

سیستم تصمیم یار مراقبت های پوستی شخصی سازی شده با استفاده از یادگیری ماشین

مهشید یعقوب زاده^۱

^۱ سازمان متبوع نویسنده اول

FirstAuthor@Email

انتخاب کرد. همچنین خود آکنه نیز با توجه به میزان شدت روی پوست به چند دسته تقسیم می شوند تا بلکه به راحتی بتوان آن ها را مورد بررسی قرار داد. در بخش Prediction و Recommendation نیز با کمک مقاله ها، کتاب ها و نظرات پزشکان معتبر دنیا دیتاستی از توصیه های مراقبت های پوستی همراه با سن، نحوه ی مصرف و ساخت، جمع آوری شده است تا تمامی نکات برای به دست آوردن نتیجه ی خوب و داشتن پوستی صورت زیبا، رعایت شود. سپس از متد Natural language processing استفاده شده است تا کلمات کلیدی و کلاس های دیتاست مورد نظر مشخص شوند و در نهایت الگوریتم های Random XBG Classifier، Voting Classification، Logistic Regression، Forest و همچنین از ROC و Confusion Matrix و شبکه ی عصبی CNN+SVM آموزش و مورد بررسی قرار گرفته شده است.

۲. کارهای دیگران

ابتدا ایده ی این کار به صورت شناسایی موضع ها مانند آکنه، لک، چین و چروک و... بوده است که دقت کار در ابتدا بسیار پایین بوده اما با گذر زمان محققان در جهت رفع مشکلات دقت پایین و همچنین گسترش این نوع سیستم ایده ها و مقالات زیادی را ارائه کرده اند. از جمله می توان به مقاله ی Xiaoping Wu و همکارانش [۱] در سال ۲۰۱۹ اشاره کرد؛ با توجه به نگاه اجمالی که از این کار در فصل یک داشتیم این مقاله در جهت شناسایی آکنه، روش جدیدی را ارائه می دهد تا بلکه مقداری از دقت پایین مقاله ها و پروژه های سال قبل جبران شود. اما این سوال برای ما مطرح می شود که اگر فقط از روی تصاویر گرفته شده ی سلفی، آکنه ها را با کمک روش object detection شناسایی بشوند و بر اساس روش Hayashi شمرده و لیبل گذاری بشوند، در نهایت دقت بالایی را به ما می دهند؟ پاسخ این سوال خیر می باشد؛ چرا که آکنه ها متاسفانه مانند اجزای صورت دارای لبه نیستند که بتوان به راحتی تشخیص داد پس می توانیم ویژگی های بیشتری را برای گسترش بیشتر تحقیق و افزایش دقت شناسایی آکنه در نظر بگیریم. در

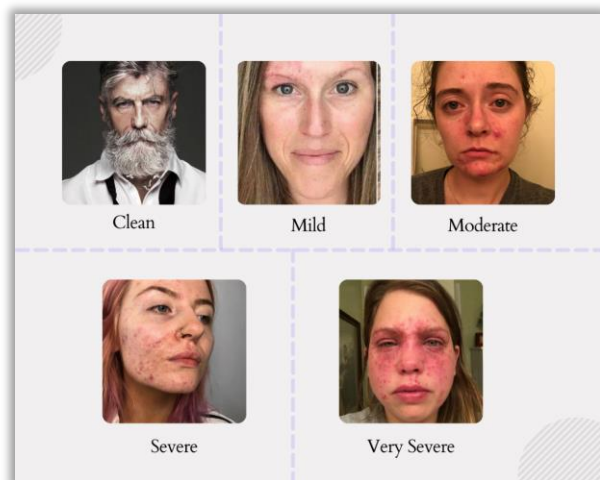
چکیده- در این paper با تمرکز بر مسئله ی زیبایی پوست و با به کارگیری از تحلیل و طبقه بندی چهره های دارای آکنه و تشخیص سن و جنسیت، این سیستم ارائه شده است به همین منظور از شبکه های عصبی عمیق و پردازش تصویر در بخش طبقه بندی اطلاعات، همچنین با کمک علم داده و تعداد مختلفی از الگوریتم های یادگیری ماشین، مورد بررسی و آموزش قرار گرفته شده است. در نهایت قسمت تشخیص سن و جنسیت با بهره گیری از پردازش تصویر ساخته شده است. در این حوزه از چندین مجموعه داده های مختلف چهره استفاده شده است و همچنین با بهره گیری از مقاله ها و کتاب های پزشکان معتبر دنیا مجموعه داده ای در زمینه های مراقبت های پوستی با توجه به سن های مختلف جمع آوری و مورد استفاده قرار گرفته شده است.

