

به نام خداوند متعال

دوره یادگیری عمیق

عنوان مطالعه: خودکار سازی فرآیند LV Segmentation

استاد: خانم مهندس زهرا امینی

ارائه کننده: عبدالرحمان نورزاد

(بورسیه دوره Deep Learning مجموعه هوش مصنوعی – فیلاگر)

۳۰ شهریور ۱۴۰۳

مقاله اول:

Automated Left Ventricle Segmentation in Echocardiography using YOLO: A Deep Learning Approach for Enhanced Cardiac Function Assessment

MadanKumar Balasubramani, Chih-Wei Sung, Mu-Yang Hsieh, Edward Pei-Chuan Huang *,
Jiann-Shing Shieh *, Maysam Abbod

Posted Date: 21 May 2024

doi: 10.20944/preprints202405.1196.v1

چکیده: این مطالعه به اهمیت تقسیم‌بندی دقیق بطن چپ (LV) در تصاویر اکوکاردیوگرام (Echo) پرداخته و به بررسی استفاده از مدل YOLO (You Only Look Once) برای این کار می‌پردازد. روش‌های سنتی تقسیم‌بندی معمولاً زمان‌بر و وابسته به ارزیابی‌های فردی هستند. یادگیری عمیق به عنوان ابزاری کارآمد برای تقسیم‌بندی خودکار تصاویر پزشکی شناخته می‌شود و می‌تواند در دقت و سرعت بهبودهایی ارائه دهد. YOLO به دلیل معماری پیشرفته‌اش، تعادل خوبی بین سرعت و دقت فراهم می‌کند و با استفاده از لایه‌های کانولوشن دیلاته و بلوک‌های bottleneck، ویژگی‌ها را استخراج می‌کند. این مطالعه تأکید دارد که با تنظیم دقیق یک مدل YOLO پیش‌آموزش‌دیده بر روی یک دیتاست معتبر، می‌توان به تقسیم‌بندی مؤثر LV دست یافت. نتایج نشان‌دهنده عملکرد عالی مدل با میانگین دقت در یک آستانه IoU ۵۰٪ برابر ۹۸٫۳۱٪ بود و میانگین دقت ۷۵٫۲۷٪ در آستانه‌های مختلف IoU است. این پیشرفت می‌تواند به تسریع و استانداردسازی تحلیل تصاویر Echo کمک کند و در نتیجه به تصمیم‌گیری بالینی و مراقبت بهتر از بیماران منجر شود.

جدول جمع‌بندی چکیده

عنوان	خلاصه
موضوع	تقسیم‌بندی خودکار LV در تصاویر Echo
روش	استفاده از مدل YOLO برای تقسیم‌بندی
مزایا	سرعت و دقت بالا با استفاده از یادگیری عمیق
نتایج	mAP50: 98.31% و mAP50: 95: 75.27%
تأثیر	تسریع در تحلیل و تصمیم‌گیری بالینی

۱. مقدمه: بیماری‌های قلبی-عروقی (CVD) به عنوان یکی از اصلی‌ترین علل مرگ و میر در جهان شناخته می‌شوند و سالانه حدود ۱۷,۹ میلیون نفر را تحت تأثیر قرار می‌دهند. ارزیابی دقیق عملکرد قلبی برای تشخیص به‌موقع، طبقه‌بندی ریسک و راهنمایی تصمیمات درمانی در بیماران مبتلا به CVD ضروری است. اکوکاردیوگرافی به عنوان یک روش تصویربرداری غیرتهاجمی، نقش مهمی در ارزیابی قلبی ایفا می‌کند. تقسیم‌بندی بطن چپ (LV) در تصاویر اکو، فرآیند ترسیم مرزهای حفره LV است که برای اندازه‌گیری پارامترهای قلبی مختلف، از جمله کسری خروجی بطن چپ (LVEF)، ضروری است. تقسیم‌بندی LV به طور سنتی توسط متخصصان قلب به صورت دستی انجام می‌شود که این فرآیند زمان‌بر و مستعد خطاهای بین ناظران است. محدودیت‌های این روش دستی باعث شده تا تحقیقات به سمت تکنیک‌های خودکار تقسیم‌بندی با استفاده از یادگیری ماشین و به‌ویژه یادگیری عمیق متمایل شود. شبکه‌های عصبی کانولوشنی (CNNs) در تحلیل تصاویر پزشکی، از جمله تقسیم‌بندی LV، موفقیت‌های قابل توجهی را نشان داده‌اند.

با وجود پیشرفت‌های انجام‌شده در تقسیم‌بندی LV با استفاده از یادگیری عمیق، چالش‌هایی نیز وجود دارد. تصاویر اکوکاردیوگرافی دارای تنوع قابل توجهی در کیفیت تصویر، زاویه‌های دریافت تصویر و ویژگی‌های بیماران هستند که این موارد موجب چالش در عمومیت مدل‌ها می‌شود. همچنین، عدم توازن کلاس در تصاویر می‌تواند منجر به اولویت دادن مدل‌ها به تقسیم‌بندی پس‌زمینه به جای delineation دقیق LV شود.

هدف این تحقیق بررسی پتانسیل YOLOv8 برای تقسیم‌بندی خودکار LV در اکوکاردیوگرافی است. فرضیه ما این است که با بهره‌گیری از نقاط قوت YOLOv8، می‌توان مدلی توسعه داد که دارای مزایای زیر باشد:

- دقت بهبود یافته: دستیابی به مقادیر بالای ضریب تشابه دایس (DSC) و ضریب IoU.
- عمومیت بهتر: نمایش عملکرد قوی در برابر تصاویر اکوکاردیوگرافی متنوع.
- کارایی محاسباتی: حفظ زمان‌های استنتاج سریع‌تر نسبت به مدل‌های سنتی مبتنی بر CNN.

جدول جمع‌بندی مقدمه

بخش	خلاصه
موضوع	تقسیم‌بندی خودکار LV در اکوکاردیوگرافی
چالش‌ها	تنوع تصاویر، عدم توازن کلاس، زمان‌بر بودن روش‌های دستی
هدف	بررسی پتانسیل YOLOv8 در تقسیم‌بندی خودکار LV
مزایا	دقت بالا، عمومیت بهتر، کارایی محاسباتی

۲. کارهای مرتبط

اکوکاردیوگرام‌ها به عنوان تصاویر اولتراسوند قلب، نقشی اساسی در تشخیص و مدیریت بیماری‌های قلبی دارند. تحلیل این تصاویر به‌طور سنتی به تقسیم‌بندی دستی ساختارهای آناتومیکی وابسته بود که فرآیندی زمان‌بر و مستعد خطا بود. با این حال، ظهور یادگیری عمیق، دوره‌ای جدید از تقسیم‌بندی خودکار را به ارمغان آورده است که می‌تواند نحوه تحلیل اکوکاردیوگرام‌ها را متحول کند.

پیشرفت‌های اولیه: روش‌های ابتدایی یادگیری عمیق از معماری‌هایی نظیر **U-Net** استفاده کردند که پایه‌گذار تقسیم‌بندی خودکار ساختارهای قلبی از جمله **LV** بودند. این مدل‌ها نتایج امیدوارکننده‌ای به دست آوردند، اما در **Capturing Variations** موجود در آناتومی اکوکاردیوگرام‌ها با محدودیت‌هایی مواجه بودند.

توسعه و بهینه‌سازی: تحقیقات به‌طور مداوم این تکنیک‌ها را با اضافه کردن مکانیسم‌های توجه و استراتژی‌های افزایش داده مانند چرخش و مقیاس‌بندی بهبود دادند. همچنین، پیشرفت‌ها در معماری‌های یادگیری عمیق، مانند اتصالات باقی‌مانده و بلوک‌های متراکم، کارایی آموزش را بهبود بخشید و ممکن است به عملکرد بهتر تقسیم‌بندی منجر شود.

جایگاه فعلی: در حال حاضر، با در دسترس بودن داده‌های اکوکاردیوگرافی **3D**، محققان در حال بررسی **CNN**های سه‌بعدی و روش‌های تقسیم‌بندی حجمی هستند تا روابط فضایی بین ساختارها را به‌طور مؤثرتری درک کنند. تکنیک‌های یادگیری ضعیف نیز به‌عنوان روشی برای کاهش بار انطباق با استفاده از نشانه‌های کمتر دشوار مورد توجه قرار گرفته‌اند.

چالش‌ها و موانع: با وجود پتانسیل عظیم یادگیری عمیق، چالش‌هایی نظیر تنوع کیفیت تصاویر اکوکاردیوگرام و نیاز به داده‌های بزرگ و با کیفیت با برچسب پیکسل-محور وجود دارد. همچنین، تفسیر و توضیح نتایج مدل‌های یادگیری عمیق برای اعتماد به کاربردهای بالینی ضروری است.

مزایا: توسعه‌ی تقسیم‌بندی خودکار می‌تواند به متخصصان بهداشت کمک کند تا به‌طور مؤثرتری حجم‌های آناتومیکی را کمی‌سازی کرده و عملکرد قلب را ارزیابی کنند. این می‌تواند منجر به بهبود دقت تشخیصی بیماری‌های قلبی شود و همچنین نتایج تقسیم‌بندی می‌تواند راهنمایی برای روش‌های درمانی کم‌تهاجمی باشد.

آینده‌ی یادگیری عمیق در تحلیل اکوکاردیوگرام: یادگیری عمیق به عنوان یک ابزار انقلابی برای تحلیل اکوکاردیوگرام، پتانسیل تبدیل تحلیل اکوکاردیوگرام‌ها به روشی کارآمدتر و بهبود تصمیم‌گیری بالینی را دارد. بهبود مداوم در این زمینه به نتایج بهتری برای بیماران منجر خواهد شد.

جدول جمع‌بندی کارهای مرتبط

بخش	خلاصه
موضوع	تقسیم‌بندی خودکار LV در اکوکاردیوگرافی
پیشرفت‌ها	استفاده از U-Net ، اضافه کردن مکانیسم‌های Attention و استراتژی‌های افزایش داده
چالش‌ها	تنوع کیفیت تصاویر، نیاز به داده‌های با کیفیت و تفسیر نتایج
مزایا	دقت بالاتر، ارزیابی مؤثر عملکرد قلب، هدایت درمان‌های کم‌تهاجمی
آینده	پیشرفت در یادگیری عمیق و بهبود تحلیل اکوکاردیوگرام‌ها

۳. مواد و روش‌ها

۳-۱. دیتاست: مجموعه داده‌های این مطالعه شامل تصاویر اکوکاردیوگرافی بیماران مراجعه‌کننده به اورژانس در بیمارستان‌های مختلف بود. متخصصان قلب با دقت مرزهای LV را در این تصاویر مشخص کردند تا به‌عنوان برچسب‌های درست (ground truth) برای آموزش و ارزیابی مدل‌های یادگیری عمیق برای تقسیم‌بندی LV استفاده شوند.

در مجموع ۴۷۸۱ مورد تصویر اکوکاردیوگرافی انتخاب و به سه زیرمجموعه تقسیم شد: ۳۸۷۷ مورد برای آموزش، ۴۷۷ مورد برای اعتبارسنجی و ۴۲۷ مورد برای آزمایش. این تقسیم‌بندی با استفاده از روش نمونه‌برداری تصادفی (random sampling approach) انجام شد تا توزیع عادلانه‌ای در داده‌های آموزش، اعتبارسنجی و آزمایش ایجاد شود.

تصاویر با استفاده از روش‌های مختلف اکوکاردیوگرافی (2D, 3D, stress echo) و پلن‌های تصویربرداری به دست آمدند که نشان‌دهنده تنوع داده‌ها در شرایط بالینی واقعی است. این مجموعه داده عمده شامل مواردی با کیفیت تصویر متفاوت، سطح نویز، آثار و چالش‌های ذاتی مرتبط با اکوکاردیوگرافی اورژانسی بود.

برای هر مورد، متخصصان قلب به‌صورت دستی مرزهای اندوکاردیوم حفره LV را در طول چرخه قلبی کامل مشخص کردند و ماسک‌های دقیق تقسیم‌بندی در سطح نمونه تولید کردند.

۲-۳. معماری YOLOv8: توسط Ultralytics توسعه یافته و از یک شبکه عصبی کانولوشنی تشکیل شده است که به دو قسمت اصلی تقسیم می‌شود: backbone و head.

ستون فقرات YOLOv8، یک نسخه تغییر یافته از معماری CSP Darknet53 است و شامل ۵۳ لایه کانولوشن است.

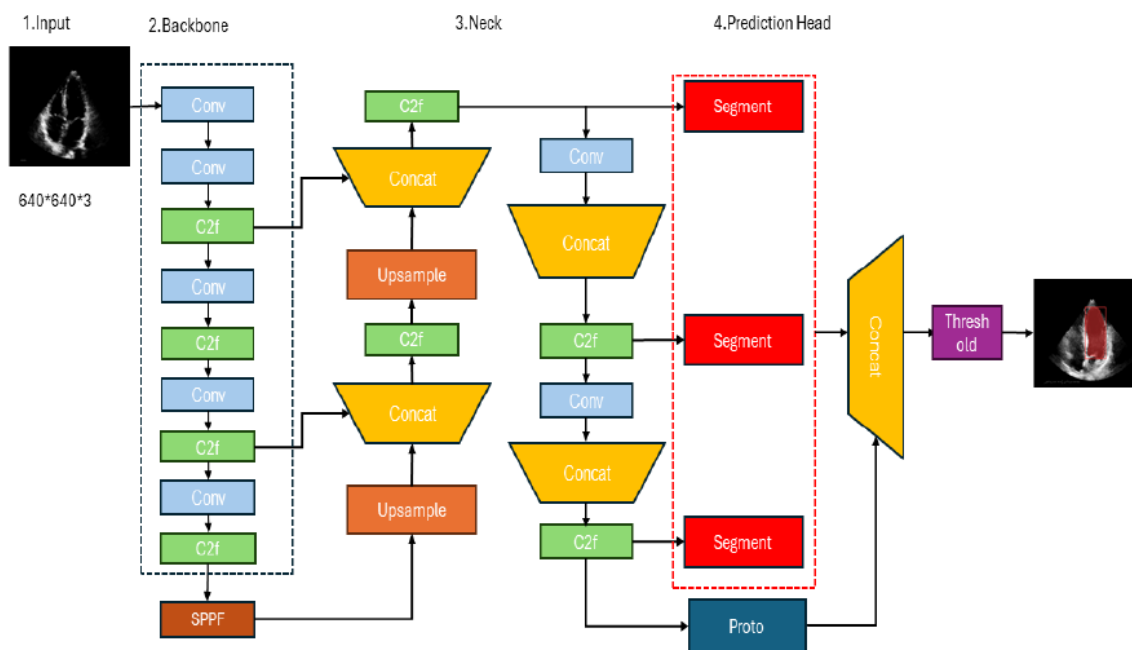


Figure 1. Architecture of YOLOv8.

Backbone به استخراج ویژگی‌ها از تصویر ورودی می‌پردازد و یک نقشه ویژگی تولید می‌کند.

Neck یا شبکه هرم ویژگی (FPN) مسئول ترکیب ویژگی‌ها از مقیاس‌های مختلف خروجی Backbone است و به مدل کمک می‌کند تا اطلاعات با وضوح بالا و پایین را به دست آورد.

Head آخرین جزء مدل تقسیم‌بندی YOLO است که مسئول تولید ماسک‌های تقسیم‌بندی است. این قسمت از ویژگی‌های ترکیبی Neck استفاده می‌کند و با اعمال لایه‌های کانولوشن، خروجی نهایی تقسیم‌بندی را تولید می‌کند.

۳-۳. معماری پیشنهادی: معماری پیشنهادی برای تقسیم‌بندی نمونه شامل چندین ماژول و تکنیک است که شامل موارد زیر است:

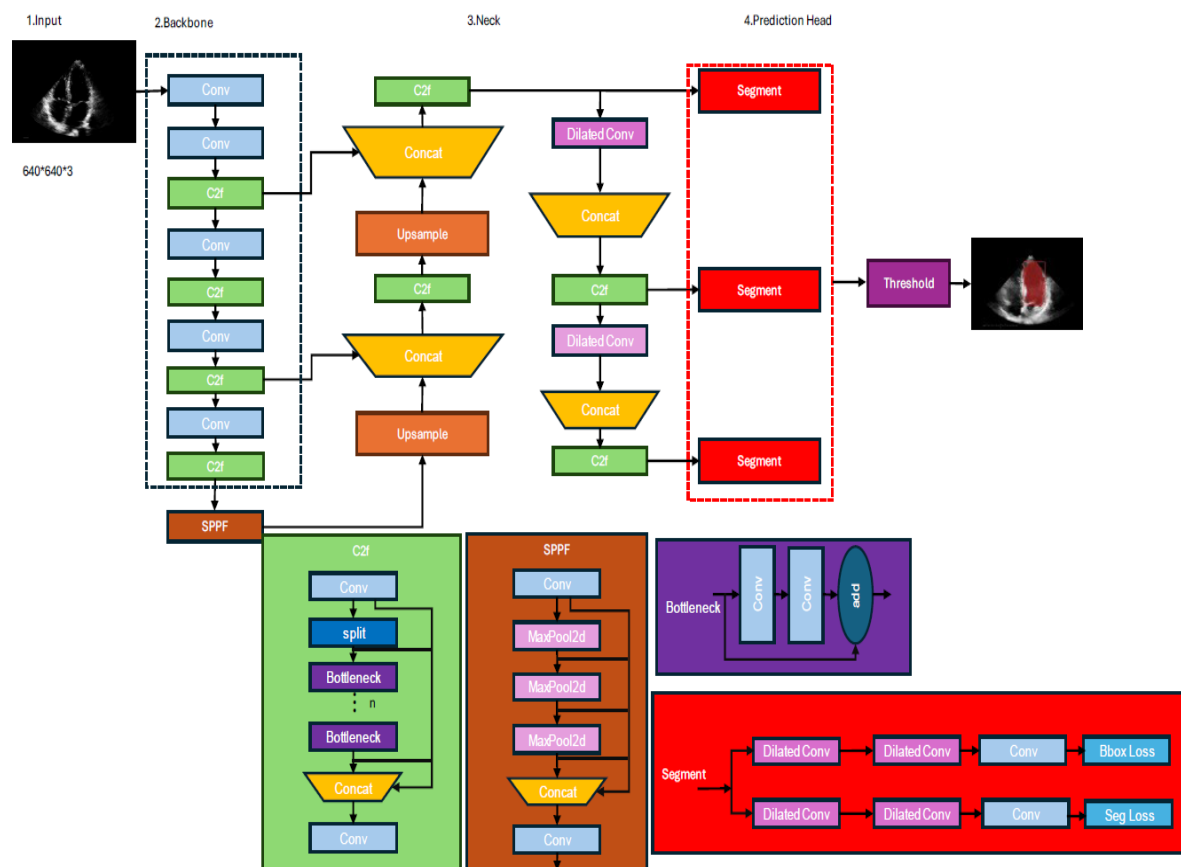


Figure 2. Improved Architecture of YOLOv8 for left ventricle segmentation.

۳-۳-۱. کانولوشن گشاد (Dilated Convolution): این تکنیک به شبکه اجازه می‌دهد تا زمینه‌های وسیع‌تری را بدون افزایش پیچیدگی محاسباتی یا تعداد پارامترهای قابل یادگیری درک کند.

Dilated convolution is an advanced technique employed in convolutional neural networks (CNNs) for tasks such as image segmentation and object detection.

۳-۳-۲. ماژول (Class-to-Fortitude) C2F: این ماژول توانایی مدل در تمایز بین کلاس‌های مختلف را بهبود می‌بخشد و ویژگی‌ها را برای بهبود عملکرد تقسیم‌بندی ترکیب می‌کند.

۳-۳-۳. ماژول **SPPF (Spatial Pyramid Pooling Fortitude)**: این ماژول به منظور ثبت ویژگی‌های چند مقیاسی طراحی شده است و به مدل کمک می‌کند تا اشیای مختلف را در مقیاس‌های متفاوت شناسایی کند.

۳-۳-۴. **ماژول تقسیم‌بندی**: این ماژول مسئول تولید ماسک‌های تقسیم‌بندی نمونه از نقشه‌های ویژگی بهبود یافته است و از لایه‌های کانولوشن موازی استفاده می‌کند.

تابع هزینه: برای وظایف تقسیم‌بندی تصویر پزشکی، ترکیب **Dice Loss** و **Robust TLoss** به عنوان تابع هزینه استفاده می‌شود تا هم به چالش‌های مقادیر پرت و هم به بهینه‌سازی همپوشانی بین ماسک‌های پیش‌بینی شده و واقعی (**ground truth**) پرداخته شود.

با ترکیب این دو اصطلاح هزینه، تابع کلی هدفمند است تا تعادل بین مدیریت مقادیر پرت و بهینه‌سازی همپوشانی را ایجاد کند.

کاربردهای بالینی: توانایی **YOLOv8** در انجام تشخیص‌ها به طور زمان واقعی برای کاربردهای بالینی بسیار مهم است، جایی که نتایج سریع تقسیم‌بندی برای تشخیص و تصمیم‌گیری‌های درمانی ضروری است.

جدول جمع‌بندی مواد و روش‌شناسی

عنوان	خلاصه
دیتاست	تصاویر اکوکاردیوگرافی با مرزهای LV مشخص شده توسط متخصصان قلب
معماری YOLOv8	Backbone: CSP Darknet53 Neck: FPN Head: تولید ماسک‌های تقسیم‌بندی
معماری پیشنهادی	شامل کانولوشن Dilated، ماژول C2F، ماژول SPPF و ماژول تقسیم‌بندی
تابع هزینه	ترکیب Dice Loss و Robust TLoss برای بهینه‌سازی تقسیم‌بندی
کاربردهای بالینی	سرعت بخشیدن به فرآیند تشخیص و تصمیم‌گیری درمانی

۳-۴. معیارهای ارزیابی

۳-۴-۱. معیار IoU

همپوشانی بر اتحادیه (IoU) یک معیار کلاسیک برای ارزیابی عملکرد مدل در تشخیص اشیاء است. این معیار نسبت همپوشانی و اتحادیه بین جعبه محدود پیش‌بینی شده‌ی ماسک و جعبه محدود واقعی را محاسبه می‌کند.

IoU به صورت زیر تعریف می‌شود:

$$IoU(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

که در آن A نمایانگر ناحیه پیش‌بینی شده‌ی ماسک و B نمایانگر ناحیه واقعی حاوی شیء است. هرچه مقدار IoU بیشتر باشد، عملکرد مدل بهتر است و مقادیر بالاتر نشان‌دهنده‌ی تفاوت کمتر بین جعبه‌های پیش‌بینی شده و واقعی است.

۳-۴-۲. دقت میانگین (mAP)

دقت میانگین (mAP) به هر دو مؤلفه Precision (اشیاء شناسایی شده صحیح) و Recall (شناسایی همه مثبت‌های واقعی) در آستانه‌های مختلف IoU توجه می‌کند. این معیار تصویری جامع از عملکرد تشخیص مدل ارائه می‌دهد.

mAP@50 به‌طور خاص به mAP محاسبه شده با آستانه IoU برابر ۰,۵ اشاره دارد، به این معنا که دقت‌های تشخیص جایی که جعبه پیش‌بینی شده یا ماسک تقسیم‌بندی حداقل ۵۰٪ با واقعیت (ground truth) همپوشانی دارد را نمایش می‌دهد.

mAP@0.5:0.95 به میانگین دقت در بازه‌ای از آستانه‌های IoU اشاره دارد که معمولاً از ۰,۵ تا ۰,۹۵ با افزایش ۰,۰۵ تعریف می‌شود. این معیار تحلیل عمیق‌تری از عملکرد مدل تحت درجات مختلف همپوشانی بین پیش‌بینی‌ها و واقعیت فراهم می‌آورد. میانگین بالای mAP در این بازه نشان‌دهنده عملکرد خوب مدل حتی با همپوشانی‌های کمتر از حد ایده‌آل است.

۳-۴-۳. منحنی Precision-recall (P-R Curve)

P-R Curve منحنی‌ای است که recall را به‌عنوان محور x و precision را به‌عنوان محور y ترسیم می‌کند. هر نقطه نمایانگر یک مقدار آستانه متفاوت است و تمامی نقاط به‌صورت یک منحنی متصل می‌شوند. مقادیر precision و recall بر اساس معادلات زیر محاسبه می‌شوند:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

که در آن:

- **TP (مثبت صحیح):** پیش‌بینی‌ای که به‌عنوان کلاس مثبت قضاوت شده و صحیح است.
- **FP (مثبت نادرست):** پیش‌بینی‌ای که به‌عنوان کلاس مثبت قضاوت شده ولی نادرست است.
- **FN (منفی نادرست):** پیش‌بینی‌ای که به‌عنوان کلاس منفی قضاوت شده ولی نادرست است.

این معیارها به ارزیابی دقیق‌تری از عملکرد مدل در تشخیص و تقسیم‌بندی اشیاء کمک می‌کنند.

۴. نتایج و بحث‌ها

در این مطالعه، آموزش و تست شبکه بر روی یک ایستگاه کاری با مشخصات زیر انجام شد:

Intel Core i7-11700K @3.6GHZ

NVIDIA GeForce RTX 4090 24G GPU

Windows 11 operating system

در فرآیند آموزش، از مدل YOLOv8 پیش‌آموزش‌دیده (pre-trained YOLOv8 model) استفاده شده است که از مجموعه داده‌ی MS COCO (Microsoft Common Objects in Context) val2017 به دست آمده است.

۴-۱. مقایسه عملکرد مدل‌های مختلف YOLO

جدول ۱ به مقایسه پنج مدل YOLOv8 (نشان‌گذاری شده با n، s، m، l و x) از نظر Precision، Recall، mAP50 و mAP50-95 می‌پردازد. این مدل‌ها به‌طور خاص بر توانایی شناسایی بطن چپ (LV) در تصاویر پزشکی ارزیابی شده‌اند. Precision و Recall مؤثر بودن مدل در طبقه‌بندی صحیح پیکسل‌ها را اندازه‌گیری می‌کند، در حالی که mAP50 و mAP50-95 دقت کلی تشخیص را با آستانه‌های مختلف هم‌پوشانی (IoU) ارزیابی می‌کنند.

Table 1. Performance comparison of various YOLO models.

Model	size (pixels)	Precision	Recall	mAP50	mAP50-95	params (M)	FLOPs (B)
YOLOv8n-seg	416	0.97247	0.95840	0.96064	0.75742	3.4	12.6
YOLOv8s-seg	416	0.97306	0.96771	0.97887	0.75604	11.8	42.6
YOLOv8m-seg	416	0.97363	0.97692	0.97957	0.75818	27.3	110.2
YOLOv8l-seg	416	0.97338	0.97899	0.97964	0.75626	46	220.5
YOLOv8x-seg	416	0.97572	0.97907	0.98005	0.75784	71.8	344.1
YOLOv8n-seg	640	0.97448	0.97456	0.97973	0.75875	3.4	12.6
YOLOv8s-seg	640	0.97651	0.97571	0.98164	0.76066	11.8	42.6
YOLOv8m-seg	640	0.9768	0.97894	0.98271	0.75816	27.3	110.2
YOLOv8l-seg	640	0.97583	0.97770	0.98263	0.75821	46	220.5
YOLOv8x-seg	640	0.97654	0.97921	0.98269	0.75852	71.8	344.1
YOLOv8n-seg	1280	0.97651	0.97907	0.98154	0.75671	3.4	12.6
YOLOv8s-seg	1280	0.97654	0.97907	0.97932	0.75164	11.8	42.6
YOLOv8m-seg	1280	0.97657	0.97907	0.98108	0.75491	27.3	110.2
YOLOv8l-seg	1280	0.9766	0.97907	0.98126	0.75542	46	220.5
YOLOv8x-seg	1280	0.97661	0.97907	0.98071	0.75409	71.8	344.1

Table 1 compares five YOLOv8 models (likely denoted by si-e - n, s, m, l-x - signifying increasing complexity) across various metrics. These models are evaluated on their ability to identify the LV in medical images. Precision and Recall measure the model's effectiveness in correctly classifying pixels, while mAP50 and mAP50-95 assess overall detection accuracy with varying strictness (Intersection over Union thresholds).

۴-۱-۱. تجزیه و تحلیل جدول ۱

مدل‌های بزرگتر با پارامترهای بیشتر (l و x) معمولاً نمرات mAP بالاتری را نشان می‌دهند که نشان‌دهنده دقت بهتر در تقسیم‌بندی است. این در حالی است که این افزایش دقت به قیمت افزایش نیاز محاسباتی با FLOPs بالاتر انجام می‌شود.

همچنین با بررسی تأثیر اندازه تصویر ورودی، مشاهده می‌شود که در برخی مدل‌ها، افزایش وضوح تصویر می‌تواند به بهبود جزئی در دقت تقسیم‌بندی منجر شود. تصاویر با وضوح بالاتر اطلاعات بیشتری از ساختار LV فراهم می‌آورند.

۴-۱-۲: نتیجه‌گیری کلیدی

تبادل بین پیچیدگی مدل و کارایی محاسباتی نکته مهمی است. در حالی که مدل‌های بزرگتر در mAP بهتر عمل می‌کنند، هزینه بالای محاسباتی آن‌ها را برای کاربردهای زمان واقعی یا دستگاه‌های با منابع محدود نامناسب می‌کند. به عنوان مثال، مدل YOLOv8x-seg بالاترین نمرات mAP را دارد اما نیاز به بیش از ۷۱ میلیون پارامتر و ۳۴۴ میلیارد FLOPs دارد. در مقابل YOLOv8n-seg با ۳,۴ میلیون پارامتر و ۱۲,۶ میلیارد FLOPs، راه حل سریع‌تری را ارائه می‌دهد، هرچند با کاهش جزئی در دقت تقسیم‌بندی.

۴-۲. مقایسه با توابع هزینه مختلف

Table 2. Comparison of proposed model with various loss functions.

Model	Precision	Recall	mAP50	mAP50-95
Yolov8	0.93532	0.94171	0.95777	0.62485
RobustTLoss [53]	0.95284	0.95581	0.96487	0.61742
DiceLoss [54]	0.94971	0.93488	0.96788	0.60437
Ours	0.98359	0.97561	0.98724	0.75876

جدول ۲ به مقایسه چهار مدل با توابع هزینه متفاوت می‌پردازد. مدل پیشنهادی با نمرات فوق‌العاده‌ای در precision (0.98359)، recall (0.97561)، mAP50 (0.98724) و mAP50-95 (0.75876) بهترین عملکرد را نشان می‌دهد. این عملکرد عالی به ویژه در mAP50-95 مشهود است، که نشان‌دهنده قابلیت‌های قوی و پایدار تشخیص اشیاء در سناریوهای مختلف است.

۴-۲-۱. اهمیت مدل پیشنهادی

مدل پیشنهادی به خوبی تعادل بین **precision** و **recall** را حفظ کرده و به یک معیار جدید برای عملکرد تشخیص اشیاء تبدیل شده است. عملکرد مستمر و بالای این مدل در مقادیر **mAP50** و **mAP50-95** نشان دهنده اعتبار و دقت این روش برای تقسیم بندی LV است.

۴-۳. تحلیل ماتریس ابهام (Confusion matrix)

پس از آموزش، ماتریس ابهام با تحلیل نتایج تولید شد. این تحلیل نشان می دهد که مدل برای کلاس مثبت عملکرد خوبی دارد، اما یک نمونه به عنوان مثبت نادرست و یک نمونه به عنوان منفی نادرست شناسایی شده است. این نشان دهنده قابلیت های خوب مدل است، اما همچنین نیاز به بهینه سازی بیشتر برای کاهش اشتباهات شناسایی دارد.

۴-۳-۱. تأثیر کیفیت داده های آموزشی

کیفیت و کمیت داده های آموزشی تأثیر قابل توجهی بر عملکرد مدل دارند. بهینه سازی بیشتر و اعتبارسنجی الگوریتم های تقسیم بندی خودکار برای دستیابی به نتایج بالینی قابل اعتماد ضروری است.

۴-۴. ارزیابی های بصری

تصاویر پیش بینی شده از تقسیم بندی بطن چپ در شکل ۷ به وضوح نشان می دهند که مدل پیشنهادی در تشخیص اشکال و سطوح مختلف موفق عمل کرده است. تقسیم بندی دقیق و مستحکم LV از تصاویر اکو به ویژه در کاربردهای بالینی نظیر ارزیابی عملکرد قلب و نظارت بر پیشرفت بیماری حیاتی است.

در نهایت، بهینه سازی مستمر مدل و اعتبارسنجی آن در شرایط بالینی واقعی می تواند به بهبود قابل توجهی در نتایج بالینی منجر شود.

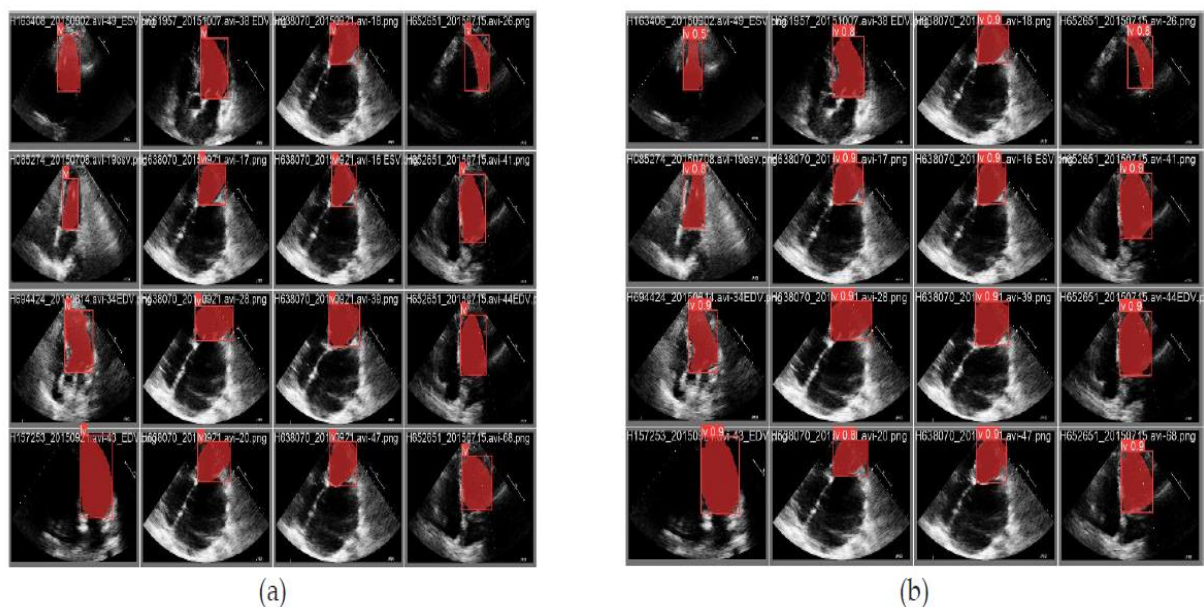


Figure 7. Predicted Left ventricle segmentation. (a) Ground Truth annotated by the experts (b) Segmentation of frames by model.

۵. نتیجه گیری

تقسیم‌بندی دقیق و قابل اعتماد بطن چپ از تصاویر اکوکاردیوگرام یک وظیفه حیاتی در تصویربرداری و تشخیص قلبی-عروقی است. این مطالعه نشان‌دهنده پتانسیل مدل‌های تقسیم‌بندی خودکار در به تصویر کشیدن شکل و موقعیت کلی حفره بطن چپ با بالاترین نمره mAP50 معادل ۹۸,۳۱٪ و mAP50:95 معادل ۷۵,۲۷٪ است. با این حال، برخی از موارد ناهماهنگی یا عدم دقت نشان‌دهنده نیاز به بهینه‌سازی و اعتبارسنجی بیشتر این الگوریتم‌ها است.

تعیین دقیق مرزهای بطن چپ برای ارزیابی‌های کمی مانند اندازه‌گیری حجم‌ها (measuring volumes)، کسری خروج (ejection fraction)، و ناهنجاری‌های حرکتی دیواره منطقه‌ای (regional wall motion abnormalities) ضروری است. این اندازه‌گیری‌ها نقش حیاتی در تشخیص و نظارت بر وضعیت‌های مختلف قلبی-عروقی، از جمله نارسایی قلبی، بیماری‌های دریچه‌ای و انفارکتوس میوکارد ایفا می‌کند.

۵-۱. زمینه‌های تحقیقاتی آینده

برای پیشرفت در حوزه تقسیم‌بندی اکوکاردیوگرام بطن چپ، چندین حوزه از کارهای آینده می‌تواند مورد بررسی قرار گیرد:

- **داده‌های بزرگ‌تر و متنوع‌تر:** آموزش مدل‌های تقسیم‌بندی با استفاده از داده‌های بزرگ‌تر و متنوع‌تر، که شامل پاتولوژی‌های مختلف، مدالیته‌های تصویربرداری و پروتکل‌های تصویربرداری است، می‌تواند قابلیت‌های عمومی و استحکام آن‌ها را افزایش دهد.
- **گنجاندن اطلاعات زمانی:** اکوکاردیوگرام‌ها چرخه‌های دینامیک قلبی را ضبط می‌کنند. استفاده از اطلاعات زمانی از طریق ادغام شبکه‌های عصبی بازگشتی یا تکنیک‌های مدل‌سازی زمانی می‌تواند دقت و ثبات تقسیم‌بندی را در طول فریم‌ها بهبود بخشد.
- **کوانت‌سازی عدم قطعیت (Uncertainty quantification):** توسعه روش‌هایی برای کوانت‌سازی عدم قطعیت یا اعتماد پیش‌بینی‌های تقسیم‌بندی می‌تواند بینش‌های ارزشمندی برای پزشکان فراهم کرده و به فرآیندهای تصمیم‌گیری کمک کند.

با پرداختن به این چالش‌ها و فرصت‌ها، حوزه تقسیم‌بندی اکوکاردیوگرام بطن چپ می‌تواند به سمت راه‌حل‌های دقیق‌تر، قابل‌اعتمادتر و کاربردی‌تر در بالین پیش برود، که در نهایت منجر به بهبود مراقبت از قلب و نتایج بیماران خواهد شد.

نوآوری‌های مقاله: (توسط خواننده – نورزاد)

۱. استفاده از YOLO: این مقاله از مدل YOLO برای انجام segmentation بطن چپ استفاده می‌کند که معمولاً برای شناسایی اشیاء در تصاویر مورد استفاده قرار می‌گیرد. این رویکرد باعث افزایش سرعت و دقت در پردازش تصاویر پزشکی می‌شود.

۲. خودکارسازی فرآیند: مقاله بر روی خودکارسازی کامل فرآیند segmenting تمرکز دارد که این موضوع به کاهش زمان و هزینه‌های مرتبط با تحلیل‌های دستی کمک می‌کند.

۳. بهبود دقت در تشخیص: به کارگیری تکنیک‌های یادگیری عمیق و بهینه‌سازی مدل برای بهبود دقت در شناسایی مرزهای بطن چپ از دیگر نوآوری‌های این مقاله است.

۴. ارزیابی عملکرد در شرایط بالینی: این مقاله به ارزیابی عملکرد مدل در شرایط واقعی بالینی و مقایسه آن با روش‌های سنتی پرداخته که می‌تواند به اعتبارسنجی نتایج کمک کند.

۵. گزارش نتایج عملکرد: ارائه نتایج قابل قبول و مقایسه آن با داده‌های موجود از اهمیت زیادی برخوردار است که نوآوری‌های جدید را به خوبی در بستر موجود قرار می‌دهد.

این نوآوری‌ها باعث می‌شود که مقاله در زمینه استفاده از یادگیری عمیق برای تحلیل تصاویر پزشکی به ویژه در کاربردهای کاردیولوژیک به یک منبع معتبر تبدیل شود.

پیشنهاداتی برای بهبود مقاله (توسط خواننده – نورزاد):

۱. افزایش تنوع داده‌ها: جمع‌آوری یک دیتاست بزرگتر و متنوع‌تر از تصاویر اکوکاردیوگرافی با شرایط مختلف بیماران می‌تواند به بهبود عملکرد مدل کمک کند. این شامل تصاویر با انواع بیماری‌های قلبی، سنین و جنسیت‌های مختلف است.

۲. توسعه معماری مدل: بررسی و آزمایش معماری‌های دیگر یادگیری عمیق مانند ResNet یا EfficientNet به جای YOLO، که ممکن است به دقت بالاتر و کارایی بهتر منجر شود.

۳. استفاده از تکنیک‌های پیش‌پردازش و پس‌پردازش: اعمال تکنیک‌های پیش‌پردازش مانند افزایش تصاویر (data augmentation) و روش‌های پس‌پردازش برای بهبود دقت segmentation می‌تواند مؤثر باشد.

۴. استفاده از چندین متریک ارزیابی: علاوه بر متریک‌های معمول مانند DSC و JSC، استفاده از متریک‌های دیگر مانند Hausdorff distance یا precision-recall برای ارزیابی بهتر عملکرد مدل.

۵. مقایسه با روش‌های دیگر: اضافه کردن بخش مقایسه با روش‌های سنتی و دیگر مدل‌های پیشرفته می‌تواند به وضوح نشان دهد که مدل YOLO چگونه عملکرد بهتری دارد.

۶. **تحلیل خطا:** تحلیل دقیق مواردی که مدل در آن‌ها دقت کمتری دارد، می‌تواند به بهبود بیشتر الگوریتم و طراحی مجدد مدل کمک کند.

۷. **ارزیابی زمان واقعی:** بررسی عملکرد مدل در زمان واقعی و ارائه راهکارهایی برای استفاده در محیط‌های بالینی به ویژه در جراحی‌های قلبی می‌تواند ارزش کاربردی مقاله را افزایش دهد.

۸. **گنجاندن نظرات بالینی:** همکاری با پزشکان و متخصصان برای به دست آوردن نظرات آنها در مورد نتایج و کاربردهای بالینی می‌تواند به بهبود پروژه کمک کند و اعتبار آن را افزایش دهد.

این بهبودها می‌تواند به افزایش دقت، کارایی و قابلیت استفاده از مدل در شرایط واقعی بالینی کمک کند.

***** **پایان بررسی مقاله اول** *****

مقاله دوم:

Fully Automatic Left Ventricle Segmentation Using Bilateral Lightweight Deep Neural Network

Muhammad Ali Shoaib 1,2, Joon Huang Chuah 1, Raza Ali 1,2, Samiappan Dhanalakshmi 3, Yan Chai Hum 4, Azira Khalil 5 and Khin Wee Lai 6, *

Published: 1 January 2023 <https://doi.org/10.3390/life13010124>

<https://www.mdpi.com/journal/life>

خلاصه بخش چکیده

این مقاله به بررسی تقسیم‌بندی خودکار بطن چپ (LV) می‌پردازد که برای به‌دست‌آوردن اندازه‌گیری‌های کمی قلب ضروری است. تقسیم‌بندی LV در عمل بالینی غالباً به‌صورت نیمه‌خودکار انجام می‌شود که این امر وابستگی به اپراتور و عوارضی را به همراه دارد. با وجود پیشرفت‌های اخیر در یادگیری عمیق، مدل‌های موجود هنوز از نظر دقت و سرعت با چالش‌هایی روبرو هستند. هدف این تحقیق توسعه یک مدل تقسیم‌بندی سبک و یک مرحله‌ای برای تقسیم‌بندی سریع و دقیق LV از تصاویر اکوکاردیوگرافی ۲ بعدی است.

در این مدل از یک شبکه بنیادی برای به‌دست‌آوردن ویژگی‌های سطح پایین و بالا استفاده شده و دو بلوک موازی به نام‌های واحد ویژگی فضایی و واحد ویژگی کانالی برای بهبود این ویژگی‌ها به کار رفته است. ویژگی‌های تصفیه‌شده در یک واحد یکپارچه ادغام شده تا LV تقسیم‌بندی شود. نتایج نشان می‌دهد که این مدل دارای بهترین مقادیر شاخص تشابه دایس (۰,۹۴۴۶)، هم‌پوشانی (۰,۸۴۴۵) و دقت (۰,۹۷۴۲) است. همچنین زمان پردازش این مدل نسبت به مدل‌های شناخته‌شده دیگر بهتر است.

جدول جمع‌بندی بخش چکیده

ویژگی	توضیحات
موضوع	تقسیم‌بندی خودکار بطن چپ (LV)
چالش‌ها	مرزهای نامشخص و وابستگی به اپراتور
هدف تحقیق	توسعه مدل سبک و یک مرحله‌ای برای تقسیم‌بندی سریع و دقیق LV
روش	استفاده از شبکه بنیادی و بلوک‌های موازی (واحد ویژگی فضایی و کانالی)
نتایج	-دایس: ۰,۹۴۴۶ -هم‌پوشانی: ۰,۸۴۴۵ -دقت: ۰,۹۷۴۲
مزیت	زمان کمتر برای آموزش و تقسیم‌بندی نسبت به مدل‌های دیگر

۱. خلاصه بخش مقدمه

بیماری‌های قلبی-عروقی (CVD) یکی از علل اصلی مرگ و میر در جهان هستند و برای ارزیابی ویژگی‌های تشخیصی قلب، اندازه‌گیری حجم‌های قلب ضروری است. بطن چپ (LV) مهم‌ترین بخش برای اندازه‌گیری این پارامترها به شمار می‌آید. تکنیک‌های تصویربرداری مانند اکوکاردیوگرافی به‌عنوان روش‌های قابل‌اعتماد در تشخیص و مدیریت CVD شناخته می‌شوند و اکوکاردیوگرافی به‌دلیل غیرتهاجمی و کاربرپسند بودن، روش غالب برای ارزیابی قلب محسوب می‌شود.

با وجود مزایای اکوکاردیوگرافی، تقسیم‌بندی خودکار LV هنوز در عمل بالینی معمول نیست و معمولاً از روش‌های نیمه‌خودکار یا دستی استفاده می‌شود که ممکن است منجر به تغییرات قابل توجه در نتایج شود. چالش‌هایی مانند کنتراست پایین و ناهمگنی روشنایی در تصاویر اکوکاردیوگرافی، نیاز به روش‌های متعددی برای حذف نویز و بهبود ویژگی‌های آناتومیک دارند.

در سال‌های اخیر، یادگیری عمیق (DL) به‌عنوان یک روش پردازش تصویر خودکار به‌طور فزاینده‌ای در بخش پزشکی مورد توجه قرار گرفته است. این مطالعه به پیشنهاد یک تکنیک DL برای تقسیم‌بندی LV از تصاویر اکوکاردیوگرافی پرداخته و مزایای زیر را مطرح می‌کند:

- یک روش تقسیم‌بندی سریع و دقیق با مدل DL سبک و یک مرحله‌ای.
- استخراج داده‌های فضایی و زمینه‌ای از یک مدل بنیادی سبک.
- تصفیه هم‌زمان ویژگی‌های فضایی و زمینه‌ای توسط دو واحد مجزا.
- استفاده از یک واحد یکپارچه برای ترکیب مؤثر ویژگی‌ها.

جدول جمع‌بندی بخش مقدمه

ویژگی	توضیحات
موضوع	اهمیت اندازه‌گیری حجم‌های قلب و بطن چپ (LV) در تشخیص بیماری‌های قلبی-عروقی (CVD)
چالش‌ها	وابستگی به روش‌های نیمه‌خودکار و دستی، کنتراست پایین و ناهمگنی روشنایی در تصاویر اکوکاردیوگرافی
تکنیک پیشنهادی	استفاده از یادگیری عمیق (DL) برای تقسیم‌بندی خودکار LV
مزایای روش	-مدل DL سبک و یک مرحله‌ای -استخراج هم‌زمان ویژگی‌ها -یکپارچه‌سازی مؤثر ویژگی‌ها

۲. خلاصه بخش مرور ادبیات

در این بخش، تکنیک‌های مختلفی که برای تقسیم‌بندی LV مورد استفاده قرار گرفته‌اند، بررسی شده است. روش‌هایی از جمله مدل‌های فعال، مدل‌های تغییر شکل‌دهنده و یادگیری ماشین، نتایج قابل قبولی را به همراه داشته‌اند، اما عمدتاً به ویژگی‌های دست‌ساز وابسته‌اند. در مقابل، یادگیری عمیق (DL) پیشرفت قابل توجهی در تقسیم‌بندی تصاویر پزشکی نشان داده است.

مدل‌های DL مانند شبکه‌های کانولوشن کامل (FCN) و U-Net به عنوان مدل‌های تقسیم‌بندی یک مرحله‌ای، موفقیت‌های زیادی در وظایف تقسیم‌بندی تصاویر به دست آورده‌اند. این مدل‌ها از معماری‌های کدگذار-کدگشا استفاده می‌کنند که ویژگی‌ها را استخراج کرده و برای تولید ماسک‌های تقسیم‌بندی به کار می‌روند. به علاوه، برخی مدل‌های توجه‌محور (Attention-based) یک مرحله‌ای نیز برای تقسیم‌بندی تصویر پزشکی پیشنهاد شده‌اند که عملکرد بالایی دارند اما ممکن است زمان بیشتری برای آموزش و پردازش نیاز داشته باشند.

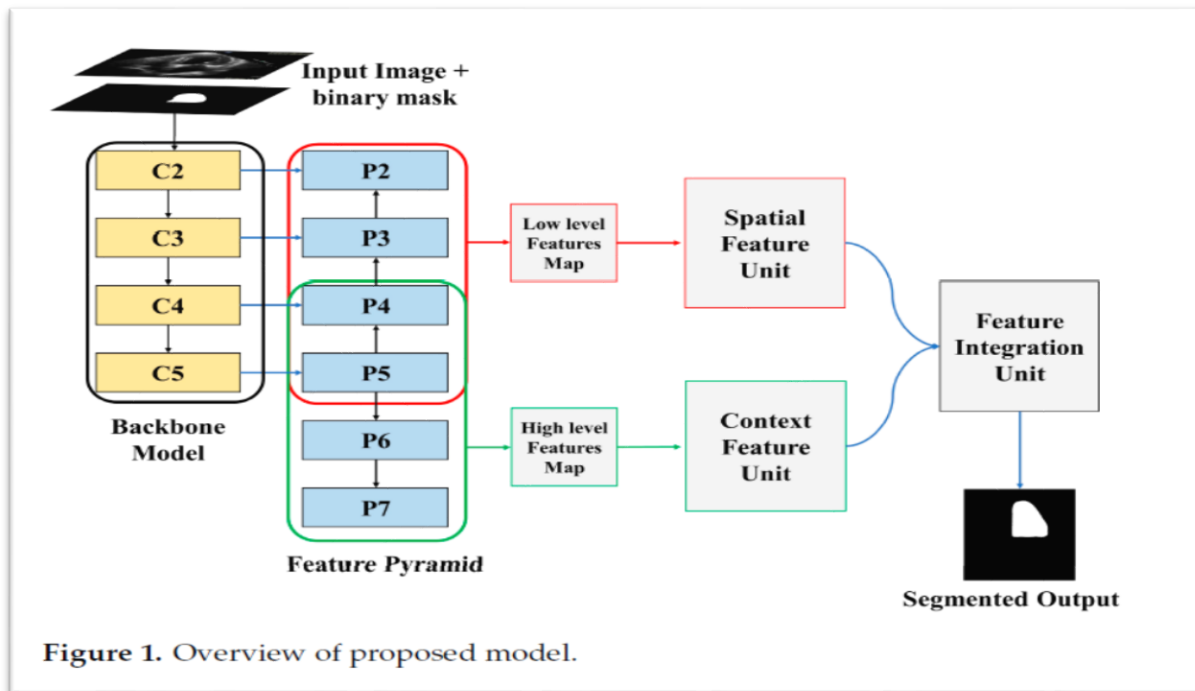
مدل‌های دو مرحله‌ای مانند Mask R-CNN نیز در تقسیم‌بندی LV موفق عمل کرده‌اند و در مقایسه با مدل‌های یک مرحله‌ای دقت بهتری را ارائه می‌دهند. با این حال، این مدل‌ها برای پردازش‌های زمان واقعی کندتر هستند. به طور کلی، مدل‌های مبتنی بر شبکه‌های عصبی کانولوشن (CNN) به طور قابل توجهی نسبت به روش‌های غیر DL در تقسیم‌بندی LV بهبود یافته‌اند، اما توازن بین سرعت و دقت همچنان چالش برانگیز است.

جدول جمع‌بندی بخش مرور ادبیات

ویژگی	توضیحات
روش‌های قبلی	مدل‌های فعال، مدل‌های تغییر شکل‌دهنده، یادگیری ماشین
چالش‌ها	وابستگی به ویژگی‌های دست‌ساز در تکنیک‌های غیر DL
مدل‌های DL	- FCN و U-Net مدل‌های یک مرحله‌ای با موفقیت بالا - مدل‌های توجه‌محور (Attention): عملکرد خوب اما زمان‌بر
مدل‌های دو مرحله‌ای	Mask R-CNN: دقت بالاتر نسبت به یک مرحله‌ای‌ها اما کندتر
نتیجه‌گیری	مدل‌های CNN پیشرفت قابل توجهی در تقسیم‌بندی LV نشان داده‌اند، اما توازن بین سرعت و دقت همچنان چالش برانگیز است

۳. خلاصه بخش روش شناسی (Methodology)

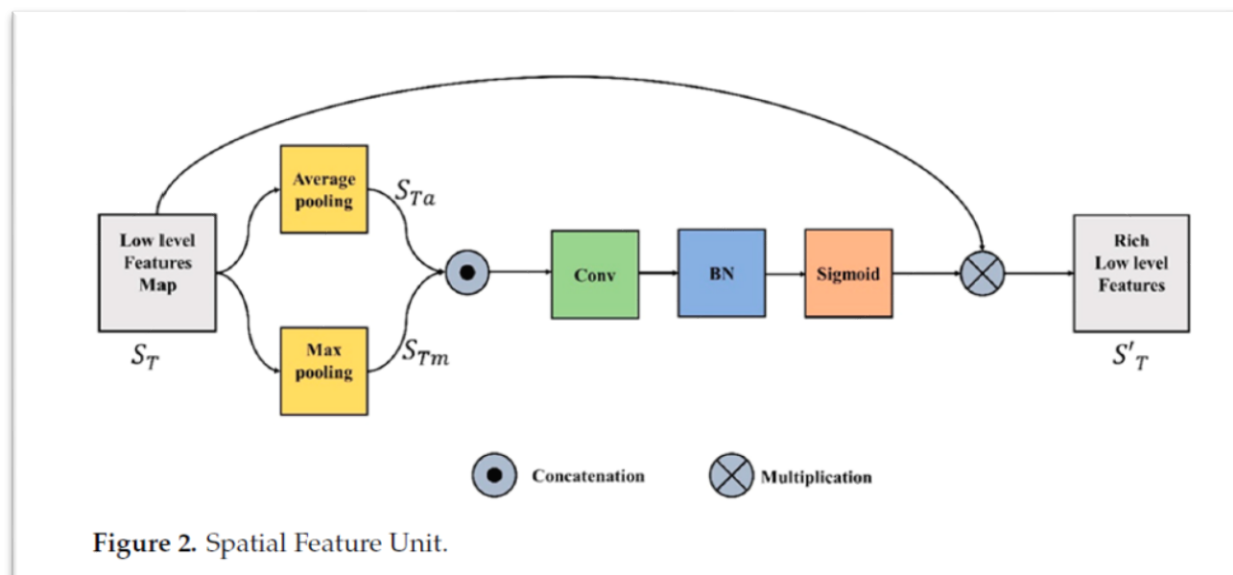
مدل تقسیم‌بندی LV پیشنهادی در این بخش شامل مراحل مختلفی است که در شکل ۱ نشان داده شده است. تصاویر آموزشی به همراه ماسک‌های باینری به مدل بنیادی و شبکه هرم ویژگی (FPN) ارسال می‌شوند که ویژگی‌های سطح پایین و بالا را استخراج می‌کند. سپس ویژگی‌های استخراج‌شده در دو واحد ویژگی فضایی و کانالی پردازش می‌شوند و در نهایت، خروجی این دو واحد در واحد یکپارچه‌سازی ویژگی‌ها ترکیب می‌شود تا LV تقسیم‌بندی شود.



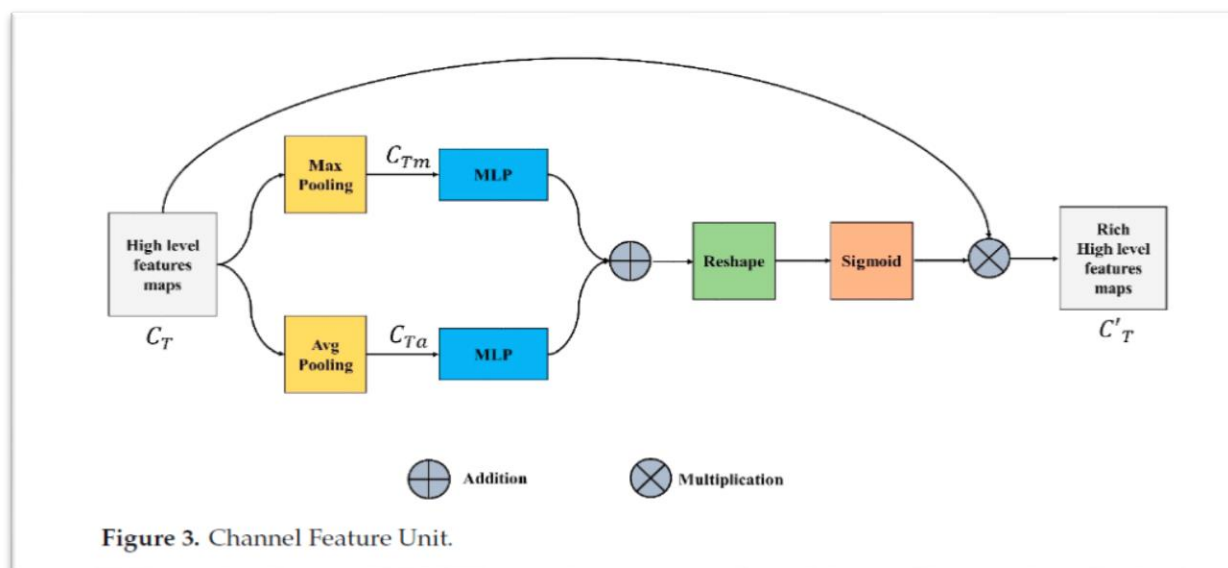
۳-۱. مدل بنیادی و شبکه هرم ویژگی

مدل بنیادی (Backbone Model) استفاده شده **ResNet50** است که به‌عنوان یک شبکه عصبی با توانایی train در تعداد زیادی لایه، ویژگی‌های تصویر را از طریق FPN استخراج می‌کند. این شبکه با ایجاد یک مسیر اضافی، ویژگی‌های سطح بالا را از لایه‌های بالایی به لایه‌های پایینی منتقل می‌کند. این کار منجر به تولید نقشه‌های ویژگی در مقیاس‌های مختلف می‌شود.

۳-۲. واحد ویژگی فضایی



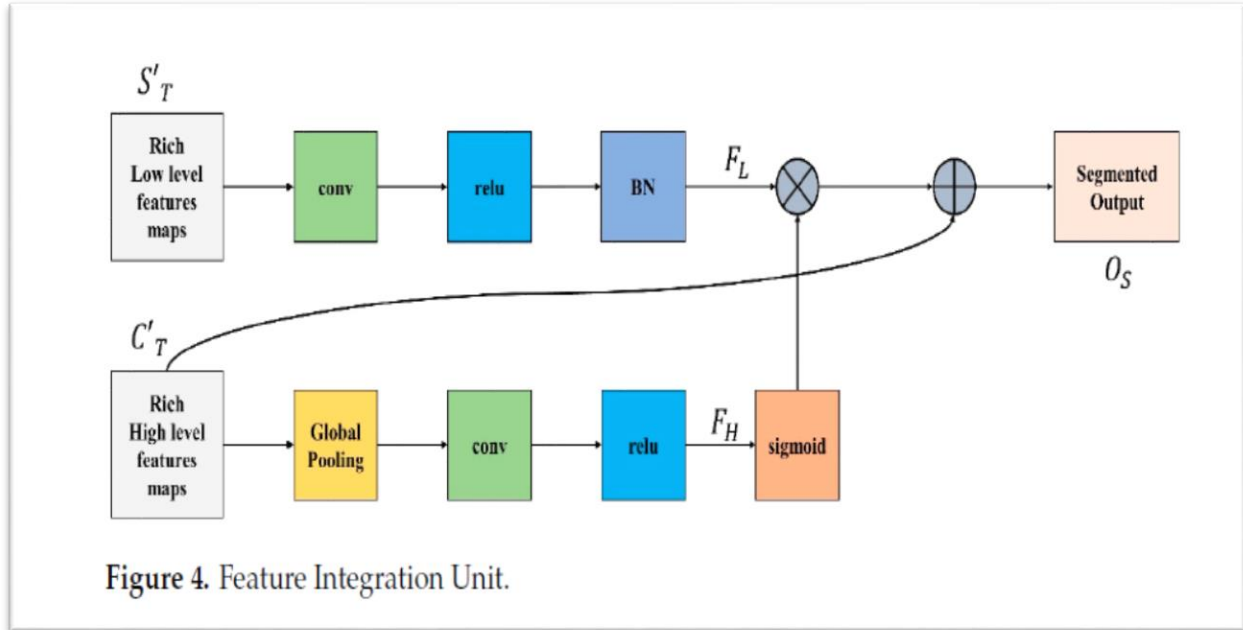
۳-۳. واحد ویژگی کانالی



ویژگی‌های سطح پایین و بالا به‌طور مجزا توسط این دو واحد پردازش می‌شوند. واحد ویژگی فضایی به پردازش ویژگی‌های فضایی و واحد ویژگی کانالی به پردازش ویژگی‌های کانالی اختصاص دارد.

۴-۳. واحد یکپارچه سازی ویژگی ها

این واحد مسئول ترکیب خروجی های دو واحد قبلی است تا در نهایت LV به طور دقیق تقسیم بندی شود.



۵-۳. Loss Function

تابع ضرر در تکنیک های تقسیم بندی مبتنی بر DL نقش حیاتی دارد. در این مطالعه، از تابع دایس (Dice Similarity Coefficient) به عنوان تابع ضرر استفاده شده که هدف آن کاهش عدم تطابق بین نتایج واقعی و پیش بینی شده است.

$$Loss = 1 - \frac{2 \sum_{k=1}^N \sum_{l=1}^M g_l^k s_l^k}{\sum_{k=1}^N \sum_{l=1}^M g_l^k + \sum_{k=1}^N \sum_{l=1}^M s_l^k}$$

جدول جمع‌بندی بخش روش‌شناسی

ویژگی	توضیحات
مدل پیشنهادی	مدل تقسیم‌بندی LV با استفاده از ResNet 50 و شبکه هرم ویژگی (FPN)
مراحل اصلی	۱. استخراج ویژگی‌ها (سطح پایین و بالا) ۲. پردازش توسط واحدهای ویژگی فضایی و کانالی ۳. یکپارچه‌سازی ویژگی‌ها
مدل بنیادی (Backbone Model)	ResNet 50
عملکرد FPN	تولید نقشه‌های ویژگی در مقیاس‌های مختلف و انتقال ویژگی‌های سطح بالا به لایه‌های پایین‌تر
واحد یکپارچه‌سازی ویژگی‌ها	ترکیب خروجی‌های واحدهای فضایی و کانالی برای تقسیم‌بندی نهایی
تابع خسارت	استفاده از تابع دیس برای کاهش عدم تطابق بین نتایج واقعی و پیش‌بینی‌شده

۴. خلاصه بخش آزمایشات

در این بخش، داده‌های استفاده‌شده برای آموزش و آزمایش مدل، سخت‌افزار، هاپرپارامترها و معیارهای ارزیابی برای تحلیل عملکرد مدل‌های یادگیری عمیق توضیح داده شده است.

۴-۱. دیتاست

داده‌ها از مؤسسه قلب ملی در کوالالامپور مالزی به‌دست آمده و شامل ۶۰۰۰ تصویر دو بعدی از اکوکاردیوگرافی چهار حفره‌ای است. این تصاویر با استفاده از سیستم اولتراسوند Philips IE33 و ترانسدایوسر S5-1 گرفته شده‌اند. ابعاد هر تصویر ۸۰۰ در ۶۰۰ پیکسل با وضوح ۰,۳×۰,۳ میلی‌متر و نرخ فریم بین ۳۰ تا ۱۰۰ هرتز است. پس‌زمینه‌های نامرتبط از تصاویر برش داده شده و اندازه تصاویر به ۵۱۲×۵۱۲ کاهش یافته است. داده‌ها به سه بخش تقسیم شده‌اند: ۷۰٪ برای آموزش، ۱۵٪ برای تست و ۱۵٪ برای اعتبارسنجی. ماسک‌های باینری از LV با تأیید متخصصان پزشکی تولید شده‌اند.

۲-۴. آموزش شبکه

آموزش شبکه با استفاده از یک workstation با پردازنده **Dell Core i7 Xeon E5-2620** و کارت گرافیک **11 GB Nvidia GeForce GTX 1080Ti** انجام شده است. از بهینه‌ساز مبتنی برگرادیان نزولی (Stochastic Gradient Descent) با مومنتوم 0.9 استفاده شده و نرخ یادگیری اولیه 10^{-4} و نرخ کاهش وزن 0.001 تنظیم شده است. هر مدل برای 50 دوره با اندازه دسته 32 و شافل پس از هر دوره آموزش داده شده است.

۳-۴. معیارهای ارزیابی

مدل آموزش‌دیده بر اساس 15% از تصاویر آزمایش که برای تست رزرو شده‌اند، ارزیابی می‌شود. ماسک‌های باینری تقسیم‌بندی شده تولید شده توسط مدل، با ماسک‌های باینری واقعی مقایسه می‌شوند. عملکرد مدل پیشنهادی با استفاده از معیارهای رایجی مانند **Dice similarity coefficient (DSC)**، **intersection over union (IoU)**، دقت (accuracy)، **Recall**، **Precision** و **Specificity** ارزیابی و مقایسه می‌شود.

جدول جمع‌بندی بخش آزمایشات

ویژگی	توضیحات
داده‌ها	6000 تصویر دو بعدی از اکوکاردیوگرافی چهار حفره‌ای، 70% آموزش، 15% تست، 15% اعتبارسنجی
منبع داده‌ها	مؤسسه قلب ملی کوالالامپور، مالزی
ابزارهای تصویربرداری	سیستم اولتراسوند Philips IE33 با ترانسدایوسر S5-1
ابعاد تصاویر	800×600 پیکسل (کاهش‌یافته به 512×512)
سخت‌افزار آموزش	Dell Core i7 Xeon E5-2620 CPU Nvidia GeForce GTX 1080Ti GPU
هایپرپارامترها	– نرخ یادگیری: 10^{-4} – مومنتوم: 0.9 – نرخ کاهش وزن: 0.001
دوره‌های آموزشی	50 دوره با اندازه دسته 32
معیارهای ارزیابی	Specificity, Precision, Recall, Accuracy, IoU, DSC

۵. خلاصه بخش نتایج

در این بخش، یافته‌ها و نتایج تجربی تقسیم‌بندی LV ارائه شده و عملکرد مدل پیشنهادی با مدل‌های معروف یادگیری عمیق مانند Mask R-CNN، FCN و DeepLab مقایسه شده است.

۵-۱. نتایج عددی

جدول ۱ مقایسه عملکرد تقسیم‌بندی مدل پیشنهادی را با استفاده از پنج معیار ارزیابی تقسیم‌بندی نمایش می‌دهد. مدل پیشنهادی برای تصاویر آزمایشی به مقادیر زیر دست یافته است:

- ضریب تشابه دایس (DSC): 0.9446
- تقاطع بر اتحاد یا همپوشانی (IoU): 0.8845
- دقت (Accuracy): 0.9742
- یادآوری (Recall): 0.9889
- خاصیت (Specificity): 0.8357

مدل DeepLab اگرچه زمان کمتری نسبت به سایر مدل‌ها می‌گیرد، اما عملکرد کلی آن پایین‌ترین است. Mask R-CNN به دلیل طراحی دو مرحله‌ای‌اش، عملکرد بهتری نسبت به FCN و DeepLab داشته، اما مدل پیشنهادی بهترین نتایج را در بین تمام معماری‌های تقسیم‌بندی ارائه کرده است.

جدول ۱: میانگین و انحراف معیار معیارهای ارزیابی

مدل	DSC	IoU	Accuracy	Recall	Specificity	Precision
Proposed	0.9446 ± 0.0319	0.8845 ± 0.0536	0.9742 ± 0.0114	0.9889 ± 0.0081	0.8357 ± 0.0817	0.9828 ± 0.0088
Mask R-CNN	0.9239 ± 0.0232	0.8651 ± 0.0457	0.9612 ± 0.0116	0.9760 ± 0.0084	0.8252 ± 0.0845	0.9806 ± 0.0087
FCN	0.8758 ± 0.0439	0.7978 ± 0.0536	0.9560 ± 0.0115	0.9617 ± 0.0178	0.8468 ± 0.0587	0.9818 ± 0.0095
DeepLab	0.8485 ± 0.0409	0.7448 ± 0.0620	0.9513 ± 0.0092	0.9593 ± 0.0167	0.7403 ± 0.0787	0.9728 ± 0.0098

در جدول ۲، زمان صرف شده برای آموزش و تقسیم‌بندی LV از تصاویر آزمایشی برای هر مدل ارائه شده است.

جدول ۲: زمان آموزش و زمان تقسیم‌بندی مدل‌ها

مدل	زمان آموزش	زمان تقسیم‌بندی تصاویر آزمایشی
Proposed	۵۳ دقیقه	24.86 ثانیه
Mask R-CNN	۱۲ ساعت ۳۷ دقیقه	41.27 ثانیه
FCN	۱۰ ساعت ۵۷ دقیقه	31.97 ثانیه
DeepLab	۶ ساعت ۴۲ دقیقه	20.82 ثانیه

۵-۲. نتایج بصری

شکل ۵ نتایج بصری خروجی تقسیم‌بندی تمامی مدل‌ها را ارائه می‌دهد. دقت تقسیم‌بندی DeepLab به اندازه کافی خوب نیست و خطوط مرزی آن دقیق نیست. در حالی که Mask R-CNN و FCN خطوط مرزی دقیقی دارند، اما شکل LV به اندازه مدل پیشنهادی دقیق نیست. مدل پیشنهادی نمای دقیقی از فرم LV و خطوط مرزی آن ارائه داده است.

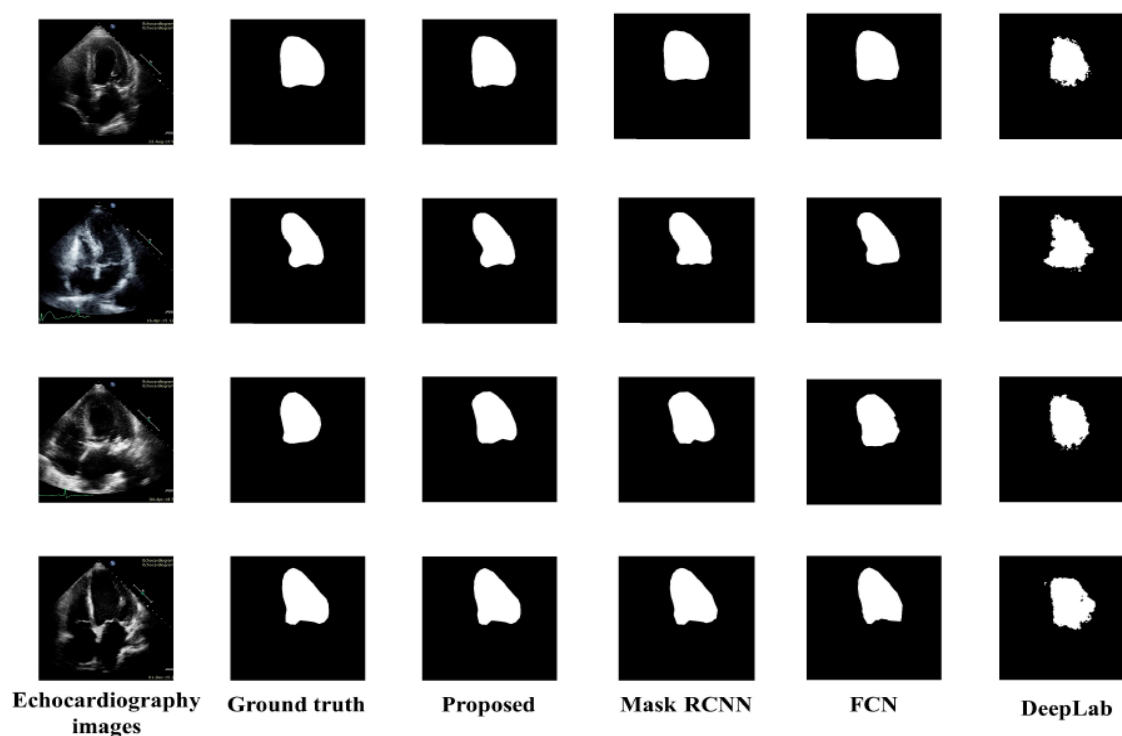


Figure 5. Visual comparisons of the proposed model to various segmentation models.

جدول جمع‌بندی بخش نتایج

ویژگی	توضیحات
عملکرد مدل پیشنهادی	DSC: 0.9446 IoU: 0.8845 Accuracy: 0.9742 Recall: 0.9889 Specificity: 0.8357
مقایسه مدل‌ها	Mask R-CNN بهتر از FCN و DeepLab است و اما مدل پیشنهادی بهترین نتایج را دارد.
زمان آموزش	Proposed: ۵۳ دقیقه Mask R-CNN: ۱۲ ساعت و ۳۷ دقیقه FCN: ۱۰ ساعت و ۵۷ دقیقه DeepLab: ۶ ساعت و ۴۲ دقیقه
زمان تقسیم‌بندی	Proposed: 24.86 ثانیه Mask R-CNN: 41.27 ثانیه FCN: 31.97 ثانیه DeepLab: 20.82 ثانیه
نتایج بصری	مدل پیشنهادی بهترین دقت را در شناسایی فرم و خطوط مرزی LV ارائه می‌دهد

۶. بحث

در این تحقیق، یک مدل تقسیم‌بندی سبک و کارآمد پیشنهاد شده است که عملکرد آن و زمان‌های آموزش و تست آن با مدل‌های DeepLab ، FCN و Mask R-CNN مقایسه شده است. همان‌طور که در جدول ۱ مشاهده می‌شود، مدل پیشنهادی بالاترین سطوح عملکرد را از نظر DSC ، IoU ، Accuracy ، Recall ، Precision و Specificity به دست آورده است. این مدل از ویژگی‌های استخراج شده با استفاده از مدل ResNet 50 بهره‌برداری کرده و سپس ویژگی‌های فضایی و کانالی را به‌طور جداگانه پردازش می‌کند تا حداکثر استفاده از اطلاعات استخراج شده را به دست آورد. مدل پیشنهادی به‌طور قابل توجهی زمان کمتری برای آموزش و تقسیم‌بندی تصاویر آزمایشی نسبت به FCN و Mask R-CNN نیاز دارد. این امر به دلیل پردازش موازی تشکیل ویژگی‌های فضایی و کانالی است که پس از استخراج ویژگی‌ها از مدل پایه انجام می‌شود.

تحلیل آماری

علاوه بر این، تحلیل آماری معیارهای ارزیابی نیز ارائه شده است. برای تحلیل و مقایسه توزیع مقادیر معیارهای ارزیابی برای تمام تصاویر آزمایشی، از نمودار جعبه‌ای استفاده شده است. شکل ۶ نمودار جعبه‌ای مقادیر DSC را برای تصاویر آزمایشی نشان می‌دهد. مدل پیشنهادی در مقایسه با سایر روش‌های تقسیم‌بندی، در زمینه حداقل و حداکثر مقدار، میانه و چارک‌های بالا و پایین عملکرد بهتری دارد. همچنین، مدل پیشنهادی تعداد کمتری از نقاط دورافتاده (outlier) را نسبت به دیگر مدل‌ها نشان می‌دهد، که نشان‌دهنده ثبات عملکرد آن در کل مجموعه داده‌های آزمایشی است.

نتایج نموداری

شکل‌های ۷ تا ۱۱ نمودارهای جعبه‌ای برای IoU، Accuracy، Recall، Precision و Specificity را نشان می‌دهند. از این نمودارها می‌توان استنباط کرد که رفتار مدل پیشنهادی برای این معیارها به‌طور کلی مشابه DSC است. علاوه بر این، عملکرد بالاتر مدل پیشنهادی با پایین‌تر بودن skewness نمودار جعبه‌ای آن نسبت به skewness نمودارهای جعبه‌ای سایر مدل‌ها تأیید می‌شود.

این بخش می‌تواند به بهبود و روشن‌تر شدن نقاط قوت و ضعف مدل‌های مختلف کمک کند و بر اهمیت انتخاب مدل‌های بهینه برای تقسیم‌بندی LV تأکید کند.

۷. نتیجه‌گیری

در این مطالعه، یک شبکه تقسیم‌بندی سبک برای تقسیم‌بندی خودکار بطن چپ (LV) از تصاویر اکوکاردیوگرافی طراحی و اجرا شد. با استفاده از ResNet50 به عنوان مدل پایه، یک شبکه هرمی ویژگی برای استخراج ویژگی‌های سطح پایین و بالا به کار گرفته شد. واحد ویژگی فضایی (SFU) با استفاده از فرآیندهای تجمیع، پیچش، نرمال‌سازی دسته‌ای و سیگموئید، ویژگی‌های فضایی تولید شده توسط مدل پایه را بهبود می‌بخشد. به‌طور مشابه، اطلاعات کانالی با استفاده از تجمیع گلوبالی و پرسپترون چند لایه در واحد ویژگی کانالی تصفیه می‌شود. با استفاده از واحد ادغام ویژگی‌ها، این دو نوع ویژگی با سطوح مختلف ترکیب می‌شوند.

مدل پیشنهادی نه تنها از نظر دقت تقسیم‌بندی بلکه از نظر سرعت نیز عملکرد خوبی داشت. زمان لازم برای تقسیم‌بندی تصاویر آزمایشی و آموزش مدل کمتر از سایر مدل‌های معروف تقسیم‌بندی مانند Mask FCN, R-CNN و DeepLab بود. مدل پیشنهادی بالاترین معیارهای ارزیابی را با DSC برابر با ۰,۹۴۴۶, IoU برابر با ۰,۸۴۴۵, Accuracy برابر با ۰,۹۷۴۲, Recall برابر با ۰,۹۸۸۹, Precision برابر با ۰,۹۸۲۸ و Specificity برابر با ۰,۸۳۵۷ به دست آورد.

جمع‌بندی کلی مقاله

بخش	خلاصه
چکیده	پیشنهاد یک مدل تقسیم‌بندی خودکار برای LV با دقت و سرعت بالا.
مقدمه	اهمیت تقسیم‌بندی LV در تشخیص بیماری‌های قلبی و نیاز به مدل‌های خودکار.
مرور ادبیات	بررسی تکنیک‌های مختلف و پیشرفت‌های یادگیری عمیق در تقسیم‌بندی.
متدولوژی	استفاده از ResNet50 و ویژگی‌های فضایی و کانالی برای بهبود دقت.
آزمایشات	توصیف داده‌ها، تنظیمات آموزشی و معیارهای ارزیابی.
نتایج	مقایسه عملکرد مدل پیشنهادی با مدل‌های دیگر و نمایش زمان‌های آموزش و تست
بحث	تحلیل آماری و بررسی ثبات عملکرد مدل.
نتیجه‌گیری	تأکید بر عملکرد برتر مدل پیشنهادی از نظر دقت و سرعت.

***** پایان بررسی مقاله دوم *****

Deep Learning Based Automatic Left Ventricle Segmentation from the Transgastric Short-Axis View on Transesophageal Echocardiography: A Feasibility Study

Yuan Tian 1, †, Wenting Qin 2, †, Zihang Zhao 2, Chunrong Wang 1, Yajie Tian 1, Yuelun Zhang 1, Kai He 1, Yuguan Zhang 1, Le Shen 1, Zhuhuang Zhou 2, * and Chunhua Yu

Published: 31 July 2024

<https://doi.org/10.3390/diagnostics14151655>

<https://www.mdpi.com/journal/diagnostics>

خلاصه‌ای از چکیده

این مطالعه بر روی اتوماسیون جداسازی بطن چپ از نماهای کوتاه‌محور ترانس‌گاستریک (TSV) در اکوکاردیوگرافی ترانس‌ازوفازال (TEE) با استفاده از تکنیک‌های یادگیری عمیق تمرکز دارد. **هدف** این است که یک روش کارآمدتر و عینی برای ارزیابی قلبی-عروقی در مدیریت پری‌عملیاتی فراهم کند، زیرا جداسازی دستی در حال حاضر زمان‌بر و وابسته به تجربه پزشک است.

نکات کلیدی:

❖ داده‌ها: مجموعه‌ای از ۱۳۸۸ تصویر TSV از ۴۵۱ بیمار جمع‌آوری شد که پس از پیش‌پردازش و افزایش داده، به ۳۳۳۶ تصویر تبدیل شد.

❖ مدل‌های آزمایش‌شده: چهار مدل یادگیری عمیق مورد استفاده قرار گرفت:

- U-Net
- Attention U-Net
- U-Net++
- UNeXt

❖ معیارهای عملکرد: دقت جداسازی با استفاده از ضریب شباهت جکارد (JSC) و ضریب شباهت دایس (DSC) ارزیابی شد.

❖ نتایج:

- بهترین عملکرد: مدل‌های Attention U-Net و U-Net++ بالاترین امتیاز معیارهای ارزیابی (JSC (86.02%) و DSC (92.00%)) را کسب کردند.
- پارامترهای شبکه UNeXt: کمترین پارامترها (۱,۴۷ میلیون) را داشت.

- **زمان آموزش و استنتاج:** مدل U-Net کمترین زمان آموزش (۶۴۲۸,۶۵ ثانیه) و سریع‌ترین زمان استنتاج در هر تصویر (۱۰۱,۷۵ میلی‌ثانیه) را داشت.
- **چالش‌ها:** مدل Attention U-Net در موارد دشوار مانند مرزهای impaired و آرتیفکت‌های عضله پایپلا به خوبی عمل کرد.

این مطالعه نتیجه‌گیری می‌کند که یادگیری عمیق می‌تواند به طور قابل توجهی فرآیند جداسازی را در TEE بهبود بخشد و یک جایگزین امیدوارکننده برای ارزیابی‌های قلبی-عروقی فراهم کند.

۱. مقدمه

طبق آخرین آمار سازمان بهداشت جهانی، سالانه بیش از ۳۰۰ میلیون عمل جراحی در سراسر جهان انجام می‌شود. اکوکاردیوگرافی ترانس‌ازوفازال (TEE)، یک تکنیک ارزیابی قلبی-عروقی با استفاده از پروب ترانس‌ازوفازال، به طور فزاینده‌ای به بخشی جدایی‌ناپذیر از مدیریت پری‌عملیاتی تبدیل شده است. TEE در تسهیل تصمیم‌گیری در طول جراحی‌ها و مدیریت همودینامیک بیماران بحرانی مؤثر است و این مزایا آن را در مقایسه با اکوکاردیوگرافی ترانس‌توراکیک (TTE) در بسیاری از جراحی‌ها به یک انتخاب مناسب تبدیل کرده است.

ارزیابی عملکرد و ساختار بطن چپ با استفاده از TEE به منظور پاسخ به مشکلات شایع و تهدیدکننده حیات در حین عمل انجام می‌شود. نماهای کوتاه‌محور ترانس‌گاستریک (TSV) امکان ارزیابی سریع و دقیق عملکرد بطن چپ را فراهم می‌کنند و با تنظیم عمق پروب، دید دقیقی از آناتومی لایه‌ای دیواره بطن چپ در طول ضربان قلب ارائه می‌دهند. با این حال، ارزیابی بطن چپ در TEE یک فرایند زمان‌بر و وابسته به تجربه است و نیاز به تخصص دارد.

با پیشرفت‌های اخیر در هوش مصنوعی پزشکی، الگوریتم‌های یادگیری عمیق به عنوان یک گزینه مکمل برای ارزیابی‌های قلبی-عروقی پری‌عملیاتی مطرح شده‌اند. در حالی که بسیاری از مطالعات به بهبود استفاده از یادگیری عمیق در ارزیابی بطن چپ اشاره کرده‌اند، بیشتر این مطالعات به TTE مربوط می‌شوند. تحقیقات کمی بر روی تصاویر TEE و تمرکز بر نماهای طولی قلب انجام شده است. با توجه به پیشرفت‌های الگوریتم‌های U-Net و انواع آن از سال ۲۰۱۵، جداسازی تصاویر پزشکی با استفاده از یادگیری عمیق به بهبود قابل توجهی در دقت محاسباتی، حساسیت و کارایی دست یافته است.

هدف این مطالعه ارزیابی قابلیت یادگیری عمیق برای جداسازی خودکار بطن چپ از طریق بررسی اعتبار الگوریتم‌های مختلف U-Net است. در ابتدا، یک مجموعه داده بزرگ از تصاویر TSV از ۴۵۱ بیمار جمع‌آوری شد. پس از پیش‌پردازش تصاویر و افزایش داده، مجموعه آموزشی برای آموزش الگوریتم‌های مختلف U-Net برای جداسازی بطن چپ استفاده شد و مجموعه اعتبارسنجی برای بررسی بیش‌برازش (overfitting) به کار رفت. در نهایت، مجموعه آزمایشی برای ارزیابی و مقایسه‌ی عملکرد جداسازی الگوریتم‌های U-Net استفاده شد.

۲. مواد و روش‌ها

در این بخش، نمودار جریان روش خودکار جداسازی بطن چپ از تصاویر TSV TEE با استفاده از چهار مدل شبکه عصبی عمیق شامل U-Net، U-Net++، Attention U-Net و UNeXt نشان داده شده است. ابتدا، فریم‌های انتهایی دیاستول (EDF) و انتهایی سیستول (ESF) از ویدیوهای TEE در یک چرخه قلبی استخراج و به تصاویر خاکستری تک‌کاناله تبدیل شدند. سپس، تصاویر ESF و EDF تک‌کاناله به اندازه مشخصی تغییر سایز داده شدند. این تصاویر تغییر سایز داده شده به مدل‌های شبکه عصبی عمیق آموزش‌دیده ورودی داده شدند تا جداسازی بطن چپ پیش‌بینی شود، که در نهایت به اندازه اصلی تغییر سایز داده شد تا جداسازی نهایی بطن چپ به دست آید.

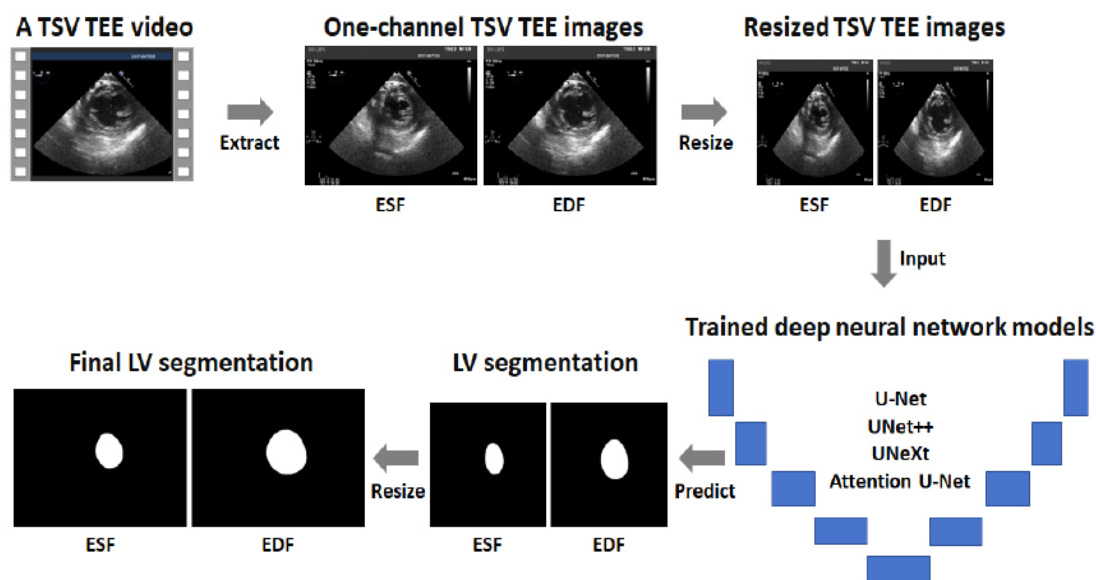


Figure 1. Flow chart of the proposed automatic LV segmentation method for TSV TEE images using deep neural network models. The deep learning models employed were U-Net [18], UNet++ [19], Attention U-Net [20], and UNeXt [21]. LV: left ventricle; TSV: transgastric short-axis view; TEE: transesophageal echocardiography; ESF: end-systolic frame; EDF: end-diastolic frame.

۲-۱. ثبت نام بیماران و تشکیل مجموعه داده

این مطالعه بازنگری شده توسط کمیته اخلاق بیمارستان دانشگاه پزشکی پکن مورد تأیید قرار گرفت. بیمارانی که در این مطالعه شرکت کردند، باید شرایط زیر را داشته باشند:

❖ انجام عمل جراحی قلب تحت بیهوشی عمومی در بیمارستان دانشگاه پزشکی پکن بین جولای ۲۰۱۵ تا اکتبر ۲۰۲۳

❖ انجام TEE پری عملیاتی با استفاده از اسکنر اولتراسوند Philips iE33 و ترانسدوسر X7-2t.

بیمارانی که دارای ناهنجاری‌های شناخته شده بطن چپ به دلیل بیماری‌های قلبی مادرزادی بودند یا ویدیوهای ذخیره شده از TSV نداشتند، از مطالعه خارج شدند. در مجموع، ۱۰۷۶ ویدیوی TSV از ۴۵۱ بیمار در این مطالعه شامل شدند. برخی از بیماران دارای دو ویدیوی TSV مجزا بودند که پیش و پس از عمل جراحی قلب ثبت شده بودند. از این تعداد، ۳۸۲ ویدیو به دلیل معیارهای زیر واجد شرایط نبودند: تکراری بودن، پوشش کمتر از یک چرخه قلبی، نداشتن مرز مشخص بطن چپ یا وجود نویز شدید.

فریم‌های EDF و ESF از هر ویدیو TSV استخراج و مجموعه داده‌ای شامل ۱۳۸۸ تصویر تشکیل شد. برای کاهش هزینه محاسباتی شبکه عصبی عمیق، تصویر B-mode دو بعدی سه کاناله TSV TEE به یک تصویر خاکستری تک کاناله تبدیل شد. در نهایت، ۱۳۸۸ تصویر دو بعدی تک کاناله TSV TEE به عنوان مجموعه داده تجربی این مطالعه به دست آمد.

مجموعه داده به نسبت ۸:۱:۱ به طور تصادفی به مجموعه‌های آموزشی، اعتبارسنجی و آزمایشی تقسیم شد. مجموعه آموزشی برای آموزش مدل‌های شبکه عصبی عمیق استفاده شد و مجموعه اعتبارسنجی برای بررسی احتمال بیش‌برازش به کار رفت. در نهایت، مجموعه آزمایشی برای ارزیابی عملکرد مدل‌های آموزش دیده استفاده شد. به طور خاص، مجموعه آموزشی شامل ۱۱۱۲ تصویر TSV TEE از ۵۵۶ ویدیو، مجموعه اعتبارسنجی شامل ۱۳۸ تصویر TSV TEE از ۶۹ ویدیو، و مجموعه آزمایشی شامل ۱۳۸ تصویر دیگر از ۶۹ ویدیو بود. داده‌های تصویری یک بیمار در هیچ یک از مجموعه‌ها تداخل نداشتند تا از نشت داده جلوگیری شود. جداسازی دستی بطن چپ برای هر یک از ۱۳۸۸ تصویر TSV TEE توسط دو متخصص بیهوشی انجام و توسط یک متخصص بیهوشی ارشد تأیید شد و به عنوان حقیقت مرجع در نظر گرفته شد.

نمودارها و تصاویری که در این بخش ذکر شده‌اند، چالش‌های اصلی جداسازی بطن چپ در تصاویر TSV TEE را نشان می‌دهند، از جمله فقدان مرز بطن چپ و تداخل ناشی از عضله پاپیلاری.

۲-۲. پیش پردازش و افزایش داده‌ها

برای انجام کاهش هزینه محاسباتی مدل‌های یادگیری عمیق، هر یک از ۱۳۸۸ تصویر TSV TEE به اندازه 256×256 پیکسل با استفاده از درون‌یابی مکعبی کاهش اندازه یافتند. بنابراین، اندازه تصویر ورودی برای شبکه‌های عصبی عمیق 256 (ارتفاع تصویر) \times 256 (عرض تصویر) \times 1 (کانال تصویر) برای هر دو مرحله آموزش و آزمایش بود. به دلیل محدودیت داده‌های تجربی، افزایش داده‌ها بر روی تصاویر مجموعه آموزشی انجام شد که شامل چرخش تصادفی از 0° تا 90° ، برش افقی و برش عمودی بود. افزایش داده‌ها (Data Augmentation) می‌تواند به کاهش بیش‌برازش مدل شبکه عصبی عمیق کمک کند و قدرت تعمیم مدل را بهبود بخشد. افزایش داده‌ها فقط بر روی مجموعه آموزشی انجام شد و بر روی مجموعه اعتبارسنجی یا آزمایش اعمال نشد. پس از انجام افزایش داده‌ها، اندازه مجموعه آموزشی به ۳۳۳۶ تصویر افزایش یافت.

۲-۳. مدل‌های شبکه عصبی عمیق

در این مطالعه، چهار شبکه عصبی عمیق برای جداسازی بطن چپ در تصاویر TSV TEE استفاده شد.

U-Net: به عنوان رایج‌ترین و ساده‌ترین مدل جداسازی در جداسازی تصاویر پزشکی شناخته می‌شود که از ساختار شبکه U شکل برای به دست آوردن اطلاعات زمینه‌ای و مکان استفاده می‌کند. این مدل شامل یک رمزگذار و یک رمزگشا است که ارتباطات پرش بین رمزگذار و رمزگشا دارد.

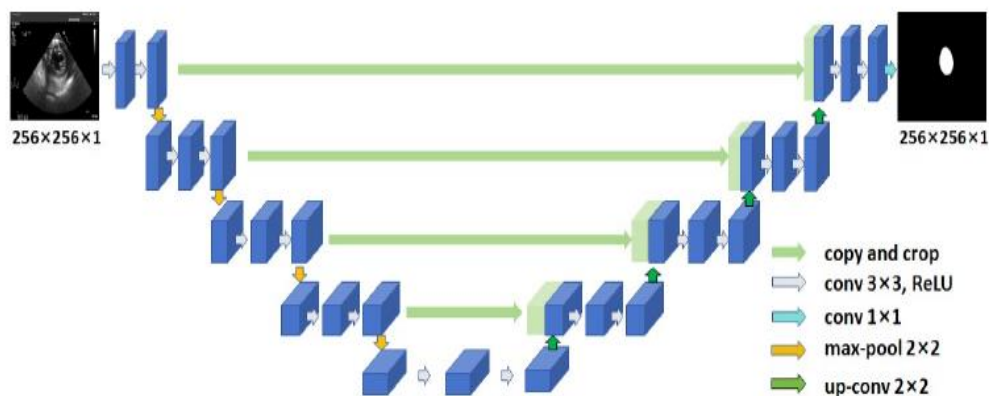


Figure 4. The U-Net network architecture for LV segmentation in TSV TEE images. LV: left ventricle; TSV: transgastric short-axis view; TEE: transesophageal echocardiography; ReLU: rectified linear unit; conv: convolution.

Attention U-Net: یک توسعه از معماری کلاسیک U-Net است که مکانیزم توجه را به U-Net اضافه می‌کند. این مدل می‌تواند وزن نواحی محلی مورد علاقه را به تدریج تقویت کرده و نواحی غیر مرتبط در تصویر ورودی را سرکوب کند.

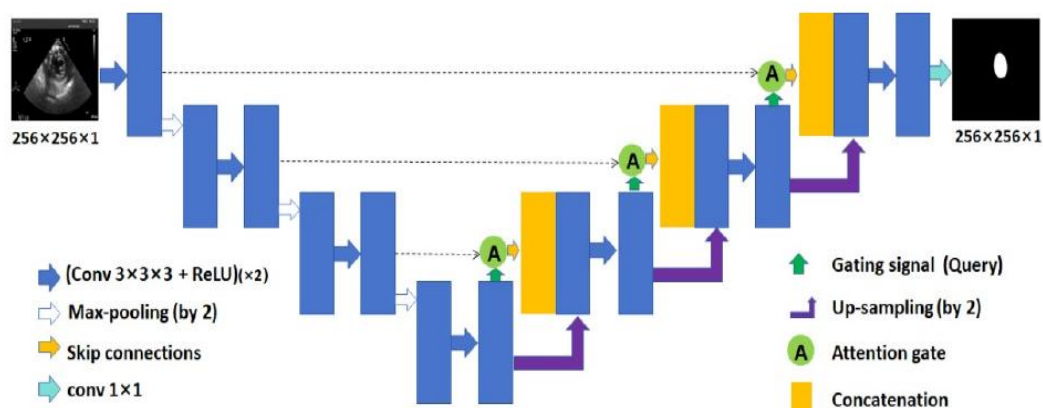


Figure 5. The Attention U-Net network architecture for LV segmentation in TSV TEE images. LV: left ventricle; TSV: transgastric short-axis view; TEE: transesophageal echocardiography; ReLU: rectified linear unit; conv: convolution.

U-Net++: به عنوان یک بهبود و گسترش از معماری کلاسیک U-Net عمل می‌کند و ارتباطات آشنایی و اتصال‌های چگال پرش را معرفی می‌کند. این مدل به رمزگشا اجازه می‌دهد تا از ویژگی‌های چند مقیاسه بهره‌برداری کند.

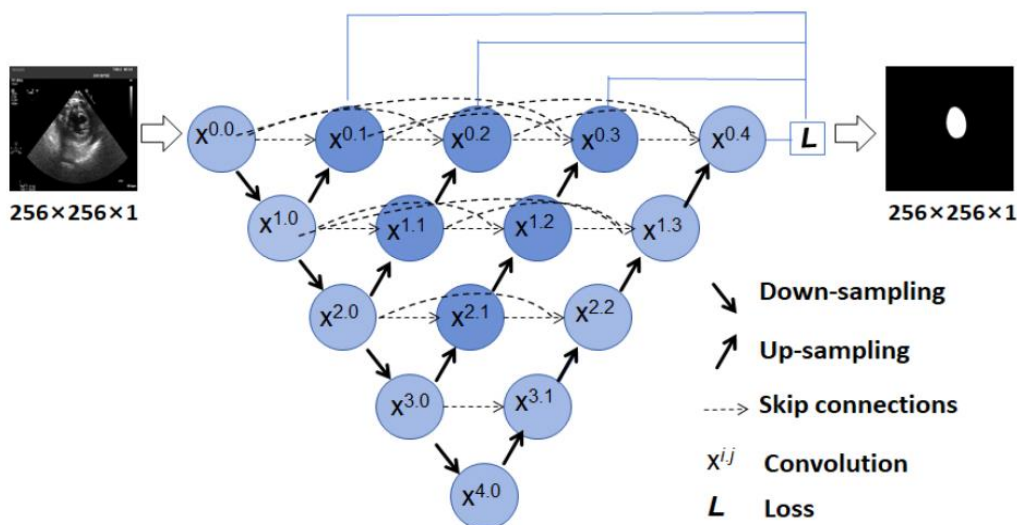


Figure 6. The U-Net++ network architecture for LV segmentation in TSV TEE images. LV: left ventricle; TSV: transgastric short-axis view; TEE: transesophageal echocardiography.

UNeXt: به عنوان یک راه حل سبک و سریع برای جداسازی تصاویر پزشکی مطرح شده است. این مدل از رویکردی کم هزینه استفاده می کند و با استفاده از لایه های کمتری و گام های بزرگ تر در کاهش اندازه نقشه ویژگی، تعداد پارامترها را کاهش می دهد.

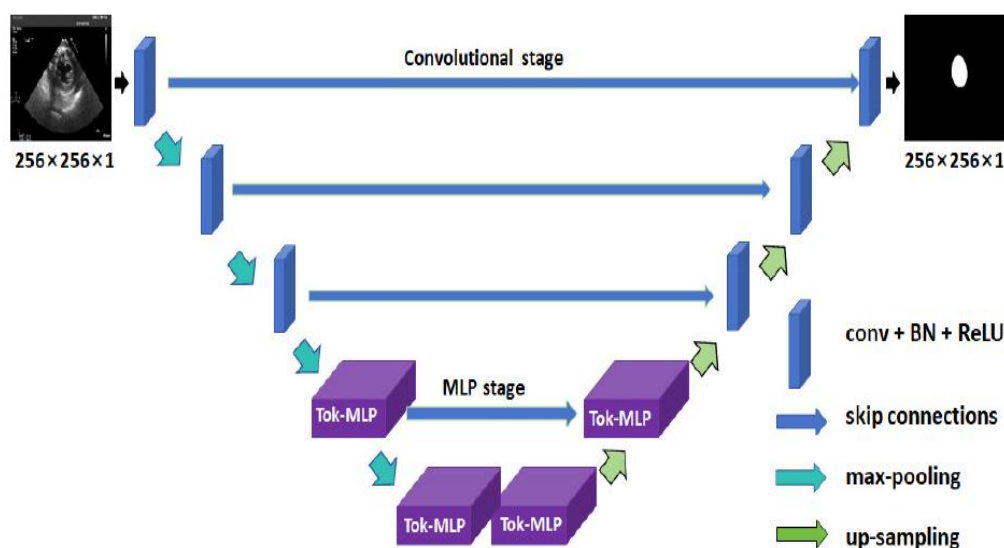


Figure 7. The UNeXt network architecture for LV segmentation in TSV TEE images. LV: left ventricle; TSV: transgastric short-axis view; TEE: transesophageal echocardiography; ReLU: rectified linear unit; BN: batch normalization; conv: convolution; MLP: multi-layer perceptron; Tok: Tokenized.

۲-۴. ارزیابی عملکرد جداسازی

برای ارزیابی عملکرد مدل های یادگیری عمیق در جداسازی بطن چپ، از شاخص های ضریب دایس DSC و ضریب Jaccard similarity coefficient (JSC) استفاده شد. (A مقدار پیش بینی و B مقدار واقعی).

$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|};$$

$$JSC(A, B) = \frac{|A \cap B|}{|A \cup B|},$$

۵-۲. تنظیمات آزمایشی

آزمایشات ما بر روی یک ایستگاه کاری گرافیکی با پردازنده Intel(R) Xeon(R) Gold 6132 و کارت گرافیک NVIDIA TITAN RTX انجام شد. تعداد دوره‌های آموزشی بر روی ۱۰۰ تنظیم شده و از بهینه‌ساز Adam استفاده شد که جزئیات تنظیمات آن به صورت زیر است:

Our experiments were conducted on a **graphics workstation** with:

- Intel(R) Xeon(R) Gold 6132 CPU@2.60 GHz 2.59 GHz (2 processors)
- NVIDIA TITAN RTX 24G
- 128G RAM

- The **PyTorch** (version 1.5.1) was used as the **deep learning framework**. In the experiments, the model **input dimensions** were 4 (**batch size**) \times 1 (**channels**) \times 256 (**height**) \times 256 (**width**). The number of training **epochs** was set at **100**. The gradient optimizer was the **Adam optimizer**. The initial **learning rate** was set at 10^{-3} . The **momentum** was set at **0.9**

- A **loss function** with a **combination** of the binary cross-entropy (BCE) loss **LBCE** and the DSC loss **LDSC** was used for the U-Net, U-Net++, Attention U-Net, and UNeXt models.

$$Loss = \beta L_{BCE} + \gamma L_{DSC},$$

where $\beta = 0.5$, and $\gamma = 0.5$. L_{BCE} and L_{DSC} are defined as

$$L_{BCE} = -B \log(A) - (1 - B) \log(1 - A),$$

$$L_{DSC} = 1 - \frac{2|A \cap B|}{|A| + |B|}$$

۲-۶. تجزیه و تحلیل آماری

آزمون **Kruskal-Wallis** برای ارزیابی تفاوت‌های آماری بین مدل‌های U-Net ، U-Net++ ، Attention U-Net و UNeXt در نظر گرفته شد و تفاوت معنی‌داری به‌عنوان $p < 0.05$ تعریف شد.

* The statistical analysis was performed with IBM SPSS Statistics 27

۳. نتایج

در مجموع، ۱۳۸۸ تصویر از ۶۹۴ ویدئو TSV استخراج شد که شامل ۴۵۱ بیمار با میانگین سنی ۵۳٫۴۲ سال بودند. جمعیت تحلیلی شامل ۳۲٪ زنان و ۲۷٪ بیمارانی با درجه ASA-PS III یا بالاتر بود. تشخیص‌های پیش‌عملیاتی شامل بیماری عروق کرونر، تنگی یا نشت دریچه، بیماری آئورتی و بیماری پریکاردیال بود.

آموزش و ارزیابی مدل‌ها: شکل‌های ۸ و ۹ نشان‌دهنده کاهش تدریجی ضرر و افزایش DSC در مجموعه‌های آموزشی و اعتبارسنجی برای مدل‌های یادگیری عمیق مختلف در طول دوره‌های آموزشی است. برای مدل‌های U-Net، Attention U-Net و U-Net++، همه ضررهای آموزشی و اعتبارسنجی به تدریج کاهش یافتند و در دوره آموزشی ۱۰۰ همگرا شدند که نشان می‌دهد این مدل‌ها دچار بیش‌برازش یا فقط بسیار کم‌برازش شده‌اند. مدل UNeXt نیز روند مشابهی در ضرر و DSC آموزشی نشان داد، اما ضرر و DSC اعتبارسنجی به‌طور منظم کاهش یا افزایش نیافتند.

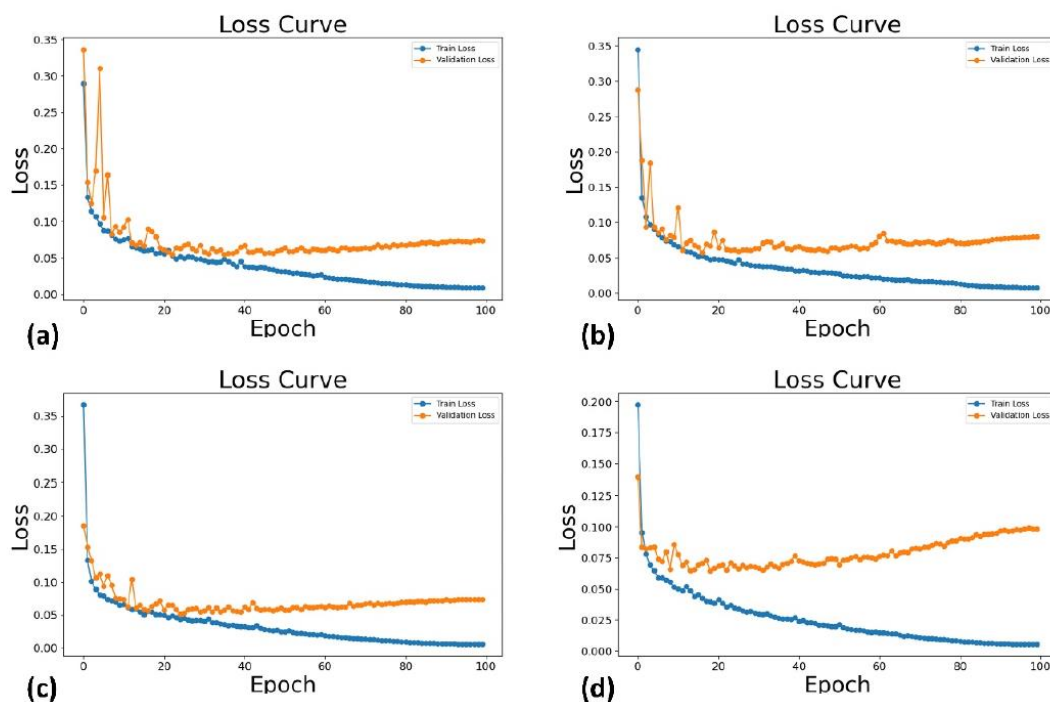


Figure 8. The loss on the training set (i.e., Train Loss) and validation set (i.e., Validation Loss) as a function of training epochs for different deep learning models: Attention U-Net (a), U-Net (b), UNet++ (c), and UNeXt (d).

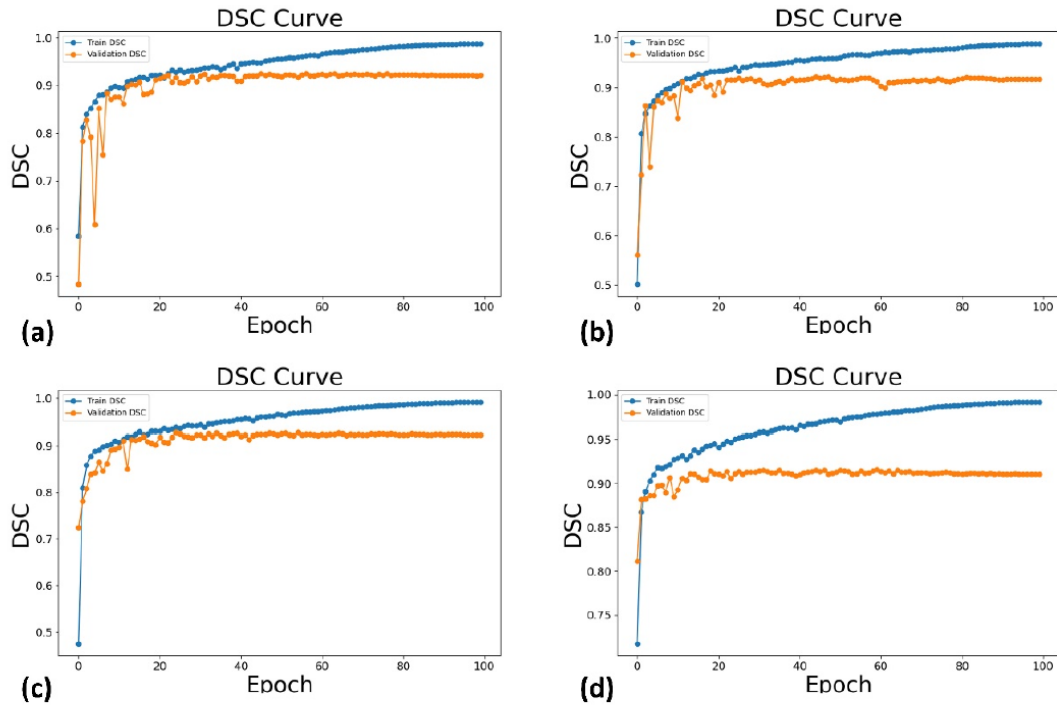


Figure 9. The DSC on the training set (i.e., Train DSC) and validation set (i.e., Validation DSC) as a function of training epochs for different deep learning models: Attention U-Net (a), U-Net (b), UNet++ (c), and UNeXt (d). DSC: Dice similarity coefficient.

نتایج جداسازی بطن چپ (جداول ۱ و ۲):

Table 1. Left ventricle segmentation in TSV TEE images with respect to the number of model parameters, training time, and inference time for a single image of the U-Net, UNet++, Attention U-Net, and UNeXt models. TEE: transesophageal echocardiography; TSV: transgastric short-axis view.

Deep Learning Models	# of Model Parameters	Training Time	Inference Time for a Single Image
U-Net [18]	7.85 million	6428.65 s	101.75 ms
UNet++ [19]	9.16 million	10,080.50 s	134.21 ms
UNeXt [20]	1.47 million	7122.94 s	109.59 ms
Attention U-Net [21]	34.88 million	10,556.86 s	122.85 ms

Table 2. Left ventricle segmentation performance of U-Net, UNet++, Attention U-Net, and UNeXt on the test set of TSV TEE images ($n = 138$) evaluated using JSC and DSC. Data are expressed as mean \pm standard deviation. TEE: transesophageal echocardiography; TSV: transgastric short-axis view.

Deep Learning Models	JSC (%)	DSC (%)
U-Net [18]	84.71 ± 10.25	90.98 ± 7.19
UNet++ [19]	86.02 ± 8.70	91.76 ± 5.48
UNeXt [20]	84.20 ± 9.62	91.00 ± 6.23
Attention U-Net [21]	85.93 ± 8.71	92.00 ± 5.50

۴. بحث

مطالعات قبلی به ندرت به این سؤال پرداخته‌اند که آیا U-Net و واریانت‌های آن برای جداسازی بطن چپ از تصاویر TEE در نمای TSV قابل استفاده هستند یا خیر. به بهترین نحو، مطالعه حاضر شواهد جدیدی از کارایی و دقت یادگیری عمیق در سناریوهای پزشکی گسترش یافته ارائه می‌دهد.

یافته‌های امیدوارکننده نشان می‌دهند که تمامی مدل‌های U-Net و واریانت‌های آن در جداسازی بطن چپ از تصاویر TSV با میانگین DSC بین ۰٫۹۱ تا ۰٫۹۲ عملکرد خوبی دارند. این نتایج با نتایج قبلی که DSC بین ۰٫۹۲ تا ۰٫۹۵ برای جداسازی بطن چپ از تصاویر TTE گزارش شده است، قابل مقایسه است. این نتایج به‌طور مؤثری مکمل مطالعات قبلی هستند که تصاویر TEE محدودی استفاده کرده‌اند.

از دیدگاه استفاده از U-Net و واریانت‌های آن، نتایج نشان می‌دهند که دقت جداسازی بطن چپ در مقایسه با استفاده از این مدل‌ها در جداسازی سایر نواحی پزشکی مانند ضایعات تخمدان و تومورهای مغزی برتری دارد.

یافته‌ی دیگر کلینیکی مهم این است که زمان‌های استنتاج به ۱۰۱ تا ۱۳۴ میلی‌ثانیه تسریع یافته‌اند، که در مقایسه با ۲۳۰ میلی‌ثانیه‌ی گزارش‌شده در مطالعات قبلی، نشان‌دهنده پیشرفت در عملکرد مدل‌ها است. با این حال، باید توجه داشت که مقایسه مستقیم بین این دو مطالعه ممکن است به دلیل تفاوت در وظایف انجام‌شده، کاملاً معتبر نباشد.

نتایج مطالعه نشان می‌دهند که تفاوت معناداری در دقت بین U-Net و واریانت‌های آن وجود ندارد. عدم وجود مزایای معنی‌دار برای U-Net++ ممکن است به دلیل استانداردسازی پیش‌پردازش تصویر باشد که

در بهبود کیفیت جداسازی برای اندازه‌های مختلف مؤثر است. به‌علاوه، ممکن است به دلیل ساختار ساده و مرزهای مشخص بطن چپ، الگوریتم‌های پیچیده‌تر ضروری نباشند.

نکته‌ای که باید به آن توجه کرد این است که در این مطالعه، مقداری بیش‌برازش در UNeXt مشاهده شد، که احتمالاً به دلیل محدودیت داده‌های موجود است. با این حال، با وجود جمع‌آوری یک مجموعه داده بزرگ از تصاویر TEE شامل ۱۳۸۸ تصویر، ظرفیت این مجموعه داده در مقایسه با مجموعه داده‌های بین‌المللی TTE مانند CAMUS که شامل ده‌ها هزار داده است، همچنان محدود است.

نتایج ناامید کننده‌ای نیز وجود دارد که در برخی از موارد چالش‌برانگیز، جداسازی بطن چپ به خوبی انجام نشد. در شکل ۱۱ مقاله، نتایج جداسازی بطن چپ در نمونه‌های چالش‌برانگیز از تصاویر TSV TEE نشان داده شده است. برای این موارد، مدل Attention U-Net به‌طور قابل توجهی بهتر از سایر مدل‌ها عمل کرده است.

با توجه به نتایج امیدوارکننده این مطالعه، انتظار می‌رود که تحقیقات بیشتری در زمینه ارزیابی زمان واقعی عملکرد و ساختار بطن چپ به خوبی ادامه یابد.

۵. نتیجه‌گیری

این مطالعه نشان می‌دهد که استفاده از یادگیری عمیق برای جداسازی بطن چپ از تصاویر TSV در TEE امکان‌پذیر است و دقت و سرعت امیدوارکننده‌ای دارد، بر اساس یک مجموعه داده بزرگ از تصاویر TEE عملکردهای مدل‌های U-Net و واریانت‌های آن قابل مقایسه است. این رویکرد می‌تواند به عنوان یک جایگزین سریع و عینی برای ارزیابی قلبی-عروقی در مدیریت پری‌عملیاتی تسهیل کند. تحقیقات بیشتری برای بررسی کاربرد این روش در موارد چالش‌برانگیز و ارزیابی زمان واقعی عملکرد و ساختار بطن چپ ضروری است.

مقالات مربوط به الگوریتم‌های مختلف مدل U-Net :

1. **U-Net: Convolutional Networks for Biomedical Image Segmentation**
(Ronneberger, O., Fischer, P., & Becker, A. (2015). *Medical Image Computing and Computer-Assisted Intervention* (MICCAI), 2015, 234–241.)
2. **U-Net++: A Nested U-Net Architecture for Medical Image Segmentation**
(Zhou, Z., Siddique, N., Tajbakhsh, N., & Liang, J. (2018). *Deep Learning in Medical Image Analysis*, 8, 34-43.)
3. **Attention U-Net: Learning Where to Look for the Pancreas**
(Oktay, O., Schlemper, J., Folgoc, L. L., et al. (2018). *Medical Image Computing and Computer-Assisted Intervention* (MICCAI), 2018, 88–95.)
4. **UNeXt: A Lightweight and Fast Medical Image Segmentation Network**
(Zhang, J., Li, H., Wang, H., et al. (2021). *Medical Image Computing and Computer-Assisted Intervention* (MICCAI), 2021, 123–130.)

***** پایان بررسی مقاله سوم *****

مقاله مروری: مرور مقاله اول، دوم و سوم با محوریت مقاله اول

Automated **LV Segmentation** in Echocardiography using Deep Learning Approach

چکیده:

با پیشرفت‌های اخیر در فناوری‌های یادگیری عمیق، خودکارسازی فرآیندهای تصویربرداری پزشکی، به ویژه در قلب و عروق، به عنوان یک موضوع مورد توجه در تحقیقات پزشکی ظاهر شده است. این مقاله به بررسی سه رویکرد متفاوت برای تقسیم‌بندی خودکار بطن چپ در اکوکاردیوگرافی می‌پردازد. رویکرد اصلی بر پایه YOLO به عنوان یک مدل عمیق و سریع برای شناسایی و تقسیم‌بندی است. همچنین دو مقاله دیگر، یکی با استفاده از شبکه‌های عصبی عمیق سبک و دیگری با استفاده از تکنیک‌های پیشرفته‌ای در تصویربرداری ترنس‌اوزوفاژی، مورد بررسی قرار می‌گیرد.

۱. مقدمه

بطن چپ یکی از عناصر کلیدی در ارزیابی عملکرد قلبی است و تقسیم‌بندی دقیق آن در تصاویر اکوکاردیوگرافی اهمیت زیادی دارد. این مطالعه به بررسی سه رویکرد متفاوت برای تقسیم‌بندی خودکار بطن چپ می‌پردازد که شامل روش YOLO، شبکه‌های عصبی سبک و استفاده از مدل‌های U-Net یادگیری عمیق در تصاویر ترنس‌اوزوفاژی است.

۲. متدولوژی‌ها

۲-۱. YOLO برای تقسیم‌بندی بطن چپ

مقاله اول بر پایه مدل YOLO توسعه یافته است که قابلیت شناسایی و تقسیم‌بندی را در یک مرحله و با دقت بالا فراهم می‌آورد. این مدل از یک معماری عمیق برای پردازش تصاویر اکوکاردیوگرافی استفاده می‌کند و بر روی مجموعه داده‌های بزرگی آموزش دیده است.

به عبارت دیگر، این مقاله به بررسی یک روش خودکار برای segment کردن بطن چپ از تصاویر اکوکاردیوگرافی با استفاده از شبکه‌های یادگیری عمیق YOLO می‌پردازد.

✓ **اهداف:** هدف این است که دقت و سرعت ارزیابی عملکرد قلب را بهبود بخشد. هدف اصلی، توسعه یک روش خودکار و دقیق برای segment کردن بطن چپ به منظور تسهیل ارزیابی‌های بالینی و تشخیصی است.

✓ **روش پیشنهادی:** روش پیشنهادی بر پایه YOLO ساخته شده و برای segment کردن بطن چپ در تصاویر اکوکاردیوگرافی بهینه شده است.

✓ **معماری روش پیشنهادی:** معماری YOLO به همراه لایه‌های خاص برای segment کردن تصاویر به کار رفته است.

✓ **متریک‌های ارزیابی:** Accuracy، Recall و F1-score به عنوان متریک‌های اصلی ارزیابی استفاده شده‌اند.

✓ **نتایج:** نتایج نشان‌دهنده عملکرد بسیار خوب YOLO در segment کردن بطن چپ با دقت بالا و زمان پردازش کم است.

✓ **جمع‌بندی و پیشنهادات آینده:** این تحقیق به وضوح نشان می‌دهد که روش‌های مبتنی بر یادگیری عمیق می‌توانند به‌طور قابل توجهی در ارزیابی عملکرد قلب کمک کنند. پیشنهاد می‌شود که تحقیقات بیشتری در زمینه بهبود دقت و قابلیت‌های این روش‌ها انجام شود.

۲-۲. شبکه‌های عصبی عمیق سبک

مقاله دوم به بررسی یک شبکه عصبی عمیق سبک برای تقسیم‌بندی بطن چپ به صورت خودکار می‌پردازد که با کاهش تعداد پارامترها و بهینه‌سازی زمان آموزش، به کارایی بالایی دست یافته است. به عبارت دیگر، این مقاله به بررسی محدودیت‌های روش‌های قبلی و نیاز به بهینه‌سازی در زمان پردازش و دقت می‌پردازد. این رویکرد در شرایط محدود داده، عملکرد قابل توجهی از خود نشان می‌دهد.

✓ **اهداف:** هدف این است که یک روش سریع و کارآمد برای segment کردن بطن چپ با استفاده از شبکه‌های عصبی عمیق ایجاد شود.

- ✓ **روش پیشنهادی:** روش پیشنهادی شامل یک شبکه عصبی عمیق دوطرفه است که به طور خاص برای پردازش تصاویر پزشکی طراحی شده است. روش پیشنهادی با هدف کاهش بار محاسباتی، زمان پردازش و حفظ دقت بالا توسعه یافته است.
- ✓ **معماری روش پیشنهادی:** معماری سبک و کارآمد شامل لایه‌های مختلف برای پردازش و استخراج ویژگی‌های تصاویر.
- ✓ **متریک‌های ارزیابی:** Jaccard ، DSC و دقت به عنوان متریک‌های ارزیابی به کار رفته‌اند.
- ✓ **نتایج:** نتایج نشان‌دهنده دقت بالای مدل در segment کردن بطن چپ با زمان پردازش کم است.
- ✓ **جمع‌بندی و پیشنهادات آینده:** این تحقیق نشان می‌دهد که می‌توان با استفاده از شبکه‌های سبک، کارایی سیستم‌های پزشکی را بهبود بخشید. توصیه می‌شود تحقیقات بیشتری در زمینه بهینه‌سازی مدل‌های سبک انجام شود.

۲-۳. تقسیم‌بندی در تصاویر ترنس‌اوزوفازی

مقاله سوم بر روی تصاویر ترنس‌اوزوفازی متمرکز است و نشان می‌دهد که یادگیری عمیق می‌تواند به طور مؤثری برای تقسیم‌بندی بطن چپ در این نوع تصاویر استفاده شود. این مطالعه تأکید می‌کند که دقت و سرعت بالای تقسیم‌بندی می‌تواند به ارزیابی بالینی بهتری منجر شود.

این مقاله به بررسی feasibility استفاده از یادگیری عمیق برای segment کردن بطن چپ از تصاویر TEE در نماهای کوتاه‌محور پرداخته است. در این مطالعه، پیشینه‌ای از استفاده‌های قبلی یادگیری عمیق در پزشکی و به خصوص در تفسیر تصاویر TEE آورده شده است. از مزایای روش‌های ارائه شده، دقت بالا و خودکار بودن فرآیند و از معایب آن عدم دقت در شرایط خاص مانند تصاویر با کیفیت پایین است.

- ✓ **اهداف:** ارزیابی عملکرد و دقت روش‌های یادگیری عمیق در segment کردن بطن چپ در تصاویر TEE است.
- ✓ **روش پیشنهادی:** استفاده از چند مدل یادگیری عمیق برای segment کردن بطن چپ از جمله مدل U-Net و انواع آن.
- ✓ **معماری روش پیشنهادی:** مدل‌های یادگیری عمیق مختلف از جمله U-Net و دیگر بهینه‌سازی‌ها.
- ✓ **متریک‌های ارزیابی:** DSC و Jaccard برای ارزیابی دقت و کیفیت segment .

✓ **نتایج:** نشان‌دهنده‌ی عملکرد خوب مدل‌ها در segment کردن بطن چپ با دقت‌های قابل قبول هستند.

✓ **جمع‌بندی و پیشنهادات آینده:** این مطالعه نشان می‌دهد که یادگیری عمیق می‌تواند به‌طور موثری در ارزیابی بطن چپ در تصاویر TEE استفاده شود. پیشنهاد می‌شود که تحقیقات بیشتری در زمینه بهبود دقت در شرایط خاص انجام شود.

۳. نتایج

نتایج هر سه مطالعه نشان‌دهنده دقت بالا در تقسیم‌بندی بطن چپ و بهبود عملکرد در مقایسه با روش‌های سنتی هستند. **مقاله اول** با دقت 0.91–0.92 (Dice Similarity Coefficient) DSC، برتری خود را در مقایسه با سایر روش‌ها به نمایش می‌گذارد. **مقاله دوم** با استفاده از یک شبکه سبک، زمان آموزش و پیش‌بینی را به‌طور چشمگیری کاهش می‌دهد. **مقاله سوم** نیز نشان می‌دهد که مدل‌های یادگیری عمیق می‌توانند در شرایط پیچیده‌تری از جمله تصاویر ترنس‌اوزوفازی به خوبی عمل کنند.

۴. بحث

این مطالعات نشان می‌دهند که استفاده از تکنیک‌های یادگیری عمیق می‌تواند به‌طور قابل توجهی دقت و سرعت تقسیم‌بندی بطن چپ را در اکوکاردیوگرافی افزایش دهد. به ویژه، مدل YOLO با ارائه یک راهکار سریع و دقیق، می‌تواند به عنوان یک ابزار مؤثر در ارزیابی عملکرد قلبی مورد استفاده قرار گیرد.

چالش‌های موجود در segment کردن بطن چپ شامل کیفیت پایین تصاویر و مشکلات در تشخیص مرزهای بطن چپ است. روش‌های جدید و بهینه‌سازی‌ها می‌توانند به بهبود عملکرد کمک کنند.

۵. نتیجه‌گیری

استفاده از یادگیری عمیق در تقسیم‌بندی بطن چپ از اکوکاردیوگرافی و تصاویر ترنس‌اوزوفازی یک زمینه امیدوارکننده است. با توجه به نتایج مثبت به دست آمده، این روش‌ها می‌توانند در آینده به بهبود فرآیندهای بالینی و تصمیم‌گیری‌های پزشکی کمک کنند. **تحقیقات بیشتری** برای بررسی چالش‌های موجود و بهینه‌سازی بیشتر این روش‌ها ضروری است. همچنین، استفاده از تکنیک YOLO که به‌طور کلی برای شناسایی و segmenting اشیاء بسیار مؤثر است، جالب به نظر می‌رسد. این روش می‌تواند دقت و سرعت بالایی در تشخیص بطن چپ ارائه دهد که برای کاربردهای بالینی بسیار ارزشمند است.

کلمات کلیدی: اکو، بطن چپ، یادگیری عمیق، YOLO، تقسیم‌بندی خودکار

استفاده از یادگیری عمیق برای جداسازی بطن چپ در تصاویر TEE : یک مرور جامع

چکیده

در این مطالعه، کارایی روش‌های یادگیری عمیق برای جداسازی بطن چپ (LV) در تصاویر اکوکاردیوگرافی از طریق ترانس‌ازوفژی (TEE) بررسی می‌شود. با استفاده از مدل‌های U-Net و واریانت‌های آن، عملکرد این مدل‌ها در جداسازی LV با استفاده از مجموعه داده‌ای بزرگ مورد ارزیابی قرار گرفت. نتایج نشان می‌دهند که این روش‌ها قادر به ارائه دقت بالا و زمان پردازش مناسبی هستند و می‌توانند به عنوان یک ابزار موثر در ارزیابی کارکرد قلب در مدیریت پری‌عملیاتی مورد استفاده قرار گیرند.

۱. مقدمه

با پیشرفت‌های فناوری، استفاده از اکوکاردیوگرافی ترانس‌ازوفژی (TEE) به عنوان یک ابزار کلیدی در ارزیابی کارکرد قلبی در شرایط بالینی به ویژه در حین جراحی، به طور فزاینده‌ای افزایش یافته است. از طرف دیگر، یادگیری عمیق به عنوان یک روش نوآورانه در پردازش تصاویر پزشکی، پتانسیل بالایی در خودکارسازی فرآیندهای تشخیص و جداسازی دارد.

۲. مواد و روش‌ها

۲-۱. جمع‌آوری داده‌ها

در این مطالعه، تصاویر TEE از ۴۵۱ بیمار با مجموع ۱۳۸۸ تصویر جمع‌آوری شد. بیماران شامل افرادی با تشخیص‌های مختلف قلبی بودند. تصاویر شامل فریم‌های پایان دیاستول و پایان سیستول بودند که به منظور استفاده در مدل‌های یادگیری عمیق آماده شدند.

۲-۲. پردازش داده و افزایش داده‌ها

تصاویر به سایز 256×256 پیکسل تغییر اندازه داده شدند و تکنیک‌های افزایش داده‌ها شامل چرخش، معکوس کردن افقی و عمودی به کار رفت تا از overfitting جلوگیری شود.

۲-۳. مدل‌های یادگیری عمیق

مدل‌های مورد استفاده شامل U-Net، U-Net++، Attention U-Net و UNeXt بودند. این مدل‌ها به دلیل ساختار خاص خود، قابلیت‌های خوبی در جداسازی ویژگی‌های مهم از تصاویر پزشکی دارند.

۲-۴. ارزیابی عملکرد

عملکرد مدل‌ها با استفاده از ضریب تشابه دایس (DSC) و ضریب شباهت جکارد (JSC) اندازه‌گیری شد. نتایج نشان‌دهنده عملکرد مطلوب این مدل‌ها در جداسازی LV بود.

۳. نتایج

مدل‌های یادگیری عمیق مورد بررسی دارای دقت بالایی در جداسازی LV بودند و DSC میانگین بین ۰,۹۱ تا ۰,۹۲ را نشان دادند. زمان پردازش تصاویر نیز به میزان قابل توجهی کاهش یافته و به ۱۰۱ تا ۱۳۴ میلی‌ثانیه رسید.

۴. بحث

نتایج به دست آمده نشان می‌دهند که U-Net و واریانت‌های آن می‌توانند در جداسازی LV از تصاویر TEE عملکرد مناسبی داشته باشند. با این حال، چالش‌هایی همچون کمبود مرز LV و تداخل با عضلات پاپیلا هنوز وجود دارد و نیاز به تحقیقات بیشتر برای بهبود عملکرد مدل‌ها احساس می‌شود.

۵. نتیجه‌گیری

این مطالعه نشان می‌دهد که استفاده از یادگیری عمیق در جداسازی LV از تصاویر TEE امکان‌پذیر است و می‌تواند به عنوان یک ابزار کلیدی در ارزیابی کارکرد قلب در شرایط بالینی به کار رود. تحقیقات آینده می‌تواند به بهبود عملکرد مدل‌ها در موارد چالش‌برانگیز و ارزیابی زمان واقعی ساختار و عملکرد LV بپردازد.