

طراح: حسام اسدالهزاده، ادیب رضایی، مسعود طهماسبی فرد، نوید دهبان

•

مهلت تحویل: جمعه ۱۹ خرداد ۱۴۰۲، ساعت ۲۳:۵۹

مقدمه	2
شبکههای عصبی کانولوشنی - (Convolutional Neural Networks (CNNs	2
نعریف مسئله	4
معرفى مجموعه داده	4
بخش صفر: آماده کردن داده	5
بخش اول: طبقه بندی تصاویر با استفاده از یک شبکهی CNN	6
بخش دوم: آموزش شبکه با دادههای نامتوازن	8
بخش سوم: استفاده از Data Augmentation برای ایجاد توازن بین تعداد داده ها	9
بخش چهارم: تاثیر روشهای Regularization در فرآیند آموزش	10
منابع	11

شبکههای عصبی کانولوشنی - Convolutional Neural Networks (CNNs)

شبکههای عصبی کانولوشنی یا پیچشی (CNN) نوعی مدل یادگیری عمیق هستند که به طور گسترده و به خصوص برای پردازش تصویر و پدیو استفاده می شوند. این مدلها برای تجزیه و تحلیل و شناسایی موثر الگرهای بصری با استفاده از مفهرم Convolution طراحی شدهاند. CNN ها معمولا از چندین لایه شامل لایههای کانولوشن، لایههای ادغام و لایههای کاملا متصل تشکیل می شوند. لایههای کانولوشن، فیلترهایی را روی تصویر ورودی اعمال می کنند و ویژگیها را در مکانهای مختلف تصویر یا ویدیو استخراج می کنند. لایههای ادغام ابعاد فضایی ویژگیها را کاهش می دهند و لایههای کاملاً متصل بر اساس ویژگیهای استخراج شده، عملیات طبقه بندی را انجام می دهند. این معماری سلسله مراتبی، CNN ها را قادر می سازد تا الگوهای بصری پیچیده را به طور خود کار یاد بگیرند. این قابلیت مدلهای کانولوشنی، آنها را در کارهایی مانند طبقه بندی تصویر، تشخیص اشیا و بخش بندی قصویر بسیار پرقدرت می سازد. CNN ها در حوزههای مختلف به موفقیتهای بزرگی دست یافتهاند و به پیشرفت در زمینهی بینایی ماشین کمک چشمگیری کردهاند. تاریخچه این شبکهها به سالهای ۱۹۸۰ و ۱۹۹۰ برمی گردد، زمانی که با وجود فراگیر شدن رایانهها و روشهای جدید پردازش تصویر مانند تبدیل فوریه و تصویر، به پردازش تصویر به صورت هانند تبدیل فوریه و Small Discrete Fourier Transform و ... همچنان محدودیتهای زیادی در این حوزه وجود داشت. اما با پیدایش شبکههای کانولوشنی، این مدلها توانستند با نوآوریهای خود در زمینه تشخیص الگو و تصویر، به پردازش تصویر به صورت خودکار و با کارایی بالا بپردازند.

یکی از اولین مدلهای شبکه کانولوشنی، مدل LeNet نام دارد که توسط یان لی کان (Yann LeCun) در سال ۱۹۹۸ ارائه شد. این مدل برای تشخیص و تمایز اعداد دستنوشته مورد استفاده قرار گرفت و در مسابقات مشهور تشخیص اعداد دستنوشته (MNIST) با استفاده از کامپیوتر به خوبی عمل کرد 4. مدل LeNet از لایههای کانولوشنی، لایههای ادغام و لایههای کاملا متصل تشکیل شدهاند. این مدل با استفاده از ترکیب این لایهها و استخراج ویژگیهای هر یک از اعداد، قادر به تشخیص و تمایز آنها می شود. LeNet بنیان گذار و الگوی اصلی برای توسعه شبکههای کانولوشنی در سالهای بعدی بود و نقش مهمی در پیشرفت علوم و تکنولوژی مرتبط با بینانی ماشین و تشخیص الگو ایفا کرد.

یکی از رویدادهای تاریخی قابل توجه در توسعه CNN ها، چالش تشخیص تصویری در مقیاس بزرگ ImageNet در سال 2012 بود. مدل برنده، معروف به AlexNet، یک CNN عمیق بود که به طور قابل توجهی از الگوریتمهای بینایی ماشین سنتی بهتر عمل کرد.

