



به نام خدا



دانشگاه تهران
دانشکده مهندسی برق و کامپیوتر
فرایندهای اتفاقی

گزارش پروژه نهایی

نام و نام خانوادگی	منا داوری	فاطمه سبحانی
شماره دانشجویی	۸۱۰۱۹۴۴۵۷	۸۱۱۹۹۴۱۲۱
تاریخ ارسال گزارش	۱۳۹۷/۲/۲۶	

1-1 چکیده

در این پروژه یک شبکه عصبی مصنوعی را پیاده سازی کردیم. برای پیاده سازی شبکه عصبی از روش آموزشی پس انتشار در معماری سه لایه استفاده کردیم. سپس برای الگوریتم بهینه سازی از روش گرادیان کاهشی و گرادیان کاهشی تصادفی کمک گرفتیم. برای تابع فعال سازی نرون ها دو روش خطی و غیر خطی سیگمویدی را به کار بردیم.

1-2 ارائه ی روش

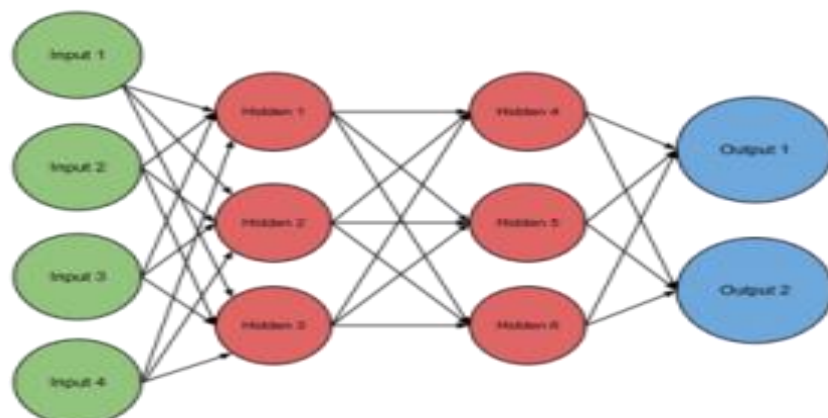
شبکه پس انتشار (Back Propagation Network)، نوعی شبکه عصبی چندلایه با تابع انتقال غیرخطی و قاعده یادگیری Widrow-Hoff می باشد. از بردار ورودی و هدف در راستای آموزش این نوع شبکه برای تقریب زدن یک تابع، یافتن رابطه بین ورودی و خروجی و دسته بندی ورودی ها استفاده می شود. این شبکه با دارا بودن بایاس، یک لایه سیگموید و یک لایه خروجی خطی، توانایی تخمین هر تابعی با تعداد نقاط ناپیوستگی محدود را داراست.

BP، یک الگوریتم استاندارد با کاهش شیب می باشد که در آن وزن های شبکه در جهت خلاف شیب تابع کارایی حرکت می کنند. لغت پس انتشار به رفتار شبکه BP در محاسبه شیب در شبکه های غیرخطی چندلایه اشاره دارد. الگوریتم های مختلفی وجود دارند که بر مبنای این الگوریتم استاندارد عمل می کنند. از جمله این الگوریتم ها می توان به الگوریتم گرادیان تویم و روش های نیوتن اشاره نمود.

۱- درمورد مراحل Feed Forward و Feed Backward در روش پس انتشار توضیح دهید .

