Sign Language Translator

نام درس: پروژه
نام استاد: دکتر محمدمهدی گیلانیان صادقی
نام دانشجو: نرگس جلیلیان
شماره دانشجویی: ۴۰۰۲۱۴۴۰۸۹۳۰۰۱

Imported Dependencies

```
import cv2
from cvzone.HandTrackingModule import HandDetector
from cvzone.ClassificationModule import Classifier
import numpy as np
import math
import time
```

گزارش مسیر پروژه

ابتدا جهت مکانیابی و تشخیص دست ها از یک detector استفاده میکنیم و بعد حرکات مختلف دست که هر کدام نشانگر حروف الفبا ناشنوایان هستند را دسته بندی میکنیم. پروژه شامل ۲ script است، script اول جهت تشخیص دست، crop کردن و جمع آوری عکس ها برای train است و script بعدی جهت تست داده است. با استفاده از cvzoneدست راست و چپ همراه با اسکلت + مفاصل را تشخیص میدهیم. در عکس اصلی y ابتدای طول و y+h انتهای طول، x ابتدای عرض و x+w ابتدای عرض است که bounding box را به ما می‌دهد. جهت بهتر نمایش دادن دست ها حاشیه offset = 20 در نظر گرفتم. وقتی داده را میخواهیم دسته بندی کنیم باید عکس ها را هم اندازه به شکل مربع crop کنیم. طول و عرض هر عکس برای نشان دادن حروف متفاوت است و ممکن است بعد از cropشدن اطلاعات زیادی از دست برود. به عنوان راه حل با استفاده از np یک ماتریس ones درست میکنم که مربعی سفید ۳۰۰×۳۰۰ به نام imgwhite و نوع داده uint8 است و هر اندازه عکس که داشتیم را در وسط مربع قرار می‌دهیم. در اینجا ریاضی بسیار مهم است و خطا در ۱ پیکسل هم به ما error میدهد. چک میکنیم اگر طول بزرگ تر از عرض بود مقدار طول را ۳۰۰ می‌گذاریم و محاسبه میکنیم که چقدر نیاز است عرض را stretch کنیم تا portionصحیح حفظ شود و در ادامه عکس را در وسط مربع قرار دهیم.

محاسبه عرض

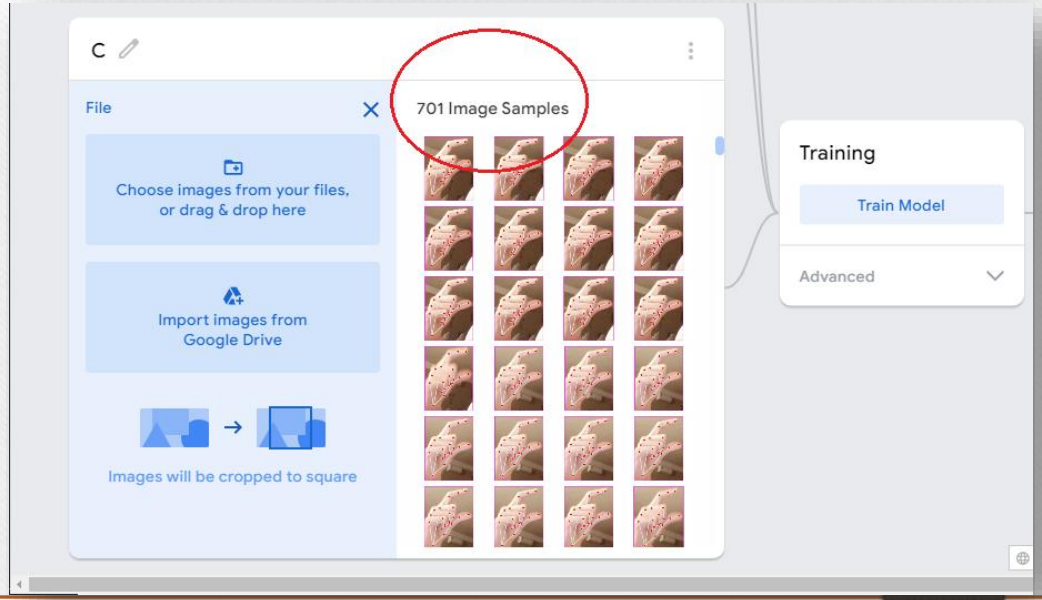
ابتدا h/w میکنیم و بعد حلقه if میسازیم که اگر خروجی بزرگ تر از ۱ باشد یعنی طول بزرگ تر است، در غیر این صورت عرض بزرگ تر است. اگر طول بزرگ تر بود سایز عکس را ÷ طول میکنیم وk (ثابت) را به ما میدهد، با استفاده از ceil هنگام محاسبه سایز عرض خروجی k × عرض را به بزرگ ترین عدد گرد میکنیم مثلا اگر خروجی ۴.۵ شد، ۵ در نظر میگیریم. در آخر عکس crop که عرضش محاسبه شده (wcal) و طول آن ۳۰۰ است را به عنوان عکس resize شده قرار می‌دهیم. حالا برای وسط قرار گرفتن عکس resize شده باید سایز (عکس –wcal)÷ ۲ کنیم تا فاصله ای به نام wgapبرای عرض ایجاد شود. در مربع imgwhite طول ۳۰۰ است، ابتدای عرض wgap و انتهای عرض wcal+ wgap

