



تمرین ششم درس شبکه های عصبی

شبکه FCGAN

Discriminator

- لایه ورودی: تصویری با ابعاد $3 \times 32 \times 32$ را دریافت می کند
- لایه پنهان:
سه لایه پنهان از نوع Dense که به ترتیب ۲۰۴۸، ۱۰۲۴، ۵۱۲، ۲۵۶ نورون دارند و تابع فعالیت آنها LeakyReLU می باشد.
- لایه خروجی
لایه Dense با یک نورون با تابع فعالیت sigmoid که با دو مقدار صفر و یک در خروجی، جعلی/واقعی بودن تصویر دریافتی را مشخص می کند.

در شکل زیر ساختار discriminator به نمایش گذاشته شده است.

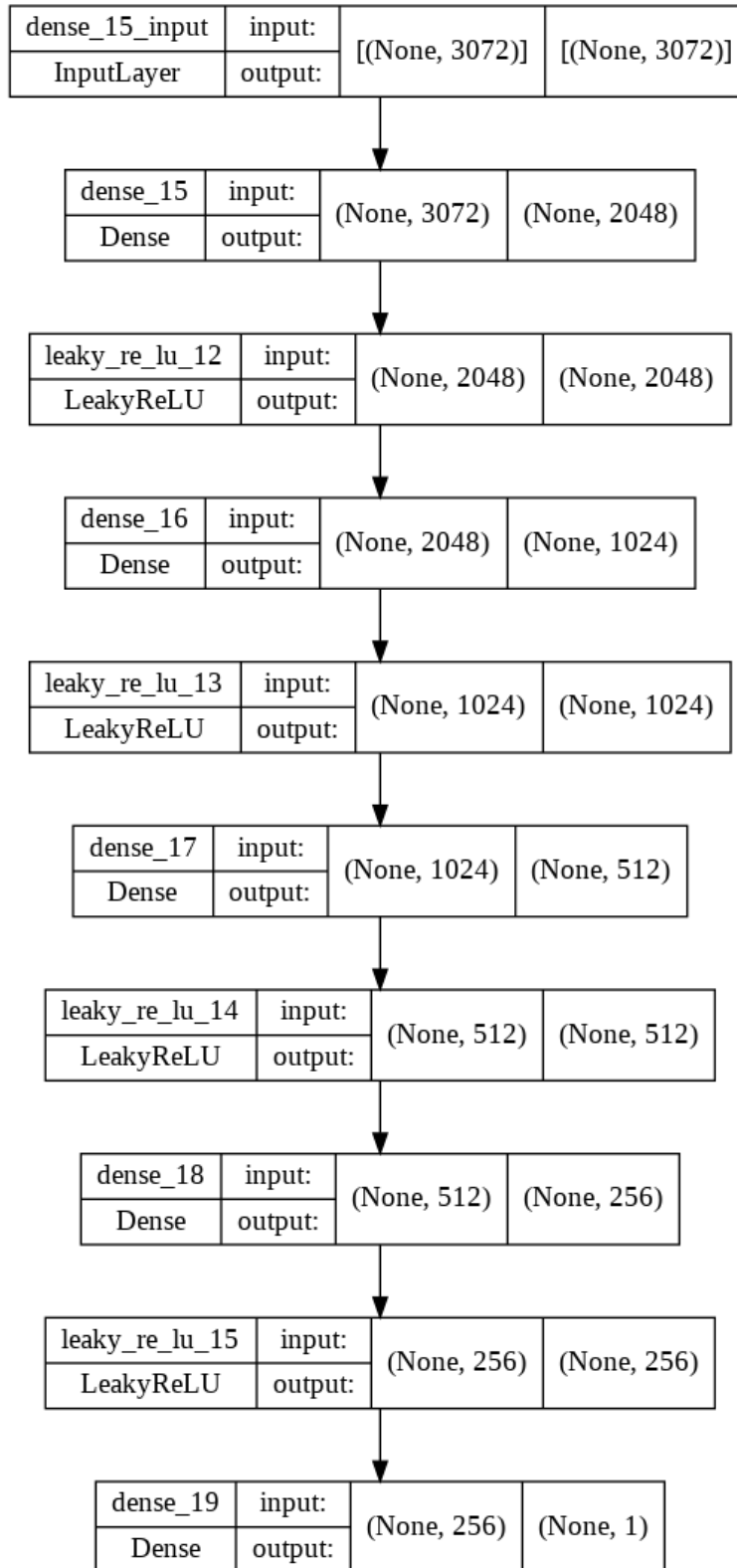
Generator

- لایه ورودی: ۱۰۰ نورون (برای دریافت یک بردار نویزی به طول ۱۰۰)
- لایه پنهان:
سه لایه پنهان از نوع Dense که به ترتیب ۲۵۶، ۵۱۲، ۱۰۲۴، ۲۰۴۸ نورون دارند و تابع فعالیت آنها LeakyReLU می باشد
- لایه خروجی
لایه Dense $3 \times 32 \times 32 = 3072$ نورون می باشد
تابع فعالیت لایه خروجی تابع tanh می باشد که مقادیر بردار خروجی را به بازه $[-1, 1]$ تبدیل میکند. (در فرایند نرمالسازی داده ها، مقادیر هر پیکسل به بازه $[-1, 1]$ تبدیل شده اند)

