

**گزارش پروژه‌ی پایانی مبانی بینایی کامپیوتر**

**استاد نظری**

**حمید الهی نراقی – ۹۷۲۴۳۰۰۵**

**خواسته‌ی پروژه:**

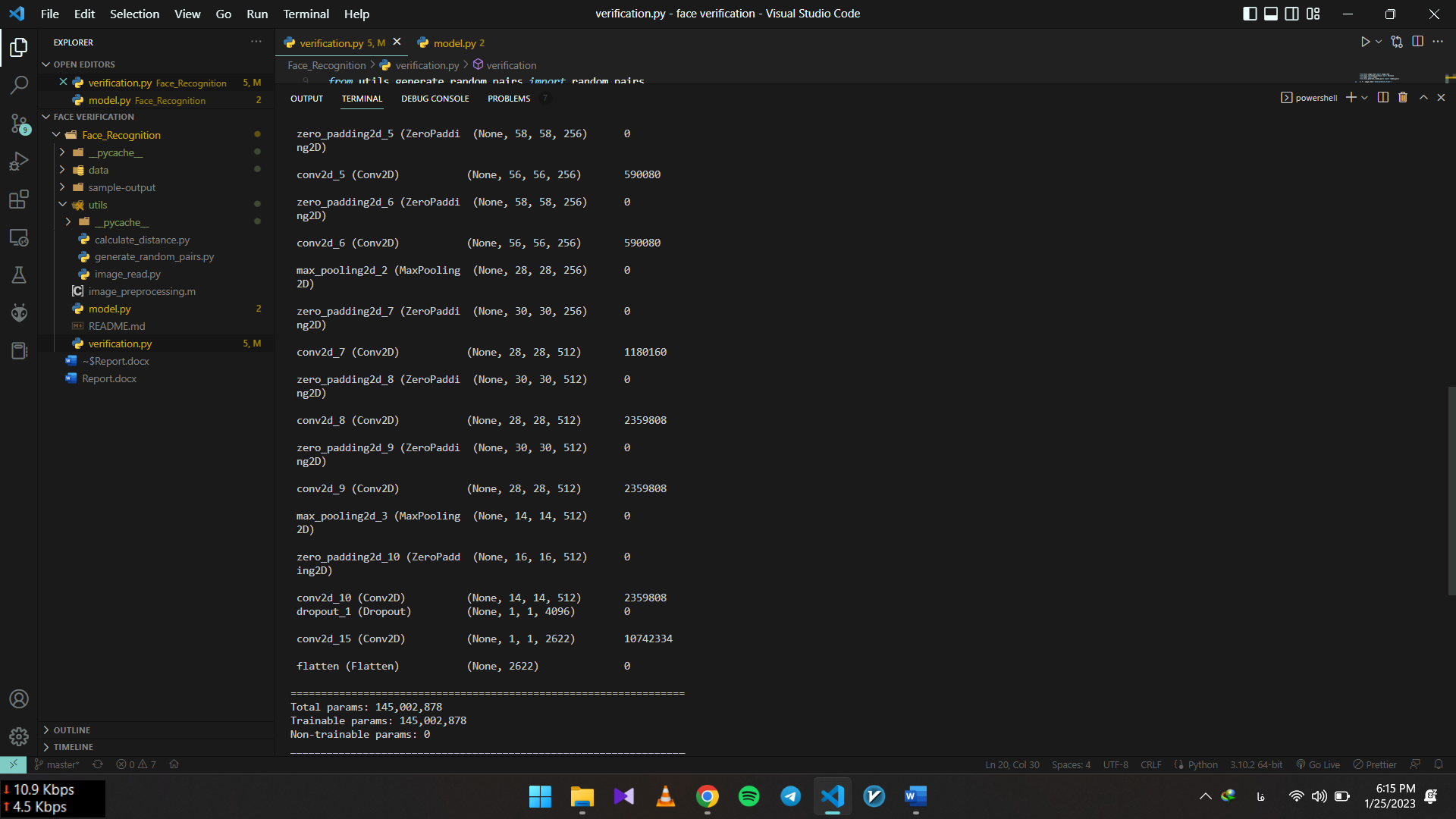
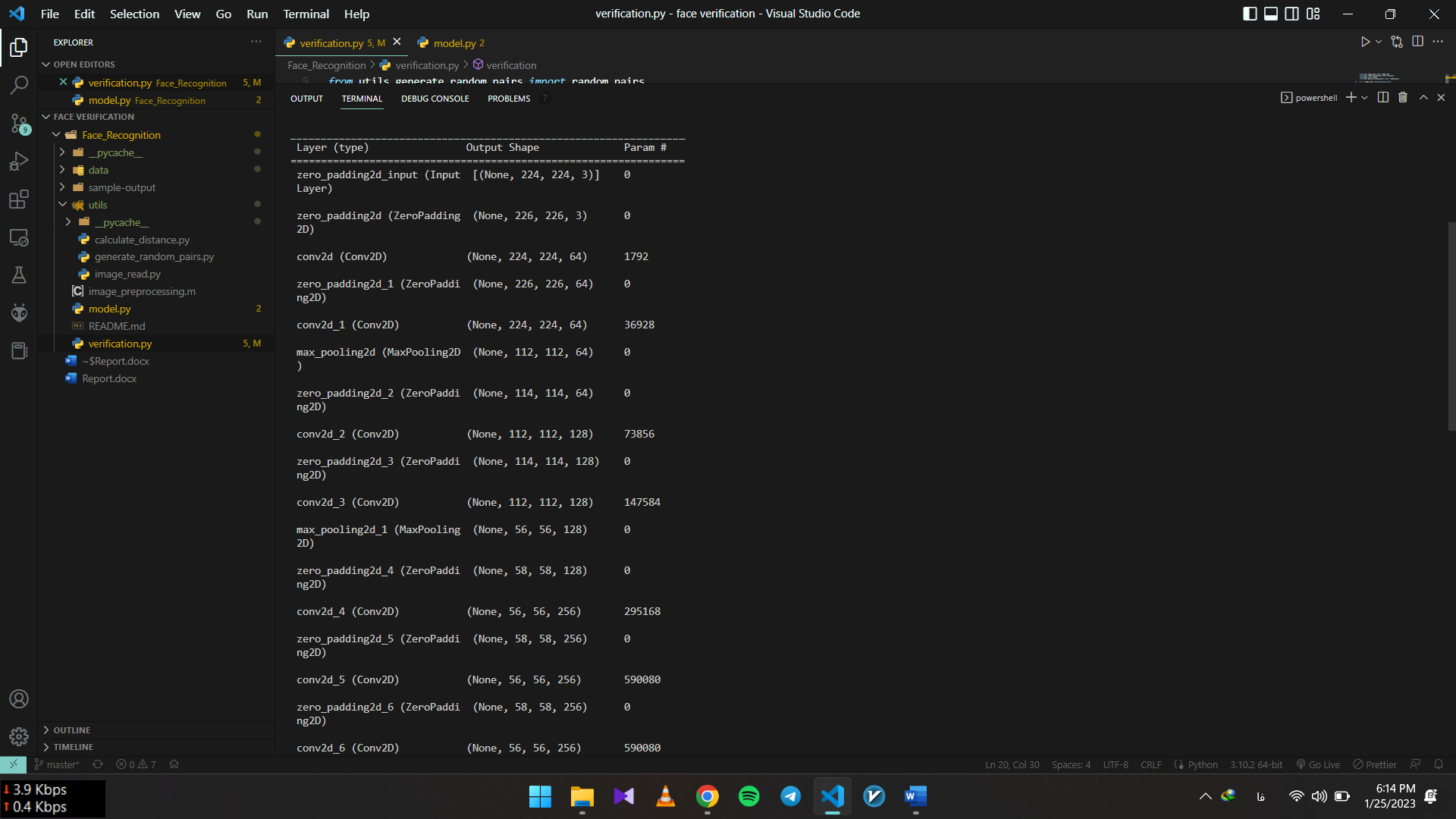
در این پروژه می‌بایست سیستمی پیاده‌سازی می‌شد که دو عکس را به عنوان ورودی بگیرد و فارغ از تشخیص هویت افراد٬ تشخیص دهد که آیا دو تصویر مربوط به یک نفر هست یا خیر.

