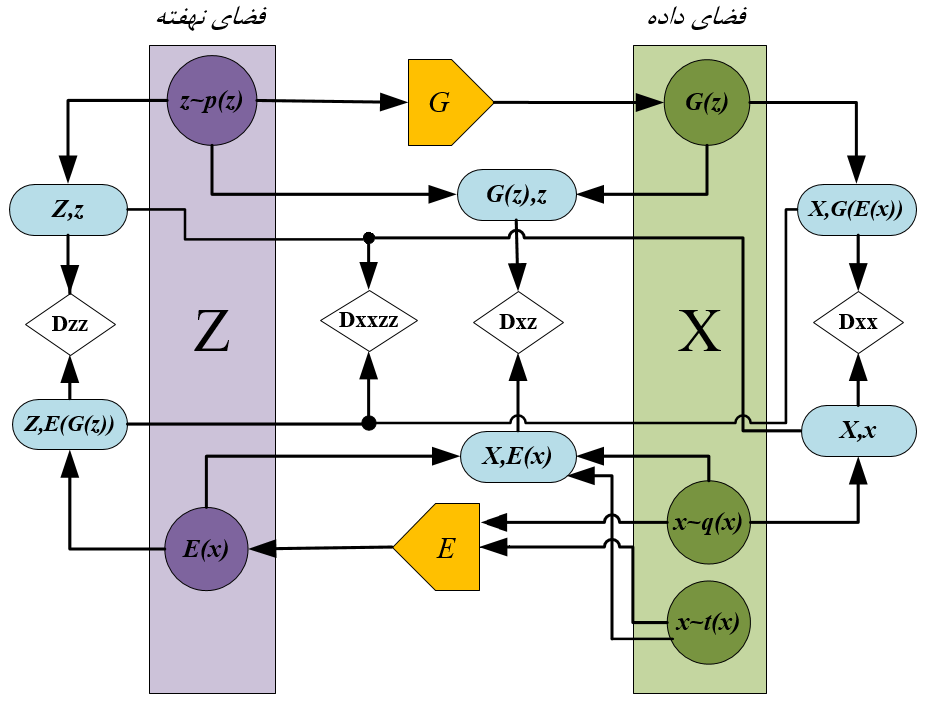
بیان ریاضی مسئله JCC بدین صورت می‌باشد. به ازای هر متغیر از فضای ورودی شبکه ابتدا کدگذار نگاشت معکوس به فضای نهفته را تخمین می‌زند که معادل می‌باشد. در مرحله بعد بازنمایی بدست آمده را به شبکه مولد وارد می‌کنیم تا بازسازی شبکه از متغیر ورودی تولید کند. سپس همین بازسازی را بار دیگر به شبکه کدگذار می‌دهیم تا بازسازی در فضای نهفته را نیز محاسبه شود یعنی . در این جریان، انتظار منطقی از هر شبکه مدعی بازسازی این است که دو متغیر و همچنین دو متغیر تا حد امکان کمترین اختلاف را داشته باشند. جریان اطلاعات در یک چرخه به صورت خلاصه در شکل 3-4 آمده است.

به منظور تقویت قدرت تشخیص تمایزگر اطلاعات حاصل از روند دگردیسی داده ورودی در تمامی مراحل چرخه، باید توسط تمایزگر قابل دسترس باشد. چرخه مورد نظر در این مسئله شامل سه گام متوالی است، در کارهای قبلی از خروجی‌‌های چرخه اطلاعات به طور کامل در شبکه استفاده نمی‌شد و این انتظار در جریان آموزش مدل­های مختلف ارائه شده با فرمولاسین مطرح­شده، مورد بررسی قرار نگرفته است. در حقیقت، مدل­های پیشین ارائه شده به بررسی مستقل این دو جفت متغیر در فضای جداگانه می­پرداخت و تمایزگر دید کاملی از جریان اطلاعات و وضعیت داده در هر دو فضا داده ورودی و نهفته به طور همزمان نداشت. قابل توجه است که این متغیرها پیش از این نیز محاسبه می­شد و این چرخه در مدل­های قبلی وجود داشت، اما مورد استفاده قرار نمی­گرفت. و از نظر بار محاسباتی، هزینه جدیدی به شبکه تحمیل نشده است.

برای حل این مسئله در این­جا شبکه تمایزگر *Dxxzz* پیشنهاد شده است. این شبکه هر چهار متغیر ( را به عنوان ورودی می­گیرد و تلاش می­کند تا از این مجموعه ورودی، ویژگی­های جدید و عمیق­تری ( به نسبت تمایزگرهای تک گامی *Dxx* و *Dzz* ) استخراج کند. این شبکه با در اختیار گرفتن خروجی کل چرخه داده، دید جامعی از وضعیت تمام قسمت‌های شبکه دارد و با استفاده از تمام این خروجی­ها به طور همزمان، به ویژگی­های ترکیبی قوی­تری برای تمیز نمونه‌های ناهنجار از نمونه داده‌های هنجار دست می­یابد.

این روند آموزش با همه مزیت‌هایی که نسبت به مدل‌های پیشین دارد، اما از یک مشکل اساسی چشم‌پوشی کرده است؛ مشکل استلزام بازسازی ضعیف. تعریف دقیق این مشکل بدین ترتیب می‌باشد: در تمامی روال‌های آموزش مدل‌های تشخیص ناهنجاری بازسازی دقیق نمونه‌های هنجار با کمترین خطا به مدل آموزش داده می‌شود، در مرحله آزمایش نمونه‌های ناهنجار و هنجار به مدل داده می‌شود، و همیشه فرض می­شود که برای نمونه‌های هنجار میزان اختلاف تصویر ورودی با تصویر بازسازی شده کم و برای نمونه‌های ناهنجار این اختلاف زیاد خواهد بود. در برخی موارد فرض فوق صحیح نیست و نمونه بازسازی شده داده ناهنجار، میزان اختلاف کمی با نمونه ورودی دارد و به همین سبب تشخیص آن به عنوان نمونه ناهنجار دشوار خواهد بود. در واقع در مدل‌های ارائه شده پیشین هیچ استلزامی برای بازسازی ضعیف نمونه ناهنجار وجود ندارد. علت وقوع این امر نگاشت تنک از فضای داده ورودی به فضای نهفته است. در حالت عادی آموزش فضای داده ورودی تنها به قسمت فضای کوچکی از نگاشت می‌شود، و در نتیجه نمونه‌گیری از فضای نهفته به منظور نگاشت دوباره به فضای داده ورودی تنک خواهد بود. در زمان مواجهه با نمونه‌های هنجار این امر مشکلی ایجاد نخواهد کر چرا که فضای متناسب زد برای نمونه‌های هنجار به خوبی مدل شده است ولی در نمونه‌های ناهنجار با توجه به اینکه مدل تا به حال چنین داده‌های را ندیده است ممکن است نمونه را به نقطه‌ای ناشناخته از فضای نهفته نگاشت کند و در نتیجه بازسازی نمونه ناهنجار نیز ممکن است به نقطه‌ای دورتر از فضای همیشگی بازسازی نگاشت شود. نگاشت به دست آمده از این فرایند هیچ ضمانتی برای ایجاد بازسازی مورد علاقه از نمونه ناهنجار ارائه نمی‌دهد.

بلوک به منظور پوشش حداکثری فضای نهفته در ساختاری کلی مدل تعبیه شده است. هدف از تعبیه این بلوک تولید نمونه‌های جدید در فضای داده ورودی و سپس نگاشت آن به فضای نهفته است. انتظار می‌رود در این روند فضای نهفته به شکل مناسب‌تری نسبت به کارهای قبلی پوشش داده شود. نتایج عملی نمایانگر صحت تئوری ارائه شده در این قسمت است. در نهایت شمای کلی مدل پیشنهادی RCALAD در شکل 3-4 قابل مشاهده است.

