|  |  |
| --- | --- |
| نام و نام خانوادگی | **پویا شیخ الاسلامی** |
| شماره دانشجویی | **810100394** |
| تاریخ ارسال گزارش | **۱۴۰2.۰4.۰۱** |

|  |  |  |
| --- | --- | --- |
|  | **به نام خدا**  **دانشگاه تهران**  **دانشکده‌ مهندسی برق و کامپیوتر** |  |
| **درس شبکه‌های عصبی و یادگیری عمیق**  **تمرین اول** | | |

**فهرست**

[**پاسخ 1**. **شبکه های رمزگذار-رمزگشا مولد** 1](#_Toc138442117)

[۱-۱. مجموعه دادگان مقاله 1](#_Toc138442118)

[2-۱. انجام PCA و ISOMAP 1](#_Toc138442119)

[3-۱. رمزگذار – رمزگشا 3](#_Toc138442120)

[4-۱. خود رمزگذار متغیر((**Variational AutoEncoder** 6](#_Toc138442128)

[5-۱. کاوش در فضای latent 8](#_Toc138442130)

[6-۱. کاوش در فضای latent 12](#_Toc138442131)

[**پاسخ ۲** **- شبکه ی متخاصم مولد** 14](#_Toc138442132)

[۱-۲. **بارگذاری دادهها و شبکه ی** ResNet 14](#_Toc138442133)

[الف: رسم نمودار دقت برای دادههای آموزشی و اعتبار سنجی و ارزیابی مدل 16](#_Toc138442134)

[ب: ترسیم ماتریس آشفتگی 17](#_Toc138442135)

[2-۲. **شبکهی** Conditional DCGAN 18](#_Toc138442136)

[الف: گزارش ساختارها و پارامترهای مورد استفاده برای شبکههای مولد و تفکیککننده 20](#_Toc138442137)

[ب: رسم و تفسیر نمودار Loss برای مولد و تفکیک کننده 20](#_Toc138442138)

[پ: تولید 2000 نمونه توسط مولد به ازای هر کلاس داده و نمایش آنها 20](#_Toc138442139)

[ت: ارائه راهکار جهت بهتر شدن خروجی مولد و پایدارسازی 21](#_Toc138442140)

[3-۲. **طبقه بندی به کمک داده های تولید شده توسط مولد** 21](#_Toc138442141)

**شکل‌ها**

شکل 1. نمایی کلی از دیتاست ورودی 1

**جدول‌ها**

[جدول 1. عنوان جدول نمونه 1](#_Toc115453324)

# **پاسخ 1**. **شبکه­های رمزگذار-رمزگشا مولد**

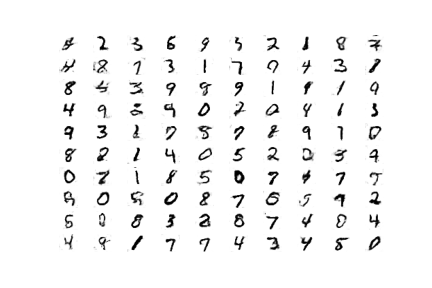
در این پرسش هدف پیاده سازی یک Variational Auto Encoder و مقایسه قدرت آن در کاهش ابعاد

با روش­های پیشین مانندPCA و ISOMAPو Decoder-Encoderاست.

۱-۱. مجموعه دادگان مقاله

مجموعه داده استفاده شده در این بخش، مجموعه داده Fashion\_mnist می­باشد که یک مجموعه داده با 60000 داده آموزشی و 10000 داده تست به 10 کلاس متفاوت مانند T-shirt/Top، Trouser، pullover، Dress و ... است. تصاویر این مجموعه داده دارای سایز 28\*28 پیکسل و خاکستری می­باشند.

برای مرحله پیش پردازش داده، همان­طور که در شکل زیر دیده می­شود،



شکل 1. نمایی کلی از دیتاست ورودی

2-۱. انجام PCA و ISOMAP

**هدف: در این بخش هدف بررسی تعاریف PCA و ISOMAP، مقایسه شباهت و تفاوت­های آن­ها و سپس کاهش ابعاد دیتای بالا و آموزش آن به کمک KNN و تست است.**

**به صورت کلی، PCA یک روش قدرتمند برای کاهش ابعاد داده­های یک دیتاست با شرط حفظ اطلاعات اصلی تصاویر، است. مراحل کاهش ابعاد در روش PCA به این ترتیب است که ابتدا مقادیر اولیه متغیرهای ورودی را استاندارد نموده، کواریانس را برای بررسی میزان هم­بستگی بین آن­ها محاسبه می­نماید. سپس مقادیر و بردار ویژه بخش های اصلی را محاسبه و با استفاده از آن­ها بردار ویژگی­های اجزای اصلی ساخته می­شود. در نهایت به کمک بردار ویژگی نهایی، برخی قسمت­ها که حاوی اطلاعات مهم هستند بازنویسی می­شوند. در این روش رابطه بین ویژگی­های کاهش یافته با متغیرهای اصلی، یک رابطه خطی است.**

**از طرفی دیگر،** ISOMAP **یک روش کاهش ابعاد غیر خطی بر اساس تئوری** Spectral **است که با ساخت یک شبکه** Neighborhood **از گراف فاصله برای تخمین فاصله­ی** geodesic **بین هر دو جفت نقطه و کاهش ابعاد بر اساس آن­ها استفاده می­نماید. در مجموع در هر دو روش بیان شده، تلاش می­شود با حفظ اطلاعات اصلی ، ابعاد ورودی کاهش پیدا کند که این مساله به نوعی باعث کاهش مقدار دقت می­شود اما در مقابل داده­ها ساده­تر شده و تجزیه و تحلیل و پردازش آن­ها راحت تر انجام می­گیرد.**

**اصلی ترین تفاوت** PCA **و** ISOMAP **در این می­باشد که چنان­چه رابطه بین غیر خطی در مجموعه داده ورودی وجود داشته باشد،** PCA **نمی­تواند مدل مناسبی برای کاهش ابعاد آن ارائه دهد و بهتر است از** ISOMAP **استفاده گردد، زیرا** PCA **برای موادی مناسب­تر است که رابطه خطی بین داده­ها وجود دارد. از طرفی روش** ISOMAP **به جای فاصله سنج اقلیدسی از روش فاصله سنج geodesic استفاده می­نماید که همین مساله به نوعی آن را در برابر داده­ای غیرخطی مقاوم می­نماید.**