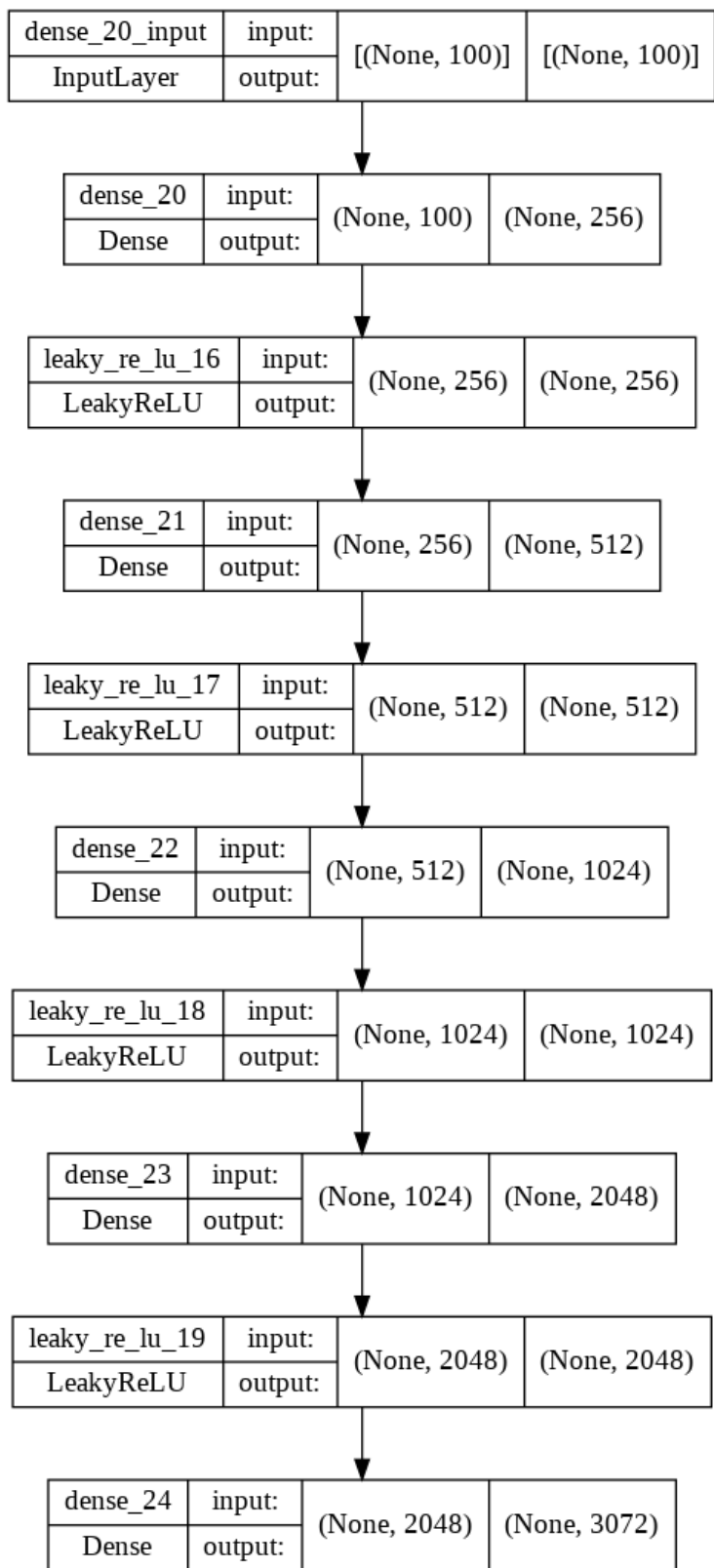
در شکل زیر ساختار generator به نمایش گذاشته شده است.

GAN

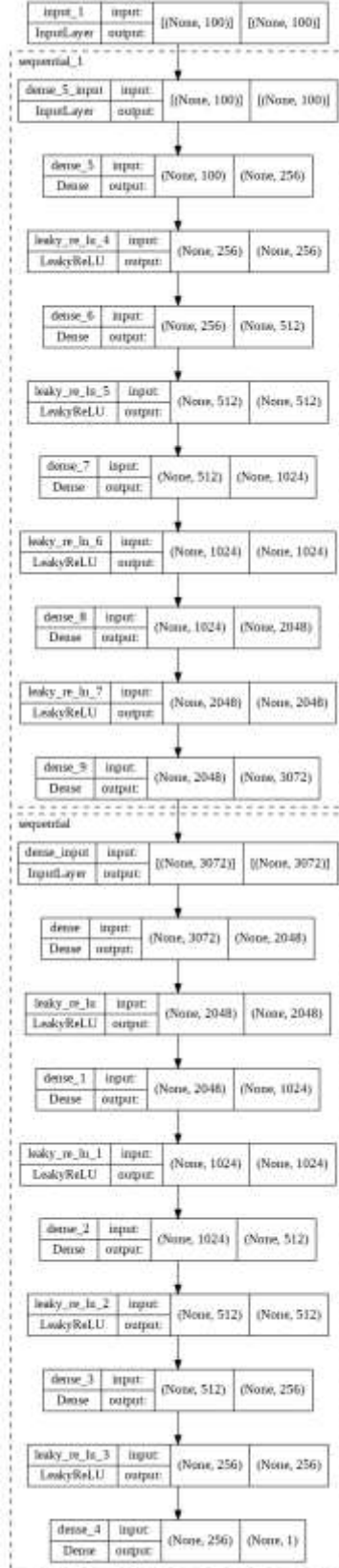
ترکیبی از اتصال دو شبکه Generator و Discriminator طراحی شده در مرحله قبل میباشد. شکل زیر ساختار این شبکه را نشان می دهد.



