

دستیار کشف خواب‌آلودگی راننده

Apr 20, 2022

علی قربان‌پور - ۹۶۱۰۵۹۹۴

امین روانبخش - ۹۶۱۰۹۷۲۵

علی قدیرلی - ۹۶۱۰۵۹۸۳

شروین جهانبخش - ۹۶۱۰۵۶۷۲

گروه ۸

آزمایشگاه سخت‌افزار - دکتر اجلالی

دانشکده مهندسی کامپیوتر دانشگاه صنعتی شریف

مقدمه

مسائلی که به جان انسان‌ها وابسته است و بر کیفیت زندگی و سلامت آن‌ها اثر می‌گذارد همیشه مورد توجه خاص حوزه‌های مختلف تکنولوژی بوده است. به همین علت پروژه‌ی جاری که دستیار است برای هشدار به رانندگان خواب‌آلود می‌تواند در جهت این مهم قدم بردارد. در مراحل پیشین به تفصیل توضیح داده‌ایم که اهمیت این پروژه چیست، چگونه می‌خواهیم این پروژه را به انجام رسانیم و در نهایت گام‌های پروژه را به دقت شرح دادیم. در این مرحله که در ادامه‌ی کارهای مرحله‌ی پیشین است، می‌خواهیم به صورت عملیاتی مدل یادگیری ماشین اصلی را بر روی دادگان انتخاب شده پیاده کنیم تا بدین صورت بخش نرم‌افزاری پروژه تکمیل و آماده‌ی بهره‌برداری گردد.

اهداف فاز سوم

در فاز سوم که در ادامه‌ی فاز پیشین می‌باشد، می‌خواهیم ابتدا پردازشی بر روی دادگان انتخاب شده در مرحله‌ی قبل انجام دهیم. سپس بعد از آماده‌سازی این مجموعه دادگان، یک مدل یادگیری ماشین را از پایه پیاده‌سازی کنیم. در مرحله‌ی بعد این مدل را با مجموعه دادگان در دست آموزش می‌دهیم و سعی می‌کنیم با تنظیم بهینه‌ی پارامترهای آن به بالاترین دقت ممکن دست بیاوریم. در نهایت نیز مدل آموزش داده‌شده را ذخیره می‌کنیم تا بتوانیم در گام نهایی از آن استفاده کرده و بر روی سخت‌افزار مد نظر خود پیاده‌سازی کنیم.

پیاده‌سازی

برای پیاده‌سازی شبکه‌های عصبی، یکی از بهترین و قوی‌ترین کتابخانه‌ها tensorflow می‌باشد. با استفاده از این کتابخانه و مجموعه دادگان¹ مدل مورد نظر را پیاده‌سازی کردیم. رویکرد این پیاده‌سازی بدین صورت بود که ابتدا در هر تصویر سعی بر آن است که چشم فرد را در تصویر تشخیص دهیم. سپس از روی تصویر جداشده‌ی چشم به این مورد پی ببریم که چشم فرد باز است یا بسته. در نهایت با بررسی وضعیت چشم در هر تصویر مشخص کنیم که آیا این فرد خواب‌آلود است یا خیر. برای این مهم شبکه‌های عصبی پیچشی از قدرت بالایی در تحلیل تصاویر برخوردار هستند. بعد از بسیاری آزمون و خطا در به نتیجه رسیدن ساختار شبکه‌ی عصبی، به شبکه‌ی زیر رسیدیم که برای ما در نهایت نتایج قابل قبولی را کسب کرد و مدل نهایی ما در این پیاده‌سازی شد. ساختار مدل نهایی به صورت زیر است:

```
tf.keras.Sequential([
    tf.keras.layers.Resizing(256,256),
    tf.keras.layers.Rescaling(1./255),
    tf.keras.layers.RandomFlip(mode="horizontal"),
    tf.keras.layers.RandomRotation(0.2,
    fill_mode='reflect',interpolation='bilinear'),
    tf.keras.layers.Conv2D(32, 3, activation='relu'),
    tf.keras.layers.MaxPooling2D(),
```