از مزایا و معایب PCA می­توان به موارد زیر اشاره نمود که ابعاد داده ها را کاهش می­دهد و تجسم داده های پیچیده را در فضای با ابعاد کمتر آسان­تر می­کند. PCA یک تکنیک بدون نظارت است که به داده های برچسب­دار نیاز ندارد. این امر آن را برای تجزیه و تحلیل داده­های اکتشافی مفید می­نماید.هم­چنین PCA ساختار داده­ها را تا حد امکان حفظ نموده و خروجی­های قابل تفسیر ارائه می­نماید همچنین می توان آن را به راحتی با استفاده از ابزارها و کتابخانه­های استاندارد پیاده سازی و تفسیر نمود. با این حال، PCA محدودیت­ها و اشکالاتی برای تجسم داده­ها نیز دارد. یکی از معایب اصلی PCA این است که می تواند برخی از اطلاعات و جزئیات را هنگام کاهش ابعاد داده­ها از دست بدهد که این مساله می‌تواند منجر به ساده‌سازی یا تحریف بیش از حد داده‌ها شده و شناسایی موارد پرت یا ناهنجاری را دشوارتر نماید. برای مثال، ممکن است برخی از الگوها یا روندهای ظریف را که فقط در ابعاد بالاتر قابل مشاهده هستند،از بین بروند. یکی دیگر از معایب PCA این است که می تواند به نقاط پرت حساس باشد که این حساسیت می­تواند کیفیت و پایداری نتایج را تحت تاثیر قرار دهد. بنابراین، باید قبل از اعمال PCA، داده ها را با دقت پیش پردازش و عادی سازی نمود. هم­چنین در کنار تمامی مزایا و معایب بالا که برای ISOmaP نیز برقرار هستند، از دیگر مزایای ISOMAP می­توان به کاربرد آن برای دیتاهای با رابطه غیر خطی و تعداد کلاس بالا نیز اشاره نمود و در مقابل داشتن پارامتر و تاثیر انتخاب آن بر روی نتیجه نهایی می­تواند از معایب این روش کاهش بعد باشد.

اکنون مطابق خواسته سوال، در ابتدا از روش­های گفته شده برای کاهش ابعاد داده­ی Fashion-Mnist استفاده شده ­است که در مجموع 784 ویژگی داده به 20 الی 30 ویژگی کاهش یافته­اند. سپس داده­های کاهش یافته به شبکه KNN داده شده­اند تا طبقه­بندی صورت گرفته و نتایج به دست آیند.

شبکه KNN یا همان K-nearest neighbors همان­طور که در شکل زیر مشاهده می­گردد، یک الگوریتم برای سنجش فاصله­ی بین هر دو داده بین نمونه­های آموزشی و تست است که تلاش می­کند تا کمترین فاصله ها وجود داشته باشند. هم­چنین K بسته به تعداد کلاس­ها می­تواند مقداری بین 1 تا n را انتخاب نماید.



شکل 2:مثالی کلی از اعمال KNN بر روی یک دیتاست

در نهایت نتایج را داریم:

برای pca نتایج طبقه بندی knn به صورت زیر می باشد

دقت بر روی داده های آموزش ۰.۹

دقت بر روی داده های تست ۰.۸۶

برای isomap نیز نتایج به صورت زیر است:

دقت بر روی داده های آموزش ۰.۷

دقت بر روی داده های تست ۰.۶۸

3-۱. رمزگذار – رمزگشا

**هدف**: در این بخش هدف ساختن رمزگذار-رمزگشا با استفاده از Dense Layer و Concolutional Layer ها برای آموزش مجموعه داده گفته شده، کاهش بعد آن به وسیله رمزگذار و آموزش و طبقه بندی به وسیله KNN است.

**رمزگذار خودکار(Autoencoder) یک مدل شبکه عصبی می­باشد که از آن برای کاهش ابعاد و فشرده سازی داده­ها و بازسازی مجدد آن­ها استفاده می­شود. رمزگذارهای خودکار از دو بخش رمزگار و رمزگشا تشکیل می­شوند که بخش رمزگذار اقدام به کاهش بعد و فشرده­سازی نموده و رمزگشا نیز خروجی رمزگذار را بازسازی می­نماید.**

**به صورت کلی ، Autoencoder ها بر خلاف PCA ها که برای داده­های خطی عملکرد خبی دارند و برای داده­های غیرخطی موثر نیستند، به راحتی می­توانند برای داده­های غیرخطی و پیچیده موثر واقع شده و با داده­های غیرخطی به خوبی کار کنند.**

**اکنون جهت برآورده سازی اهداف سوال، ابتدا یک رمزگذار خودکار برای لایه Dense و یک رمزگذار خودکار برای CNN نوشته که ورودی آن­ها داده­های ورودی و خروجی نیز یک تصویر می­باشد که اختلاف دو تصویر ورودی و خروجی اندازه گرفته می­شود.**

**در واقع در این مسیر بیان شده، ابعاد تصویر 2 بار کاهش پیدا می­کند به این ترتیب که تصاویر در مدل رمزگذار خودکار آموزش داده شده، ابعاد کاهش یافته و باز تولید می­شوند. در نهایت خروجی این بخش به عنوان ورودی به شبکه KNN برای آموزش مورد استفاده قرار گرفته و نتایج بررسی می­گردند.**

**برای بخش Dense با دو مقدار نهفته نتایج طبقه بندی به صورت زیر می باشد:**

**با ابعاد نهفته ۲۰ بر روی داده های آموزش ۷۲درصد و داده های تست ۶۸**

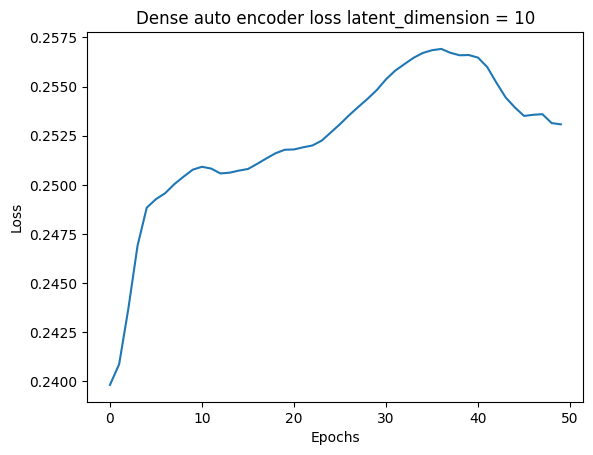
**با اباد نهفته ۱۰ بر روی داده های آموزش ۶۲درصد و داده های تست ۵۹درصد**

