



به نام خدا
دانشگاه تهران
دانشکده مهندسی برق و کامپیوتر



درس شبکه‌های عصبی و یادگیری عمیق

تمرین دوم

محمد امین نورائی یگانه	نام دستیار طراح	پرسش ۱
amin.nooraei@ut.ac.ir	رایانامه	
مائده صادقی	نام دستیار طراح	پرسش ۲
Maisa.sadeqi@gmail.com	رایانامه	

فهرست

- پرسش ۱. تشخیص بیماران مبتلا به COVID-۱۹ با استفاده از تصاویر X-Ray ۳
- ۱-۱. معرفی مقاله و تمرین ۳
- ۲-۱. تحلیل دیتاست ۳
- ۳-۱. پیش پردازش داده‌ها ۴
- ۴-۱. آماده‌سازی مدل ۴
- ۵-۱. آموزش و ارزیابی مدل ۵
- ۶-۱. یادگیری انتقالی ۵
- پرسش ۲. پیاده سازی یک سیستم طبقه بندی خودرو با استفاده از VGG16 و SVM ۷
- ۱-۲. مقدمه ۷
- ۲-۲. پیش‌پردازش داده‌ها ۸
- ۳-۲. استخراج ویژگی‌ها ۹
- ۴-۲. آموزش و ارزیابی مدل ۹
- ۵-۲. تحلیل نتایج ۱۰
- ۶-۲. امتیازی ۱۱

فهرست اشکال

شکل ۱. گام های پیاده سازی ۸

شکل ۲. نمایش گرافیکی عملکرد به دست آمده توسط مدل های مختلف برای مجموعه داده تویوتا ۱۰

پرسش ۱. تشخیص بیماران مبتلا به COVID-۱۹ با استفاده از تصاویر X-Ray

۱-۱. معرفی مقاله و تمرین

در دوران کرونا، یکی از چالش‌های اساسی جامعه پزشکی، تشخیص سریع و به‌موقع افراد مبتلا به COVID-۱۹ بود. در برخی موارد، تراکم بالای مراجعه‌کنندگان به مراکز درمانی به حدی بود که عملکرد پزشکان در بررسی و تشخیص افراد مبتلا را کاهش می‌داد یا موجب می‌شد بیماران برای ملاقات و بررسی تصاویر X-Ray توسط پزشک، ساعات زیادی منتظر بمانند. در این [مقاله](#)، یک گروه پژوهشی به طراحی و ارزیابی یک مدل CNN پرداخته است تا با آموزش بر روی تصاویر X-Ray افراد مبتلا به COVID-۱۹ و افراد سالم، تشخیص بیماری را با دقت بالا و به‌صورت سریع انجام دهد. در این پژوهش، از دو مجموعه داده (دیتاست) استفاده شده است که امروزه دیگر در دسترس نیستند. به همین دلیل، ما از دیتاست موجود در [Kaggle](#) استفاده می‌کنیم. این دیتاست از نظر تعداد کلاس‌های طراحی شده برای تشخیص، با مقاله اصلی کمی متفاوت است و شامل سه کلاس می‌باشد. هدف از این تمرین، پیاده‌سازی مدل ارائه‌شده در مقاله و تغییر آن بر مبنای نیازهای خود، با استفاده از دیتاست جدید است. همچنین، در این تمرین قصد داریم از مفهوم Transfer Learning که پیش‌تر در فصل سوم درس با آن آشنا شده‌اید، استفاده کنیم. می‌خواهیم بررسی کنیم که چرا باوجود مدل‌های از پیش آموزش‌دیده‌شده در حوزه Image Classification، پژوهشگران همچنان به طراحی مدل‌های جدید برای کاربردهای مختلف می‌پردازند.

۲-۱. تحلیل دیتاست

(۱۰ نمره)

در مواجهه با یک دیتاست، لازم است شناخت دقیقی از آن پیدا کنیم تا بتوانیم آن را برای مدل خود آماده‌سازی کنیم. بنابراین، اولین مرحله در پیاده‌سازی و آموزش هر مدل، شناخت کامل داده‌هایی است که به مدل ارائه می‌شود و فراخوانی آن‌ها. ازجمله اطلاعاتی که باید به دست آوریم، نوع توزیع داده‌ها و فرمت آن‌ها است. در این بخش، ابتدا دیتاست را فراخوانی کرده و سپس به بررسی موارد زیر می‌پردازیم:

- شامل چه کلاس‌هایی است؟
- فرمت یا فرمت‌های داده‌های موجود در دیتاست چیست؟
- توزیع داده‌ها در هر کلاس به چه صورت است؟
- هیستوگرام تعداد داده‌های هر کلاس را برای هر دو مجموعه Train و Test رسم کنید.