کلمات کلیدی- الگوریتم های یادگیری ماشین، پردازش تصویر، شبکه های عصبی، یادگیری عمیق

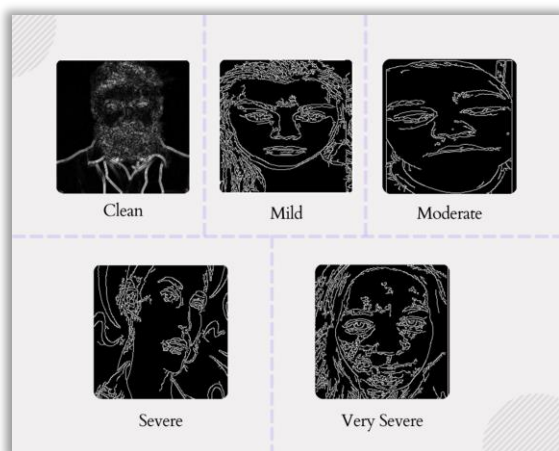
۱. مقدمه

در این پروژه تمامی تلاش ها برای ساخت سیستمی با دقت بالا متمرکز شده است. ابتدا با جمع آوری دیتاستی از چهره ها با انواع مختلف سطوح شدت آکنه ی ولگاریس به صورت ورودی به سیستم داده می شود. به علت تشابه زیاد میان دو سطح شدت آکنه severe و very severe از سه شبکه ی CNN-Res Net، CNN-، SVM، CNN-Random Forest در بخش Classification تست و آموزش داده شده اند. در بخش بعدی سن و جنسیت با کمک روش های پردازش تصویر شناسایی می شود. باید این را در نظر داشته باشیم که بسیاری از ماسک ها و مراقبت های طبیعی خانگی می توانند تأثیرات مثبت به سزایی نسبت به ماسک های آماده روی صورت داشته باشند البته اگر طبق نحوه ی مصرف، از آن ها استفاده شود و قبل از استفاده حتما پوست صورت کاملاً تمیز باشد. همچنین بایستی در نحوه ی استفاده از آن ها احتیاط کرد و حتماً با توجه به نوع پوست، رنج سنی و نوع آکنه ماسک مناسب را

کاهش پیچیدگی تصویر و پردازش سریع تر آن و حذف نویز می باشد که با توجه به شدت میزان آکنه برای حذف نویز از فیلتر متناسب با آن استفاده می کنیم که فیلتر های Canny و Gaussian استفاده شده است .



شکل ۱. تصاویر صورت در سطوح مختلف شدت آکنه



شکل ۲. تصاویر صورت در سطوح مختلف شدت آکنه پس از انجام پیش پردازش

قبل از هر چیزی باید ای را بدانید که به علت شباهت زیاد دو سطح severe و very sever بر روی این دو نوع تصاویر بایستی دقت بیشتری به خرج داد . برای به دست آوردن نتایج بهتر در حوزه ی طبقه بندی داده ها از سه شبکه ی زیر استفاده و بررسی می کنیم.

• RandomForest

شبکه ی CNN – Random Forest به این صورت عمل می کند که feature های ماتریس gray تصاویر به صورت خودکار به وسیله ی CNN استخراج می گردند . همچنین dropout و zero

