

اندازه‌گیری دور سر نوزاد با استفاده از تکنیک‌های یادگیری عمیق

فرزاد زارع زاده^{۱*}، علی حسینی نوه^۲، زهره حبیبی^۳

^۱ دانشجوی کارشناسی ارشد مهندسی نقشه‌برداری، فتوگرامتری - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی

خواجه‌نصیرالدین طوسی

farzadzare3511@gmail.com

^۲ استادیار دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه‌نصیرالدین طوسی

ali_hosseini_naveh@yahoo.com

^۳ استادیار دانشکده پزشکی - دانشگاه علوم پزشکی و خدمات بهداشتی درمانی تهران

z-habibi@sina.tums.ac.ir

(تاریخ دریافت آذر ۱۳۹۷، تاریخ تصویب اسفند ۱۳۹۷)

چکیده

اندازه‌گیری دور سر نوزادان و پایش روند رشد دور سر، نقش مهمی را در تشخیص برخی بیماری‌ها ایفا می‌کند. با توجه به مشکلات اندازه‌گیری تماسی از قبیل انتقال بیماری، عفونت و راحت نبودن نوزاد در هنگام اندازه‌گیری، رفتن به‌سوی اندازه‌گیری‌های غیر تماسی امری اجتناب‌ناپذیر می‌باشد. هدف از این پژوهش، ارائه روشی غیر تماسی و تصویر مینا به‌منظور اندازه‌گیری دور سر نوزاد می‌باشد. در این پژوهش الگوریتمی ارائه شد که دور سر نوزاد را با استفاده از یک تصویر گرفته‌شده از بالای سر نوزاد و شاخص مقیاس موجود در کنار سر محاسبه می‌نماید. گام اول در محاسبه محیط سر، شناسایی و قطعه‌بندی سر نوزاد در تصویر می‌باشد که در این پژوهش از دو روش MaskR-CNN و CRF-RNN برای شناسایی و قطعه‌بندی سر نوزاد استفاده شد. بعداز آن پیکسل‌های تشکیل‌دهنده دور سر، با ترکیب الگوریتم لبه یاب Canny و الگوریتم‌های ریخت‌شناسی در تصویر مشخص شد. در مرحله بعد رزولوشن مکانی پیکسل‌ها در واقعیت، با توجه به برچسب مقیاس موجود در تصویر محاسبه شد. درنهایت با توجه به رزولوشن مکانی به‌دست‌آمده و تعداد پیکسل‌های تشکیل‌دهنده دور سر، محیط سر محاسبه شد. ارزیابی‌ها نشان می‌دهد که روش MaskR-CNN با دقت کلی ۹۸/۸ درصد روش مناسب‌تری از روش CRF-RNN برای شناسایی و قطعه‌بندی سر در تصویر می‌باشد. همچنین با مقایسه نتایج به‌دست‌آمده از الگوریتم ارائه‌شده، با مقادیر واقعی به‌دست‌آمده به‌وسیله متر نواری بر روی ۱۰ تصویر، مشخص شد خطای روش ارائه‌شده در حدود ۱ تا ۳ درصد می‌باشد.

واژگان کلیدی: اندازه‌گیری غیر تماسی، یادگیری عمیق، شبکه عصبی پیچشی، تشخیص شیء، قطعه‌بندی

۱- مقدمه

امروزه پیشرفت تکنولوژی و رفتن به سمت اتوماسیون در تمام عرصه‌ها اجتناب‌ناپذیر می‌باشد. رشته پزشکی و تجهیزات آن نیز از این امر مستثنا نیست. یکی از علومی که می‌تواند در این امر به رشته پزشکی کمک کند، علم فتوگرامتری می‌باشد. از اولین روزهای پیدایش علم فتوگرامتری، فتوگرامتریست‌هایی بودند که تحقیقاتی در حوزه اندازه‌گیری‌های پزشکی انجام داده‌اند. به‌عنوان مثال نگین اکبری و همکاران [۱] روشی برای آشکارسازی رگ‌های زیرپوستی با استفاده از مدل‌سازی سه‌بعدی با کمک دوربین میکرو لنز مادون‌قرمز ارائه دادند یا آرش فتحی و همکاران [۲] روشی به‌منظور پایش درجه حرارت محدوده گوش هنگام مکالمه با تلفن همراه با استفاده از فتوگرامتری برد کوتاه مبتنی بر تصاویر حرارتی ارائه نمودند. علی‌رغم فعالیت‌های فتوگرامتری زیادی که در حوزه پزشکی شده است فتوگرامتری پزشکی هنوز نتوانسته دنیای پزشکی را متحول کند [۳].

بر اساس آمارهای رسمی منتشرشده توسط سازمان ثبت‌احوال کشور روزانه حدود چهار هزار نوزاد در کشور متولد می‌شود [۴]. رشد و نمو طبیعی نوزاد، نشانه صحت و سلامت او در حال و آینده است. یکی از مهم‌ترین شاخصه‌های رشد و نمو طبیعی نوزاد در سال اول تولد، اندازه دور سر نوزاد می‌باشد. این شاخصه به دلیل اینکه ۸۰ درصد رشد مغز نوزاد در سال اول تولد صورت می‌گیرد از اهمیت بسیار بالایی برخوردار می‌باشد و اگر مشکلی ایجاد شود جبران‌ناپذیر خواهد بود. از طرف دیگر بیماری‌هایی وجود دارند که به پایش روند رشد دور سر بسیار نیاز دارند. این بیماری‌ها به دو دسته میکروسفالی^۱ و ماکروسفالی^۲ تقسیم می‌شوند. در دسته ماکروسفالی بیماری به نام هیدروسفالی^۳ وجود دارد [۵]. هیدروسفالی در پزشکی به وضعیتی گفته می‌شود که مایع مغزی-نخاعی تولیدشده در مغز، بیش‌ازاندازه در بطن‌های مغزی و دیگر حفره‌های مغز انباشته می‌گردد و منجر به افزایش حجم آن‌ها می‌شود. این افزایش حجم اشغال‌شده توسط مایع در سیستم عصبی مرکزی و به‌ویژه مغز می‌تواند ناشی

از اختلال در توزیع، تشکیل، جریان یا جذب مایع مغزی-نخاعی باشد و پی آمد آن افزایش فشار درون جمجمه و آسیب‌های مغزی و حتی مرگ است. هیدروسفالی در کودکان زیر دو سال به خاطر باز بودن درزهای استخوان جمجمه منجر به بزرگ شدن غیرطبیعی جمجمه می‌شود [۵]. به‌طورمعمول نحوه شناسایی، تشخیص و کنترل این بیماری‌ها به شش صورت انجام می‌شود که عبارت‌اند از تشخیص به‌صورت ارزیابی بصری، اندازه‌گیری مداوم دور سر نوزاد با متر نواری، اندازه‌گیری شاخص‌های سفالومتری^۴ با کولیس، روش پلاژیوسفالو متری^۵، تصویربرداری رادیولوژی و تصویربرداری سه‌بعدی می‌باشد [۶]؛ که در این تحقیق روش اندازه‌گیری دور سر مدنظر می‌باشد. در حال حاضر برای اندازه‌گیری دور سر نوزادان در تمام بیمارستان‌ها از نوار مدرج پلاستیکی یا فلزی استفاده می‌شود که این نوار مدرج را بین موهای سر و ابروی کودک قرار می‌دهند و دور سر را در مقطع پیشانی و برجستگی پس سر اندازه‌گیری می‌کنند. به‌عبارت‌دیگر با داشتن دید از بالا می‌توان گفت که بزرگ‌ترین منحنی دور سر نوزاد با استفاده از این نوار مدرج اندازه‌گیری می‌شود.

با توجه به مشکلات اندازه‌گیری تماسی از جمله احتمال انتقال بیماری از یک کودک به کودک دیگر، احتمال عفونت به دلیل اندازه‌گیری‌های لازم بعد از عمل و همچنین راحت نبودن نوزاد و عامل اندازه‌گیر در هنگام اندازه‌گیری، استفاده از یک روش غیر تماسی به‌منظور اندازه‌گیری محیط دور سر نوزاد ضروری می‌باشد.

با توجه به تحقیقات صورت گرفته تاکنون روشی برای اندازه‌گیری محیط دور سر نوزاد توسط محققین ارائه نشده است در سال ۲۰۱۶ Mummolo و همکاران [۷] روشی برای ردیابی شاخصه‌های سفالومتری در فضای سه‌بعدی با استفاده از تکنیک Tele Motion Tracking ارائه نمودند؛ و در سال ۲۰۱۷ Garcia و همکاران [۸] روشی جدید برای ارزیابی تغییر شکل جمجمه نوزاد با استفاده از تصاویر ارائه داد.

هدف از این پژوهش ارائه روشی غیر تماسی و تصویر مینا می‌باشد که بتواند محیط دور سر نوزاد را اندازه‌گیری نماید و چشم‌انداز این تحقیق در آینده تهیه یک نرم‌افزار

۴ cephalometry
۵ pelagia cephalometry

۱ microcephaly
۲ Macrocephaly
۳ hydrocephaly

در سال ۲۰۱۴ Byoung-Keon Park و همکاران [۱۰] یک روش جدید برای اندازه‌گیری شکل بدن نوزادان از روی تصاویر دارای نویز با هزینه پایین ارائه نمودند. در سال ۲۰۱۵ Nir Sokolover و همکاران [۱۱] یک سیستم استریو توسعه دادند که بعد از کالیبره شدن با تست فیلد مربوطه آماده عکس‌برداری از نوزاد می‌شود و پس از عکس گرفتن از نوزاد در این سیستم قسمت‌های خاصی از بدن نوزاد به صورت نقاط کلیدی در نظر گرفته می‌شود، به‌عنوان مثال بالای سر، زیر چانه، بالای لگن، سر زانو و کف پا به‌عنوان نقاط کلیدی در نظر گرفته می‌شود و فاصله بین آن‌ها به دست می‌آید و بعد از آن تمام این فواصل باهم جمع شده و قد نوزاد را ارائه می‌دهد مزیت این سیستم این است که دیگر برای به دست آوردن قد نوزاد نیاز به کشیدن پاهای نوزاد برای اندازه‌گیری نیست و در این روش راحتی و آرامش نوزاد در نظر گرفته شده است. در سال ۲۰۱۷ نگین اکبری و همکاران [۱۲] از فتوگرامتری برد کوتاه برای اندازه‌گیری قطر مردمک چشم انسان جهت بررسی بیماری آنیزوکوریا (ناهمسانی مردمک چشم‌ها)، استفاده نمودند. برای این کار یک سیستم اندازه‌گیری در قالب اجزاء اصلی طراحی شبکه، فیلم‌برداری از عارضه در تاریکی مطلق تحت نور، مدل‌سازی سه‌بعدی و استخراج و اندازه‌گیری قطر مردمک دو چشم طراحی و پیاده‌سازی گردیده است. نتایج نشان‌دهنده قابلیت سیستم معرفی‌شده در دستیابی به مدل‌سازی چشم انسان به‌وسیله تصاویر و استخراج مردمک چشم‌ها از تصاویر و اندازه‌گیری قطر آن‌ها برای تشخیص بیماری می‌باشد. در تمام این روش‌ها اجزای بدن مدنظر می‌باشد و هیچ‌کدام از آن‌ها اندازه‌گیری‌هایی بر روی سر انجام نداده‌اند.