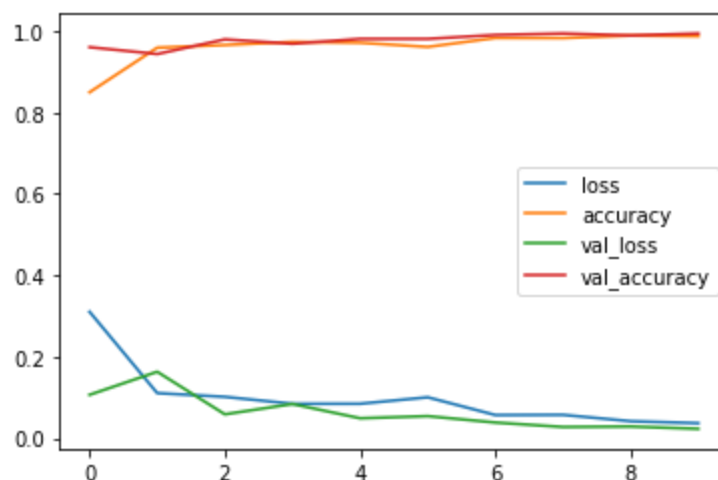
¹ <https://www.kaggle.com/datasets/dheerajperumandla/drowsiness-dataset>

```
tf.keras.layers.Conv2D(32, 3, activation='relu'),
    tf.keras.layers.MaxPooling2D(),
    tf.keras.layers.Conv2D(32, 3, activation='relu'),
    tf.keras.layers.MaxPooling2D(),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(128, activation='relu'),
    tf.keras.layers.Dense(2, activation='softmax')
])
```

با استفاده از ساختار بالا و با الگوریتم بهینه‌سازی Adam، توانستیم در زمان معقولی بر روی مجموعه‌ی دادگان خود مدل را آموزش داده و به مدل نهایی برسیم. تمامی پیاده‌سازی‌ها به صورت عمومی برای مشاهده قابل دسترسی هستند و از لینک² می‌توانید جزئیات پیاده‌سازی را مشاهده کنید.

نتایج

در ابتدا روند تغییرات دقت یادگیری مدل در مراحل آموزش را مشاهده می‌کنید.




همانطور که در نمودار فوق مشاهده می‌کنید، دقت پیش‌بینی مدل بر روی مجموعه‌ی سنجش (validation) به مرور بهتر شده و مدل ما توانسته است این مجموعه دادگان را یاد بگیرد.

همچنین روند تغییرات دقت تشخیص را نیز در جدول زیر می‌توانید مشاهده کنید:

Epoch 1/10

² <https://colab.research.google.com/drive/1kX8lvd5XW0YVy6BeawVOM7AqPStOT8Mb?usp=sharing>



```
264s 3s/step - loss: 0.3107 - accuracy: 0.8496 - val loss: 0.1067 -  
val accuracy: 0.9603
```

```
Epoch 2/10
```

```
246s 2s/step - loss: 0.1109 - accuracy: 0.9592 - val loss: 0.1633 -  
val accuracy: 0.9435
```

```
Epoch 3/10
```

```
247s 2s/step - loss: 0.1017 - accuracy: 0.9657 - val loss: 0.0588 -  
val accuracy: 0.9796
```

```
Epoch 4/10
```

```
248s 2s/step - loss: 0.0847 - accuracy: 0.9734 - val loss: 0.0841 -  
val accuracy: 0.9688
```

```
Epoch 5/10
```

```
247s 2s/step - loss: 0.0848 - accuracy: 0.9712 - val loss: 0.0491 -  
val accuracy: 0.9808
```

```
Epoch 6/10
```

```
247s 2s/step - loss: 0.1006 - accuracy: 0.9610 - val loss: 0.0543 -  
val accuracy: 0.9808
```

```
Epoch 7/10
```

```
247s 2s/step - loss: 0.0571 - accuracy: 0.9830 - val loss: 0.0386 -  
val accuracy: 0.9904
```

```
Epoch 8/10
```

```
248s 2s/step - loss: 0.0575 - accuracy: 0.9827 - val loss: 0.0277 -  
val accuracy: 0.9940
```

```
Epoch 9/10
```

```
255s 3s/step - loss: 0.0419 - accuracy: 0.9889 - val loss: 0.0288 -  
val accuracy: 0.9892
```

```
Epoch 10/10
```

```
252s 2s/step - loss: 0.0371 - accuracy: 0.9873 - val loss: 0.0232 -  
val accuracy: 0.9940
```