**برای بخش CNN نیز نتایج برای ۲ نوع فیلتر سایز ۱۶ و ۳۲ داریم:**

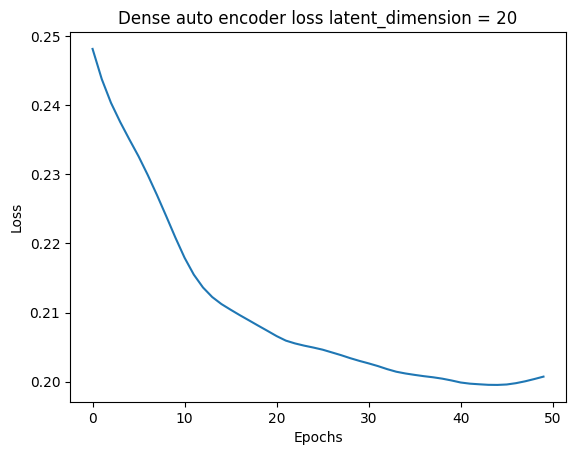
**برای فیلتر سایز ۱۶ بر روی داده های آموزش ۸۶ درصد و داده های تست ۸۰**

**برای فیلتر سایز ۳۲ داده های آموزش ۸۶ و داده های تست ۸۱**

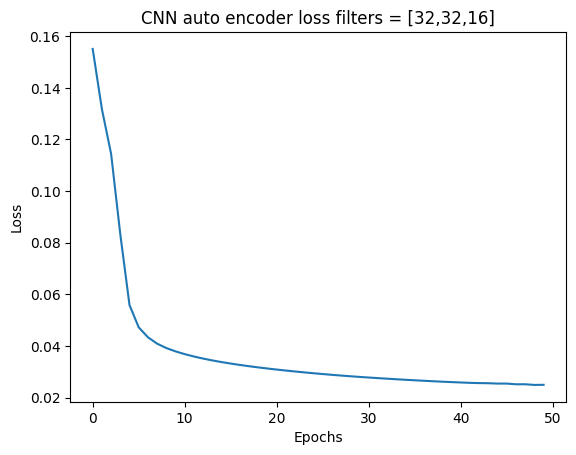
هم­چنین نمودار خطا و دقت شبکه را به فرم زیر داریم:

****

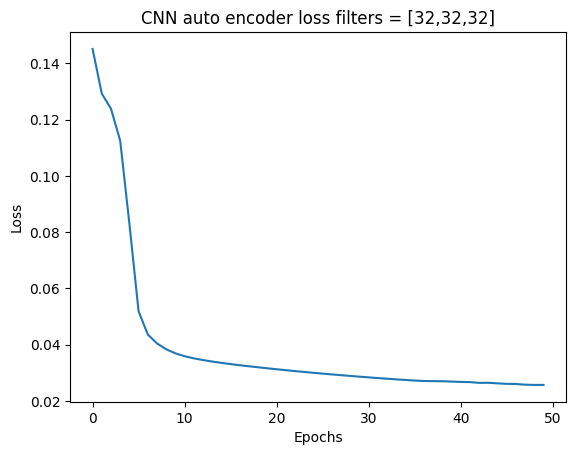
**شکل3: خطای رمزگذار-رمزگشای خودکار بر روی لایه Dense با بعد پنهان برابر 10**

****

**شکل4:خطای رمزگذار-رمزگشای خودکار بر روی لایه Dense با بعد پنهان برابر 10**

****

شکل5:خطای رمزگذار-رمزگشای خودکار بر روی لایه CNN با فیلتر 32و 32و 16

****

شکل6:خطای رمزگذار-رمزگشای خودکار بر روی لایه CNN با فیلتر 32و 32و 32

**که همانطور که مشاهده میشود روش CNN نتایج بهتری به ما می­دهد زیرا از یک لایه کانولوشنی استفاده نموده است که قدرت بیشری در شناسایی الگوها و ویژگی­های کلی و داخلی داده­های ورودی دارد.**

4-۱. خود رمزگذار متغیر((**Variational AutoEncoder**

**هدف: در این بخش، هدف استفاده از رمزگذار خودکار متغیر جهت کاهش بعد، با استفاده از لایه Dense و Convolutional و در نهایت آموزش شبکه KNN و بررسی نتایج می­باشد.**

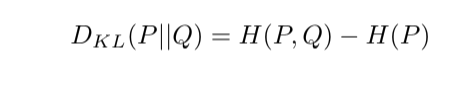
رمزگذار خودکار متغیر (VAE) یک الگوریتم هوش مصنوعی مولد است که از یادگیری عمیق برای تولید محتوای جدید، تشخیص ناهنجاری‌ها و حذف نویز استفاده می‌کند. در واقع VAE پارامترهای یک توزیع از پیش تعریف شده در فضای پنهان(latent space) را برای هر ورودی می­دهد، سپس برای این توزیع پنهان یک محدودیت تحمیل کرده و آن را مجبور می­کند که نرمال شود. این محدودیت باعث ایجاد نظم و انسجام در فضای پنهان می­گردد. در واقع VAE ها بر خلاف اتو انکودر ها بردار ویژگی را در فضای latent به صورت رندوم تغییر نمی دهند بلکه پارامتر های mu , sigma را با کمک الگوریتم EBP یاد گرفته و از این توزیع یک نمونه تصادفی انتخاب می کنند.این ویژگی در واقع variation variety نامیده می شود و باعث می شود تا ورودی های یکسان بردار های انکودینگ متنوعی داشته باشند و لذا امکان ایجاد نمونه های غیر واقعی فراوان و متنوعیی را فراهم می کنند.

پس بنابراین تابع خطا در این روش دو هدف دارد، بازسازی ورودی و توزیع نرمال فضای پنهان و در نتیجه تابع هزینه به عنوان مجموع هزینه بازسازی و شباهت سنجیده می-شود. خطای بازسازی حاصل میانگین مجذور تفاوت ورودی و خروجی و خطای شباهت واگرایی KL بین توزیع فضای پنهان و حالت استاندارد با میانگین صفر و واریانس یک است.

