



به نام خدا

آزمایشگاه یادگیری و بینایی ماشین

آزمایش نهم

تشخیص رفتار حرکتی انسان و پیاده سازی به صورت آنلاین

هدف:

هدف از این پژوهش، تشخیص رفتار حرکتی انسان با استفاده از دیتاست KTH و با روش‌های متداول بینایی کامپیوتر است. در انتها، الگوریتم توسعه داده شده باید بتواند به صورت برخط رفتار حرکتی شما را تشخیص دهد.

روش:

به این منظور ابتدا باید دیتاست مورد نظر را به دو دسته آموزش و آزمون تقسیم کنید. با استفاده از ویدئوهای آموزش، حرکت فرد را تشخیص داده و ویژگی‌های منتخب را استخراج کنید. سپس یک طبقه‌بند را انتخاب کنید و آن را آموزش دهید. حال در ویدئوهای تست رفتار حرکتی فرد را تشخیص داده و چاپ کنید.

1. پایگاه داده‌ی KTH

دیتاست شامل ۶۰۰ ویدئو با نرخ 25 فریم بر ثانیه از انسان‌ها می‌باشد که شش رفتار حرکتی مشت‌زنی^۱، تکان دادن دست^۲، دست زدن^۳، دویدن نرم^۴، دویدن^۵، قدم زدن^۶ را انجام می‌دهند.

دیتاست را برای سه حرکت تکان دادن دست، دست زدن و دویدن از اینجا [دانلود](#) کنید.

2. محاسبه شار نوری و ویژگی‌ها

1.2 با در دست داشتن هر فریم از ویدئو و فریم قبل از آن، میتوان ویژگی‌های مرتبط با شار نوری را به دست آورد. در نتیجه با داشتن یک ویدئو با تعداد N فریم، یک مجموعه N-1 تایی از ویژگی‌های شار نوری به دست خواهد آمد. این ویژگی‌ها را برای تمامی ویدئوها محاسبه و ذخیره کنید.

نکته: برای کاهش هزینه محاسباتی می‌توانید ویدئوها را downsample کنید و تاثیر کاهش نرخ فریم برثانیه را روی شارنوری بررسی کنید و نرخ مناسب را انتخاب کنید.

2.2 در اینجا میخواهیم از تکنیک کیسه لغات^۷ استفاده کنیم که یکی از محبوب‌ترین روش‌ها در ارائه‌ی ویژگی برای تشخیص رفتار انسان است. این تکنیک به همراه روش‌های دیگر استخراج ویژگی به این صورت استفاده میشود که ابتدا ویژگی‌های اولیه از نمونه‌ها استخراج میشود. سپس به کمک تکنیک‌های خوشه بندی ویژگی‌های استخراج شده به خوشه‌های مختلف تقسیم می‌شوند. و در انتها برای به دست آوردن بردار ویژگی یک نمونه، تعداد رخداد ویژگی‌های آن در هر خوشه شمرده میشود (یعنی برداری با بعد تعداد خوشه‌ها به دست می‌آید). در این پروژه نمونه‌ها همان ویدئوها و ویژگی‌های اولیه همان شار نوری محاسبه شده در هر دو فریم متوالی است. بنابراین برای یافتن بردار ویژگی مربوط به هر ویدئو باید تعداد رخداد فلوهای مربوطه در هر خوشه مشخص

¹ Boxing

² Handwaving

³ Handclapping

⁴ Jogging

⁵ Running

⁶ Walking

⁷ Bag of Words

شود. به عبارت دیگر رشته BoW را می‌توانید به صورت یک هیستوگرام تصور کنید که فرکانس رخداد هر خوشه را در هر ویدئو نشان می‌دهد.

داده‌ها را به دو دسته آموزش و آزمون تقسیم کنید (20٪ به عنوان داده آزمون). حال برای داده‌های آموزش مرحله‌ی زیر را اجرا کنید.

3. آموزش طبقه بند

برای ویژگی‌های به دست آمده در مرحله 2.1 رشته BoW را با خوشه‌بند K-means به دست آورید. تعداد خوشه‌ها را در سه حالت 180، 200 و 220 در نظر بگیرید و عملکرد نهایی را بر روی داده‌های آزمون مقایسه کنید و بهترین تعداد خوشه را گزارش کنید.

در این مرحله با طبقه بند SVM که پیش‌تر با آن آشنا شدید و با آزمودن کرنل‌های مختلف برای آن، مدل خود را بر مجموعه دادگان آموزش، آموزش دهید. عملکرد نهایی را بر روی دادگان آزمون برای هر کدام از کرنل‌ها مقایسه و بهترین مدل را گزارش دهید. (مرحله ۴)

نکته: توجه داشته باشید که برای ارزیابی مدل خود باید دقت بر دادگان آزمون را گزارش کنید که در فرآیند خوشه بندی، تشکیل BoW و آموزش SVM وارد نشده بودند.

4. مرحله آزمون

حال دقت هر مدل را بر روی دادگان آزمون گزارش کنید. بهترین مدل را معرفی کنید (یعنی بهترین تعداد خوشه‌ها برای BoW و کرنل مورد انتخاب شده برای SVM).

نکته: خوشه‌بندی و تعیین دسته‌هایی که در ساخت رشته‌ی BoW استفاده می‌شوند، باید با استفاده از دادگان آموزش به دست آیند. پس تا اینجا برای دادگان آزمون برای هر ویدئو، تنها ویژگی‌های مرحله 1.2 (شار نوری) محاسبه شده‌اند. حال باید ببینید در هر ویدئو، هر بردار شار نوری در کدام خوشه قرار می‌گیرد و هیستوگرام را بر این اساس تشکیل دهید و مدل خود را تست کنید.

5. تشخیص بر خط

بهترین مدل را که در مرحله 4 انتخاب کردید، در نظر بگیرید. حال ویدئو را از وبکم خوانده و رفتار حرکتی خود را با استفاده از مدل انتخابی تشخیص دهید و روی صفحه چاپ کنید.

نیازمندی‌ها:

- پایگاه داده KTH
- کتابخانه openCV

تحويل:

گزارش کار خود را که شامل موارد زیر است، تا تاریخ 22 دی ماه در درس‌افزار با نام student name_ID.zip

ارسال کنید:

- نتایج مرحله‌ی آموزش بر داده‌ی آموزش برای هر مدل (ماتریس درهم ریختگی⁸ و دقت تشخیص رفتار) و مقایسه آن‌ها
- نتایج مرحله‌ی آزمون برای هر مدل (ماتریس درهم ریختگی و دقت تشخیص رفتار) و مقایسه آن‌ها
- نتایج مرحله‌ی تشخیص بر خط (شامل چند تصویر از عملکرد)

⁸ Confusion Matrix