

آشکارسازی توده‌ها در تصاویر اولتراسوند خودکار سه‌بعدی پستان با استفاده از Inception 3D U-Net بهبودیافته

سپیده برکت رضایی^۱، امین ملک‌محمدی^۲، احسان‌اله کوزه‌گر^۳، معصومه سلامتی^۴ و محسن سریان^۵

چکیده

سرطان پستان عامل اصلی مرگ‌ومیر ناشی از سرطان در زنان اکثر کشورهای جهان است. تشخیص سرطان پستان در مراحل اولیه تأثیر زیادی در کاهش نرخ مرگ‌ومیر دارد. تصویربرداری اولتراسوند خودکار سه‌بعدی پستان (3D ABUS) نوعی تصویربرداری است که اخیراً در کنار ماموگرافی برای تشخیص زودهنگام سرطان پستان استفاده می‌شود. حجم سه‌بعدی تولیدشده در این تصویربرداری شامل تعداد زیادی برش است. رادیولوژیست برای یافتن توده باید تمام برش‌های تصویر را بررسی کند که کاری زمان‌بر با احتمال خطای زیاد است. امروزه برای کمک به رادیولوژیست‌ها در آشکارسازی توده، سیستم‌های آشکارساز کامپیوتری (CAD) بسیاری ارائه شده است. در این مقاله، معماری 3D U-Net با قرار دادن دو نوع ماژول Inception تغییریافته در بخش کدگذار بهبود داده شده و برای آشکارسازی توده در تصاویر 3D ABUS استفاده شده است. در ماژول پیشنهادی اول که در اولین لایه کدگذار قرار می‌گیرد، ویژگی‌های سه‌بعدی متنوع با دو میدان دید متفاوت تولید می‌شود. در ماژول دوم که در لایه‌های بعدی کدگذار قرار می‌گیرد، ویژگی‌های خطی و ویژگی‌های صفحه‌ای استخراج می‌شود. مجموعه داده مورد استفاده دارای ۶۰ حجم 3D ABUS از ۴۳ بیمار و شامل ۵۵ توده است. شبکه پیشنهادی به حساسیت ۹۲٫۹٪ و میانگین FP برابر با ۲۲٫۷۵ به ازای هر بیمار دست یافته است.

کلید واژه‌ها

تصویربرداری اولتراسوند خودکار سه‌بعدی پستان، سیستم آشکارساز کامپیوتری، شبکه عصبی کانولوشنی سه‌بعدی، آشکارسازی توده، Inception

۱ - مقدمه

در میان زنان، سرطان پستان شایع‌ترین نوع سرطان در اکثر کشورها (۱۵۹ از ۱۸۵ کشور) و همچنین علت اصلی مرگ و میر ناشی از سرطان در بین زنان در ۱۱۰ کشور جهان است [۱]. تشخیص سرطان پستان در مراحل اولیه می‌تواند به طور چشمگیری میزان مرگ و میر ناشی از آن را کاهش دهد. برای این منظور از روش‌های مختلف تصویربرداری پستان استفاده می‌شود. در بین این روش‌ها ماموگرافی به عنوان استاندارد فعلی برای تصویربرداری پستان شناخته شده است. اما ماموگرافی به دلیل استفاده از تابش یون‌ساز، برای زنان زیر ۳۰ سال و زنان باردار مناسب نیست و همچنین برای تشخیص سرطان در زنان با پستان‌های متراکم چندان مؤثر نیست. به همین دلیل تصویربرداری اولتراسوند پستان (دستی یا خودکار) به عنوان مکمل ماموگرافی استفاده می‌شود [۲].

^۱ دانشجوی دکتری هوش مصنوعی، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران.

رایانامه: s_barekat@comp.iust.ac.ir

^۲ دانشجوی کارشناسی ارشد هوش مصنوعی، دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران.

رایانامه: a_malekmohammadi@comp.iust.ac.ir

^۳ دانشکده فنی و مهندسی شرق گیلان، دانشگاه گیلان، گیلان، ایران.

رایانامه: kozegar@guilan.ac.ir

^۴ پژوهشگاه رویان، پژوهشکده زیست‌شناسی و علوم پزشکی تولید مثل جهاد دانشگاهی، مرکز تحقیقات پزشکی تولیدمثل، گروه تصویربرداری تولیدمثل، تهران، ایران.

رایانامه: dr.m.salamati@gmail.com

^۵ دانشکده مهندسی کامپیوتر، دانشگاه علم و صنعت ایران، تهران، ایران.

رایانامه: soryani@iust.ac.ir

در این مقاله یک شبکه عصبی کانولوشنی اصلاح شده مبتنی بر 3D U-Net، برای آشکارسازی توده‌ها در تصاویر 3D ABUS ارائه شده است. برای بخش کدگذار این شبکه، دو نوع ماژول Inception طراحی شده است.

سازمان‌دهی مقاله به این صورت است که در بخش ۲، کارهای مرتبط را به صورت خلاصه مرور کرده‌ایم. در بخش ۳، مجموعه داده را معرفی و در بخش ۴، روش پیشنهادی را شرح داده‌ایم. نتایج آزمایش‌ها در بخش ۵ آورده شده است. در بخش ۶، بحث و در بخش ۷ نتیجه‌گیری را بیان کرده‌ایم.

۲- کارهای مرتبط

یکی از شبکه‌های عصبی کانولوشنی که به صورت گسترده در پردازش تصاویر پزشکی سه‌بعدی استفاده می‌شود، 3D U-Net [۱۴] است. این شبکه مشابه 2D U-Net [۱۱] شامل یک مسیر کدگذاری برای تحلیل کل حجم و یک مسیر کدگشایی برای تولید بخش بندی با رزولوشن کامل است. این شبکه حجم سه‌بعدی را به عنوان ورودی دریافت و توسط عملیات سه‌بعدی، شامل کانولوشن، ادغام حداکثر^۱ و لایه‌های up-convolution، آن‌ها را پردازش می‌کند. در مسیر کدگذاری، هر لایه شامل دو کانولوشن $3 \times 3 \times 3$ و به دنبال هر کدام یک واحد یکسوساز خطی^۲ (ReLU)، سپس یک ادغام حداکثر $2 \times 2 \times 2$ با گام ۲ در هر بُعد است. در مسیر کدگشایی، هر لایه شامل یک up-convolution $2 \times 2 \times 2$ با گام ۲ در هر بُعد، سپس دو کانولوشن $3 \times 3 \times 3$ و به دنبال هر کدام یک واحد ReLU است. اتصال‌های پرشی^۳ از لایه‌های با رزولوشن یکسان در مسیر کدگذاری، ویژگی‌های اساسی رزولوشن بالا را برای مسیر کدگشایی فراهم می‌کنند. در لایه آخر یک کانولوشن $1 \times 1 \times 1$ تعداد کانال‌های خروجی را به تعداد برجسب‌ها کاهش می‌دهد. در این شبکه، در هر لایه از کدگذار تعداد کانال‌ها نسبت به لایه قبل دو برابر می‌شود. همچنین این شبکه از نرمال‌سازی دسته‌ای^۴ (BN) قبل از هر ReLU، به منظور همگرایی سریع‌تر استفاده می‌کند.

نا وانگ و همکاران [۱۸] از 3D U-Net به عنوان معماری پایه برای آشکارسازی توده در تصاویر 3D ABUS استفاده کردند. آن‌ها یک تابع هزینه آستانه با افزودن یک لایه نقشه آستانه به انتهای شبکه برای فراهم کردن آستانه تطبیقی سطح و کسل به منظور اصلاح نقشه احتمال و دستیابی به حساسیت زیاد و مثبت کاذب کم، ارائه کردند. نقشه آستانه پیشنهادی می‌تواند با استفاده از اطلاعات حاصل از ویژگی‌های یادگرفته‌شده، اطلاعات برجسب و نقشه احتمال پیش‌بینی شده، آستانه تطبیقی سطح و کسل را برای رده‌بندی و کسل‌ها به سرطانی و غیر سرطانی فراهم کند و به تعادل

تصویربرداری اولتراسوند دستی^۱ محدودیت‌هایی دارد: زمان‌بر و وابسته به کاربر است، قابلیت استانداردسازی و تولید مجدد ندارد. اما تصویربرداری اولتراسوند خودکار پستان^۲ (ABUS) بر این محدودیت‌ها غلبه کرده است [۳]. علاوه بر این، تصویربرداری ABUS سه‌بعدی قابلیت تولید حجم‌های سه‌بعدی برای هر پستان را دارد. به دلیل کیفیت پایین این تصاویر و خطای رادیولوژیست هنگام بررسی تعداد زیاد برش‌های تصاویر، وجود یک سیستم آشکارساز کامپیوتری (CAD) می‌تواند کمک زیادی به رادیولوژیست در تشخیص توده‌ها نماید و احتمال تشخیص نادرست و یا عدم تشخیص را تا حد زیادی کاهش دهد. ون‌زلست و همکاران [۴] نشان دادند که استفاده از سیستم آشکارساز کامپیوتری برای تصاویر ABUS می‌تواند بدون تحت تأثیر قرار دادن صحت، زمان خواندن رادیولوژیست را کاهش دهد و همچنین از عدم تشخیص توده‌های بدخیم پستان توسط رادیولوژیست جلوگیری کند.