شکل ۱: ساختار discriminator مدل FCGAN



شکل ۱: ساختار Generator مدل DCGAN



شکل ۱: ساختار کلی مدل DCGAN

آموزش شبکه

برای آموزش شبکه، ابتدا تعدادی تصاویر جعلی توسط Generator تولید میشود سپس ترکیبی از این تصاویر جعلی با برچسب ۰ و تصاویر واقعی با برچسب ۱ به Discriminator داده میشود تا این شبکه برای تشخیص تصاویر جعلی از واقعی آموزش ببیند. در مرحله بعد آموزش Discriminator غیر فعال میشود و بخش Generator آموزش ببیند. به این صورت که یک بردار نویز (با توزیع تصادفی نرمال) تولید میشود و به عنوان ورودی به شبکه داده میشود.

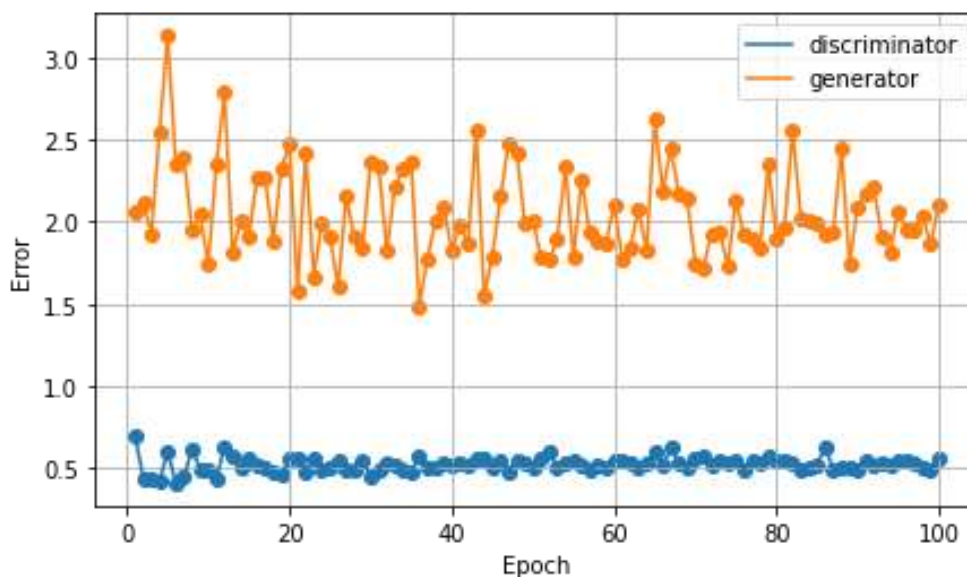
Generator یک تصویر در خروجی تولید می‌کند و Discriminator که قبلاً آموزش دیده جعلی بودن این تصویر را تشخیص می‌دهد و بدین ترتیب وزن لایه‌های Generator به گونه‌ای آپدیت می‌شود که میزان خطا را کاهش دهد و بتواند تصاویر نزدیک به واقعی تولید کند تا حدی که Discriminator نتواند جعلی بودن تصاویر را تشخیص دهد.

توابع هزینه

برای شبکه متمایز کننده از تابع هزینه binary cross entropy استفاده می‌شود زیرا این شبکه مشابه یک دسته بند باینری عمل میکند که تصاویر جعلی و واقعی را از هم متمایز میکند. برای شبکه مولد نیز از تابع هزینه binary cross entropy استفاده می‌شود زیرا این شبکه نیز فیدبکی دریافت میکند مبنی بر اینکه آیا توانسته یک تصویر جعلی نزدیک به واقعی تولید کند و شبکه متمایزکننده را فریب دهد یا خیر. این استدلال برای الگوی DCGAN نیز برقرار است

