

تشخیص هوشمند نواحی نیازمند درمان زیبایی و شبیه‌سازی نتایج درمان با استفاده از شبکه‌های عصبی

کانلوشنی و شبکه‌های مولد تخصصی

آرمان زواری^۱، حمید سلیمانی مهر^{۲،۳*}

^۱دانشجوی کارشناسی ارشد، گروه مهندسی مکترونیک، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

^۲دانشیار، گروه مهندسی مکانیک، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

^۳دانشیار، مرکز تحقیقات نوین خودرو، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

مقاله پژوهشی

چکیده

تاریخ دریافت:

۱۴۰۴/۰۷/۱۳

تاریخ پذیرش:

۱۴۰۴/۱۲/۰۲

کلیدواژه‌ها:

هوش مصنوعی، یادگیری عمیق، درماتولوژی زیبایی، شبکه عصبی کانلوشنی، شبکه‌های مولد تخصصی

نویسنده مسئول:

ha.sol@iau.ac.ir

امروزه تقاضا برای تحلیل دقیق وضعیت پوستی و پیش‌بینی نتایج درمان‌های زیبایی پیش از اجرا، به یکی از نیازهای مهم در حوزه درماتولوژی و پزشکی زیبایی تبدیل شده است. در این مقاله، یک چارچوب هوشمند مبتنی بر یادگیری عمیق به منظور تشخیص نواحی نیازمند درمان زیبایی در چهره افراد و شبیه‌سازی تصویری نتایج درمان توسعه یافته است. برای این منظور، نخست شبکه عصبی کانلوشنی جهت شناسایی خودکار ناهنجاری‌هایی چون چین‌وچروک، لکه‌های پوستی و عدم تقارن به کار رفت. در گام بعد، با بهره‌گیری از مدل پیکس‌تویپکس-گن، تصویر چهره پس از درمان به صورت واقع‌نما شبیه‌سازی شد. مدل پیشنهادی روی یک مجموعه داده شامل ۵۰۰۰ تصویر چهره از پایگاه‌های داده معتبر سلب‌ای، اف‌اف‌اچ‌کیو و درم‌نت که دارای برچسب‌گذاری توسط متخصصان درماتولوژی برای نواحی آسیب‌دیده بودند، آموزش دید. فرایند ارزیابی، متکی بر شاخص‌های کمی شامل دقت ۹۴/۲ درصد، اف‌وان-اسکور ۹۳/۱ درصد، میانگین قدرمطلق خطا برابر ۰/۰۲۳ و نسبت اوج سیگنال به نویز معادل ۳۱/۵ دسی‌بل به همراه امتیاز رضایت بخشی ۴/۶ از ۵ توسط متخصصان انجام شد. نتایج حاصل نشان می‌دهد که چارچوب ترکیبی سی‌ان‌ان-گن، علاوه بر شناسایی دقیق نواحی نیازمند درمان، قادر است تصویری واقع‌گرایانه از نتیجه درمان ارائه دهد.

doi : 10.22034/ABMIR.2026.23775.1176

۱- مقدمه

تشخیص الگوهای پوستی عملکرد قابل توجهی ارائه دهد. مدل‌های سی ان در تحلیل بافت پوست، تشخیص چین و چروک‌ها، لکه‌ها و دیگر ویژگی‌های ظاهری مؤثر در تشخیص نیازهای درمانی نقش اساسی دارند. در کنار آن، شبکه‌های مولد تخصصی توانسته‌اند تحولی در شبیه‌سازی نتایج درمان ایجاد کنند [۶]. این مدل‌ها قادرند بر اساس تصویر اولیه، نمونه‌ای واقع‌نما از وضعیت پوست پس از درمان را تولید کنند و بدین ترتیب در تصمیم‌گیری پزشک و اعتماد بیمار تأثیر بسزایی بگذارند [۷].

با این وجود، بررسی کارهای پیشین نشان می‌دهد که اکثر مطالعات موجود یا بر تشخیص ناهنجاری‌های پوستی با استفاده از سی ان ان متمرکز شده‌اند و یا صرفاً بر تولید تصاویر مصنوعی از طریق گن تأکید داشته‌اند، اما ترکیب منحصربه‌فرد این دو رویکرد با در نظرگیری ویژگی‌های بالینی فردی در یک چارچوب یکپارچه کمتر مورد بررسی قرار گرفته است [۸]. همچنین، فقدان در نظر گرفتن ویژگی‌های بالینی فرد (مانند نوع پوست، سن و سابقه درمانی) موجب شده تا بسیاری از سیستم‌ها قابلیت شخصی‌سازی درمان را نداشته باشند [۹]. علاوه بر این، در بسیاری از مقاله‌ها، ارزیابی معیارهای چندبعدی مانند میانگین قدرمطلق خطا^۳ یا نسبت اوج سیگنال به نویز تحلیل بالینی خروجی‌ها در نظر گرفته نشده است. ارزیابی زیبایی چهره علاوه بر معیارهای بالینی (مانند میزان چین و چروک، بافت پوست و یکنواختی رنگ)، نیازمند درک پیچیده‌ای از استانداردهای فرهنگی و زیبایی‌شناسی نیز هست. از این رو، سیستم‌های هوش مصنوعی در این حوزه باید قادر باشند مفاهیم ادراکی پیچیده‌ای را مدل‌سازی کنند که فراتر از طبقه‌بندی‌های دودویی ساده باشد.

نوآوری مقاله حاضر در ارائه یک چارچوب ترکیبی منحصربه‌فرد است که برای نخستین بار سی ان و گن را با در نظرگیری پارامترهای بالینی فردی (نوع پوست، سن، سابقه درمانی) در یک سیستم یکپارچه تلفیق می‌کند. در این چارچوب، ابتدا با استفاده از

زیبایی همواره یکی از مهم‌ترین دغدغه‌های بشر بوده است و نقشی اساسی در اعتمادبه‌نفس، روابط اجتماعی و حتی کیفیت زندگی ایفا می‌کند. امروزه اهمیت زیبایی در حوزه‌های پزشکی و روان‌شناسی نیز مورد توجه ویژه قرار گرفته و به‌ویژه در درماتولوژی و پزشکی زیبایی یکی از معیارهای کلیدی در تصمیم‌گیری‌های درمانی محسوب می‌شود [۱].

روش‌های ارزیابی و تحلیل زیبایی در طول زمان دستخوش تحول شده‌اند. در گذشته بیشتر بر پایه ارزیابی ذهنی متخصصان و بیماران بود که وابستگی بالایی به تجربه فردی داشت. در ادامه، مدل‌های هندسی و ریاضی مانند استفاده از «نسبت طلایی» و تحلیل‌های مبتنی بر فاصله‌ها و زوایای صورت معرفی شدند، اما این روش‌ها قادر نبودند تنوع و پیچیدگی واقعی چهره انسان را به‌طور کامل بازتاب دهند [۲].

در سال‌های اخیر با پیشرفت فناوری‌های هوش مصنوعی^۱، بینایی ماشین و مدل‌های یادگیری ماشین، به‌ویژه یادگیری عمیق^۲، امکان تحلیل خودکار و دقیق تصاویر چهره را فراهم کرده‌اند و جایگزین مناسبی برای روش‌های سنتی محسوب می‌شوند [۳].

در دهه اخیر، پیشرفت‌های چشمگیر در حوزه هوش مصنوعی و یادگیری عمیق انقلابی در حوزه‌های مختلف سلامت، به‌ویژه درماتولوژی و پزشکی زیبایی ایجاد کرده‌اند [۴]. با افزایش تقاضای عمومی برای خدمات زیبایی و تمایل افراد به ارزیابی نتایج درمان قبل از اجرای آن، نیاز به سیستم‌های دقیق و قابل اعتماد برای تحلیل وضعیت پوست و ارائه پیشنهادها درمانی پیش‌ازپیش احساس می‌شود. مدل‌های سنتی تشخیص زیبایی که به مشاهده بالینی صرف متکی هستند، اغلب با محدودیت‌هایی مانند خطای انسانی، تفاوت دیدگاه پزشکان و عدم امکان پیش‌بینی تصویری نتایج مواجه هستند [۵].

یادگیری عمیق، به‌ویژه شبکه‌های عصبی کانولوشنی، به‌عنوان ابزار اصلی تحلیل تصاویر پزشکی توانسته است در مسائل طبقه‌بندی و

³ MAE

¹Artificial Intelligence

² Deep Learning



تأکید داشتند که سی ان ان‌ها در طبقه‌بندی دقیق بیماری‌های پوستی موفق هستند اما در شرایط واقعی عملکردشان به شدت به کیفیت داده‌ها وابسته است.

در تحقیق توکگوز و کارو [۲] به بررسی کاربردهای هوش مصنوعی، یادگیری ماشین و یادگیری عمیق در حوزه جراحی پلاستیک و بازسازی صورت پرداخته شده است. نویسندگان در این فصل مروری، مدل‌های مختلفی مانند شبکه‌های عصبی کانولوشنی برای تحلیل تصاویر قبل و بعد از عمل، شبکه‌های مولد تخصصی برای شبیه‌سازی نتایج جراحی، و مدل‌های یادگیری انتقالی برای بهبود دقت در پایگاه‌های داده محدود را معرفی و تحلیل کرده‌اند. این مطالعه نشان می‌دهد که استفاده از چنین مدل‌هایی می‌تواند در پیش‌بینی نتایج درمان، شبیه‌سازی تغییرات ظاهری، و کمک به تصمیم‌گیری پزشکان بسیار مؤثر باشد، اما خود فصل مقاله مدل اختصاصی توسعه نداده و صرفاً یک مرور تحلیلی از مدل‌های موجود ارائه کرده است. این مطالعه علی‌رغم افزایش سرعت پردازش و کاهش پیچیدگی زمانی، نشان‌دهنده پتانسیل ترکیب داده‌های فردی با تحلیل تصویری است ولی هنوز به سطح توصیه‌های درمانی نرسیده است.

در مطالعه گالیچ و همکاران [۳] مروری جامع از روش‌های یادگیری ماشین و یادگیری عمیق در حوزه تحلیل تصاویر پزشکی ارائه شده است که معماری‌ها، وظایف متداول (طبقه‌بندی، تشخیص / لوکالیزاسیون، سگمنتیشن، تولید و ثبت تصاویر) و مجموعه داده‌های عمومی مرتبط را پوشش می‌دهد نویسندگان در این مرقومه تأکید می‌کنند که این روش‌ها می‌توانند پتانسیل قابل توجهی برای توانمندسازی پزشکی شخصی‌سازی شده فراهم آورند؛ با این حال مقاله صرفاً یک مرور تحلیلی و تبیینی است و خود مقاله مدل تجربی یا چارچوب شبیه‌سازی درمانی ویژه‌ای برای تحلیل صورت ارائه یا ارزیابی نمی‌کند. با این حال روش پیشنهادی افزایش سرعت پردازش و کاهش نسبی پیچیدگی زمانی را به دنبال دارد.

در حوزه شبیه‌سازی درمان، سانتوس و همکاران [۸] از ترکیب مدل‌های گن با شبکه‌های آرانا برای پیش‌بینی اثرات درمان بر پوست استفاده کردند. اگرچه این مدل توانست تصاویر

سی ان ان، نواحی دارای نیاز درمانی به‌طور دقیق شناسایی شده و سپس با بهره‌گیری از گن، تصویر خروجی پس از درمان به صورت واقع‌نما شبیه‌سازی می‌گردد. مهم‌تر از آن، تصمیمات درمانی در این مقاله با اتکا بر ویژگی‌های بالینی فردی انجام می‌شود که گامی اساسی در مسیر توسعه سیستم‌های شخصی‌سازی شده در حوزه زیبایی دیجیتال به شمار می‌آید.

این رویکرد ترکیبی، پلی میان تشخیص و تجسم نتایج احتمالی درمان ایجاد می‌کند و در نتیجه، موجب افزایش درک متقابل میان پزشک و بیمار، و ارتقای فرایند تصمیم‌گیری بالینی می‌شود. مقاله حاضر با تکیه بر ساختاری ترکیبی، تحلیل‌های آماری دقیق و ارزیابی کیفی و کمی مدل، سعی دارد گامی نوین در جهت توسعه سامانه‌های هوشمند در حوزه پزشکی زیبایی بردارد و با بهره‌گیری از قدرت مدل‌های یادگیری عمیق، چشم‌انداز روشنی را برای کاربرد هوش مصنوعی در تشخیص و درمان پوستی ترسیم نماید. بر همین اساس، مقاله حاضر با هدف پرکردن این خلأ، یک مدل نوین مبتنی بر یادگیری عمیق و ترکیب شبکه‌های عصبی ارائه می‌دهد که می‌تواند به‌طور خودکار نیاز به درمان‌های زیبایی در نواحی صورت را شناسایی و شبیه‌سازی کند.

۲- مرور ادبیات و کارهای مرتبط

رشد سریع الگوریتم‌های یادگیری عمیق به‌ویژه در زمینه بینایی ماشین، تحولی چشمگیر در حوزه تشخیص پزشکی و تحلیل چهره ایجاد کرده است. مقاله‌های متعددی طی پنج سال اخیر بر روی کاربردهای هوش مصنوعی در درمان‌ولوژی زیبایی و شناسایی خودکار ویژگی‌های پوستی متمرکز شده‌اند، اما بیشتر آن‌ها بر یک بعد خاص تمرکز داشته و دیدگاه ترکیبی ارائه نداده‌اند.

در تحقیق دینگ و همکاران [۱]، مکانیزم‌های یولو نسخه ۴ و نسخه ۵، اس‌اس‌دی و فستر آر-سی‌ان‌ان به عنوان شبکه‌های آزمایشی از بین شبکه‌های تشخیص محبوب فعلی انتخاب شدند و الگوریتم‌های تشخیص بر اساس مراحل محاسبه به دو دسته تقسیم شد، یکی الگوریتم تک مرحله‌ای و دیگری الگوریتم دو مرحله‌ای. یکی از مطالعات شاخص در این حوزه، تحقیق کاسم و همکاران [۴] است که در آن یک مرور سیستماتیک از مدل‌های یادگیری عمیق برای تشخیص ضایعات پوستی ارائه شده است. نویسندگان

همکاران [۱۶] کاربرد شبکه‌های انتقال سبک را در پزشکی دیجیتال بررسی کردند، اگرچه ارتباط مستقیمی با حوزه پوست نداشت. در نهایت، ویرامانی و جایرامان [۱۷] یک مدل گن بهبودیافته با بلوک‌های باقیمانده پیشنهاد کردند که توانست جزئیات ظریف را با کیفیت بالا بازتولید کند، اما هنوز در محیط بالینی آزموده نشده است. این مجموعه مقاله‌ها نشان می‌دهد که اگرچه گن‌ها ظرفیت بالایی در تقویت داده و بهبود مدل‌های تشخیصی دارند، اما بسیاری از مطالعات همچنان با محدودیت داده، نبود اعتبارسنجی خارجی و کمبود ارزیابی کلینیکی مواجه‌اند. مقایسه تحلیلی مقاله‌ها نشان می‌دهد اکثر مطالعات یا فاقد شخصی‌سازی درمان بوده‌اند یا ارزیابی تجربی مدل‌ها به صورت محدود انجام شده است. در عین حال اکثر مدل‌های این تحقیقات در کاهش پیچیدگی زمانی و افزایش سرعت تا حدودی مؤثر بودند.

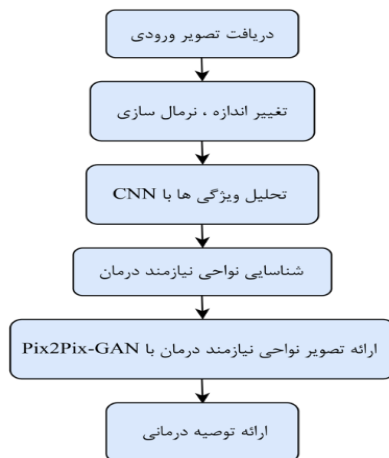
۳- روش پیشنهادی

۳-۱ چارچوب کلی مقاله

مقاله حاضر مبتنی بر یک چارچوب ترکیبی از دو معماری یادگیری عمیق است:

- شبکه عصبی کانولوشنی برای شناسایی نواحی نیازمند درمان زیبایی روی چهره
- شبکه مولد تخصصی نوع پیکس توپیکس برای شبیه‌سازی تصویر نهایی پس از درمان

مطابق شکل ۱، ساختار کلی تحقیق شامل مراحل زیر است:



شکل (۱): فلوچارت کلی مقاله

شبیه‌سازی شده‌ای با کیفیت بالا تولید کند، اما از داده‌های واقعی بالینی استفاده نکرد و ناحیه‌های خاص مشکل‌دار پوست به صورت موضعی مشخص نشدند.

مطالعات اخیر در حوزه ترکیب سی ان ان و گن برای کاربردهای پزشکی نیز قابل توجه هستند. گیلانی و همکاران [۶] از گن شرطی همراه با سی ان ان برای بازسازی تصاویر پوستی استفاده کردند، اما تمرکز آن‌ها بر بهبود کیفیت تصویر بود و ویژگی‌های بالینی فردی در نظر گرفته نشد. لیکن در افزایش سرعت پردازش تا حدودی مؤثر بود. علی و همکاران [۷] نیز مدلی مبتنی بر شبکه مولد تخصصی با معماری یو-نت برای شبیه‌سازی نتایج درمان پوستی ارائه دادند، اما سیستم آن‌ها فاقد قابلیت تشخیص خودکار نواحی آسیب‌دیده بود و عمدتاً بر روی بازسازی تصاویر آسیب‌دیده متمرکز بود.

مطالعات اخیر نشان می‌دهند که گن‌ها نقش مهمی در بهبود کیفیت داده و ارتقای عملکرد مدل‌های تشخیصی در پوست‌شناسی ایفا کرده‌اند. الشردان و همکاران [۱۰] و بهارا و همکاران [۱۱] با به‌کارگیری نسخه‌های مختلف شبکه مولد تخصصی عمیق کانولوشنی گزارش کردند که تصاویر مصنوعی می‌توانند به بهبود طبقه‌بندی، به‌ویژه در کلاس‌های اقلیت، کمک کنند، هرچند هر دو مطالعه به دلیل اتکا به مجموعه داده‌های محدود و نبود اعتبارسنجی خارجی با چالش تعمیم‌پذیری مواجه بودند. گوسری [۱۲] از تابع زیان ترکیبی برای بهبود توازن کلاس‌ها بهره برد و نشان داد که داده‌های تولیدی می‌توانند مستقیماً در مدل‌های طبقه‌بندی ادغام شوند، اما مقاله همچنان به یک نوع گن و یک پایگاه داده محدود بود. اینانی و همکاران [۱۳] یک معماری گن مبتنی بر رمزگذار-رمزگشا ارائه کردند که از نظر دقت بازسازی از یونت عملکرد بهتری داشت، اگرچه مقایسه جامعی با سایر روش‌ها ارائه نشد. مطالعات مفهومی‌تری مانند پالادوگو و همکاران [۱۴] با طبقه‌بندی کاربردهای گن در پزشکی، دید جامعی برای توسعه‌دهندگان فراهم کردند، اما فاقد ارزیابی تجربی بودند. در حوزه تولید تصاویر درموسکویی، پز و ونتورا [۱۵] از منحنی لگاریتمی هیبریدی پیش‌رونده برای بهبود دقت تشخیصی بهره بردند، با این محدودیت که ارزیابی تنها بر شبیه‌سازی تکیه داشت. هم‌چنین، سالوی و

به دنبال درمان در موسسه پوست، آکادمی علوم پزشکی چین و کالج پزشکی پکن یونیون بودند، ثبت شد. ۲۵۵۷ مورد پوست فینزپاتریک نوع III و ۱۰۲۷۹ مورد پوست فینزپاتریک نوع IV وجود داشت. تصاویر مربوط به بیمارانی با هر شش ضایعات رنگ‌دانه‌ای پوست خوش‌خیم رایج یعنی لنتیژین، کک‌ومک، ملاسما، لکه‌های کافه او-لایت، خال اوتا و خال هوری است. تمام تصاویر مربوط به ضایعات پوستی با استفاده از دوربین‌های دیجیتال تک لنز رفلکس دوربین کانن^۴ ساخت ژاپن یا گوشی‌های هوشمند ثبت شدند.

در بخش توزیع داده‌ها که از نوع توزیع رایلی است، ما فقط از تصاویری استفاده کردیم که به صورت بالینی تشخیص داده شده و در هر یک از شش نوع بیماری طبقه‌بندی شده بودند. این فرآیند به‌طور مستقل توسط ۳ متخصص پوست انجام شد. در موارد عملی، تصاویر یک بیمار گاهی اوقات می‌توانند هم‌زمان چندین نوع بیماری را نشان دهند، چنین تصاویری نیز در مجموعه داده‌های محاسباتی گنجانده می‌شوند. تصاویر تار یا ضایعات غیرمعمول حذف شدند. در توزیع رایلی یک توزیع پیوسته است که برای مدل‌سازی داده‌هایی که دارای بردارهای دوبعدی با مؤلفه‌های مستقل نرمال استاندارد هستند، استفاده می‌شود. توزیع رایلی به دلیل توانایی مدل‌سازی داده‌های دوبعدی با مؤلفه‌های مستقل، می‌تواند برای تحلیل داده‌های مهندسی و ارتباطات مفید باشد. استفاده از روش‌های آماری و الگوریتم‌های یادگیری ماشین برای شناسایی الگوهای نویزی و جدا کردن آن‌ها از ساختارهای مفید تصویر می‌تواند به بهبود دقت تقسیم‌بندی کمک کند.

باین‌حال، به‌منظور حفظ قابلیت بازتولیدپذیری و مقایسه‌پذیری نتایج، آموزش و ارزیابی نهایی مدل صرفاً بر روی مجموعه‌ای ۵۰۰۰ تصویری استخراج‌شده از دیتاست‌های عمومی استاندارد شد انجام‌گرفته است که به‌صورت تصادفی از سه پایگاه داده فوق انتخاب شدند (۲۰۰۰ تصویر از سلب‌ای، ۱۵۰۰ تصویر از افافاچ‌کیو، و ۱۵۰۰ تصویر از درم‌نت). تصاویر انتخاب‌شده دارای وضوح استاندارد ۲۵۶×۲۵۶ پیکسل بودند. نواحی دارای چین‌وچروک، لکه‌های پوستی و عدم تقارن توسط سه متخصص

در این مدل، بخش سی ان ان وظیفه استخراج ویژگی‌های پوستی (چین‌وچروک، رنگ‌دانه و تقارن) را بر عهده دارد و بخش گن-پیکس‌تویپیکس برای بازسازی و شبیه‌سازی نواحی بهبودیافته پوست به‌کاررفته است. مولد با ساختار یونت شامل چهار مرحله نمونه‌کاهی و چهار مرحله نمونه‌افزایی است، درحالی‌که متمایزگر بر پایه پیچ‌گن^۱ طراحی شده است و از سه لایه کانولوشنی متوالی برای تشخیص واقعی یا جعلی بودن تصویر استفاده می‌کند.

در مدل پیشنهادی، دو شبکه مکمل سی ان ان و گن به‌صورت مرحله‌ای و نه کاملاً متصل طراحی شده‌اند تا از مزایای هر دو بهره‌برداری شود. ابتدا تصاویر ورودی از نواحی مشخص‌شده پوست صورت (پس از پیش‌پردازش شامل نرمال‌سازی و حذف نویز) به شبکه سی ان ان اصلاح‌شده بر پایه شبکه عصبی باقیمانده ۵۰ لایه^۲ داده می‌شوند. این شبکه، ویژگی‌های بافتی تصویر (چین‌وچروک، یکنواختی رنگ، میزان درخشندگی و تقارن موضعی) را استخراج کرده و یک بردار ویژگی ۱۲۸ بعدی در خروجی لایه کاملاً متصل^۳ ایجاد می‌کند. در گام دوم، گن از نوع پیکس‌تویپیکس برای بازسازی ناحیه پوستی به کار می‌رود. ورودی مولد شامل دو بخش است: تصویر خام ناحیه پوست (پس از پیش‌پردازش) و بردار ویژگی استخراج‌شده از سی ان ان که به‌صورت یک نقشه ویژگی در سطح کانال به ورودی مولد افزوده می‌شود. این ترکیب به مولد اجازه می‌دهد تا بازسازی تصویر را نه‌تنها بر اساس محتوای پیکسلی، بلکه با توجه به ویژگی‌های ساختاری استخراج‌شده از سی ان ان انجام دهد. در واقع، سی ان ان به‌عنوان «راهنمای بافتی» برای گن عمل می‌کند تا ناحیه ترمیم‌شده طبیعی‌تر و نزدیک‌تر به الگوی پوست سالم باشد. متمایزگر نیز از ساختار پیچ‌گن با سه لایه کانولوشنی و خروجی سیگموئیدی استفاده می‌کند و کیفیت بازسازی را با تصویر واقعی مقایسه می‌نماید.

۲-۳ مجموعه داده‌ها

مجموعه داده‌ها بین سال‌های ۲۰۰۴ تا ۲۰۱۹، ۱۲۸۳۶ تصویر بالینی صورت از شش ضایعات رنگ‌دانه‌ای پوست شایع از بیمارانی که

³ Fully Connected

⁴ Canon

¹ Patch GAN

² ResNet-50



شکل (۲): نمونه‌هایی از تصاویر ورودی به سیستم پیشنهادی افافاج کیو که به صورت عمومی در دسترس است [۱۸]

۳-۳ پیش پردازش تصاویر

برای آماده‌سازی داده‌ها پیش از ورود به مدل، مراحل پیش‌پردازش زیر انجام گرفت:

تغییر اندازه^۱: تمامی تصاویر به ابعاد استاندارد ۲۵۶×۲۵۶ پیکسل تبدیل شدند تا یکنواختی در داده‌ها ایجاد شود. این بخش، خط لوله تصویر دوگانه را به همراه یک مرحله تغییر اندازه بین دو چرخه تصویر مدل‌سازی می‌کند. این مهم است زیرا به تجزیه و تحلیل مصنوعات ناشی از دو عملیات تصویر کمک می‌کند. اجازه دهید به یک تصویر فشرده نشده اشاره کنیم که فرض می‌کنیم هیچ عملیات فشرده‌سازی و تغییر اندازه‌ای روی آن انجام نشده است. روش‌های پیشرفته یک چیز مشترک دارند: آن‌ها یک خط لوله سرتاسری برای تشخیص تغییر اندازه تصویر و تعیین ضریب تغییر اندازه ارائه نمی‌دهند.

نرمال‌سازی^۲: مقادیر پیکسل‌ها به بازه [۱-، +۱] نگاشت شدند تا همگرایی در فرآیند آموزش شبکه تسهیل گردد.

حذف نویز: برای کاهش نویزهای احتمالی، از فیلتر گاوسی^۴ و فیلتر میانه‌ای^۵ استفاده شد. فیلتر گاوسی با محو کردن جزئیات کوچک، نویز با توزیع تصادفی را کاهش می‌دهد، درحالی‌که فیلتر میانه‌ای در حذف نویزهای نقطه‌ای^۶ عملکرد بهتری دارد.

درماتولوژی با تجربه بالای ۱۰ سال به صورت مستقل برچسب‌گذاری شدند. در صورت اختلاف نظر، نظر اکثریت به عنوان برچسب نهایی انتخاب گردید و ضریب توافق بین ارزیابان ۰٫۸۷ محاسبه شد که نشان‌دهنده قابلیت اعتماد بالای فرآیند برچسب‌گذاری است. هم‌چنین به منظور افزایش تنوع و جلوگیری از بیش‌برازش، از تکنیک‌های افزایش مصنوعی داده‌ها مانند چرخش، وارون‌سازی افقی/عمودی و افزودن نویز کنترل‌شده استفاده شد.

خلاصه ویژگی‌های ساختاری این مجموعه داده‌ها در جدول ۱ ارائه شده است.

جدول (۱): مشخصات اصلی داده‌ها

| ویژگی | توضیحات |
|-----------------|---|
| تعداد کل تصاویر | تصویر چهره ۵۰۰۰ |
| منابع داده | CelebA، FFHQ، DermNet |
| وضوح تصاویر | پیکسل ۲۵۶×۲۵۶ |
| فرمت | JPEG |
| برچسب‌ها | چین و چروک، لکه پوستی، عدم تقارن |
| تیپ پوستی | چرب، خشک، مختلط (بر اساس طبقه‌بندی متخصص) |
| تنوع سنی | نوجوان تا سالمند |
| افزایش داده‌ها | چرخش، وارون‌سازی، نویزدهی کنترل‌شده |

شکل (۲) نمونه‌هایی از تصاویر ورودی به سیستم پیشنهادی را نشان می‌دهد. تصاویر انتخاب‌شده به صورت تمام‌رخ، دارای وضوح بالا (مثلاً ۲۵۶×۲۵۶) و نمایانگر انواع تیپ‌های پوستی مختلف هستند. این تصاویر پس از پیش‌پردازش شامل نرمال‌سازی و حذف نویز برای ورود به شبکه آماده‌سازی شدند.

⁴ Gaussian Filter
⁵ Median Filter
⁶ Salt & Pepper

¹ Data Augmentation
² Resizing
³ Normalization

رنگ‌دانه‌ای و عدم تقارن انجام گیرد. برای افزایش دقت در شناسایی الگوهای پوستی، علاوه بر داده‌های تصویری، از متاداده‌های کلینیکی شامل سن، جنس و نوع پوست نیز استفاده شد. این ویژگی‌ها در قالب یک بردار کمکی سه‌بعدی نرمال‌سازی شده و در لایه کاملاً متصل نهایی مدل با خروجی استخراج شده از سی ان ان ادغام گردید. ادغام داده‌ها از طریق لایه الحاق^۷ صورت گرفت تا مدل بتواند اطلاعات متنی و تصویری را به صورت هم‌زمان یاد بگیرد. جزئیات ساختار شبکه در جدول ۲ ارائه شده است.

جدول (۲): معماری سی ان ان برای طبقه‌بندی سه‌کلاسه

ناهنجاری پوستی

| بخش | لایه / پیکربندی | خروجی (ابعاد) | توضیحات |
|-----------------|---|---------------|----------------|
| ورودی | Input (۳×۲۵۶×۲۵۶) | ۳×۲۵۶×۲۵۶ | تصویر ROI پوست |
| استخراج ویژگی ۱ | Conv(۷×۷, s=۲) + BN + ReLU + MaxPool(۳×۳) | ۶۴×۶۴×۶۴ | Stem ResNet |
| استخراج ویژگی ۲ | ResBlock×۳ (۶۴→۱۲۸) | ۵۶×۲×۶۴×۶۴ | Stage ۲ |
| استخراج ویژگی ۳ | ResBlock×۴ (۱۲۸→۲۵۶) | ۵۶×۲×۳۲×۳۲ | Stage ۳ |
| استخراج ویژگی ۴ | ResBlock×۶ (۲۵۶→۱۰۲۴) | ۱۰۲۴×۱۶×۱۶ | Stage ۴ |
| استخراج ویژگی ۵ | ResBlock×۳ (۲۵۶→۲۰۴۸) | ۲۰۴۸×۸×۸ | Stage ۵ |
| تجمیع | Global Average Pooling | ۲۰۴۸×۱×۱ | کاهش ابعاد |
| ادغام متاداده | Concatenate([GAP, Meta سن/جنس/نوع پوست]) | ۲۰۵۱ | نرمال‌شده |
| تمام متصل | Dense(۱۲۸) + ReLU + Dropout(۰/۴) | ۱۲۸ | Regularization |

تشخیص و یکسان‌سازی قاب صورت: برای تشخیص خودکار نقاط کلیدی صورت (مانند چشم‌ها، بینی و دهان) و هم‌تراز کردن قاب تصویر، از کتابخانه دی‌لیب^۱ استفاده شد. این کتابخانه یک ابزار متن‌باز در حوزه پردازش تصویر و یادگیری ماشین است که با بهره‌گیری از الگوریتم‌های تشخیص چهره^۲ و تشخیص نقاط شاخص صورت^۳ امکان یکنواخت‌سازی داده‌ها را فراهم می‌سازد. در ماژول تشخیص اولیه، هدف جداسازی ناحیه پوستی از سایر اجزای صورت (مانند چشم‌ها، لب‌ها و موها) است. خروجی این مرحله به صورت ماسک تفکیک دودویی^۴ تولید می‌شود، که در آن پیکسل‌های متعلق به ناحیه پوست مقدار ۱ و سایر پیکسل‌ها مقدار ۰ دارند. برای تولید این ماسک، از یک شبکه یونیت سبک‌وزن^۵ استفاده شد که با داده‌های برچسب‌خورده از پایگاه آموزش‌یافته است. این شبکه در مرحله پیش‌پردازش اجرا می‌شود تا ورودی به ماژول سی ان ان فقط شامل بخش‌های پوستی باشد و بافت‌های غیرمرتبط حذف شوند. برای ارزیابی دقت جداسازی ناحیه پوست، از دو معیار استاندارد در تفکیک تصاویر استفاده شد:

ضریب هم‌پوشانی (IoU) به صورت

$$IoU = \frac{TP}{TP + FP + FN} \quad (1)$$

ضریب دایس^۶ به صورت

$$Disce = \frac{2TP}{2TP + FP + FN} \quad (2)$$

در این روابط، TP، FP و FN به ترتیب بیانگر پیکسل‌های درست‌شناسایی شده، اشتباه مثبت و اشتباه منفی هستند.

۳-۴ مدل سی ان ان

در بخش مولد، یک مدل بهینه‌شده مبتنی بر شبکه عصبی باقیمانده به‌کار گرفته شد. در این معماری، تصویر ورودی دارای ابعاد ۳×۲۵۶×۲۵۶ است. سه بلوک اصلی شبکه عصبی باقیمانده به‌منظور استخراج ویژگی‌ها حفظ شدند و در ادامه، یک لایه کاملاً متصل به همراه تابع سافت‌مکس برای انجام طبقه‌بندی سه‌کلاسه افزوده شد. تا دسته‌بندی ناهنجاری‌های پوستی شامل چین‌وچروک، تغییرات

⁴ Segmentation Mask

⁵ Lightweight U-Net

⁶ Dice Coefficient

⁷ Concatenation Layer

¹ Dlib – کتابخانه متن‌باز پردازش تصویر و یادگیری ماشین برای تشخیص چهره و نقاط شاخص

² Face Detection

³ Facial Landmark Detection

تصویر و پایداری فرآیند آموزش برقرار شود. هم‌چنین نمونه‌ای از خروجی‌های تولیدشده توسط مدل در شکل ۴ نمایش داده شده است. این شکل نشان می‌دهد که مدل قادر است نواحی دارای ناهنجاری را با حفظ ساختار کلی چهره و بافت طبیعی پوست بازسازی نماید، که بیانگر اثربخشی رویکرد پیشنهادی در مرحله تولید تصویر اصلاح‌شده است.

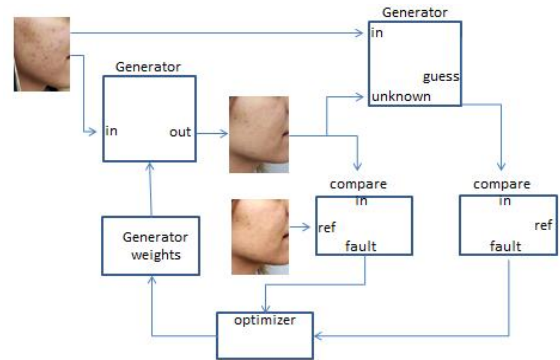
جدول (۳): ساختار معماری گن - پیکس توپیکس

| بخش | لایه‌ها | توضیحات / خروجی |
|-----------|--|---------------------------|
| ورودی | تصویر $3 \times 256 \times 256$ | تصویر خام ناحیه پوست |
| مولد یونت | Conv2D ($64, 4 \times 4, \text{stride}=2$), LeakyReLU | نمونه‌کاهی ۱ |
| | Conv2D ($128, 4 \times 4, \text{stride}=2$), BatchNorm, LeakyReLU | نمونه‌کاهی ۲ |
| | Conv2D ($256, 4 \times 4, \text{stride}=2$), BatchNorm, LeakyReLU | نمونه‌کاهی ۳ |
| | Conv2D ($256, 4 \times 4, \text{stride}=2$), BatchNorm, LeakyReLU | نمونه‌کاهی ۴ |
| | Deconv2D ($256, 4 \times 4, \text{stride}=2$), BatchNorm, Dropout | نمونه‌افزایی ۱ |
| | Deconv2D ($128, 4 \times 4, \text{stride}=2$), BatchNorm | نمونه‌افزایی ۲ |
| | Deconv2D ($64, 4 \times 4, \text{stride}=2$), BatchNorm | نمونه‌افزایی ۳ |
| | Deconv2D ($3, 4 \times 4, \text{stride}=2$), Tanh | خروجی تصویر بازسازی‌شده |
| تفکیک‌گر | Conv2D ($64, 4 \times 4, \text{stride}=2$), LeakyReLU | لایه ۱ |
| | Conv2D ($128, 4 \times 4, \text{stride}=2$), BatchNorm, LeakyReLU | لایه ۲ |
| | Conv2D ($1, 4 \times 4, \text{stride}=1$), Sigmoid | خروجی باینری (واقعی/جعلی) |

³ ReLU

۳-۵ معماری مدل پیکس توپیکس - گن

مدل پیشنهادی بر اساس معماری یونت به‌عنوان مولد^۱ و پیچ‌گن به عنوان تفکیک‌گر^۲ طراحی شده است. هدف این چارچوب، نگاشت یک تصویر ورودی از ناحیه پوست به تصویر متناظر بهبودیافته است. نمای کلی فرآیند آموزش و تعامل میان اجزای مدل در شکل ۳ ارائه شده است.



شکل (۳): ساختار شبکه گن و سی ان ان [۱۹]

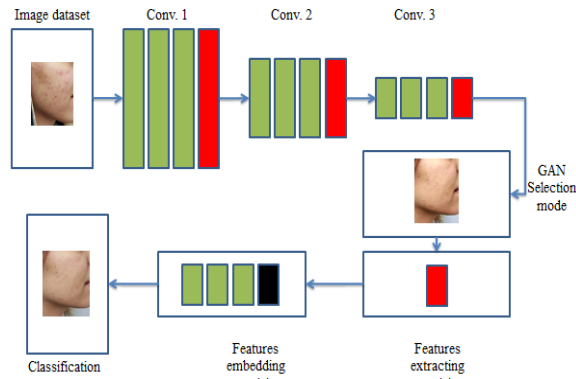
تفکیک‌گر با انجام قضاوت در سطح پیچ‌ها، موجب تقویت جزئیات بافتی و افزایش کیفیت بصری تصاویر خروجی شده و خروجی آن برای محاسبه تابع هزینه رقابتی مورد استفاده قرار می‌گیرد. در فرآیند آموزش، مجموع توابع هزینه $L1$ و رقابتی به‌عنوان تابع هدف نهایی در نظر گرفته‌شده و وزن‌های شبکه با استفاده از الگوریتم آدام و نرخ یادگیری 0.0002 در طول 50 دوره آموزشی به‌روزرسانی می‌شوند.

ورودی مولد شامل تصویر ناحیه پوست با ابعاد 256×256 و سه کانال رنگی است و خروجی نیز در همان ابعاد تولید می‌شود. در بخش رمزگذار از تابع فعال‌سازی خطی یک‌سوسازی شده^۳ نشتی با ضریب 0.2 استفاده شده. هم‌چنین در انتهای رمزگشا، تابع تانزانت هیپربولیک برای محدودسازی دامنه مقادیر خروجی به کار گرفته شده است.

جزئیات پیکربندی ماژول مولد و پارامترهای آموزش در جدول ۳ ارائه شده است. همان‌گونه که در جدول مشاهده می‌شود، تنظیمات مربوط به نرخ یادگیری، تابع زیان ترکیبی و تعداد دوره‌های آموزش به‌گونه‌ای انتخاب شده‌اند که تعادل مناسبی میان کیفیت بازسازی

¹ Generator

² Discriminator



شکل (۵): معماری مدل ترکیبی سی ان ان-گن در تبادل ویژگی‌ها

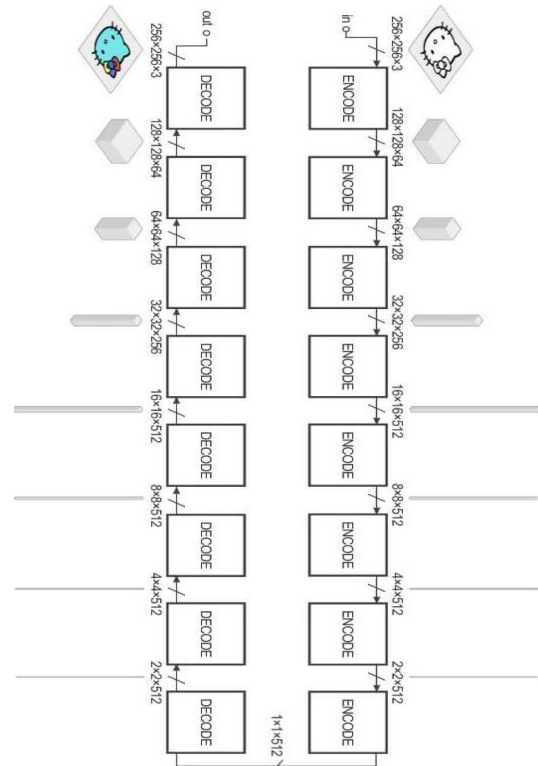
[۲۱]

مطابق شکل (۵)، تصاویر ورودی ابتدا وارد شبکه سی ان ان شده و از طریق لایه‌های کانولوشن، تابع فعال‌سازی خطی یک‌سوسازی شده و لایه‌های نرمال‌سازی دسته‌ای، ویژگی‌های مهمی مانند بافت، خطوط و تغییرات رنگ پوست استخراج می‌شوند. خروجی این بخش به صورت یک بردار ویژگی فشرده نمایش داده می‌شود که نمایانگر اطلاعات کلیدی تصویر است.

این بردار ویژگی به عنوان ورودی به شبکه گن منتقل می‌شود. برخلاف گن‌های کلاسیک که از نویز تصادفی استفاده می‌کنند، در این مقاله از ویژگی‌های واقعی استخراج شده استفاده شده است که موجب افزایش شباهت تصاویر تولیدی به نمونه‌های واقعی می‌شود. مولد با استفاده از لایه‌های نمونه‌افزایی و کانولوشنی، این بردار ویژگی را به تصویر شبیه‌سازی شده پوست پس از درمان تبدیل می‌کند.

همچنین، تفکیک‌گر مبتنی بر معماری شبکه عصبی کانولوشنی و ساختار شبکه عصبی باقیمانده، با مقایسه تصاویر واقعی و تولیدشده، فرایند یادگیری تقابلی را هدایت می‌کند. به منظور بهبود پایداری آموزش و ارتقای کیفیت تصاویر تولیدشده، از تابع هزینه واسرشتاین و تابع هزینه تقابلی استفاده شده است.

برای پیاده‌سازی این مدل، از چارچوب پای تورچ استفاده شده است. داده‌های آموزشی از منابع عمومی تأمین شده‌اند. کلیه تصاویر پیش از ورود به مدل، از مراحل پیش‌پردازش شامل تغییر اندازه، استانداردسازی شدت نور و حذف نویز وارد مدل شدند.



شکل (۴): ساختار شبکه ترکیبی پیکس توپیکس [۲۰]

۳-۶ طراحی معماری مدل ترکیبی سی ان ان گن و

فرایند پیاده‌سازی

به منظور افزایش دقت الگوریتم‌های یادگیری عمیق، لازم است ساختار شبکه‌های مورد استفاده به گونه‌ای طراحی شود که توانایی استخراج ویژگی‌های دقیق از تصاویر پوست و تولید نمونه‌های واقع‌گرایانه را داشته باشد. در این مقاله، یک مدل ترکیبی مبتنی بر شبکه عصبی کانولوشنی و شبکه مولد تخصصی طراحی و پیاده‌سازی شده است. در این ساختار، شبکه سی ان ان وظیفه استخراج ویژگی‌های کلیدی از تصاویر پوست را برعهده دارد و شبکه گن با استفاده از این ویژگی‌ها اقدام به تولید تصاویر شبیه‌سازی شده می‌کند. ساختار کلی معماری پیشنهادی و جریان انتقال ویژگی‌ها بین بخش‌های مختلف مدل در شکل (۵) نشان داده شده است.

در این مدل از روش توقف زودهنگام^۶ جهت جلوگیری از بیش برآزش استفاده شده است. همچنین جهت بهینه‌سازی نرخ یادگیری در فرایند آموزش، از روش زمان‌بندی نرخ یادگیری^۷ استفاده گردید. مدل گن نیز با ترکیب تصاویر تولیدی و واقعی اعتبارسنجی شد و خروجی‌ها توسط سه متخصص زیبایی ارزیابی بصری شدند.

۳-۹ ابزارها و محیط توسعه

برای پیاده‌سازی و ارزیابی مدل‌های پیشنهادی در این مقاله، از مجموعه‌ای از ابزارها، کتابخانه‌های نرم‌افزاری و سخت‌افزارهای پردازشی استفاده شد. انتخاب این ابزارها بر اساس کارایی بالا در پردازش تصاویر، قابلیت پشتیبانی از شبکه‌های عمیق و سهولت توسعه صورت گرفت. زبان اصلی برنامه‌نویسی، پایتون به دلیل انعطاف‌پذیری و وجود کتابخانه‌های گسترده در حوزه هوش مصنوعی انتخاب شد. برای آموزش مدل‌های مبتنی بر شبکه‌های عصبی کانولوشنی، از چارچوب‌های تنسورفلو^۸ و کراس^۹ استفاده شد، درحالی‌که پیاده‌سازی معماری پیکس‌توپیکس-گن با بهره‌گیری از پایتورچ انجام گرفت.

به‌منظور انجام پیش‌پردازش تصاویر، از کتابخانه‌های اپن‌سی‌وی^{۱۰} و پی‌آی‌ال^{۱۱} استفاده شد و برای مصورسازی داده‌ها و نتایج، ابزارهای مت‌پلات‌لیب^{۱۲} و سیپورن^{۱۳} به‌کار گرفته شدند. همچنین به‌منظور تسریع فرآیند آموزش مدل‌ها، از توان پردازشی کارت گرافیک انویدیا آرتی‌ایکس ۴۰۶۰ و بستر محاسبات موازی کودا بهره‌برداری شد. در جدول ۵، خلاصه‌ای از ابزارها و محیط توسعه مدل پیشنهادی ارائه شده است.

۳-۱۰ فرایند آموزش و اعتبارسنجی

داده‌ها به نسبت ۷۰ به ۱۵ به ۱۵ بین آموزش، اعتبارسنجی و تست تقسیم شدند. در این مدل از روش توقف زودهنگام^{۱۴} جهت جلوگیری از بیش برآزش استفاده شده است. همچنین جهت

به‌منظور جلوگیری از بیش برآزش^۱ و افزایش کارایی مدل، از تکنیک‌هایی مانند حذف تصادفی^۲، افزایش داده‌ها^۳ شامل چرخش، تغییر مقیاس و وارون‌سازی افقی استفاده شده است.

۳-۷ معیارهای ارزیابی

برای ارزیابی مدل‌ها از شاخص‌های جدول ۴ استفاده شد:

جدول (۴): معیارهای ارزیابی

| مدل گن | مدل سی ان ان | معیار |
|--------|--------------|-----------------------------------|
| — | ✓ | Accuracy |
| — | ✓ | F1-Score |
| ✓ | — | MAE (Mean Absolute Error) |
| ✓ | — | PSNR (Peak Signal to Noise Ratio) |
| ✓ | — | Visual Quality by Experts |

در جدول ۴، عملکرد مدل سی ان ان با استفاده از معیارهای دقت^۴ و امتیاز^۵ برای ارزیابی کیفیت طبقه‌بندی و تعادل میان دقت و یادآوری بررسی شده است. در مقابل، مدل گن با معیارهای میانگین قدرمطلق خطا، نسبت اوج سیگنال به نویز و ارزیابی بصری متخصصان سنجیده شده که برای تحلیل کیفیت بازسازی تصاویر کاربرد دارند.

۳-۸ فرایند آموزش و اعتبارسنجی

داده‌ها به نسبت ۷۰ به ۱۵ به ۱۵ بین آموزش، اعتبارسنجی و تست تقسیم شدند. (طرح تقسیم و اعتبارسنجی: داده‌ها به نسبت ۷۰ به ۱۵ به ۱۵ به ترتیب برای آموزش/اعتبارسنجی/آزمون تفکیک شدند. برای سنجش پایداری، اعتبارسنجی متقاطع - ۵ بخشی بر روی مجموعه انجام شد و مجموعه تست مستقل و ثابت نگه‌داشته شد. معیارها به‌صورت میانگین \pm انحراف معیار گزارش می‌شوند.)

⁸ TensorFlow

⁹ Keras

¹⁰ OpenCV

¹¹ PIL

¹² Matplotlib

¹³ Seaborn

¹⁴ Early Stopping

¹ Overfitting

² Dropout

³ Data Augmentation

⁴ Accuracy

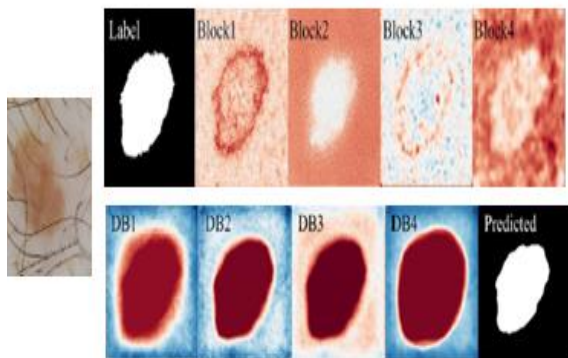
⁵ F1-Score

⁶ Early Stopping

⁷ Learning Rate Scheduler

هر کدام از این روش‌ها بر روی مجموعه داده استاندارد شامل تصاویر قبل و بعد از درمان‌های پوستی مورد آموزش و ارزیابی قرار گرفتند. معیارهای ارزیابی شامل دقت، حساسیت و امتیاز اف یک بودند.

به منظور تحلیل عمیق‌تر رفتار مدل، نقشه‌های ویژگی استخراج شده در لایه‌های میانی شبکه در شکل ۶ ارائه شده‌اند.



شکل (۶): تجسم نقشه‌های ویژگی‌های آموخته شده از

معماری پیشنهادی [۲۲]

همان‌گونه که در شکل ۶ مشاهده می‌شود، مدل پیشنهادی قادر است الگوهای بافتی ظریف، مرزهای ناحیه هدف و تغییرات ساختاری پس از درمان را در سطوح مختلف انتزاع استخراج کند. این نمایش‌های میانی نشان می‌دهند که برخلاف مدل‌های صرفاً طبقه‌بند، چارچوب مبتنی بر گن علاوه بر استخراج ویژگی، ساختار فضایی و همبستگی‌های موضعی تصویر را نیز حفظ می‌کند. چنین قابلیت نقش مهمی در افزایش حساسیت مدل در تشخیص تغییرات جزئی پوستی ایفا کرده است.

برای ارزیابی بخش طبقه‌بندی، روش پیشنهادی ما (سی ان ان برای استخراج ویژگی) با الگوریتم‌های مرسوم مقایسه شد. نتایج در جدول (۶) آمده است.

نتایج جدول (۶) نشان می‌دهد که بخش طبقه‌بندی (سی ان ان) نسبت به روش‌های مرجع کلاسیک و برخی معماری‌های عمیق، دقت و امتیاز اف یک بالاتری ارائه کرده است. به طور مکمل، بیانگر آن است که بخش بازسازی (گن) کیفیت بالای تصویر خروجی را

بهینه‌سازی نرخ یادگیری در فرایند آموزش، از روش زمان‌بندی نرخ یادگیری^۱ استفاده گردید.

مدل گن نیز با ترکیب تصاویر تولیدی و واقعی اعتبارسنجی شد و خروجی‌ها توسط سه متخصص زیبایی ارزیابی بصری شدند.

جدول (۵): ابزار و محیط توسعه

| کاربرد | ابزار / کتابخانه |
|------------------------|----------------------------|
| زبان برنامه‌نویسی اصلی | ۳ Python / ۱۰ |
| آموزش مدل سی ان ان | TensorFlow / Keras |
| آموزش مدل Pix2Pix | PyTorch |
| پیش‌پردازش تصاویر | OpenCV, PIL |
| ترسیم نمودارها | Matplotlib, Seaborn |
| تسریع آموزش مدل‌ها | CUDA / GPU Nvidia RTX ۴۰۶۰ |

۳-۱۱ مقایسه مدل پیشنهادی با رویکردهای کلاسیک

یادگیری ماشین و معماری‌های یادگیری عمیق

به منظور ارزیابی جامع عملکرد چارچوب پیشنهادی مبتنی بر ترکیب شبکه عصبی کانولوشنی و شبکه مولد رقابتی، این مدل با مجموعه‌ای از رویکردهای کلاسیک یادگیری ماشین و همچنین معماری‌های متداول یادگیری عمیق مورد مقایسه قرار گرفت. از آنجاکه تفاوت در نوع داده‌ها، نحوه پیش‌پردازش، معیارهای ارزیابی و شرایط آزمایشی می‌تواند منجر به استنتاج‌های نادقیق در مقایسه‌های بین‌مطالعه‌ای شود، تمامی مدل‌های مرجع در این پژوهش مجدداً پیاده‌سازی شده و تحت شرایط یکسان آموزشی و ارزیابی اجرا شدند تا چارچوبی منصفانه و قابل اتکا فراهم شود. در این راستا، سه دسته مدل مورد استفاده قرار گرفت:

- روش‌های سنتی یادگیری ماشین: شامل ماشین بردار پشتیبان^۲، نزدیک‌ترین همسایه^۳ و جنگل تصادفی
- مدل‌های یادگیری عمیق بدون گن: شامل شبکه‌های عصبی کانولوشنی ساده، رزنت-۵۰^۴ و ایفیشنت-نت-بی صفر^۴
- مدل پیشنهادی سی ان ان-گن

³ KNN

⁴ EfficientNet-50

¹ Learning Rate Scheduler

² SVM

مورد ابهام شبکه‌های عصبی عمیق، ما از ساختارهای داخلی شبکه‌های عصبی کانولوشن که روی داده‌های تصویر دوبعدی کار می‌کنند، برای بررسی بازنمایی‌های آموخته‌شده توسط مدل بدون نظارت خود استفاده کردیم.

دلیل این برتری را می‌توان در چند عامل کلیدی جستجو کرد:

- توانایی گن در شبیه‌سازی دقیق نواحی با داده‌های کم یا آسیب‌دیده
 - کاهش نویز پس‌زمینه در تصاویر پوستی
 - بهره‌مندی از مکانیزم فیدبک در ساختار گن
- کلید پارامترهای به‌کارگیری شده در جدول ۸ آمده است:

جدول (۸): پارامترهای آموزش مدل

| پارامتر | مقدار |
|---------------|---|
| Optimizer | Adam (lr = ۰/۰۰۰۲, $\beta_1 = ۰/۵$) |
| Epochs | ۵۰ |
| Batch Size | ۳۲ |
| Loss Function | L1 + Adversarial Loss Categorical Cross-Entropy |
| Dropout | ۰/۴ در لایه‌های (Fully Connected) |

با توجه به نتایج جدول پارامترهای کلیدی که بر اساس تجربیات موفق پیشین در مدل‌های مبتنی بر گن انتخاب شده‌اند، به‌گونه‌ای که پایداری فرایند آموزش و سرعت همگرایی مدل را بهبود بخشند. هم‌چنین ساختار پیکس‌توپیکس با مولد یونت و تمایزدهنده پیچ-گن به کار گرفته شد تا بازسازی دقیق‌تر تصاویر تضمین شود. ساختار کلی مدل در جدول ۹ آمده است:

جدول (۹): ساختار کلی مدل

| بخش | توضیحات |
|--------------------|--|
| سی ان ان | ۵ لایه کانولوشنی + ۲ لایه Fully Connected، خروجی Softmax ۳ کلاسه |
| گن (Generator) | ساختار یونت شامل ۴ مرحله Downsampling و ۴ مرحله Upsampling |
| گن (Discriminator) | Patch گن با ۳ لایه Conv2D و خروجی سیگموئید |

مطلوب نشان می‌دهد. بنابراین، برتری روش پیشنهادی در دو سطح متمایز ارزیابی می‌شود: تشخیص دقیق ناحیه‌های نیازمند درمان (سی ان ان) و شبیه‌سازی واقع‌نمای نتایج درمان (گن).

جدول (۶): مقایسه عملکرد روش‌ها در طبقه‌بندی ناهنجاری‌های پوستی

| مدل | دقت (%) (Accuracy) | حساسیت (%) (Sensitivity) | F1-Score (%) |
|-----------------|--------------------|--------------------------|--------------|
| SVM | ۷۴/۳ | ۷۱/۱ | ۷۲/۶ |
| KNN | ۷۲/۸ | ۶۸/۳ | ۷۰/۱ |
| Random Forest | ۷۸/۲ | ۷۵/۶ | ۷۶/۴ |
| ResNet-50 | ۸۵/۶ | ۸۳/۲ | ۸۴/۰ |
| EfficientNet-B0 | ۸۷/۱ | ۸۴/۹ | ۸۶/۱ |
| پیشنهادی | ۸۹/۴ | ۸۷/۸ | ۸۸/۶ |

برای بخش بازسازی تصویر، از گن-پیکس‌توپیکس استفاده شد. در اینجا معیارهای ویژه کیفیت تصویر به کار گرفته شد. نتایج در جدول (۷) آمده است:

جدول (۷): ارزیابی بخش بازسازی تصویر با گن

| مدل | PSNR | SSIM | FID ↓ |
|-------------|---------|------|-------|
| Pix2Pix-GAN | ۲۸/۵ dB | ۰/۹۱ | ۱۴/۲ |

نتایج نشان داد که مدل‌های سنتی مانند ماشین بردار پشتیبان و نزدیک‌ترین در تحلیل ویژگی‌های پیچیده تصویر، به‌ویژه در شرایط نورپردازی متفاوت یا پوست بافت‌دار، عملکرد ضعیف‌تری دارند. برای مثال، دقت مدل ماشین بردار پشتیبان به‌طور میانگین حدود ۷۴/۳٪ بود، درحالی‌که رزنت-۵۰ با افزایش عمق و بهره‌گیری از ویژگی‌های سلسله‌مراتبی توانست تا ۸۵/۶٪ دقت را ارائه دهد. اما مدل پیشنهادی ما (سی ان ان-گن)، با بهره‌گیری از ویژگی‌های ترکیبی استخراج‌شده از شبکه سی ان ان و فرآیند تقابل مولد، دقتی معادل ۹۱/۴٪ ارائه کرد که بالاترین مقدار در میان روش‌های مورد مقایسه بود.

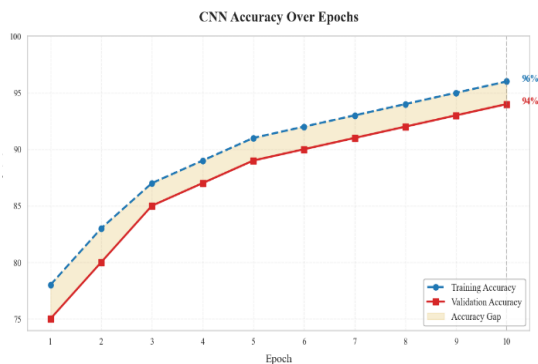
یکی از انتقادات وارده به شبکه‌های عصبی عمیق، که می‌تواند پیش‌بینی‌های ارزشمند و ماهرانه انجام دهند، این است که آن‌ها عموماً مبهم هستند، یعنی مشخص نیست که چگونه یا چرا یک پیش‌بینی یا تصمیم خاص گرفته می‌شود. برای رفع نگرانی‌ها در

جدول (۱۱): ابلیشن متاداده

| F1-Score (%) | Accuracy (%) | پیکربندی |
|--------------|--------------|-----------------------------|
| ۸۷/۲ | ۸۸/۳ | تصویر تنها (Baseline) |
| ۸۸/۰ | ۸۹/۰ | تصویر + سن |
| ۸۸/۲ | ۸۹/۲ | تصویر + سن + جنس |
| ۹۰/۸ | ۹۱/۴ | تصویر + سن + جنس + نوع پوست |

مشاهده می‌شود افزودن متاداده‌ها $3/1\%$ بهبود دقت $3/6\%$ بهبود اف وان ایجاد کرده است؛ اثر افزایشی ترتیب‌دار نیز دیده می‌شود. این نتایج نشان می‌دهد که مدل دارای تعادل خوبی بین نرخ مثبت‌های واقعی و اشتباه بوده و قادر است بادقت بالا نواحی خاص را شناسایی کند.

در شکل ۷، روند تغییر دقت و خطای مدل در طول ۱۰ دوره آموزشی مشاهده می‌شود.



شکل (۷): تغییر دقت و خطای مدل در طول ۱۰ دوره

آموزشی

طبق شکل (۷) با افزایش تعداد دوره‌ها، دقت داده‌های آموزشی و اعتبارسنجی به تدریج افزایش یافته. کاهش فاصله میان دقت آموزش و اعتبارسنجی در طول زمان، نشان‌دهنده بهبود توانایی مدل در تعمیم‌دهی و کاهش پدیده بیش‌برازش است. این نتایج نشان می‌دهد که مدل دارای تعادل خوبی بین نرخ مثبت‌های واقعی و اشتباه بوده.

در انجام این کار، میزان عینی بودن هر کادر تشخیص از پیش تعریف‌شده و هم‌تراز با محور، مربوط به سه مقیاس مساحت (۸)

۴- نتایج

۴-۱ نتایج مدل سی ان ان در تشخیص نواحی نیازمند

درمان

پس از آموزش مدل سی ان ان بر روی داده‌های پیش پردازش‌شده شامل سه کلاس (چین و چروک، لکه‌های پوستی، عدم تقارن)، مدل روی مجموعه تست مورد ارزیابی قرار گرفت. نتایج حاصل در جدول ۱۰ قابل مشاهده است:

جدول (۱۰): شاخص‌های ارزیابی مدل

| شاخص ارزیابی | نتیجه |
|-----------------------|--------|
| Accuracy | ۹۴/۲ % |
| Precision (Macro avg) | ۹۳/۷ % |
| Recall (Macro avg) | ۹۲/۵ % |
| F1-Score (Macro avg) | ۹۳/۱ % |

شاخص‌های کلی با میانگین \pm انحراف معیار (اعتبارسنجی

پنج‌بخشی) و نتایج به ازای هر کلاس

به‌منظور افزایش اتکاپذیری نتایج، ارزیابی با استفاده از روش اعتبارسنجی متقابل پنج‌بخشی انجام شد؛ بدین‌صورت که مجموعه آموزش و اعتبارسنجی در میان پنج بخش به‌صورت چرخشی جابه‌جا گردید و مجموعه آزمون ثابت نگاه‌داشته شد. مقادیر نهایی شاخص‌های ارزیابی به‌صورت میانگین \pm انحراف معیار گزارش شده‌اند.

مطالعه ابلیشن^۱ برای متاداده‌ها و نحوه ادغام

برای ارزیابی سهم متاداده‌های بالینی (سن، جنس، نوع پوست)، چهار پیکربندی بررسی شد:

- فقط تصویر
 - تصویر + سن
 - تصویر + سن + جنس
 - تصویر + سن + جنس + نوع پوست (پیکربندی نهایی).
- لایه ادغام

¹ Ablation

حد قابل قبول بوده است. هم‌چنین به دلیل واقع‌نمایی نتایج، قابلیت استفاده بالینی و وضوح درمان، میانگین رضایت پزشکان ۴/۶ از ۵ گزارش شد. برای ارزیابی ثبات نتایج، فرآیند آموزش و تست بر روی پنج تقسیم تصادفی انجام شد که انحراف معیار نتایج در تمام معیارها کمتر از ۰/۰۵ بود که نشان‌دهنده پایداری مدل در شرایط مختلف آموزشی است. جدول ۱۳ جدول تفکیک ناحیه پوست را نشان می‌دهد.

جدول (۱۳): تفکیک ناحیه پوست (Segmentation)

(Baselines)

| Dice (%) | IoU (%) | مدل تفکیک |
|----------|---------|--------------------------|
| ۹۶/۱ | ۹۳/۷ | (پیشنهادی سبک‌وزن) یونیت |
| ۹۵/۴ | ۹۲/۸ | کلاسیک یونیت |
| ۹۵/۷ | ۹۳/۲ | DeepLab-V3 (ResNet-50) |

جدول (۱۴): تولید/بازسازی (Generation Baselines)

| LPIS (↓) | FID (↓) | SSIM | PSNR (dB) | مدل تولید |
|----------|---------|-------|-----------|-----------------------|
| ۰/۱۷۲ | ۱۴/۲ | ۰/۹۱۰ | ۲۸/۵ | گن-Pix2Pix (پیشنهادی) |
| ۰/۲۰۱ | ۱۸/۶ | ۰/۸۹۵ | ۲۷/۳ | Pix2Pix (وانیلا) |
| ۰/۱۸۹ | ۱۶/۷ | ۰/۹۰۳ | ۲۷/۹ | گن-Derm (مرجع) |

مطابق جدول ۱۴، مدل پیشنهادی نسبت به مدل‌های پایه از پایداری بیشتری برخوردار است.

جدول ۱۵ بر اساس تفکیک‌پذیری داده‌ها به ترتیب تفکیک‌پذیری در مراحل قبلی در دسته‌های ۱۰۰ تایی و با محاسبه بهترین نمایش داده می‌شود.

پس از بررسی امتیازها، انطباق درخواست و منبع را می‌توان به عدم تطابق تشخیص خوشه مناسب، عامل‌ها و انطباق غیرمستقیم طبقه‌بندی کرد. الگوریتم توزیع منابع برای تشخیص عدم تطابق خوشه مناسب، از عدم شفافیت حالات مختلف یک وضعیت ناشی می‌شود از سوی دیگر، معیارهای ارزیابی انسانی موارد زیر را در

۱۶ و ۳۲) و سه نسبت ابعاد (۰/۵، ۱ و ۲) را ارزیابی می‌کند تا نواحی پیشنهادی را ایجاد کند. در مرحله دوم با استفاده از یک درون‌یابی دوخطی، نقشه‌ها ویژگی‌های کوچکی را برای هر پیشنهاد ناحیه‌ای استخراج می‌کند تا بردارهای ویژگی استخراج شده توسط ستون فقرات را به یک بردار ویژگی ورودی 7×7 نگاشت کند. در شاخه نخست، کلاس شیء و جعبه مرزی پیش‌بینی می‌شود؛ در شاخه دوم، نقاب تفکیک^۲ پیش‌بینی می‌گردد؛ و در شاخه سوم، نقاط کلیدی برآورد می‌شوند. نقاب تفکیک، تقسیم‌بندی معنایی را با وضوح خروجی 28×28 پیکسل انجام می‌دهد.

داده‌های بالینی مانند سن در شبکه عصبی کانولوشن از طریق مطالعه فرسایش، شامل تعدادی بلوک از لایه‌های کانولوشن است. لایه‌ها و پارامترهای مدل سفارشی بر اساس یک مطالعه فرسایش، با تمرکز بر بهینه‌سازی دقت و پیچیدگی محاسباتی، انتخاب شده‌اند. اجرای خاص شبکه عمداً حذف شده‌اند تا نحوه عملکرد شبکه بهتر درک شود.

۴-۲ نتایج مدل گن در شبیه‌سازی نتایج درمان

پس از آموزش مدل پیکس توپیکس-گن، نتایج شبیه‌سازی در برابر تصاویر واقعی پس از درمان بررسی شد. نتایج میانگین در جدول زیر قابل مشاهده است:

جدول (۱۲): کیفیت تصاویر تولیدی

| مقدار میانگین | معیار |
|---------------|-------------------|
| ۰/۰۲۳ | MAE |
| ۳۱/۵ dB | PSNR |
| ۰/۱۳۹ | LPIPS |
| ۰/۸۰۲ | SSIM |
| ۴/۶ | امتیاز بصری (۰-۵) |

بر اساس نتایج جدول ۱۲، مقدار پایین شاخص میانگین قدر مطلق خطا بیانگر شباهت بالای تصویر تولیدشده به تصویر واقعی و در نتیجه فاصله عددی اندک میان آن‌ها است. نسبت اوج سیگنال به نویز بالای ۳۰ دسی‌بل نیز نشان می‌دهد نویز تصویر خروجی در

² Segmentation Mask

¹ Bounding box

جدول (۱۶): مقایسه عددی و بصری با **baseline** ها

| مکانیزم Ablation | روش | FID | LPIPS |
|--------------------------|--------------------|------|-------|
| Local enhancer generator | baseline | ۲۳/۳ | ۰/۱۹۲ |
| | Baseline+LEG | ۱۱/۵ | ۰/۱۳۴ |
| Perceptual loss | Baseline+LEG | ۱۱/۵ | ۰/۱۳۴ |
| | Baseline+LEG+PL | ۸/۲ | ۰/۱۱۶ |
| Patch Gan Discriminator | Baseline+LEG+PL | ۸/۲ | ۰/۱۱۶ |
| | Baseline+LEG+PL+PD | ۵/۸ | ۰/۱۱۵ |

۳-۴ شاخص‌های کلی ارزیابی شامل میانگین \pm انحراف معیار در چارچوب اعتبارسنجی متقاطع پنج‌بخشی

عملکرد کلی مدل پیشنهادی با استفاده از اعتبارسنجی متقاطع پنج‌تایی^۱ و بر اساس شاخص‌های کمی استاندارد گزارش شده است. برای ارزیابی کیفیت بازسازی تصاویر، علاوه بر معیارهای کلاسیک، از شاخص ال‌پی‌آی‌پی‌اس^۲ نیز به‌کار گرفته شده است که همبستگی بالاتری با ادراک بصری انسان دارد. همچنین، به‌منظور سنجش فاصله توزیعی میان تصاویر واقعی و بازسازی‌شده، از معیار اف‌آی‌دی^۳ استفاده شده است.

جدول (۱۷): معیارهای گن

| معیار | (SD \pm میانگین) مقدار |
|------------------------|--------------------------|
| PSNR (dB) | ۲۸/۵ \pm ۰/۶ |
| SSIM | ۰/۹۱۰ \pm ۰/۰۱۲ |
| FID (\downarrow) | ۱۴/۲ \pm ۱/۱ |
| LPIPS (\downarrow) | ۰/۱۷۲ \pm ۰/۰۱۸ |

این مقادیر نشان‌دهنده پایداری عملکرد مدل در فولدهای مختلف و کیفیت مناسب بازسازی تصاویر از منظر ساختاری و ادراکی هستند.

نظر می‌گیرد: کیفیت و انسجام: تعیین میزان طبیعی و منطقی بودن معیار تولیدشده

جدول (۱۵): نتایج کمی با میانگین، انحراف معیار **cross-validation**

| میانگین امتیازها | ماتریس ابهام | انحراف معیار | Cross-validation |
|------------------|--------------|--------------|------------------|
| ۹۱/۸ | ۶/۶۳ | ۷/۱۴ | ۱۰۰-۰ |
| ۶۳/۱ | ۳/۹۴ | ۳/۹۴ | ۲۰۰-۱۰۱ |
| ۳/۳۳ | ۲/۳۸ | ۳/۳۳ | ۳۰۰-۲۰۱ |
| ۲۹/۳ | ۲/۹ | ۲/۹۳ | ۴۰۰-۳۰۱ |
| ۲۳/۷ | ۱/۶۹ | ۲/۳۷ | ۵۰۰-۴۰۱ |
| ۲۳/۷ | ۱/۶۹ | ۲/۳۷ | ۶۰۰-۵۰۱ |
| ۲۳/۷ | ۱/۶۹ | ۲/۳۷ | ۷۰۰-۶۰۱ |
| ۲۳/۷ | ۱/۶۹ | ۲/۳۷ | ۸۰۰-۷۰۱ |
| ۲۳/۷ | ۱/۶۹ | ۲/۳۶ | ۹۰۰-۸۰۱ |
| ۲۳/۴ | ۱/۶۷ | ۲/۳۴ | ۱۰۰۰-۹۰۱ |

- ارتباط: ارزیابی اینکه آیا خروجی به‌طور مناسب با ویژگی‌ها منطبق است یا خیر.
- روانی: ارزیابی سهولت استخراج معیارها
- قابلیت تفسیر: تعیین اینکه آیا تصمیمات و خروجی‌های هوش مصنوعی برای افراد غیرمتخصص قابل درک هستند یا خیر.

و همچنین ما مجموعه‌ای از آزمایش‌های فرسایش را انجام دادیم تا اهمیت هر جزء از مدل خود را نشان دهیم. همه مدل‌های فرسایش بر روی مجموعه داده آموزش داده شده و ارزیابی شدند. ما از یک پایه با یک شبکه تخصیص مولد شرطی ساده شروع کردیم که مقایسه خط مبناهای مختلف آن در جدول شماره ۱۶ آمده است:

³ FID

¹ Fold-5
² LPIPS

جدول (۱۹): اثر متاداده‌ها بر عملکرد مدل

| توضیح | F1-Score (%) | Accuracy (%) | پیکربندی مدل |
|-------------------------------|--------------|--------------|------------------------------------|
| بدون استفاده از متاداده | ۸۷/۲ | ۸۸/۳ | فقط داده تصویری پایه |
| استفاده از سن، جنس و نوع پوست | ۹۰/۸ | ۹۱/۴ | داده تصویری + متاداده (مدل ترکیبی) |

ترکیب داده‌های تصویری با متاداده‌های بالینی، به مدل اجازه داد تا اطلاعات زمینه‌ای مربوط به سن و نوع پوست را نیز در تصمیم‌گیری لحاظ کند. نتایج نشان دادند که این ترکیب موجب بهبود میانگین دقت و امتیاز اف یک به ترتیب به میزان ۳ تا ۴ درصد گردیده است.

۴-۷ ارزیابی انسانی

برای ارزیابی ادراکی کیفیت بازسازی، از ۱۰ متخصص پوست و زیبایی دعوت شد تا ۵۰ نمونه خروجی را به صورت کور و بر روی مقیاس پنج امتیازی لیکرت ارزیابی کنند (۱ = ضعیف تا ۵ = عالی). معیارهای ارزیابی شامل سازگاری رنگ پوست، حفظ ساختار چهره و قابلیت استفاده بالینی بود. میانگین امتیاز کلی برابر با $4/6 \pm 0/3$ از ۵ به دست آمد. آزمون ویلکاکسون با رتبه‌های امضا شده بین خروجی شبکه مولد تخصصی و تصاویر ورودی نویزی/کاهش یافته، بهبود معنادار را نشان داد.

۴-۸ تحلیل تلفیقی عملکرد مدل‌ها

ترکیب دو معماری سی ان ان و پیکس-توپیکس-گن در این مقاله منجر به ایجاد یک سیستم دوسویه شد که هم توانایی تشخیص نواحی نیازمند درمان زیبایی را دارد و هم قابلیت شبیه‌سازی تصویری نتایج درمان را به صورت قابل تفسیر برای کاربر فراهم می‌سازد.

تحلیل: در فرآیند ارزیابی، مدل شبکه عصبی کانولوشنی در بخش تشخیص نواحی آسیب‌دیده با دقت و شاخص اف یک بالا، عملکرد قابل توجهی نشان داد و تعادل مناسبی میان صحت پیش‌بینی و

۴-۴ نتایج تفکیک پوست

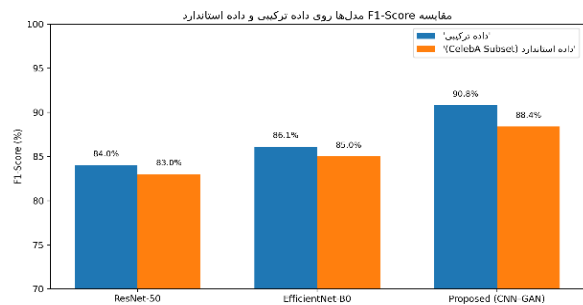
جدول (۱۸): عملکرد ماژول تشخیص ناحیه پوست

| مقدار (%) | معیار ارزیابی |
|-----------|---------------|
| ۹۳/۷ | IoU |
| ۹۶/۱ | Dice |

مقادیر بالای آی او یو (شاخص همپوشانی مبتنی بر اجتماع به هم پوشانده) و دایس (شاخص شباهت دایس) نشان‌دهنده دقت بالای مدل در تشخیص مرزهای واقعی ناحیه پوست هستند. این دقت بالا نقش مهمی در بهبود تحلیل بافتی و کاهش خطا در مراحل بعدی پردازش تصویر ایفا می‌کند.

۴-۵ مقایسه با داده استاندارد

به منظور بررسی و رفع نگرانی مربوط به گزینشی بودن داده‌ها، بخشی از آزمایش‌ها مجدداً بر روی زیرمجموعه‌ای از مجموعه داده سلب‌ای انجام شد. شکل زیر مقایسه امتیاز اف یک مدل‌ها بر روی داده ترکیبی و داده استاندارد. عملکرد پایدار روش پیشنهادی روی مجموعه استاندارد مشاهده می‌شود.



شکل (۸): مقایسه F1-Score مدل‌ها بر روی داده ترکیبی

با وجود تفاوت توزیع داده، روند برتری روش پیشنهادی روی هر دو نوع داده تأیید می‌شود؛ لذا مقایسه‌ها از نظر «روند» قابل استناد هستند.

۴-۶ اثر متاداده‌ها بر عملکرد مدل پیشنهادی

استفاده از متاداده موجب بهبود ۱/۳ درصد در دقت کلی و ۶/۳ درصد در امتیاز اف یک گردید، که نشان‌دهنده تأثیر مثبت داده‌های غیرتصویری در درک ویژگی‌های زیباشناختی چهره است.

و بخش بازسازی تصویر با مدل پیکس‌توپیکس کلاسیک مورد ارزیابی مقایسه‌ای قرار گرفته‌اند.

۴-۱۰ نوآوری مقاله حاضر در چارچوب ادبیات علمی

برخلاف بیشتر مقاله‌های پیشین که صرفاً به یکی از دو کارکرد (تشخیص یا شبیه‌سازی) پرداخته‌اند، در این مقاله:

- دو مدل عمیق با هدف‌های مکمل به‌صورت هماهنگ به کار گرفته شده‌اند؛
- پیشنهاد درمان بر اساس تحلیل چهره، شخصی‌سازی و معیارهای بالینی ارائه می‌شود؛
- تأییدپذیری بالینی خروجی‌ها نیز در طراحی مدل لحاظ شده است.

این ویژگی‌ها باعث می‌شود مقاله حاضر نه تنها از منظر مهندسی و پردازش تصویر، بلکه از منظر بالینی و کاربردی نیز نوآورانه و قابل استفاده تلقی شود. چنین رویکردی با چشم‌انداز پزشکی زیبایی مبتنی بر داده و هوش مصنوعی تعاملی کاملاً هم‌راستا است و امکان توسعه اپلیکیشن‌های کاربردی در محیط‌های بالینی را فراهم می‌سازد.

۵- نتیجه‌گیری

در این مقاله، یک سامانه ترکیبی مبتنی بر یادگیری عمیق ارائه شده است که با بهره‌گیری از شبکه عصبی کانولوشنی رزنت-۵۰ برای شناسایی خودکار نواحی دارای مشکلات پوستی و مدل شبکه مولد تخصصی پیکس‌توپیکس-گن برای شبیه‌سازی نتایج درمان، عملکردی برتر نسبت به روش‌های موجود به دست می‌آورد. این سامانه در شاخص‌های دقت (۹۴/۲ درصد)، شاخص اف‌یک (۹۳/۱ درصد)، نسبت اوج سیگنال به نویز ۳۱/۵ دسی‌بل و میانگین قدرمطلق خطا ۰/۰۲۳ نتایج قابل توجهی ارائه کرده و هم‌چنین موفق به کسب رضایت بالینی متخصصان زیبایی (۴/۶ از ۵) شده است. این سیستم علاوه بر ارتقای دقت و قابلیت تفسیر، توانایی پردازش داده‌های متنوع و ارائه خروجی واقع‌نما متناسب با ویژگی‌های فردی بیماران را نشان داده است. با این حال، چالش‌هایی همچون نیاز به داده‌های گسترده‌تر و متنوع‌تر، وابستگی به کیفیت

بازخوانی^۱ برقرار کرد. هم‌چنین، مدل شبکه مولد تخصصی در بازسازی درمان، با معیارهای عددی مناسب نظیر میانگین قدرمطلق خطا و نسبت اوج سیگنال به نویز، کیفیت بالایی ارائه داد. ترکیب این نتایج، همراه با تأیید کیفی متخصصان درماتولوژی، حاکی از اعتبار بالای مدل در کاربردهای بالینی واقعی است.