- توضیح دهید که بالانس بودن کلاس‌های دیتاست چه مزیتی در شبکه‌های عصبی دارد؟
- در صورت نامتعادل بودن (عدم بالانس) دیتاست، چه راهکاری پیشنهاد می‌دهید؟ (راهکارهای پیشنهادی شامل معرفی ابزارها و کتابخانه‌های مختلف نیز باشد)

۱-۳. پیش‌پردازش داده‌ها

(۲۰ نمره)

در این قسمت به آماده‌سازی داده می‌پردازیم. پس از رسم هیستوگرام توزیع داده‌ها در دو پوشه train و test کمترین تعداد نمونه در سه کلاس را در هر پوشه مشخص و به همان تعداد به صورت تصادفی برای هر دسته نمونه برداشته و آماده کنید. در این مرحله باید این داده‌ها را آماده برای آموزش و ارزیابی مدل کنید. سپس در مرحله بعد از داده‌های آماده‌شده مرحله قبل برای Data Augmentation استفاده کنید. در مرحله اول توضیح دهید که Data Augmentation چیست و کاربرد آن چیست. سپس به وسیله روش‌های Data Augmentation دیتاست به دست آمده در مرحله قبل را به ۴ تا ۶ برابر تبدیل کنید و توضیح دهید از چه روش‌هایی برای این کار و به چه علت استفاده کرده‌اید؟ دیتاست جدید تشکیل شده را ذخیره کرده تا در مراحل بعد از آن‌ها استفاده کنید.

همچنین در این مرحله لازم است که یک نمونه از داده‌های دیتاست را انتخاب و با اعمال روش‌های مختلف داده افزایی نمونه‌های جدید تولید شده از نمونه جدید را نمایش دهید.

۱-۴. آماده‌سازی مدل

(۲۰ نمره)

در این قسمت لازم است به مقاله مراجعه کرده و مدل CNN ارائه شده را بررسی و پیاده کنید. این مدل و جزئیات آن در جدول شماره ۳ و شکل شماره ۴ ارائه شده است. پس از پیاده‌سازی مدل از آن Summary گرفته و با جزئیات ارائه شده در مقاله چک کنید. همچنین سائز ورودی تصاویر را مطابق با مدل تغییر دهید و خروجی مدل را مطابق دیتاست خود تغییر دهید. در مقاله نرخ یادگیری بیان نشده و از نرخ یادگیری متغیر در طول آموزش استفاده شده است درباره این روش تحقیق و در صورت نیاز از آن استفاده کنید در غیر این صورت اگر از یک نرخ یادگیری ثابت استفاده می‌کنید دلیل آن را بیان و ذکر کنید سایر نرخ‌ها چه تأثیری بر آموزش داشته است؟ (بیان علت به وسیله نمودار خطا و دقت دارای اعتبار بیشتری خواهد بود).

۱-۵. آموزش و ارزیابی مدل

(۳۰ نمره)

مدل را مطابق با جدول ۳ و در ۱۵ الی ۱۰۰ epoch آموزش دهید. (بسته به سیستم خود مختارید ولی سعی کنید در epoch بیشتری آموزش دهید تا به یک نقطه پایداری از دقت برسید. ۳۰ epoch نیز کافی است بیشتر از آن اختیاری است). نرخ یادگیری را باید خودتان به دست آورید که با توجه به دیتاست و مدل چه نرخ مناسب است و چرا؟ از ۳۵ درصد داده‌های train برای اعتبارسنجی استفاده کنید.