سیستم‌های CAD را می‌توان به دو نوع تقسیم کرد: سنتی و مبتنی بر یادگیری عمیق. اصلی‌ترین تفاوت بین این دو نوع سیستم در نحوه استخراج ویژگی‌ها است. در سیستم‌های CAD سنتی، ویژگی‌ها توسط انسان تعریف می‌شوند. اما در سیستم‌های CAD مبتنی بر یادگیری عمیق، ویژگی‌ها توسط شبکه به صورت خودکار استخراج می‌شوند [۵]. در [۶] کوزه‌گر و همکاران سیستم‌های CAD سنتی برای آشکارسازی توده‌ها در تصاویر 3D ABUS مرور کردند. امروزه سیستم‌های CAD مبتنی بر یادگیری عمیق در تجزیه و تحلیل تصاویر پزشکی از سیستم‌های سنتی پیشی گرفته‌اند [۷].

شبکه عصبی کانولوشنی^۳ (CNN) یک نوع مدل یادگیری عمیق است که توانایی زیادی در تجزیه و تحلیل تصاویر پزشکی نشان داده است [۸]. یاپ و همکاران [۹] از سه CNN دوبعدی (LeNet [۱۰]، U-Net دوبعدی [۱۱] و FCN-AlexNet [۱۲]) برای آشکارسازی توده‌ها در تصاویر اولتراسوند دوبعدی استفاده کردند. استفاده مستقیم از شبکه‌های عصبی کانولوشنی دوبعدی برای آشکارسازی توده‌ها در حجم‌های 3D ABUS نمی‌تواند به طور مؤثر از اطلاعات سه‌بعدی ارائه شده در این حجم‌ها استفاده کند، بنابراین ممکن است حساسیت^۴ و تعداد مثبت کاذب^۵ (FP) رضایت بخشی بدست نیاید [۱۳]. تاکنون از شبکه‌های عصبی کانولوشنی سه‌بعدی 3D U-Net [۱۴]، 3D RPN [۱۵]، VGG-16 [۱۶] و 3D DenseNet [۱۷] به عنوان شبکه‌های پایه برای آشکارسازی توده‌ها در تصاویر 3D ABUS استفاده شده است.

^۱ Max Pooling
^۲ Rectified Linear Unit
^۳ Skip Connection
^۴ Batch Normalization

^۱ Handheld Ultrasound
^۲ Automated Breast Ultrasound
^۳ Convolutional Neural Networks
^۴ Sensitivity
^۵ False Positive

این مقاله، تصاویر 3D ABUS به صورت داده‌های متوالی در نظر گرفته می‌شود که هر توالی از سه برش مجاور در تصویر ورودی کدگذار تشکیل می‌شود. نویسندگان یک ماژول حافظه طولانی کوتاه‌مدت دوطرفه^۹ (BLSTM) همراه با ماژول توجه SC برای مدل‌سازی همبستگی بین برش‌ها طراحی کردند که از اطلاعات بین برشی برای کمک به بخش‌بندی و کاهش مثبت‌های کاذب استفاده می‌کند.

یو و همکاران [۲۶] به منظور مکان‌یابی توده‌ها در تصاویر 3D ABUS، یک شبکه آشکارساز سه بعدی مبتنی بر شبکه پیشنهاد ناحیه^{۱۱} (RPN) سه‌بعدی ارائه کردند. سپس از یک شبکه رده‌بند برای رده‌بندی توده‌ها به خوش‌خیم یا بدخیم استفاده کردند.

چیانگ و همکاران [۲۷] با استفاده از یک 3D VGG-16 ساده شده، احتمال تومور بودن هر حجم موردعلاقه^{۱۱} (VOI) در تصاویر 3D ABUS را تخمین زدند. مون و همکاران [۲۸] در شبکه 3D VGG-16 ساده شده [۲۷] با dropout با احتمال ۰٫۲۵، در لایه‌های max-pooling و از dropout با احتمال ۰٫۵، در همه لایه‌های کاملاً متصل استفاده کردند. همچنین آن‌ها به جای تابع فعال‌سازی ReLU از Leaky ReLU استفاده کردند زیرا واحدهای ReLU هنگامی که گرادیان صفر است، فعال نیستند. در این مقاله به منظور کاهش واریانس و بایاس مدل و در نتیجه بهبود عملکرد، از یادگیری گروهی^{۱۳} استفاده شده است. در روش گروهی چند مدل با یکدیگر ترکیب می‌شوند تا نتیجه بهتری تولید کنند. در این روش برای ساخت یک مدل آشکارساز یا رده‌بند قوی‌تر، از روشهای مختلف ترکیب، مانند میانگین‌گیری، رأی‌گیری و ترکیب کردن توسط یادگیری استفاده می‌شود. در این مقاله به دلیل سادگی و کارایی، از میانگین‌گیری به عنوان رویکرد ترکیب استفاده می‌شود. از این رو، هر VOI به دو مدل پیشنهادی 3-D VGG-16 ساده شده و 3-D DenseNet [۱۷] داده می‌شود و در انتها میانگین نتایج حاصل به عنوان خروجی گروهی نهایی برای آن VOI در نظر گرفته می‌شود.

۳- مجموعه داده

در این مقاله از تصاویر 3D ABUS استفاده کردیم. این تصاویر از مرکز پزشکی دانشگاه رادبود (نایمخن^{۱۴}، هلند)، مرکز یرون بوش زیکنهوس^{۱۵} (دن بوش^{۱۶}، هلند)، بیمارستان مرکزی فالون (فالون، سوئد) و انیستیتو جولز بوردت^{۱۷} (بروکسل، بلژیک) تهیه شده‌اند.

خوبی بین حساسیت بالا و مثبت کاذب کم برسد. آن‌ها همچنین یک روش نظارت عمیق متراکم^۱ (DDS) طراحی نمودند که با یادگیری مؤثر ویژگی‌های متمایزکننده چند مقیاسی از همه لایه‌ها و تقویت جریان گرادیان کل شبکه، موجب بهبود قابل توجه حساسیت می‌شود. آن‌ها از آنتروپی متقاطع متعادل‌شده رده^۲ برای مقابله با مسئله نمونه‌های آموزشی مثبت محدود و تابع هزینه همپوشانی^۳ برای یادگیری بهتر بازنمایی‌های متمایزکننده سرطان، به صورت ترکیبی به عنوان تابع هزینه DDS استفاده کردند.

یی وانگ و همکاران [۱۹]، در شبکه [۱۸] از کانولوشن‌های گسترش‌یافته سه‌بعدی^۴ برای جمع‌آوری سیستماتیک اطلاعات زمینه‌ای چندمقیاسی استفاده کردند که برای بهبود حساسیت آشکارسازی توده‌های کوچک مفید است. آن‌ها برای بهبود عملکرد DDS، در تابع هزینه ترکیبی به جای تابع هزینه آنتروپی متقاطع متعادل‌شده رده، از تابع هزینه دایس برای مقابله با مسئله محدود بودن نمونه‌های آموزشی مثبت استفاده کردند.

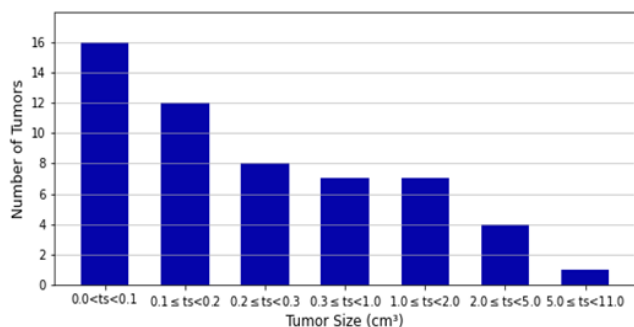
فیقایان وانگ و همکاران [۲۰]، [۲۱] یک مدل آشکارساز براساس معماری U-Net 3D ارائه کردند که در آن به منظور جلوگیری از مسئله ناپدید شدن گرادیان، تمام لایه‌های کانولوشنی ساده با بلوک‌های باقیمانده جایگزین شدند. همچنین آن‌ها برای ترکیب اطلاعات با وضوح بالا و اطلاعات با وضوح پایین از اتصال پرشی توجه^۵، به منظور تأکید بیشتر مدل به نواحی مفید مکانی و بهبود توانایی مکان‌یابی استفاده کردند.

