

شبکه عصبی تمرین چهارم

سحر محمدی ۴۰۲۱۳۹۰۱۰۹

هدف این تمرین پیاده سازی transfer learning با استفاده از مدل از پیش آموزش دیده شده face net و fine tune کردن آن برای دیتاستی شامل تصاویر هنرپیشه‌هاست. همچنین در مرحله بعد پیاده سازی یک face verifier است.

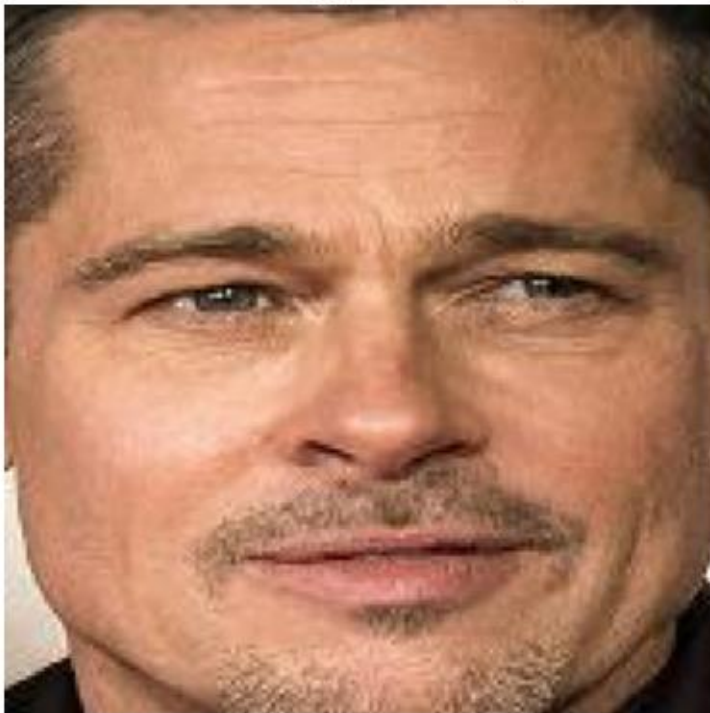
شرح پیاده سازی

ابتدا مدل InceptionResnetV1 را لود کردیم.

سپس با استفاده از pythorch دیتاست را لود کردیم و به تنسور پای تورچ تبدیل کردیم. ۸۰ درصد برای آموزش و ۲۰ درصد برای تست قرار دادیم.

یک تابع برای نمایش تصویر و یک تابع برای mapping لیبل‌ها تعریف شد. سپس چند نمونه از تصاویر را به همراه لیبل آن‌ها نمایش دادیم.

Label: (3, 'Brad Pitt')



Label: (0, 'Alexandra Daddario')