- نمودار خطا و دقت در طول روند آموزش را رسم کنید.
- مدل را بر روی داده‌های test تست کرده و عملکرد را گزارش کنید. (راهکاری برای بهتر کردن آن دارید؟ حتی اگر لازم است معماری مدل را تغییر دهید)
- مدل را با معیارهای Accuracy، Precision، f1-score و recall ارزیابی و هرکدام از این معیارها را به صورت مفهومی توضیح دهید که هرکدام می‌تواند چه دیدی از مدل آموزش دیده به طراح آن دهد.
- Confusion Matrix را تشکیل داده و بررسی کنید کدام کلاس راحت تر و کدام سخت تر تشخیص داده شده‌اند؟

لازم به ذکر است که تمامی خروجی‌های ذکر شده را تحلیل کرده و کامل در گزارش بیاورید. چیزی که در این تمرین مهم است میزان و نوع تحلیل شما از چالش‌های شما در آموزش این مدل خواهد بود. به علت عوض شدن دیتاست نیازی نیست که به خروجی‌های مقاله دست پیدا کنید ولی لازم است که عملکرد مدل شما مناسب باشد و قدرت تعمیم خوبی ثبت کرده باشد.

۱-۶. یادگیری انتقالی

(۲۰ نمره)

در این بخش به کاربرد یادگیری انتقالی (Transfer Learning) خواهیم پرداخت. همان طور که در درس ذکر شد از این تکنیک برای استفاده از مدل‌های از پیش آموزش دیده در حوزه‌های مختلف استفاده می‌شود. لطفاً این تکنیک را توضیح داده و ذکر کنید چه بخشی از مدل استفاده شده استفاده و چه بخشی حذف می‌شود؟ چرا؟ پس از آن لازم است که بر روی دیتاست اولیه که در بخش قبل تشکیل دادید دو مدل VGG۱۶، MobileNetV۲ را آموزش داده و عملکرد آن‌ها را نیز مطابق بخش قبل ثبت کنید. سپس در یک جدول به طور کامل عملکرد هر سه مدل را ارزیابی و تحلیل کنید. مزایا و معایب هر مدل را ذکر و

به‌طور خاص ذکر کنید باوجود چنین مدل‌های از پیش آموزش‌دیده‌ای چرا هنوز نیاز است که پژوهشگران به طراحی مدل‌هایی شبیه به مدل ارائه‌شده در مقاله بپردازند؟

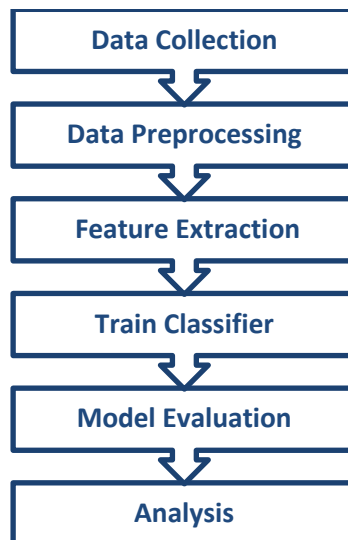
پیش ۲. پیاده‌سازی یک سیستم طبقه‌بندی خودرو با استفاده از SVM و VGG16

۲-۱. مقدمه

با توجه به سرعت تکامل هوش مصنوعی، تکنیک‌های یادگیری عمیق و یادگیری ماشین به‌طور گسترده‌ای برای طبقه‌بندی تصاویر استفاده می‌شود. طبقه‌بندی وسایل نقلیه، به‌ویژه، به یک کاربرد اساسی در سیستم‌های حمل‌ونقل هوشمند، جمع‌آوری خودکار عوارض و نظارت بر امنیت تبدیل شده است. با این حال، مدل‌های یادگیری عمیق سنتی اغلب به مقادیر زیادی داده برچسب‌گذاری شده و قدرت محاسباتی برای دستیابی به دقت بالا نیاز دارند. برای پرداختن به این موضوع، یک رویکرد ترکیبی از استخراج ویژگی‌های عمیق و طبقه‌بندی یادگیری ماشین می‌تواند راه‌حل مؤثری ارائه دهد. با افزایش تنوع مدل‌های خودرو، نیاز به روش‌های دقیق و کارآمد برای شناسایی و طبقه‌بندی آن‌ها بیشتر شده است.