یی وانگ و همکاران [۱۳] برای بهبود شبکه 3D U-Net، دو نوع بلوک Inception و یک بلوک نظارت عمیق ترکیبی پیشنهاد کردند. آن‌ها همچنین یک تابع هزینه نامتقارن^۶ جدید مبتنی بر شاخص تورسکی^۷، برای کمک به ایجاد تعادل بین نواحی مثبت کاذب و منفی کاذب^۸ (FN) و در نتیجه بهبود حساسیت آشکارسازی توده‌های کوچک ارائه کردند.

پن و همکاران [۲۲] یک شبکه جدید برای بخش‌بندی توده‌ها در تصاویر 3D ABUS طراحی کردند که مشابه معماری U-Net از دو مسیر کدگذاری و کدگشایی تشکیل شده است. آن‌ها از VGG16 [۲۳]، AlexNet [۲۴] و ResNet18 [۲۵] به عنوان کدگذار استفاده کردند. به دلیل اختلاف بین اطلاعات معنایی و اطلاعات مکانی، تلفیق ساده ویژگی‌های سطح بالا و ویژگی‌های سطح پایین کمتر می‌تواند مؤثر باشد. به همین دلیل در این مقاله برای بهبود اثربخشی، یک ماژول توجه مکان-کانال (SC) برای ترکیب ویژگی‌های سطح پایین و سطح بالا طراحی شده است. در

^۹ Bi-directional Long Short-Term Memory
^{۱۰} Region Proposal Network
^{۱۱} Volume of Interest
^{۱۲} Overfitting
^{۱۳} Ensemble Learning
^{۱۴} Nijmegen
^{۱۵} Jeroen Bosch Ziekenhuis
^{۱۶} Den Bosch
^{۱۷} Jules Bordet Institute

^۱ Densely Deep Supervision
^۲ Class-Balanced Cross Entropy Loss
^۳ Overlap Loss
^۴ 3D Dilated Convolutions
^۵ Attention Skip Connection
^۶ Asymmetric Loss
^۷ Tversky index
^۸ False Negative



شکل (۱): توزیع اندازه توده‌ها در مجموعه داده

۴- روش پیشنهادی

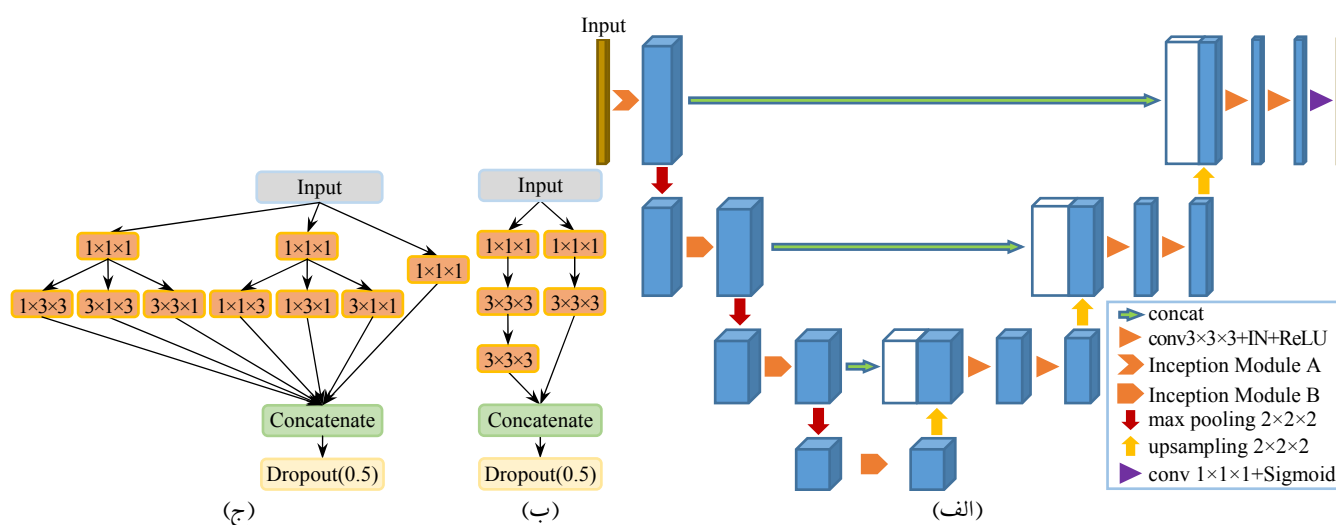
در این مقاله به منظور آشکارسازی توده‌ها در تصاویر 3D ABUS، شبکه‌ای مبتنی بر Inception 3D U-Net پیشنهاد کردیم که در شکل (۲-الف) نشان داده شده است. همانطور که مشاهده می‌شود، در بخش کدگذار این شبکه از دو نوع ماژول Inception بهبودیافته استفاده شده است. این ماژول‌ها در شکل (۲-ب و ج) نشان داده شده‌اند. ساده‌ترین راه برای بهبود عملکرد شبکه‌های عصبی عمیق، افزایش اندازه آن‌ها شامل افزایش عمق (تعداد سطوح شبکه) و عرض (تعداد واحدها در هر سطح) است. اما افزایش یکنواخت اندازه شبکه، دو اشکال عمده دارد. شبکه با اندازه بزرگتر، به معنای تعداد پارامترهای بیشتر است که باعث می‌شود شبکه، بیشتر در معرض بیش‌برازش قرار گیرد، به خصوص اگر تعداد نمونه‌های برچسب‌دار در مجموعه آموزشی کم باشد. اشکال دیگر افزایش یکنواخت اندازه شبکه، افزایش چشمگیر استفاده از منابع محاسباتی است. به عنوان مثال، در یک شبکه عصبی عمیق، اگر دو لایه کانولوشنی زنجیر شوند، هر افزایش یکنواخت در تعداد فیلترهای آن‌ها، منجر به افزایش توان دو محاسبات می‌شود. اگر ظرفیت اضافه‌شده به صورت ناکارآمد استفاده شود (مثلاً اگر اکثر وزن‌ها در نهایت نزدیک به صفر باشند)، بیشتر محاسبات هدر می‌رود. از آنجایی که بودجه محاسباتی همیشه محدود است،

این تصاویر توسط دو نوع دستگاه تصویربرداری ABUS به نام Somo-V و ACUSON S2000 گرفته شده‌اند.

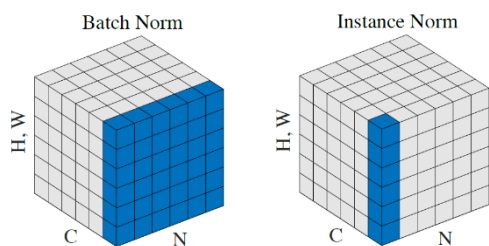
تصاویر سیستم Somo-V دارای حداکثر ابعاد ۱۴,۶ سانتی‌متر در ۱۶,۸ سانتی‌متر روی صفحه کرونال و حداکثر عمق ۴,۶۸ سانتی‌متر هستند. مبدل این سیستم از فرکانس ۸ یا ۱۰ مگاهرتز استفاده می‌کند. هر حجم تولیدشده توسط این سیستم دارای حداقل اندازه ۰,۹۲ میلی‌متر (در امتداد مبدل) در ۰,۱۳ میلی‌متر (در جهت عمق) در ۰,۶ میلی‌متر (در جهت جاروب کردن) است.

تصاویر سیستم ACUSON S2000 دارای حداکثر ابعاد ۱۵,۴ سانتی‌متر در ۱۶,۸ سانتی‌متر روی صفحه کرونال و حداکثر عمق ۶ سانتی‌متر هستند. فرکانس مبدل این سیستم بین ۵ تا ۱۴ مگاهرتز متغیر است و می‌تواند براساس اندازه پستان تنظیم شود. هر حجم تولیدشده توسط این سیستم دارای حداقل اندازه ۰,۷ میلی‌متر (در جهت عمق) در ۰,۲۱ میلی‌متر (در امتداد مبدل) در ۰,۰۷ میلی‌متر (در جهت عمق) در ۰,۵۲ میلی‌متر (در جهت جاروب کردن) است. در این مقاله، همه حجم‌ها با استفاده از درون‌یابی خطی نمونه‌گذاری شده‌اند تا به طور یکنواخت هر بُعد و کسل‌ها ۰,۶ میلی‌متر شود.