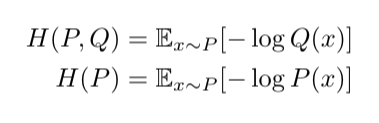
تفاوت اصلی مدل رمزگذار خودکار و رمزگذار خودکار متغیر در مدل خطای آن­ها است. در رمزگذار خودکار متغیر علاوه بر بازسازی، یک اصطلاح واگرایی KL نیز اضافه دارد و میانگین و واریانس را نیز استخراج می­نماید در حالی که رمزگذار خودکا تنها بخش بازسازی را بررسی می­نماید. افزودن عبارت واگرایی KL تضمین می کند که توزیع آموخته شده نزدیک به توزیع قبلی است، که برای تولید نمونه های متنوع و معنی دار ضروری است. هم­چنین رمزگذار خودکار داده­ها را به یک نقطه و رمزگذار خودکار متغیر به یک توزیع نرمال چندمتغیره نگاشت می­نماید. در واقع عملکرد اتو انکودر ها به این صورت می باشد که ورودی را به صورت unsupervised که می تواند عکس و یا هرچیز دیگری باشد به طور کامل گرفته، پس از کاهش بعد و حذف correlation ها ان را به سایز واقعی رسانده و خروجی ان هم سایز ورودی اصلی ما می باشد و در این روش یادگیری به کمک سنجش میزان فاصله می باشد. در حالی که VAE در واقع یک generative system می باشد و در واقع فرض می کند داده های ورودی ان نوعی توزیع احتمالی می باشند و تلاش می کند تا پارامترهای توزیع را پیدا نماید. و در نتیجه با این روش دیتا را مدل نموده و می توان ورودی های جدید با ان تولید کرد .

حال همان­طور که در بخش هدف به آن اشاره شده است، مطابق زیر یک بار رمزگذار خودکار متغیر را با استفاده از لایه Dense و یک بار با استفاده از لایه Convolutional داریم که در آن­ها ابعاد تصاویر به 20و 30 کاهش پیدا کرده­اند. در نهایت خروجی آن­ها به شبکه KNN داده شده است تا نتایج بررسی گردد.

در مدل زیر، نوع تابع loss مورد استفاده KL divergence می باشد که مزیت ان این می باشد که به ما این اطمینان را می دهد که داده های تولیدی ما توسط انکودر از داده های اصلی ما دور نمی شوند . فرمول ان به صورت زیر می باشد :

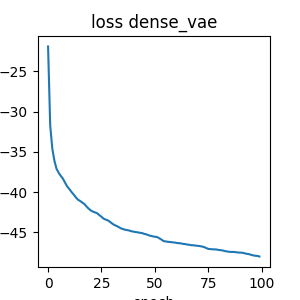
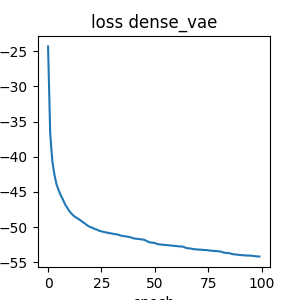
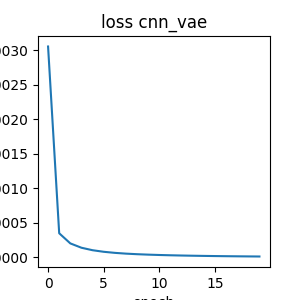
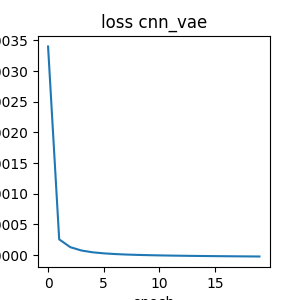


که در واقع داریم:



و P, Q هر کدام خروجی واقعی و تخمینی ما می باشند.

هم­چنین نمودار خطا شبکه را به فرم زیر داریم:



شکل 7 : تصویر خطا در VAE و با لایه های Dense و CNN

**همان­ن طور که در بالا نیز اشاره شد، این روش رمزگذاری خودکار نسبت به قسمت قبل تر نتایج بهتر و دقیق تری به ما می­تواند بدهد.**

5-۱. کاوش در فضای latent

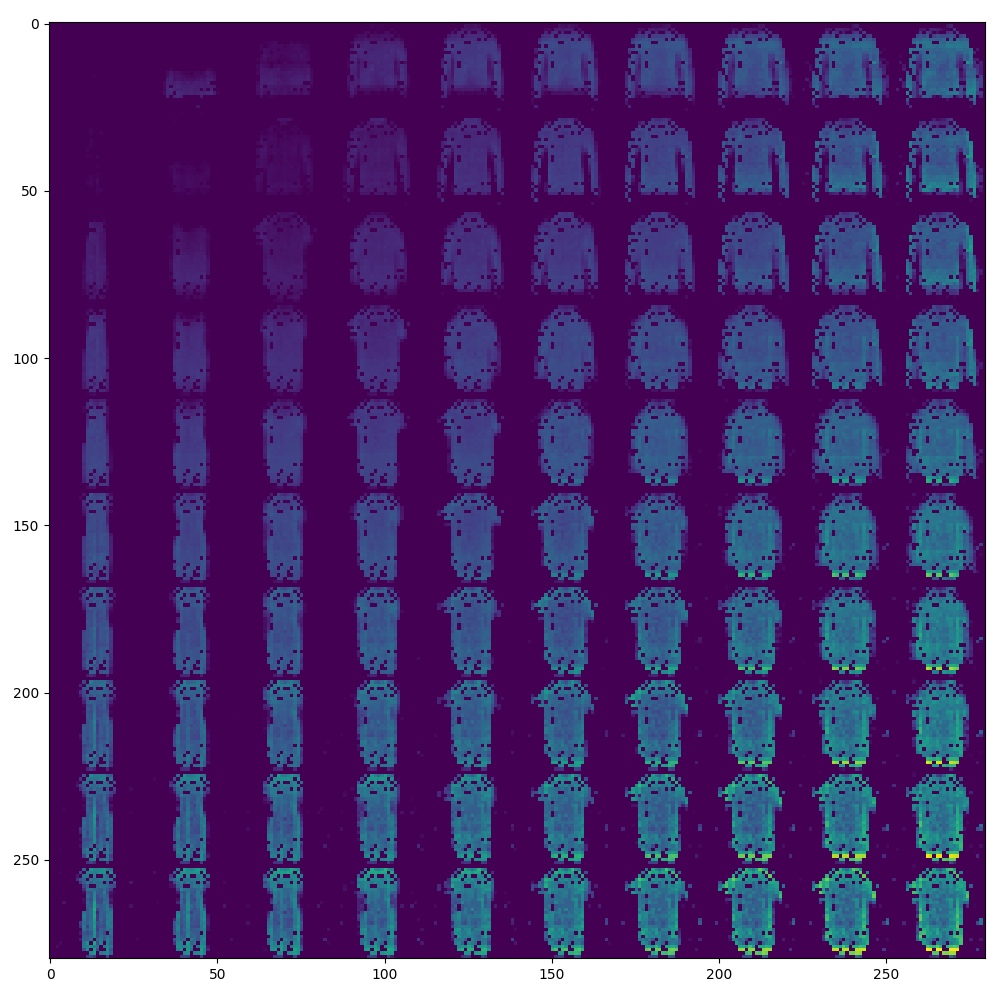
**هدف: در این بخش هدف ساخت یک grid از اعداد به کمک np.linspace و فرستادن آن به decoder در شبکه VAE و بررسی و تحلیل نتایج است. سپس داده­های آموزش به رمزگذار شبکه VAE داده شده و نمودار نقطه­ای داده­های آموزش در فضای پنهان ترسیم می­گردد.**