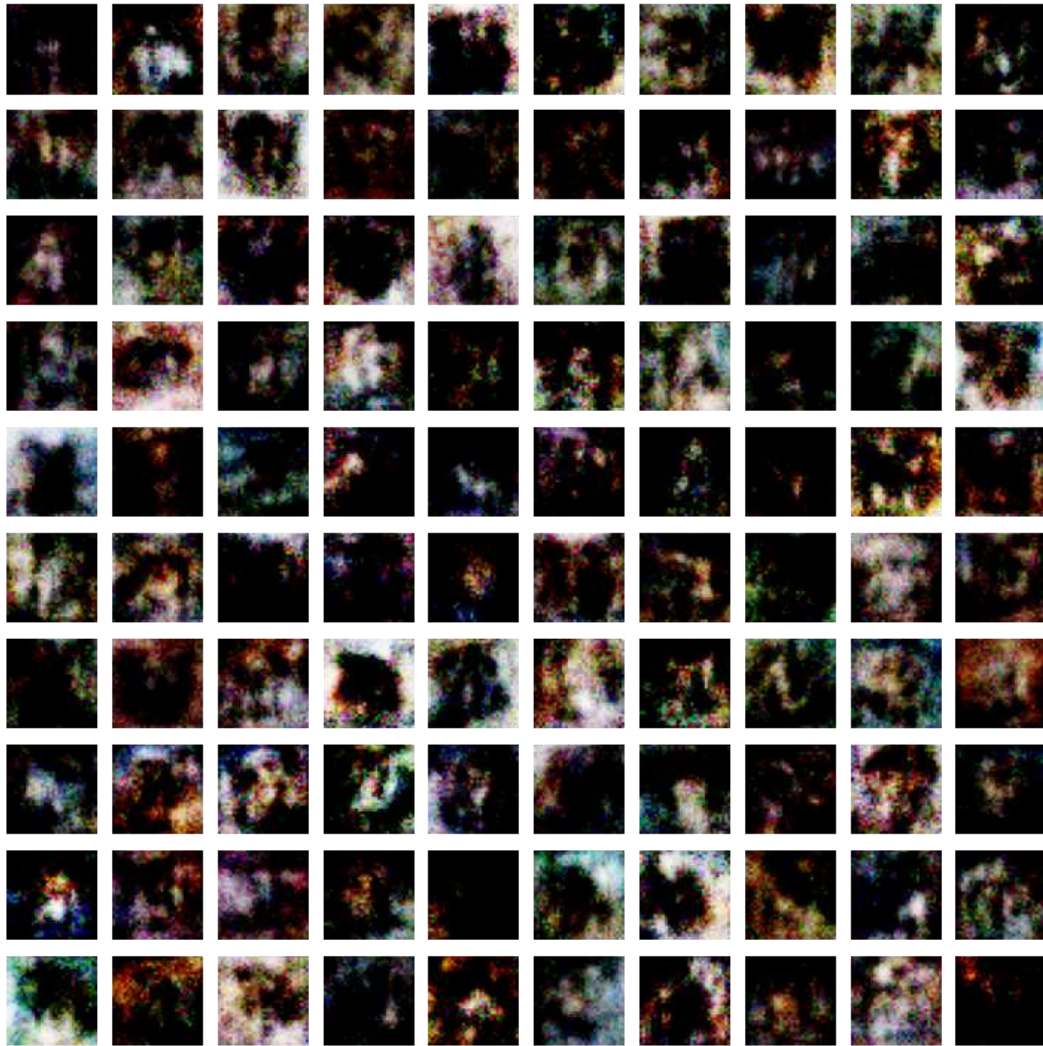
نتایج

نمودار زیر میزان تغییرات خطای Generator و Discriminator را در حین آموزش شبکه نشان میدهد:



شکل ۱: نمودار خطای DCGAN

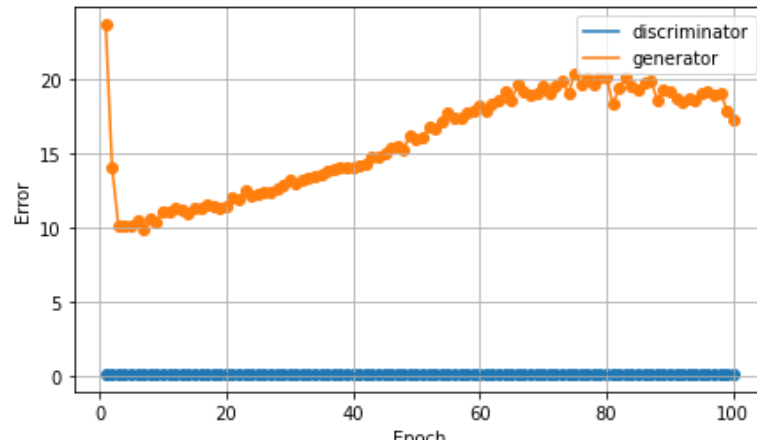
خروجی شبکه پس از ۱۰۰ epoch آموزش به صورت زیر است:



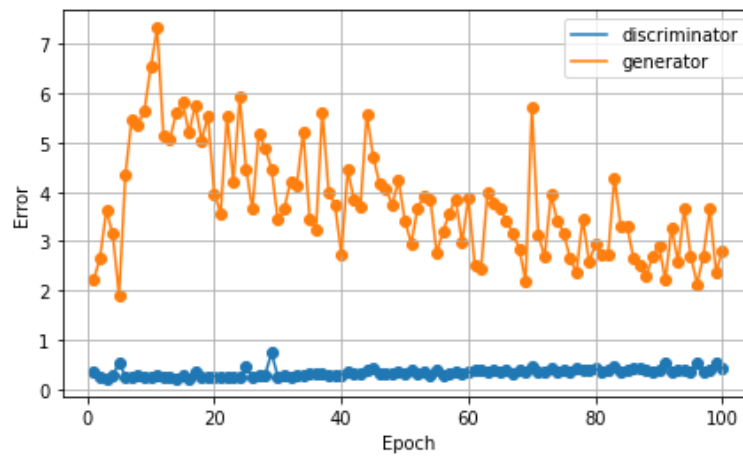
شکل ۱: ۱۰۰ تصویر تولید شده توسط fcgan پس از epoch ۲۰۰

