



دانشگاه صنعتی امیر کبیر
(پلی تکنیک تهران)

گزارش پروژه درس تصویر پردازش رقمی

استاد درس جناب آقای دکتر رحمتی

محسن عبادپور

شماره دانشجویی: ۴۰۰۱۳۱۰۸۰

نیمسال دوم سال تحصیلی ۱۴۰۰-۱۴۰۱

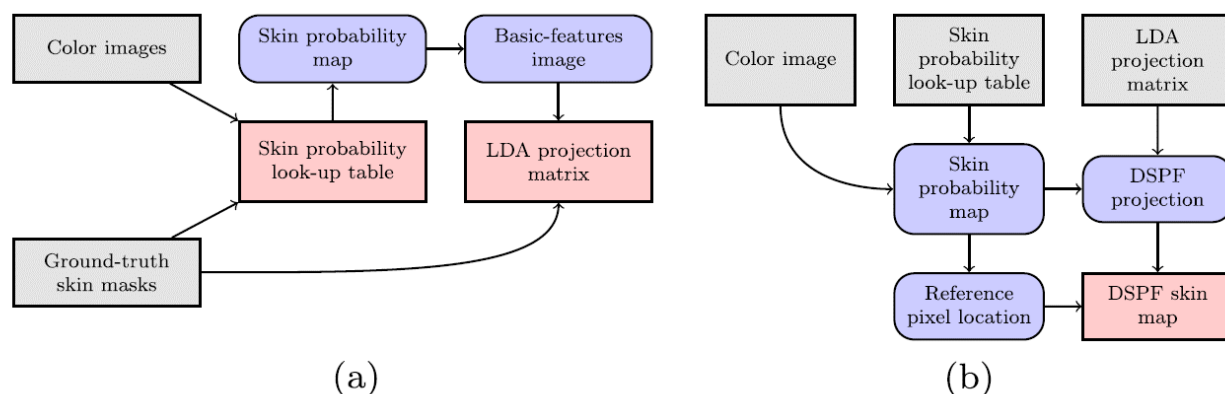
فهرست مطالب

۱- مقدمه	۳
۲- تئوری و روش ارائه شده	۳
۲-۱- آموزش دسته بند بیز و اعمال فیلتر های مکانی	۳
۲-۲- تحلیل تمایز گر خطی (LDA) و فضای DSPF	۴
۲-۳- آنالیز مکانی پیسکل های پوستی با بهره از تبدیل فاصله (distance transform)	۵
۳- پیاده سازی و آموزش	۶
۳-۱- مجموعه داده و پیش پردازش های اعمالی	۶
۳-۲- آموزش دسته بند بیز و محاسبه BFIs	۶
۳-۳- اعمال LDA و محاسبه DSPFs	۷
۳-۴- محاسبه DSPF skin map و آنالیز مکانی آن (DTCD)	۸
۴- نتایج	۹
۴-۱- تحلیل عددی نتایج	۹
۴-۱-۱- بحث و تحلیل بصری نتایج	۱۰

تشخیص ناحیه پوستی انسان در کاربرد های مختلف می تواند حائز اهمیت باشد که از آن میتوان به مباحثی چون تشخیص بیماری های پوستی، سانسور و دسته بندی تصاویر پورنوگرافی و... اشاره نمود. بنده در زمان جست و جو موضوع انتخابی را با کلیدواژه های فیلتر های مکانی (Spatial Filters) و تشخیص پوست (Skin detection) یافتیم که هدف آن ادغام دسته بند بیز، فیلتر های مکانی آماری، تحلیل تمایز گر خطی (LDA) و آنالیز مکانی با بهره از تبدیل فاصله (distance transform) برای تشخیص ناحیه پوستی می باشد؛ در ادامه و گام به گام مراحل مقاله و پیاده سازی مورد اشاره قرار خواهد گرفت.

۲- تئوری و روش ارائه شده

روش پیشنهادی در این مقاله مشتمل بر سه مبحث می باشد که در ادامه هر کدام مورد بررسی قرار خواهد گرفت؛ ضمناً فلوچارت روش پیشنهادی در شکل زیر قابل ملاحظه است. فاز a مربوط به فاز آموزش و فاز b مربوط مرحله شناسایی می باشد:



۲-۱- آموزش دسته بند بیز و اعمال فیلتر های مکانی

اولین مرحله آموزش یک دسته بند بیز دودویی مبتنی بر هیستوگرام تصاویر آموزشی می باشد که رابطه آن در زیر آورده شده و در درس شناسایی آماری الگو نیز مورد بررسی قرار گرفته است؛ حاصل این آموزش استخراج یک look-up table (LUT) می باشد که بوسیله آن میتوان به ازای پیکسل های ورودی احتمال قرار گرفتن شان در کلاس پوست را بدست آورد؛ اگر این فرآیند را برای همه پیکسل ها یک تصویر انجام دهیم صفحه احتمال پوستی (skin prob. map) متناظر با آن بدست خواهد آمد که برای هر پیکسل متناظر یک ویژگی برای دسته بندی محسوب می شود.

$$P(C_s|v) = \frac{P(v|C_s)P(C_s)}{P(v|C_s)P(C_s) + P(v|C_{ns})P(C_{ns})}$$

حال چهار فیلتر مکانی آماری شامل کمینه، میانه، انحراف معیار و اختلاف مقدار کمینه و بیشینه بروی skin prob. map با سه سایز کرنل مختلف ۵، ۹ و ۱۳ اعمال می‌شود تا ۱۲ ویژگی جدید متناظر با هر پیکسل حاصل شود و با skin prob. Map در مجموع ۱۳ ویژگی آماده پردازش شود. این ۱۳ ویژگی، ویژگی‌های پایه تصویر (BIFs) basic features image نامیده می‌شود.

۲-۲- تحلیل تمایزگر خطی (LDA) و فضای DSPF

همانگونه که در درس شناسایی آماری الگو مورد بحث قرار گرفته است LDA یک روش با نظارت خطی و مبتنی بر دادگان پراکندگی کلاس‌ها برای تبدیل فضای ویژگی داده‌ها به فضایی است که تمایزپذیری بین کلاسی افزایش و داخل کلاسی کاهش یابد؛ برای این منظور دو ماتریس within(Sw) و between(Sb) محاسبه شده و سپس با استفاده از بردارهای ویژه ماتریس $Sw^{-1}Sb$ عملیات تبدیل انجام می‌پذیرد؛ روابط موجود در زیر آورده شده است:

$$S_B = \sum_{i=1}^K (\mu_i - \mu)(\mu_i - \mu)^T \quad S_W = \sum_{i=1}^K \sum_{\mathbf{u}_k \in K_i} (\mathbf{u}_k - \mu_i)(\mathbf{u}_k - \mu_i)^T$$

برای محاسبات بهتر، بردارهای ویژه در یک ماتریس (LDA Projection Matrix) کنار یکدیگر قرار گرفته و در BIFs های مرحله قبل ضرب میشود تا داده‌ها در فضای جدید حاصل شود؛ در فضای جدید ابعادی که حاصل از بردار ویژه با مقدار ویژه بزرگتر می‌باشد تفکیک پذیری بهتری را ارائه می‌کند لذا در این مرحله میتوان از تعداد اندکی از فضای ویژگی جدید استفاده نمود. (کوچک‌تر مساوی ۱۳) فضای ویژگی جدید discriminative skin-presence features (DSPF) نامیده می‌شود.

اگر همه‌ی دو مرحله فوق را به جای skin prob. map با luminance تصویر انجام میدادیم ویژگی‌های حاصل را DSPF-L خواهیم نامید و ابعاد می‌تواند کوچکتر مساوی ۱۳ باشد؛ همچنین اگر هر دو را توأمان اعمال نماییم ویژگی‌های حاصل را DSPF-L+ خواهیم نامید و ابعاد میتواند کوچکتر مساوی ۲۵ باشد. روش پیشنهادی اصلی بر اساس ویژگی‌های DSPF می‌باشد، لذا ادامه بحث و پیاده سازی نیز بر اساس همین فرض می‌شود.

حال بر اساس DSPF بایستی صفحه پوستی (DSPF skin map) متناظر با آن را بدست آورد. برای این منظور از رابطه زیر استفاده می‌شود که در آن V_i بعد lam بردار DSPF می‌باشد و x متناظر با هر کدام از پیکسل‌ها بوده و r نیز پیکسل مرجع برای تشخیص پوست می‌باشد. محاسبه پیکسل مرجع بدین نحو محاسبه می‌شود که عمل erosion با سایز ۱۵ (بزرگتر از سایر کرنل‌های مورد استفاده در BIFs) روی skin prob. map اعمال می‌شود و سپس پیکسلی که بالاترین احتمال برای تخصیص به کلاس پوست می‌باشد استخراج می‌شود؛ بردار متناظر با آن پیکسل در DSPF به عنوان مرجع (r) برای محاسبه قرار می‌گیرد؛ DSPF skin map به عبارتی نشان‌دهنده‌ی فاصله‌ی پیکسل‌ها با محتمل‌ترین پیکسل پوستی در فضای DSPF می‌باشد (همان D در رابطه زیر).

$$D(x) = \left[\sum_{i=1}^m \left(v_i^{(x)} - v_i^{(r)} \right)^2 \right]^{1/2}$$

در نهایت صفحه پوستی حاصل را توسط بزرگترین پیکسل نرمال می کنیم (تا D_N حاصل شود). حال با تنظیم یک حد آستانه می توان با DSPF skin map پیکسل های پوستی را پیش بینی نمود که تا همین حد نیز خطای تشخیص پیکسل های پوستی نسبت به skin prob. map پیشفرض کاهش می یابد و این نشان میتوان که صفحه پوستی حاصل جداپذیری خوبی بین پیکسل ها دو کلاس ایجاد می کند اما میتوان با بهره از آنالیز مکانی که در بخش بعد توضیح داده می شود، میتوان همچنان نتیجه را ارتقا داد که در پیاده سازی نیز مرحله بعد نیز انجام شده است.

۲-۳- آنالیز مکانی پیکسل های پوستی با بهره از تبدیل فاصله (distance transform)

در مقاله دیگر و قبل تری از نویسندگان همین مقاله (لینک) یک روش انتشار محور و بر اساس تبدیل ارائه شده است و هدف آن انتشار ناحیه پوستی از پیکسل های با احتمال بسیار بالا به سایر پیکسل ها می باشد. (به عبارتی بنده دو مقاله را پیاده سازی نموده ام :)

برای این منظور ابتدا پیکسل هایی که احتمال بسیار بالایی برای تخصیص به کلاس پوست می باشد را به عنوان دانه (seed) انتخاب می کنیم سپس دانه کنار های نواحی همبندی را به وجود می آورند. از ناحیه هایی که تعداد پیکسل های آن کمتر از ۱۰٪ بزرگترین ناحیه باشد صرف نظر می کنیم. سپس به ازای هر کدام از پیکسل های غیر دانه، نزدیک ترین مسیر به یک دانه ی مفروض محاسبه شده و روابط زیر برای آن و در قبال پیکسل های مسیر (۰: دانه مرزی تا L-1: پیکسل مجاور غیردانه) محاسبه می شود که در آن H مربوط به کانال H در فضای رنگی HSV و Y مربوط به luminance تصویر می باشد و a نیز یک ضریب پنالتی به ازای همسایه های قطری و معادل با رادیکال (۲) می باشد.

$$C(x) = \sum_{i=0}^{l-1} \rho(p_i \rightarrow p_{i+1})$$

$$\rho(x \rightarrow y) = \rho_l(x, y) \cdot [1 + \rho_p(x \rightarrow y)],$$

$$\rho_l(x, y) = \alpha_{diag} \cdot (|Y(x) - Y(y)| + |H(x) - H(y)|),$$

$$\rho_p(x \rightarrow y) = \begin{cases} 1 - D_N(y) & \text{for } D_N(y) > P_\beta, \\ \infty & \text{for } D_N(y) \leq P_\beta. \end{cases}$$

تابع C به ازای هر کدام از پیکسل های موجود در تصویر که غیر دانه هستند محاسبه می شود تا در نهایت صفحه پوستی جدید با عنوان distance transform in a combined domain (DTCD) که مبتنی بر فاصله در DSPF skin map می باشد حاصل شود و پیش بینی و تنظیم حد آستانه در این حالت انجام پذیرد.

۳- پیاده سازی و آموزش

این پیاده سازی بصورت بهینه و چند رشته ای (multi-thread) انجام شده است تا زمان مورد نیاز برای اجرا به حداقل ترین حالت برسد.

۳-۱- مجموعه داده و پیش پردازش های اعمالی

در مقاله از دو مجموعه داده استفاده شده است. اولی و اصلی با عنوان ECU می باشد که مربوط به تشخیص نواحی پوستی بوده و شامل صحنه های مختلف می باشد. برای دریافت این دیتاست بایستی یک فرم موافقت با خط مشی های نویسنده پر شده و ایمیل می شد که بنده نیز چنین کردم اما با وجود دوبار یادآوری پاسخی دریافت نکردم. دیتاست دوم مربوط به تشخیص حالت دست با عنوان HGR می باشد که عمومی بوده و شامل ۸۹۹ تصویر می باشد که من نیز آن را دانلود کرده و از آن برای پیاده سازی مدل استفاده کرده ام.

برای دیتاست جایگزین برای ECU با هدف تشخیص ناحیه پوستی که بصورت پیکسلی نبوده و grand truth مکانی آن موجود باشد (تا بتوان BFIs بدست آورد) جست و جو کردم که صرفا به یک دیتاست با عنوان Pratheepan رسیدم که تعداد آن کلا در حد ۵۰ تصویر بسیار چالشی بوده و توجیه آن برای آموزش و تست روی این روش وجود نداشت چرا که در توضیحات و در فروم ها اشاره شده بود که این دیتاست مناسب آموزش نبوده و صرفا مناسب بنچمارک روش های پیشنهادی می باشد.

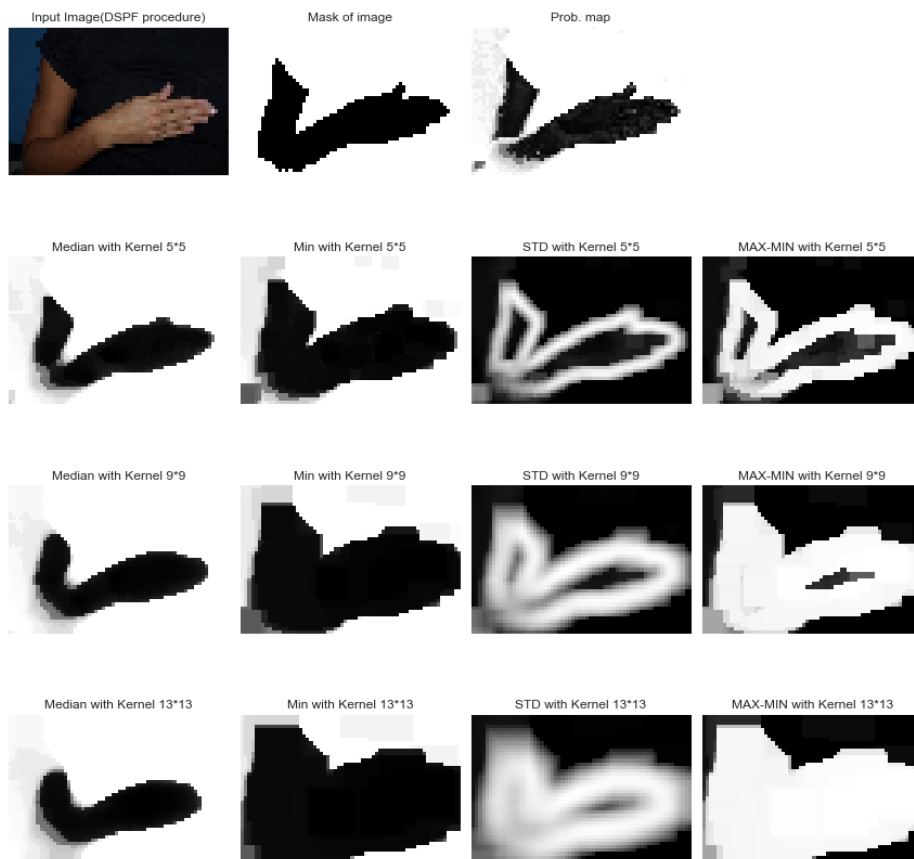
در مقاله برای اینکه حجم پردازش ها بالا نرود، از تصاویر و از نواحی با نرخ ۱/۱۵ نمونه برداری شده و ابعاد تصاویر کاهش یافته است. بنده در پیاده سازی خود از نرخ ۱/۱۰ استفاده کرده و تصاویر را نسبت به مقاله بزرگتر بدست آورده ام. (از جایی که چند رشته ای پیاده سازی کرده ام نگرانی چندانی در خصوص این افزایش وجود نداشت)

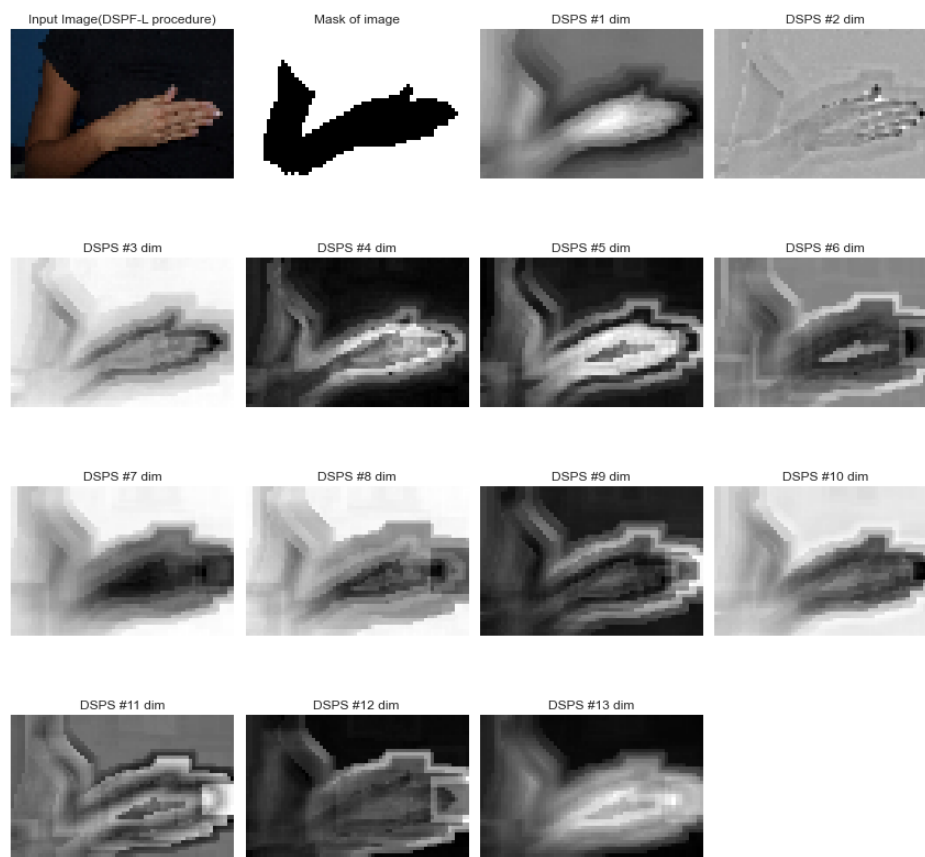
۳-۲- آموزش دسته بند بیز و محاسبه BFIs

طبق توضیحات ارائه شده در مقاله، برای محاسبه هیستوگرام و احتمال های مربوط به بیز از تعداد bins برابر با ۶۴ استفاده شده و طبق رابطه ذکر شده در بخش ۲-۱ فرآیند تولید بردار ویژگی های پایه ی تصویر (BFIs) انجام می شود. برای این بخش نیز پیاده سازی بصورت چند رشته ای انجام و برای دو نمونه تصویر محاسبه شده ی آن بصورت صفحه بعد حاصل شده است. برای چگونگی ارتباط بین احتمال سه کانال رنگی RGB توضیحی در متن مقاله صورت نگرفته است لذا ضرب احتمال سه کانال به عنوان خروجی دسته بند بیز فرض شده است.

۳-۳- اعمال LDA و محاسبه DSPFs

ماتریس $Sw^{-1}Sb$ و ماتریس متشکل از بردار های ویژه آن (مرتب نزولی بر اساس مقدار ویژه متناظر) محاسبه شده و ماتریس نگاشت LDA projection matrix حاصل شده است. حال فضای BFIs توسط این ماتریس به فضای DSPS نگاشت می شود که خروجی آن برای نمونه تصویر مرحله قبل در صفحه بعد قابل ملاحظه است. همانطور که در تصویر قابل ملاحظه بوده و در مقاله نیز ذکر شده است صرفاً ۱ یا ۲ بُعد ابتدایی که متناظر با دو بردار ویژه اولی می باشد تمایز پذیری خوبی بین پیکسل های پوستی و غیرپوستی ایجاد می کند و بهتر است از همه ۱۳ بُعد جدید استفاده نشود. همانند پیاده سازی مقاله، از دو بُعد ابتدایی برای ادامه ی روش پیشنهادی استفاده شده است.



