به نام خدا

دانشگاه تهران

پردیس دانشکده‌های فنی

دانشکده برق و کامپیوتر

درس شبکه‌های عصبی و یادگیری عمیق

مدرس: دکتر احمد کلهر

تمرین شماره ۲

فروردین ماه ۱۴۰۵

فهرست

تشخیص چهره با استفاده از تابع زیان فاصله بین کلاسی **۳**

بخش اول: سوالات تئوری **۴**

زیربخش اول: تابع زیان طبقه بندی کلاسیک(۱۰ نمره) ٤

زیربخش دوم: توابع زیان و فاصله بین کلاسی (۱۰ نمره) ٤

زیربخش سوم: ارتباط فاصله زاویه ای به تشخیص چهره (۱۰ نمره) ٤

زیربخش چهارم: تابع زیان IAM (۱۰ نمره) ٥

زیربخش پنجم: دیگر توابع زیان دیگر (۱۰ نمره) ٥

بخش دوم: سوالات پیاده سازی **۶**

زیربخش اول: آماده سازی داده(۱۰ نمره) ٦

زیربخش دوم: مدل پایه CNN با Softmax (۱۰ نمره) ٧

زیربخش سوم: پیاده‌سازی Angular Margin Loss (۲۰ نمره) ٨

زیربخش چهارم: مقایسه نهایی، تحلیل و ارتباط با مقاله IAM (۲۰ نمره) ٩

نکات تحویل **۱۱**

تشخیص چهره با استفاده از تابع زیان فاصله بین کلاسی

افـزایش واریـانس بین‌کلاسـی و کـاهش فاصـله درون‌کلاسـی دو دغدغـه و هـدف اصـلی در حـوزه تشـخیص چهـره هسـتند. در ایـن تمـرین، مـا یـک تـابع هزینـه جدیـد بـا نـام Inter-class Angular Margin (IAM) Loss معرفـی مـی‌کنیم کـه هـدف آن افـزایش واریـانس بین‌کلاسـی اسـت. بـرخلاف روش‌هـای موجـود کـه حاشـیه بین‌کلاسـی را بـه یـک مقـدار ثابـت محـدود می‌کننـد، تـابع هزینـه IAM به‌صـورت تطبیقـی زوایـای کوچـک‌تر بـین کلاس‌هـا را بـا جریمـه بیشـتری مواجـه می‌کند و با موفقیـت حاشـیه زاویـه‌ای بـین کلاس‌هـا را افـزایش می‌دهـد؛ امـری کـه می‌توانـد بـه‌طور قابل توجهی قدرت تمایز ویژگی‌های چهره را بهبود بخشد.

تـابع هزینـه IAM می‌توانـد بـه‌راحتی بـه‌عنوان یـک جملـه منظم‌سـاز (regularization term) بـه تـابع هزینـه پرکـاربرد Softmax و نسـخه‌های جدیـدتر آن اضـافه شـود تـا عملکـرد آن‌هـا بـیش از پـیش بهبـود یابـد. همچنـین، بـازه مناسـب ابرپـارامتر منظم‌سـازی را از دیـدگاه پس‌انتشـار خطـا (backpropagation) تحلیـل و بررسـی مـی شـود. بـه‌منظور نمـایش عملـی، مـدل پیشـنهادی بـر روی یک دیتاست کوچک آموزش داده شده و مورد ارزیابی قرار گرفته است.

بخش اول: سوالات تئوری

زیربخش اول: تابع زیان طبقه بندی کلاسیک(۱۰ نمره)

تابع Softmax معمولی را بنویسید و توضیح دهید:

- چگونه طبقه‌بندی انجام می‌دهد؟
- چرا Softmax به تنهایی تضمین‌کننده فاصله زیاد بین کلاس‌ها نیست؟
- تابع زیان scale چیست و چگونه تعریف می شود؟
- چرا تابع زیان scaled معرفی میشود؟ چه مشکلی را نمی تواند حل کند؟

زیربخش دوم: توابع زیان و فاصله بین کلاسی (۱۰ نمره)

در مقاله بیان شده که IAM Loss باعث افزایش inter-class variance می‌شود. توضیح دهید:

- منظور از inter-class variance چیست؟
- چرا افزایش آن برای مسائل تشخیص چهره مهم است؟
- چرا فاصله زاویه‌ای (angular margin) مهم‌تر از فاصله اقلیدسی ساده است؟

زیربخش سوم: ارتباط فاصله زاویه ای به تشخیص چهره (۱۰ نمره)

اگر ویژگی‌ها نرمال‌سازی شوند (روی کره واحد باشند)، نشان دهید:

$$cos(\theta) = \frac{w^T x}{\| w \| \| x \|}$$

چگونه به فاصله زاویه‌ای مرتبط است و چرا این نمایش برای تشخیص چهره مناسب‌تر است؟

زیربخش چهارم: تابع زیان IAM (۱۰ نمره)