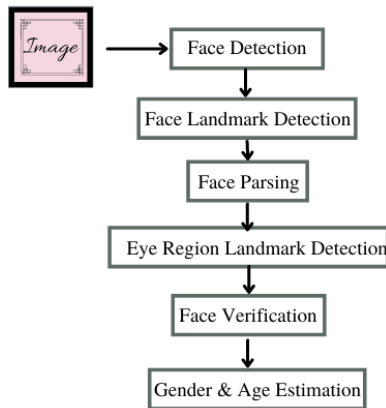
حقیقت باید از دیتاست localization استفاده کنیم که محل آکنه ، طول و عرض آکنه و مختصات مرکز آکنه را مشخص کند تا object detection دقت بالایی داشته باشد ! ولی آنقدر آکنه ها در تصاویر گرفته شده کوچک هستند که این روش بسیار سخت آموزش پیدا می کند و یاد می گیرد . من از این روش در قسمت شناسایی و کلاس بندی آکنه به ۴ دسته ی mild , moderate , severe و very severe برای ساخت پروژه ی خودم الهام گرفته ام با این تفاوت که شخص مخاطب اطلاعات پوست خود را که در کدام دسته ی شدت آکنه قرار دارد ، به ورودی سیستم وارد کند . اما درست در همان سال Tingting Zhao و همکارانش [۲] تصمیم گرفتند که این معضل تخمین دقت پایین در بخش شدت آکنه را رفع کنند ، پس از روش texture استفاده کردند اما باز هم خطای قابل توجهی را داد پس با شرکت Microsoft همکاری کردند و شبکه ی عصبی Res Net را برای بخش classification وارد این روش کردند . در این روش به دلیل اینکه label کل صورت را به عنوان لیبل برای هر skin patch استفاده شده ، باعث ایجاد noise اضافه به بخش training skin می شود ، پس بایستی هر یک از skin patch ها را به صورت جداگانه لیبل گذاری شود . مشکل دیگری که وجود دارد این است که طبق گفته ها ی قبل شناسایی آکنه به وسیله تصاویر گرفته شده کار بسیار سختی است پس باید موارد دیگری نیز برای شناسایی اضافه کنیم.

۳. روش تهیه مقاله

سیستم تصمیم یار مراقبت های پوستی شخصی سازی شده به علت داشتن حجم بالایی محاسباتی در هر بخش نیازمند سیستمی با قدرت بالا می باشد ؛ بنابراین برای رفع چنین مشکلی هر بخش به صورت جداگانه اجرا می شود . پس به طور کلی می توان این سیستم را از سه جهت بررسی کرد .

۱-۳- کلاس بندی و طبقه بندی داده ها

ابتدا شخص کاربر عکس چهره ی خودش را به سیستم می دهد همچنین دسته بندی سطح آکنه ی پوست خود را نیز بر اساس مقیاس شدت به سیستم وارد می کند . طبق طبقه بندی آکادمی پوست [۳] در سال ۱۹۹۰ برای آکنه های ولگاریس پوست بر اساس شدت در این پروژه نیز مقیاس درجه بندی clear , mild , moderate , severe و very severe را برای این نوع آکنه در نظر گرفته شده است . به طور کلی قبل از بررسی پیش پردازش تصویر این را باید بدانید که چارچوب کار این بخش برای تحلیل داده های ورودی الزامی می باشد . در حقیقت سیستم بینایی ماشین به تحلیل تصاویر آکنه و همچنین طبقه بندی و دسته بندی اطلاعات کمک به سزایی را می کند . در حالت کلی پیش پردازش تصویر شامل تغییر اندازه ی تصاویر ، تبدیل فضای رنگی از فضای RGB به grayscale برای



شکل ۳. دیاگرام کلی تخمین سن و جنسیت

۳-۳ پیش بینی و پیشنهاد دهنده ی مراقبت های پوستی

در این بخش پس از ورود دیتاست توصیه های مراقبت های پوستی به سیستم از آنجایی که اطلاعات دیتاست شامل داده های متنی که ساخت نیافته هستند می باشد ، به دلیل چنین پیچیدگی از متد پردازش زبان طبیعی استفاده می کنیم تا کلمات کلیدی را از دیتاست استخراج کنیم که به وسیله ی آن بتوانیم به راحتی پیشنهاداتی را بدهیم از آنجایی که در دیتاست acne فقط اطلاعات بخش How to make و Description برای پیشنهاد دهنده مهم هستند پس ابتدا باید این دو بخش را جدا کنیم . علائم نگارشی و نماد ها (مانند / * _ @ = : " { } ^ < % ! ؟) باید از دیتا حذف شود . سپس مرحله ی Tokenization باید صورت گیرد . در این مرحله متن باید به واحد هایی مثل جمله یا کلمه تبدیل شوند . Stop word ها حذف می شوند . سپس ستون age که سن های متناسب برای هر توصیه ی مراقبت زیبایی به ما می دهد را مرتب می کنیم . به این صورت که برای هر سن در نظر گرفته شده ی دیتاست باید با کمک روش میانگین گیری یک عدد در نظر بگیریم زیرا که دیتا به ما یک عدد به خصوص نداده است و فقط بازه ی سنی را مشخص کرده است . تا این مرحله encoding داده صورت می گیرد . از آنجایی که توصیه های مراقبت های پوستی دیتاست مورد نظر به صورت متعادل در دیتا

padding نیز برای بهینه سازی ساختار CNN و جلوگیری از کاهش overfit شدن کمک می کند . روش Random Forest نیز قدرت زیادی در توانایی generalization دارد و classification را به خوبی انجام می دهد .