**فضای پنهان در واقع یک فضای چند بعدی انتزاعی تعریف می­شود که یک نمایش داخلی معنادار از رویدادهای مشاهده شده بیرونی را رمزگذاری می­کند. از ویژگی­های اصلی این فضا این است که نمونه­هایی که در دنیای بیرونی مشابه هستند در فضای نهان نزدیک به یکدیگر قرار می­گیرند.**

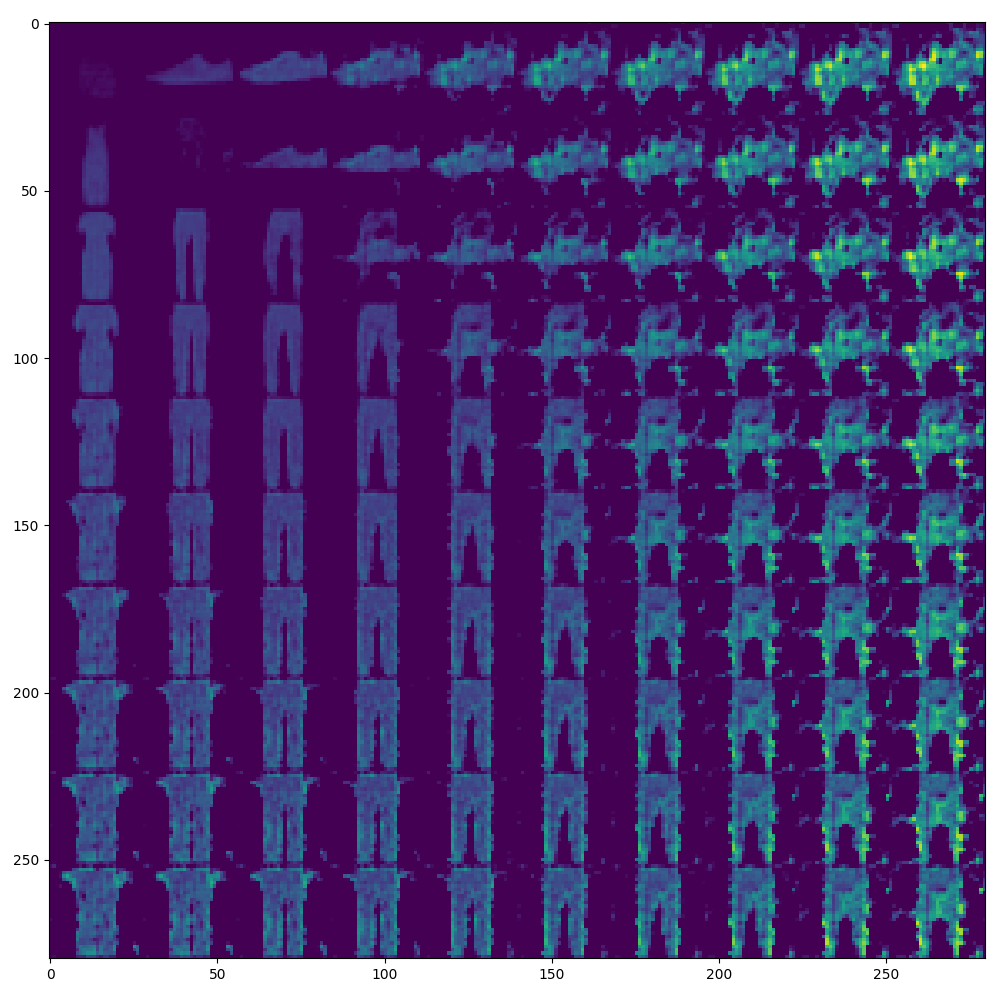
**برای درک بهتر این فضا می­توان به نگاه انسانی و درک مغز انسان­ها از رویدادها و مشاهدات بیرونی اشاره نمود. برای مثال یک انسان در هنگاه تشخیص یک سگ، تمام جزئیات آن را در نظر نمی­گیرد چرا که پیش­تر یک نمایش نهان داخلی از ظاهر کلی سگ را حفظ کرده و بای تشخیص از آن استفاده می­نماید و به همین ترتیب در شبکه­ها نی فضای پنهان تلاش می­نماید تا درک فشرده­ای از داده­های مد نظر را از طریق یک نمایش فضایی در اختیار مدل قرار دهد.**

**حال در رمزگذارهای خودکار، تصویر ورودی در ابتدا به یک فضای پنهانی تصویر شده و در نهایت به کمک رمزگذار یک بردار پنهان با ابعاد کمتری نسبت به تصویر اولیه، ایجاد می­گردد.**

**اکنون همان­طور که در صورت سوال اشاره شده است، ابتدا به کمک np. Linspace یک grid از اعداد ساخته و با اعمال تصاویر به دیکودر شبکه VAE اقدام به بررسی و تحلیل خروجی ها می­نماییم. لذا داریم:**

****

شکل 8: خروجی دیکودر dense

****

شکل 9: خروجی دیکودرCNN

6-۱. کاوش در فضای latent

**هدف : بررسی Diffusion models و مقایسه شیوه عملکرد آن­ها با VAE**

**مدل­های Diffusion یک سری مدل­های مولد عمیق هستند که یک زنجیره مارکوف از مراحل انتشار تعریف می­کنند تا به آرامی نویز تصادفی گوسی را به داده­های اموزش موجود در روند انتشار رو به جلو اضافه نموده و سپس اقدام به یادگیری روند معکوس که شامل حذف نویز و یا روند diffusion معکوس است می­نماید تا داده­های نمونه دلخواه را از نویزها بسازد. اضافه نمودن تدریجی نویزهای تصادفی تا زمانی ادامه خواهددداشت که نویز سفید به دست بیاید و در نتیجه این فرآیند غیر قابل یادگیری، به صورت معمول در طی 1000 مرحله طول می­کشد. در واقع، مدل­های Diffusion می­توانند داده­های مشابه آن­چه که بر روی آن­ها آموزش دیده­اند، تولید نمایند. برای مثال اگر مدل بر روی تصاویر گربه آموزش ببیند، می­تواند تصاویر واقعی مشابهی از گربه­ها ایجاد نماید.**