² Fully-Connected

¹ Pooling

³ Segmentation

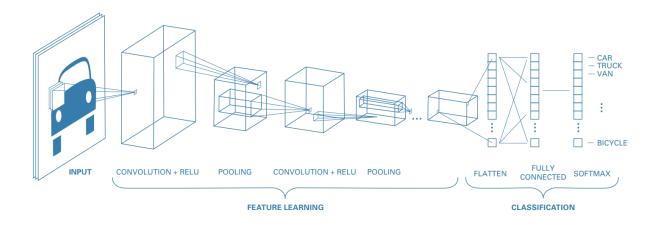
Convolutional Network Demo from 1989

ImageNet Large Scale Visual Recognition Challenge

موفقیت AlexNet، پتانسیل CNN ها را در وظایف طبقه بندی تصاویر نشان داد و باعث ایجاد علاقه بیشتر و در نتیجه گسترش پژوهشها در این زمینه شد.

از آن زمان، CNN ها به پیشرفتهای متعددی دست یافتهاند. در سال 2015، مدل VGGNet عملکرد چشمگیری را در وظایف طبقهبندی تصاویر نشان داد و معماری آن با لایههای کانولوشنی متعدد به طور گستردهای مورد استفاده قرار گرفت. نقطه عطف دیگر، معرفی معماری ResNet در سال 2015 بود که به چالش آموزش شبکههای بسیار عمیق پرداخت و منجر به پیشرفتهای قابل توجهی در دستیابی به مدلهایی با دقت بالا شد. علاوه بر این، CNN ها در پیشرفت تکنیکهای تشخیص اشیا، با مدلهایی مانند Faster R-CNN که به سرعت و دقت قابل توجهی دست می یابند، نقش مهمی داشتهاند. CNN ها با تمرکز بر حوزههایی مانند مکانیسمهای توجه و بادگیری انتقالی و تفسیر پذیری به تکامل خود ادامه می دهند تا مرزهای وظایف بینایی ماشین را حتی بیش از این،

پیش ببرند.



شکل ۱. ساختار کلی یک شبکه ی کانولوشنی

⁷ Transfer Learning

⁶ Attention

⁸ Interpretability

تعريف مسئله

در این تمرین، در بخش اول به پیاده سازی یک شبکه ی کانولوشنی برای طبقه بندی تصاویر، با استفاده از کتابخانه Keras می پردازید. در بخش دوم و سوم اثر نامتوازن و بودن تعداد داده های آموزشی در کلاس های مختلف را بررسی می کنید و از تکنیک های Regularization برای حل اثرات این موضوع استفاده می کنید. در بخش آخر نیز به بررسی روش های مختلف overfitting در آورنید یادگیری و جلوگیری از overfitting می پردازید. برای راحتی استفاده از کتابخانه ها و تسریع فرآیند آموزش، می توانید از سرویس آموزش می توانید از سرویس GPU بسیار سریع تر از GPU بسیار سریع تر از GPU بسیار سریع تر از COlab بسیار سریع تر از COlab بسیار سرویس GPU را روی GPU قرار دهید. دقت داشته باشید که مدت زمان استفاده از GPU پیش پردازش و تعریف مدل خود را روی CPU استفاده از GPU سرویس GPU به پایان نرسد.

معرفي مجموعه داده

مجموعه داده CIFAR-10 مجموعهای از تصاویر است که معمولاً برای آموزش الگوریتمهای یادگیری ماشین و بینایی ماشین استفاده می شود. این دیتاست یکی از پرکاربردترین مجموعههای داده برای تحقیقات یادگیری ماشین است. مجموعه داده CIFAR-10 شامل می شود. این دیتاست یکی از پرکاربردترین مجموعههای داده برای تحقیقات یادگیری ماشین است. مجموعه داده تورباغه، می تورباغه، می تصویر رنگی ۳۲×۳۲ در ۱۰ کلاس مختلف است. ۱۰ کلاس مختلف هواپیما، ماشین، پرنده، گربه، آهو، سگ، قورباغه، اسب، کشتی و کامیون را نشان می دهد. از هر کلاس ۶۰۰۰ تصویر وجود دارد.