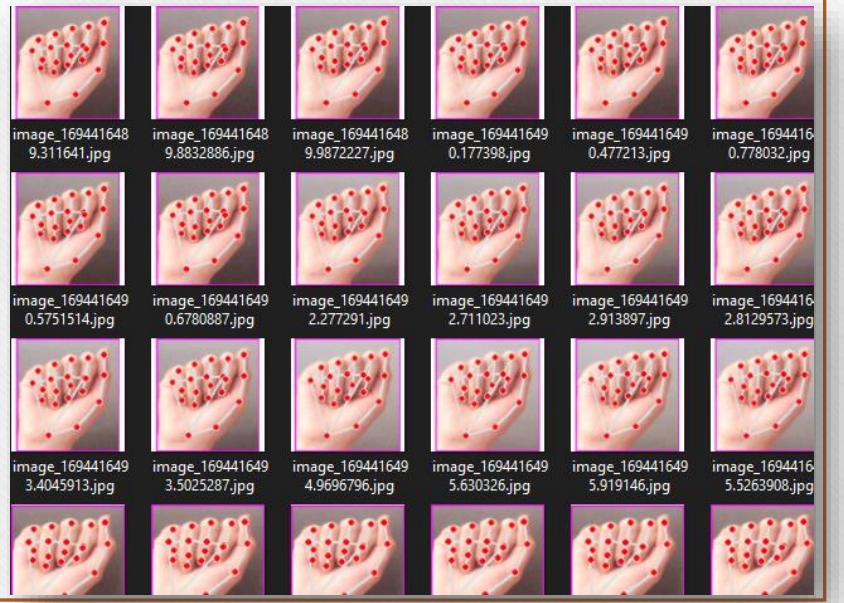
محاسبه طول

اگر عرض بزرگ تر بود سایز عکس را ÷ عرض میکنیم وk (ثابت) را به ما میدهد، با استفاده از ceilهنگام محاسبه سایز طول خروجی k × طول را به بزرگ ترین عدد گرد میکنیم مثلا اگر خروجی ۴.۵ شد، ۵ در نظر میگیریم. در آخر عکس crop که طولش محاسبه شده (hcal) و عرض آن ۳۰۰ است را به عنوان عکس resize شده قرار می‌دهیم. حالا برای وسط قرار گرفتن عکس resize شده باید سایز (عکس –hcal)÷ ۲ کنیم تا فاصله ای به نام hgap برای طول ایجاد شود. در مربع imgwhite ابتدای طول hgap و انتهای طول hcal + hgap، عرض ۳۰۰ است.

ذخیره عکس

حالا نوبت ذخیره عکس ها برای dataset است، کلیدی به نام s در نظر میگیریم که هنگام فشردن آن عکس ها در پوشه ی مورد نظر ذخیره شوند. شمارنده ای داریم که به ما نشان میدهد از هر حروف چقدر عکس گرفته میشود، در اینجا برای هر حرف حدودا ۷۰۰ تا عکس گرفته شده است تا بعد ازtrain شدن دقیق تر عمل کند.





Training with Google teachable machine

بعد از جمع آوری عکس ها با استفاده از google teachable machine آن ها را train میکنیم و مدل train شده keras را دانلود میکنیم. فایل zip به ما میدهد که آن را در پوشه ای جدید به نام model ذخیره و بعد در فایل testفرامیخوانیم. باید Tensorflow نصب داشته باشیم، در فایل testهمان کُد های فایل datacollection را کپی میکنیم، بعضی خط ها را کامنت میکنیم و در ابتدا classifier را از cvzone میاییم import میکنیم و آدرس فایل های keras و labels را در classifier می‌دهیم. جهت گرفتن index و prediction عکس کُد اضافه کردیم، چون نمیخواهیم اسکلت و مفاصل مشخص شود draw=false قرار میدهیم و به جایش مستطیل بنفشی را برای کادر دست به همراه label نمایش میدهیم.