تعداد لایه‌های مختلف

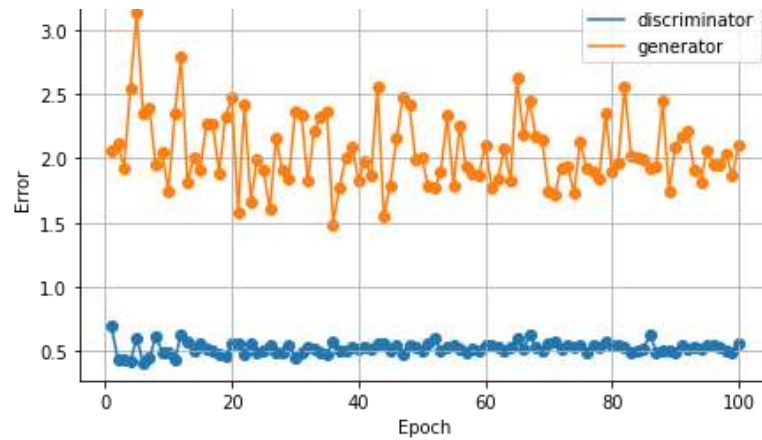
با تغییر تعداد لایه‌های پنهان به بررسی تاثیر تعدادهای های مدل مولد بر عملکرد مدل می‌پردازیم. با افزایش تعداد لایه‌ها، توانایی شبکه مولد در تولید تصاویر متنوع افزایش یافته و خطای مولد کاهش می‌یابد. در زیر نمودار ۳ حالت که لایه پنهان آنها به ترتیب ۵، ۴ و ۳ لایه مخفی دارد به نمایش گذاشته شده است.



شکل ۱: نمودار خطای DCGAN با ۳ لایه پنهان



شکل ۱: نمودار خطای DCGAN با ۴ لایه پنهان



شکل ۱: نمودار خطای DCGAN با ۵ لایه پنهان

متعادل سازی

یکی از راه‌های ایجاد تعادل، اضافه کردن یک نویز گاوسی به تصویر موجود در مجموعه داده است.

افزودن میزان کمی نویز به تصاویر دیتاست، باعث افزایش بالانس بین مولد و متمایزکننده میشود. همچنین کیفیت تصاویر تولید شده توسط مولد افزایش می‌یابد. اما وقتی تصاویر با احتمال ۰,۱ یا بیشتر از آن نویزی می‌شوند، و میزان نویز افزایش می‌یابد نه تنها باعث افزایش خطای مولد نسبت به متمایزکننده شده، بلکه کیفیت تصاویر تولید شده توسط مولد بسیار کاهش می‌یابد.

شبکه DCGAN

Discriminator

- لایه ورودی: تصاویری با ابعاد 32×32 را دریافت می‌کند
- لایه پنهان:
چهار لایه پنهان از نوع Conv2D که به ترتیب $256, 128, 128, 64$ کرنل 3×3 دارند و تابع فعالیت آنها LeakyReLU می‌باشد.
- لایه خروجی
لایه Dense با یک نورون با تابع فعالیت sigmoid که با دو مقدار صفر و یک در خروجی، جعلی/واقعی بودن تصویر دریافتی را مشخص می‌کند.

در شکل زیر ساختار discriminator به نمایش گذاشته شده است.