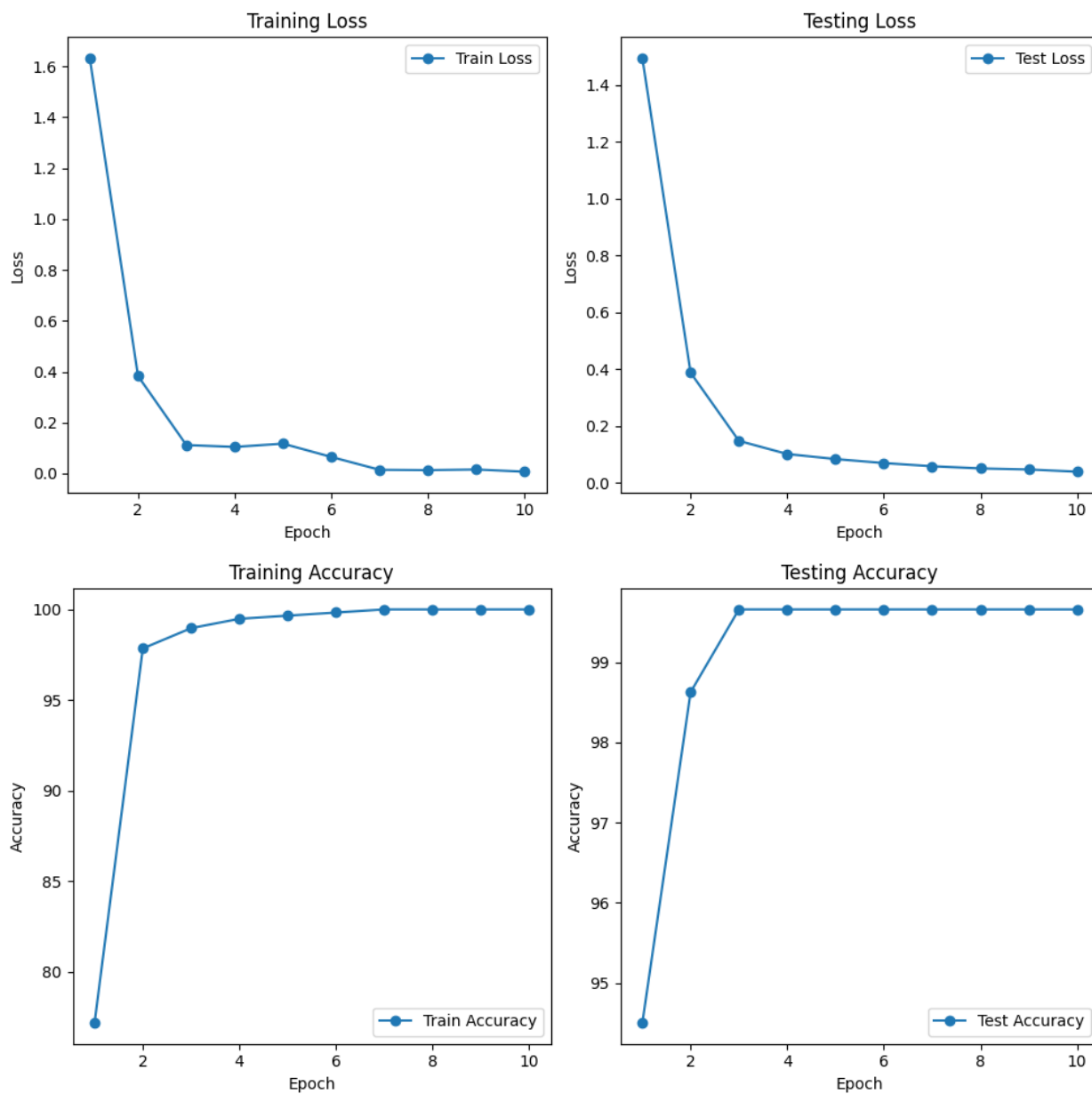


در مرحله بعد، یک شبکه MLP ساختیم که به انتهای مدل اضافه شود با ورودی به اندازه بردار embedding یعنی ۵۱۲ و یک لایه مخفی با ۲۵۶ نرون و خروجی به اندازه کلاس‌های دیتاست یعنی ۱۴.

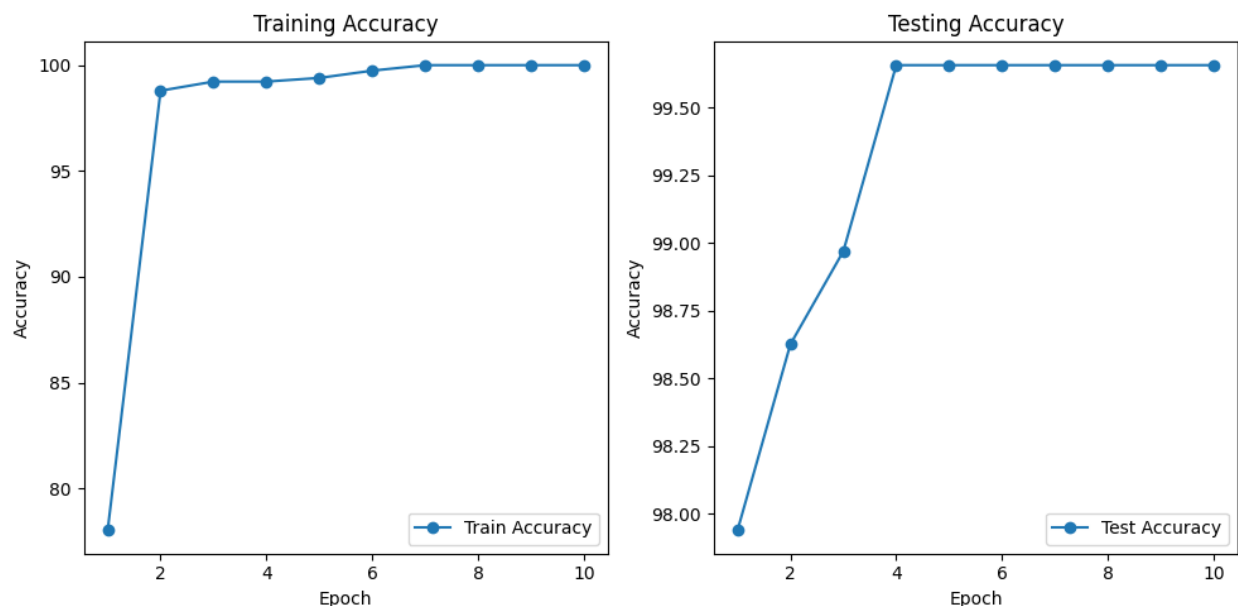
برای بخش face verifier ابتدا یک بار داده‌های تست را وارد مدل کردیم که خروجی embedding به دست بیاید. همان طور که خواسته شده بود یک بخش از داده‌های تست به عنوان مشخص در نظر گرفته شد، از داده ۰ تا ۲۸ که از همه کلاس‌ها موجود باشد. و بقیه داده‌ها هم غیر مشخص.

