

تشخیص چهره به کمک شبکه عصبی CNN

محمد لشکری ۴۰۰۱۱۲۰۸۷

۱ چکیده

در این پروژه یک شبکه CNN^۱ را برای یادگیری ویژگی های عکس جهت تشخیص چهره انسان ها پیاده سازی کردیم. این شبکه روی ۲۴۴ عکس آموزش دیده که ۲/۰٪ آن به عنوان مجموعه اعتبارسنجی جدا شده است. مدل نهایی دارای دقت ۹۸/۰٪ روی مجموعه آموزشی و ۹۷/۰٪ روی مجموعه اعتبارسنجی است.

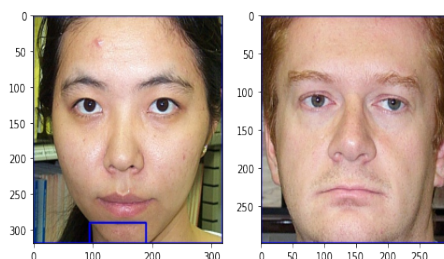
۲ مقدمه

تشخیص چهره یکی از مهم ترین کاربردهای یادگیری ماشین است که در سیستم های پیشنهاد دهنده، احراز هویت در دستگاه های الکترونیکی و استخراج اطلاعات استفاده می شود. امروزه روش های مختلفی برای حل این مسئله به کمک شبکه های عصبی گوناگون ارائه شده است. هدف این پروژه، پیاده سازی یک شبکه CNN برای طبقه بندی چهره افراد است. یکی از خواص این شبکه استخراج ویژگی از دادگان است تا زاویه عکس، ویژگی های فرعی در عکس و بزرگ نمایی و کوچک نمایی در مقیاس کم روی نتیجه تأثیر نداشته باشد.

۳ دادگان

مجموعه داده هایی که برای این پروژه استفاده شده است شامل ۲۴۴ عکس به عنوان مجموعه آموزشی و ۶۴ عکس به عنوان دادگان تست از چهره های آقایان و خانم ها است. این مجموعه که نمونه آن را در شکل ۱ مشاهده می کنید شامل عکس هایی با ۱۶ برچسب است که هر برچسب نشان دهنده یک چهره متفاوت است. مجموعه داده از این لینک قابل دریافت است.

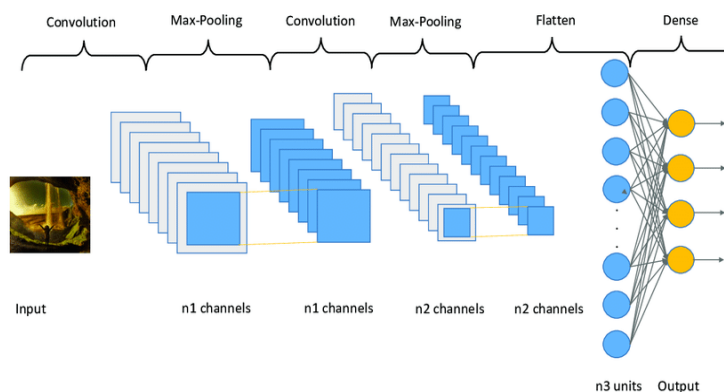
¹Convolutional Neural Network



شکل ۱: نمونه دادگان

۴ شبکه CNN

همان‌طور که گفته شد این شبکه‌ها به کمک استخراج ویژگی از عکس، مدل پیش‌بینی را می‌سازند. ابتدا با استفاده از یک لایه پیچیدگی^۲ ویژگی‌های مهم از عکس استخراج و سپس به کمک یک لایه ترکیب^۳ ویژگی‌های فرعی حذف می‌شود تا ویژگی‌های مهم‌تر نمایان شوند. خروجی این مرحله خود یک عکس است که آن را feature map می‌نامیم. این عمل به صورت بازگشتی روی خروجی مراحل قبل انجام می‌شود [۱]. feature map نهایی مسطح شده و به عنوان ورودی به یک پرسپترون چندلایه داده می‌شود که تعداد نورون‌های لایه آخر آن به تعداد برجسب‌های عکس‌های موجود در دادگان است. شکل زیر نشان‌دهنده نقشه راه این شبکه است [۲]:



شکل ۲: نقشه راه CNN

^۲Convolution Layer

^۳Pooling Layer

۵ آزمایش

برای پیاده سازی ابتدا ابعاد هر پیسکل از بازه $[0, 255]$ به $[0, 1]$ مقیاس شده است و روش‌های مختلف افزایش داده روی دادگان اعمال شده است. در معماری شبکه ابتدا یک لایه پیچیدگی با ۳۲ هسته و سپس یک لایه ترکیب آمده است. در پی آن یک لایه پیچیدگی با ۶۴ هسته و یک لایه ترکیب آمده است که بعد هر هسته ۳ و بعد ترکیب ۲ است. در نهایت یک لایه قویاً همبند^۴ با ۶۴ نورون قبل از لایه انتهایی شبکه آمده است. شبکه با استفاده از الگوریتم آدام با ۳۲ دسته^۵، نسبت ۰/۲ برای مجموعه اعتبارسنجی و در ۳۰ مرحله^۶ آموزش دیده است. نتایج برای مجموعه اعتبارسنجی و آموزشی در جدول ۱ قابل مشاهده است. مدل تا مرحله

مرحله	دقت آموزش	دقت اعتبارسنجی
۵	۰/۱۲	۰/۰۹
۱۰	۰/۵۳	۰/۷۰
۱۵	۰/۸۸	۰/۷۹
۲۰	۰/۹۷	۰/۹۰
۲۵	۰/۹۵	۰/۹۵
۳۰	۰/۹۸	۰/۹۷

جدول ۱: نتایج

۱۰ دچار underfitting و بعد از آن شروع به یادگیری عمیق‌تر کرده و در نهایت در مرحله ۲۰ با دقت خوبی همگرا شده است. همچنین برای دو نمونه از دادگان تست که در ۱ قابل مشاهده است مدل به درستی پیش‌بینی را انجام داد. تصویر اول متعلق به کلاس face3 و عکس دوم متعلق به face2 است.

۶ نتیجه‌گیری

همانطور که در بخش ۵ اشاره شد مدل با دقت خوبی همگرا شده است و قابل قبول است. در این پروژه از CNN برای حل این مسئله استفاده شده است. اما تشخیص چهره به کمک شبکه‌های عصبی بازگشتی مانند RNN, LSTM, Bi-LSTM قابل انجام است. همچنین برای آنکه مدل به دست آمده کاربردی‌تر باشد می‌توان از روش‌هایی استفاده کرد که حافظه کمتری مصرف کنند تا مدل بتواند روی دستگاه‌های مختلف مانند موبایل و تبلت نیز کارا باشد.

⁴Fully Connected

⁵batch

⁶Epochs

مراجع

- [1] <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>
- [2] García-Ordás, M. T., Benítez-Andrades, J. A., García-Rodríguez, I., Benavides, C., & Alaiz-Moretón, H. (2020). Detecting respiratory pathologies using convolutional neural networks and variational autoencoders for unbalancing data. *Sensors*, 20(4), 1214.