

به نام خدا



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

درس شبکه‌های عصبی
استاد صفابخش

تمرین هفتم

علیرضا مازوچی
۴۰۰۱۳۱۰۷۵

سوال ۱

مدل GAN از دو بخش مولد (Generator) و متمایزگر (Discriminator) تشکیل شده است. بخش مولد هدفش ایجاد داده مشابه با داده‌های واقعی و بخش متمایزگر هدفش شناسایی و تفکیک داده‌های واقعی از داده‌های تقلبی است. مولد برای ساخت تصاویر تقلبی از داده‌های نویز استفاده می‌کند و در شبکه خود این نویز را تبدیل به یک تصویر با ابعاد و شرایط اولیه مشابه با تصاویر واقعی می‌کند. متمایزگر یک تصویر را می‌گیرد و میزان واقعی بودن آن را می‌سنجد. یعنی متمایزگر عملاً یک دسته‌بند است. تابع خطا در مدل GAN بدین شکل تعریف می‌شود که هر چه تصاویر تقلبی تولیدشده توسط مولد واقعی‌تر باشد و متمایزگر را به خطا بیاندازد به مولد هزینه کمتری داده می‌شود و هرچه متمایزگر بهتر بتواند تصاویر را تفکیک کند یعنی بفهمد تصاویر واقعی، واقعی هستند و تصاویر تقلبی، تقلبی هزینه کمتری برای آن در نظر گرفته می‌شود. این تابع به شرح زیر است:

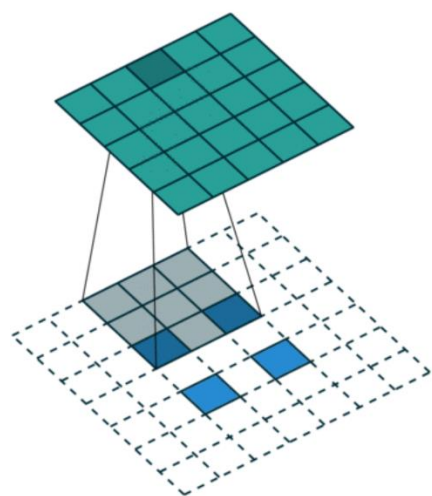
$$\min_{\theta_g} \max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z))) \right]$$

همانطور که در این تابع مشخص است، یک مدل سعی در بیشینه کردن تابع هزینه و یک مدل سعی در کمینه کردن آن دارند؛ به همین دلیل نوعی تقابل میان این دو مدل وجود دارد.

آموزش هر دو شبکه نیز همزمان انجام می‌گیرد. در ابتدا مولد تصاویر تصادفی ایجاد می‌کند؛ تصاویری که به راحتی از نوع واقعی قابل تفکیک است. به همین دلیل متمایزگر می‌تواند به آسانی آن را تشخیص دهد و مولد جریمه می‌شود. به واسطه این جریمه مولد سعی می‌کند تا تصاویر شبیه‌تر به واقعیت ایجاد کند. در همین حال که مولد قدرت‌مندتر می‌شود، متمایزگر نیز مجبور می‌شود تا ویژگی‌های مناسب‌تری از تصویر را استخراج کند. بین ترتیب متمایزگر هم قدرتمندتر می‌شود. هر چه مولد قوی‌تر شود، متمایزگر قوی‌تر می‌شود و هرچه متمایزگر قوی‌تر شود، مولد قوی‌تر. حاصل یک مولد با توان ایجاد داده‌های تقلبی با شباهت زیاد به واقعیت و یک متمایزگر با قدرت زیاد تشخیص تصاویر واقعی از تقلبی خواهد بود.

سوال ۲

لایه معکوس کانولوشن عملکردی مخالف لایه کانولوشن دارد؛ یعنی با دریافت یک نقشه ویژگی یک تصویر را می‌سازد. با نگاهی دیگر می‌توان دید که یک لایه کانولوشن اطلاعات پیکسل‌های مختلف را در یک پیکسل خلاصه می‌کند ولی یک لایه معکوس کانولوشن اطلاعات یک پیکسل را به پیکسل‌های مختلف گسترش می‌دهد.



برای ایجاد پیکسل جدید باید ورودی را حاشیه‌گذاری (Padding) کرد و سپس عمل کانولوشن را انجام داد. این بیت‌های حاشیه‌ای هم در میان پیکسل‌های ورودی و در روی مرزها قرار می‌گیرند. مانند تصویر روبرو:

سوال ۳

در سوال ۱ تابع هزینه کلی ارائه شده است. از روی این تابع می‌توان توابع هزینه را به صورت جداگانه برای هر یک از دو مدل ایجاد کرد. از آنجایی که هر مدل کنترل مستقیمی بر روی پارامترهای مدل دیگر ندارد، برای ایجاد توابع هزینه صرفاً باید توجه کرد که پارامتر مدل مذکور در کدام قسمت تابع هزینه کلی نقش داشته است. با این تفاسیر برای مدل متمایزگر تابع هزینه زیر بدست می‌آید:

$$\max_{\theta_d} \left[\mathbb{E}_{x \sim p_{data}} \log D_{\theta_d}(x) + \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z))) \right]$$

برای مدل مولد نیز تابع هزینه زیر:

$$\min_{\theta_g} \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z)))$$

تابع هزینه‌ای که برای مدل مولد پیشنهاد شده است می‌تواند بهینه‌تر شود و به فرم زیر در آید:

$$\min_{\theta_g} \mathbb{E}_{z \sim p(z)} \log(1 - D_{\theta_d}(G_{\theta_g}(z)))$$

این تابع از تابع قبلی از این جهت بهتر است که در اوایل آموزش مقدار گرادیان زیاد است و مدل مولد می‌تواند سریع‌تر آموزش ببیند.