

به نام خدا

تمرین سری ششم - درس مبانی بینایی کامپیوتر

سید محمد علی نخاری - شماره دانشجویی: 99521496

سوال چهارم)

الف) stride یا گام به میزان جابجا شدن فیلتر در هر مرحله از کانولوشن اشاره میکند. در واقع با استفاده از stride مشخص میکنیم که در هر مرحله فیلتر، چقدر به سمت راست یا پایین حرکت کند. اگر مقدار آن برابر 1 باشد، یعنی فیلتر باید بر روی همه پیکسل ها و همسایگی آن ها قرار بگیرد و همه نقاط تصویر را پوشش دهد. در این حالت ویژگی های بیشتری از تصویر استخراج شده و ابعاد تصویر کاهش نیابد. (اگر از padding استفاده کنیم) اگر مقدار گام را بزرگتر از 1 قرار دهیم، فیلتر با گام های بیشتری به سمت چپ و راست حرکت میکند و نقاطی از تصویر که فیلتر در آنها اعمال میشود کمتر شده و در نتیجه ویژگی های کمتری از تصویر استخراج میشود. در این حالت ابعاد تصویر ورودی هم کمتر میشود.

تفاوت مفهوم stride با مفهوم pooling در این است که در پولینگ ما در یک ناحیه های مشخص از تصویر ویژگی هایی نظیر بیشترین مقدار و یا میانگین مقادیر پیکسل ها استخراج کرده و در خروجی قرار میدهم و با این کار اندازه تصویر ورودی را در بعد مکانی کاهش میدهم اما در stride ما از یکسری از پیکسل ها صرف نظر میکنیم و مقادیر آنها را تاثیر نمیدهم. پس میتوان گفت در pooling با وجود کم شدن اندازه تصویر، ویژگی های بیشتری از تصویر استخراج میشود.

ب) 1- برای لایه های میانی معمولا از تابع فعالسازی ReLu یا Leaky ReLu استفاده میکنیم؛ چرا که این توابع به دلیل سادگی در محاسبات و اینکه مقادیر بزرگتر از صفر در آنها به مقدار خاصی میل نمیکند (اشباع نمیشوند) برای لایه های میانی مناسب هستند. (از ناپدید شدن گرادیان هم به نحوی جلوگیری میشود) برای لایه آخر که باید عمل دسته بندی را انجام دهیم استفاده از تابع

softmax که احتمال تعلق به هر کلاس را با مقادیر احتمالاتی و عددی بین 0 و 1 مشخص میکند برای این کار مناسب است.

2- توابع Categorical Crossentropy و یا Sparse Categorical Crossentropy به عنوان تابع ضرر استفاده میشوند. این توابع برای محاسبه خطای خروجی مدل در حال آموزش با توجه به برچسب های واقعی داده ها مناسب هستند. تابع اول برای برچسب هایی به فرمت one-hot بیشتر بکار میرود.

3- برای پاسخ به این سوال ابتدا توضیح کوتاهی در مورد دو معیار precision و recall میدهیم. معیار صحت (precision) بر روی درستی تشخیص های "بلی" توسط الگوریتم تمرکز دارد. به عبارت دیگر در این معیار مشخص میکنیم که چند درصد از "بله" های مدل آموزش دیده شده، صحیح بوده اند. اما معیار پوشش (recall) بر خلاف معیار صحت، تمرکزش بر روی داده هایی است که واقعا "بله" بودند یعنی نسبت داده هایی که شبکه آن ها را معیوب تشخیص داده به کل داده هایی که واقعا برچسب معیوب خورده اند.

در این مثال برای اینکه بخواهیم تعداد محصولات معیوب که به دست مشتری میرسد را به حداقل برسانیم، باید از معیار "precision" استفاده کنیم. دلیل آن هم این است که همانطور گفتیم معیار صحت، به دنبال داده هایی است که توسط شبکه برچسب "بله" خورده اند که در این مثال یعنی محصولاتی که توسط شبکه آموزش دیده شده معیوب اعلام میشوند. با این کار اگرچه ممکن است شبکه با درصدی دچار خطا شود و محصول سالم را معیوب اعلام کند اما تا حدی از رسیدن محصول معیوب به دست مشتری جلوگیری میکند.

به عنوان مثال فرض کنید تعدادی محصول معیوب و سالم داریم. در صورت استفاده از معیار صحت حتی اگر مقدار پایینی داشته باشد، ممکن است محصول سالم را به اشتباه معیوب اعلام کند و کالا به دست مشتری نرسد اما اگر از معیار پوشش استفاده کنیم در صورتیکه پوشش پایینی داشته باشیم ممکن است محصولی واقعا معیوب باشد اما به اشتباه آن را سالم اعلام کنیم و به دست مشتری برسد. پس بهتر است در همین مساله هایی از معیار صحت برای عملکرد مدل استفاده کنیم بهتر است.

ج) 3- شبکه های عصبی کانولوشنی معمولا در "تحلیل جدول مربوط به مشتریان یک فروشگاه برای پیش بینی رفتار بعدی هر مشتری" موفق نخواهند بود. دلیل آن هم این است که شبکه های عصبی کانولوشنی معمولا برای کاربرد هایی مثل پردازش تصویر و یا تشخیص الگو های مکانی طراحی شده اند. اما برای تحلیل داده ها و پیش بینی مدل های دیگری مثل شبکه های عصبی بازگشتی عملکرد بهتری خواهند داشت.

د) شبکه های عصبی کانولوشنی در کنار مزایای زیادی که دارند، معایبی نیز دارند:

1- نیاز به داده های آموزشی زیاد: برای اینکه این شبکه ها بتوانند عملکرد خوبی داشته باشند نیاز به داده های آموزشی زیاد دارند که این میتواند زمان آموزش را افزایش دهد و نیاز به منابع محاسباتی بیشتری داشته باشد.

2- شبکه های عصبی کانولوشنی معمولا در زمینه پردازش تصویر و مسائل دسته بندی کاربرد دارد.

3- با توجه به پیچیده بودن مدل های طراحی شده برای شبکه های کانولوشنی نیاز به منابع سخت افزاری زیادی دارند.

4- شبکه های عصبی کانولوشنی به تغییرات کوچک در تصاویر حساسیت دارند. به عنوان مثال اگر نور یک تصویر مشخص کم یا زیاد باشد شاید به خوبی نتواند آن را در دسته بندی مناسبی قرار دهد.