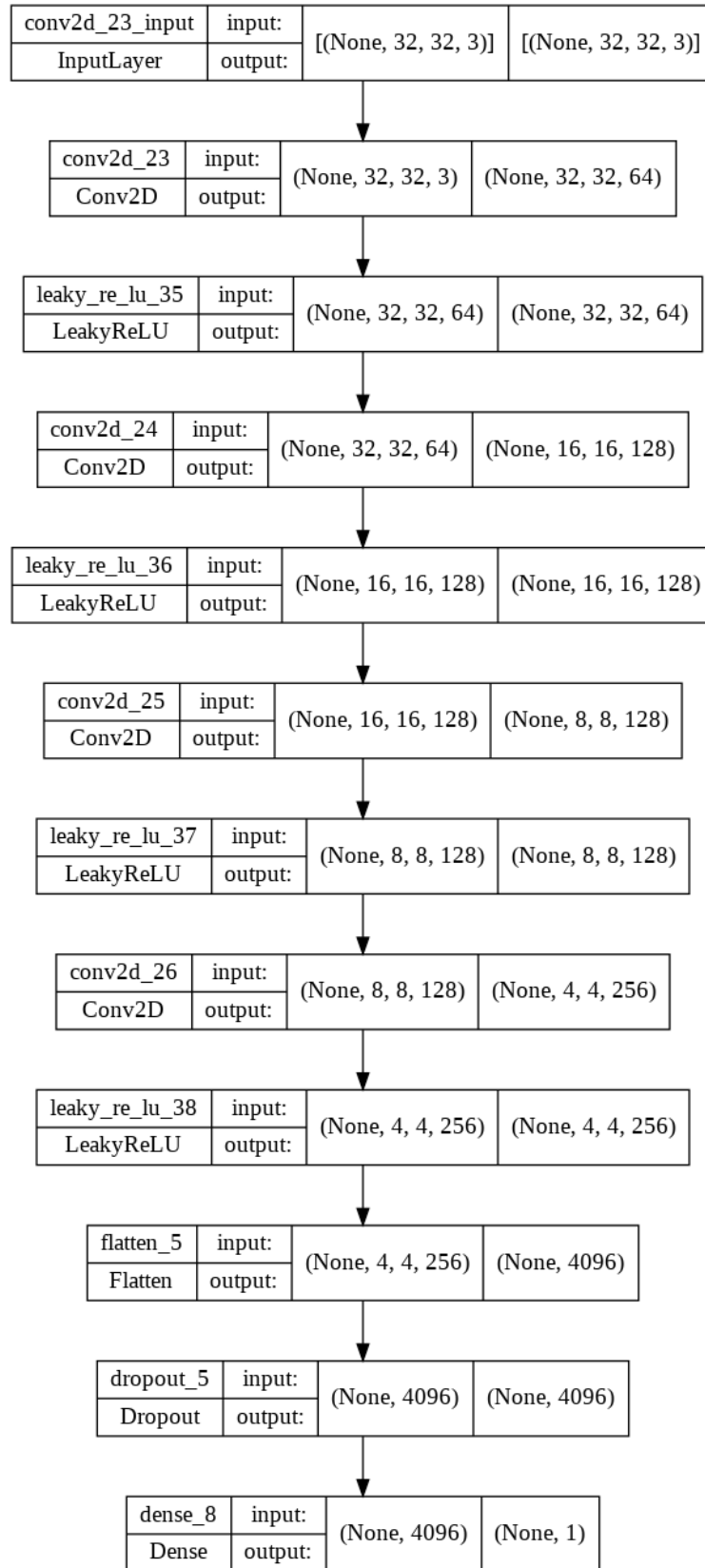
Generator

- لایه ورودی: ۱۰۰ نورون (برای دریافت یک بردار نویزی به طول ۱۰۰)
- لایه پنهان:
یک لایه Dense با $256 \times 4 \times 4$ نورون و تابع فعالیت LeakyRelu
سه لایه Conv2DTranspose که هر یک ۱۲۸ کرنل 4×4 دارند و تابع فعالیت هر دو LeakyReLU می‌باشد.
- لایه خروجی
لایه Conv2D که از کرنل‌های 3×3 استفاده میکند و تعداد کرنل‌ها برابر است با تعداد کانال‌های تصاویر دیتاست. این لایه تصاویری با ابعاد $32 \times 32 \times 3$ تولید می‌کند.

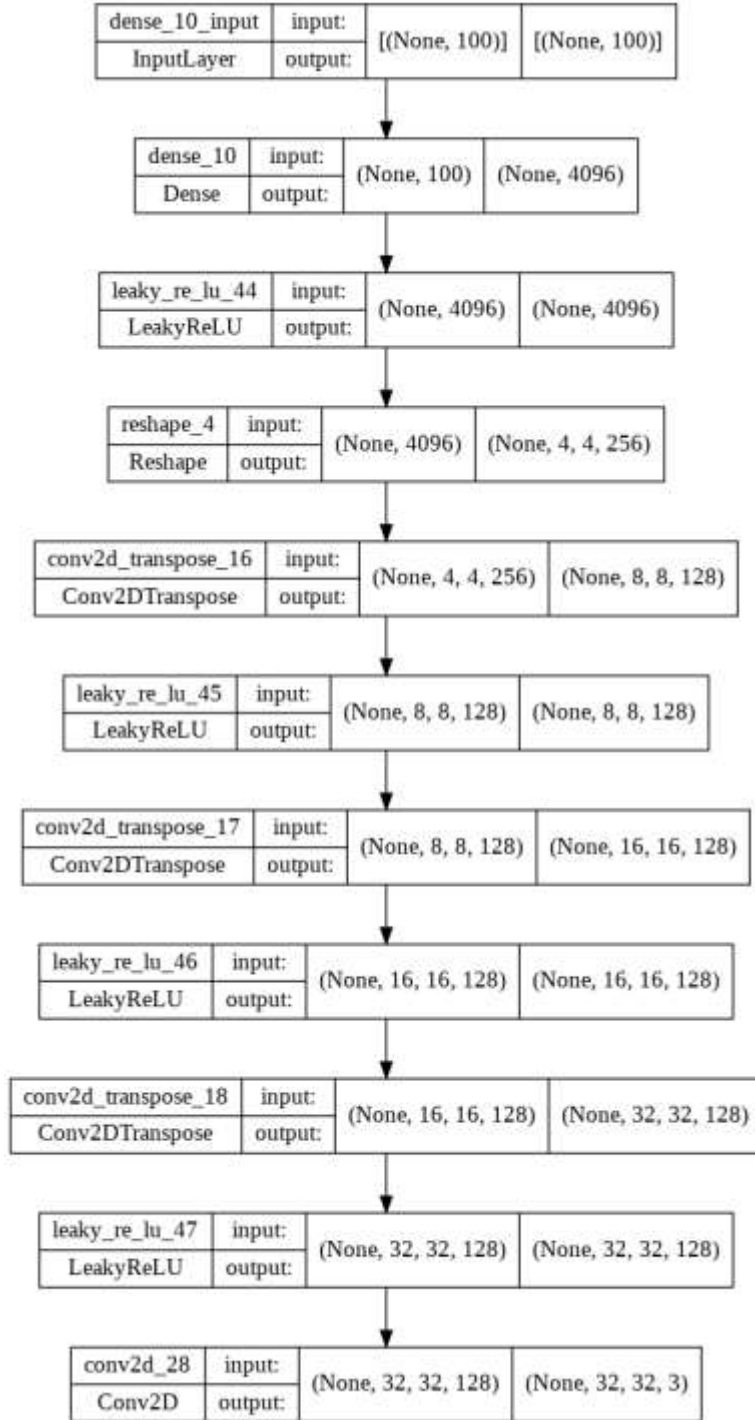
در شکل زیر ساختار generator به نمایش گذاشته شده است.

