

## شبکه های عصبی برای تشخیص الفبای انگلیسی دست نویس

علی اکبر ذبیحی

دانشگاه خوارزمی

ZABIHIALI1998@GMAIL.COM

### چکیده

در این گزارش قرار است سیستمی طراحی شود که با بررسی حروفی که به صورت دستی نوشته شده اند را تشخیص دهد و حروف را از داخل متن دستی خارج کند که ورودی یک تصویر دارای متن که از پیش پردازش عبور می کند و به عنوان ورودی به سیستم داده می شود و حرف مورد نظر در تصویر شناسایی و به عنوان خروجی باز گردانده می شود .

### [ ۱ ] مقدمه

در این گزارش که امیدی است برای ساخت سیستم هایی قوی تری که قدرت تشخیص متن های دستی که با حالت ها مختلف نوشته شده را شناسایی و بررسی نماید هر فرد با فرد دیگری از نظر نوشتاری با هم تفاوت هایی دارند که این تفاوت ها در سبک نوشتاری و شکل حروف و اندازه های حروف که هر یک به نوعی پیچیدگی خاصی را به تصویر اضافه می کند .

میزان اهمیت این کار بسیار زیاد است به خاطر کاربردی که دارد از نظر بررسی چک های بانکی و از نظر بررسی پاسپورت ها و از نظر بررسی فرم های اداره ای ..... .

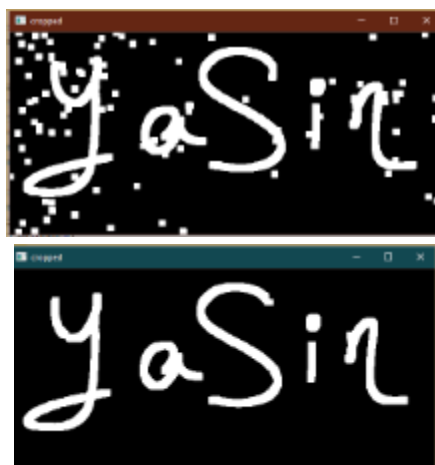
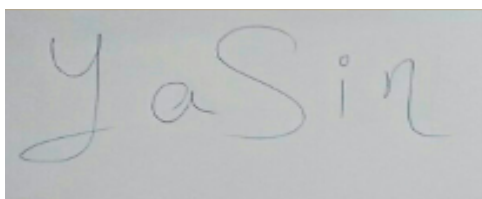
به گونه ای این قابلیت باعث می شود که ماشین قابلیت درک دید و قابلیت تشخیص حروف و متن ها تا بتواند کار های همانند انسان بر روی متن های دستی صورت دهد .

حتی می توان حالتی را ایجاد کرد که متن دستی نوشته شده از نظر درک بصری انسان غیر قابل تشخیص می باشد ولی با سیستم می توان به صورت احتمالی پیشبینی که صورت می دهد مشخص کرد که حروف به کار رفته در کلمه نشان دهنده چه حرف و کلمه ای به عنوان خروجی می باشد.

در مورد شبکه عصبی طراحی شده تعداد لایه ها و تعداد نرون های عصبی و مقدار دهی وزن ها و تابع بهینه ساز و تابع خطا نیز یکی از مهم ترین پارامتر ها در ایجاد یک شبکه عصبی مناسب می باشد.

#### [۲] پیش پردازش

یکی از مهم ترن بخش های این سیستم بخش پیش پردازش قبل از اعمال به شبکه می باشد زیرا در این مرحله یک نوع مقدمه ای که تصویر ورودی را برای بررسی در داخل سیستم و این امر بدین صورت می باشد که ابتدا تصویر مورد نظر را دریافت می کنیم و با از بین بردن نویز های تصویر و افزایش بهبود کیفیت تصویر در مرحله بعدی قطعه قطعه کردن تصویر اولیه به تصاویر کوچک تر می باشد. که این عمل به بهبود کیفیت تصویر و درک بهتر شبکه از کلمه موجود در تصویر می شود.



شکل ۱. در تصویر بالا تصویر اصلی می باشد که با حالت خاصی نوشته شده است در تصویر دوم که بدون استفاده از پیش پردازش و نویز های تصویر را ملاحظه خواهیم کرد که در تصویر سوم تصویر اول را از تابع پیش پردازش که بهبود کیفیت تصویر را مشاهده می کنیم.

