بسمه‌تعالی

شرح پروژه: طبقه‌بندی چندکلاسه‌ی گونه گل با استفاده از شبکه‌های عصبی کانولوشنی

مهلت انجام: ۱۴۰۵/۰۵/۱۴

هدف از این پروژه، آشنایی با مفهوم یادگیری عمیق و کاربرد آن در طبقه‌بندی تصاویر چندکلاسه با استفاده از شبکه‌های عصبی کانولوشنی است. در این پروژه از دیتاست Oxford Flowers102 شامل تصاویر گونه‌های مختلف گل استفاده می‌شود. این پروژه شامل مراحل پیش‌پردازش تصویر، طراحی و آموزش شبکه عصبی کانولوشنی و مقایسه دو رویکرد آموزش مدل است.

شکل ۱ نمونه تصاویر مجموعه داده مورد استفاده (برگرفته از یک مقاله[1])

پیش از انجام پروژه، درنظرگیری نکات زیر الزامی است:

- استفاده از زبان برنامه‌نویسی پایتون و فریمورک پایتورچ توصیه می‌شود.
- داده‌ی مورد استفاده در این پروژه Oxford Flowers102[2] (قابل‌دسترس در کتابخانه تورچ‌ویژن) است که شامل ۸۱۸۹ تصویر رنگی از ۱۰۲ گونه مختلف گل می‌باشد. این دیتاست به صورت پیش‌فرض دارای سه بخش آموزش، اعتبارسنجی و آزمون است.

[1] https://www.researchgate.net/publication/357066811_Deep_Transfer_Learning_for_Biology_Cross-Domain_Image_Classification

[2] https://www.kaggle.com/deadskull7/fer2013

مراحل انجام این پروژه به شرح زیر است.

مرحله ۱: آماده‌سازی داده‌ها و پیش‌پردازش

- استفاده از دیتاست FER2013 که شامل تصاویر رنگی با اندازه‌های متفاوت با ۷ کلاس احساس مختلف نظیر خشم[۳]، تنفر[۴]، ترس[۵]، شادی[۶]، غم[۷]، شگفتی[۸] و حالت خنثی[۹] است.
- پیش‌پردازش[۱۰] شامل موارد زیر است.
  - نرمال‌سازی: اعمال نرمال‌سازی برای مقیاس‌بندی مقدار پیکسل‌ها (مثلاً در بازه‌ی [۰,۱] یا با میانگین و واریانس استاندارد) به انضمام تغییر ابعاد تصاویر به ۶۴*۶۴ پیکسل برای تمامی آزمایش‌ها الزامی است.
- انتخاب حداقل دو روش آگمنتیشن[۱۱] به منظور داده‌افزایی و بررسی اثر آن‌ها بر عملکرد مدل الزامی است.
- تقسیم داده‌ها به بخش‌های آموزش[۱۲]، اعتبارسنجی[۱۳] و تست[۱۴] (که این مهم برای مجموعه داده مذکور در این پروژه به صورت پیش‌فرض محقق می‌باشد). شایان ذکر است، تنظیم اَبَرپارامترهای هر روش به کمک داده‌های اعتبارسنجی و مقایسه نهایی میان تمام آزمایش‌های انجام‌شده براساس داده تست صورت گیرد.

[3] Angry
[4] Disgust
[5] Fear
[6] Happy
[7] Sad
[8] Surprise
[9] Neutral
[10] Pre-process
[11] Augmentation
[12] Train set
[13] Validation set
[14] Test set

مرحله ۲: آموزش مدل از صفر[۱۵]

- طراحی و پیاده‌سازی یک شبکه عصبی کانولوشنی با حداقل ۳ لایه کانولوشن، هر لایه شامل:
  - لایه کانولوشن
  - نرمال‌سازی بسته[۱۶]
  - فعال‌سازی ReLU
  - MaxPooling
- لایه‌های تماماً متصل[۱۷] در انتها برای خروجی ۷ کلاس
- استفاده از کراس‌انتروپی[۱۸] به عنوان تابع هزینه[۱۹]
- انتخاب بهینه‌ساز آدام[۲۰] با نرخ یادگیری[۲۱] مناسب
- آموزش مدل روی داده‌های پیش‌پردازش‌شده و فاقد آن به تفکیک
- گزارش دقت نهایی مدل روی داده تست در تمام ارزیابی‌های انجام‌شده و گزارش جدولی کامل از تمام آزمایش‌های انجام‌شده، هر ردیف متناظر با یک آزمایش باشد.

[15] Training from scratch
[16] Batch normalization
[17] Fully connected layers
[18] Cross entropy
[19] Loss function
[20] Adam optimizer
[21] Learning rate

مرحله ۳: آموزش مدل با استفاده از انتقال یادگیری[۲۲]

- استفاده از مدل پیش‌آموزش‌دیده[۲۳] هم‌چون رزنت[۲۴]۱۸ یا موبایل‌نت۲[۲۵] (موجود در کتابخانه تورچ‌ویژن[۲۶]) – درج توضیحات پیرامون علت انتخاب مدل الزامی است.
- فریز کردن[۲۷] لایه‌های اولیه‌ی مدل از پیش‌آموزش‌دیده (بَک‌بُن[۲۸]: غیر از لایه‌های تماماً متصل انتهایی)
- تغییر لایه‌های تماماً متصل انتهایی به منظور تطبیق با ۱۰۲ کلاس متناظر با برچسب‌های ارائه‌شده در داده
- آموزش فقط لایه‌های انتهایی جدید بر روی دیتاست Flowers102 و فریزنمودن بخش بَک‌بُن
- اعمال پیش‌پردازش‌ها مشابه نکات ذکرشده در مرحله ۱
- مقایسه دقت مدل انتقالی با یک مدل آموزش داده شده از صفر (با درنظرگیری معماری کاملاً مشابه براساس معماری منتخب از کتابخانه‌ی تورچ‌ویژن)

[22] Transfer learning
[23] Pre-trained model
[24] ResNet18
[25] MobileNetV2
[26] Torchvision
[27] Freeze
[28] Backbone

مرحله ۴: تحلیل و گزارش‌گیری

به منظور رصد بهتر آزمایش‌ها و اقدامات انجام‌شده، نیاز است متناسب با هر آزمایش شرح گزارشی از جزئیات و مشاهدات به عمل آمده درج گردد. در انتها به منظور مقایسه بهتر نتایج کسب‌شده، جدول مقایسه‌ای از نتیجه‌ی هر آزمایش و تحلیل مربوطه ارائه گردد. اهم تحلیل‌های مورد انتظار در هر بخش عبارتند از:

- رسم نمودارهای تغییرات تابع هزینه[۲۹] و دقت[۳۰] در طول آموزش برای هر دو روش به ازای داده‌ی آموزش و اعتبارسنجی در قالب یک نمودار مشترک
- رسم ماتریس کانفیوژن[۳۱] برای هر آزمایش بر روی داده تست (درنظرگیری برچسب‌گذاری درایه‌های ماتریس الزامی است)
- بررسی اثر پیش‌پردازش‌ها و داده‌افزایی
- تحلیل پیچیدگی محاسباتی، زمان آموزش و دقت

[29] Loss function
[30] Accuracy
[31] Confusion matrix