• CNN SVM

در چنین روشی سیستم پیشنهادی طبقه بندی شبکه ی CNN را به وسیله ی مدل کلاس بندی SVM تغییر می دهد . در حقیقت این روش هم ویژگی ها را به طور خودکار از تصاویر به دست می آورد و هم طبقه بندی ویژگی ها را انجام می دهد .

• CNN Resnet

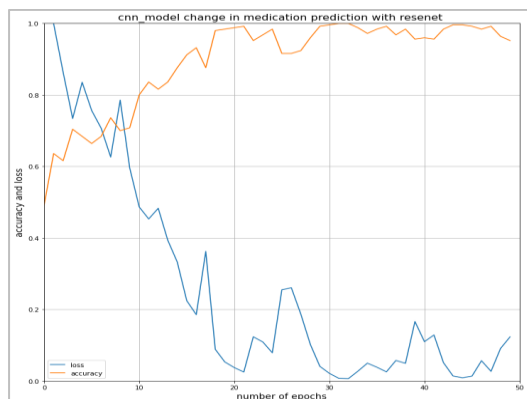
این شبکه ترکیبی از CNN و Resnet می باشد که در حقیقت عملیات در بخش شبکه ResNet به گونه ای رخ می دهد که می تواند عملیات را لایه به لایه خروجی را مستقیماً به ورودی وصل کند . یعنی می تواند^۱ Skip connection انجام دهد .

۲-۳ تخمین سن و جنسیت

برای ساخت پیشنهاد دهنده ای با قدرت و دقت کافی نمی توان فقط یک پارامتر طبقه بندی آکنه را در نظر گرفت . انواع مختلف مراقبت های پوستی می توانند به عوامل مختلفی از جمله سن ، جنسیت ، موقعیت آب و هوایی ، نوع بافت پوست و ... بستگی داشته باشد و بهتر است متناسب با این پارامتر ها مراقبت های پوستی پیشنهاد داده شود ؛ بنابراین در این مرحله ما تشخیص سن و جنسیت را نیز وارد سیستم کرده ایم . وهله ی اول بایستی تصویر ورودی را بعد از انجام پیش پردازش های لازم ، در بخش تشخیص چهره وارد کنیم تا فقط صورت را از عوامل مختلف و پس زمینه جدا سازیم و موقعیت یابی کنیم . برای برش صورت به جای آنکه feature های تصاویر از جمله scale ، translation و ... را از بین برده و تغییر داده شود landmark های صورت را از تصویر ورودی تشخیص داده می شود که از دو شبکه ی Facial landmark localization و FRN ساخته شده است [۴] سپس تجزیه ی چهره با کمک شبکه ی ResNet انجام می شود تا اجزای مختلف صورت مثل مو ، چشم ، دهان و ... با کمک روش تجزیه چهره محاسبه گردد . [۵] در مرحله ی بعد landmark های ناحیه ی چشم با کمک متد GazeMap شناسایی می شود [۶] و در نهایت face verification رخ می دهد تا با استفاده از شبکه ی Mobile Net و روش های یادگیری ماشین سن و جنسیت تخمین زده شود .

^۱ shortcut connection

همچنین از الگوریتم یادگیری عمیق مثل ترکیب شبکه ی CNN و Resnet نیز استفاده شده است که برای تحلیل این شبکه نمودار زیر که نشان دهنده ی فرآیند دقت و خطا در شبکه ی CNN – Resnet می باشد ، را در نظر گرفته ایم . وقتی هر epoch زده می شود ، مدل باید مراحل مختلف را طی کند تا بتواند به يك ساختار ثابتی پیدا کند در نتیجه همان طور که مشاهده می کنید خطا که خط آبی است کاهش پیدا می کند و دقت نیز که خط زرد است افزایش می یابد .



شکل ۴. نمودار دقت و خطای CNN - ResNet

به طور مثال برای تعداد ۱۰۰ epoch دقت % ۶۰ و خطا ۱/۲ درصد محاسبه شده است . پس اگر با مدل هایی که قبلا داشتیم محاسبه شود متوجه می شویم با وجود این که مدل ها ساده تر بودند اما دقت بالاتری را در نهایت به ما داده اند در نتیجه عملکرد بهتری را نسبت به این شبکه ی ترکیبی داشتند .

Algorithm	Test Accuracy	Train Accuracy	Total
XGB Classifier	68.52	88.79	78.65 (+/- 0.1)
XGB Classifier + Random Over Sampling	96.78	99.50	97.00 (+/- 0.1)
Logistic Regression	68.65	68.75	68.00 (+/- 0.1)
RandomForest	97.22	99.50	96.00 (+/- 0.1)
Voting Classifier	—	—	97.00 (+/- 0.1)

جدول ۲. دقت پیش بینی داده ها

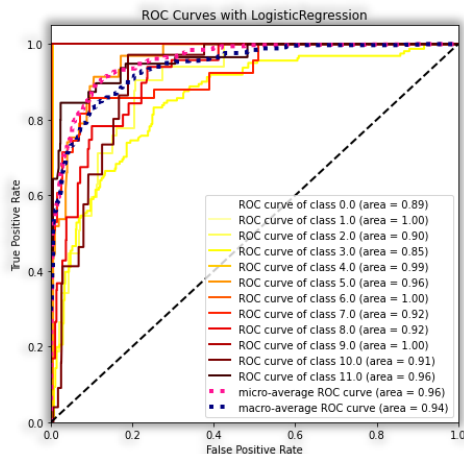
توزیع نشده است این مجموعه داده از نوع Imbalance data می باشد . این مسئله بر روی طبقه بندی و در نهایت پیش بینی اثر منفی می گذارد و دقت کار را کاهش می دهد . پس می توان داده ی بیشتری جمع کرد تا تعداد داده ها برابر شود اما چنین روشی واقعا زمانبر می باشد . برای رفع چنین مشکلی از resampling data استفاده می شود . بعد از آنکه داده ها encode شدند آن ها را به داده های train و test تبدیل می کنیم . برای این بخش از يك شبکه و چند الگوریتم را بررسی و آموزش داده ایم تا بهترین دقت را به دست بیاوریم . برای هر بخش دقت محاسبه می شود تا بتوان به راحتی مقایسه کرد . ابتدا از الگوریتم XGB classifier استفاده می شود . سپس Logistic reg و در نهایت Random forest و Voting classifier که تا کنون دو الگوریتمی که خوب پیش بینی کرده اند XGB Classifier و Random Forest هستند پس این دو را در این الگوریتم قرار داده می شوند تا بررسی شوند . حال از شبکه های عصبی مصنوعی استفاده می کنیم که دارای لایه های مختلف Convolution , Dense layer و Dropout هستند . این لایه ها را کنار هم و پشت سر هم می چینیم تا بتوانیم صفحات مختلف لایه ها را test و بررسی کنیم . در حقیقت در این بخش از ترکیب شبکه ی عصبی CNN با الگوریتم SVM استفاده کرده ایم . همان طور که گفته شد این روش هم ویژگی ها را از داده استخراج می کند و هم آن ها را کلاس بندی و پیش بینی می کند .

۴.نتایج

در جدول زیر می توانید میزان دقت هر يك از روش های به بررسی شده در مرحله ی طبقه بندی داده ها مشاهده کنید .

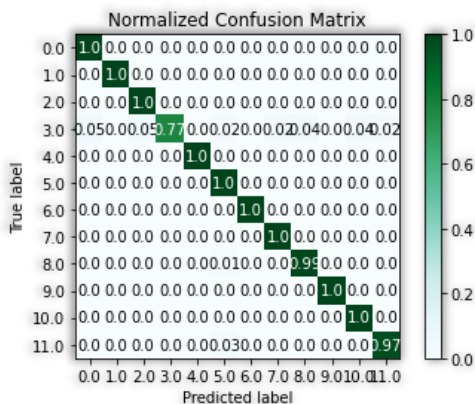
Algorithm	Test	Train
Random forest	50.67	99.93
Random forest with random search CV	48.00	93.14
SVM	57.33	100.00

