

پروژه پایانی

فاز اول:

داده های Mnist شامل ۶۰ هزار تصویر ۲۸*۲۸ پیکسل اعداد ۰ تا ۹ دستنویس در مجموعه train و ۱۰ هزار داده در مجموعه test است.

(۱) با تعلیم یک شبکه عصبی با نورونهای Sigmoid در لایه های مخفی و لایه خروجی، روی داده های train و test، خطا را (خطای طبقه بندی) بدست آورید. لایه خروجی را SoftMax در نظر بگیرید. برای تابع هدف نیز Crossentropy را به عنوان تابع زیان انتخاب کنید. با تعداد نورون و تعداد لایه های متفاوت تست کنید. کدام ساختار بهتر است؟ توجه تان برای آن چیست؟

(۲) تابع فعالیت نورون ها را از Sigmoid به ReLu تغییر دهید. برای تعداد لایه های زیادتری الگوریتم را مشابه مورد ۱ اجرا کنید. برای هر کدام از آزمایشاتی که انجام می دهید منحنی فراگیری (learning curve) را برای loss و accuracy نمایش دهید.

فاز دوم:

در این فاز قصد داریم از دیتاست [The Oxford-IIIT Pet Dataset](#) استفاده کنیم. این دیتاست شامل ۳۷ کلاس داده می باشد که هر کلاس به صورت تقریبی شامل ۲۰۰ دیتا می باشد.

با استفاده از ۳ مدل resnet-۱۸ و mobile-net v۲ و google-net قصد داریم فیچر هایی را استخراج کنیم و سپس با استفاده از الگوریتم SVM و کرنل RBF کلاس بندی کنیم.

برای مدل resnet-۱۸ فیچر های خروجی از avgpool و بلاک های ۱ و ۳ مورد بررسی می باشد.

برای مدل google-net فیچر های خروجی از لایه های inception(۳b) و inception(۴e) و avgpool مورد بررسی بر اساس شکل زیر می باشد.

type	patch size/ stride	output size	depth	# 1×1	# 3×3 reduce	# 3×3	# 5×5 reduce	# 5×5	pool proj	params	ops
convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								

برای مدل mobile-net نیز فیچر های خروجی avgpool و خروجی کانولشون اول و لایه میانی آن بررسی می‌شود.

۸۰ درصد داده ها را به عنوان داده های train و ۲۰ درصد را به عنوان داده های test در نظر بگیرید.

برای این فاز نیاز است که ROC curve و Confusion Matrix را به ازای هر بخش به دست آورید.

فاز سوم:

در این بخش قصد داریم با استفاده از دیتاست و سه مدل فاز قبلی که از پیش آموزش دیده هستند از لایه avgpool هر مدل فیچر استخراج کنید. سپس با استفاده از الگوریتم خوشه بندی k-means فیچر های به دست آمده از هر کلاس را خوشه بندی می‌کنیم. پس از خوشه بندی چندین تصویر از دیتاست انتخاب کرده و برحسب دو معیار شباهت آن ها را با خوشه های به دست آمده می‌سنجیم و خوشه مشابه تر را به عنوان نماینده آن در کنار تصویر داده شده نمایش می‌دهیم و با یکدیگر مقایسه می‌کنیم.

دقت شود که در این پروژه فاز ۲ و ۳ ام پروژه نیاز به آموزش شبکه نیست.

در صورتی که پردازش دیتا ها زمان بر بود میتوانید از گوگل کولب استفاده کنید و فرمت ران تایم آن را به gpu تغییر دهید در اینصورت میتوانید از گرافیک برای پردازش و بررسی آزمایشات استفاده کنید. در صورتی که باز هم مشکل پردازشی داشتید میتوانید از PCA برای کاهش ابعاد فیچر ها استفاده کنید که در این صورت نیاز به پردازش های بالا نباشد.

موفق باشید.