در سال ۲۰۱۰ Sen Wong و همکاران [۱۳] به‌منظور اندازه‌گیری توپوگرافی سطح سر از ۸ دوربین ارزان‌قیمت در ترکیب با تکنیک‌های سرشکنی باندل و سلف کالیبراسیون استفاده نمودند. بدین منظور تارگت‌هایی را بر روی سر گذاشته و از آن‌ها تصویربرداری می‌شود سپس با استفاده از تکنیک‌های شناسایی تارگت به شناسایی تارگت‌ها می‌پردازد و توپوگرافی سطح سر را اندازه‌گیری می‌نماید. در سال ۲۰۱۲ Bruce King و همکاران [۱۴]

تلفن همراه به‌منظور اجرای آنی این روش می‌باشد تا پزشکان که افرادی غیرمتخصص در حوزه فتوگرامتری می‌باشند بتوانند از این روش به‌راحتی استفاده نمایند. در این روش ابتدا یک تصویر از بالای سر نوزاد اخذ می‌شود. سپس شناسایی و قطعه‌بندی سر نوزاد در این تصویر انجام می‌گیرد. در مرحله بعد لبه‌های دور سر استخراج شده و از طرف دیگر با شناسایی برچسب مقیاس و مشخص بودن اندازه آن در واقعیت، رزولوشن مکانی هر پیکسل در واقعیت مشخص شده و با استفاده از مقیاس به‌دست‌آمده و تعداد پیکسل‌های تشکیل‌دهنده دور سر، محیط دور سر محاسبه می‌شود. با استفاده از این روش می‌توان اندازه‌گیری محیط دور سر نوزاد را به‌صورت غیر تماسی انجام داد که در این صورت ضدعفونی کردن متر در هر بار استفاده برای اندازه‌گیری سر نوزادان مختلف کنار گذاشته می‌شود و احتمال انتقال بیماری‌های پوستی از یک نوزاد به نوزاد دیگر به صفر می‌رسد و همچنین در هنگام اندازه‌گیری، هردوی اپراتور اندازه‌گیر و نوزاد در هنگام اندازه‌گیری مشکلات قبلی (تکان خوردن نوزاد، بی‌تابی و گریه کردن، رنجش اپراتور اندازه‌گیر به دلیل تکان‌های پی‌درپی سر نوزاد) را نخواهند داشت.

در این مقاله ابتدا مروری بر تحقیقات گذشته صورت گرفته است. بعد از آن مروری کلی بر روش‌های قطعه‌بندی انجام‌گرفته، سپس روش پیشنهادی در این تحقیق به‌تفصیل بیان‌شده است و بعد از آن پیاده‌سازی روش موردنظر شرح داده می‌شود و در ادامه ارزیابی نتایج به‌دست‌آمده شرح داده شده است و در نهایت نتیجه‌گیری انجام شده است.

۲- مروری بر تحقیقات گذشته

همان‌طور که اشاره شد، در تحقیقات صورت گرفته تاکنون روشی برای اندازه‌گیری محیط دور سر نوزاد با استفاده از تصویر توسط محققین ارائه نشده است؛ اما این بدین معنا نیست که تاکنون هیچ محقق، روشی برای اندازه‌گیری اجزای دیگر نوزاد یا به‌طور کل اجزای بدن انسان ارائه نداده باشد. به‌عنوان مثال در سال ۲۰۱۳ Jamie Shotton و همکاران [۹] یک روش جدید برای پیش‌بینی دقیق و سریع موقعیت سه‌بعدی بدن انسان از روی تک تصویر عمق بدون نیاز به اطلاعات فریم قبلی ارائه نمودند.

یک سیستم اندازه‌گیری غیر تماسی بر پایه تصویر ارائه دادند که متشکل از ۴ دوربین رایانه^۱ می‌باشد که برای اندازه‌گیری اجزای صورت نوزاد استفاده می‌شود. نکته اساسی در این کار استفاده از دوربین رایانه ارزان قیمت به جای استفاده از دوربین‌های دیجیتال گران قیمت بود که هزینه ساخت را بسیار پایین آورده است. در این روش‌ها اندازه‌گیری بر روی سر، صورت گرفته ولی هدف این روش‌ها به دست آوردن شاخص‌های سفالومتری نمی‌باشد. در سال ۲۰۱۴ Pogulwar و همکاران [۱۵] با استفاده از عکس اقدام به اندازه‌گیری شاخص‌های سفالومتری نمودند و توانستند حدود ۶۰ درصد از اندازه‌گیری‌های سفالومتری را با استفاده از عکس به دست آورند. در سال ۲۰۱۶ نیز Mummolo و همکاران روشی برای ردیابی شاخص‌های سفالومتری در فضای سه‌بعدی با استفاده از تکنیک Tele Motion Tracking ارائه نمودند. برای انجام عکس‌برداری از دو دوربین و یک لامپ زرد رنگ و تابش مکرر مادون قرمز استفاده نمودند و نتایج حاصل از این روش را با نتایج حاصل از سفالومتری رادیوگرافی مقایسه نمودند و به این نتیجه رسیدند که اندازه‌های حاصل از دو روش از همبستگی معناداری برخوردار می‌باشد و در انتها این تکنیک را یک روش قابل قبول برای مطالعه مورفولوژی صورت و همچنین برای نظارت در حین درمان بیمار معرفی نمودند. در سال ۲۰۱۷ Garcia و همکاران روشی جدید برای ارزیابی تغییر شکل جمجمه نوزاد با استفاده از تصاویر ارائه داد. در این روش با استفاده از یک گوشی هوشمند ویدیویی با سرعت آهسته در ۵ حالت مختلف سر ضبط شد و بعد از پردازش‌های لازم مدل سه‌بعدی سر تهیه گردید و نتایج آن با اندازه‌گیری شاخص‌های سفالومتری با استفاده از کولیس مورد مقایسه قرار گرفت که نتایج به دست آمده با اندازه‌گیری‌های انجام شده با کولیس مطابقت داشت. در تمام این روش‌هایی که به منظور به دست آوردن شاخص‌های سفالومتری می‌باشد از چندین تصویر استفاده شده است و مدل سه‌بعدی از سر ایجاد شده است که این روش‌ها علاوه بر نیاز به تعداد تصاویر بالا، به دانش فتوگرامتری نیز نیاز دارد که این مسئله کار با این روش‌ها را برای پزشکان سخت می‌کند.

در این پژوهش نیاز به توسعه دستگاه نمی‌باشد بلکه یک الگوریتم توسعه داده شده تا بتواند محیط دور سر نوزاد که یکی از شاخص‌های سفالومتری می‌باشد را به دست آورد. اولین گام در محاسبه محیط دور سر نوزاد با استفاده از تک تصویر، شناسایی و قطع‌بندی سر نوزاد در تصویر می‌باشد.

۳- روش‌های قطع‌بندی

تاکنون روش‌های زیادی برای قطع‌بندی تصاویر ارائه شده است. در یک تقسیم‌بندی کلی می‌توان روش‌های قطع‌بندی را به دودسته سنتی و روش‌های مبتنی بر یادگیری عمیق تقسیم‌بندی نمود [۱۶].

Cheng و همکاران [۱۷] الگوریتم‌های قطع‌بندی سنتی را به ۷ دسته تقسیم نمودند که عبارت‌اند از: روش‌های آستانه‌گذاری هیستوگرام، روش‌های مبتنی بر خوشه‌بندی، روش‌های مبتنی بر ناحیه، روش‌های آشکارسازی لبه، روش‌های فازی، روش‌های مبتنی بر شبکه‌های عصبی، روش‌های مبتنی بر فیزیک. در این پژوهش به الگوریتمی نیاز است که در تصاویر با شرایط نوری مختلف و همچنین پس‌زمینه‌های گوناگون، فقط سر را به صورت دقیق قطع‌بندی کند و نیز از سرعت قابل قبولی برخوردار باشد. سرعت الگوریتم به دلیل چشم‌انداز این تحقیق در آینده یعنی استفاده از روش به صورت آنی قابل توجه می‌باشد.

به‌رحال با پیشرفت‌های اخیر در زمینه یادگیری عمیق بسیاری از روش‌های سنتی منسوخ شده بنابراین برای انجام پروژه‌های دقیق باید از روش‌های نوین که بر روی مجموعه داده‌های^۲ معروف اجرا شده و به‌دقت و سرعت قابل قبولی دست‌یافته استفاده نمود.

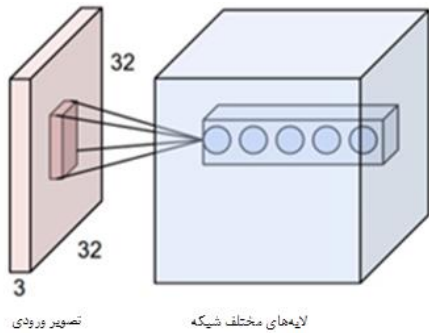
قبل از بررسی الگوریتم‌های قطع‌بندی مبتنی بر یادگیری عمیق لازم است تا یک توضیح کوتاه در مورد خود یادگیری عمیق داده شود.

در یادگیری عمیق شبکه‌های عصبی چندلایه وجود دارد. در این شبکه‌ها یادگیری چندلایه ویژگی‌ها صورت می‌گیرد. یادگیری چندلایه ویژگی‌ها به این معنی است که به‌عنوان مثال در شکل (۱)، یک شبکه عصبی عمیق وجود دارد که شامل یک‌لایه ورودی، سه لایه مخفی و یک‌لایه خروجی می‌باشد. ورودی این شبکه تصاویر چهره می‌باشد و

^۲ Data set

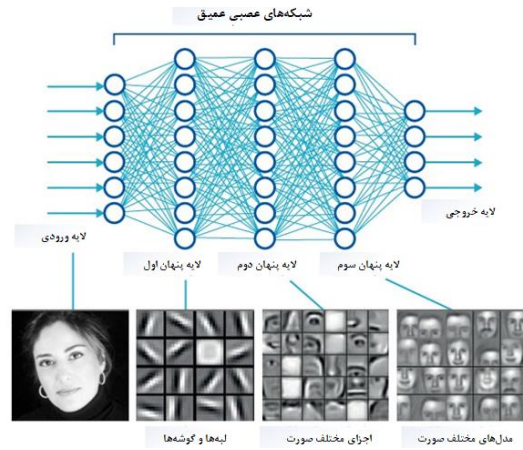
^۱ webcam

عمیق هستند که در آن‌ها چندین لایه با روشی قدرتمند آموزش می‌بینند [۱۹]. این روش بسیار کارآمد بوده و یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی کامپیوتر است. به‌طور کلی، یک شبکه عصبی پیچشی از سه لایه اصلی تشکیل می‌شود که عبارت انداز: لایه پیچشی^۱، لایه ادغام^۲ و لایه تماماً متصل^۳. لایه‌های مختلف وظایف مختلفی را انجام می‌دهند. لایه ورودی شامل مقادیر پیکسل‌های خام تصویر ورودی هستند. لایه پیچشی خروجی نورون‌هایی که به نواحی محلی در ورودی متصل هستند را محاسبه می‌کند. عمل محاسبه هم از طریق ضرب نقطه‌ای، بین وزن‌های هر نورون و ناحیه‌ای که نورون‌های لایه پیچشی به لایه ورودی متصل هستند صورت می‌گیرد. نمونه‌ای از اتصال محلی نورون‌ها در شکل (۳) مشاهده می‌شود. لایه ادغام ابعاد توده ورودی (تصویر) را کاهش می‌دهد. لایه تماماً متصل وظیفه محاسبه امتیاز دسته‌ها را دارد. همان‌طور که از اسم این لایه مشخص است هر نورون در این شبکه با تمام نورون‌ها در لایه قبل از خود ارتباط دارد. نتیجه کار این لایه یک بردار امتیاز می‌باشد که به هر دسته یا کلاس موجود یک امتیاز می‌دهد. با این روش، شبکه پیچشی، مقادیر پیکسل‌های خام تصویر اصلی را، لایه به لایه به امتیاز دسته‌ها در انتهای شبکه تبدیل می‌کند. در این شبکه‌ها از تابع فعال‌ساز نیز استفاده می‌شود. این توابع در واقع یک قسمت در شبکه‌های عصبی هستند که ورودی آن‌ها یک عدد (کوچک یا بزرگ در بازه‌ی دلخواه) است و خروجی آن‌ها معمولاً یک عدد بین ۰ و ۱، یا -۱ و ۱ است. در واقع این توابع یک عدد ورودی را به یک بازه مشخص (مثلاً -۱ تا ۱) تبدیل می‌کنند. نمونه‌ای از شبکه‌های عصبی پیچشی در شکل (۴) مشاهده می‌شود.

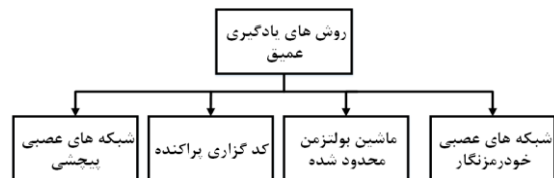


۱ convolution layer
 ۲ pooling layer
 ۳ fully connected layer

خروجی آن کلاس مربوط به تصویر موردنظر می‌باشد. تصویر ورودی به‌صورت برداری از پیکسل‌ها، وارد شبکه می‌شود. در لایه پنهان اول شبکه یاد می‌گیرد که ویژگی‌های سطح پایین را استخراج کند. منظور از ویژگی‌های سطح پایین لبه‌ها و گوشه‌ها می‌باشد که این ویژگی‌ها معنادار نیستند. ویژگی‌های به‌دست‌آمده در لایه اول به‌عنوان ورودی در لایه بعدی استفاده می‌شود و در لایه بعدی ویژگی‌های پیچیده‌تر و معنادارتر که از ترکیب لبه‌ها و گوشه‌ها استخراج شده است تشخیص داده می‌شود که در این مثال اجزای مختلف صورت می‌باشد؛ و به همین ترتیب در لایه بعدی ویژگی‌های سطح بالاتری مثل مدل‌های مختلف صورت تشخیص داده می‌شود و با توجه به کاربرد هر شبکه تعداد لایه‌های آن متفاوت می‌باشد. در نهایت در لایه خروجی کلاس تصویر ورودی مشخص می‌شود.

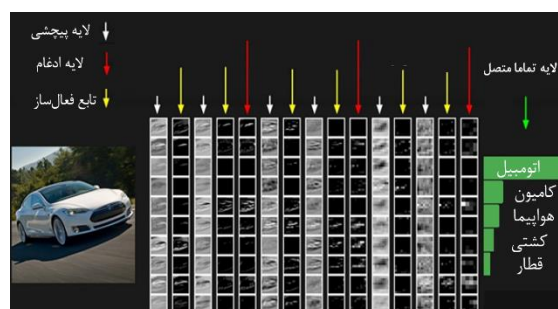


طی سال‌های اخیر، یادگیری عمیق به‌صورت گسترده در حوزه بینایی کامپیوتر مورد مطالعه قرار گرفته است و به همین دلیل، تعداد زیادی از روش‌های مرتبط با آن به وجود آمده است. به‌طور کلی، این روش‌ها بر اساس روش پایه‌ای که از آن مشتق شده‌اند به ۴ دسته مختلف تقسیم می‌شوند که می‌توان آن‌ها را در شکل (۲) مشاهده نمود.



در این پژوهش از شبکه‌های عصبی پیچشی (CNN) استفاده شده است که یکی از مهم‌ترین روش‌های یادگیری

این نوع از معماری شبکه‌های عصبی که در بالا توضیح داده شد برای طبقه‌بندی استفاده می‌شود. برای تبدیل این شبکه‌های طبقه‌بندی به یک شبکه قطع‌بندی لازم است تا لایه‌های تماماً متصل این شبکه‌ها به لایه‌های پیچشی تبدیل شوند. این کار باعث می‌شود که خروجی شبکه به جای این که امتیازات مربوط به کلاس‌ها باشد یک تصویر قطع‌بندی شده باشد.



شکل ۴- نمونه‌ای از شبکه عصبی پیچشی

بعد از تعریف روش‌های مختلف یادگیری عمیق لازم است روش‌های قطع‌بندی مبتنی بر یادگیری عمیق شرح داده شود. الگوریتم‌های قطع‌بندی مبتنی بر یادگیری عمیق را می‌توان به سه دسته کلی تقسیم‌بندی نمود [۱۶]:

۳-۱- قطع‌بندی مبتنی بر شبکه تماماً پیچشی

در این روش از شبکه‌های تماماً پیچشی استفاده می‌شود و این روش توانایی اخذ تصاویر با ابعاد متفاوت را به‌عنوان ورودی دارد. این شبکه‌ها به این دلیل که تنها دارای لایه‌های پیچشی و ادغام می‌باشند توانایی قطع‌بندی تصاویر ورودی با اندازه‌های دلخواه را دارند. این شبکه‌ها اگر به‌صورت مستقیم مورد استفاده قرار گیرند با گسترش چندین لایه متناوب پیچشی و ادغام، رزولوشن تصویر قطع‌بندی شده خروجی کم می‌شود و منجر به ایجاد مرزهای تار می‌شود و اگر به تنهایی و بدون روش‌های بهبود قطع‌بندی مورد استفاده قرار گیرد، نتایج قطع‌بندی آن ضعیف می‌باشد.

۳-۲- قطع‌بندی نظارت‌شده ضعیف

در بیشتر روش‌های قطع‌بندی معنایی تولید تصاویر ماسک شده به‌منظور استفاده در آموزش شبکه‌ها به‌صورت دستی انجام می‌گیرد که اگر تعداد تصاویر بالا باشد این کار بسیار سخت و طاقت‌فرسا می‌باشد؛ بنابراین برخی روش‌های

نظارت‌شده ضعیف اخیراً ارائه شده‌اند که برای انجام قطع‌بندی معنایی از جعبه‌های محاطی تفسیر شده استفاده می‌کنند و تولید تصاویر ماسک با دخالت اپراتوری کمتر انجام می‌شود. به‌عنوان مثال روش Boxsup جعبه محاطی تفسیر شده را به‌عنوان یک نظارت بر آموزش شبکه و بهبود متناوب ماسک‌های برآورد شده برای قطع‌بندی معنایی به کار می‌برد [۲۰]. آموزش این شبکه‌ها به دلیل پیچیده بودن الگوریتم‌های ارائه شده بسیار سخت می‌باشد. منظور از قطع‌بندی معنایی این است که هر قطعه در تصویر معنای خاصی دارد به عبارت دیگر هر قطعه نشان‌دهنده یک شی خاص می‌باشد بدین معنی که هر قطعه برجسی دارد که مشخص می‌شود مربوط به چه شی‌ای می‌باشد.

۳-۳- قطع‌بندی مبتنی بر منطقه

در این روش‌ها به جای اینکه الگوریتم بر روی تک تک پیکسل‌ها اجرا شود مناطقی که احتمال وجود شی در آن‌ها وجود دارد از تصویر استخراج می‌شوند و قطع‌بندی تنها بر روی مناطق حاوی شی صورت می‌گیرد و پیش‌بینی‌های منطقه مبنای پیش‌بینی‌های پیکسل تبدیل می‌شود و برجسب‌گذاری پیکسل بر اساس بالاترین امتیاز منطقه‌ای که شامل آن پیکسل می‌باشد انجام می‌شود [۲۱]. شبکه‌های عصبی پیچشی منطقه‌ای یک مثال برای روش‌های منطقه مبنای می‌باشد. به‌طور خاص شبکه‌های عصبی پیچشی منطقه‌ای ابتدا از یک جستجوی انتخابی برای استخراج تعداد زیادی از مناطق پیشنهادی استفاده می‌کند و سپس ویژگی‌های شبکه‌های عصبی پیچشی را برای هر یک از آن‌ها محاسبه می‌کند در نهایت هر منطقه را با استفاده از ماشین‌های بردار پشتیبان خطی^۱ خاص، طبقه‌بندی می‌کند. در مقایسه با ساختار شبکه‌های عصبی پیچشی سنتی که عمدتاً برای طبقه‌بندی استفاده می‌شوند شبکه‌های عصبی پیچشی منطقه‌ای می‌تواند برای کارهای پیچیده‌تر مثل شناسایی اشیاء یا قطع‌بندی تصویر استفاده شود. علاوه بر این، شبکه‌های عصبی پیچشی منطقه‌ای می‌تواند بر روی هر یک از ساختارهای شبکه‌های عصبی پیچشی مثل AlexNet، VGG، GoogLeNet و ResNet ایجاد شود [۲۲].

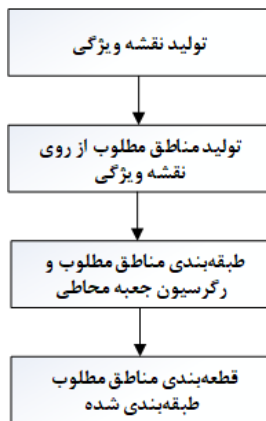
برای قطع‌بندی تصاویر، شبکه‌های عصبی پیچشی منطقه‌ای دو نوع ویژگی برای هر منطقه استخراج می‌کند:

^۱ Linear support vector machine

قطعه‌هایی با مرزهای کاملاً مشخص و باکیفیت بالا تبدیل کند. یکی از راه‌ها برای بکار بردن میدان‌های تصادفی شرطی به‌منظور بهبود نتایج برچسب‌گذاری، بکار بردن آن‌ها در یک مرحله جداگانه از آموزش شبکه‌های عصبی پیچشی به‌عنوان مرحله پس پردازش می‌باشد. مسلماً به دلیل اینکه میدان‌های تصادفی شرطی با شبکه عمیق ادغام نشده‌اند قدرت آن‌ها به‌طور کامل تأثیرگذار نیست. در این روش شبکه عمیق در طول آموزش از وجود میدان‌های تصادفی شرطی بی‌خبر می‌باشد. روش CRF-RNN ترکیبی از نقاط قوت هر دو روش شبکه عصبی پیچشی و میدان تصادفی شرطی مبتنی بر مدل‌های گرافیکی در یک چارچوب واحد می‌باشد. به‌طور خاص نتایج تقریبی زمینه‌های متوسط را برای میدان تصادفی شرطی متراکم با پتانسیل جفتی گوسین به‌عنوان یک شبکه عصبی بازگشتی ایجاد شده که قادر است خروجی‌های زمخت به‌دست‌آمده از شبکه‌های عصبی پیچشی را در یک حرکت روبه‌جلو درحالی‌که خطاهای مختلف را به عقب برمی‌گرداند در طول آموزش شبکه‌های عصبی پیچشی تصحیح نماید. به‌عبارت‌دیگر میدان تصادفی شرطی را در یک مرحله جداگانه اعمال نکرده بلکه آن را با شبکه‌های عصبی پیچشی ادغام نموده و در آموزش شبکه شرکت می‌دهد. مهم‌تر از همه با این فرمول‌بندی کل شبکه عمیق شامل شبکه عصبی پیچشی و شبکه عصبی بازگشتی برای نتیجه دادن میدان تصادفی شرطی می‌تواند با استفاده از الگوریتم‌های پس انتشار معمول به‌طور کامل آموزش ببیند.

۳-۲-۳-۲ روش Mask R-CNN

این روش شامل چهار مرحله می‌باشد؛ که مراحل آن در شکل (۵) مشاهده می‌شود و در ادامه هرکدام از مراحل توضیح داده می‌شود.



شکل ۵- روند کلی روش Mask R-CNN

ویژگی‌های کامل منطقه و ویژگی‌های پس‌زمینه که این کار باعث عملکرد بهتر آن می‌شود. باین حال این روش سه مشکل برای کارهای قطعه‌بندی دارد. (۱) این ویژگی‌ها برای قطعه‌بندی مناسب نیستند. (۲) ویژگی‌ها شامل اطلاعات مکانی دقیق برای تولید مرزهای دقیق نیستند. (۳) ایجاد پیشنهاد‌های مبتنی بر قطعه‌بندی زمان زیادی را صرف می‌کند و تا حد زیادی بر عملکرد نهایی تأثیر می‌گذارد. برای حل این مشکلات روش‌های بسیار خوبی از جمله SDS [۲۳]، Hypercolumns [۲۴]، CRF-RNN [۲۵] و Mask R-CNN [۲۶] ارائه شده است. با توجه به بررسی‌های صورت گرفته مشخص شد که روش Mask R-CNN و CRF-RNN به دلیل اینکه برای آموزش نیاز به سخت‌افزار با هزینه کمتری دارند و دارای دقت مناسبی برای قطعه‌بندی می‌باشند و همچنین از سرعت قابل قبولی برخوردار می‌باشند برای این پژوهش انتخاب شدند. در ادامه به شرح این دو روش پرداخته می‌شود.

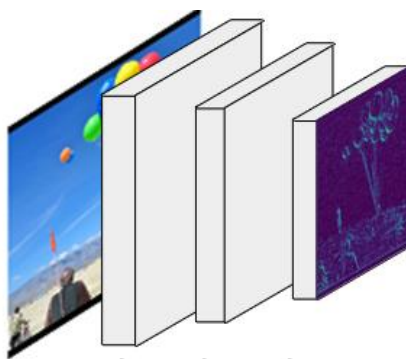
۳-۳-۱ روش CRF-RNN^۱

این روش فرم جدیدی از شبکه‌های عصبی ارائه می‌دهد که ترکیبی از شبکه‌های عصبی پیچشی و میدان تصادفی شرطی مبتنی بر مدل‌های گرافیکی احتمالی می‌باشد. شبکه‌های عصبی پیچشی دارای فیلترهای پیچشی با زمینه‌های پذیرش بزرگ می‌باشند و از این‌رو در کارهای برچسب‌گذاری پیکسلی، خروجی‌هایی تولید می‌کنند که گوشه‌های قطعه‌های تولید شده نرم نیستند. همچنین حضور لایه‌های ادغام بیشینه باعث می‌شود خروجی‌هایی با مرزهای نه‌چندان دقیق و لکه مانند تولید شود از طرف دیگر فقدان قیده‌های نرم‌کننده که منجر به ایجاد برچسب‌های دقیق بین پیکسل‌های مشابه منجر به نتایج ضعیف و ایجاد مناطق نادرست در تصاویر خروجی می‌شود. استفاده از مدل‌های گرافیکی احتمالی به‌عنوان روشی برای بهبود دقت روش‌های برچسب‌گذاری پیکسلی در کنار تکنیک‌های یادگیری عمیق می‌تواند مؤثر باشد. به‌طور خاص میدان‌های تصادفی مارکوف و میدان‌های تصادفی شرطی به موفقیت‌های فراوانی در این زمینه دست یافته‌اند و به یکی از موفق‌ترین مدل‌های گرافیکی در زمینه بیانی ماشین تبدیل شده‌اند. زمینه تصادفی شرطی قادر است که قطعه‌های زمخت و ضعیف را به

^۱ Conditional random fields as recurrent neural network

۳-۲-۱- تولید نقشه ویژگی

ستون اصلی این روش یک شبکه عصبی پیچشی معروف در دنیای یادگیری عمیق، یعنی ResNet50 یا ResNet101 می‌باشد که به‌عنوان یک استخراج‌کننده ویژگی عمل می‌کند. فرق این دو شبکه در تعداد لایه‌های آن‌ها می‌باشد. لایه‌های اولیه ویژگی‌های سطح پایین (لبه‌ها و گوشه‌ها) را تشخیص می‌دهند و لایه‌های بعد به‌طور پیوسته ویژگی‌های سطح بالاتری را تشخیص می‌دهند. با عبور تصویر از شبکه ResNet، از یک تصویر $3 \times 1024 \times 1024$ به یک نقشه ویژگی به شکل $32 \times 32 \times 3$ تبدیل می‌شود. نمای ساده‌ای از نقشه ویژگی تولید شده از تصویر را می‌توان در شکل (۶) مشاهده نمود، که تصویر بعد از عبور از لایه‌های مختلف به یک نقشه ویژگی تبدیل می‌شود. بعد سوم نقشه ویژگی در واقع تعداد فیلترهایی است که در تصویر مورد نظر در لایه‌های مختلف ضرب پیچشی شده‌اند. این نقشه ویژگی، ورودی برای مراحل بعد می‌باشد. اتفاقی که در این شبکه می‌افتد بسیار مشابه با مراحل توضیح داده‌شده در مورد شکل می‌باشد با این تفاوت که تعداد لایه‌های این شبکه بسیار بالاتر می‌باشد و همچنین لایه خروجی که یک لایه تماماً متصل برای کلاسه‌بندی تصویر می‌باشد در اینجا حذف شده است. درحالی‌که ستون اصلی توصیف‌شده در بالا به‌خوبی کار می‌کند می‌توان آن را با شبکه هرمی ویژگی (FPN) که می‌تواند اشیاء با مقیاس‌های مختلف را بهتر نمایش دهد بهبود بخشید. شبکه هرم ویژگی، هرم استخراج ویژگی استاندارد را با اضافه کردن یک هرم دوم که ویژگی‌های سطح بالا را از هرم اول می‌گیرد و آن‌ها را به لایه پایین‌تر انتقال می‌دهد بهبود می‌بخشد. با انجام این کار، به ویژگی‌های هر سطح اجازه می‌دهد تا به هر دو ویژگی پایین و بالاتر دسترسی داشته باشند [۲۷].

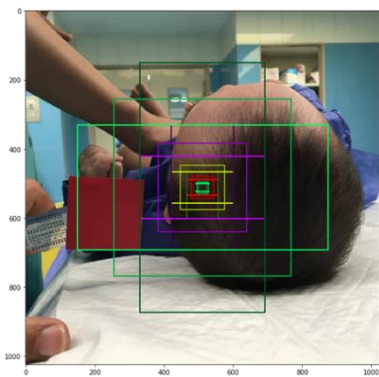


شکل ۶- نمای ساده از نقشه ویژگی تولید شده از تصویر

۳-۲-۲- تولید مناطق مطلوب از روی نقشه ویژگی

در این مرحله باید مناطق پیشنهادی (مناطق که احتمالاً شیئی در آن‌ها وجود دارد، منظور از شیئی، اجسام، عوارض، حیوانات و غیره می‌باشد، به‌طور کلی هر چیزی که ممکن است در تصویر وجود داشته باشد) از نقشه ویژگی استخراج شوند برای این کار از شبکه پیشنهاد منطقه (RPN) استفاده می‌شود. شبکه پیشنهاد منطقه یک شبکه عصبی سبک می‌باشد که تصویر را با یک پنجره لغزان جستجو می‌کند و مناطقی را که حاوی اشیاء می‌باشند را پیدا می‌کند. مناطقی را که شبکه پیشنهاد منطقه بر روی آن‌ها جستجو می‌کند لنگر نامیده می‌شود که بر روی سطح تصویر پخش می‌شود. نمونه‌ای از لنگرها را می‌توان در شکل (۷) مشاهده نمود.

در عمل تعداد بالایی لنگر با ابعاد و اندازه‌های متفاوت سطح تصویر را می‌پوشاند. تعداد بالای لنگر با ابعاد و اندازه‌های متفاوت به‌منظور قابل‌تعمیم بودن شبکه برای شناسایی و قطع‌بندی اشیاء با ابعاد و اندازه‌های مختلف در تصویر می‌باشد. حال این سؤال مطرح می‌شود که شبکه پیشنهاد منطقه با چه سرعتی می‌تواند تمام این لنگرها را جستجو کند؟ در حقیقت خیلی سریع این کار انجام می‌شود. پنجره لغزان که توسط ماهیت پیچشی شبکه پیشنهاد منطقه اداره می‌شود به آن اجازه می‌دهد که جستجوی تمام مناطق را به‌صورت موازی (بر روی واحد پردازش گرافیکی) انجام دهد. علاوه بر این شبکه پیشنهاد منطقه به‌صورت مستقیم بر روی تصویر جستجو نمی‌کند بلکه جستجوی خود را بر روی نقشه ویژگی تهیه‌شده از قبل انجام می‌دهد. این کار باعث می‌شود تا شبکه پیشنهاد منطقه دوباره از ویژگی‌های استخراج‌شده به‌طور مؤثری استفاده کند و از محاسبات تکراری جلوگیری کند. با این بهینه‌سازی با توجه به مقاله Faster- CNN [۲۸] که آن را معرفی می‌کند شبکه پیشنهاد منطقه در حدود ۱۰ میلی‌ثانیه اجرا می‌شود. در Mask R-CNN با توجه به اینکه از تصاویر بزرگ‌تر و لنگرهای بیشتری استفاده می‌کند زمان اجرای آن کمی طولانی‌تر می‌باشد.



شکل ۷- چند نمونه از لنگرهای پخش شده روی تصویر

یک مشکل کوچکی که وجود دارد این است که طبقه‌بندی کننده‌ها اندازه متغیر ورودی را به خوبی مدیریت نمی‌کنند. این‌ها معمولاً به یک ورودی با اندازه ثابت نیاز دارند؛ اما با توجه به مرحله تصحیح جعبه محاطی در شبکه پیشنهاد منطقه، کادرهای مناطق مطلوب می‌توانند اندازه‌های مختلفی داشته باشند، این‌جایی است که شبکه ادغام مناطق مطلوب وارد کار می‌شود؛ بنابراین وظیفه شبکه ادغام مناطق مطلوب، برش یک قسمت از نقشه ویژگی و تغییر اندازه آن به یک اندازه ثابت می‌باشد. این کار در اصل مشابه با برش یک قسمت از تصویر و تغییر اندازه آن می‌باشد اما در جزئیات پیاده‌سازی آن تفاوت‌هایی وجود دارد. در مقاله Mask R_CNN یک روش به نام تنظیم مناطق مطلوب پیشنهاد شده است که در نقاط مختلف نقشه ویژگی، نمونه می‌گیرد و یک درون‌یابی دوخطی روی آن انجام می‌دهد. به جای این روش پیشنهادی، در این پژوهش به راحتی از تابع برش و تغییر اندازه موجود در تنسورفلو استفاده شده است. تا اینجا مرحله تشخیص شیئی انجام می‌شود.

۴-۳-۲-۴- قطعه‌بندی مناطق مطلوب

بخش ماسک در حقیقت یک شبکه پیچشی است که مناطق مثبت انتخاب شده توسط طبقه‌بندی کننده مناطق مطلوب را می‌گیرد و ماسک‌هایی را برای آن‌ها تولید می‌کند. ماسک‌های تولید شده دارای رزولوشن پایین ۲۸×۲۸ پیکسل می‌باشند؛ اما آن‌ها ماسک‌های نرمی هستند که توسط اعداد اعشاری نمایش داده می‌شوند و در مقایسه با ماسک‌های باینری جزئیات بیشتری را حفظ می‌کنند. این ماسک‌های کوچک باعث می‌شوند که بخش ماسک بخش سبکی باشد. در حقیقت در طول آموزش شبکه، ماسک‌های داده‌های آموزشی به اندازه ۲۸×۲۸ کاهش مقیاس داده می‌شوند و در زمان ارائه نتایج، ماسک‌های پیش‌بینی شده به اندازه جعبه محاطی مناطق مطلوب افزایش مقیاس داده می‌شوند و ماسک نهایی را برای هر شیئی ارائه می‌دهد.

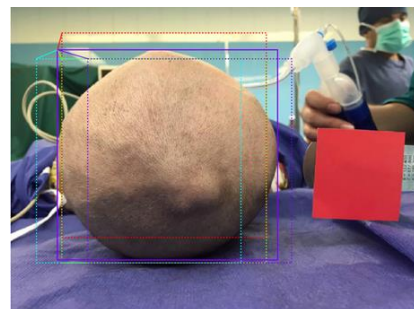
۴- روش پیشنهادی

همان‌طور که اشاره شد هدف از این تحقیق ارائه روشی غیر تماسی با استفاده از یک تصویر به منظور اندازه-

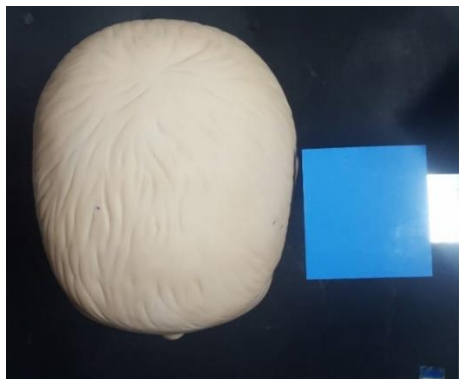
شبکه پیشنهاد منطقه دو خروجی برای هر لنگر تولید می‌کند. ۱- کلاس لنگر: یکی از دو کلاس پیش‌زمینه و دیگری پس‌زمینه می‌باشد که کلاس پیش‌زمینه به این اشاره دارد که احتمالاً یک شیئی در کادر وجود دارد. ۲- تصحیح جعبه محاطی: یک لنگر پیش‌زمینه که لنگر مثبت هم نامیده می‌شود ممکن است به‌طور کامل بر روی شیئی متمرکز نشود بنابراین شبکه پیشنهاد منطقه یک دلتا (تغییرات در x ، y ، عرض و ارتفاع) برای تصحیح جعبه لنگر به منظور انطباق هر چه بهتر به شیئی برآورد می‌کند. با استفاده از پیش‌بینی‌های شبکه پیشنهاد منطقه، لنگرهایی که احتمال وجود شیئی در آن‌ها بیشتر است انتخاب می‌شوند و موقعیت و اندازه آن‌ها تصحیح می‌گردد. اگر چندین لنگر باهم همپوشانی زیادی داشته باشند یکی از آن‌ها با بالاترین امتیاز پیش‌زمینه نگه‌داشته می‌شود و از بقیه صرف‌نظر می‌گردد. پس از آن پیشنهاد نهایی یعنی پیشنهاد منطقه مطلوب وجود خواهد داشت که به مرحله بعد منتقل می‌شود. به ازای هر تصویر ورودی، مناطق مطلوب استخراج می‌شود.

۳-۳-۲-۳- طبقه‌بندی مناطق مطلوب و رگرسیون جعبه محاطی

این مرحله بر روی مناطق مطلوب پیشنهادی توسط شبکه پیشنهاد منطقه اجرا می‌شود. این مرحله نیز برای هر منطقه مطلوب دو خروجی تولید می‌کند. ۱- کلاس: کلاس شیئی در هر منطقه مطلوب، برخلاف شبکه پیشنهاد منطقه که تنها دو کلاس داشت این شبکه عمیق‌تر می‌باشد و قابلیت طبقه‌بندی مناطق به کلاس‌های مخصوص (شخص، درخت، ماشین و ...) را دارد. همچنین می‌تواند یک کلاس پس‌زمینه ایجاد کند که منجر به از بین رفتن منطقه مطلوب شود. ۲- تصحیح جعبه محاطی: که بسیار شبیه به نحوه کار در شبکه پیشنهاد منطقه هست و هدف آن تصحیح بیشتر موقعیت و اندازه جعبه محاطی برای هر چه بهتر در برگرفتن شیئی می‌باشد. نمونه‌ای از منطقه مطلوب را می‌توان در شکل (۸) مشاهده نمود.



شکل ۸- نمونه منطقه مطلوب و تصحیح جعبه محاطی



شکل ۱۰- نمونه برچسب مقیاس



شکل ۱۱- نحوه تصویربرداری از سر نوزاد

بعد از اخذ تصویر باید سر نوزاد در تصویر شناسایی شود و به صورت پیکسل به پیکسل قطعه‌بندی شود. برای شناسایی و قطعه‌بندی تصویر نیاز به روشی می‌باشد که بتواند تحت هر شرایطی مثل شرایط نوری مختلف، کیفیت‌های تصویر متفاوت، وجود هر شیئی اضافی دیگر در تصویر و غیره به بهترین نحو عمل کند.

۴-۱- قطعه‌بندی پیکسل‌های سر در تصویر

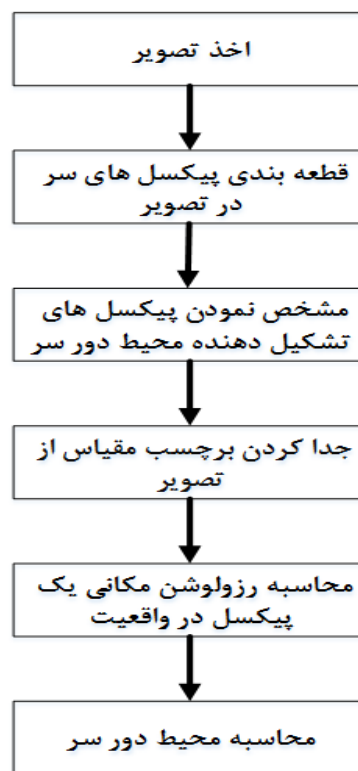
به‌منظور قطعه‌بندی پیکسل‌های سر در تصویر از دو روشی که در بخش قبل توضیح داده شد استفاده می‌شود.

۴-۲- مشخص نمودن پیکسل‌های تشکیل‌دهنده

دور سر

پس از شناسایی و قطعه‌بندی سر نوزاد در تصویر لازم است تصویر به‌دست‌آمده، تبدیل به یک تصویر باینری شود که پیکسل‌های این تصویر یا جزء سر نوزاد می‌باشند یا جزء پس‌زمینه، بعد از انجام این کار اگر پیکسل‌های اضافی در تصویر موجود باشد باید حذف شوند.

گیری محیط دور سر نوزاد می‌باشد. روند کلی روش پیشنهادی در این پژوهش در شکل (۹) نشان داده شده است که در ادامه هر یک از مراحل این روش پیشنهادی به‌تفصیل شرح داده می‌شود. اولین گام در این پروژه گرفتن یک تصویر از بالای سر نوزاد می‌باشد به‌گونه‌ای که محور نوری دوربین تا حد امکان عمود به شئی (سر نوزاد) باشد همچنین به‌منظور حل مسئله مقیاس یک برچسب مقیاسی نیز باید در تصویر موجود باشد، این برچسب مقیاس از جنس مقوا می‌باشد و با استفاده از کولیس به‌صورت دقیق با ابعاد ۵ سانتی‌متر در ۵ سانتی‌متر برش داده شده است. نمونه برچسب مقیاس را می‌توان در شکل (۱۰) و نحوه تصویربرداری را در شکل (۱۱) مشاهده نمود. نکته دیگری که وجود دارد این است که با توجه به نزدیک بودن ایستگاه تصویر برداری به عارضه و همچنین عدم مسطح بودن سطح عارضه و وجود تغییرات عمق نسبت به دوربین، طبیعتاً مقیاس تصویر برداری در بخش‌های مختلف عارضه هدف متفاوت است بنابراین باید حتمی شاخص مقیاس در تراز ارتفاعی منحنی دور سر قرار گیرد تا فاصله دوربین تا منحنی مورد نظر و شاخص مقیاس برابر شود و اندازه پیکسل‌های تشکیل دهنده منحنی دور سر و شاخص مقیاس برابر شود.



شکل ۹- روند کلی روش پیشنهادی

تصویر حاصل کاملاً یکدست نمی‌باشد یا به اصطلاح سوراخ‌هایی در تصویر و لبه‌های تصویر وجود دارد و همچنین به خاطر استفاده از تکنیک گشودن مرزهای تصویر تخریب می‌شوند و این عوامل بر روی دقت کار تأثیر منفی خواهد گذاشت بنابراین برای جبران آن لازم است که این بار بر روی تصویر گسترش و بلافاصله بعد از آن سایش صورت گیرد در واقع تکنیک بستن بر روی تصویر اعمال شود تا سوراخ‌های موجود، پر گردد و همچنین مرزهای تخریب شده به حالت اول بازگردند. در عملیات بستن ابتدا سوراخ‌های موجود در تصویر پر می‌گردد و سپس تصویر به اندازه اولیه برمی‌گردد.

حال تصویر آماده است تا لبه‌های آن استخراج شود بنابراین لازم است بر روی تصویر لبه یابی صورت گیرد. با ارزیابی ۴ فیلتر لبه یاب با نام‌های Canny, Prewitt, Zerocross و Laplacian of Gaussian بر روی ۲۰ تصویر نمونه مشخص شد که فیلتر لبه یاب کنی [۲۹] مناسب ترین لبه‌ها را در تصویر استخراج می‌کند.

۴-۳- جدا کردن برچسب مقیاس از تصویر

به دلیل استفاده از یک برچسب مقیاس به رنگ خاص، به راحتی می‌توان با گذاشتن یک حد آستانه، برچسب مقیاس را از تصویر جدا نمود و سپس اگر پیکسل‌های اضافی در تصویر موجود بود بر روی تصویر عمل باز کردن اعمال می‌شود و بعد از آن برای داشتن تصویری یکدست، بر روی تصویر مورد نظر عملیات بستن اجرا می‌شود. حال با اعمال یک فیلتر لبه یاب لبه‌های این برچسب مقیاسی استخراج می‌شود.

۴-۴- محاسبه رزولوشن مکانی یک پیکسل در واقعیت

پس از مشخص شدن پیکسل‌های تشکیل دهنده لبه برچسب مقیاس، می‌توان تعداد این پیکسل‌ها را به دست آورد. با محاسبه تعداد این پیکسل‌ها در واقع تعداد پیکسل‌های تشکیل دهنده محیط برچسب مقیاس مشخص می‌شود. از طرف دیگر محیط برچسب مقیاس در واقعیت مشخص می‌باشد؛ بنابراین با استفاده از این دو پارامتر می‌توان به رزولوشن مکانی یک پیکسل در واقعیت دست یافت، بنابراین داریم:

به منظور حذف پیکسل‌های اضافی موجود در تصویر نیاز است که از تکنیک گشودن^۱ استفاده نمود. تکنیک گشودن خود از دو تکنیک گشایش^۲ و سایش^۳ تشکیل می‌شود. منظور از عملیات گشایش عملیاتی است که باعث افزایش ابعاد اجزای داخل تصویر به اندازه یک یا چند پیکسل می‌گردد. در اثر این عمل ممکن است نقاطی که از یک تصویر باینری در اثر عواملی چون تأثیر نویز یا اعمال حد آستانه نامطلوب جا افتاده است، تصحیح گردند. مثلاً ممکن است دو جزء از تصویر به یکدیگر متصل گردند. الگوریتم اعمال فیلتر گشایش بدین صورت است که تمامی نقاط سیاه تصویر بررسی شده در صورتی که حداقل یکی از همسایگان انتخابی نقطه مورد بررسی سفید باشند، نقطه مزبور نیز سفید خواهد شد در غیر این صورت سیاه باقی خواهد ماند.

عملیات سایش دقیقاً عکس عملیات گشایش است. در این عملیات معمولاً نقاط ناخواسته تصویر باینری حذف می‌شوند و سایر اجزا تصویر نیز به اندازه یک یا چند پیکسل نازک تر خواهند شد. عملاً تمامی نقاط سفید تصویر بررسی شده در صورتی که حداقل یکی از همسایگان انتخابی آن سیاه باشد، آن نقطه نیز سیاه خواهد شد.

از ترکیب‌های مختلف دو عملیات افزایش و فرسایش می‌توان عملیات دیگری ایجاد کرد. مهم‌ترین این عملیات، عملیات گشودن و بستن است. در عملیات گشودن اجزایی از تصویر باینری که از یک اندازه تعیین شده کوچکتر باشند حذف می‌شوند بدون آنکه ابعاد سایر اجزا تغییر کند. در عملیات بستن نیز نواحی جا افتاده تصویر باینری بدون تغییر در ابعاد سایر اجزا ترمیم می‌گردند. عملاً در صورتی که ابتدا عملیات فرسایش و سپس افزایش بر یک تصویر باینری اعمال شود، نتیجه، عملیات گشودن خواهد بود اما اگر ابتدا افزایش و سپس فرسایش اعمال گردد، عملیات بستن حاصل خواهد شد.

در عملیات گشودن اجزایی از تصویر باینری که از یک اندازه تعیین شده کوچکتر باشد حذف می‌شود بدون آنکه ابعاد سایر اجزا تغییر کند؛ بنابراین با اعمال عملیات گشودن بر روی تصویر پیکسل‌های اضافی در تصویر حذف می‌گردد.

^۱ Opening

^۲ Dilation

^۳ Erosion

$$P.S = \frac{N.P.S}{E.S} \quad (1)$$

که P.S اندازه یک پیکسل تصویر در واقعیت برحسب میلی‌متر می‌باشد، N.P.S تعداد پیکسل‌های تشکیل‌دهنده محیط برچسب مقیاس می‌باشد و E.S محیط برچسب مقیاس برحسب میلی‌متر می‌باشد.

۴-۵- محاسبه محیط دور سر

پس از آن که لبه سر نوزاد در تصویر مشخص شد و تعداد پیکسل‌های تشکیل‌دهنده آن نیز تعیین شد و از طرف دیگر رزولوشن مکانی یک پیکسل در واقعیت مشخص شد بنابراین می‌توان با استفاده از فرمول زیر به محیط سر نوزاد دست یافت.

$$E.H = P.S \times N.P.H \quad (2)$$

که E.H محیط سر نوزاد می‌باشد، P.S رزولوشن مکانی یک پیکسل در واقعیت می‌باشد و N.P.H تعداد پیکسل‌های موجود در دور سر می‌باشد.

۵- پیاده‌سازی

بعد از تعیین تئوری مسئله که در بخش قبل بیان شد لازم است تا تجهیزات و داده‌های مورد استفاده و روش پیاده‌سازی الگوریتم پیشنهادی نیز شرح داده شود.

۵-۱- تجهیزات و داده‌های مورد استفاده

در بخش شناسایی و قطعه‌بندی سر نوزاد در تصویر که با استفاده از روش‌های یادگیری عمیق انجام می‌شود نیاز به سخت‌افزار قوی برای آموزش می‌باشد؛ بنابراین از یک سیستم با پردازنده مرکزی Intel Core i7-7700 و کارت گرافیک NVIDIA GeForce GTX 1050 Ti استفاده گردید، در صورت استفاده از سیستم ضعیف ممکن است روند آموزش به کلی مختل شود و یا این روند بسیار طولانی شود. همچنین برای آموزش شبکه یادگیری عمیق به منظور شناسایی و قطعه‌بندی سر نوزاد در تصویر از ۲۵۰ تصویر استفاده شد. برای پیاده‌سازی روش‌های Mask R_CNN و CRF-RNN از کدهای موجود در سایت github [۳۰] استفاده شد ولی تغییرات اساسی در کدها با توجه به مجموعه داده ایجاد شده صورت گرفت.

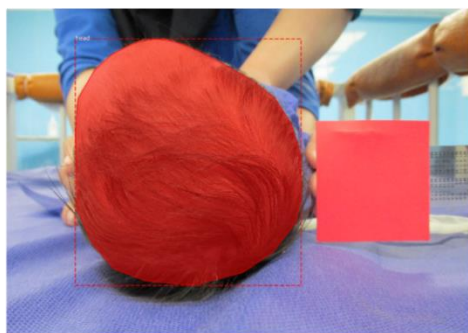
۵-۲- قطعه‌بندی پیکسل‌های سر در تصویر

به منظور قطعه‌بندی پیکسل‌های سر در تصویر نیاز بود تا الگوریتم Mask R_CNN به طور اختصاصی برای این کار آموزش داده شود. اولین لازمه آموزش یک الگوریتم، داشتن یک مجموعه داده می‌باشد تا بتوان الگوریتم مورد نظر را بر روی این مجموعه داده، آموزش داد. به طور معمول برای آموزش این شبکه‌ها از مجموعه داده‌های موجود در اینترنت که هر کدام از آن‌ها دارای تعداد زیادی تصاویر مختلف می‌باشند استفاده می‌شود ولی در این مورد خاص که نیاز است شبکه برای شناسایی سر نوزاد در تصویر، آموزش داده شود باید یک مجموعه داده اختصاصی به این منظور تولید گردد. برای ایجاد این مجموعه داده به مرکز طبی کودکان مراجعه شد و تعداد ۱۵۰ تصویر از سر نوزاد تهیه گردید و تعداد ۱۰۰ تصویر نیز در اینترنت و همچنین با استفاده از کتابخانه imgaug [۳۱] موجود در پایتون تهیه گردید. با استفاده از این تصاویر یک مجموعه داده تهیه شد تا بتوان به عنوان داده‌های آموزش^۱ و اعتبار سنجی^۲ در فرایند آموزش الگوریتم مورد نظر، مورد استفاده قرار گیرد. برای ساخت این مجموعه داده از نرم‌افزار VGG Image Annotator (VIA) استفاده شد. نحوه تولید ماسک برای داده‌های آموزشی به صورت دستی و توسط کاربر با مشخص کردن محدوده‌های ماسک با استفاده از ابزارهای موجود در این نرم‌افزار صورت گرفت. حال این سؤال مطرح می‌شود که آیا با این تعداد کم تصاویر شبکه به خوبی آموزش داده می‌شود؟ در پاسخ به این سؤال باید گفت که به جای اینکه شبکه از ابتدا آموزش داده شود می‌توان از وزن‌های به دست آمده از آموزش شبکه بر روی مجموعه داده COCO استفاده نمود، این مجموعه داده یک مجموعه داده استاندارد می‌باشد که دارای ۱۲۰ هزار تصویر می‌باشد و از این مجموعه داده برای آموزش الگوریتم‌ها و همچنین برای تست الگوریتم‌های آموزش داده شده استفاده می‌شود. درست است که تصاویر مجموعه داده COCO شامل تصاویر سر نوزاد نمی‌باشد ولی شبکه قادر است ویژگی‌های فراوانی که در خیلی از تصاویر مشترک می‌باشند را یاد بگیرد؛ بنابراین می‌توان این شبکه از پیش آموزش داده شده را یکبار دیگر با تعداد تصاویر بسیار کمتر آموزش داد. آموزش

^۱ Train
^۲ validation



شکل ۱۰- تصویر سر از نوزاد



شکل ۱۱- تصویر حاصل از شناسایی و قطعه‌بندی

۵-۳- مشخص نمودن پیکسل‌های تشکیل دهنده محیط دور سر

پس از شناسایی و قطعه‌بندی سر در تصویر با استفاده از آستانه گذاری مناسب، تصویر موردنظر تبدیل به یک تصویر باینری گردید به نحوی که پیکسل‌های موجود در تصویر یا جزء سر، یا جزء پس‌زمینه می‌باشند. این آستانه‌گذاری دلیل شناسایی درجات خاکستری مشابهی که ممکن است در تصویر وجود داشته باشند و متعلق به پیکسل‌های سر نوزاد نباشند صورت می‌گیرد. می‌توان تصویر حاصل را در شکل (۱۴) مشاهده نمود.

اما همان‌طور که در شکل مشاهده می‌شود تعدادی از پیکسل‌های این تصویر به‌اشتباه در کلاس سر قرار گرفته‌اند بنابراین لازم است تا این پیکسل‌های اضافی حذف شوند بنابراین بر روی تصویر موردنظر عملیات گشودن صورت گرفت که می‌توان نتیجه آن را در شکل (۱۵) مشاهده نمود. عنصر ساختاری عملگر گشودن مورد استفاده یک ساختار دایره‌ای با شعاع ۳ پیکسل می‌باشد.

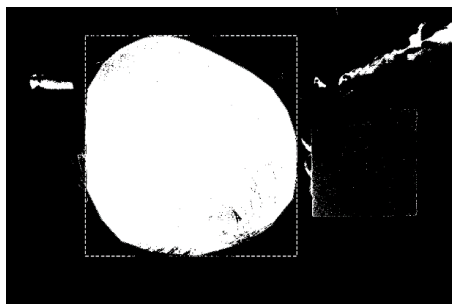
اما همان‌طور که در شکل (۱۵) مشاهده می‌شود تصویر دارای سوراخ‌هایی می‌باشد بنابراین برای داشتن یک تصویر یکدست نیاز است تا بر روی تصویر عملیات بستن صورت گیرد. تصویر حاصل از اعمال عملیات بستن در شکل (۱۶) مشاهده می‌شود.

این شبکه در محیط لینوکس تحت کتابخانه معروف یادگیری ماشین یعنی تنسورفلو صورت گرفت. به دلیل محاسبات ریاضی و به‌طور خاص ماتریسی در مقیاس زیاد، آموزش یک شبکه عصبی عمیق با پردازنده‌های معمولی تقریباً غیرممکن خواهد بود. به همین دلیل آموزش این شبکه عصبی عمیق بر روی پردازنده گرافیکی صورت گرفت. روش مورد استفاده در این آموزش روش پس انتشار خطا می‌باشد. در این آموزش تعداد دسته^۱ تصاویر ورودی به الگوریتم^۲ بود و مقدار مومنتوم یادگیری^۳ بر روی ۰/۹ و همچنین نرخ یادگیری^۴ بر روی ۰/۰۰۱ و گام‌ها در هر دوره^۵ بر روی ۱۰۰ قرار گرفت و همچنین به‌منظور جلوگیری از دقیقاً مدل کردن داده‌های آموزشی و توانایی تعمیم به مثال‌های دیگر پارامتر مجازات وزن‌ها^۵ نیز در نظر گرفته شد و مقدار آن بر روی ۰/۰۰۰۱ قرار گرفت و در نهایت مقیاس‌های لنگرهای شبکه پیشنهاد منطقه به ترتیب ۳۲، ۶۴، ۱۲۸، ۲۵۶ و ۵۱۲ با توجه به شیء موردنظر در تصویر در نظر گرفته شد. می‌توان یک نمونه از تصویر ورودی به الگوریتم آموزش داده‌شده و خروجی آن را در شکل (۱۲، ۱۳) مشاهده نمود.

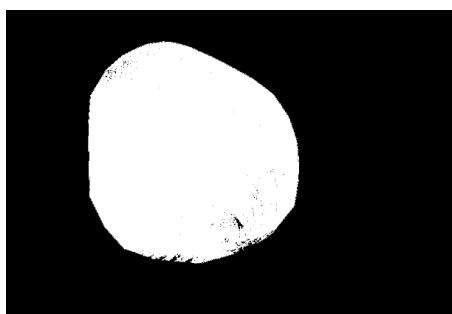
در روش دوم نیز به‌منظور شناسایی و قطعه‌بندی پیکسل‌های سر در تصویر نیاز است تا الگوریتم CRF-RNN آموزش داده شود. به‌منظور ایجاد شرایط برابر، از همان ۲۵۰ تصویر استفاده‌شده در روش Mask_R_CNN، برای ایجاد مجموعه داده استفاده شد که در اینجا از نرم‌افزار RectLabel برای برچسب‌گذاری و ایجاد مجموعه داده استفاده شد. آموزش این شبکه نیز در محیط لینوکس و تحت کتابخانه تنسورفلو صورت گرفت. مقداردهی بخش اول شبکه با استفاده از وزن‌های موجود شبکه FCN-8 انجام شد. در نهایت شبکه یک‌بار دیگر بر روی مجموعه داده تهیه‌شده آموزش دید که در فرایند آموزش از الگوریتم پس انتشار استفاده شد. نرخ یادگیری ۰/۰۰۱ و مقدار مومنتوم ۰/۹۹ در نظر گرفته شد. برای جلوگیری از مواجه شدن با محدودیت حافظه نیز تعداد دسته تصاویر ورودی به الگوریتم برابر ۱ در نظر گرفته شد.

۱ batch size
 ۲ learning momentum
 ۳ learning rate
 ۴ step per epoch
 ۵ weight decay

حال که لبه‌های تصویر ترمیم شد و تصویر یک‌دست گردید باید لبه‌های تصویر استخراج گردد. به همین دلیل فیلتر لبه یاب کنی بر روی تصویر اعمال گردید و لبه‌های تصویر مورد نظر استخراج شد که می‌توان خروجی آن را در شکل (۱۷) مشاهده نمود. این تصویر لبه نشانگر پیکسل-های تشکیل‌دهنده محیط دور سر نوزاد می‌باشد.