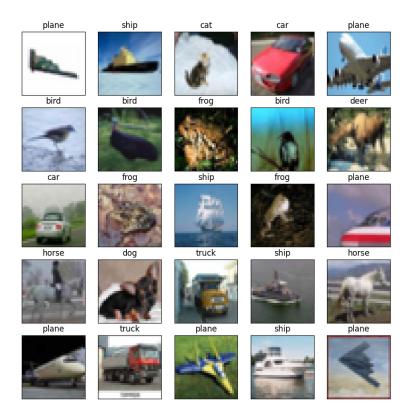
airplane	<u></u>
automobile	ar 🖏 🗟 🕵 🚾 😻 🐷 🖹 🚍 💖
bird	
cat	🕸 👺 😂 🐼 🎉 🙋 🗸 🥒 🤛
deer	
dog	B. (4) 10 10 10 10 10 10 10 10 10 10 10 10 10
frog	
horse	
ship	in the second se
truck	

⁹ Unbalanced Data

بخش صفر: آماده کردن داده

ديتاست CIFAR-10 را از tensorflow.keras.datasets لود كنيد:

- یک تصویر از هر کلاس در دیتاست را نمایش دهید.
- تعداد دادههای هر کلاس در مجموعه داده train و test را نمایش دهید.
 - عملیات Normalization را برای کل تصاویر انجام دهید.
- برچسبهای train و test را با استفاده از to_categorical به حالت OneHot تبدیل کنید.



شكل ۲. تعدادى از تصاوير مجموعه داده CIFAR-10

بخش اول: طبقه بندی تصاویر با استفاده از یک شبکهی CNN

در بخش اول این پروژه، میخواهیم از یک شبکه کانولوشنی برای طبقه بندی تصاویر استفاده کنیم.

همانطور که توضیح داده شد، شبکههای کانولوشنی، با استفاده از لایههای کانولوشنی، لایههای ادغام و لایههای کاملا متصل، قادرند اطلاعات مربوط به الگوها و ویژگیهای مختلف در تصویر را استخراج و استفاده کنند. لایه کانولوشنی با استفاده از عملیات کانولوشن، فیلترها را بر روی تصویر اعمال کرده و نمایش ¹⁰ جدیدی از تصویر را ایجاد می کند که شامل ویژگیهای محلی است. این لایهها به صورت مکرر در سراسر شبکه استفاده می شوند تا ویژگیهای سطح بالاتر را استخراج کنند.

لایههای ادغام به منظور کاهش ابعاد تصویر و حذف اطلاعات بیاهمیت، استفاده میشوند. این لایهها با استفاده از معیارهایی مانند ¹¹ عا میانگین گیری² ، به کاهش ابعاد ویژگیهای استخراج شده میپردازند.

لایههای کاملا متصل، در نهایت با استفاده از ویژگیهای استخراج شده توسط لایههای کانولوشنی و ادغام، تصمیم گیری نهایی را برای طبقهبندی تصویر انجام میدهند. این لایهها مشابه لایههای معمولی در شبکههای عصبی عمل میکنند و خروجی نهایی را تولید میکنند که شامل احتمالهای مربوط به تعلق تصویر به هر کلاس است.

در ادامه یک نمونه کد برای تعریف یک شبکه کانولوشنی با استفاده از کتابخانهی keras آمده است. توجه کنید که الزامی برای استفاده از این معماری خاص برای این بخش وجود ندارد و میتوانید مدل زیر را تغییر داده و شبکه ی کانولوشنی خود را تعریف کنید.

```
from keras.models import Sequential
from keras.layers import Dense, Conv2D, MaxPooling2D, Dropout, Activation, Input, Flatten
from keras.optimizers import Adam
cnn = Sequential()
cnn.add(Conv2D(32, (3, 3), padding='same', input_shape=x_train.shape[1:]))
cnn.add(Activation('relu'))
cnn.add(Conv2D(32, (3, 3)))
cnn.add(Activation('relu'))
cnn.add(MaxPooling2D(pool_size=(2, 2)))
cnn.add(Conv2D(64, (3, 3), padding='same'))
cnn.add(Activation('relu'))
cnn.add(Conv2D(64, (3, 3)))
cnn.add(Activation('relu'))
cnn.add(MaxPooling2D(pool size=(2, 2)))
cnn.add(Flatten())
cnn.add(Dense(512))
```

1

¹⁰ Representation

¹¹ Max Pooling

¹² Average Pooling

در فاز اول پروژه با بهینهساز SGD آشنا شدید. این روش بهینهسازی با استفاده از روش گرادیان کاهشی به بهینهسازی وزنها و پارامترهای شبکه عصبی میپردازد تا Loss کلی کاهش یافته و دقت افزایش یابد. بهینهسازهای قدرتمندتر دیگری نیز در کتابخانه و دومت وجود دارند. یکی از معروف ترین بهینهسازها، بهینهساز Adam میباشد.

? به طور مختصر نحوه کار بهینهساز Adam و تفاوت آن با بهینهساز SGD را توضیح دهید.

- حال یک مدل Convolutional با استفاده از داده های train و Adam optimizer آموزش دهید و عملکرد شبکه را روی
 داده های test بررسی کنید و معیارهای Accuracy, Recall, Precision و F1-Score را گزارش کنید.
- نکته ۱: دقت کنید که ورودی شبکهی عصبی در این فاز، تصاویر هستند که ساختار چند بعدی دارند. همانطور که در کد نمونه مشاهده می کنید، برای شبکهی کانولوشنی، نیاز است که پس از اتمام لایههای کانولوشنی، ویژگیهای استخراج شده توسط فیلترها Pense شده و وارد لایههای Dense برای طبقه بندی شوند.
- نکته ۲: در تمامی مراحل تمرین، پس از اتمام تمامی epoch ها باید نمودار مقدار loss و loss بر حسب epoch را در مرحله می توانید یک تابع برای این کار تعریف کرده و هر بار تابع خود را فراخوانی کنید.
- نکته ۳: در تمامی مراحل تمرین، پس از اتمام تمامی epoch ها باید معیارهای precision, recall, F1 برای دادههای تمرین و تست گزارش شود ۱۰ .

-

¹³ Stochastic Gradient Descent

¹⁴ مى توانيد از تابع classification_report كتابخانه sklearn براى اين منظور استفاده كنيد.

بخش دوم: آموزش شبکه با دادههای نامتوازن

وجود تعداد دادههای نامتوازن می تواند باعث کاهش کارایی و دقت مدل در پیش بینیها شود. وقتی دادهها در دسته بندیهای مختلف توزیع نامتوازنی دارند، مدل تمایل به پیش بینی کلاس اکثریت خواهد داشت و نتایج اشتباهی برای کلاسهای اقلیت تولید می کند. برای روشن تر شدن، فرض کنید که در یک مسئله تشخیص تصویر، دو کلاس داریم: "گربه" و "سگ". اکنون فرض کنید که ۹۰٪ از تصاویر داده شده برای آموزش، تصاویر سگ است و ۱۰٪ تصاویر گربه. در این حالت، شبکه عصبی به سادگی می تواند همیشه کلاس سگ را پیش بینی کند و با این روش دقت ۹۰٪ را داشته باشد. اما واقعیت این است که شبکه عصبی وظیفه یادگیری ویژگیهای هر دو کلاس را دارد و نمی تواند به درستی تمایز بین تصاویر گربه و سگ را تشخیص دهد.

برای مقابله با این نامتوازنی داده، روشهای مختلفی وجود دارد. یک رویکرد متداول برای حل این مسئله استفاده از تکنیکهای نمونه برای مقابله با این نامتوازنی داده، روشهای مختلفی وجود دارد. یک رویکرد متداول برای حل این مسئله استفاده از طریق تکرار نمونه برای مجدد این مجدد کارس اقلیت را افزایش می دهیم (از طریق تکرار نمونه های مصنوعی مشابه نمونه های اصلی) تا تعادلی بین کلاس ها برقرار شود.

همچنین، معیارهای مختلفی مانند AUC (مساحت زیر نمودار ROC) و ماتریس درهمریختگی می توانند برای ارزیابی صحت مدل در حالت نامتوازن مورد استفاده قرار گیرند که در این پروژه مورد توجه ما نمی باشند. این معیارها می توانند برای نشان دادن کارایی و عملکرد مدل در تشخیص کلاسهای کمتر و بررسی نقاط ضعف و قوت مدل در مواجهه با دادههای نامتوازن مورد استفاده قرار بگیرند. در این بخش به بررسی تاثیر وجود تعداد دادههای نامتوازن در مجموعه داده خود می پردازیم.

