بسم الله الرحمن الرحیم

دانشگاه علم و صنعت ایران

پروژه پایانی درس مبانی بینایی کامپیوتر

هدف این پروژه، توسعه یک الگوریتم Anti-spoofing برای تشخیص زنده بودن[۱] چهره در ویدیو می‌باشد. که در آن یک فریم از ویدیو را گرفته و تشخیص دهید تصویر زنده شخص می‌باشد یا نه. به‌طور مثال تصویر دارای گریم یا پرینت شده از وی است یا تصویر از یک دستگاه دیگر درحال پخش است یا نه.

این الگوریتم‌ها در دو دسته‌ی کلی زیر قرار می‌گیرند:

- روش‌های مبتنی بر استخراج ویژگی (فرکانس، بافت تصویر، اطلاعات عمقی، علایم حیاتی شخص)
- روش‌های مبتنی بر یادگیری عمیق (مبتنی بر دامنه، یادگیری انتقالی، ادغام ویژگی‌ها)

برای انجام این پروژه، ابتدا شما باید تعدادی ویدیو دارای چهره (حداقل ۱۰ مورد) در شرایط متفاوت تهیه کنید. به‌طور مثال شامل موارد زیر:

حالت‌های مختلف غیر زنده بودن، حالت‌های متفاوت چهره، سن، جنسیت، جهت زاویه سر[۲]، وضعیت نورپردازی[۳]، حجاب اسلامی و ...

سپس می‌باست برای هر دسته فوق یک مدل به انتخاب خود را آموزش دهید.

برای آموزش نیز از مجموعه داده‌های عمومی مثل موارد زیر می‌توانید استفاده کنید:

CASIA-FASD [link] [paper]

CASIA-SURF [link] [paper]

OULU-NPU [link] [paper]

برای ارزیابی بر روی مجموعه‌داده‌ای که خودتان تهیه کردید می‌باست موارد زیر را نیز تست بفرمایید:

- کل تصویر را به مدل دهید.
- ناحیه چهره را در تصویر برش بزنید و آن را به مدل دهید.
- از تصویر ورودی در حوزه فرکانس ویژگی دربیاورید و آن‌ها را به شبکه دهید.

[۱] liveness
[۲] Pose
[۳] Illumination

در انجام این پروژه مجاز به استفاده از تمام ابزارهای خوانده شده در درس هستید. همچنین، در صورت استفاده از ابزارهایی که در درس مطرح نشده است لازم است به جزئیات آن مسلط باشید. اجرای پروژه در قالب گروه‌های دو نفره خواهد بود. لطفا اسامی اعضای تیم خود را به ایمیل mrmohammadi@iust.ac.ir ارسال بفرمائید.

خروجی کار شما عبارت است از:

- یک گزارش کامل از تمام کارهایی که در این پروژه انجام داده‌اید و نتایجی که بدست آورده‌اید و ارزیابی میکنید.
- یک کد کامل که بتواند بر روی سیستم اجرا شود و نتایج را بر روی یک مجموعه داده جدید ذخیره کند.
- کد شما باید دارای یک فایل با نام main.ipynb باشد (که قابل پیاده‌سازی بر روی کولب باشد) و با اجرای آن، تمام مراحل مورد نیاز انجام شوند.
- برای مجموعه داده جدید (شامل مواردی که خودتان جمع‌آوری کرده‌اید)، یک فایل متنی با نام dataset.txt در کنار فایل main.ipynb قرار داده خواهد شد که در هر سطر آن نام یک ویدیو نوشته شده است، مشابه با شکل زیر:

```
../dataset/spoof1.mp4
../dataset/not-spoof1.mp4
```

- برنامه می‌بایست ویدیوهای موجود در این آدرس را خوانده و عملیات تشخیص زنده بودن در یک فریم از ویدیوها را انجام دهد.
- در نهایت خروجی کد شما دو فایل متنی مشابه با نام predictions_feature.csv و predictions_deep.csv برای دو روش فوق خواهد بود که شامل ۴ ستون به شرح زیر است:
  1) نام فایل ویدیو
  2) احتمال زنده بودن در فریم اصلی
  3) احتمال زنده بودن در تصویر چهره برش زده شده
  4) احتمال زنده بودن با استفاده از ویژگی‌های حوزه فرکانس

به‌طور مثال:

filename	liveness_score	liveness_score_crop	liveness_score_frequency
spoof1.mp4	0.15	0.1	0.2
not-spoof1.mp4	0.8	0.7	0.75

نکاتی که باید رعایت فرمایید:

- به‌ازای هر ویدیو باید یک سطر در فایل خروجی وجود داشته باشد.