در این پروژه، شما یک مدل ترکیبی مبتنی بر VGG^{۱۶} و SVM^۱ برای طبقه‌بندی خودروها پیاده‌سازی خواهید کرد. مدل VGG^{۱۶} به‌عنوان استخراج‌کننده ویژگی‌های تصویر عمل می‌کند، درحالی‌که SVM به‌عنوان یک طبقه‌بند مؤثر، داده‌های استخراج‌شده را تحلیل و دسته‌بندی می‌کند. هدف این روش، افزایش دقت و کاهش نرخ خطا در طبقه‌بندی تصاویر خودروهای برند تویوتا است. در انتها با مدل‌هایی که در مطالعات پیشین مورد استفاده قرار گرفته‌اند آشنا خواهید شد و ضمن پیاده‌سازی آن‌ها، در مورد نحوه‌ی عملکرد و تفاوت‌ها یک توضیح جامع ارائه خواهید کرد. به‌طور مختصر در این سؤال گام‌های ذیل طی خواهد شد:

^۱ Support vector machines



شکل ۱. گام‌های پیاده‌سازی

۲-۲. پیش‌پردازش داده‌ها

(۳۰ نمره)

در گام نخست از طریق [لینک](#)^۱ داده‌شده تصاویر مربوط به خودروهای تویوتا را دانلود کنید. سپس لازم است که نمودار فراوانی آماری دیتاست داده‌شده را نمایش دهید و پیش از پیش‌پردازش‌های گفته‌شده ذیل، برچسب‌های^۲ خودروها را به مقادیر عددی تبدیل کنید.

- از بین کلاس‌های خودروهای این برند، به‌دلخواه ۱۰ مدل خودرو (۱۰ کلاس) را برای مراحل بعدی انتخاب کنید.
- طبق مقاله اندازه‌ی تصاویر را به ۲۲۴x۲۲۴ تغییر دهید.
- پس از تغییر سائز تصاویر به ابعاد خواسته‌شده، از طریق نرمال‌سازی مقادیر پیکسلی را بین ۰ و ۱ تبدیل کنید.
- دیتاست داده‌شده دارای تعداد محدودی تصویر برای هر کلاس است. با توجه به نمودار فراوانی که در مراحل قبل نمایش داده‌اید، در صورت بالانس نبودن داده‌ها ابتدا راهکاری برای بالانس کردن دیتاست ارائه دهید و ضمن ذکر مزیت روش اتخاذشده به باقی روش‌ها، آن را پیاده کنید. همچنین طبق توصیه‌ی مقاله، برای تعمیم بهتر مدل می‌توانید از داده‌افزایی^۳ استفاده کنید.

^۱ <https://www.kaggle.com/datasets/occultainsights/toyota-cars-over-۲۰k-labeled-images>

^۲ Labels

^۳ Data Augmentation

- در گام بعد به منظور شروع آموزش داده‌های موجود در ۱۰ کلاس انتخاب شده را به دودسته‌ی آموزشی و تست با نسبت ۸۰-۲۰ تقسیم کنید و ابعاد هر دسته را گزارش کنید.

۲-۳. استخراج ویژگی‌ها

(۳۰ نمره)

در این بخش ابتدا مدل پیش‌آموزش یافته‌ی VGG^۱ را بدون لایه‌های Fully Connected بارگذاری کنید. سپس به منظور استخراج ویژگی‌ها، تصاویر را از آخرین لایه‌ی کانولوشنی عبور دهید و برای هر تصویر ویژگی‌ها را ذخیره کنید. در گام بعد، ویژگی‌های استخراج شده را به بردارهای یک‌بعدی^۱ تبدیل کنید. مراحل گفته شده را برای مدل Alex Net هم انجام دهید و مجدداً ویژگی‌ها را ذخیره کنید.

۲-۴. آموزش و ارزیابی مدل

(۳۰ نمره)

۱. با ویژگی‌های استخراج شده در مرحله‌ی قبل دو مدل معرفی شده در مرحله‌ی قبل را آموزش دهید و سپس عملکرد مدل را بر روی داده‌های آزمایشی^۲، ارزیابی کنید. معیارهای ارزیابی درواقع همان معیارهای جدول ۴ مقاله خواهد بود که بایستی نحوه‌ی عملکرد هر معیار را توضیح دهید و مقادیر حاصل شده را گزارش دهید.
 ۲. اصلی‌ترین تفاوت بین Alex Net و VGG Net را فارغ از نتایج بیان کنید.
 ۳. در مطالعات پیشین مقاله برای مقایسه بهتر و جامع‌تر نتایج، یک مدل CNN ارائه کرده است که در این مرحله آن را پیاده‌سازی می‌کنید و متعاقباً نتایج را گزارش کنید.
 ۴. مقاله برای بهبود دقت مدل‌های قبلی، یک مدل پیشنهادی با ترکیب VGG و SVM ارائه کرده است؛ بدین صورت که SVM با ویژگی‌های استخراج شده در مرحله‌ی قبل به صورت خطی^۳ آموزش داده می‌شود و به عنوان یک classifier عمل می‌کند. برای این مدل نیز همانند مدل‌های قبل نتایج را روی داده‌های تست امتحان کنید و گزارش دهید.
- به طور کلی در این بخش لازم است نتایج را برای چهار مدل داخل مقاله گزارش دهید.