GAN

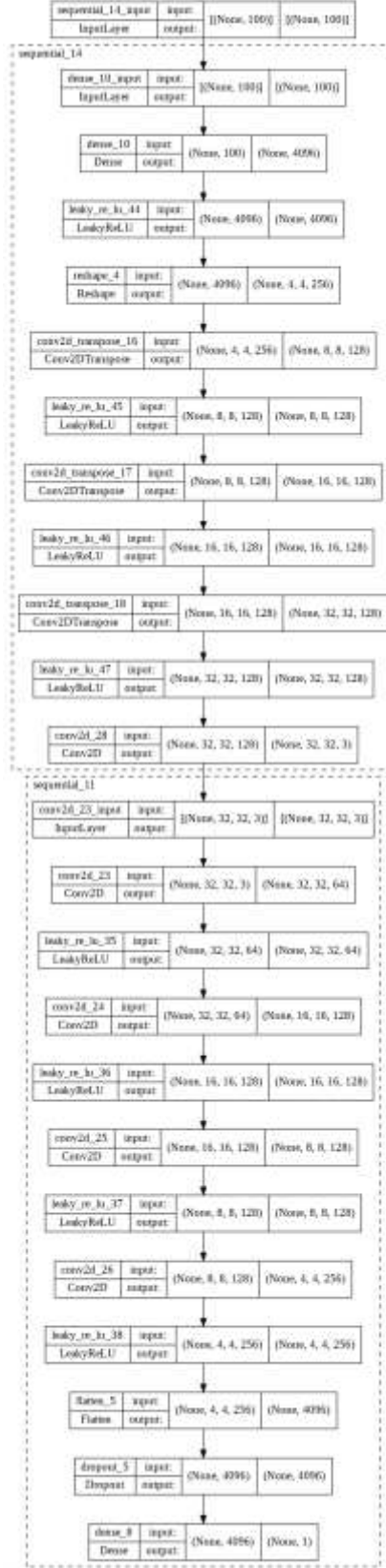
ترکیبی از اتصال دو شبکه Generator و Discriminator طراحی شده در مرحله قبل میباشد. شکل زیر ساختار این شبکه را نشان می‌دهد.