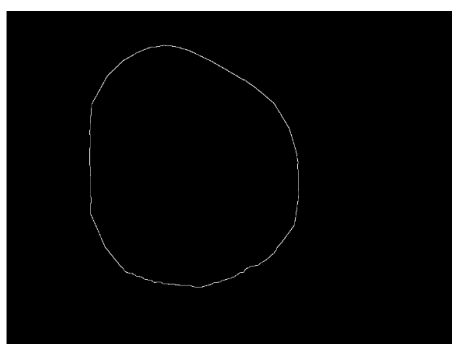
شکل ۱۲- تصویر باینری سر نوزاد



شکل ۱۳- تصویر بعد از حذف پیکسل‌های اضافی



شکل ۱۴- تصویر بعد از ترمیم لبه‌ها



شکل ۱۵- تصویر محیط تشکیل‌دهنده سر نوزاد

۴-۵- جدا کردن برجسب مقیاس از تصویر

برای محاسبه محیط دور سر نوزاد نیاز است که مقیاس تصویر مشخص شود بنابراین با گذاشتن یک حد آستانه، برجسب مقیاس از تصویر جدا می‌گردد و برای حذف پیکسل‌هایی که به‌اشتباه در کلاس برجسب مقیاس قرار گرفته‌اند عملیات گشودن انجام گرفت و پس از آن برای ترمیم لبه‌ها و پر کردن سوراخ‌ها عملیات بستن صورت گرفت و در نهایت تصویر حاصل لبه یابی شد.

۵-۵- محاسبه رزولوشن مکانی یک پیکسل در واقعیت

با توجه به تعداد پیکسل‌های تشکیل‌دهنده محیط برجسب مقیاس و معلوم بودن محیط آن در واقعیت و همچنین فرمول (۹) در بخش قبل محاسبه رزولوشن مکانی یک پیکسل به‌صورت خودکار صورت می‌گیرد.

۶-۵- محاسبه محیط دور سر

در نهایت با مشخص بودن رزولوشن مکانی هر پیکسل و همچنین شمارش تعداد پیکسل‌های تشکیل‌دهنده محیط دور سر نوزاد، محیط سر نوزاد با استفاده از فرمول (۱۰) به دست می‌آید. با توجه به اینکه اندازه سر نوزادان در یک محدوده خاص می‌باشد تمامی مراحل فوق به‌صورت خودکار انجام می‌شود و کاربر نقشی در مراحل اجرایی ندارد.

۶- ارزیابی

به دلیل اینکه مرحله شناسایی و قطعه‌بندی تصویر از اهمیت بالایی برخوردار می‌باشد و تأثیر بسزایی در نتایج پایانی دارد بخش ارزیابی این تحقیق به دو بخش تقسیم می‌شود ابتدا ارزیابی الگوریتم مورد استفاده در شناسایی و قطعه‌بندی مورد ارزیابی قرار می‌گیرد و سپس نتایج محاسبه محیط دور سر نیز مورد ارزیابی قرار خواهد گرفت.

۶-۱- ارزیابی الگوریتم شناسایی و قطعه‌بندی

در این مرحله نتایج روش Mask R-CNN و روش CRF-RNN مورد ارزیابی قرار خواهد گرفت تا مشخص شود کدام روش برای انجام این پژوهش مناسب‌تر می‌باشد.

علاوه بر معیارهای ارزیابی بالا می‌توان به صورت بصری نیز نتایج این دو روش را مورد بررسی قرارداد. همان‌طور که در شکل‌های (۱۸، ۱۹، ۲۰) مشاهده می‌شود نتایج قطعه‌بندی روش Mask R_CNN در تصاویر مختلف بهتر از روش CRF-RNN می‌باشد.

به منظور ارزیابی بیشتر این دو الگوریتم، توانایی آن‌ها در تصاویر چالشی نیز مورد بررسی قرار گرفت به منظور ارزیابی بیشتر این الگوریتم‌ها از چالش‌هایی که این الگوریتم‌ها در واقعیت با آن مواجه می‌باشند استفاده شد. برخی از این چالش‌ها عبارت انداز اندازه‌گیری سرهایی که دارای تغییر شکل شدید می‌باشند، اندازه‌گیری سرهای نوزادان بلافاصله بعد از عمل که در آن شرایط سر باندپیچی شده می‌باشد و اندازه‌گیری سرهایی که دارای رنگ تیره می‌باشند و همچنین مدتی از زمان عمل جراحی آن‌ها گذشته بنابراین بر روی سر جای زخم وجود دارد. این دو الگوریتم بر روی این تصاویر اجرا شد و نتایج آن در شکل‌های (۲۱، ۲۲، ۲۳) مشاهده می‌شود. همان‌طور که در شکل‌های (۲۱، ۲۲، ۲۳) مشخص می‌باشد روش Mask R_CNN به خوبی این چالش‌ها را پشت سر گذاشته ولی روش CRF-RNN نتوانست قطعه‌بندی را به خوبی انجام دهد. می‌توان دلیل موفقیت بیشتر روش Mask R_CNN را استفاده از شبکه پیشنهادی منطقه (PRN) دانست که با استفاده از تعداد بسیار بالای لنگر با ابعاد و اندازه‌های متفاوت در سطح تصویر توانست به راحتی با سرعت و دقت بالا کل تصویر را جستجو کرده و مناطقی که احتمال وجود شئی در آن‌ها وجود دارد را ارائه دهد. از طرف دیگر در بخش ماسک به دلیل استفاده از ماسک‌های نرمی که توسط اعداد اعشاری نمایش داده می‌شود و نسبت به ماسک‌های باینری جزئیات بیشتری را حفظ می‌کنند منجر به موفقیت بیشتر روش Mask R_CNN شده است.

روشی که برای ارزیابی نتایج این دو الگوریتم در نظر گرفته شده است استفاده از درایه‌های ماتریس درهم‌ریختگی یا ماتریس ابهام می‌باشد (جدول شماره (۱))؛ که با استفاده از درایه‌های این ماتریس می‌توان مقدار شاخص‌های ارزیابی یعنی نرخ تشخیص درست، نرخ تشخیص اشتباه و دقت کلی را به دست آورد.

در جدول (۱) TP تعداد پیکسل‌هایی است که هم در تصویر مرجع و هم در تصویر خروجی به درستی به عنوان سر نوزاد قطعه‌بندی شده است. FP تعداد پیکسل‌هایی است که در تصویر مرجع به عنوان سر قطعه‌بندی نشده است ولی در روش موردنظر به اشتباه به عنوان سر قطعه‌بندی شده است. FN تعداد پیکسل‌هایی است که در تصویر مرجع قطعه‌بندی شده است ولی در روش موردنظر به اشتباه قطعه‌بندی نشده است. TN تعداد پیکسل‌هایی است که در تصویر مرجع و در روش موردنظر به عنوان سر قطعه‌بندی نشده است. حال با استفاده از این موارد می‌توان سه معیار ارزیابی را به صورت زیر به دست آورد.

$$\text{True Positive Rate} = \frac{TP}{TP+FP} \quad (۳)$$

$$\text{False Negative Rate} = \frac{FN}{TP+FP} \quad (۴)$$

$$\text{Overall Accuracy} = \frac{TP+TN}{TP+FN+TN+FP} \quad (۵)$$

برای ارزیابی و مقایسه، ۱۸ تصویر مشترک به هر دو الگوریتم داده شد و معیارهای ذکر شده برای تمام این ۱۸ تصویر حاصل از دو الگوریتم به دست آمد که می‌توان میانگین نتایج ارزیابی حاصل از دو روش بر روی این ۱۸ تصویر را در جدول (۲) مشاهده نمود.

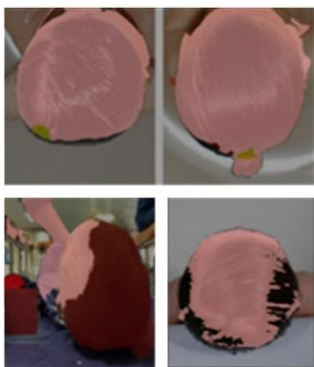
همان‌طور که در جدول (۲) مشاهده می‌شود روش Mask R_CNN نتایج به مراتب بهتری از روش CRF-RNN به دست آورده است.

۱- ماتریس ابهام

تصویر مرجع			مجموع	نتایج حاصل از روش موردنظر
مجموع	قطعه‌بندی نشده	قطعه‌بندی شده		
TP + FP	FP	TP	قطعه‌بندی شده	
TN + FN	TN	FN	قطعه‌بندی نشده	
FP + TN + TP + FN	FP + TN	TP + FN	مجموع	

۲- میانگین نتایج ارزیابی دو روش قطعه‌بندی

شاخص ارزیابی روش	نرخ تشخیص درست (%)	نرخ تشخیص اشتباه (%)	دقت کلی (%)
Mask R_CNN	۹۹/۱	۰/۸	۹۸/۸
CRF-RNN	۸۲/۲	۱۷/۵	۸۱/۹



شکل ۲۰- خروجی روش CRF-RNN



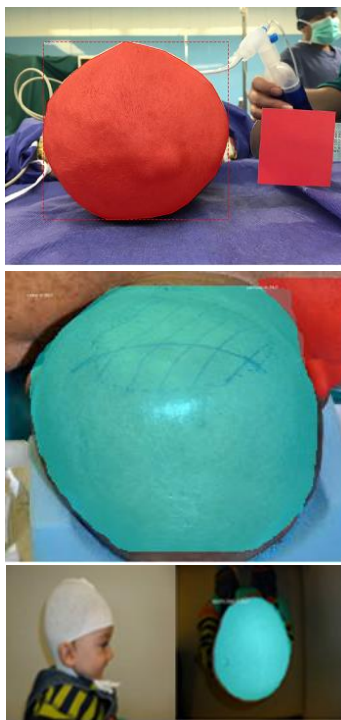
شکل ۱۹- خروجی روش Mask R_CNN



شکل ۱۸- تصاویراز سر نوزاد



شکل ۲۳- خروجی روش CRF-RNN



شکل ۲۲- خروجی روش Mask R_CNN



شکل ۲۱- تصاویر چالشی

۲-۶- ارزیابی دقت روش تخمین محیط سر

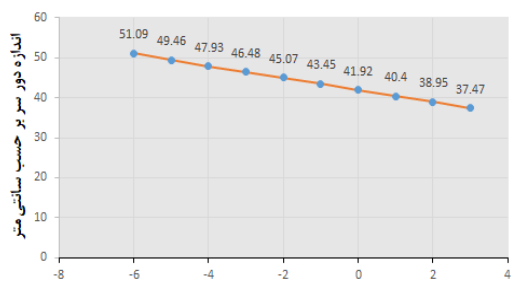
به‌منظور ارزیابی نتایج پایانی یعنی اندازه محیط سر نوزاد، ۱۰ تصویر از سر نوزادان مختلف در مرکز طبی کودکان تهیه گردید، همچنین محیط سر این نوزادان نیز با متر نواری اندازه‌گیری شد، سپس بر روی هر کدام از این تصاویر الگوریتم پیشنهادی اجرا گردید و اندازه محیط سر هر یک از نوزادان محاسبه شد. برای ارزیابی نتایج

محاسبه‌شده توسط الگوریتم پیشنهادی نتایج به‌دست‌آمده را با مقدار واقعی (اندازه‌گیری شده با متر نواری) مقایسه شد بنابراین برای هر کدام از این تصاویر مقدار خطای مطلق که میزان اختلاف مقدار اندازه‌گیری شده با مقدار واقعی می‌باشد محاسبه گردید، اما همان‌طور که مشخص است خطای مطلق نشان‌دهنده دقت آزمایش نمی‌باشد و باید خطای نسبی اندازه‌گیری‌ها که نسبت خطای مطلق به اندازه واقعی می‌باشد نیز محاسبه شود تا دقت کار مشخص



شکل ۲۴- تصویر ماکت سر نوزاد

در این آزمایش مشخص شد که روند ایجاد خطای جابجا گرفتن برچسب مقیاس به صورت خطی می باشد و هرچقدر برچسب مقیاس دورتر از محل درست خود نسبت به مرکز عدسی دوربین عکس برداری قرار گیرد محیط دور سر بزرگتر از واقعیت محاسبه می شود و هرچقدر این برچسب مقیاس نزدیکتر از محل درست خود نسبت به مرکز عدسی دوربین عکس برداری قرار گیرد محیط دور سر کوچکتر از واقعیت محاسبه می شود. روند این خطا در شکل (۲۵) به صورت یک نمودار مشخص شده است.