- ابتدا ۸۰ درصد دادههای کلاس اول و دوم (یعنی هواپیما و ماشین) را به صورت تصادفی انتخاب کرده و از مجموعه دادهی
 آموزش خود حذف کنید (داده ی تست بدون تغییر باقی می ماند).
- شبکهی کانولوشنی مشابه مدلی که در بخش اول استفاده کردید، تعریف کنید و این شبکه را روی دادههای نامتوازنی که
 ایجاد کردید آموزش دهید.
- نتایج Recall و F-1 روی داده ی تست کلاس هواپیما و ماشین را با بخش اول مقایسه کنید. دلیل تغییر این مقادیر را شرح
 دهید.

-

¹⁵ re-sampling

¹⁶ Confusion Matrix

بخش سوم: استفاده از Data Augmentation برای ایجاد توازن بین تعداد دادهها

همانطور که اشاره شد، یک راه برای مقابله با مشکل تعداد دادههای نامتوازن در مجموعه داده، استفاده از روشهای تقویت داده یا Data Augmentation است. این روشها برای افزایش تنوع دادهها و ایجاد تعادل بین کلاسها استفاده می شوند.

یکی از روشهای متداول برای افزایش تعداد دادهها در شبکههای عصبی کانولوشنی، اعمال تبدیلات جزئی بر روی تصاویر است. مثلاً میتوان تصاویر را به صورت تصادفی برش داد، شدت روشنایی را تغییر داد، آنها را به صورت تصادفی برش داد، شدت روشنایی را تغییر داد و یا با اعمال نویز به تصاویر، تنوع بیشتری از تصاویر موجود به دست آورد. بدین ترتیب شبکه عصبی میتواند الگوهای مختلف را یاد بگیرد و مشکل نامتوازن بودن تعداد دادههای کلاسهای مختلف تا حدی برطرف شود.

در این بخش، باید این تبدیلات را به صورت تصادفی بر روی تصاویر مجموعه داده ی آموزش اعمال کنید. با این کار، تعداد دادههای هواپیما و ماشین را متعادل تر می کنیم و شبکه عصبی قادر خواهد بود تا الگوهای هر دو کلاس را به خوبی یاد بگیرد.

نکته مهمی که باید در نظر داشته باشیم، این است که Data Augmentation را فقط بر روی دادههای آموزش اعمال می کنیم و دادههای تست را بدون هیچگونه تغییری باقی می گذاریم. در واقع، ما فقط تصاویر اصلی را در مجموعه داده ی آموزش داریم و تبدیلات تصادفی را هر بار بر روی تصاویر اعمال می کنیم تا دادههای تقویت شده ای بسازیم.

- با استفاده از روشهایی مانند RandomRotation ،RandomFlip ،RandomZoom یا هر روش دیگری که به نظر شما
 مناسب است، تعداد دادههای کلاس هواپیما و ماشین را افزایش داده و دوباره مجموعه دادهای متوازن ایجاد کنید.
- شبکهی کانولوشنی مشابه مدلی که در بخش اول و دوم استفاده کردید، تعریف کنید و این شبکه را روی دادههای تقویت شدهای که ایجاد کردید آموزش دهید.
- نتایج Recall و F-1 روی داده ی تست کلاس هواپیما و ماشین را با بخش دوم مقایسه کنید. نتایج به دست آمده را تحلیل کنید.

بخش چهارم: تاثیر روشهای Regularization در فرآیند آموزش

روشهای Regularization در فرآیند آموزش شبکه عصبی مورد استفاده قرار می گیرند تا از بیش برازش ¹⁷ جلوگیری کنند. دو روش معروف در این زمینه Dropout و Batch Normalization هستند که به ترتیب در ادامه توضیح داده خواهند شد:

- 1. Dropout: این روش یکی از تکنیکهای مؤثر برای کاهش بیش برازش در شبکههای عصبی است. این روش در فرآیند آموزش، به صورت تصادفی برخی از نورونها را در هر مرحله غیرفعال می کند. به این ترتیب، هر نورون مجبور است الگوهای مفید را در حضور نورونهای دیگر یاد بگیرد و به این ترتیب میزان وابستگی نورونها به هم و همچنین میزان وابستگی شبکه به هر نورون کاهش می یابد. این باعث می شود که شبکهی عصبی توانایی تعمیم پذیری بهتری پیدا کند و کمتر دچار بیش برازش شود. Dropout با ضریب احتمال Dropout خاصی در ساختار شبکه اعمال شود.
- 2. Batch Normalization: این تکنیک، یک روش استانداردسازی ورودی های هر بچ (batch) در یک لایه است. در هر بچ از داده ها، میانگین و واریانس ورودی ها محاسبه شده و سپس ورودی ها به صورت استانداردسازی شده با استفاده از میانگین و واریانس محاسبه شده، وارد لایه ی بعدی می شوند. این روش باعث می شود که توزیع ورودی ها در هر لایه بهبود یابد و بیش برازش کاهش یابد. علاوه بر این، Batch Normalization نیز می تواند به عنوان یک لایه Batch Norm در ساختار شبکه اعمال شود.

در نتیجه، استفاده از Dropout و Batch Normalization در شبکه عصبی، می تواند تأثیر مثبتی بر عملکرد شبکه در فرآیند آموزش داشته باشد. این روشها موجب کاهش بیش برازش، بهبود تعمیم پذیری و عملکرد مدل در مجموعه داده های تست می شوند. با اعمال این روشها، شبکه عصبی قادر خواهد بود الگوهای مفید را یاد بگیرد و همچنین مقاومت بیشتری در برابر داده های نویزی و تغییرات کوچک در ورودی ها خواهد داشت.

در این بخش، باید از لایه Dropout بین لایههای Dense و لایه Batch Normalization بین لایههای کانولوشنی در شبکه خود استفاده کنید.

- شبکه ی کانولوشنی مشابه مدلهای قبلی، تعریف کنید و لایههای Dropout و BatchNorm را با توجه به توضیحات ارائه شده به مدل خود اضافه کنید. این شبکه را روی دادههای اصلی آموزش دهید (از دادههای نامتوازن یا augment شده در این بخش استفاده نکنید).
 - نتایج Recall و F-1 روی داده ی تست را با بخش اول مقایسه کنید. نتایج به دست آمده را تحلیل کنید.

overfitting

منابع

برای یادگیری شروع کار با Keras میتوانید از این لینک کمک بگیرید.

برای آموزش شیوه استفاده از Google Colab می توانید از این لینک یا این لینک استفاده نمایید.

در این لینک دربارهی Adam Optimizer مطالعه کنید.

در این لینک دربارهی روشهای مختلف regularization مطالعه کنید.

نكات ياياني

- دقت کنید که هدف پروژه تحلیل نتایج و تاثیر عوامل مختلف است؛ بنابراین از ابزارهای تحلیل داده بطور مثال نمودارها استفاده کنید و توضیحات مربوط به هر بخش از پروژه را بطور خلاصه و در عین حال مفید در گزارش خود ذکر کنید.
- نتایج و گزارش خود را در یک فایل فشرده با عنوان AI-CA5-P2-<#SID>.zip تحویل دهید. محتویات پوشه باید شامل فایل های مورد نیاز برای اجرای آن باشد. تحلیل و نمایش خروجیهای فایل html و فایلهای مورد نیاز برای اجرای آن باشد. تحلیل و نمایش خروجیهای خواسته شده بخشی از نمره این تمرین را تشکیل میدهد. از نمایش درست خروجیهای مورد نیاز در فایل مطمئن شوید.
- توجه داشته باشید که علاوه بر ارسال فایلهای پروژه ، این پروژه به صورت حضوری نیز تحویل گرفته خواهد شد. بنابراین تمام بخشهای پروژه باید قابلیت اجرای مجدد در زمان تحویل حضوری را داشته باشند. همچنین در صورت عدم حضور در تحویل حضوری نمرهای دریافت نخواهید کرد.
- در صورتی که سوالی در مورد پروژه داشتید بهتر است در فروم یا گروه درس مطرح کنید تا بقیه از آن استفاده کنند؛ در غیر
 این صورت با طراحان در ارتباط باشید.
 - هدف از تمرین، یادگیری شماست. لطفا تمرین را خودتان انجام دهید.

موفق باشيد.