بسم الله الرحمن الرحیم

شرکت مهندسی نرم افزاری هلو

گزارش مربوط به تولید مجموعه های داده­ای مربوط به Pose Command

کاری از امیرعلی نسیمی

* **تولید مجموعه­ی داده**

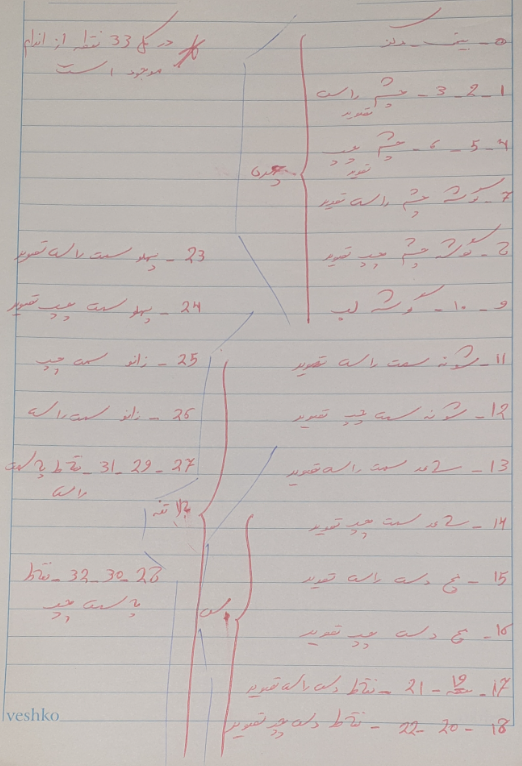
جهت تولید مجموعه داده مربوط به شناسایی فرمان­های مربوط به افراد، ابتدا 3 عمل کلی در نظر گرفته شد. عملیات مورد نظر به صورت:

* + سلام: تکان دادن دست راست
  + اعلام خطر: بالا بودن دو دست بصورتیکه کف هر دو دست روبروی دوربین باشد
  + پخش موسیقی: تکان دادن دست ها در بالای سر

جهت جمع آوری مجموعه داده مورد نظر هر فرد به مدت حدوداً یک ثانیه (به تعداد 30 فریم) هر کدام از حرکات را انجام می­دهد. از هر حرکت حدوداً 20 بار در هر بخش انجام می­شود. در این حالت بطور کلی در حدود 600 نمونه از هر فرد بدست آمده است.

با توجه به اینکه دوربین دارای لنز ماهی است و با توجه به شرایط مربوط به استخراج ویژگی از هر فرد از الگوریتم شناسایی الگوی بدن استفاده شده است؛ بطور دقیق­تر در این حالت کلیه نقاط اندام (122 نقطه) و نقاط کف دست چپ و راست ( هر کدام 63 نقطه) استخراج شده است. کلیه موارد بصورت بردار، تصویر فرد و تصویر کلی ذخیره می­شود.

ترتیب نقاط اندام هر فرد بصورت تصویر صفحه بعد قابل مشاهده می­باشد. از آنجائیکه تعداد نقاط بسیار زیاد می­باشد لذا از ذکر نقاط مربوط به چهره (در حدود 400 نقطه) و نقاط کف دست چپ و راست (در کل 126 نقطه) صرف نظر شده است.



* **مسئله استخراج ویژگی**

با توجه به اینکه مجموعه دادهای تولید شده دارای ویژگی های بسیار کم – حدود 192 بوده است لذا مسئله استخراج ویژگی­های مناسب بسیار حائز اهمیت است. در این خصوص از روش های افزایش نویز استفاده شد تا تعداد نمونه های آموزشی افزایش داشته باشد. از طرفی با توجه به اینکه ویژگی­های استخراج شده باید بصورتی طراحی شده باشد که حالات خاصی که هر فرد انجام می دهد را از حالات دیگر تمیز دهد. در این حالت از شبکه کانولوشن یک بعدی استفاده شده است تا ویژگی­های مربوطه را بهتر شناسایی کند.

* **استفاده از ویژگی­های تفکیک کننده**

جهت آموزش شبکه­ی عصبی ابتدا از کل مجموعه های داده­ای استفاده شد. سپس نقاط کر به پایین حذف گردید و در نهایت نقاط کف دست چپ و راست. با توجه به این مسئله در هر مرحله تعداد ویژگی­های استفاده شده از 192 به 156 در آخرین مرحله به 30 نقطه کاهش یافت. با توجه به کم بودن تعداد ویژگی­های مربوط به هر فرد، امکان آموزش شبکه­ی عصبی بطور کامل وجود ندارد؛ بطور دقیق­تر زمانی که نمونه­ها بصورت تصادفی در حین آموزش شرکت داشته باشند آنگاه دقت بر نمونه­های آموزشی بسیار کم بود – 45 درصد اما در حالتی که داده­ها تصادفی داده نشوند آنگاه دقت به 96 درصد افزایش می­یابد. یکی از مسئل مربوطه در خصوص مدل مذکور، عدم تشخیص حرکات ناشناس می­باشد که برای این مسئله نیاز به داده­های بیشتر می­باشد.