مجموعه داده موردنظر شامل ۶۰ حجم 3D ABUS از ۴۳ بیمار است. میانگین ابعاد این تصاویر $70 \times 26 \times 277$ و کسل است. ۳۴ بیمار دارای ۱ توده، ۶ بیمار دارای ۲ توده و ۳ بیمار دارای ۳ توده هستند. به این ترتیب مجموعه داده فعلی شامل ۵۵ توده است. توزیع اندازه توده‌ها در شکل (۱) نشان داده شده است. یکی از چالش‌هایی که برای آشکارسازی توده‌های این مجموعه داده وجود دارد، تنوع اندازه توده‌ها است. در این مجموعه داده کوچکترین توده 0.13 cm^3 و بزرگترین توده 8.42 cm^3 است. یکی دیگر از چالش‌ها این است که اکثر توده‌های این مجموعه داده خیلی کوچک هستند و آشکارسازی آن‌ها بسیار سخت است. اندازه ۷۸٪ از توده‌ها زیر 1 cm^3 و ۲۹٪ از توده‌ها زیر 0.1 cm^3 است.



شکل (۲): (الف) معماری Inception 3D U-Net پیشنهادی، (ب) ماژول Inception نوع A و (ج) ماژول Inception نوع B. در ماژول‌های Inception اندازه کرنل‌های کانولوشن مشخص شده است. پس از هر کانولوشن، نرمال‌سازی دسته‌ای و ReLU اعمال می‌شود.



شکل (۳): نرمال‌سازی دسته‌ای و نمونه‌ای [۳۲]. هر زیرشکل، یک نقشه ویژگی با N دسته و C کانال است. H و W محورهای مکان هستند. پیکسل‌های آبی با میانگین و واریانس یکسان، که براساس مقادیر همه پیکسل‌های آبی محاسبه شده است، نرمال می‌شوند.

از آنجایی که تعداد داده‌ها در مجموعه داده ما، کم است، به منظور کاهش احتمال بیش‌برازش، از Dropout [۳۳] استفاده کردیم. Dropout به این صورت عمل می‌کند که در حین آموزش، درصدی از واحدهای لایه موردنظر به صورت تصادفی نادیده گرفته می‌شوند. همانطور که در شکل (۲) مشاهده می‌شود، در انتهای هر ماژول Inception، Dropout با نرخ ۰٫۵ قرار دادیم.

در این پژوهش در مرحله پیش‌پردازش، ۷ برش اول و ۷ برش آخر هر حجم حذف شده‌اند زیرا این برش‌ها مربوط به پوست و پشت دنده‌ها هستند و نمی‌توانند حاوی توده باشند [۳۴]. یکی دیگر از پیش‌پردازش‌هایی که در این مقاله انجام می‌دهیم، نرمال‌سازی سطوح خاکستری است. علت این است که تصاویر موجود توسط دستگاه‌های متفاوتی گرفته شده و در نتیجه سطوح خاکستری بافت‌های مختلف (بافت پستان، بافت چربی و ...) از یک بیمار به بیمار دیگر متفاوت است.

در این مقاله برای آموزش و آزمایش شبکه‌ها، با استفاده از روش پنجره کشویی^۲ حجم سه‌بعدی اولیه را به حجم‌های کوچکتر کوچکتر تقسیم کردیم. در هنگام آموزش، اندازه گام پنجره کشویی را نصف اندازه پنجره قرار دادیم، به این ترتیب تکه‌های^۳ سه‌بعدی حاصل با یکدیگر همپوشانی خواهند داشت. همچنین در مرحله آموزش، به منظور متعادل‌سازی تعداد تکه‌های حاوی تومور و تکه‌های فاقد تومور، به تعداد تکه‌های حاوی تومور از بین تکه‌های فاقد تومور به صورت تصادفی انتخاب کردیم. در مرحله آزمایش، نتایج حجم‌های کوچک را، دقیقاً مطابق با تقسیم حجم اولیه، به یکدیگر متصل کردیم و نتیجه سه‌بعدی نهایی را بدست آوردیم.

اکثر تومورهای موجود در مجموعه داده ما بسیار کوچک هستند. به همین دلیل، به منظور آشکارسازی بیشتر تومورهای کوچک، از اندازه تکه کوچک ($40 \times 40 \times 8$ و کسل) برای تقسیم حجم سه‌بعدی به حجم‌های کوچکتر استفاده می‌کنیم. البته استفاده از این اندازه تکه کوچک، موجب تولید تعداد زیادی FP می‌شود. برخی از این FPها دارای ویژگی‌ای هستند که مشخص می‌کند آن

توزیع کارآمد منابع محاسباتی به افزایش بی‌قید و شرط اندازه ترجیح داده می‌شود، حتی زمانی که هدف اصلی بهبود عملکرد باشد. یک روش برای حل این دو مشکل، افزایش عمق و عرض شبکه با استفاده از ماژول‌های Inception است [۲۹]. در این ماژول‌ها، با استفاده از اندازه‌های مختلف کرنل کانولوشن و به هم چسباندن نتایج آن‌ها عرض شبکه افزایش می‌یابد و قبل از این کانولوشن‌ها، از کانولوشن $1 \times 1 \times 1$ به منظور کاهش تعداد پارامترها استفاده می‌شود. به همین منظور، در این مقاله از ماژول‌های Inception برای افزایش عمق و عرض شبکه و در نتیجه بهبود عملکرد شبکه استفاده کردیم.

یکی از بخش‌های مهم شبکه، لایه اول آن است که ویژگی‌ها را مستقیماً از روی حجم اصلی بدست می‌آورد و به لایه‌های بعدی منتقل می‌کند. بنابراین اگر ویژگی‌های بهتر و کامل‌تری در لایه اول محاسبه شوند، نتیجه نهایی بهبود خواهد یافت. به همین منظور در لایه اول کدگذار از ماژول A استفاده کردیم. ماژول A از دو مسیر تشکیل شده است که یکی شامل یک کانولوشن $3 \times 3 \times 3$ و دیگری شامل دو کانولوشن $3 \times 3 \times 3$ است. به این ترتیب در لایه اول، ویژگی‌های سه‌بعدی متنوع با دو میدان دید متفاوت تولید می‌شود. در لایه‌های بعدی بخش کدگذار، از ماژول B استفاده کردیم. در ماژول B از کانولوشن‌های $1 \times 1 \times 3$ ، $1 \times 3 \times 1$ ، $3 \times 1 \times 1$ برای استخراج ویژگی‌های خطی و از کانولوشن‌های $3 \times 3 \times 3$ ، $1 \times 3 \times 3$ و $3 \times 3 \times 1$ برای استخراج ویژگی‌های صفحه‌ای استفاده کردیم.

در شبکه‌های عصبی، ورودی هر لایه تحت تأثیر پارامترهای لایه‌های قبلی است. به همین دلیل در حین آموزش شبکه‌های عصبی، با تغییر پارامترهای لایه‌های قبلی، توزیع ورودی به لایه فعلی تغییر می‌کند. بنابراین لایه‌ها باید به طور مداوم با توزیع جدیدی مطابقت پیدا کنند. این مسئله موجب کاهش سرعت آموزش می‌شود. با استفاده از نرمال‌سازی ورودی لایه‌ها، یعنی با ثابت قرار دادن میانگین و واریانس ورودی لایه‌ها، می‌توان این مشکل را حل کرد. یکی از روش‌های معمول نرمال‌سازی، نرمال‌سازی دسته‌ای (BN) است که نرمال‌سازی ویژگی‌ها را با محاسبه میانگین و واریانس در هر دسته کوچک از ورودی انجام می‌دهد [۳۰]. یکی دیگر از روش‌های نرمال‌سازی، نرمال‌سازی نمونه‌ای^۱ (IN) است که نرمال‌سازی را برای هر ورودی به صورت صورت جداگانه انجام می‌دهد [۳۱]. در شکل (۳)، تفاوت این دو نوع نرمال‌سازی نشان داده شده است. از آنجایی که داده‌های موردنظر این مقاله تفاوت زیادی با یکدیگر دارند، در اینجا نرمال‌سازی نمونه‌ای بهتر از نرمال‌سازی دسته‌ای عمل می‌کند. به همین دلیل در شبکه پیشنهادی از نرمال‌سازی نمونه‌ای استفاده کردیم.