جدول ۱. دقت کلاس بندی داده ها



شکل ۶. ROC Curved. الگوریتم Logistic Regression

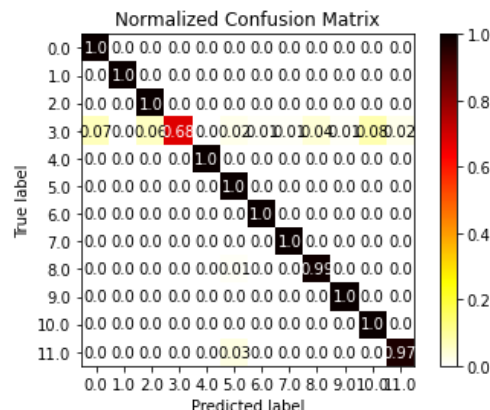
برای بررسی الگوریتم RandomForest در این بخش ماتریس Confusion ترسیم شده حاکی از دقت بالایی نتیجه‌ی این روش دارد:



شکل ۷. Confusion Matrix. الگوریتم RandomForest

در این روش با توجه به ماتریس اکثر پیش‌بینی‌ها برای ۱۲ کلاس با احتمال ۱۰۰٪ درست پیش‌بینی شده است ولی در برخی کلاس‌ها به طور مثال کلاس آخر با احتمال ۰/۰۳ خطا در پیش‌بینی ما وجود دارد. همچنین نشان می‌دهد که اگر کلاسی را پیش‌بینی می‌کنیم با چه احتمالی آن کلاس با کلاس دیگری اشتباه پیش‌بینی می‌شود. به عنوان مثال در کلاس آخر (کرم) احتمال ۳٪ نیز در نقطه‌ی کلاس پنجم که اسکراب است، داریم. این مسئله به این معنی است که با احتمال ۳٪ کرم را با اسکراب اشتباه پیش‌بینی می‌کنیم. اما این قضیه چه کمکی به ما می‌کند؟ به ما کمک می‌کند که زمانی که می‌خواهیم یک توصیه‌ی مراقبت زیبایی را به مشتری ارائه بدهیم، می‌توانیم بگوییم اگر من کرم را به شما

همان طور که ملاحظه می‌کنید الگوریتم XGB Classifier بسیار خوبی را به ما می‌دهد. برای بررسی بهتر نتایج خروجی هر یک از الگوریتم‌ها را تحلیل کرده ایم و نمودارهای زیر را ترسیم کرده ایم و دو معیار آنالیز و تحلیل را برای نتایج خروجی در نظر گرفته ایم. به طور مثال در بخش الگوریتم XGB Classifier برای بررسی بهتر یک Confusion Matrix را در نظر گرفته ایم:



شکل ۵. Confusion Matrix. الگوریتم XGB Classifier

به طور مثال اولین کلاس به نام facemask را اگر بخواهیم بررسی کنیم، خواهیم دید ۱/۰ درست پیش‌بینی کرده است. این ماتریس نشان می‌دهد با احتمال ۱۰۰٪ می‌تواند تمامی کلاس‌ها به غیر از کلاس سوم، هشتم و آخر را درست پیش‌بینی کند.

برای تحلیل نتیجه‌ی الگوریتم Logistic Regression از نمودار ROC Curved استفاده شده است. همان طور که ملاحظه می‌کنید اعداد خوبی در این بخش به دست نیامده است پس برای استفاده در سیستم اصلاً مناسب نمی‌باشد. همان طور که خطوط را در شکل مشاهده می‌کنید میانگین خطوط به سمت False Positive Rate متمایل شده‌اند و فقط کلاس‌های دوم، ششم و نهم عدد یک به دست آورده‌اند که نشان‌دهنده‌ی این است که تعداد کمی از کلاس‌ها صد درصد درست هستند.