**راهبرد:**

برای ساخت چنین سیستمی از یک مدل از پیش آموزش داده شده (pre-trained) به نام vggFace استفاده کردم. در این شبکه‌ی عصبی که از نوع پیچشی (CNN) است به ازای هر عکس٬ یک بردار متناظر در فضای nبعدی تولید می‌کند که با مقایسه فاصله‌ی این نقاط می‌توان تشخیص داد که تا چه حد شبیه به یکدیگر هستند. به این شکل که هرچه فاصله دو نقطه در فضای nبعدی کمتر و نزدیک‌تر به صفر باشد٬ تصاویر متناظر با این نقاط نیز شبیه‌تر به هم هستند.

**پیاده‌سازی:**

با استفاده از زبان پایتون و با کمک کتابخانه‌ی keras یک مدل ترتیبی طبق معماری مدل معروف vggFace ساختم و لایه‌های کانوولوشنی و max-pooling را به ترتیب در نظر گرفتم. معماری کلی مدل به شرح زیر است:



در تمامی لایه‌ها از تابع فعال‌سازی relu و در لایه‌ی آخر پس ار لایه‌ی faltten از تابع فعالسازی softmax استفاده شده است. در نهایت وزن‌های مدل را به کمک دستور زیر در مدل ساخته شده load کردم:

 model.load\_weights('data/vgg\_face\_weights.h5')

**سنجش:**

برای سنجش صحت مدل٬ توابع کمکی random\_pairs و Distance را به ترتیب برای ساخت جفت‌ عکس‌های ورودی و محاسبه فاصله‌ی اقلیدسی دو نقطه نوشتم. در نهایت در کلاس verification اقدام به ساخت شئ مدل کردم و ۱۰ جفت عکس‌ را برای دادن ورودی به مدل به صورت رندوم انتخاب کردم. با هر جفت عکسی که به مدل داده می‌شود٬ می‌توان فاصله‌ی بین نقاط متناظر آن‌ها را محاسبه کرد و میزان شباهت دو چهره را محاسبه کرد. مرز 0.55 به عنوان threshold در نظر گرفته شده و فاصله‌ی کمتر از آن به معنای شباهت دو چهره است. در ادامه تشخیص مدل برای شباهت یا عدم شباعت برخی از جفت‌ عکس‌های ورودی آورده شده است:

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |