

بسمه تعالی

گزارش کار تمرین اول یادگیری عمیق

تمرین 1:

طبق داده های موجود در جدول و با توجه به دو کلاسه بودن مساله احتمال شرطی طبق لاپلاس و با در نظر گرفتن ضریب الفای یک به صورت ضرب کسر ها به دست آوردم و مقدار احتمال 6 عبارت داده شده محاسبه شد.

در ادامه کلاس مربوط به هر داده به دست آورده شد و دسته بندی شدند

سوال 3:

Mnsit:

دیتاستی است قدیمی شامل تصاویر 28×28 پیکسل و 10 رقم مختلف دارد از 0 تا 9. دارای 60 هزار تصویر آموزشی و 10 هزار تصویر تست است. برای خواندن این دیتاست از لود کردن آن استفاده میشود

CIFAR-10:

مجموعه تصاویری است که معمولاً برای آموزش الگوریتم های یادگیری ماشین و بینایی رایانه استفاده می شود. این مجموعه داده یکی از پرکاربردترین مجموعه های داده برای تحقیقات یادگیری ماشین است. مجموعه داده شامل 60,000 تصویر رنگی 32×32 در 10 کلاس مختلف است

کلاس مختلف هواپیما ، ماشین ، پرند ، گربه ، آهو ، سگ ، قورباغه ، اسب ، کشتی و کامیون را نشان می دهد. از هر کلاس 10 6000 تصویر وجود دارد

الگوریتم های رایانه ای برای تشخیص اشیا در عکس ها اغلب با مثال بهتر یاد می گیرند که می تواند برای یادگیری نحوه تشخیص اشیا به کامپیوتر استفاده شود. از آنجا که تصاویر موجود با وضوح کم (32×32) است ، این مجموعه داده می تواند به محققان اجازه دهد تا به سرعت خیلی بالا الگوریتم های مختلف را امتحان کنند تا ببینند چه عواملی برای یادگیری ماشین مفید است

زیرمجموعه ی مجموعه از جمع اوری 80 میلیون عکس که تمامی آن ها با استفاده از لیبیل (هشتگ) نام گذاری شده است. هنگام ایجاد این مجموعه داده ، از دانشجویان استفاده شد و برای برچسب گذاری به تک تک دانشجویان به ازای هر لیبیل پول پرداخت می شد

Fer2013:

حاوی تقریباً 30000 تصویر صورت با عبارات مختلف با اندازه محدود به 48×48 است و برچسب های اصلی آن را

می توان به 7 نوع تقسیم کرد: 0 = عصبانی، 1 = انزجار، 2 = ترس، 3 = شاد، 4 = غمگین. 5=سورپرایز 6=خنثی. عبارت Disgust تصویر 600 را دارد ، در حالی که برچسب های دیگر هر کدام نزدیک به 5000 نمونه دارند

با استفاده از این تابع یک ارایه نامیای یا یک وکتور که شامل اعداد صحیح است را نشان میدهد که قابل تبدیل به صورت ماتریسی است. با توجه به اکتیویشن لایه اخر نباید به صورت گسسته به دست آیند این متد روی وزن شبکه بی اثر است . با استفاده از تابع می توانیم هر کدام از کلاس ها را به یک ستون 0 و 1 تبدیل می نمایم

X_train متغیر

به عنوان متغیر مستقل و متغیر

y_train

به عنوان متغیر وابسته در مدل نقش دارند

در مدل اول یکس ترین دیتای آموزشی است بعد اول تعداد داده آموزشی، بعد دوم و سوم ابعاد تصویر و ایگرگ ترین لیبل های دیتای آموزشی است بعد اول تعداد لیبل ها و بعد دوم کلاس است

در مدل دوم یکس ترین در بعد اول داده آموزشی دارد بعد دوم هم ابعاد تصویر را نشان میدهد. بعد چهارم هم سه پارامتری است.

ایگرگ ترین در بعد اول حاوی تعدادی لیبل است و بعد دوم ان کلاس دارد.

در مدل سوم یکس ترین در بعد اول دیتا بعد دوم و سوم ابعاد تصویر و بعد چهارم کانال است.

ایگرگ ترین هم بعد اول لیبل بعد دوم کلاس است.