- تابع زیان معرفی شده در مقاله را توضیح دهید.
- توضیح دهید اگر دو کلاس زاویه کمی با هم داشته باشند، IAM loss چه تغییری در گرادیان ایجاد می‌کند؟
- نحوه آموزش با این تابع زیان را به همراه الگوریتم توضیح دهید.
- نحوه تعیین هایپرپارامتر ها برای این تابع زیان چگونه است؟ هر نوع مقدار از این هایپرپارامتر چه تاثیری دارد؟

زیربخش پنجم: دیگر توابع زیان دیگر (۱۰ نمره)

- تابع زیان sphare چیست و چه کاری را انجام میدهد؟
- تابع زیان cos چیست و چه کاری انجام می دهد؟
- تابع زیان arch چیست و چه اصلاحاتی را انجام میدهد؟
- تفاوت تابع زیان معرفی شده در مقاله با shpare-face و norm-face در چیست؟
- چرا مقاله تابع زیان IAM را و با چه هدفی معرفی کردند؟

بخش دوم: سوالات پیاده سازی

در این بخش قصد داریم یک مدل تشخیص چهره را پیاده سازی کرده و برای آموزش یک مدل CNN استفاده کنیم اما آنجایی که پیاده سازی بر روی تصاویر اشخاص واقعی نیازمند منابع بسیار پردازشی است، از ست داده toy که خود شما تولید خواهید نمود استفاده خواهیم کرد تا عملیات طبقه بندی را انجام دهیم. دقت کنید در تمامی این بخش، هم روی داده های تست و هم روی داده های آموزش، خواسته ها را گزارش کنید.

زیربخش اول: آماده سازی داده(۱۰ نمره)

برای داده ها، به ۱۰ کلاس (نمایان گر ۱۰ شخص متفاوت) که هر کلاس ۲۰۰۰ نمونه خواهد داشت نیاز است. مقادیر هر کلاس از توزیع نرمال با میانگین در بازه 5 و 5- و واریانس 1.5 تولید خواهند شد. ابعاد هر داده، ۱۴ در ۱۴ فرض می شود. شماره کلاس ها را از ۱ تا ۱۰ بزنید. داده ها را به ۸۰ درصد آموزش و ۲۰ درصد تست، با توزیع یکسان از کلاس ها (stratified) تقسیم کنید. می توانید اگر اضافه کردن لایه های batch normalization باعث افزایش دقت شود، آن را در میان لایه های کانولوشنی استفاده کنید. در بهینه ساز از weight decay با مقدار یک صدم نرخ یادگیری بهره ببرید.

- ماتریس کوواریانس ارتباط کلاس ها را رسم کنید.
- هیستوگرام توزیع مقادیر در یک نمونه در هر کلاس را نشان دهید.
- دقت طبقه بندی را گزارش کنید.
- با استفاده از طبقه بند میانگین، جدول confusion matrix را رسم کنید.

زیربخش دوم: مدل پایه CNN با SOFTMAX (۱۰ نمره)

در این زیربخش، یک مدل CNN استاندارد به عنوان baseline پیاده‌سازی می‌شود تا عملکرد آن با مدل مبتنی بر IAM loss در ادامه مقایسه شود.

۱. طراحی مدل CNN

- شبکه کانولوشنی ساده با ساختار پیشنهادی خود مقاله پیاده سازی کنید (MNIST network).
- تعداد پارامتر ها و ساختار را گزارش کنید.

۲. آموزش مدل

با تابع هزینه CrossEntropyLoss تعداد epoch حداقل ۳۰ و بهینه ساز adam و batch size دلخواه، مدل اولیه را آموزش دهید.

- تابع زیان برای آموزش و تست را در طول epoch ها رسم کنید.
- نمودار دقت روی آموزش و تست را در طول دوره آموزش رسم کنید.
- دقت نهایی را گزارش کنید.
- تاثیر تغییر batch size چیست؟

۳. ارزیابی مدل

برای داده‌های آموزش و تست، دقت، Confusion matrix را گزارش کنید.

- آیا overfitting رخ داده است؟
- آیا کلاس‌های نزدیک از نظر میانگین بیشتر با هم اشتباه می‌شوند؟

۴. تحلیل فضای ویژگی

خروجی لایه قبل از Softmax را استخراج کنید. با استفاده از PCA، داده‌ها را به ۲ بعد نگاشت کنید و نمودار پراکندگی کلاس‌ها را رسم کنید.

- آیا کلاس‌ها به‌خوبی از هم جدا شده‌اند؟
- چه میزان هم‌پوشانی وجود دارد؟

زیربخش سوم: پیاده‌سازی ANGULAR MARGIN LOSS (۲۰ نمره)

در این بخش، مدل CNN را بدون تغییر معماری، فقط با تغییر تابع هزینه بهبود می‌دهیم تا تأثیر margin زاویه‌ای بررسی شود.

۱. نرمال‌سازی ویژگی و وزن‌ها

خروجی embedding (لایه ۱۲۸ بعدی) را نرمال‌سازی L2 کنید. وزن‌های لایه نهایی را نیز نرمال‌سازی کنید. نشان دهید که خروجی logits برابر cos(θ) است. رابطه ریاضی را در گزارش بنویسید.