در نهایت مدل نهایی ما با دقت ۹۹ درصد بر روی مجموعه‌ی دادگان مان می‌تواند بسته یا باز بودن چشم را تشخیص دهد که این دقت، مقدار قابل توجهی است و برای پروژه‌ی ما کفایت می‌کند.

چالش‌ها

برای انجام این فاز که تمرکز اصلی آن بر روی بخش نرم‌افزاری پروژه و علی‌الخصوص بحث یادگیری ماشین بود، چالش‌های متعددی داشتیم که به شرح برخی از آن‌ها در این بخش می‌پردازیم:

1. نخستین چالش این بود که از حالت دهان در یادگیری مدل نهایی استفاده کنیم یا خیر. برای تشخیص این مورد باید در درجه اول به پیچیدگی مدل خود می‌افزودیم که سعی ما بر این است که تا جای امکان از این مورد جلوگیری کنیم. زیرا مدل هر چه پیچیده‌تر شود، منابع محاسباتی برای یادگیری آن و همچنین سرعت آن در پاسخگویی به یک ورودی بیشتر می‌شود که این مورد نهایی با توجه به قالب پروژه که ذاتاً یک پروژه real-time محسوب می‌شود منافات دارد. همچنین با افزودن تشخیص حالت دهان که در حال خمیازه است یا خیر، دقت محاسباتمان به میزان قابل توجهی پایین می‌آید، در حد ۱۰ درصد که این مورد نیز توجیه مناسبی نداشت. به همین علت در نهایت تصمیم بر آن شد که در مدل نهایی فقط از حالت چشم بهره ببریم و با استفاده از حالت چشم به تشخیص خواب‌آلوده بودن یا نبودن راننده بپردازیم.
2. چالش بعدی که از چالش‌های ذاتی پروژه‌های یادگیری ماشین است تنظیم هاپرپارامترها و توابع بهینه‌ساز به گونه‌ی مناسب بود تا بتوانیم دقت مدل را در حد قابل توجهی به میزان بالا حفظ کنیم. برای این مورد نیز با جست‌وجوی گسترده در گوگل و آزمودن مقادیر مختلف به پارامترهای نهایی دست یافتیم که در مدل نهایی ارائه شده موجود می‌باشند.

چالش‌های نام برده بارزترین چالش‌هایی بودند که با آن‌ها روبرو شدیم. مواردی جزئی‌تر نیز وجود داشتند که به تفصیل در گزارش پایانی ذکر خواهند شد. در اینجا به دلیل کم بودن اهمیت‌شان از ذکر آن‌ها خودداری می‌کنیم.

گام بعدی

گام بعدی، گام نهایی پروژه می‌باشد به این صورت که با در دست داشتن مدل آموزش‌دیده و یک سخت‌افزار مناسب، مدل مد نظر را روی آن پیاده‌سازی می‌کنیم. همچنین ورودی این سخت‌افزار به نحوی که در فاز بعدی تصمیم گرفته می‌شود تصاویر راننده در حال رانندگی خواهد بود که در صورتی که مدل ما تشخیصش بر آن بود که چشمان راننده به مدت طولانی بسته است، با یک چراغ یا صدای بوق این مورد را به راننده اخطار دهد تا از خطرات احتمالی این مورد جلوگیری کند.