**به صورت کلی عملکرد مدل­های VAE بااستفاده از یک رمزگذار و رمزگشا و بر اساس کاهش ابعاد تصویر و بازسازی تصویر کاهش یافته می­باشد در حالی که عملکرد مدل­های Diffusion با استفاده از افزودن تدریجی نویز تا رسیدن به نویز سفید و بازسازی مجدد تصویر است.**

**مدل­های Diffusion به دلیل تدریجی بودن افزودن نویز، کندتر از مدل­های VAE هستند اما در مقابل همین قابلیت مرحله به مرحله بودن افزودن نویز قابلیت تولید تصویر به تصویر را صرفا با بهره­برداری از نویز می­دهد. هم­چنین به صورت کلی رمزگذار در VAE توزیع کد نهان را پیش­بینی می­کند و ممکن است تصاویری با کدهای نهان مشابه نیز وجود داشته باشند که این هم­پوشانی باعث میان­گیری از ورودی­های مشابه و ایجاد نمونه­های تار و کاهش دقت می­گردد اما در مقابل در Diffusion** تصاویر نویزی به عنوان کدهای پنهان عمل می­کنند و هم اندازه تصاویر ویرودی می­باشند و لذا در این حالت تمامی نمونه­ها کیفیت بالایی خواهند داشت.

# **پاسخ ۲** **- شبکه ی متخاصم مولد**

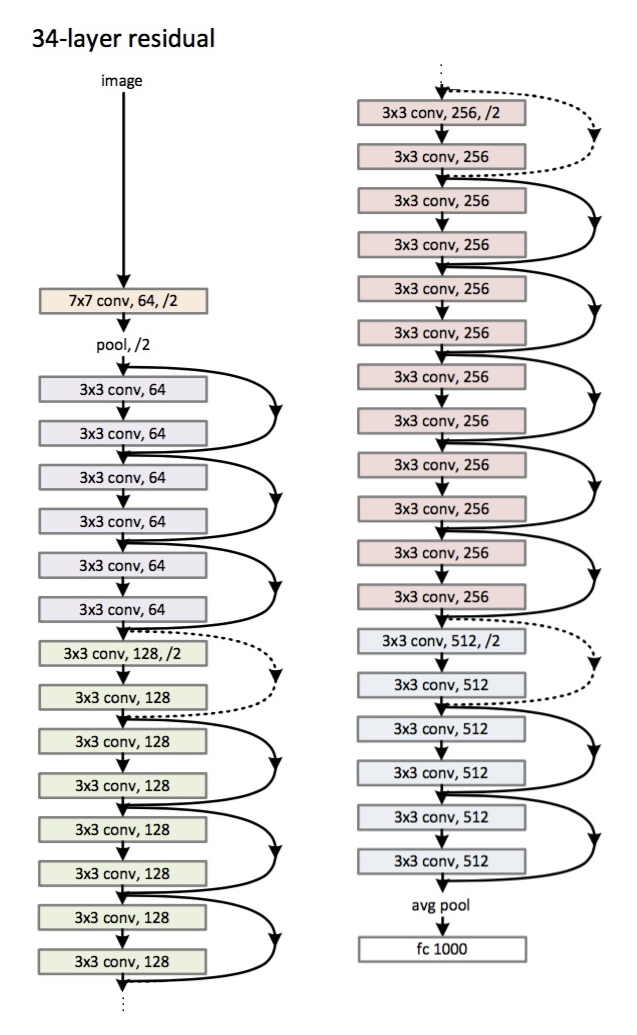
در بسیاری از مسائل طبقه­بندی ممکن است به دلیل محدودیت­های که برای داده­های آموزش وجود دارد(مانند محدود بودن تعداد نمونه­ها، توزیع نامتوازن کلاس­های داده و ...)، مدل نهایی به دقت مناسب نرسد. ازاین رو می­توان از شبکه­های متخاصم مولد برای تولید نمونه­های جدید استفاده کرد و محدودیت­های موجود در داده­های آموزش را تا حدی برطرف نمود. در این تمرین هدف استفاده از یک Conditional Deep Convolutional GAN جهت تولید داده­هایی است که برای آموزش و طبقه­بندی در یک شبکه­ی دیگر به کار گرفته می­شوند.

## ۱-۲. **بارگذاری داده­ها و شبکه­ی** ResNet

هدف: بارگذاری و پیش پردازش داده BreastMNIST و آموزش یک شبکه با معماری ResNet توسط آن.

داده­های انتخابی در این بخش، داده کلی MedMNIST هستند که شامل 12 دیتاست با تصاویر دو بعدی و سه بعدی است که کاربردهای مختلفی در زمینه­های متفاوتی مانند طبقه­بندی، رگرسیون و ... دارند. یکی از داده­های این دیتاست بزرگ، دیتاست BreastMNIST می­باد که در این سوال از ان استفاده می­گردد. این مجموعه داده حاصل تصاویر سونوگرافی سینه است، دارای 2 کلاس باینری، 780 نمونه با سایز 28 \* 28 که 546 داده آن، داده­های آموزشی، 78 داده اعتبار سنجی و 156 داده ارزیابی می­باشند.

حال هدف استفاده از شبکه ResNet می­باشد. در واقع همان­طور که در تصویر زیر مشاهده می­گردد، شبکه Resnet یک مدل یادگیری عمیق است که وزن­های لایه­های Residual با ارجاع داده شدن به لایه­های ورودی آموزش داده می­شوند. در واقع این شبکه شامل لایه­های بسیار زیادی است که یک سری پرش­های میان لایه­ای نیز دارد تا به کمک ان­ها خروجی را پیش­بینی نماید در واقه این مراحل برگشتی کمک می­کنند تا با تعداد لایه­های کمتر به نتایج بهتری رسید. هم­چنین نقطه چین ها به کاهش بعد کمک می­کنند. این مدل به مدل­های یادگیری با ده­ها و صدها لایه امکان می­دهد تا به راحتی آموزش ببینند و هرچه مسیر رو به جلو رفته و شبکه عمیق تر می­شود، دقت آن بهتر و بهتر گردد.