۲. تعریف ANGULAR MARGIN LOSS

یک margin زاویه‌ای به logits کلاس صحیح اضافه کنید:

$$\hat{y} = \cos(\theta_y) - m$$

که:

$$m \in [0.1, 0.5]$$

فقط برای کلاس صحیح اعمال شود. از CrossEntropy روی logits اصلاح‌شده استفاده کنید و پیاده سازی کنید.

۳. آموزش مدل با ANGULAR MARGIN LOSS

با تنظیمات و شبکه قسمت قبل با استفاده از تابع زیان جدید، آموزش را انجام دهید.

- تابع زیان برای آموزش و تست را در طول epoch ها رسم کنید.
- نمودار دقت روی آموزش و تست را در طول دوره آموزش رسم کنید.
- دقت نهایی را گزارش کنید.
- تاثیر تغییر batch size چیست؟

۴. تحلیل فضای EMBEDDING

دوباره PCA بگیرید و نمودار پراکندگی کلاس‌ها را رسم کنید

- فاصله بین مراکز کلاس‌ها را مقایسه کنید.
- با حالت آموزش با soft-max مقایسه کنید.

زیربخش چهارم: مقایسه نهایی، تحلیل و ارتباط با مقاله IAM (۲۰ نمره)

در این بخش تحلیل تفاوت CNN و CNN + Angular Margin و اتصال نتایج به ایده‌های مقاله IAM انجام می شود.

۱. مقایسه کمی

جدول زیر را برای train و test تکمیل کنید:

معیار	Softmax CNN	Angular Margin CNN
Accuracy		
Macro Precision		
Macro Recall		
Mean inter-class distance		
Mean intra-class distance		

فاصله‌ها را در فضای embedding محاسبه کنید.

۲. تحلیل زاویه‌ای کلاس‌ها

- مرکز هر کلاس در فضای embedding را محاسبه کنید.
- زاویه بین مراکز کلاس‌ها را رسم کنید.
- نشان دهید کدام مدل زاویه‌های بزرگ‌تری بین کلاس‌ها ایجاد می‌کند.

۳. تحلیل گرادیان و رفتار آموزشی (امتیازی)

به صورت مفهومی توضیح دهید:

- Angular Margin Loss چگونه نمونه‌های سخت (hard samples) را بیشتر جریمه می‌کند؟
- چرا این رفتار مشابه توضیح مقاله IAM است؟
- آیا راه حلی برای اینکه هم مزایا روش های مبتنی بر زاویه و هم مبتنی بر کلاس نهایی را ترکیب کنیم وجود دارد؟ چگونه؟

۴. تحلیلی نهایی

به سؤالات زیر پاسخ دهید:

- آیا بهبود Accuracy الزاماً زیاد بوده است؟ چرا؟
- چرا حتی با Accuracy مشابه، embedding مدل margin-based بهتر است؟
- اگر این داده ساده‌تر نبود، کدام روش برتری بیشتری داشت؟
- چرا مقاله IAM این روش را به‌عنوان regularization معرفی می‌کند؟
- مدل آموزش دیده برحسب margin در چه کارکردی ممکن است موثرتر و مفید تر باشد؟

نکات تحویل

- مهلت ارسال این تمرین تا پایان روز "دوشنبه ۱۴ اردیبهشت ماه" خواهد بود.
- این زمان قابل تمدید نیست و درصورت نیاز می‌توانید از گریس استفاده کنید.
- در نظر داشته باشید که حداکثر مهلت آپلود تمرین در سامانه تا ۷ روز پس مهلت تحویل است و پس از آن سامانه بسته خواهد شد.
- پیاده سازی با زبان برنامه نویسی پایتون باید باشد و کدهای شما باید قابل اجرا بوده و به همراه گزارش آپلود شوند.
- انجام این تمرین به صورت یک نفره می‌باشد.
- در صورت مشاهده هر گونه تشابه در گزارش کار یا کدهای پیاده‌سازی، این امر به منزله تقلب برای طرفین در نظر گرفته خواهد شد.
- استفاده از کدهای آماده بدون ذکر منبع و بدون تغییر به منزله تقلب خواهد بود و نمره تمرین شما **صفر** در نظر گرفته می‌شود.
- در صورت رعایت نکردن فرمت گزارش کار نمره گزارش به شما تعلق نخواهد گرفت.
- تحویل تمرین به صورت **دستنویس** قابل پذیرش نیست.
- تمامی تصاویر و جداول مورد استفاده در گزارش کار باید دارای توضیح (caption) و شماره باشند.
- بخش زیادی از نمره شما مربوط به گزارش کار و روند حل مسئله است.
- لطفا گزارش ، فایل کدها و سایر ضمائم مورد نیاز را با فرمت زیر در سامانه بارگذاری نمائید.
- HW2 _[Lastname]_[StudentNumber].zip
- در صورت وجود سوال و یا ابهام میتوانید از طریق رایانامه زیر با موضوع NNDL_HW2 با دستیاران آموزشی در ارتباط باشید:
  - دستیار تمرین

    ehsan.forootan@ut.ac.ir

با آرزوی سلامتی و موفقیت روزافزون