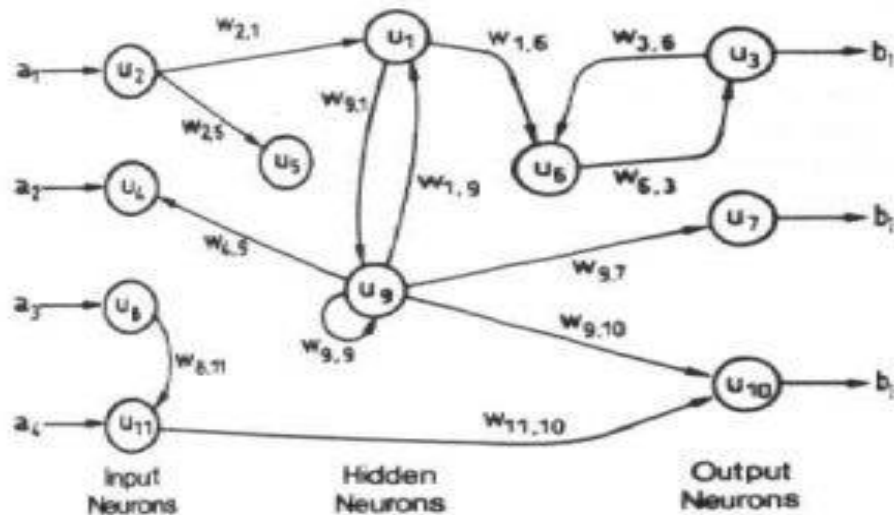
شبکه های پیش خور (FeedForward Neural Network)

شبکه های پیش خور، شبکه هایی هستند که مسیر پاسخ در آن ها همواره رو به جلو پردازش شده و به نرون های لایه های قبل خود باز نمی گردد. در این نوع شبکه به سیگنال ها تنها اجازه عبور از مسیر یکطرفه (از ورودی تا خروجی) داده می شود. بنابراین بازخورد یا فیدبک وجود ندارد به این معنی که خروجی هر لایه تنها بر لایه بعد اثر میگذارد و در لایه ی خودش تغییری ایجاد نمی کند.



شبکه‌های پس‌خور (FeedBack Neural Network)

تفاوت شبکه‌های پس‌خور با شبکه‌های پیش‌خور در آن است که در شبکه‌های برگشتی حداقل یک سیگنال برگشتی از یک نرون به همان نرون یا نرون‌های همان لایه یا نرون‌های لایه‌های قبل وجود دارد و اگر نرونی دارای فیدبک باشد بدین مفهوم است که خروجی نرون در لحظه حال نه تنها به ورودی در آن لحظه بلکه به مقدار خروجی خود نرون در لحظه‌ی گذشته نیز وابسته است.



۲- در مورد یادگیری‌های باناظر، بدون ناظر و تشدید یادگیری اختصاراً توضیح دهید.

یادگیری با ناظر (Supervised learning)

در یادگیری با ناظر به الگوریتم یادگیری، مجموعه‌ای از زوج داده‌ها داده می‌شود. هر داده یادگیری شامل ورودی به شبکه و خروجی هدف است. پس از اعمال ورودی به شبکه، خروجی شبکه با خروجی هدف مقایسه می‌گردد و سپس خطای یادگیری محاسبه شده و از آن جهت تنظیم پارامترهای شبکه (وزن‌ها)، استفاده می‌گردد. به گونه‌ای که اگر دفعه بعد به شبکه همان ورودی را دادیم، خروجی شبکه به خروجی هدف نزدیک گردد.

یادگیری تشدید

یادگیری تشدید حالت خاصی از یادگیری با ناظر و یک یادگیری برخط (On-Line) از یک نگاشت ورودی-خروجی است. این کار از طریق یک پروسه سعی و خطا به صورتی انجام می‌پذیرد که شاخصی موسوم به سیگنال تشدید، ماکزیمم شود که در آن بجای فراهم نمودن خروجی هدف، به شبکه عددی که نشان‌دهنده میزان عملکرد شبکه است ارائه می‌گردد.

یادگیری بدون ناظر (UnSupervised learning)

در یادگیری بدون ناظر یا یادگیری خود سامانده، پارامترهای شبکه عصبی تنها توسط پاسخ سیستم اصلاح و تنظیم می‌شوند. به عبارتی تنها اطلاعات دریافتی از محیط به شبکه را بردارهای ورودی تشکیل می‌دهند.