آزمایشات

۱. با بچ سائز ۳۲ و ۱۰ اپوک نتیجه چنین شد:



و دقت ۹۹.۶۶ درصد در داده تست دست یافتیم. و می‌بینیم که از اپوک ۳ حدوداً به این دقت رسیده‌ایم.
با فریز کردن لایه‌های 'conv2d_4a_3x3', 'conv2d_4b_3x3' در اپوک ۴ به دقت بالا رسیدیم.



۲. Face verifier

ابتدا برای بررسی تاثیر معیار شباهت **cosine similarity** و فاصله اقلیدوسی، یک بار همه داده‌ها تست را با داده بعد از خودش سنجیدیم، اگر دو تصویر مشابه تشخیص داده شد و لیبل آن‌ها هم یکسان بود، یک عدد به **count** اضافه می‌شود و همچنین برعکس اگر دو تصویر غیر مشابه تشخیص داده شد و لیبل آن‌ها هم یکی نبود به **count** اضافه می‌شود که در نهایت بر تعداد داده‌های تست تقسیم شده تا دقت **verifier** به دست بیاید.

ابتدا بررسی **cosine similarity** که برای یافتن ترشلد مناسب که اگر از آن مقدار بیشتر بود نتیجه اینکه عضو یک کلاس هستند داده شود، ابتدا چند مورد که در یک کلاس بودند را میزان شباهت را سنجیدیم و چند مورد را امتحان کردیم.

با ترشلد ۰.۹، دقت ۰.۹۵۸ داشتیم. با ترشلد ۰.۳، دقت ۰.۹۲ و با ترشلد ۰.۶، دقت ۰.۹۷۹۳.

همچنین برای بررسی معیار فاصله اقلیدوسی نیاز است که یک ترشلد پیدا کنیم که اگر فاصله از آن میزان کمتر بود نتیجه بدهد که در یک کلاس هستند.

ابتدا با ترشلد ۱.۵ بررسی شد که دقت ۰.۱۱ به دست آمد. با ترشلد ۰.۴، دقت ۰.۹۵ و با ترشلد ۱.۰۵، دقت ۰.۹۸۲۷ که بیشترین است.

در آخر هم برای همه داده‌های تست غیر مشخص یعنی از ۲۸ تا آخر، فاصله هر داده با ۲۸ داده اول به دست آمد که بررسی شود فاصله تا هر کدام کمتر بود یعنی در آن کلاس قرار دارد. برای مشاهده روند کار چند تصویر را نمایش داده و لیبل واقعی و پیشنهادی تطبیق دهنده را نمایش دادیم.

در اینجا ۱۳ اول مربوط به لیبل واقعی و ۱۳ دوم لیبل پیشنهادی است.

13 belongs to 13

Label: (13, 'Zac Efron')



Zac Efron

8 belongs to 8

Label: (8, 'Jessica Alba')



Jessica Alba
4 belongs to 4

Label: (4, 'Claire Holt')



Claire Holt
6 belongs to 12

Label: (6, 'Hrithik Roshan')



Vijay Deverakonda

که مشاهده می‌شود این تصویری که لیبل پیشنهادی اشتباه تشخیص داده شده، هر دو کلاس مربوط به افراد هندی هستند. یا در دو مورد دیگر تصاویر که هر دو زن بوده‌اند تطبیق دهنده اشتباه کرده است. اما در مجموع بسیار خوب عمل کرده.