^۲ Sliding Window
^۳ Patch

^۱ Instance Normalization

اصلی را با استفاده از آستانه‌گذاری Otsu دودویی می‌کنیم. به این ترتیب نواحی اطراف بافت پستان که تیره‌تر هستند، مقدار صفر می‌گیرند. توده‌ها و سایه‌ها که به صورت نواحی تیره داخل بافت پستان هستند، موجب ایجاد حفره‌هایی در داخل بافت می‌شوند. با استفاده از عملگرهای ریخت‌شناسی این حفره‌ها را پر می‌کنیم [۳۶].

۵- آزمایش‌ها و نتایج

۵-۱- جزئیات پیاده‌سازی

در این مقاله، پیاده‌سازی شبکه‌ها را با استفاده از کتابخانه Keras برای Tensorflow [۳۷] انجام دادیم. در همه آزمایش‌ها روش اعتبارسنجی ضربدیری^۲ 10-fold را به کار بردیم. به منظور جلوگیری از قرار گرفتن حجم‌های مربوط به یک بیمار در دو مجموعه آموزش و آزمایش، حجم‌ها را در سطح بیمار بین foldها تقسیم کردیم.

به منظور افزایش داده‌های آموزشی، از معکوس کردن افقی^۳ روی تکه‌های حاوی تومور استفاده کردیم. برای آموزش شبکه‌ها، بهینه‌ساز Adam [۳۸] را بکار بردیم. نرخ یادگیری اولیه را ۰,۰۰۱ قرار دادیم و آن را در حین آموزش کاهش دادیم. هر شبکه را ۵۰ اپیک با استفاده از تابع هزینه دایس [۳۹] آموزش دادیم. رابطه تابع هزینه دایس به صورت (۲) است. در این رابطه، y_i و p_i به ترتیب وکسل‌های درستی مرجع و وکسل‌های ناحیه شناسایی شده را نشان می‌دهند. در این رابطه، برای جلوگیری از تقسیم بر صفر، ثابت s به صورت و مخرج اضافه شده است.

$$Dice Loss = 1 - \frac{2 \sum_{t=1}^N p_t y_t + s}{\sum_{t=1}^N p_t + \sum_{t=1}^N y_t + s} \quad (2)$$

۵-۲- معیارهای ارزیابی

در مرحله آزمایش، خروجی احتمالی شبکه را با آستانه ۰,۵ به دودویی تبدیل می‌کنیم. اگر ناحیه شناسایی شده و درستی مرجع، با یکدیگر همپوشانی داشته باشند و فاصله بین مراکز آن‌ها کوچکتر از ۱۰ میلی‌متر باشد، ناحیه شناسایی شده به عنوان مثبت صحیح (TP) و در غیر این صورت به عنوان مثبت کاذب (FP) در نظر گرفته می‌شود. معیارهای ارزیابی در این مقاله، حساسیت و میانگین FP برای هر بیمار است. حساسیت درصدی از تومورهایی است که به درستی یافت شده‌اند و طبق رابطه (۳) محاسبه می‌شود. در این رابطه، FN تعداد تومورهایی است که یافت نشده‌اند.

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

ناحیه نمی‌تواند توده باشد. بنابراین پس از دودویی کردن خروجی احتمالی شبکه با آستانه ۰,۵، با استفاده از برخی ویژگی‌ها تعداد FP را کاهش می‌دهیم. یکی از این ویژگی‌ها اندازه (تعداد وکسل) ناحیه شناسایی شده است. با استفاده از این ویژگی می‌توان نواحی خیلی کوچک که مانند نویز هستند را حذف کرد. اگر اندازه ناحیه، کمتر از نصف اندازه کوچکترین توده موجود در مجموعه داده (۳۰ وکسل) باشد، آن را حذف می‌کنیم.

چربی‌ها در تصاویر 3D ABUS معمولاً به صورت نواحی نازک و تخت دیده می‌شوند. به منظور حذف این نواحی، یک مکعب مستطیل محاطی موازی با محورهای اصلی، اطراف ناحیه شناسایی شده در نظر می‌گیریم. اگر اندازه این مکعب مستطیل در محور x یا y کمتر از ۳ وکسل یا اندازه آن در محور z (عمق) یک وکسل باشد، آن ناحیه را حذف می‌کنیم. آستانه‌های در نظر گرفته شده برابر نصف کوچکترین مقدار ویژگی مربوطه در مجموعه داده است. همچنین اگر نسبت کوچکترین محور به بزرگترین محور حجم شناسایی شده کمتر از ۰,۲ باشد، آن ناحیه را حذف می‌کنیم. آستانه این ویژگی برابر با میانگین منهای دو برابر انحراف معیار این ویژگی برای توده‌های مجموعه داده در نظر گرفته شده است.

یکی دیگر از ویژگی‌های مورد استفاده برای کاهش FP، فشردگی^۱ است که میزان دوار بودن یک ناحیه را تعیین می‌کند [۳۵]. نحوه محاسبه این معیار در رابطه (۱) آورده شده است. در این رابطه، R_{volume} حجم و R_A مساحت سطح ناحیه سه‌بعدی را نشان می‌دهد. مقدار این ویژگی در بازه $[0, 1]$ و برای یک کره ۱ است. از آنجایی که توده‌ها نسبتاً دوار هستند، نواحی که از دوار بودن فاصله دارند را می‌توان حذف کرد. آستانه این ویژگی برابر با میانگین منهای دو برابر انحراف معیار این ویژگی برای توده‌های مجموعه داده در نظر گرفته شده است.

$$R_{C_volume} = \frac{R_{volume} - R_A/6}{R_{volume} - (\sqrt[3]{R_{volume}})^2} \quad (1)$$

در این مقاله، آستانه‌ها را به صورت آماری براساس مجموعه داده موجود، به گونه‌ای تعیین کردیم تا قابلیت تعمیم شبکه به داده‌های جدید حفظ شود. اگر فرض کنیم داده جدیدی وجود داشته باشد که مقادیر ویژگی‌های مذکور برای آن کوچکتر از آستانه مربوطه باشد، این آستانه‌ها به عنوان فرضیات شبکه پیشنهادی خواهند بود. به عنوان مثال، شبکه پیشنهادی، تومورهای بزرگتر از ۳۰ وکسل تقریباً معادل ۶,۵ میلی‌متر مکعب را می‌تواند شناسایی کند.

از آنجایی که حجم اصلی به تکه‌های کوچک تقسیم می‌شود، گاهی در تکه‌هایی که روی مرز بافت پستان قرار می‌گیرند، نواحی تیره پس‌زمینه به اشتباه به عنوان توده در نظر گرفته می‌شوند. به منظور حذف چنین مواردی، نواحی شناسایی شده خارج از بافت پستان را حذف می‌کنیم. برای بدست آوردن بافت پستان، حجم

^۲ Cross Validation
^۳ Horizontally Flip

^۱ Compactness

۵-۳- نتایج آزمایش‌ها

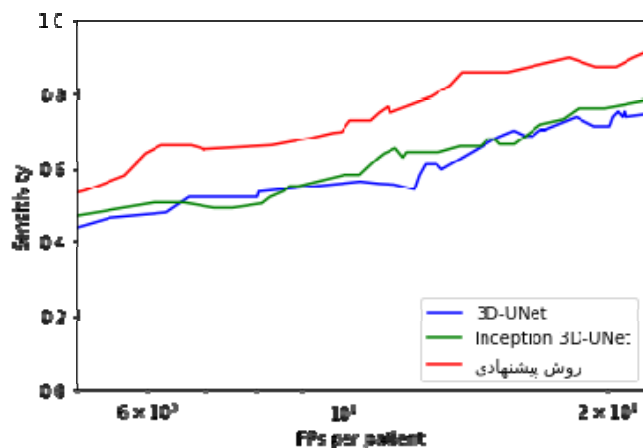
متأسفانه مجموعه داده عمومی برای ABUS 3D وجود ندارد و مجموعه داده مقاله‌های مختلف از نظر تعداد و نوع تصویربرداری بسیار متفاوت هستند. به همین دلیل نمی‌توان نتایج ارائه شده در مقاله‌ها را مستقیماً با یکدیگر مقایسه کرد. بنابراین به منظور مقایسه عادلانه، دو شبکه دیگر را پیاده‌سازی و روی مجموعه داده موجود آموزش دادیم و آزمایش کردیم.

در جدول (۱) نتایج شبکه پیشنهادی با 3D U-Net [۱۴] و Inception 3D U-Net [۱۳] مقایسه شده است. همانطور که مشاهده می‌شود، حساسیت و میانگین FP به ازای هر بیمار برای شبکه پیشنهادی بهتر از دو شبکه دیگر است. در شکل (۴) منحنی مشخصه عملیاتی فارغ از پاسخ (FROC) شبکه پیشنهادی و دو شبکه دیگر نشان داده شده است. این منحنی با تغییر آستانه مورد استفاده برای دودویی کردن نتیجه احتمالی شبکه بدست آمده است.