شکل ۲۵- نمایش تغییرات اندازه دور سر ناشی از جابجا گرفتن برچسب

بنابراین لازم است که در هنگام اندازه گیری، برچسب مقیاس، درست در محل مورد نظر قرار گیرد در غیر این صورت اندازه گیری ما به طور متوسط با خطایی به میزان ۱/۵۰ سانتی متر به ازای هر یک سانتی متر جابجا گرفتن برچسب مقیاس ایجاد خواهد شد.

۲-۲-۶- ارزیابی هندسه تصویربرداری مرکزی و هندسه تصویربرداری موازی قائم در دقت

در این ارزیابی هدف مقایسه هندسه تصویربرداری مرکزی، تک عکس که در این پژوهش از آن استفاده شده است با هندسه تصویربرداری موازی قائم می باشد؛ بنابراین تعدادی تصویر از ماکت نوزاد گرفته شد و با استفاده از

گرد و بنابراین می توان نتایج ارزیابی را در جدول (۳) مشاهده نمود. همان طور که در جدول مشاهده می شود درصد خطا نیز مشخص شده است.

۳- ارزیابی مقادیر اندازه گیری شده

درصد خطا (%)	خطای نسبی	خطای مطلق	مقدار واقعی (cm)	مقدار اندازه گیری شده (cm)
۱/۸۱%	۰/۰۱۸	۰/۷۶	۴۲/۱	۴۲/۸۶
۲/۰۱%	۰/۰۲۰	۰/۸۵	۴۲/۲۵	۴۳/۱
۱/۸۰%	۰/۰۱۸	۰/۶۷	۳۷/۲	۳۶/۵۳
۲/۱۵%	۰/۰۲۱	۰/۷۱	۳۳/۰۵	۳۳/۷۶
۱/۸۲%	۰/۰۱۸	۰/۷۰	۳۸/۵	۳۷/۸
۱/۱۹%	۰/۰۱۲	۰/۴۷	۳۹/۵	۳۹/۰۳
۱/۳۵%	۰/۰۱۴	۰/۴۴	۳۲/۵	۳۲/۹۴
۲/۴۹%	۰/۰۲۵	۱/۱	۴۴/۱۵	۴۳/۰۵
۳/۱۵%	۰/۰۳۲	۱/۱	۳۴/۹	۳۳/۸
۱/۷۵%	۰/۰۱۸	۰/۶۵	۳۷/۱	۳۷/۷۵

همان طور که در جدول (۳) مشاهده شد خطای اندازه گیری این روش در حدود ۱ تا ۳ درصد می باشد.

۶-۲-۱- ارزیابی تأثیر جابجا گرفتن برچسب مقیاس در محاسبه محیط سر

به منظور بررسی میزان تأثیر جابجا گرفتن برچسب مقیاس در تصویر از یک ماکت سر استفاده گردید که بر روی این ماکت محل دقیق قرار گرفتن برچسب مقیاس که در ارتفاع بزرگترین منحنی تشکیل دهنده سر می باشد و در واقع همان محلی است که با استفاده از متر نواری اندازه گیری می شود و سه نقطه مشخص کننده آن برآمدگی بالای دو ابرو و همچنین برآمدگی پشت سر می باشد، مشخص شد و همچنین تعداد ۹ محل دیگر که هر کدام به فاصله یک سانتی متر از هم قرار گرفته بودند مشخص شد. تصویر این ماکت را می توان در شکل (۲۴) مشاهده نمود. در هر مرحله عکس برداری که از فاصله ۵۰ سانتی متری انجام می شد، برچسب مقیاس در یکی از این نقاط قرار می گرفت بعد از اتمام عکس برداری ها محاسبه محیط دور سر برای هر یک از این تصاویر با استفاده از روش پیشنهادی صورت گرفت. پس از محاسبه محیط های سر مشخص شد که هر یک سانتی متر جابجا گرفتن برچسب مقیاس خطایی حدود ۱/۴۰ تا ۱/۶۰ سانتی متر در اندازه گیری محیط سر ایجاد می کند.

نرم‌افزار Agisoft ارتوفتو از ماکت نوزاد تهیه شد که می‌توان در شکل (۲۶) مشاهده نمود، از طرف دیگر یک تک تصویر از ماکت نوزاد گرفته شد و بر روی هر دو، الگوریتم پیشنهادی در این پژوهش اجرا شد. مقدار واقعی دور سر ۴۰/۹ سانتی‌متر می‌باشد، مقدار به‌دست‌آمده از تک تصویر ۴۱/۹۲ سانتی‌متر و مقدار به‌دست‌آمده از اجرای الگوریتم بر روی ارتوفتو ۴۱/۳ سانتی‌متر می‌باشد؛ بنابراین می‌توان نتیجه گرفت که مقداری از خطای روش پیشنهادی به دلیل استفاده از هندسه تصویربرداری مرکزی می‌باشد ولی با توجه به این‌که تهیه اورتوفتو نیازمند گرفتن تصاویر زیاد و همچنین تخصص در زمینه فتوگرامتری دارد روش تک تصویر ارجحیت دارد.



شکل ۲۶- اورتوفتو تولید شده

۷- نتیجه‌گیری

در این پژوهش روشی غیر تماسی با استفاده از یک تصویر گرفته‌شده از بالای سر نوزاد به‌منظور محاسبه محیط سر نوزاد ارائه شد که مشکلات روش قبلی را که با

استفاده از متر نواری و به‌صورت تماسی می‌باشد را ندارد. اولین گام در این روش شناسایی و قطعه‌بندی سر نوزاد در تصویر می‌باشد که برای این کار دو روش MaskR-CNN و CRF-RNN با استفاده از معیارهای استخراج‌شده از ماتریس درهم‌ریختگی مورد مقایسه قرار گرفت و مشخص شد که روش MaskR-CNN با دقت کلی ۹۸/۸ درصد روش مناسب‌تری برای این پژوهش می‌باشد در ادامه پیکسل‌های تشکیل‌دهنده دور سر، با لبه یابی سر در تصویر مشخص شد. در مرحله بعد رزولوشن مکانی پیکسل‌ها در واقعیت، با توجه به برچسب مقیاس موجود در تصویر محاسبه شد و در نهایت با توجه به رزولوشن مکانی به‌دست‌آمده و تعداد پیکسل‌های تشکیل‌دهنده دور سر محیط سر محاسبه شد. برای ارزیابی، روش ارائه‌شده بر روی ۱۰ تصویر اجرا شد و نتایج به‌دست‌آمده با مقادیر واقعی مورد مقایسه قرار گرفت و خطای این روش در حدود ۱ تا ۳ درصد برآورد شد. در ارزیابی‌های صورت گرفته نیز مشخص شد که قرار گرفتن برچسب مقیاس در محل درست خود یعنی دقیقاً در کنار محلی که اندازه آن مدنظر می‌باشد بسیار حائز اهمیت می‌باشد. برای کارهای آتی پیشنهاد می‌شود که به‌جای استفاده از برچسب مقیاس از الگوریتم‌های یادگیری عمیق برای تشخیص عمق و حل مسئله مقیاس استفاده شود، همچنین برای اندازه‌گیری سر نوزادانی که دارای مو می‌باشند پیشنهاد می‌شود از تصاویر حاصل از سنسورهای حرارتی استفاده شود.

مراجع

- [1] H. E. Akbari, F. Esmaeili, "Detection of subcutaneous vessels using 3D modeling with the aid of infrared microlens," Geomatics National Conference, 2016.
- [2] H. E. Fathi, F. Esmaeili, "Monitoring of the temperature of the ear and face when talking to a cellphone using close range photogrammetric based on thermal images," Geomatics National Conference, 2018.
- [3] A. Grün and P. Niederer, "Photogrammetry and remote sensing in medicine biostereometry and medical imaging," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 45, 1990.
- [4] "National organization for civil registration," "http://www.sabteahval.ir/", 2018.
- [5] M. Ichiyama, S. Ohga, M. Ochiai, K. Fukushima, M. Ishimura, M. Torio, et al., "Fetal hydrocephalus and neonatal stroke as the first presentation of protein C deficiency," Brain and Development, vol. 38, pp. 253-256, 2016.
- [6] M. H. Siegenthaler, "Methods to diagnose, classify, and monitor infantile deformational plagiocephaly and brachycephaly: a narrative review," Journal of chiropractic medicine, vol. 14, pp. 191-204, 2015.
- [7] S. Mummolo, A. Nota, E. Marchetti, G. Padricelli, and G. Marzo, "The 3D Tele Motion Tracking for the Orthodontic Facial Analysis," BioMed research international, vol. 2016, 2016.

- [8] I. Barbero-García, J. L. Lerma, Á. Marqués-Mateu, and P. Miranda, "Low-cost smartphone-based photogrammetry for the analysis of cranial deformation in infants," *World neurosurgery*, vol. 102, pp. 545-554, 2017.
- [9] J. Shotton, T. Sharp, M. Finocchio, R. Moore, et al., "Real-time human pose recognition in parts from single depth images," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011, pp. 1297-1304.
- [10] B.-K. Park, J. C. Lumeng, C. N. Lumeng, S. M. Ebert, and M. P. Reed, "Child body shape measurement using depth cameras and a statistical body shape model," *Ergonomics*, vol. 58, pp. 301-309, 2015.
- [11] N. Sokolover, M. Phillip, L. Sirota, A. Potruch, N. Kiryati, G. Klinger, et al., "A novel technique for infant length measurement based on stereoscopic vision," *Archives of disease in childhood*, vol. 99, pp. 625-628, 2014.
- [12] E. H. Akbari Farahani N, Esmaeili F., "Pupil Diameter Measuring for Anisocoria Detection Using Three-Dimensional Modeling by an Infrared Micro-Lens Camera.," *GEJ*, 2018.
- [13] S. Wong, S. Robson, A. Gibson, and J. Hebden, "Real-time close range web-cam photogrammetry suited to the coordination of optical topography sensors located on the human head," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 38, pp. 608-612, 2010.
- [14] B. King, "A webcam-based photogrammetric system for the measurement of facial landmarks on newborns," *Photogrammetric Engineering & Remote Sensing*, vol. 78, pp. 1285-1294, 2012.
- [15] S. Pogulwar, R. Gautam, A. Kalia, A. Hegde, and E. Ahmad, "Photographic Assessment of Cephalometric Measurements," *Int J of Craniofac Res*, vol. 2, pp. 21-26, 2014.
- [16] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *International Journal of Multimedia Information Retrieval*, vol. 7, pp. 87-93, 2018.
- [17] H.-D. Cheng, X. H. Jiang, Y. Sun, and J. Wang, "Color image segmentation: advances and prospects," *Pattern recognition*, vol. 34, pp. 2259-2281, 2001.
- [18] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27-48, 2016.
- [19] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278-2324, 1998.
- [20] J. Dai, K. He, and J. Sun, "Boxsup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1635-1643.
- [21] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580-587.
- [22] S. Gidaris and N. Komodakis, "Object detection via a multi-region and semantic segmentation-aware cnn model," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1134-1142.
- [23] B. Romera-Paredes and P. H. S. Torr, "Recurrent instance segmentation," in *European Conference on Computer Vision*, 2016, pp. 312-329.
- [24] B. Hariharan, P. Arbeláez, R. Girshick, and J. Malik, "Hypercolumns for object segmentation and fine-grained localization," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 447-456.
- [25] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, et al., "Conditional random fields as recurrent neural networks," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1529-1537.
- [26] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017, pp. 2980-2988.
- [27] T.-Y. Lin, P. Dollár, S. J. Belongie, "Feature Pyramid Networks for Object Detection," in *CVPR*, 2017, p. 4.
- [28] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91-99.
- [29] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, pp. 679-698, 1986.
- [30] W. Abdulla, "Mask R-CNN for instance segmentation on Keras and TensorFlow," *GitHub repository*, 2017.
- [31] GitHub, "Image augmentation for machine learning experiments. ," <https://github.com/aleju/imgaug>, 2018.