تقویت دیتاست و افزودن داده های تصویری تکنیکی است که با استفاده از آن می توان به طور مصنوعی و بدون تصویربرداری جدید، و تنها با ایجاد نسخه های دستکاری شده از تصاویر موجود اندازه دیتاست را افزایش داد

آموزش شبکه های عصبی مصنوعی عمیق با داده های بیشتر می تواند به داشتن شبکه های قوی تر منتهی شود تکنیک های ان می توانند گونه هایی از تصاویر را بسازند شبکه با یادگیری خصوصیات جدیدی از آن تصاویر علاوه بر تصاویر اصلی، قدرت درک گستره بیشتری از هر شی را پیدا می کند. مصنوعی تصاویر جدیدی برای آموزش تولید کند که از تصاویر اصلی به دست می آیند. این کار در هر کاربرد و دامنه ای از شبکه های یادگیری عمیق روش های مخصوص به خود را دارد؛ مثلاً در یادگیری تشخیص گفتار انسانی، صداها کمی نویز به سیگنال صوت استفاده می کنیم تا شبکه تشخیص کلمات یا احساسات را حتی با صدا های غیر شفاف نیز یاد بگیرد. اما در تشخیص وجود شی در تصویر یا پیدا کردن دقیق جای چند شی در یک تصویر، تکنیک های دیگری مثل چرخش و تغییر روشنایی و بریدن تصویر استفاده می شود

منظور ما از دگرگونی که گفتیم گستره ای از عملیات های پردازش تصاویر مانند شیفته ها، برگردان ها، بزرگنمایی ها و چیزهای خیلی بیشتر است

هدف هم افزایش اندازه دیتاست با استفاده از تصاویر جدیدی است که محتمل و باور کردنی باشند (مثلاً نمی توان برای آموزش شبکه تشخیص گربه از سگ در تصاویر، همه تصاویر را به صورت افقی برعکس کرد، چرا که سگ ها و گربه ها روی سقف و هوا حرکت نمی کنند!)

پس مشخص است که انتخاب این تکنیک باید با دقت و مطابق با محتوای دیتاست و نوع مسئله انجام شود. می توانید در این زمینه به آزمایش و خطا هم بپردازید، به این شکل که بخش کوچکی از تصاویر آموزش را با شبکه ای کوچک آموزش دهید و در این آموزش از تنظیم های مختلفی برای تولید تصاویر جدید استفاده کنید و بهبود یا عدم بهبود شبکه را به طور مداوم در نظر داشته باشید

الگوریتم های مدرن یادگیری عمیق، مثل شبکه های عصبی کانولوشنی می توانند ویژگی هایی را یاد بگیرند که مستقل از جای آن ها در تصویر است. با این وصف تقویت و افزودن داده ها با تغییر چرخش و نور و رنگ، می تواند کمک بیشتری به این یادگیری ویژگی های مستقل در تصویر بکند؛ در نهایت شبکه به خوبی یاد می گیرد شی را اگر در تصویر از چپ به راست برگردانده شده بود یا نور و شفافیت کمی داشت یا خیلی چیزهای دیگر، باز هم به درستی تشخیص دهد

از جمله کاربرد های ان:

تکنیک شیفت عمودی و افقی تصاویر

تقویت و گسترش دیتاست

برگرداندن تصاویر از چپ به راست و بالا به پایین

تکنیک چرخش تصویر به مقدار تصادفی

تکنیک بزرگنمایی تصویر به اندازه تصادفی

تکنیک تغییر روشنایی به مقدار تصادفی

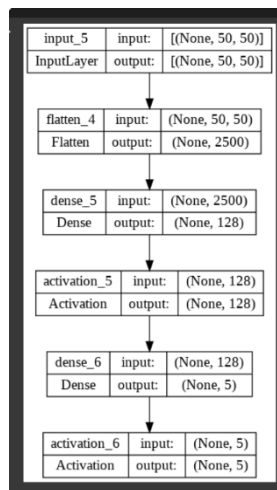
خلاصه:

را برای افزایش اندازه دیتاست استفاده کنید تا بهبود عملکرد شبکه یادگیری عمیق و تعمیم قدرت Image data augmentation تشخیص اشیا حاصل شود.

پایه سازی شده است ImageDataGenerator پشتیبانی می شود و در کلاس Keras در کتابخانه Image data augmentation

نحوه استفاده از شیفت، معکوس کردن (برگرداندن)، چرخش، تغییر روشنایی و بزرگنمایی تصاویر مرور شد

خروجی دو مدل ترسیم شده:

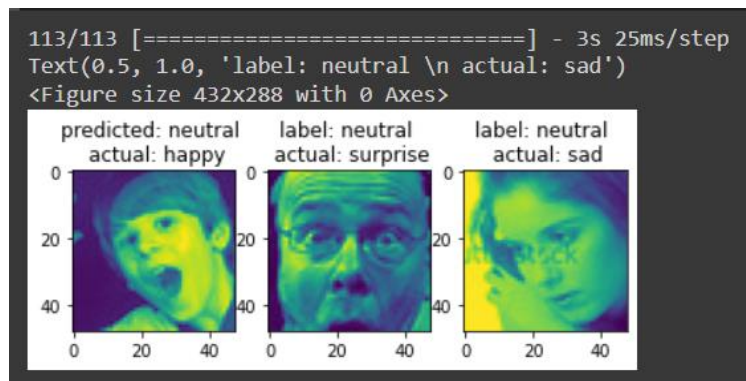
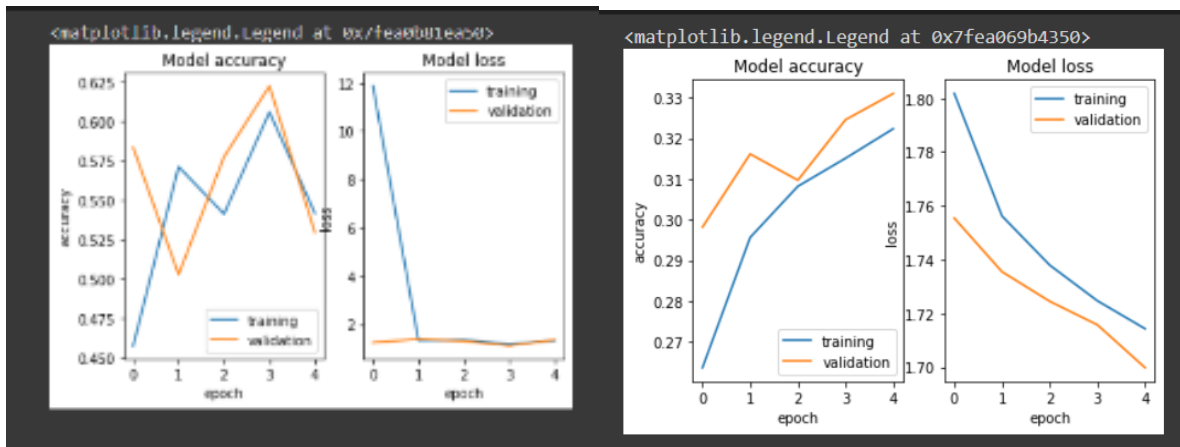


Summery:

ساختار و جزئیات هر لایه استفاده شده و تعداد نوروں ها تعداد پارامترها و پارامترهای یادگیری را اضافه می کنند

<div>Model Summary</div> <div><pre>[1] model_template_2.summary() Model: "model" Layer (type) Output Shape Param # ----- input_9 (InputLayer) [(None, 50, 50)] 0 flatten_8 (Flatten) (None, 2500) 0 dense_10 (Dense) (None, 128) 320128 activation_10 (Activation) (None, 128) 0 dense_11 (Dense) (None, 5) 645 activation_11 (Activation) (None, 5) 0 Total params: 320,773 Trainable params: 320,773 Non trainable params: 0</pre></div> <div>Plotting Model and Save it as structure_2.png</div>	<div>model_template_2.summary()</div> <div><pre>Model: "model" Layer (type) Output Shape Param # ----- input_9 (InputLayer) [(None, 50, 50)] 0 flatten_8 (Flatten) (None, 2500) 0 dense_10 (Dense) (None, 128) 320128 activation_10 (Activation) (None, 128) 0 dense_11 (Dense) (None, 5) 645 activation_11 (Activation) (None, 5) 0 Total params: 320,773 Trainable params: 320,773 Non trainable params: 0</pre></div>
---	---

Mnist & Fer2013 model:



عدم تطابق دیده میشود. و بسیاری از داده های تست به اشتباه تخمین زده شده اند. چون صورت انسان مورد بررسی است نیازمند دیتاست بسیار کامل هستیم تا دقت افزایش یابد. در نتیجه نیاز به تعداد لایه و نرون های بیشتر ضروری میشود.

منابع:

<https://open-mind.ir/1399/%D8%A7%D8%B3%D8%AA%D9%81%D8%A7%D8%AF%D9%87-%D8%A7%D8%B2-image-data-augmentation-%D8%AF%D8%B1-%DA%A9%D8%AA%D8%A7%D8%A8%D8%AE%D8%A7%D9%86%D9%87-keras/>

<https://dataio.ir/%DB%8C%D8%A7%D8%AF%DA%AF%DB%8C%D8%B1%DB%8C-%D8%B9%D9%85%DB%8C%D9%82-q7q1i0nglhx>

<https://deeplearning.ir/%D8%AF%DB%8C%D8%AA%D8%A7%D8%B3%D8%AA/>

<https://colab.research.google.com/drive/13OJEslwgMNEvDbnUaSONFTOLJ8tZnaY9#scrollTo=SyCzEMRYx>

LMC