جدول (۱): نتایج شبکه 3D U-Net، Inception 3D U-Net و روش

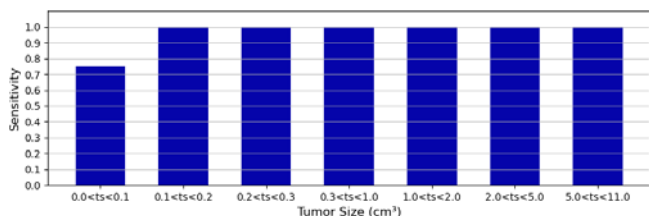
پیشنهادی، میانگین (حداقل و حداکثر)

معماری	حساسیت (%)	FP به ازای هر بیمار
3D U-Net [۱۴]	۷۶٫۹۸ (۴۰ و ۱۰۰)	۲۷/۷۱ (۴ و ۷۱٫۴)
Inception 3D U-Net [۱۳]	۸۵/۵۵ (۶۰ و ۱۰۰)	۲۸/۹۳ (۶٫۵ و ۵۹)
روش پیشنهادی	۹۲/۹۰ (۶۰ و ۱۰۰)	۲۲/۷۵ (۴ و ۴۳٫۲)

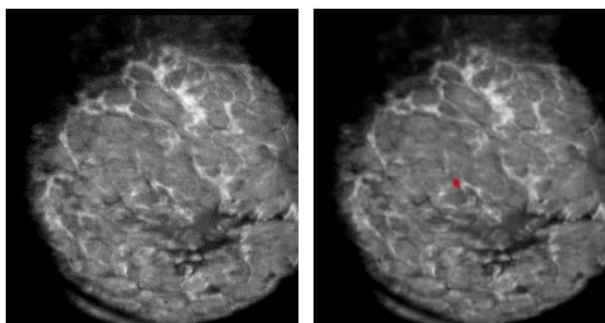


شکل (۴): منحنی مشخصه عملیاتی فارغ از پاسخ (FROC)

در شکل (۵) حساسیت شبکه پیشنهادی برای اندازه مختلف توده‌ها نشان داده شده است. همانطور که مشاهده می‌شود، توده‌هایی که یافت نشده‌اند (منفی‌های کاذب) جزو کوچکترین توده‌های مجموعه داده هستند و اندازه آن‌ها کمتر از 0.1 cm^3 است. یکی از این توده‌ها در شکل (۶) نشان داده شده است.



شکل (۵): حساسیت شبکه پیشنهادی برای اندازه مختلف توده‌ها



شکل (۶) نمونه‌ای از FN شبکه پیشنهادی. در سمت چپ، برشی از یک حجم و در سمت راست درستی مرجع با رنگ قرمز روی برش نشان داده شده است.

۶- بحث

در کار حاضر تلاش کردیم شبکه‌ای با حساسیت تا حد ممکن بالا بدست آوریم. در این شبکه میانگین FP در اولویت دوم قرار دارد زیرا در حال طراحی شبکه دیگری هستیم که نتایج این شبکه را به عنوان ورودی دریافت کند و سعی کند حساسیت را بالا نگه دارد و میانگین FP را کاهش دهد. در مجموعه داده مورد استفاده اندازه ۲۹ درصد از توده‌ها زیر 0.1 cm^3 است. آشکارسازی این توده‌های کوچک بسیار سخت است و شبکه پیشنهادی ۲۵ درصد از آن‌ها را پیدا نکرده است. به دلیل وجود این چالش نتوانستیم به حساسیت ۱۰۰ درصد دست یابیم ولی به گونه‌ای تنظیمات را انجام دادیم که تا حد ممکن، حساسیت بالا بدست آوریم. در ادامه این بخش، تأثیر عوامل مختلف را روی حساسیت و میانگین FP شبکه پیشنهادی مورد بررسی قرار می‌دهیم.

همانطور که در منحنی FROC شکل (۴) نشان داده شده است، یکی از عوامل مؤثر بر حساسیت و میانگین FP، آستانه مورد استفاده برای دودویی کردن نتیجه احتمالی شبکه است. با افزایش آستانه، حساسیت و میانگین FP کاهش می‌یابد. با توجه به هدف بدست آوردن حساسیت بالا و FP قابل قبول، در این مقاله، آستانه را ۰٫۵ در نظر گرفتیم.

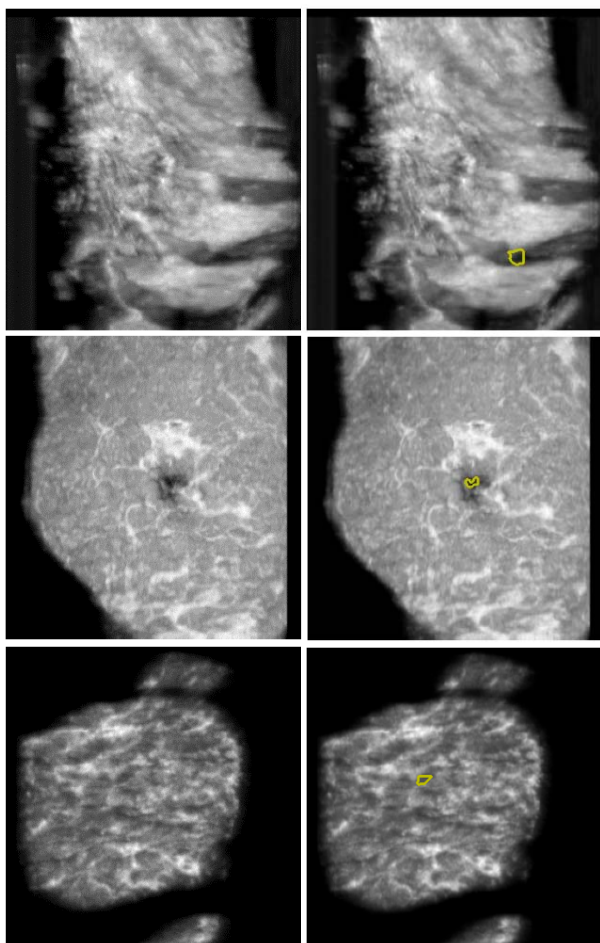
متغیر دیگری که باید مقدار مناسبی برای آن تعیین کنیم، اندازه تکه است. در جدول (۲) نتایج شبکه پیشنهادی برای چند اندازه تکه مختلف آورده شده است. با توجه به نتایج این جدول، اندازه تکه $8 \times 4 \times 4$ را برای تقسیم حجم سه‌بعدی به حجم‌های کوچکتر انتخاب کردیم. هنگامی که اندازه تکه را بزرگ کنیم، شبکه میدان دید بزرگتری دارد و در نتیجه FP کمتری بدست می‌آورد ولی امکان آشکارسازی توده‌های خیلی کوچک را از دست می‌دهد

در جدول (۵) نتیجه شبکه پیشنهادی قبل و بعد از اعمال پس‌پردازش مقایسه شده است. همانطور که مشاهده می‌شود، حساسیت قبل از پس‌پردازش بیشتر است ولی از آنجایی تعداد FP نیز خیلی زیاد است، بهتر است پس‌پردازش‌ها اعمال شوند.

جدول (۵): تأثیر پس‌پردازش روی نتایج شبکه پیشنهادی، میانگین (حداقل و حداکثر)

حساسیت (%)	FP به ازای هر بیمار	
۹۶/۵۷ (۸۰ و ۱۰۰)	۱۰۹/۵۸ (۲۸,۷۵ و ۱۶۶,۶)	بدون پس‌پردازش
۹۲/۹۰ (۶۰ و ۱۰۰)	۲۲/۷۵ (۴ و ۴۳,۲)	با پس‌پردازش

در شکل (۷) نمونه‌هایی از FPهای شبکه پیشنهادی نشان داده شده است. ردیف اول، نمونه‌ای از FP را نشان می‌دهد که داخل دنده‌ها است. با استفاده از روشی برای تشخیص دنده‌ها، می‌توان چنین مواردی را حذف کرد. ردیف دوم، نمونه‌ای را نشان می‌دهد که شبکه نوک پستان را به اشتباه توده تشخیص داده است. با استفاده از روشی برای تشخیص نوک پستان، می‌توان چنین مواردی را حذف کرد. ردیف سوم، نمونه‌ای از FP را نشان می‌دهد که مشابه توده است و به همین دلیل شبکه به اشتباه افتاده است.



شکل (۷): نمونه‌هایی از FPهای شبکه پیشنهادی. در هر ردیف، در سمت چپ برشی از یک حجم و در سمت راست نتیجه شبکه با رنگ زرد روی برش نشان داده شده است.

و حساسیت کاهش می‌یابد. علت کاهش حساسیت این است که اکثر توده‌های موجود در مجموعه داده بسیار کوچک هستند. هنگامی که از اندازه تکه کوچک استفاده می‌کنیم، شبکه قابلیت آشکارسازی توده‌های کوچک را دارد ولی چون میدان دید بزرگی ندارد، نواحی زیادی را به اشتباه توده شناسایی می‌کند. از آنجایی که در مسائل پزشکی آشکار نشدن توده‌ها به مراتب خطرناکتر از تولید FP است، اندازه تکه متناظر با بیشترین حساسیت را انتخاب کردیم.

جدول (۲): نتایج شبکه پیشنهادی برای چند اندازه تکه مختلف، میانگین (حداقل و حداکثر)

حساسیت (%)	FP به ازای هر بیمار	اندازه تکه
۹۰/۶۴ (۶۰ و ۱۰۰)	۳۷/۱۶ (۱۱,۲۵ و ۸۶,۸)	۳۲×۳۲×۸
۸۴/۴۸ (۶۰ و ۱۰۰)	۲۵/۶۶ (۱۰,۲۵ و ۴۹,۴)	۳۲×۳۲×۱۶
۹۲/۹۰ (۶۰ و ۱۰۰)	۲۲/۷۵ (۴ و ۴۳,۲)	۴۰×۴۰×۸
۸۱/۸۸ (۴۰ و ۱۰۰)	۱۸/۹۱ (۲,۷۵ و ۴۰,۶)	۴۰×۴۰×۱۶
۸۸/۰۵ (۶۰ و ۱۰۰)	۱۵/۲۵ (۶,۵ و ۲۵)	۶۴×۶۴×۸
۷۶/۶۲ (۶۰ و ۱۰۰)	۹/۷۲ (۳,۷۵ و ۱۶)	۸۰×۸۰×۸
۷۰/۶۷ (۴۲,۸۶ و ۱۰۰)	۶/۱۶ (۲,۵ و ۱۰,۲)	۱۲۰×۱۲۰×۸

توابع هزینه متفاوت که در آموزش شبکه مورد استفاده قرار گرفتند، در جدول (۳) آورده شده است. نتایج شبکه پیشنهادی هنگام استفاده از این توابع هزینه در جدول (۴) ارائه شده است. پارامترهای توابع هزینه طوری انتخاب شدند که بالاترین حساسیت را ایجاد کنند. تابع هزینه دایس بالاترین حساسیت و کمترین FP را ایجاد می‌کند و به همین دلیل این تابع هزینه را انتخاب کردیم.

جدول (۳): توابع هزینه مورد بررسی. y_i و p_i به ترتیب وکسل‌های درستی مرجع و وکسل‌های ناحیه شناسایی شده هستند.

تابع هزینه	رابطه
Binary Cross Entropy	$BCE = -\frac{1}{N} \sum_{i=1}^N [y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)]$
Weighted Binary Cross Entropy	$WBCE = -\frac{1}{N} \sum_{i=1}^N [\alpha y_i \ln(p_i) + (1 - \alpha)(1 - y_i) \ln(1 - p_i)]$
Dice Loss [۳۹]	$DL = 1 - \frac{2 \sum_{i=1}^N p_i y_i + s}{\sum_{i=1}^N p_i + \sum_{i=1}^N y_i + s}$
Combo Loss [۴۰]	$CL = \beta \left(-\frac{1}{N} \sum_{i=1}^N [\alpha y_i \ln(p_i) + (1 - \alpha)(1 - y_i) \ln(1 - p_i)] \right) - (1 - \beta) \left(\frac{2 \sum_{i=1}^N p_i y_i + s}{\sum_{i=1}^N p_i + \sum_{i=1}^N y_i + s} \right)$
Tversky Loss [۴۱]	$TV = 1 - \frac{\sum_{i=1}^N p_i y_i + s}{\sum_{i=1}^N p_i + \alpha \sum_{i=1}^N p_i (1 - y_i) + \beta \sum_{i=1}^N y_i (1 - p_i) + s}$

جدول (۴): نتایج شبکه پیشنهادی هنگام استفاده از توابع هزینه مختلف، میانگین (حداقل و حداکثر)

تابع هزینه	حساسیت (%)	FP به ازای هر بیمار
Binary Cross Entropy	۷۹/۹۵ (۵۷ و ۱۰۰)	۲۸/۸۲ (۹,۵ و ۱۰۴,۶)
Weighted Binary Cross Entropy ($\alpha = ۰,۸$)	۸۷/۵۵ (۵۷ و ۱۰۰)	۲۸/۸۵ (۱۳,۲۵ و ۶۷,۶)
Dice Loss	۹۲/۹۰ (۶۰ و ۱۰۰)	۲۲/۷۵ (۴ و ۴۳,۲)
Combo Loss ($\alpha = ۰,۸$ و $\beta = ۰,۲$)	۹۲/۶۷ (۶۰ و ۱۰۰)	۲۴/۶۷ (۴,۵ و ۴۷,۴)
Tversky Loss ($\alpha = ۰,۳$ و $\beta = ۰,۷$)	۹۱/۲۴ (۶۰ و ۱۰۰)	۳۱/۸۳ (۱۵,۵ و ۵۲,۴)

۷- نتیجه‌گیری

در این مقاله شبکه Inception 3D U-Net بهبودیافته را برای آشکارسازی توده‌ها در تصاویر اولتراسوند خودکار سه‌بعدی پستان (3D ABUS) پیشنهاد کردیم. برای این شبکه دو نوع ماژول Inception بهبودیافته طراحی کردیم. در ماژول اول که در لایه اول کدگذار استفاده می‌شود، ویژگی‌های سه‌بعدی با دو میدان دید متفاوت بدست می‌آید. در ماژول دوم که در لایه‌های بعدی کدگذار استفاده می‌شود، ویژگی‌های خطی و صفحه‌ای بدست می‌آیند. هدف اصلی این تحقیق بدست آوردن شبکه‌ای با حساسیت تا حد ممکن بالا و میانگین FP قابل قبول است. اکثر توده‌های مجموعه داده مورد استفاده، بسیار کوچک هستند. به منظور آشکارسازی بیشتر این توده‌های کوچک و در نتیجه افزایش حساسیت شبکه، در این مقاله از اندازه تکه $40 \times 40 \times 8$ برای تقسیم کردن حجم اولیه به حجم‌های کوچکتر استفاده کردیم. برای انجام آزمایش‌ها، روش اعتبارسنجی ضربدری 10-fold را در سطح بیمار به کار بردیم. شبکه پیشنهادی به حساسیت ۹۲٫۹٪ با میانگین مثبت‌های کاذب ۲۲٫۷۵ به ازای هر بیمار در آشکارسازی توده‌های تصاویر 3D ABUS دست یافت. یکی از محدودیت‌های روش پیشنهادی، پیدا نکردن توده‌های بسیار ریز است که براساس مشاهدات نویسندگان مقاله، بسیاری از این توده‌ها از دید رادیولوژیست‌ها نیز پنهان می‌مانند. دومین محدودیت، تعداد مثبت‌های کاذب نسبتاً بالا است که برای رفع این مشکل، در حال طراحی شبکه‌ای هستیم که خروجی این شبکه را دریافت کند و ضمن حفظ حساسیت بالا، تعداد مثبت‌های کاذب را کاهش دهد.