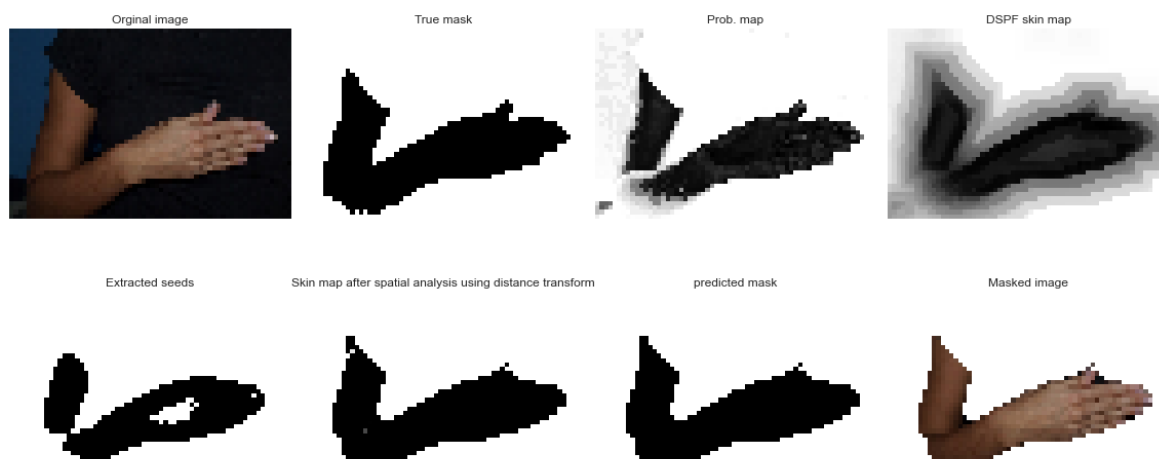


۳-۴- محاسبه DSPF skin map و آنالیز مکانی آن(DTCD)

طبق توضیحات مقاله برای محاسبه DSPF skin map ابتدا پیکسل مرجع محاسبه شده و طبق رابطه دوم در ۲-۲ اقدام شده (D) و سپس نرمال سازی اتفاق افتاده است (D_N)؛ لازم به ذکر مجدد است که همانند مقاله از دو بُعد ابتدایی DSPF استفاده شده است. حال بایستی DTCD روی DSPF skin map اعمال شود؛ برای این محاسبه نیاز به سه ثابت می باشد که اولی مربوط به دودویی کردن می باشد (P_a)، دومی مربوط به رابطه ۲-۳ می باشد (P_b) و سومی نیز مربوط به کمینه بودن پیکسل مرجع (P_{ref}) می باشد؛ ثابت سوم شرط باعث می شود تصاویر دنیای واقعی که هیچ ناحیه پوستی در آن وجود ندارد به درستی پردازش و پیش بینی شود.

مقادیر استفاده شده در مقاله که بصورت تجربی تعیین شده و به ازای نرخ نمونه برداری ۱۵/۱ می باشد بترتیب ۰.۹۶، ۰.۳ و ۰.۶ می باشد؛ در این پیاده سازی من این ثابت ها را تعیین کرده و سپس شروع به تغییرات در آن رنج کردم تا بتوانم خطای نهایی را کمینه کنم. در پیاده سازی P_{ref} را کلاً لحاظ نکردم (چون در دیتاست ما همه تصاویر دارای ناحیه پوستی بوده و تصاویر کنترل شده می باشد و بود/نبود این آستانه معنی ندارد) همچنین در این پیاده سازی برای P_a و P_b بترتیب مقادیر ۰.۸۵ و ۰.۳۴ با آزمایش های تجربی بدست آورده و استفاده شده است بگونه ای که خطا کمینه شود که در قسمت نتیجه ذکر شده است. (لازم به ذکر است که بالا بودن P_a باعث می شود اطمینان کنیم که انتشار ناحیه پوستی از محل های درستی انجام شده و نرخ False Positive نیز مینیمم شود)

با استفاده از P_a DSPF skin map دودوئی شده و دانه ها (seed) ها استخراج شده است و سپس از نواحی کمتر از ۱۰٪ بزرگترین ناحیهی دانه ها صرف نظر شده است. سپس انتشار ناحیه پوستی طبق روابط ۲-۳ روی آن اعمال شده و در نهایت mask پیش بینی شده بدست آورده شده است. برای تصویر نمونه این مرحله نیز خروجی گرفته شده و در زیر قابل ملاحظه می باشد:



۴- نتایج

در این قسمت نتایج حاصل از پیاده سازی مورد تحلیل قرار گرفته و همچنین تاثیر روش پیشنهادی بصورت بصری آنالیز خواهد شد.