۵. نتیجه گیری

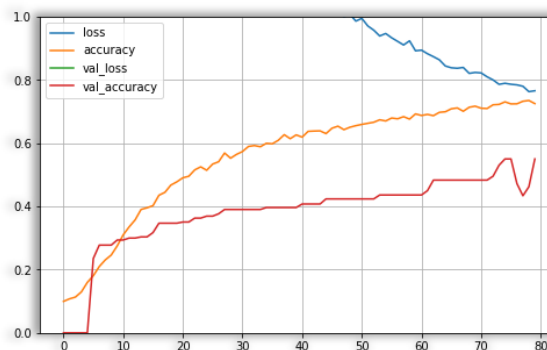
این پژوهش به منظور ارائه ی پیشنهاد توصیه ی مراقبت های زیبایی پوستی خانگی انجام شده است . در این مقاله به علت حجم بالایی داده ها و پردازش های مسائل به جای آنکه سیستم از یک منبع داده استفاده کند و روند پروژه به صورت متوالی پشت سر هم پیش برود ، هر بخش پروژه به صورت موازی و به طور جداگانه انجام شده سه مرحله ی طبقه بندی و کلاس بندی داده ها ، تخمین سن و جنسیت و پیش بینی توصیه مراقبت های پوستی ارائه شده است . و در هر مرحله الگوریتم های مختلف بررسی و آموزش داده شده است پس از این بررسی ها و آموزش های لازم ، مدل ترکیبی شبکه ی CNN – Resnet در بخش کلاس بندی و الگوریتم XGB Classifier در بخش پیش بینی به عنوان بهترین روش ها از لحاظ دقت نسبت به روش های دیگر انتخاب شد تا با دقت بالا نتیجه ارائه شود چون تمام اهداف پروژه در این است که با به کارگیری پارامترهای سن ، جنسیت و سطح شدت آکنه شخص کاربر یک پیشنهاد دهنده ی متناسب با پوست صورت وی را ارائه دهیم . مجموعه داده های Acne که حاوی ۳۳۶ تصویر اشخاص مختلف با سطوح مختلف شدت آکنه و dataset_facemask_with_target1 که حاوی ۸۴۹۶ داده از تمام توصیه های مراقبت های پوستی خانگی همراه با طرز تهیه ، نحوه ی مصرف و سن متناسب با آن ، با دقت بالا جمع آوری و مورد استفاده قرار گرفته شده است .

مراجع

۱. Wu, X., et al. *Joint Acne Image Grading and Counting via Label Distribution Learning*. in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019.
۲. T, Z., Z. H., and S.J. A, *Computer Vision Application for from Selfie Images. Assessing Facial Acne Severity*. 2019 Jul 18.
۳. Feldman, S., et al., *Diagnosis and treatment of acne*. ۲۰۰۴ American Family Physician, 69(9): p. 2123-2130.
۴. Yang, J., A. Bulat, and G. Tzimiropoulos. *Fan-face: a simple orthogonal improvement to deep face in Proceedings of the AAAI .recognition*. *Conference on Artificial Intelligence*. 2020.
۵. Lin, J., et al. *Face parsing with roi tanh-warping*. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.
۶. Park, S., A. Spurr, and O. Hilliges. *Deep pictorial gaze estimation*. in *Proceedings of the European .Conference on Computer Vision (ECCV)*. 2018.

پیشنهاد می کنم ، چنین پیشنهادی با دقت % ۹۷ می باشد و اگر می خواهید انتخاب دومی نیز داشته باشید می توانید به سراغ اسکراب بروید .

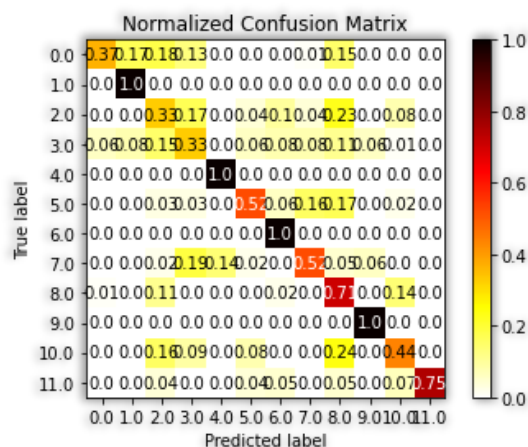
حال با پشت سر هم گذاشتن چند لایه شبکه ی عصبی مصنوعی ساختیم به طوری که بتوان لایه ها را به ترتیب تست کرد . پس از کامپایل کردن مدل با جلو رفتن epoch ها دقت حدود % ۸۳ به دست آمد که حتی از مدل های قبلی نیز کمتر شد :



شکل ۸. نمودار دقت و خطای الگوریتم یادگیری عمیق

همان طور که ملاحظه می کنید ، با ۸۰ epoch میزان درصد خطا در حال کاهش است و دقت نیز در حال افزایش است اما باز هم میزان خطا از دقت بیشتر است و باید epoch های بیشتری را برای بررسی در نظر گرفت .

سپس از یک الگوریتم ترکیبی دیگر به نام CNN – SVM استفاده شده است . دقت به دست آمده از این شبکه که حدوداً % ۷۲/۴۸ نیز نسبت به متد های دیگر کم است . با کمک ماتریس Confusion دقت این متد را بررسی کرده ایم :



شکل ۹. Confusion Matrix الگوریتم یادگیری عمیق