۳- در مورد مشکل **overfitting** توضیح دهید. برای رفع آن چه راه حل هایی وجود دارد.

overfitting ناشی از تنظیم وزن‌ها برای در نظر گرفتن مثالهای نادری است که ممکن است با توزیع کلی داده ها مطابقت نداشته باشند. تعداد زیاد وزنهای یک شبکه عصبی باعث میشود تا شبکه درجه آزادی زیادی برای انطباق با این مثالها داشته باشد.

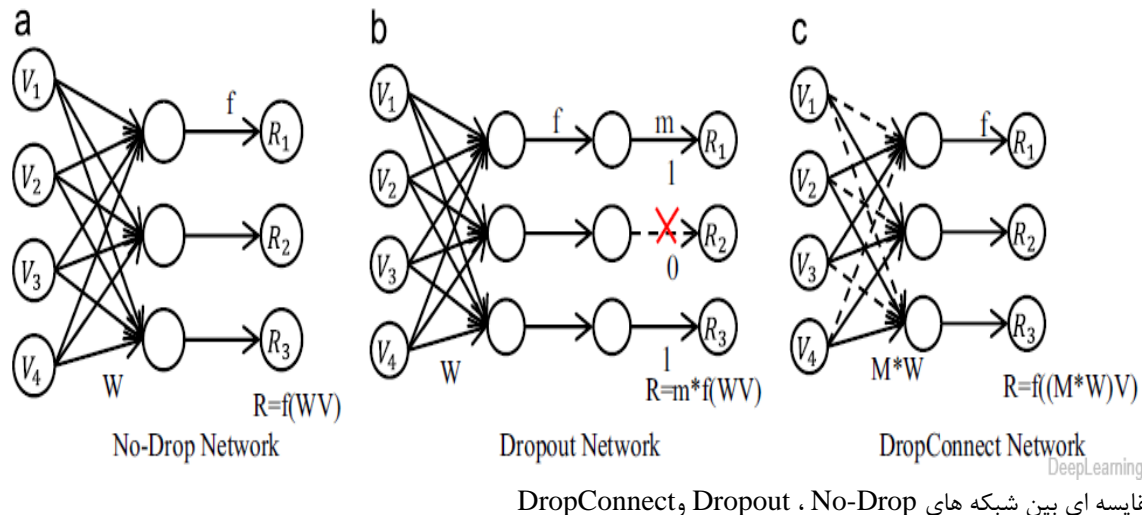
با افزایش تعداد تکرار، پیچیدگی فضای فرضیه یادگرفته شده توسط الگوریتم بیشتر و بیشتر میشود تا شبکه بتواند نویز و مثالهای نادر موجود در مجموعه آموزش را بدرستی ارزیابی نماید.

۴- در مورد هر یک از دو روش **drop out** و **L2Norm** توضیح دهید. در چه مواردی استفاده ی هر کدام از آن ها ترجیح دارد؟

روش **L2Norm** در این روش هر بار جمع توان ۲ وزن ها را ضرب در یک مقدار تجربی می کنیم و به مقدار **calculate_total_net_input** اضافه می کنیم. توجه شود که مقدار تجربی (ضریب که در جمع کل ضرب می شود) باید متناسب باشد در صورتی که زیاد باشد باعث می شود هر بار ما از مسیر اصلی برنامه پرت شویم و در صورت کم بودن تاثیر آن کم می شود.

روش **dropout** به منظور جلوگیری از **overfitting** معرفی شد. و نحوه کار آن به این صورت است که در هر مرحله از آموزش، هر نورون یا با احتمال $p-1$ (از شبکه) بیرون انداخته شده (**dropped out**) و یا با احتمال p نگه داشته میشود، بطوریکه نهایتاً یک شبکه کاهش داده شده باقی بماند. یالهای ورودی و خروجی به یک نود بیرون انداخته شده (**dropped-out node**) نیز حذف میشوند. اینگونه تنها شبکه کاهش یافته بر روی داده ها در آن مرحله آموزش خواهد دید. (بعد از این کار) نودهای حذف شده سپس به همراه وزنهای سابق آنها (قبل از حذف شدن) دوباره به درون شبکه وارد میشوند.

یک نمونه معروف مشتق شده از **Dropout** به **DropConnect** معروف است که بصورت تصادفی وزن‌ها را بجای مقادیر فعالسازی حذف میکند. آزمایشات نشان دادند که این روش میتواند پایاپای و حتی بهتر از روش **Dropout** در انواع گوناگونی از **Benchmark** های استاندارد عمل کند. هرچند با سرعت کمتر.



۵- در مورد شبکه های عصبی **recurrent** و **convolutional** تحقیق کنید و بگویید که چگونه میتوان از این شبکه ها برای همین مسئله استفاده کرد. همچنین مزایا و معایب آنها را در مقابل شبکه فعلی بررسی کنید.

شبکه های بازگشتی

شبکه پیش خوردار دور ندارد و به محض اینکه عصبی آموزش ببیند، وضعیتش ثابت می شود و با اعمال داده های ورودی به آن تغییر نمی کند یعنی اینکه حافظه ندارد. در شبکه بازگشتی می توانیم طوقه داشته باشیم همچنین یالهای از گره خروجی به گره های ورودی این شبکه در هنگام ورودی جدید به شبکه دگرگون شود (قابلیت حافظه)

شبکه شامل یک یا چند لایه از چندین نورون (گره) پنهان است که قسمتی از ورودی یا خروجی نیستند. این گره شبکه را قادر می سازند تا توابع غیر خطی پیچیده را یاد بگیرند.

شبکه بازگشتی به عنوان شبکه مجذوب کننده نامیده می شوند زیرا آنها به مقادیر مشخص جذب می شوند. بدین مفهوم که برای هر ورودی داده شده به شبکه آنی که نزدیکترین به ورودی است را به عنوان خروجی صادر می کند.

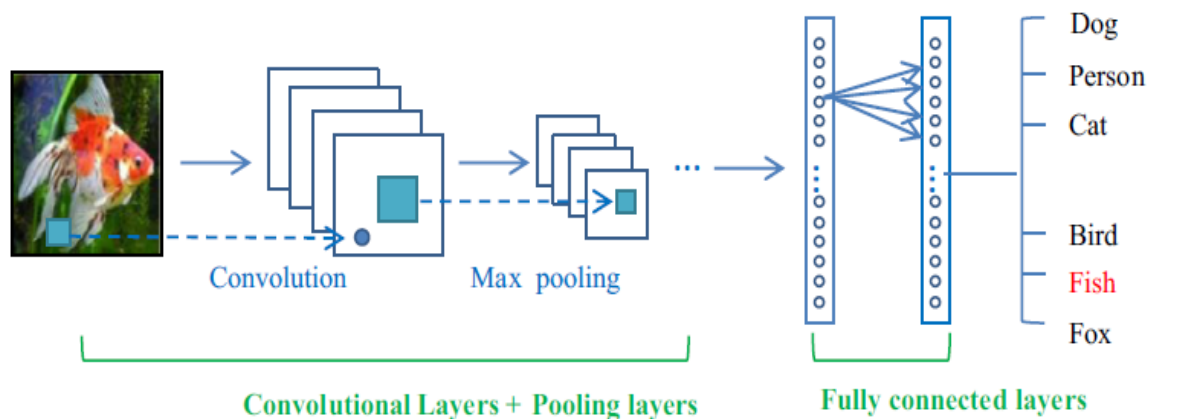
شبکه های کانولوشن

شبکه های عصبی کانولوشن (CNN) یکی از مهمترین روش های یادگیری عمیق هستند که در آنها چندین لایه با روشی قدرتمند آموزش می بینند این روش بسیار کارآمد بوده و یکی از رایجترین روشها در کاربردهای مختلف بینایی کامپیوتر است. تصویر کلی یک معماری شبکه عصبی کانولوشن در شکل زیر نمایش داده شده