در پیاده سازی این پروژه در بخش پیش پردازش باید تصویر ورودی دو قانون را رعایت کند و بعد به عنوان ورودی به شبکه داده می شود :

- باید بین گراکتر ها فاصله یا همان نیم فاصله وجود داشته باشد.
- کیفیت تصویر در بخش قطعه بندی کلمه به حروف کیفیت تصویر بسیار اهمیت دارد.

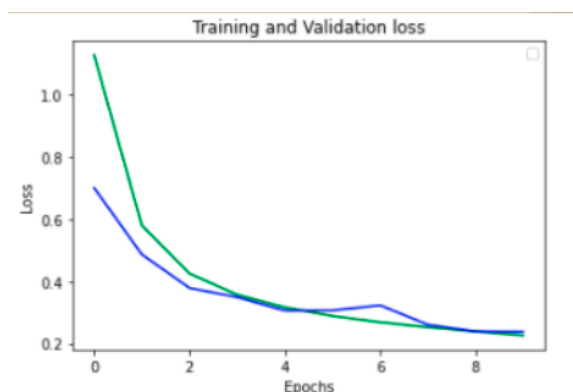
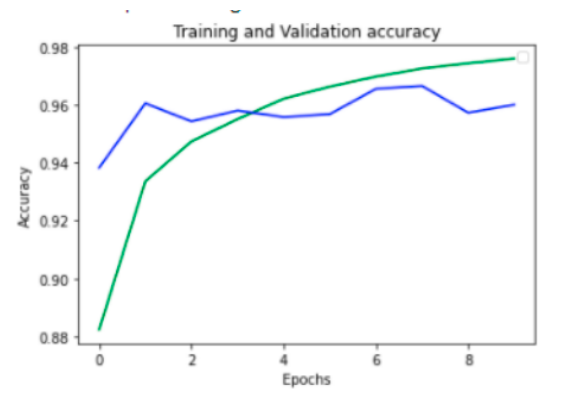
اصولی که بیان شده بسیار مهمه که در تصویر های ورودی باید رعایت شود که در همه شبکه ها همچین اصول وجود دارد که باید با فرمت خاص یا باید با کیفیت مناسب یا ..... به عنوان ورودی به شبکه اعمال شود.

## [۴] شبکه عصبی

از شبکه عصبی کانولوشنی و از داده های EMNIST برای آموزش شبکه که داده های آموزش دارای ۱۲۸۰۰۰ تصویر مختلف و دارای ۲۰۰۰۰ تصویر تست می باشد .

Layer (type)	Output Shape	Param #
conv2d_12 (Conv2D)	(None, 26, 26, 32)	320
conv2d_13 (Conv2D)	(None, 24, 24, 64)	18496
max_pooling2d_6 (MaxPooling2D)	(None, 12, 12, 64)	0
dropout_6 (Dropout)	(None, 12, 12, 64)	0
flatten_6 (Flatten)	(None, 9216)	0
dense_12 (Dense)	(None, 256)	2359552
dense_13 (Dense)	(None, 27)	6939
Total params: 2,385,307		
Trainable params: 2,385,307		
Non-trainable params: 0		

شکل ۳. در تصویر بالا ساختار شبکه کانولوشنی طراحی شده را نشان می دهد که دارای ۲ میلیون پارامتر می باشد . پس از اجرای شبکه که از تابع بهینه سازی ADAM و از تابع خطا CATEGORICAL استفاده می کنیم و این شبکه را ۱۰ بار اجرا می کنیم و هدف در این آموزش جلوگیری از بیش برآزش یعنی سعی شده دقت تست همواره بیشترین باشد زیرا در چندین بار تست شبکه به این نکته رسیدیم که هر چه اجرای شبکه وقتی میزان آموزش دارای بیشتری دقت نسبت به تست بود کلماتی که در بخش پیشبینی بدست می آمد کلمات مناسبی نبوده ولی هر چه میزان تست دقتش افزایش پیدا می کند باعث می شود که در مرحله پیشبینی حروف های بهتری پیشبینی شوند .