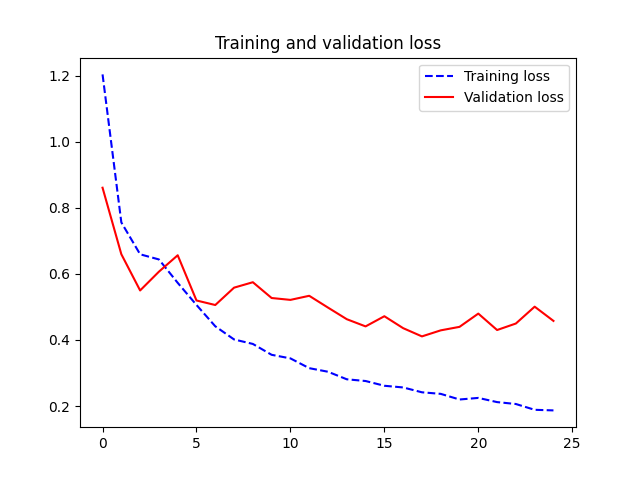
شکل 10: تصویر کلی شبکه ResNet50

اکنون مراحل پیش پردازش و اموزش شبکه را داریم:

در مرحله پیش پردازش داده­ها، از پیش پردازش خود شبکه Resnet50 استفاده شده است که ابتدا سایز داده­ها بر اساس نیاز شبکه تغییر یافته و سپس داده­ها mean substraction/scalling روشون اعمال میشه تا استاندارد یا نرمال شوند. از طرفی با توجه به اینکه ResNet در ImageNet سه بعدی است، داده­ها را ه بار کپی نموده­ایم و عمق سه به آن­ها داده شد است و سپس شبکه ResNet روی داده­ها برای Finetune شدن اعمال شده است.

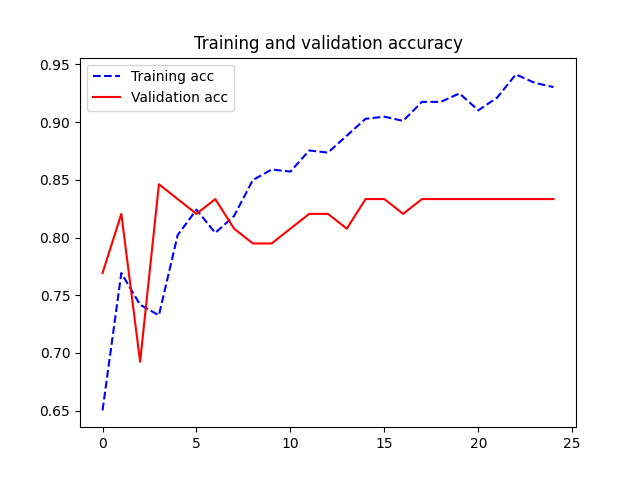
### الف: رسم نمودار دقت برای داده­های آموزشی و اعتبار سنجی و ارزیابی مدل

در اینجا، پس از آموزش داده­ها به کمک شبکه Resnet50، نمودار دقت برای داده­های آموزش و اعتبار سنجی را داریم:



شکل 11: نمودار خطای داده­های آموزش و اعتبارسنجی

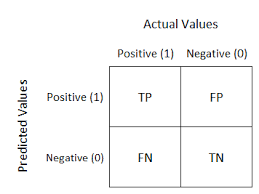
با توجه به نمودار بالا، تعداد ۲۵ ایپاک برای آموزش کافی است تا شبکه overfit نشود. اکنون بعد از ارزیابی شبکه برای داده های ارزیابی، نتایج را داریم:



شکل 12: تصویر نمودار دقت داده­های اموزش و اعتبارسنجی

### ب: ترسیم ماتریس آشفتگی

بر اساس تعریف، همان­طور که در شکل زیر مشاهده می­شود میدانیم که ماتریس آشفتگی جدولی است که برای تعریف عملکرد یک الگوریتم طبقه بندی استفاده می شود. یک ماتریس آشفتگی عملکرد یک الگوریتم طبقه بندی را تجسم و خلاصه می کند.



شکل 13: تعریف ماتریس اشفتگی

لذا داریم:

confusion\_matrix

[[11 31]

[24 90]]

precision recall f1-score support

0 0.31 0.26 0.29 42

1 0.74 0.79 0.77 114

accuracy 0.65 156

macro avg 0.53 0.53 0.53 156

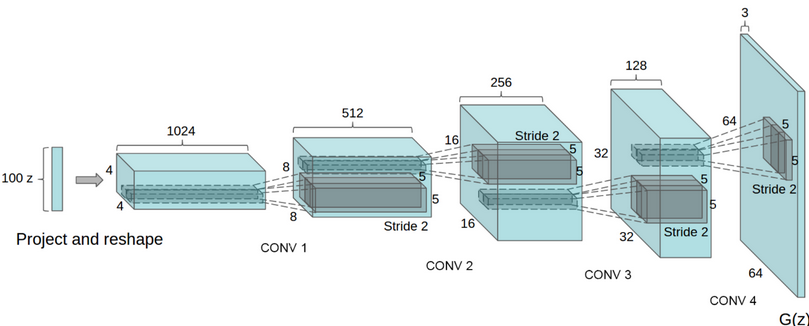
weighted avg 0.63 0.65 0.64 156

## 2-۲. **شبکه­ی** Conditional DCGAN

در واقع شبکه ی DCGAN شبکه ای بسیار موفق و دارای یک معماری ترکیبی می باشد به این صورت که در مولد و تفکیک کننده از لایه های کانولوشن بدون max pooling و بدون استفاده از لایه های fully connected استفاده شده است و در واقع به جای این لایه ها می توان از average pooling ها استفاده نمود و لذا بدین وسیله مدل پایدار تر خواهد بود اما متقابلا سرعت همگرایی ان کاهش می­یابد. در این شبکه از stride های کانولشنال و لایه­های transposed convolution استفاده می­شود و در واقع این لایه­ها به شبکه کمک می­کنند تا بتوانند عملیات­های spatial downsampling and upsampling را در تفکیک کننده و مولد انجام دهند .

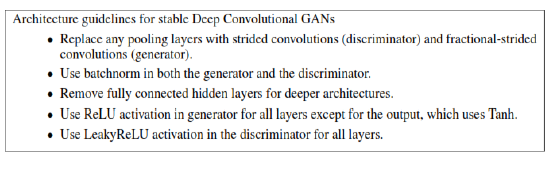
از طرفی ویژگی دیگری که این شبکه­ها دارند این می­باشد که تابع فعال­ساز در هردوی generator و discriminator تابع leaky relu می باشد که باعث افزایش رزولوشن مدل و دریافت پاسخ­های دقیق­تر می­شود.

هم چنین مبحث دیگری که در این شبکه­ها اهمیت دارد بحث نرمالیزیشن می­باشد که در واقع داده­های ورودی به کمک نرمالیزیشن دارای میانگین صفر و واریانس یک می­گردند و لذا تاثیر مقدار دهی­های اولیه به راحتی حذف شده و لذا فرایند یادگیری بهتر خواهد بود.