است. بطور کلی، یک شبکه CNN از سه لایه اصلی تشکیل میشود که عبارتند از: لایه کانولوشن، لایه Pooling و لایه تماماً متصل. لایه های مختلف وظایف مختلفی را انجام میدهد. در شکل یک معماری کلی از شبکه عصبی کانولوشن برای دسته بندی تصاویر بصورت لایه به لایه نمایش داده شده است. در هر شبکه عصبی کانولوشن دو مرحله برای آموزش وجود دارد. مرحله feed forward و مرحله backpropagation یا پس انتشار. در مرحله اول تصویر ورودی به شبکه تغذیه میشود و این عمل چیزی جز ضرب نقطه ای بین ورودی و پارامترهای هر نورون و نهایتاً اعمال عملیات کانولوشن در هر لایه نیست. سپس خروجی شبکه محاسبه میشود. در این جا به منظور تنظیم پارامترهای شبکه و یا به عبارت دیگر همان آموزش شبکه، از نتیجه خروجی جهت محاسبه میزان خطای شبکه استفاده میشود. برای اینکار خروجی شبکه را با استفاده از یک تابع خطا (loss function) با پاسخ صحیح مقایسه کرده و اینطور میزان خطا محاسبه میشود. در مرحله بعدی بر اساس میزان خطای محاسبه شده مرحله backpropagation آغاز میشود. در این مرحله گرادیانت هر پارامتر با توجه به قاعده chain rule محاسبه میشود و تمامی پارامترها با توجه به تاثیری که بر خطای ایجاد شده در شبکه دارند تغییر پیدا میکنند. بعد از بروز آوری شدن پارامترها مرحله بعدی feed-forward شروع میشود. بعد از تکرار تعداد مناسبی از این مراحل آموزش شبکه پایان میابد.

انواع لایه های شبکه CNN

در حالت کلی، یک شبکه عصبی کانولوشن یک شبکه عصبی سلسله مراتبی است که لایه های کانولوشنی آن بصورت یک در میان با لایه های pooling بوده و بعد از آنها تعدادی لایه تماماً متصل وجود دارد.



یک طرح کلی از معماری یک شبکه عصبی کانولوشن

لایه کانولوشن:

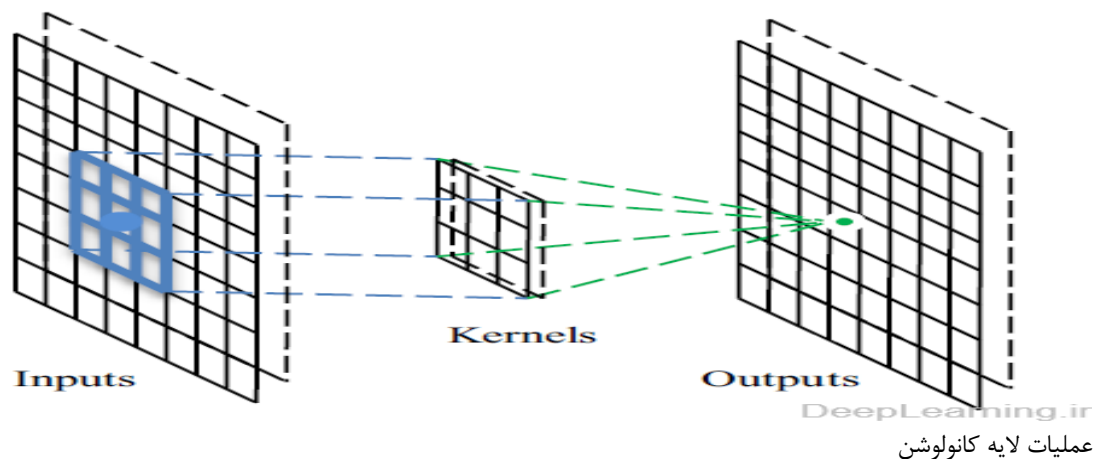
در این لایه ها، شبکه CNN از kernel های مختلف برای convolve کردن تصویر ورودی و همینطور feature map های میانی استفاده میکند و اینگونه feature map های مختلفی همانند آنچه در شکل مشاهده میکنید ایجاد میکند. انجام عملیات convolution سه فایده دارد:

۱- مکانیزم اشتراک وزن در هر feature map باعث کاهش شدید تعداد پارامترها میشود

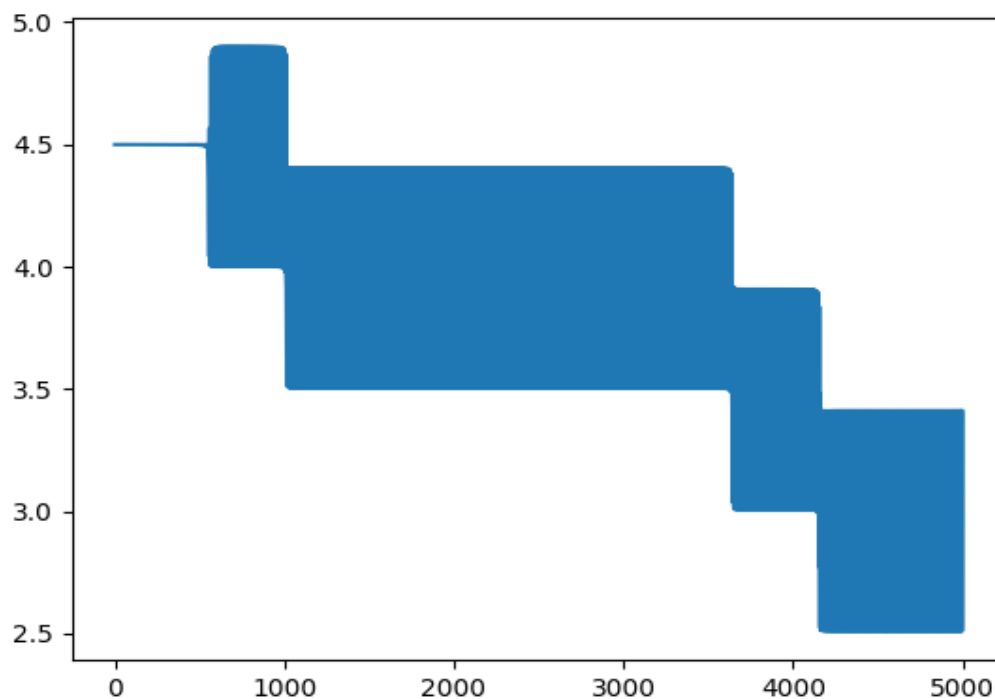
۲- اتصال محلی، ارتباط بین پیکسل های همسایه را یادمیگیرد

۳- باعث تغییر ناپذیری و ثبات نسبت به تغییر مکان شیء میشود

بواسطه فواید معرفی شده توسط عملیات convolution , بعضی از مقالات تحقیقاتی مشهور از آن جهت جایگزینی لایه های تماما متصل استفاده کردند تا با این کار سرعت فرآیند یادگیری را افزایش دهند. یکی از روشهای جالب مدیریت لایه های کانولوشنی، روش Network in Network (NIN) است که در آن ایده اصلی جایگزینی لایه کانولوشنی با یک شبکه عصبی پرسپترون کوچک است که شامل چندین لایه تماما متصل با توابع فعال سازی غیرخطی است. به این ترتیب فیلترهای خطی با شبکه های عصبی غیرخطی جایگزین میشوند. این روش باعث بدست آوردن نتایج خوبی در دسته بندی تصاویر میشود.



1-3 ارائه‌ی نتایج



1-4 تحلیل نتایج

بامشاهده‌ی نمودار دریافت میشود که تابع **lossing function** با هر بار پیشروی مقدار خطای کمتری را نشان میدهد و به سمت صفر تمایل پیدا میکند.

1-5 جمع‌بندی و نتیجه‌گیری