شکل ۴. در تصویر بالا میزان دقت در مرحله ارزشیابی و در مرحله تست را نشان می دهد که در آموزش دقتی نزدیک به ۰,۹۸ و در تست دقتی معادل با ۰,۹۶ بدست می آوریم و بر عکس در روش تابع خطا که در تصویر دوم مشخص است میزان خطا در تابع خطا کاهش یافته و همگرایی یکسانی به نقطه مینیمم دارند

## [۴] کتابخانه ها و پکیج های از پیش آماده

یکی از مهمترین کتاب خانه ها در محور (OCR) استفاده از (PYTHON TESSERACT) می باشد که برای تشخیص متن در داخل تصویر مورد استفاده قرار می گیرد به گونه ای نوعی استخراج یا تشخیص جایگاه متن در تصویر را بر عهده دارد . با توجه به وجود یک سری تابع ها و کتاب خانه ها تصویر را می خواند و مکان هایی که متن وجود دارد شناسایی و خروجی می دهد . نکته ای که وجود دارد این کتاب خانه در تشخیص متن ها دقت مناسبی ندارد بدین شکل که در

متن هایی که به صورت کامپیوتری نوشته شده دقت مناسب ولی در متن های دستی خروجی مناسبی دریافت نمی شود و این پروژه در جهت بهبود این دقت می باشد . در ادامه تصویری که گویا این شواهد می باشد :

```
import pytesseract

def ocr_core(filename):
    """
    This function will handle the core OCR processing of images.
    """
    text = pytesseract.image_to_string(Image.open(filename)) # Use pytesseract to detect the string in the image
    return text

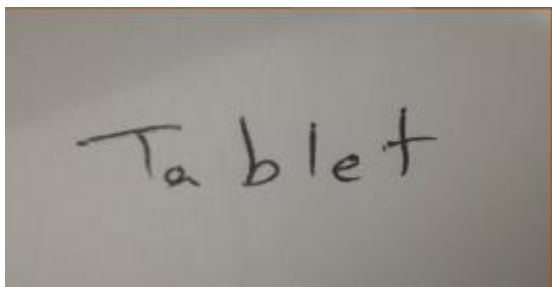
print(ocr_core('images/ocr_example_1.png'))
```

شکل ۵. در شکل بالا قطعه کدی برای استخراج کردن متن از تصویر می باشد .



```
C:\Users\Ali\anaconda3\envs\untitled\python.exe
Book
Process finished with exit code 0
```

شکل ۶. در تصویر بالا خروجی است که از قطعه کد داده شده بدست آمده که مراحل تشخیص متن و استخراج به درستی صورت گرفته .



```
C:\Users\Ali\anaconda3\envs\untitled\python.exe C:/Users/Ali,
Process finished with exit code 0
```

شکل ۷. در تصویر بالا که همان تصویری است که با دست نوشته شده وقتی به عنوان ورودی به تابع می دهیم همان طور که از خروجی تصویر دوم مشخص است قادر به حدس متن نوشته شده در تصویر نمی باشد. که این امتیازی هست که می توان به سیستم طراحی شده داد.

[۵] جمع بندی

در این مطالعه سعی شده سیستمی طراحی شود که قدرت تشخیص متن های دست نویس را دارا باشید هر فرد برای خود نوع نوشتاری مخصوصی دارد که از این رو می توان بعضی از افراد را خوش خط و بعضی ها را بد خط دسته بندی نمود و از کار های این متن های دست نویس در علم پزشکی و علوم مختلف بررسی کردیم و سر انجام سیستم طراحی شده را با یک سیستم آماده بررسی کردیم که در سیستم آماده متن هایی که به صورت کامپیوتری نوشتار شده بود را تشخیص می داد و برای تصاویر دستی دارای خطا بسیاری بود و سر انجام در مورد چرخه اجرای برنامه صحبت کردیم و به این قدرت که سیستم بتواند دست نویس های هر فرد را تشخیص دهد مقدمه است برای برابر کردن درک بصری انسان و ماشین .

مراجع

- H. Al-Yousefi and S. S. Udpa, "Recognition of handwritten Arabic characters," in Proc. SPIE 32nd Ann. Int. Tech. Symp. Opt. Optoelectric Applied Sci. Eng. (San Diego, CA), Aug. 1988
- Nadal, C. Legault, R. Suen and C.Y, "Complementary Algorithms for Recognition of totally Unconstrained Handwritten Numerals," in Proc. 10th Int. Conf. Pattern Recognition, 1990, vol. 1, pp. 434-449.