شکل ۱: ساختار discriminator مدل DCGAN



شکل ۱: ساختار generator مدل DCGAN

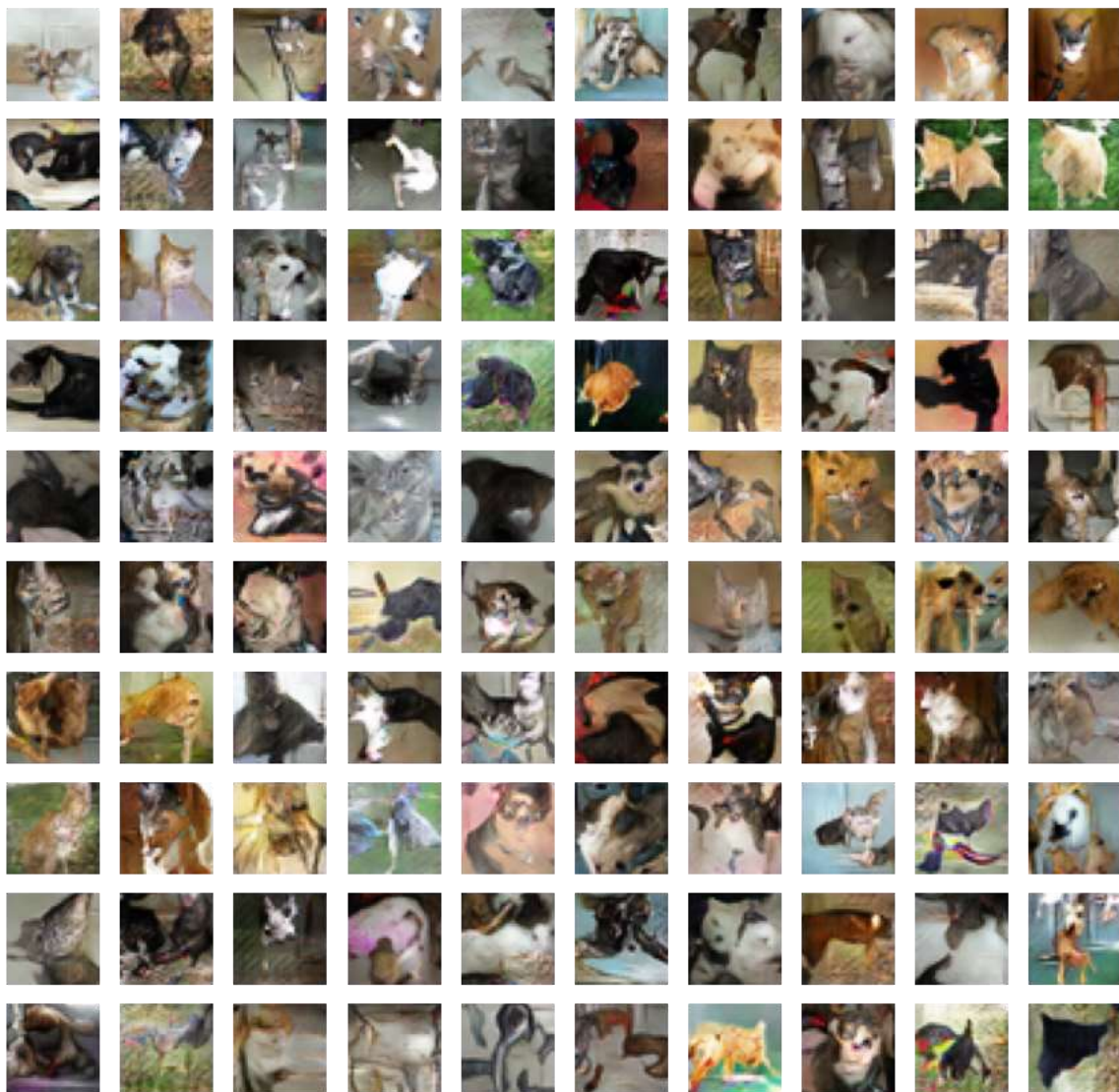


شکل ۱: ساختار کلی مدل DCGAN

آموزش شبکه DCGAN:

آموزش این شبکه نیز مشابه با شبکه FCGAN می باشد. برای آموزش شبکه، ابتدا تعدادی تصاویر جعلی توسط Generator تولید می شود سپس ترکیبی از این تصاویر جعلی با برچسب صفر و تصاویر واقعی با برچسب یک به Discriminator داده میشود تا این شبکه برای تشخیص تصاویر جعلی از واقعی آموزش ببیند. در مرحله بعد آموزش Discriminator غیر فعال میشود و بخش Generator آموزش میبیند. به این صورت که یک بردار نویز با توزیع تصادفی نرمال (تولید میشود و به عنوان ورودی به شبکه داده میشود. Generator یک تصویر در خروجی تولید میکند و Discriminator که قبلاً آموزش دیده جعلی بودن این تصویر را تشخیص میدهد و بدین ترتیب وزن لایه های Generator به گونه ای آپدیت میشود که میزان خطا را کاهش دهد و بتواند تصاویر نزدیک به واقعی تولید کند تا حدی که Discriminator نتواند جعلی بودن تصاویر را تشخیص دهد.

در شکل زیر آموزش شبکه پس از ۲۰۰ تکرار به نمایش گذاشته شده است. همان طور که مشاهده می کنید شبکه به خوبی توانسته تصاویر مناسب تولید کند



شکل ۱: ۱۰۰ تصویر تولید شده توسط dcgan پس از ۲۰۰ epoch

تعداد لایه‌های مختلف

برای این موضوع، با تغییر تعداد لایه‌های مدل مولد خروجی های $۱۶*۱۶$ ، $۳۲*۳۲$ و $۶۴*۶۴$ را مورد ارزیابی قرار دادیم و تصاویر خروجی را پس ۵۰ تکرار ثبت کردیم که در زیر مشاهده می‌کنید.



شکل ۱: ۱۰۰ تصویر تولید شده توسط dcgan پس از ۵۰ epoch با خروجی $۱۶*۱۶$



شکل ۱: ۱۰۰ تصویر تولید شده توسط dcgan پس از epoch ۵۰ با خروجی ۳۲*۳۲



شکل ۱: ۱۰۰ تصویر تولید شده توسط dcgan پس از epoch ۵۰ با خروجی ۶۴*۶۴

همان طور که در بالا قابل مشاهده است مدل با خروجی ۳۲*۳۲*۳ عملکرد بهتری نسبت به دو حالت دیگر دارد و در dcgan می توان گفت یک تعداد لایه بهینه وجود دارد که مدل های با بیشتر و کمتر از آن تعداد لایه عملکرد ضعیف تری دارند.