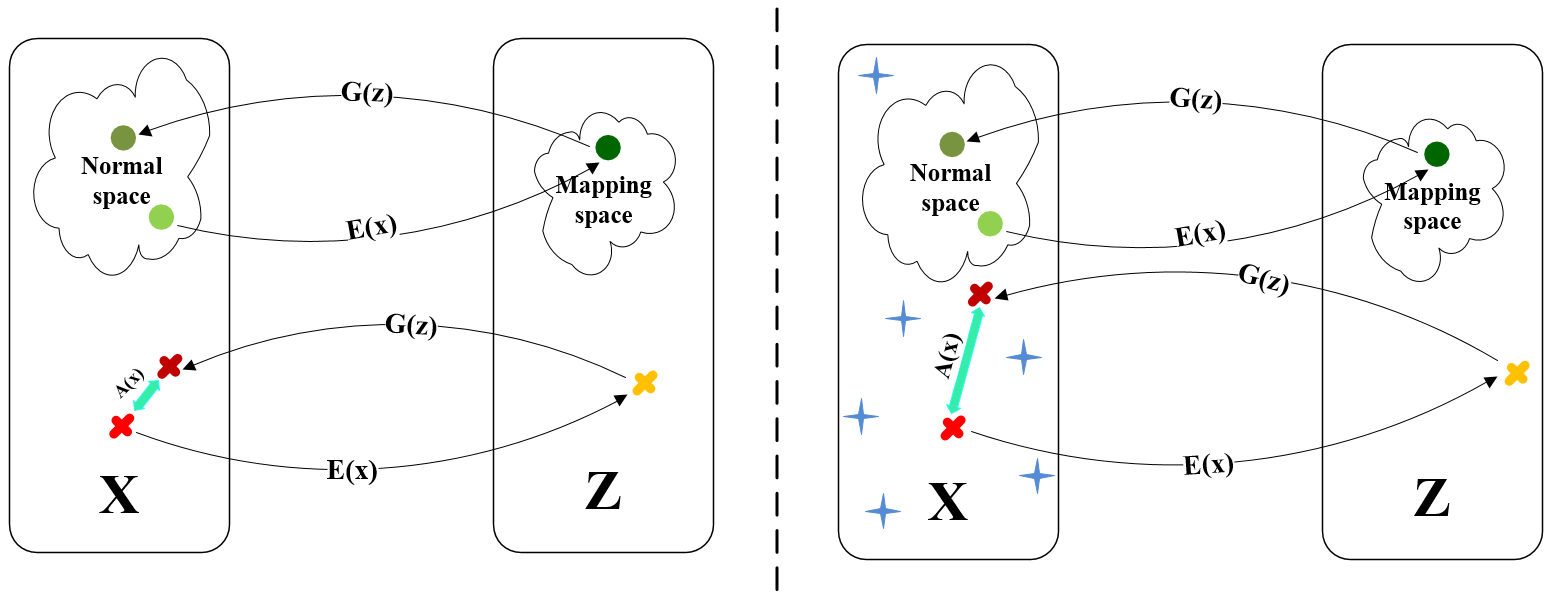


شکل3-4 معماری .RCALAD

در بخش 3-3-2 به توضیح دقیق­تر روال آموزش و بررسی جزئیات بلوک­های موجود در مدل خواهیم پرداخت.

### 2-3-3- روال آموزش

همانند سایر مدل‌های پیشین در روند آموزش مدل پیشنهادی از روال‌های آموزش تخاصمی استفاده می‌شود، بدیت ترتیب که به صورت متوالی بخش مولد و رمزگذار و پس از آن تمایزگر‌ها آموزش داده می‌شوند، یعنی به ترتیب با ثابت نگه داشتن وضعیت بخش مولد و رمزگذار، پارامترهای شبکه‌های تمایزگر را به‌روزرسانی می‌کنیم. سپس آموزش این بخش‌ها را متوقف کرده و با توجه به خروجی بهبود یافته آن‌ها، پارامترهای بخش مولد و رمزگذار به روزرسانی می‌شود و این روال بارها و بارها تکرار شده تا مدل به کیفیت خروجی مطلوب دست یابد.

هدف از آموزش تمامی مدل‌های مبتنی بر بازسازی از جمله مدل RCALAD تولید بازسازی مناسب برای داده‌های هنجار و بازسازی ضعیف برای نمونه داده‌های ناهنجار است. در اینجا برای ضمانت ایجاد بازسازی ضعیف برای داده ناهنجار، تلاش شده است تا تمامی فضای داده ورودی را به کمک نویز پوشش دهیم و مولد شبکه را به سمت بازسازی هر بیشتر نزدیک به توزیع داده هنجار متمایل کنیم. در شکل 3-3 چگونگی عملکرد این قسمت از مدل پیشنهادی و نحوه تاثیر آن در فرایند آموزش به تصویر کشیده شده است.

11شکل3-4: در این شکل x بیانگر فضای داده ورودی و z بیانگر فضای داده ورودی است. نمونه‌ها توسط مولد از فضای داده ورودی به فضای نهفته نگاشت می‌شوند و وظیفه انجام نگاشت معکوس بر عهده کدگذار است. دایره‌های سبز رنگ نماد نمونه داده‌های هنجار و ضربدر‌های قرمز نماد نمونه‌های ناهنجار هستند. بعلاوه آبی رنگ نشانگر نمونه‌های تولید شده توسط توزیع هستند که در تنها مرحله آموزش مورد استفاده قرار گرفته‌اند. فلش فیروزه‌ای مقدار امتیاز ناهنجاری را نشان می‌دهد .

همانطور که در شکل 3-4 مشاهده می‌شود در صورت عدم حضور t(x) ( در سمت چپ شکل ) در روند آموزش، امتیاز ناهنجاری برای نمونه­های غیرعادی کمتر از زمانی است که از این توزیع استفاده شده است، در واقع در تصویر سمت راست، توزیع t(x) مدل را به سمت بازسازی همه نمونه‌ها اعم از ناهنجار و هنجار به سمت توزیع داده‌‌های هنجار متمایل کرده است.

بیان ریاضی تابع هدف مدل پیشنهادی حاصل از جمع دو بخش کلی و مطابق معادله 3-14 می‌باشد.

|  |  |
| --- | --- |
| **(3 -14)** |  |

در ادامه به بررسی هدف و جزئیات ریاضی هر یک از این دو بخش خواهیم پرداخت.

اولین بخش از این تابع هدف می‌باشد. با هدف حل مسئله استلزام بازسازی (که در بخش قبلی به تعریف دقیق آن پرداختیم) و توانمندسازی شبکه پیشنهادی برای تمایز هر چه بهتر بین داده‌های هنجار و ناهنجار از توزیع جریمه استفاده شده است. ابعاد داده‌های خروجی این توزیع برابر با ابعاد داده ورودی می‌باشد و در حالتی که هیچ اطلاعات اضافه ای از داده نداشته باشیم، با توجه به این که داده­ها در مرحله پیش­پردازش نرمال شده اند، از توزیع گوسی نرمال برای پوشش کلی­تر فضا استفاده می­کنیم. البته اگر اطلاعات اضافه ای به مدل از توزیع داده­های ناهنجار داده شود، می­توان با بکارگیری به جای توزیع اعمال کرد. در حالت کلی بخش اول تابع هدف این مسئله طبق معادله 3-15 تعریف می­شود.

|  |  |
| --- | --- |
| **(3 -15)** |  |

در این معادله، داده‌های هنجار مورد استفاده در مرحله آموزش با تابع توزیع احتمال تعریف می‌شود و به عنوان تابع توزیع ورودی برای شبکه مولد در نظر گرفته شده است. در این معادله از تابع توزیع جریمه برای پوشش بهتر فضای داده ورودی و حل مشکل تنک بودن فضای نهفته استفاده می­شود. در این معادله با داده­های خروجی از توزیع و بازسازی آن­ها به عنوان نمونه­های تخاصمی برخورد می­شود و بدین ترتیب با افزودن عبارت سوم یعنی به تابع هزینه کلی، تلاش می­شود تا تمامی فضای ورودی به زیرفضا متعلق به توزیع داده­های هنجار در فضای نهفته نگاشت شود و با این روش بازسازی­های ارائه شده توسط شبکه به سمت توزیع داده هنجار ورودی متمایل شود. توضیح بیشتر این که ابتدا داده تولیدی از توزیع به شبکه کدگذار وارد می‌شود، در طی فرایند آموزش کدگذار می‌آموزد تمام فضای داده ورودی را به فضای متناسب توزیع داده هنجار در فضای نهفته نگاشت کند، در نتیجه این روند، توزیع داده ورودی شبکه مولد همواره از فضای متناسب با توزیع داده هنجار خواهد بود و خروجی حاصل نیز متعلق به توزیع متناسب با آن در فضای داده ورودی می­شود. از آن جایی که شبکه مولد نگاشت از فضای نهفته متعلق به داده هنجار به فضای داده ورودی متناسب را یاد گرفته است، تمامی فضای ورودی فارغ از هنجار یا ناهنجار بودن به فضای داده هنجار نگاشت می­شود. پس به ازای داده­های ناهنجار اختلاف بازسازی ارائه شده توسط شبکه بیشتر از داده­های هنجار خواهد بود.

**برای اثبات نظری تمایل شبکه‌های مولد و تمایزگر به سمت داده‌های هنجار ابتدا تمایزگر و مولد بهینه را آموزش می‌دهیم. توزیع توام کدگذار روی داده‌های هنجار به صورت**  و توزیع توام کدگذار روی *داده*‌*های ناشی از توزیع جریمه به صورت و است*. نقطه بهینه برای تمایزگر *Dzz* که از معادله 3-15 بدست می­آید مطابق معادله 3-16 می­باشد.

|  |  |
| --- | --- |
| **(3 - 16)** |  |

در معادله 3-16 هر دو توزیع داده ورودی و توزیع داده جریمه در نظر گرفته است. از معادله فوق می‌توان نتیجه گرفت بر خلاق GAN استاندارد که تنها روی نمونه‌های عادی آموزش می‌بیند تمایزگر مدل ارائه شده در این کار احتمال بیشتری به نمونه‌های عادی اختصاص می‌دهد، و با توجه به این که به نمونه‌های ناهنجار با کوچکتر اختصاص می‌یابد، خروجی تمایزگر بهینه برای داده­های ناهنجار کم می­شود. نتایج عملی ارائه شده در فصل چهارم گزاره‌های ارائه شده در این قسمت را پشتیبانی می‌کنند.