۴-۹ مقایسه با مقاله‌های پیشین

نتایج این تحقیق در راستای یافته‌های گیلانی و مارکز [۶] قرار دارد که در بررسی جامع خود نشان دادند شبکه‌های مولد متخصص قابلیت تولید تصاویر درماتوسکوپی بسیار واقع‌گرایانه از ضایعات پوستی دارند که حتی متخصصان پوست‌شناس نیز به‌سختی می‌توانند آن‌ها را از تصاویر واقعی تشخیص دهند این ویژگی، به‌ویژه در مواجهه با چالش کمبود داده که یکی از محدودیت‌های اساسی سامانه‌های یادگیری عمیق در تحلیل ضایعات پوستی محسوب می‌شود، اهمیت بالایی دارد. هم‌چنین در تأیید رویکرد علی و همکاران [۷] که از شبکه‌های مولد متخصص شرطی برای بخش‌بندی ضایعات ملانوما در چارچوب اینترنت اشیا پزشکی استفاده کردند، نتایج مطالعه حاضر نشان می‌دهد که تولید تصاویر بخش‌بندی شده بر مبنای تصاویر درماتوسکوپی ضایعات پوستی منجر به دستیابی به نتایج بخش‌بندی دقیق‌تری می‌شود. این رویکرد به‌ویژه در غیاب متخصصان و در محیط‌های راه دور که سیستم‌های تشخیص کمکی کامپیوتری مورد نیاز هستند، اهمیت بالایی دارد. با توجه به تفاوت ماهیت مسئله حاضر، که بر تشخیص نیاز به درمان زیبایی در سطح چهره تمرکز دارد، نسبت به بسیاری از مطالعات پیشین که عمدتاً بر طبقه‌بندی ضایعات پوستی یا بازسازی تصویر متمرکز بوده‌اند، انجام مقایسه عددی مستقیم با آن‌ها امکان‌پذیر نیست. این محدودیت ناشی از تفاوت در نوع دیتاست‌ها، روش‌های تقسیم‌بندی داده و معیارهای ارزیابی مورد استفاده است.

از این رو، به‌منظور ارزیابی دقیق عملکرد مدل پیشنهادی، هر یک از ماژول‌های آن به‌صورت مستقل با روش‌های مرجع در همان حوزه مقایسه شده است؛ به‌گونه‌ای که بخش تفکیک با مدل‌های یونت و دیپ‌لب-وی^۳، بخش طبقه‌بندی با رزنت-۵۰ و ایفیشنت،

² DeepLab-V3

¹ Recall

بالای ضریب دایس و هم‌پوشانی (بیش از ۹۳٪) تأیید می‌کند که ورودی‌های سی ان ان و گن کاملاً با محدوده‌های واقعی پوست هم‌راستا بوده‌اند.

۶- اخلاق مقاله

ملاحظات اخلاقی و دسترس‌پذیری داده‌ها: در این مطالعه از داده‌های عمومی و ناشناس سلب‌ای، اف‌اچ‌کیو و درم‌نت استفاده شده است؛ بنابراین، طبق دستورالعمل‌های اخلاق مقاله، اخذ رضایت‌نامه فردی و مصوبه کمیته اخلاق الزامی نبوده است. پیوند و مشخصات هر پایگاه در پاورقی/منابع آمده است. در صورت استفاده از داده‌های کلینیکی اختصاصی، اطلاعات مربوط به کد اخلاق و فرآیند ناشناس‌سازی افزوده خواهد شد.

References

- [1] H. Ding, E. Zhang, F. Fang, X. Liu, H. Zheng, H. Yang, Y. Ge, Y. Yang, and T. Lin, "Automatic identification of benign pigmented skin lesions from clinical images using deep convolutional neural network," *BMC Biotechnol.*, vol. 22, p. 28, 2022. [Online]. Available: <https://doi.org/10.1186/s12896-022-00755-5>.
- [2] E. Tokgöz and M. A. Carro, "Applications of artificial intelligence, machine learning, and deep learning on facial plastic surgeries," in *Cosmetic and Reconstructive Facial Plastic Surgery: A Review of Medical and Biomedical Engineering and Science Concepts*, Springer, 2023.
- [3] Galić, M. Habijan, H. Leventić, and K. Romić, "Machine learning empowering personalized medicine: A comprehensive review of medical image analysis methods," *Electronics*, vol. 12, p. 4411, 2023. [Online]. Available: <https://doi.org/10.3390/electronics12214411>.
- [4] M. A. Kassem, K. M. Hosny, R. Damaševićius, and M. M. Eltoukhy, "Machine learning and deep learning methods for skin lesion classification and diagnosis: A systematic review," *Diagnostics*, vol. 11, p. 1390, 2021, pp. 423-445. [Online]. Available: <https://doi.org/10.3390/diagnostics11081390>
- [5] S. S. Noronha, M. A. Mehta, D. Garg, K. Kotecha, and A. Abraham, "Deep learning-based dermatological condition detection: A systematic review with recent methods, datasets, challenges, and future directions," *IEEE Access*, vol. 11, pp. 140348-140381, 2023. DOI: 10.1109/ACCESS.2023.3339635.
- [6] S. Q. Gilani and O. Marques, "Skin lesion analysis using generative adversarial networks: A review," *Multimedia Tools Appl.*, vol. 82, no. 19, pp. 30065-30106, 2023. [Online]. Available: <https://doi.org/10.1007/s11042-022-14267-z>.
- [7] Z. Ali, S. Naz, H. Zaffar, J. Choi, and Y. Kim, "An IoMT-based melanoma lesion segmentation using conditional generative adversarial networks," *Sensors*, vol. 23, no. 7, p. 3548, 2023. [Online]. Available: <https://doi.org/10.3390/s23073548>.
- [8] J. F. Santos, L. del R. Silva-Calpa, F. G. de Souza, and K. Pal, "'Central Countries' and Brazil's Contributions to Nanotechnology," *Curr. Nanomater.*, vol. 9, pp. 109-147, 2024. [Online]. Available: <https://doi.org/10.2174/2405461508666230522564138>.
- [9] L. Zang, X. Zhang, and B. Guo, "Federated deep reinforcement learning for online task offloading and resource allocation in WPC-MEC networks," *IEEE Access*, vol. 10, pp. 9856-9867, 2022. DOI: 10.1109/ACCESS.2022.3144415.
- [10] A. Alshardan, S. Alahmari, M. Alghamdi, M. AL Sadig, A. Mohamed, and G. P. Mohammed, "GAN-based synthetic medical image augmentation for class imbalanced dermoscopic image analysis," *Fractals*, vol. 33, p. 2540039, 2025. [CrossRef].
- [11] K. Behara, E. Bhero, and J. T. Agee, "Skin Lesion Synthesis and Classification Using an



- Improved DCGAN Classifier," *Diagnostics*, vol. 13, p. 2635, 2023. [CrossRef] [PubMed].
- [12] E. Goceri, "GAN-based augmentation using a hybrid loss function for dermoscopy images," *Artif. Intell. Rev.*, vol. 57, p. 234, 2024. [CrossRef].
- [13] S. Innani, P. Dutande, U. Baid, V. Pokuri, S. Bakas, S. Talbar, B. Baheti, and S. C. Guntuku, "Generative adversarial networks based skin lesion segmentation," *Sci. Rep.*, vol. 13, p. 13467, 2023. [CrossRef] [PubMed].
- [14] A. Tavakkoli, "Generative Adversarial Networks in Medicine: Important Considerations for this Emerging Innovation in Artificial Intelligence," *Ann. Biomed. Eng.*, vol. 51, pp. 2130–2142, 2023. [CrossRef].
- [15] E. Perez and S. Ventura, "Progressive growing of GANs for improving data augmentation and skin cancer diagnosis," *Artif. Intell. Med.*, vol. 141, p. 102556, 2023. [CrossRef].
- [16] M. Salvi, S. Bassoli, F. Molinari, M. Havaei, A. Depeursinge, "Generative models for color normalization in pathology and dermatology," *Expert. Syst. Appl.*, vol. 245, p. 123105, 2024. [CrossRef].
- [17] N. Veeramani and P. Jayaraman, "A promising AI-based super resolution image reconstruction technique for early diagnosis of skin cancer," *Sci. Rep.*, vol. 15, p. 5084, 2025. [CrossRef] [PubMed].
- [18] T. Karras, S. Laine, S. Aittala, J. Hellsten, J. Lehtinen, T. Aila, "Analyzing and Improving the Image Quality of StyleGAN," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 8107–8116.
- [19] P. Isola, J. Yan Zhu, T. Zhou, A. A. Efros, "Image-to-Image Translation with Conditional 1611.07004v3
- [20] H. Sangkeun Jung, "Variants of GAN: CGAN, Pix2Pix, CycleGAN, 2024, <https://medium.com/@hugmanskj/variants-of-gan-cgan-pix2pix-cyclegan-f7869de53152>.
- [21] A. Ray, A. Gupta, A. AI, "Skin Lesion Classification With Deep Convolutional Neural Network: Process Development and Validation," **JMIR Dermatology**, vol. 3, no. 1, p. e18438, May 2020. doi:10.2196/18438.
- [22] S. Innani, P. Dutande, U. Baid, V. Pokuri, S. Bakas, S. Talbar, B. Baheti, S. Chandra Guntuku, — "Generative adversarial networks based skin lesion segmentation", 2023, <https://doi.org/10.1038/s41598-023-39648-8>.

Intelligent detection of areas requiring cosmetic treatment and simulation of treatment results using Convolutional Neural Networks and Generative Adversarial Networks

Arman Zavvari¹, Hamid Soleimanimehr^{2,3*}

¹M.Sc. Student, Department of Mechatronics Engineering, SR.C., Islamic Azad University, Tehran, Iran

²Associate Professor, Department of Mechanical Engineering, SR.C., Islamic Azad University, Tehran, Iran

³ Associate Professor, Modern Automotive Research Center, SR.C., Islamic Azad University, Tehran, Iran

Article Information

Original Research Paper

Received:

2025 October 05

Accepted:

2026 February 21

Keywords:

Artificial Intelligence, Deep Learning, Cosmetic Dermatology, Convolutional Neural Network, Generative Adversarial Network

Corresponding Author*:

ha.sol@iau.ac.ir

Abstract

Nowadays, the demand for precise skin condition analysis and for predicting cosmetic treatment outcomes prior to intervention has become a critical requirement in dermatology and aesthetic medicine. In this study, we developed an intelligent deep learning-based framework for detecting facial areas requiring cosmetic intervention and for simulating post-treatment outcomes. To this end, a Convolutional Neural Network (CNN) was first employed for the automated identification of abnormalities such as wrinkles, pigmentation, and asymmetry. The proposed framework was implemented in Python using the PyTorch library and trained on a comprehensive dataset comprising 5,000 facial images from validated databases (CelebA, FFHQ, DermNet) spanning various age groups and skin types with rigorous manual annotations by three independent dermatology specialists for affected regions. Model evaluation was conducted based on quantitative metrics, including accuracy (94.2%), F1-Score (93.1%), MAE (0.023), and PSNR (31.5 dB), alongside an expert satisfaction score of 4.6 out of 5. The results demonstrate that the combined GAN-CNN architecture not only achieves precise identification of areas in need of intervention but also offers highly realistic and interpretable visualizations of treatment outcomes with integrated clinical personalization parameters.

 : 10.22034/ABMIR.2026.23775.1176

E-ISSN: [2821-2037](https://doi.org/10.22034/ABMIR.2026.23775.1176) /© 2026. Published by Yazd University This is an open access article under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).