^۱ Flatten

^۲ Test

^۳ Linear

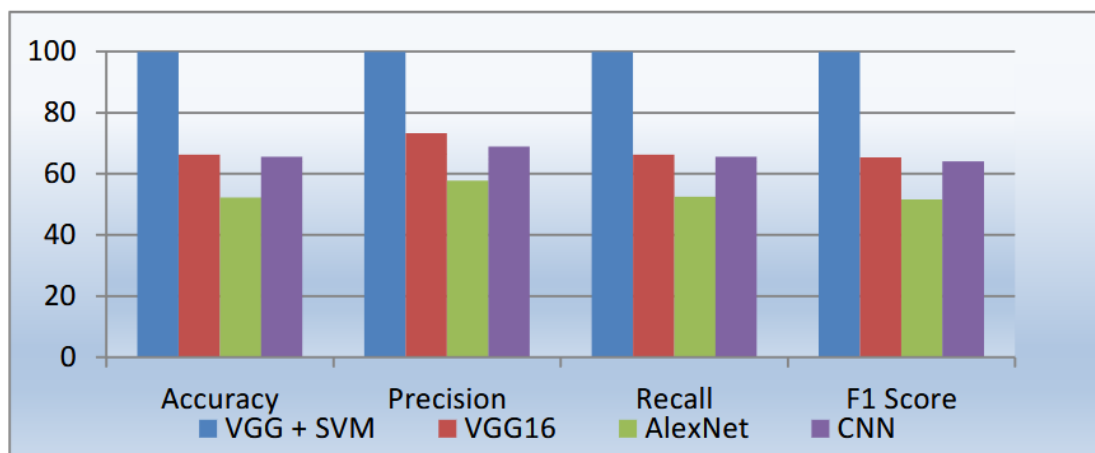
- VGG^{۱۶}
- VGG^{۱۶} + SVM
- Alex Net
- CNN

برای تحلیل عملکرد مسائل طبقه‌بندی یکی از بهترین روش‌ها، استفاده از ماتریس درهم‌ریختگی است. ابتدا نحوه‌ی تحلیل مقادیر عددی این ماتریس را مختصراً بیان کنید و این ماتریس را برای مدل‌های خواسته‌شده نمایش دهید.

۵-۲. تحلیل نتایج

(۱۰ نمره)

معیارهای عملکردی محاسبه‌شده از مرحله‌ی قبل را به‌صورت یک جدول نمایش دهید و مقایسه‌ای از نحوه‌ی عملکرد هر مدل ارائه دهید و به‌منظور تسهیل در مقایسه نمودارهایی مشابه شکل ۳ مقاله ارائه دهید.



شکل ۲. نمایش گرافیکی عملکرد به‌دست‌آمده توسط مدل‌های مختلف برای مجموعه داده توپوتا

از بین ۱۰ کلاس انتخاب‌شده بیان کنید کدام کلاس‌ها در هر مدل بهتر تشخیص داده‌شده‌اند و هر مدل در تشخیص کدام کلاس‌ها ضعیف‌تر عمل کرده است. حال با توجه به اینکه با مقایسه‌ی چهار مدل دید

کلی‌تری به دست آمده است برای بهتر عمل کردن مدل‌ها چه راهکارهایی پیشنهاد می‌کنید تا قدرت تشخیص مدل‌ها افزایش پیدا کند و به‌طور کلی مدل‌ها ارتقا پیدا کنند.

۶-۲. امتیازی

(۵ نمره)

۱. مدل پیشنهادی مقاله (VGG+SVM) را با کرنل‌های مختلف (برای مثال Linear, RBF^۱) امتحان کنید و تفاوت عملکرد آن‌ها را بیان کنید.
۲. در این مقاله قبل از شروع طبقه‌بندی^۲ استخراج ویژگی^۳ صورت گرفته است، تأثیر این عمل بر نتیجه نهایی را در مسائل طبقه‌بندی به‌ویژه این سؤال مطرح کنید.

^۱ Radial basis function

^۲ Classification

^۳ Feature extraction