شکل 14: نمای کلی شبکه DCGAN

در حالت کلی مراحلی که باید جهت بهبود کیفیت پاسخ در DCGAN ها طی کرد به صورت زیر می باشد:



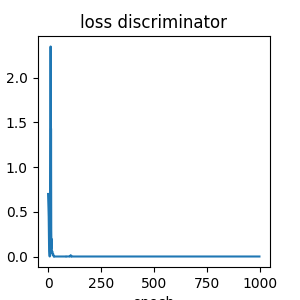
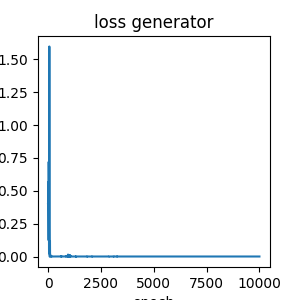
حال همان­طور که در صورت سوال گفته شده است، مراحل پیاده سازی دقیقا مانند Conditional GAN می­باشد با این تفاوت که لایه­های میانی دیگر fully connected نبوده و کانولوشن می­باشند . هم چنین لایه­ها به گونه­ای از لحاظ سایز، طراحی شده­اند که باعث می­شوند تا خروجی generator برای ما سایز 28\*28\*3 داشته باشد..

### الف: گزارش ساختارها و پارامترهای مورد استفاده برای شبکه­های مولد و تفکیک­کننده

در این بخش، ساختار کلی شبکه به فرمی است که در بالا توضیح داده شده است. در مرحله اول هر بار که یک ورودی به Generator داده می­شود، تعیین می­کنیمکه لیبل چه چیزی باشد، برای مثال یک بار میخواهیم اعداد صفر و یک بار اعداد 1 را تولید کند. پس با توجه به اینکه دیتاست مورد استفاده دو کلاسه می­باشد، پس یک نویز 98 تایی و یک one hot دوتایی به ورودی generator داده و در نهایت در descriminator دو تا خروجی مد نظر را پیش بینی می­نماییم. خروجی ها به صورت هم زمان پیش بینی کلاس ورودی و بررسی Real /Fake بودن آن­ها است.

### ب: رسم و تفسیر نمودار Loss برای مولد و تفکیک کننده

در این­جا خروجی و نمودار خطای مولد و تفکیک کننده را به صورت زیر داریم:

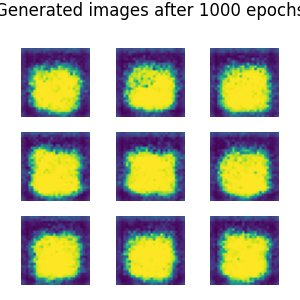


شکل 15: نمودار خطای مولد و تفکیک کننده

### پ: تولید 2000 نمونه توسط مولد به ازای هر کلاس داده و نمایش آن­ها

در این بخش، هدف تولید 2000 نمونه به ازای هر کلاس از داده مدنظر به وسیله مولد است. این نمونه ها به فرم زیر تولید میکنیم:

اکنون برای نمایش برخی نمونه ها داریم:



شکل 16: نمایش برخی نمونه ها بعد از 1000 ایپاک

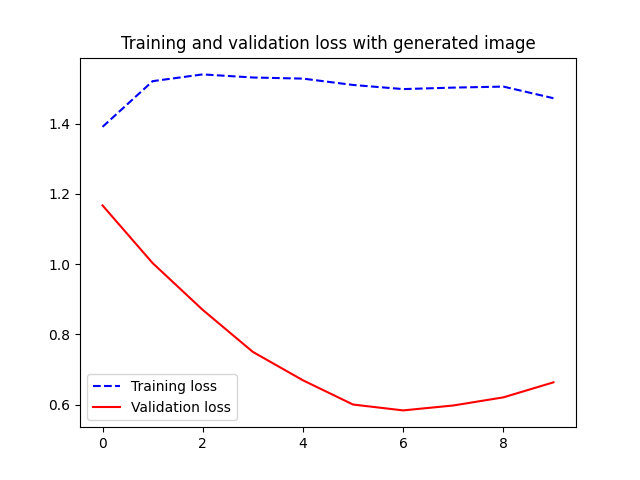
### ت: ارائه راهکار جهت بهتر شدن خروجی مولد و پایدارسازی

به صورت کلی روش­های مختلفی برای بهبود عملکرد مولد وجود دارد که برخی از آن­ها نیز در مدل بالا مورد استفاده قرار گرفته اند. برای مثال چنان چه تابع خطا Relu است می­توان از LeakyRelu استفاده نمود. روش دیگر این است که به جای ترکیب داده­های واقعی و تولید شده و آموزش دسته ای آن­ها، می­توان داده­های واقعی را جدا و داده­های تولید شده را جدا دسته بندی نمود. از طرفی می­توان به جای استفاده از مقادیر باینری صفر و یک، به متغیر تفکیک کننده نویز اضافه نمود که این روش نیز در این سوال مورد استفاده قرار گرفته است.

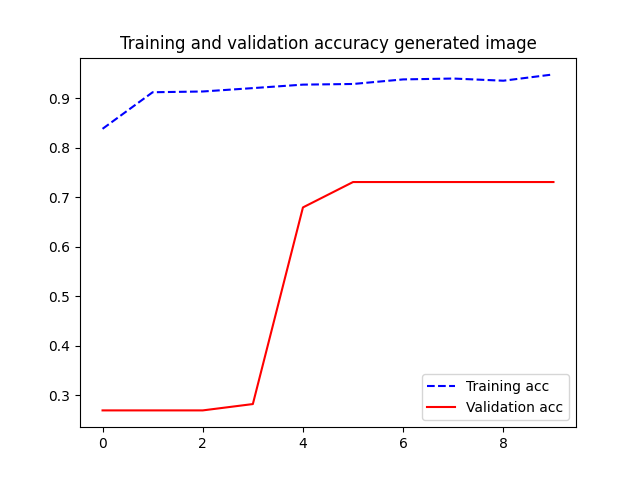
هم­چنین از دیگر روش­های بهبود عملکرد می­توان به استفاده از Drapout و یا استفاده از ConvTransposed2d برای نمونه­برداری استفاده نمود. هم­چنین می­توان Checkpoint های مدل را ذخیره و در هر بار تولید با مولد و تفکیک کننده ترکیب نمود. باید دقت نمود که با بررسی مداوم نتایج گرادیان و هنجارهای موجود و بررسی مداوم نمونه­های تولید­شده، می­توان به دقت و ثبات بهتری رسید.

## 3-۲. **طبقه بندی به کمک داده های تولید شده توسط مولد**

اکنون به وسیله­ داده­های تولید شده توسط مولد، مجدد طبقه بندی را انجام داده و نتایج را داریم:



شکل 17: نمودار خطای داده های اموزش و اعتبارسنجی در مولد



شکل 17: نمودار دقت داده های اموزش و اعتبارسنجی در مولد