بخش دوم تابع هدف مورد استفاده در تابع هزینه این مدل تلاش می­کند تا شباهت میان داده تولیدی و بازسازی آن توسط شبکه مولد برای داده­های هنجار را تضمین کند. در واقع این تابع هدف شرط پایداری چرخه را هم به صورت تک مرحله­ای و هم به صورت دو مرحله ای یعنی چرخه کامل ارضا می‌کند. این بخش مطابق معادله 3-17 فرمول­بندی می­شود.

|  |  |
| --- | --- |
| **(3 -17)** |  |

در معادله 3-17 که متعلق به بخش چرخه پایداری تابع هدف می­باشد، سطر اول مربوط به تمایزگر *Dxx* می­باشد. این تمایزگر تلاش می­کند که شرط چرخه پایداری را به صورت تک گامی با شروع از فضای داده در این مدل ایجاد کند. سطر دوم مربوط به تمایزگر *Dzz* است که با هدف ایجاد چرخه پایداری تک گامی با شروع از فضای نهفته به مدل اضافه شده است. و در نهایت دو عبارت آخر مربوط به تمایزگر جامع*Dxxzz*  است. این تمایزگر در دو گام جریان اطلاعات را دنبال می­کند و خروجی تمامی مراحل را بررسی می­کند و چرخه پایدار را برای همه مراحل نگاشت داده در مدل ایجاد می­نماید. همان­طور که در بخش قبل نیز اشاره شد، هدف از تعبیه این تمایزگر در این ساختار، حل مشکل چرخه پایداری توامان[[84]](#footnote-84) یا JCC می‌باشد که تعریف دقیق این مسئله JCC در بخش 3-3-1 اشاره شد.

### 3-3-3- تشخیص ناهنجاری

همانطور که گفته شد مدل پیشنهادی در این پروژه به منظور تشخیص ناهنجاری بر اساس بازسازی داده ورودی است. داده هنجار به صورت دقیق بازسازی می‌شود در حالی که بازسازی نمونه ناهنجار ضعیف خواهد بود. پس از بازسازی نمونه‌ها، عنصر کلیدی در تشخیص ناهنجاری تعریف امتیاز ناهنجاری با هدف محاسبه فاصله میان نمونه ورودی و خروجی بازسازی شده توسط شبکه است. اولین انتخاب در زمینه محاسبه میزان فاصله این دو داده فاصله اقلیدسی است ولی در فضای داده این معیار ممکن است به اندازه کافی قابل اتکا نباشد[6]. به عنوان مثال در زمینه تشخیص ناهنجاری در تصاویر ممکن است علی‌رغم ویژگی‌های بصری مشابه فاصله اقلیدسی زیادی داشته باشند.

در این کار به جای محاسبه فاصله میان نمونه‌ها در فضای داده ورودی، از فضای ویژگی موجود در تمایزگر استفاده می‌شود. به این منظور خروجی توابع فعالیت[[85]](#footnote-85) لایه یکی مانده به آخر به عنوان ویژگی استفاده می‌شوند. امتیاز ناهنجاری مورد استفاده به صورت زیر و با استفاده از خطای بازسازی نرم یک و مطابق معادله 3-18 تعریف می‌شود.

|  |  |
| --- | --- |
| **(3 -18)** |  |

*در معادله 3-18 بیانگرتابع فعالیت لایه یکی مانده به آخر در ساختار تمایزگر است. میزان اطمینان تمایزگر از کیفیت روند کدگذاری و بازسازی داده توسط مولد است که اگر خوب انجام شده باشد در واقع نمونه متعلق به توزیع داده هنجار است که مدل روی آن آموزش دیده است. بنابر مطالب گفته شده هر چه مقدار این معیار بیشتر باشد احتمال ناهنجاری بودن آن بیشتر است. الگوریتم محاسبه امتیاز ناهنجاری در الگوریتم 3-1 قابل مشاهده است. عملکرد مناسب این معیار در مقایسه با سایر معیارهای تشخیص ناهنجاری در فصل آینده نمایش داده شده است.*

|  |
| --- |
| Algorithm 2 Regularized Comprehensive Adverarially Learned Anomaly Detection |
| Input |