مراجع

- [7] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual Review of Biomedical Engineering*, vol. 19, no. 1, pp. 221-248, 2017.
- [8] X. Yao, X. Wang, S.-H. Wang, and Y.-D. Zhang, "A comprehensive survey on convolutional neural network in medical image analysis," *Multimedia Tools and Applications*, pp. 1-45, 2020.
- [9] M. H. Yap et al., "Automated breast ultrasound lesions detection using convolutional neural networks," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 4, pp. 1218-1226, 2018.
- [10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.
- [11] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234-241: Springer.
- [12] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431-3440.
- [13] Y. Wang et al., "3D Inception U-net with asymmetric loss for cancer detection in automated breast ultrasound," *Medical Physics*, 2020.
- [14] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2016, pp. 424-432: Springer.
- [15] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2016.
- [16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2015.
- [17] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700-4708.
- [18] N. Wang et al., "Densely deep supervised networks with threshold loss for cancer detection in automated breast ultrasound," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 641-648: Springer.
- [19] Y. Wang et al., "Deeply-supervised networks with threshold loss for cancer detection in automated breast ultrasound," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2018, pp. 641-648: Springer.
- [1] H. Sung et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209-249, 2021.
- [2] L. Wang, "Early diagnosis of breast cancer," *Sensors*, vol. 17, no. 7, p. 1572, 2017.
- [3] H. J. Shin, H. H. Kim, and J. H. Cha, "Current status of automated breast ultrasonography," *Ultrasonography*, vol. 34, no. 3, p. 165, 2015.
- [4] J. C. van Zelst et al., "Dedicated computer-aided detection software for automated 3D breast ultrasound; an efficient tool for the radiologist in supplemental screening of women with dense breasts," *European Radiology*, vol. 28, no. 7, pp. 2996-3006, 2018.
- [5] Q. Huang, F. Zhang, and X. Li, "Machine learning in ultrasound computer-aided diagnostic systems: a survey," *BioMed Research International*, p. 10, 2018.
- [6] E. Kozegar, M. Soryani, H. Behnam, M. Salamati, and T. Tan, "Computer aided detection in automated 3-D breast ultrasound images: a survey," *Artificial Intelligence Review*, pp. 1-23, 2019.

- [32] Y. Wu and K. He, "Group normalization," in *Proceedings of the European Conference on Computer Vision*, 2018, pp. 3-19.
- [33] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929-1958, 2014.
- [34] E. Kozegar, M. Soryani, H. Behnam, M. Salamati, and T. Tan, "Breast cancer detection in automated 3D breast ultrasound using iso-contours and cascaded RUSBoosts," *Ultrasonics*, vol. 79, pp. 68-80, 2017.
- [35] C. Lo et al., "Multi-dimensional tumor detection in automated whole breast ultrasound using topographic watershed," *IEEE Transactions on Medical Imaging*, vol. 33, no. 7, pp. 1503-1511, 2014.
- [36] T. Tan, B. Platel, R. Mus, L. Tabár, R. M. Mann, and N. Karssemeijer, "Computer-aided detection of cancer in automated 3-D breast ultrasound," *IEEE Transactions on Medical Imaging*, vol. 32, no. 9, pp. 1698-1706, 2013.
- [37] M. Abadi et al., "TensorFlow: a system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp. 265-283.
- [38] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [39] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image segmentation," in *Fourth International Conference on 3D Vision*, 2016, pp. 565-571: IEEE.
- [40] S. A. Taghanaki et al., "Combo loss: handling input and output imbalance in multi-organ segmentation," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 24-33, 2019.
- [41] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," in *International Workshop on Machine Learning in Medical Imaging*, 2017, pp. 379-387: Springer.
- ultrasound," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 866-876, 2019.
- [20] F. Wang et al., "Spatial attention lesion detection on automated breast ultrasound," in *International Work-Conference on Bioinformatics and Biomedical Engineering*, 2019, pp. 216-227: Springer.
- [21] F. Wang et al., "Study on automatic detection and classification of breast nodule using deep convolutional neural network system," *Journal of Thoracic Disease*, vol. 12, no. 9, p. 4690, 2020.
- [22] P. Pan, H. Chen, Y. Li, N. Cai, L. Cheng, and S. Wang, "Tumor segmentation in automated whole breast ultrasound using bidirectional LSTM neural network and attention mechanism," *Ultrasonics*, vol. 110, p. 106271, 2021.
- [23] H. Qassim, A. Verma, and D. Feinzimer, "Compressed residual-VGG16 CNN model for big data places image recognition," in *IEEE 8th Annual Computing and Communication Workshop and Conference*, 2018, pp. 169-175.
- [24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097-1105, 2012.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778.
- [26] J. Yu et al., "Computer-aided tumor diagnosis in automated breast ultrasound using 3D detection network," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2020, pp. 181-189: Springer.
- [27] T.-C. Chiang, Y.-S. Huang, R.-T. Chen, C.-S. Huang, and R.-F. Chang, "Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 240-249, 2018.
- [28] W. K. Moon et al., "Computer-aided tumor detection in automated breast ultrasound using a 3-D convolutional neural network," *Computer Methods and Programs in Biomedicine*, vol. 190, p. 105360, 2020.
- [29] C. Szegedy et al., "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1-9.
- [30] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning*, 2015.
- [31] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Improved texture networks: maximizing quality and diversity in feed-forward stylization and texture synthesis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 6924-6932.



سپیده برکت رضایی از سال ۱۳۸۶ تا ۱۳۹۰ در رشته مهندسی کامپیوتر گرایش نرم‌افزار دانشگاه خوارزمی مشغول به تحصیل بود. سپس در سال ۱۳۹۰ به عنوان استعداد برتر به طور مستقیم وارد مقطع کارشناسی ارشد در گرایش هوش مصنوعی همان دانشگاه شد و در سال ۱۳۹۲ فارغ‌التحصیل گردید. ایشان در سال

۱۳۹۵ در مقطع دکتری رشته مهندسی کامپیوتر گرایش هوش مصنوعی دانشگاه علم و صنعت ایران پذیرفته شد و هم‌اکنون در حال تحصیل است. زمینه‌های پژوهشی مورد علاقه ایشان بینایی ماشین، پردازش تصویر و یادگیری عمیق است.



امین ملک محمدی فارغ‌التحصیل رشته مهندسی کامپیوتر، گرایش فناوری اطلاعات از دانشگاه تبریز و کارشناسی ارشد مهندسی کامپیوتر، گرایش هوش مصنوعی از دانشگاه علم و صنعت ایران است. زمینه‌های پژوهشی مورد علاقه ایشان بینایی ماشین، پردازش تصاویر پزشکی و تشریح‌پذیری مدل‌های عمیق است.



احسان‌اله کوزه‌گر در سال ۱۳۸۳ در رشته مهندسی کامپیوتر-نرم‌افزار دانشگاه علم و صنعت ایران (واحد بهشهر) پذیرفته شد. سپس، در سال ۱۳۸۸، در مقطع کارشناسی ارشد رشته مهندسی کامپیوتر-گرایش هوش مصنوعی در دانشگاه علم و صنعت مشغول به تحصیل گردید و در سال ۱۳۹۰ فارغ‌التحصیل شد. پس از آن، در سال ۱۳۹۱ به عنوان استعداد برتر دانشگاه علم و صنعت به طور مستقیم وارد مقطع دکتری در گرایش هوش مصنوعی شده و در سال ۱۳۹۷ از رساله خود با درجه عالی دفاع کرد. از سال ۱۳۹۷ تاکنون به عنوان عضو هیات علمی دانشکده فنی و مهندسی شرق گیلان در دانشگاه گیلان مشغول به کار است. از زمینه‌های تحقیقاتی مورد علاقه ایشان می‌توان به یادگیری ماشین و پردازش تصویر اشاره کرد.



معصومه سلامتی فارغ‌التحصیل پزشک عمومی دانشگاه علوم پزشکی آزاد اسلامی در سال ۱۳۷۳ و فارغ‌التحصیل پرستاری تشخیصی از دانشگاه علوم پزشکی ایران در سال ۱۳۷۹ است. ایشان دوره فوق تخصصی (فلوشیپ) پستان و دوره فوق تخصصی سی‌تی اسکن را نیز گذرانده‌اند. همچنین سالهاست که در مراکز مهمی به عنوان متخصص رادیولوژی مشغول به کار هستند.



محسن سربانی در سال ۱۳۵۹ مدرک کارشناسی خود را در رشته مهندسی برق - الکترونیک از دانشگاه علم و صنعت ایران اخذ کرد. او سپس به ترتیب در سال‌های ۱۳۶۶ و ۱۳۶۹ مدارک کارشناسی ارشد و دکتری خود را از دانشکده مهندسی برق و کامپیوتر دانشگاه هریوت-وات در شهر ادینبورو در اسکاتلند با تخصص‌های "فنون دیجیتال" و "پردازش تصویر" دریافت نمود. در سال ۱۳۶۹ و پس از فراغت از تحصیل، در دانشکده فنی دانشگاه مازندران به عنوان عضو هیات علمی مشغول به کار گردید. در سال ۱۳۸۱ به گروه هوش مصنوعی و رباتیک دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت ایران پیوست و در حال حاضر در این گروه با مرتبه دانشیاری مشغول به کار است. زمینه‌های پژوهشی مورد علاقه وی شامل پردازش و تحلیل تصاویر، بینایی ماشین و کاربردهای هوش مصنوعی در این مباحث است.