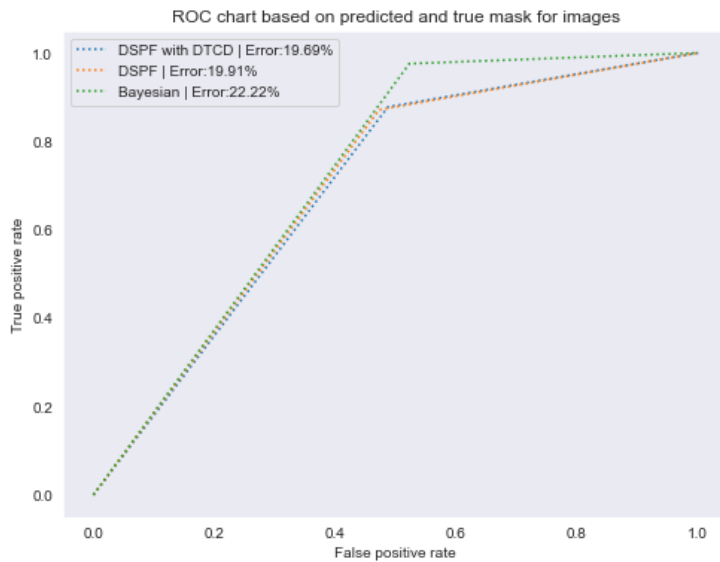
۴-۱- تحلیل عددی نتایج

در این قسمت همانند مقاله مقدار خطا برای حالت اصلی محاسبه شده است. منظور از خطا حاصل جمع مقادیر False و False Negative Positive به ازای پیکسل های پردازشی می باشد. مقدار خطای گزارش شده در مقاله برای این دیتاست (HGR) برابر با ۷۶٪ می باشد که در این پیاده سازی خطای حاصل برای همین دیتاست برابر با ۱۳.۴٪ نتیجه شده است که در حد ۶ درصد با نتیجه گزارش شده متفاوت است. اما ترتیب و نسبت نتایج حاصل در حالت دسته بندی با بیز و DSPF بدون آنالیز مکانی همانند گزارش مقاله حاصل شده است طوری

که بهترین حالت مربوط به DSPF با آنالیز مکانی بوده و سپس با اختلاف اندک DSPF بدون آنالیز مکانی بوده و با اختلاف بیشتر نیز دسته بند بیز ظاهر شده است. نمودار ROC نیز بصورت مقابل حاصل می شود:



حال اگر به جای DSPF از DSFP-L نیز استفاده نماییم، می بینیم که خطای مدل افزایش می یابد(همانند گزارش مقاله) ولی همچنان ترتیب نتایج ثابت می باشد؛ در چارت زیر نمودار ROC به همراه خطاهای هر یک آمده است:



۴-۱- بحث و تحلیل بصری نتایج

همانطور که در تصویر صفحه قبل مشخص است، گذاشتن حد آستانه‌ی احتمالی به تنهایی بروی خروجی دسته بند بیز میتواند خطای بیشتری داشته باشد بطوری که قسمت های آرنج فرد میتواند نادیده گرفته شود(فارغ از اینکه تعیین حد آستانه خود به تنهایی جای بحث فراوان دارد). حال اعمال و بدست آوردن DSPF skin map باعث می شود گسیختگی بین نواحی پوستی(قسمت تاشو آرنج) اندکی هموارتر شود و بتوان نتیجه بهتری یافت. اما در صورتی که DTCD را روی DSPF skin map اعمال نماییم می بینیم که دانه های استخراجی همگی از پیسکل ها حتمی پوست بوده و انتشار ناحیه پوستی باعث می شود نرخ FP کاهش یافته و گسیختگی بین نواحی نیز پر شود و نتیجه نیز همچنان بهتر شود(هرچند اندک). به نظر بنده ایده کلی روش پیشنهادی در این مقاله مبتنی بر کاهش نرخ FP می باشد و توانسته در زمان خود(۲۰۱۳) بهبود مناسبی را ایجاد نماید.