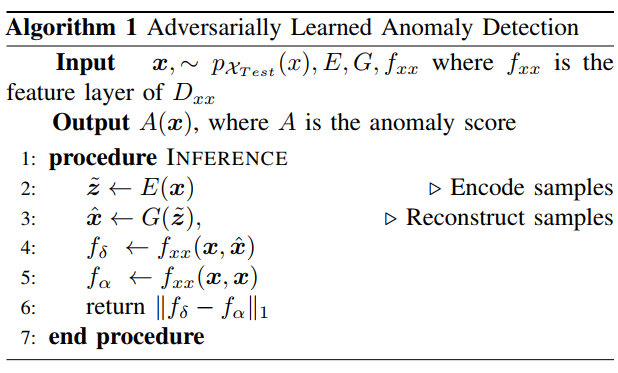
[[86]](#footnote-86)

Figure 12الگوریتم 3-1: روند محاسبه امتیاز ناهنجاری.

معیار پیشنهادی در این قسمت بر اساس روش تطبیق ویژگی[[87]](#footnote-87) و یا به اختصار FM، پایه‌گذاری شده است [11] ، در ساختار GAN استاندارد برای تطبیق ویژگی از خروجی شبکه مولد استفاده شده است و در مقاله[6] اگرچه از خروجی تمایزگر استفاده کرده است اما لاجیت‌های تولیدی توسط این شبکه از نظر اطلاعات نسبت به تمایزگر فقیرتر هستند.

در مدل ALAD تنها از خروجی تمایزگر برای شناسایی نمونه‌های ناهنجار استفاده شد و بنابر نتایج به‌دست آمده خروجی این تمایزگر فاقد اطلاعات مفید است و نمی‌توان بر اساس آن نمونه­های ناهنجار را تشخیص داد اما در این کار از جمع خروجی تمایزگر‌های ، و به عنوان معیار تشخیص ناهنجاری استفاده شد. تمام تمایزگر‌های موجود در مدل پیشنهادی تنها روی نمونه‌های هنجار آموزش دیده‌اند پس انتظار می‌رود تصویر بازسازی شده نمونه ناهنجار و همچنین بازنمایی آن در فضای نهفته که توسط کدگذار تولید می‌شود ضعیف باشد، علاوه بر این‌ها خروجی بر اساس ویژگی‌های غنی‌تری تولید می‌شود و طبیعی‌ست که از اهمیت بالایی برخوردار باشد. حال مسئله قابل بررسی این موضوع می‌باشد که آیا معیار حاوی اطلاعات مفید کافی برای تشخیص داده‌های هنجار از ناهنجار می‌باشد یا خیر. پاسخ به این سوال در حالت کلی بله می‌باشد زیرا این تمایزگرها در طی فرایند آموزش یاد می‌گیرند که به اختلاف میان دوتایی و و همچنین دوتایی و توجه کنند یعنی هر چه از و یا از فاصله بگیرد، تشخیص آن برای تمایزگرها ساده­تر می­شود. حال با افزودن توزیع به این مجموعه و تلاش برای متمایل کردن تمامی بازسازی به سمت توزیع داده­های هنجار، اختلاف بازسازی برای داده­های ناهنجار زیادتر شده و عملاً خروجی تمایزگرها یک معیار قابل اتکا برای تشخیص ناهنجاری می­شود. بیان ریاضی این معیار در معادله 3-19 آورده شده است.

|  |  |
| --- | --- |
| **(3 -19)** |  |

در فصل آینده این معیار با سایر معیار‌های تشخیص ناهنجاری مقایسه شده و در عمل کارایی این امتیاز نشان داده می­شود.

# فصل چهارم جمع‌بندي و نتيجه‌گيري جمع‌بندي و نتيجه‌گيري

در این تحقیق سعی برآن است که جمع­بندی مختصری نسبت به محاسبات تقریبی، چالش­ها، دسته­بندی­ها، سطوح پیاده­سازی، معیارهای کیفیت­سنجی و تکنیک­های حل مسائل محاسبات تقریبی ارایه شود.

سوال: مجموعه مقالات مرور شده مربوط به چه بازه زماني است و در چه مجلات/كنفرانسهايي چاپ شده­اند؟ هيستوگرام فراواني زماني و مكاني مقالات را رسم كنيد.

جواب: پاسخ این قسمت در بخش 2-5 در شکل­های 2-1 و 2-2 آمده­است.

سوال: مجموعه مقالات مرور شده از چه جنبه­هايي با يكديگر شباهت يا تفاوت دارند؟

جواب: پاسخ این قسمت در بخش 2-4 شرح داده­شده است.

سوال: دسته­بندي ارائه شده براي مقالات چه جنبه­هايي را در نظر گرفته­است و چه جنبه­هاي ديگري را مي­توانست در نظر بگيرد

جواب: پاسخ این قسمت در بخش 2-5 به طور کامل آمده­است.

سوال: چالش­هاي اصلي اين حوزه كه مقالات بيشتر سعي در رفع نمودن آنها داشته­اند كدامند؟

جواب: پاسخ این قسمت در بخش 2-2 به طور کامل بحث­شده­است.

سوال: تكنيكهاي ارائه شده در اين حوزه كدامند؟

جواب: پاسخ این قسمت در بخش 2-4 به طور کامل آورده شده­است.

سوال: به نظر شما، چه جنبه­هايي در كارهاي گذشته مغفول مانده و بايد بيشتر مورد توجه قرار مي­گرفته­است؟

جواب: پاسخ این قسمت در بخش 2-4 به طور کامل آمده­است.

سوال معيارهاي كيفيت­سنجي تكنيك­هاي ارائه شده در اين حوزه كدامند

جواب: پاسخ این قسمت در بخش 2-3 به طور کامل آمده­است.

سوال: در اين حوزه پژوهشي، از بين انرژي، توان و دما كداميك بيشتر مورد توجه است؟

جواب: پاسخ این قسمت در بخش 2-5 به طور کامل آمده­است.

سوال: در اين حوزه پژوهشي، به غير از انرژي/توان/دما، چه معيارهاي ديگري نيز مورد توجه قرار­گرفته­است؟

جواب: پاسخ این قسمت در بخش 2-5 به طور کامل آمده­است.

سوال: كاهش انرژي/توان/دما اولويت چندم بوده است؟

جواب: پاسخ این قسمت در بخش 2-5 به طور کامل آمده­است.

سوال: از چه روشها يا تكنيكهاي رياضي حل مسئله در اين حوزه استفاده شده كه به نظر شما جذاب است؟

جواب: پاسخ این قسمت در بخش 2-3 به طور کامل آمده­است.

سوال: سطح تجريد تكنيك­هاي ارائه شده در اين حوزه چيست؟ آيا اطلاعاتي از سطوح ديگر تجريد نيز در اين تكنيك­ها استفاده شده است؟

جواب: پاسخ این قسمت در بخش 4-2 به طور کامل آورده شده­است.

سوال: مهمترين و باارزشترين مقاله از بين مقالات مرور شده از نظر شما كدام است؟

جواب: پاسخ این قسمت در فصل سوم به طور کامل شرح داده­شده­است.

# منابع و مراجع

[1] S. Mittal, “A Survey Of Techniques for Approximate Computing”, 2015. ACM Comput. Survey.

[2] Hadi Esmaeilzadeh, Adrian Sampson, Luis Ceze, and Doug Burger. 2012. Neural acceleration for general purpose approximate programs. In IEEE/ACM International Symposium on Microarchitecture. 449–460.

**Abstract**

Approximate computing trades off computation quality with the effort expended and as rising performance demands confront with plateauing resource budgets, approximate computing has become, not merely attractive, but even imperative. In this paper, we present a survey of techniques for approximate computing (AC). We discuss strategies for finding approximable program portions and monitoring output quality, techniques for using AC in different processing units, processor components, memory technologies etc., and programming frameworks for AC. We classify these techniques based on several key characteristics to emphasize their similarities and differences. The aim of this paper is to provide insights to researchers into working of AC techniques and inspire more efforts in this area to make AC the mainstream computing approach in future systems.

**Key Words:** Approximate Computing Technique; quality configurability.



Amirkabir University of Technology   
(Tehran Polytechnic)

Department of Computer Engineering and Information Technology

Low Power Design   
Final Report

A Survey Of Techniques for

Approximate Computing

By

Seyed Salar Hashemi Taheri

July 2019

1. Encoder [↑](#footnote-ref-1)
2. Autoencoder [↑](#footnote-ref-2)
3. Generative adversarial networks [↑](#footnote-ref-3)
4. Joint [↑](#footnote-ref-4)
5. Latent [↑](#footnote-ref-5)
6. Tabular [↑](#footnote-ref-6)
7. Street view house number [↑](#footnote-ref-7)
8. Generative Adversarial Networks [↑](#footnote-ref-8)
9. Unsupervised anomaly detection with generative adversarial networks to guid marker [↑](#footnote-ref-9)
10. Fast unsupervised anomaly detection with generative adversarial networks [↑](#footnote-ref-10)
11. Adversarially Learned Inference [↑](#footnote-ref-11)
12. Efficient GAN-Based Anomaly Detection [↑](#footnote-ref-12)
13. Cycle consistency [↑](#footnote-ref-13)
14. ALICE [↑](#footnote-ref-14)
15. Adversarially Learned Anomaly Detection [↑](#footnote-ref-15)
16. Clustering [↑](#footnote-ref-16)
17. False positive rate [↑](#footnote-ref-17)
18. Residual [↑](#footnote-ref-18)
19. Maximum likelihood estimation [↑](#footnote-ref-19)
20. Chi-square [↑](#footnote-ref-20)
21. kernel [↑](#footnote-ref-21)
22. Support vector machine [↑](#footnote-ref-22)
23. Posterior [↑](#footnote-ref-23)
24. Decision tree [↑](#footnote-ref-24)
25. Generative Adversarial Networks (GAN) [↑](#footnote-ref-25)
26. Reinforcement learninig [↑](#footnote-ref-26)
27. Reconstruction [↑](#footnote-ref-27)
28. Principal component analysis [↑](#footnote-ref-28)
29. Variational autoencoder [↑](#footnote-ref-29)
30. Deep autoencoding gaussian mixture models [↑](#footnote-ref-30)
31. Gradient descent [↑](#footnote-ref-31)
32. Detection rate [↑](#footnote-ref-32)
33. Accuracy [↑](#footnote-ref-33)
34. Performance [↑](#footnote-ref-34)
35. Scalability [↑](#footnote-ref-35)
36. True positive [↑](#footnote-ref-36)
37. True negative [↑](#footnote-ref-37)
38. False negative [↑](#footnote-ref-38)
39. Precision [↑](#footnote-ref-39)
40. Receiver operating characteristics [↑](#footnote-ref-40)
41. Discriminator models [↑](#footnote-ref-41)
42. Backpropagation [↑](#footnote-ref-42)
43. Dropout [↑](#footnote-ref-43)
44. Piecewise linear units [↑](#footnote-ref-44)
45. Feed forward [↑](#footnote-ref-45)
46. Minimax [↑](#footnote-ref-46)
47. Scaler [↑](#footnote-ref-47)
48. Overfitting [↑](#footnote-ref-48)
49. Minibatch [↑](#footnote-ref-49)
50. Log-Liklihood [↑](#footnote-ref-50)
51. Jensen-Shanon divergence [↑](#footnote-ref-51)
52. Nash cost [↑](#footnote-ref-52)
53. True Negative Rate [↑](#footnote-ref-53)
54. Feature matching [↑](#footnote-ref-54)
55. Convolutional [↑](#footnote-ref-55)
56. Residual Loss [↑](#footnote-ref-56)
57. Discrimination loss [↑](#footnote-ref-57)
58. Sigmoid cross entropy [↑](#footnote-ref-58)
59. logits [↑](#footnote-ref-59)
60. Deep convolutional generative adversarial network [↑](#footnote-ref-60)
61. Wasserstein GAN [↑](#footnote-ref-61)
62. Residual [↑](#footnote-ref-62)
63. Mean squared error [↑](#footnote-ref-63)
64. Encoder [↑](#footnote-ref-64)
65. Tuple [↑](#footnote-ref-65)
66. Feedforward [↑](#footnote-ref-66)
67. Interpretable Representation Learning by Information Maximizing Generative Adversarial Nets [↑](#footnote-ref-67)
68. Mutual information [↑](#footnote-ref-68)
69. Posterior [↑](#footnote-ref-69)
70. Bidirectional Generative Adversarial Model [↑](#footnote-ref-70)
71. Cross entropy [↑](#footnote-ref-71)
72. Feature matching loss [↑](#footnote-ref-72)
73. Conditional entropy [↑](#footnote-ref-73)
74. Cycle Consistency [↑](#footnote-ref-74)
75. L-norm [↑](#footnote-ref-75)
76. Threshold [↑](#footnote-ref-76)
77. Invert [↑](#footnote-ref-77)
78. Lipschitz Constraints [↑](#footnote-ref-78)
79. Activation [↑](#footnote-ref-79)
80. Regularized Cycle onsistent Generative Adversarial Network for anomaly detection [↑](#footnote-ref-80)
81. Regularized Comprehensive Adversarialy Learned Anomaly Detection [↑](#footnote-ref-81)
82. Bias [↑](#footnote-ref-82)
83. Joint Cycle Consistency [↑](#footnote-ref-83)
84. Joint Cycle Consistency [↑](#footnote-ref-84)
85. Activation function [↑](#footnote-ref-85)
86. [↑](#footnote-ref-86)
87. Feature Matching [↑](#footnote-ref-87)