- برنامه باید مدیریت خطاهای احتمالی را داشته باشد. در صورتی که کد با هر گونه خطایی مواجه شود و خروجی فایل CSV تولید نشود، در ارزیابی نمره کسر خواهد شد.
- کد شما با ویدیوهایی که سایر تیم‌ها جمع‌آوری کرده‌اند اجرا می گردد.
- نتایج کار شما از لحاظ کمی (دقت و سرعت) و کیفی با نتایج دیگران مقایسه خواهد شد.
- برای ارزیابی عملکرد الگوریتم از معیار دقت [۴] استفاده خواهد شد.
- در نوشتن گزارش توجه داشته باشید که به هر مرجعی که استفاده می‌کنید (چه از لحاظ تئوری، چه از لحاظ کدنویسی) به دقت ارجاع بدهید.
- در گزارش نتایج بدست آمده در آموزش مدل و تست را مقایسه و تحلیل کنید.
- انجام پروژه حتما در هر گروه به صورت جداگانه انجام شود.
- با توجه به محدودیت‌های زمانی و سخت‌افزاری که در اجرای ایده‌های خود دارید بهتر است که پروژه را زودتر شروع کرده تا بتوانید ایده‌های خود را کامل تست کنید.

[۴] Acuuracy

بخش امتیازی:

برای انجام این قسمت (بعد از انجام بخش اول)، ابتدا باید تعدادی تصویر از کارت ملی افراد شرکت کننده در بخش قبل در زوایای مختلف تهیه نمایید و با تکنیک‌های پردازش تصویر محدوده مربوط به چهره را برش بزنید و با ویدیوهایی که قبلا تهیه کرده‌اید، احراز هویت با چهره[۵] انجام دهید. (اطلاعات فریم‌های مختلف را تجمیع کنید و در نهایت بگویید این شخص کیست). برای این منظور از دو مدل facenet۵۱۲ و arcface در فریمورک deepface استفاده کنید و آنها را در شرایط مختلف مقایسه کنید.

برای این منظور با رعایت نکات بخش قبلی می‌توانید از دستور زیر استفاده کنید.

```
result = DeepFace.verify(
  img۱_path = "person۱-۱.jpg",
  img۲_path = "۱.jpg",
  detector_backend = backends[۰],
  model_name = models[۰],
  distance_metric = metrics[۱]

)
```

برای تشخیص، تطبیق و محاسبه فاصله در این منبع می‌توان از جایگشت‌های مختلف فوق استفاده نمود. که شما مجاز به انتخاب موارد پررنگ شده هستید.

```
models = [
  "VGG-Face",
  "Facenet",
  "Facenet۵۱۲",
  "OpenFace",
  "DeepFace",
  "DeepID",
  "ArcFace",
  "Dlib",
  "SFace",
  "GhostFaceNet",
]
```

```
backends = [
  'opencv',
  'ssd',
```

[۵] Facial Authentication

```
 'dlib',
 'mtcnn',
 'fastmtcnn',
 'retinaface',
 'mediapipe',
 'yolov۸',
 'yunet',
 'centerface',
]
```

```
metrics = ["cosine", "euclidean", "euclidean_l۲"]
```

در نهایت خروجی کد شما (با نام deepface.ipynb) یک فایل متنی مشابه با نام results.csv که شامل ۴ ستون به شرح زیر است:

- نام عکس کارت ملی
- نام ویدیو
- میزان شباهت در مدل arcface
- میزان شباهت در مدل facenet۵۱۲

به‌طور مثال:

A	B	C	D
id-card-name	file_name	similarity_arcface	similarity_facenet512
person1-1.jpg	1.mp4	0.8	0.85
person1-2.jpg	2.mp4	0.15	0.2

لازم به ذکر است که تصاویر کارت ملی‌های مختلف با ویدیوهای افراد متفاوت ارزیابی خواهد شد و میزان دقت و نرخ خطای مثبت[۶] در نتایج به دست امده مهم هستند. در گزارش نیز حتما این ارزیابی را داشته باشید و نتایج به دست آمده در هر مرحله را مقایسه و تحلیل کنید و دلیل انتخاب مدل backend را بفرمایید. همچنین، نکات قوت و ضعف دو مدل استفاده شده را در حالت‌های مختلف تصاویر ارزیابی کنید.

موفق باشید.

[۶] FP rate