



دانشکده مهندسی کامپیوتر

گزارش نهایی پروژه

قطعه بندی و طبقه بندی تومورهای سرطان سینه در تصاویر ماموگرافی  
با استفاده از شبکه های مولد رقابتی و شبکه های عصبی همگشتی

در رشته مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیک

نام دانشجو:

پوریا محمدی نسب

نام دستیار:

امیررضا فاتح

استاد:

دکتر محمدرضا محمدی

تیر ماه ۱۴۰۱

## چکیده

بررسی تصاویر ماموگرافی برای پیدا کردن تومورهای سینه یکی از چالش‌های اساسی حتی برای متخصصین پزشکی و رادیولوژیست‌ها است. بنابراین استفاده از روش‌های تحلیل تصاویر برای شناسایی تومورهای سینه مورد نیاز است، زیرا این روش‌ها می‌توانند با استخراج اطلاعات مهم و حیاتی صحت تشخیص را بالا ببرند. در این پروژه برای وظیفه‌ی قطعه‌بندی تومور از چند روش مختلف از جمله معماری برپایه‌ی شبکه‌های مولد رقابتی شرطی، معماری LinkNet و معماری U-net استفاده شده است. همچنین از معماری‌های مختلفی برای طبقه‌بندی شکل تومور استفاده شده است. همچنین برای طبقه‌بندی نوع سرطان (خوش‌خیم، بدخیم، نرمال) نیز از چند معماری مختلف شبکه‌های عصبی همگشتی استفاده شده است که عموماً از یادگیری انتقالی برای انجام این وظیفه استفاده می‌کنند. نتایج نشان می‌دهد در وظیفه قطعه‌بندی تومور، شبکه U-net نتایجی مشابه شبکه مولد رقابتی معرفی شده در مقاله اصلی دارد. اما در وظیفه طبقه‌بندی شکل تومور معماری پیشنهادی ما با دقت ۴۷/۴۳٪ به طور میانگین از مدل‌های دیگر کمی بهتر عمل میکند.

**واژه‌های کلیدی:** سرطان سینه، شبکه مولد رقابتی شرطی، شبکه عصبی همگشتی، قطعه‌بندی و طبقه‌بندی تومور

## فهرست مطالب

۱	فصل ۱: مقدمه
۵	فصل ۲: مروری بر ادبیات موضوع
۶	۲-۱- قطعه‌بندی تومور .....
۹	۲-۲- طبقه‌بندی شکل تومور .....
۱۰	۲-۳- طبقه‌بندی نوع تومور .....
۱۲	فصل ۳: روش انتخابی حل مسئله
۱۳	۳-۱- قطعه‌بندی تومور .....
۱۳	۳-۱-۱- مدل cGAN پیشنهادی .....
۱۵	۳-۱-۲- مدل Link-Net .....
۱۷	۳-۱-۳- مدل U-Net .....
۱۷	۳-۲- طبقه‌بندی شکل تومور .....
۱۷	۳-۲-۱- مدل پیشنهادی مقاله .....
۱۸	۳-۲-۲- مدل ResNet50 .....
۱۹	۳-۲-۳- مدل VGG16 .....
۲۰	۳-۲-۴- مدل MobileNet .....
۲۱	۳-۲-۵- مدل پیشنهادی جدید .....
۲۲	۳-۳- طبقه‌بندی نوع تومور .....
۲۲	۳-۳-۱- مدل VGG19 .....
۲۲	۳-۳-۲- مدل ResNet50 .....
۲۲	۳-۳-۳- مدل DenseNet121 .....
۲۴	۳-۳-۴- مدل پیشنهادی جدید .....
۲۵	فصل ۴: نتایج و تفسیر آنها
۲۶	۴-۱- مجموعه داده‌گان .....
۲۶	۴-۱-۱- مجموعه داده‌گان INbreast .....
۲۶	۴-۱-۲- مجموعه داده‌گان DDSM .....
۲۷	۴-۲- نتایج قطعه‌بندی تومور .....
۲۹	۴-۳- نتایج طبقه‌بندی شکل تومور .....
۳۳	۴-۴- نتایج طبقه‌بندی نوع تومور .....
۳۵	فصل ۵: جمع‌بندی و پیشنهادها
۳۶	۵-۱- جمع‌بندی .....

۲-۵- پیشنهادها ..... ۳۶

مراجع ۳۷

## فهرست اشکال

۳	..... شکل (۱-۱) خط لوله مقاله اصلی
۶	..... شکل (۱-۲) معماری U-Net
۷	..... شکل (۲-۲) معماری Seg-Net
۸	..... شکل (۳-۲) معماری مدل معرفی شده در مرجع ۲۳
۱۰	..... شکل (۴-۲) معماری مدل Point-Net
۱۳	..... شکل (۱-۳) (a) ماموگرافی کامل (b) ماموگرافی عریض (c) ماموگرافی تنگ
۱۴	..... شکل (۲-۳) جزئیات معماری cGAN
۱۵	..... شکل (۳-۳) نحوه استفاده دو تابع ضرر Dice و BCE
۱۶	..... شکل (۴-۳) (a) معماری Link-Net (b) ساختار بلوک رمزنگار (c) ساختار بلوک رمزگشا
۱۸	..... شکل (۵-۳) معماری CNN پیشنهادی مقاله
۱۹	..... شکل (۶-۳) معماری مدل ResNet
۲۰	..... شکل (۷-۳) معماری VGG16
۲۱	..... شکل (۸-۳) عملکرد کانولوشن depth-wise
۲۱	..... شکل (۹-۳) مدل پیشنهادی طبقه‌بندی شکل تومور
۲۲	..... شکل (۱۰-۳) مدل VGG19
۲۳	..... شکل (۱۱-۳) مدل شبکه DenseNet
۲۴	..... شکل (۱۲-۳) مدل پیشنهادی طبقه‌بندی نوع تومور
۲۸	..... شکل (۱-۴) نتایج کیفی قطعه‌بندی
۲۹	..... شکل (۲-۴) نمودارهای دقت و ضرر مدل پیشنهادی مقاله
۳۰	..... شکل (۳-۴) نمودار دقت و ضرر مدل ResNet
۳۰	..... شکل (۴-۴) نمودار دقت و ضرر مدل Mobilenet
۳۱	..... شکل (۵-۴) نمودار دقت و ضرر مدل VGG16
۳۱	..... شکل (۶-۴) نمودار دقت و ضرر مدل پیشنهادی جدید
۳۲	..... شکل (۷-۴) ماتریس همبستگی مدل‌های ارزیابی شده
۳۳	..... شکل (۸-۴) نمودار دقت و ضرر مدل DenseNet
۳۳	..... شکل (۹-۴) نمودار دقت و ضرر مدل VGG19
۳۴	..... شکل (۱۰-۴) نمودار دقت و ضرر مدل ResNet50
۳۴	..... شکل (۱۱-۴) نمودار دقت و ضرر مدل پیشنهادی

## فهرست جداول

۱۶	جدول (۱-۳) نقشه ویژگی‌های ورودی و خروجی .....
۱۹	جدول (۲-۳) معماری Res-Net با عمق‌های مختلف .....
۲۰	جدول (۳-۳) مشخصات مدل‌های مختلف .....
۲۳	جدول (۳-۴) جزئیات معماری DenseNet121 .....
۲۷	جدول (۱-۴) جدول عملکرد مدلهای بررسی شده در پروژه .....
۳۲	جدول (۴-۲) جدول نتایج مدلهای بررسی شده .....
۳۴	جدول (۳-۴) دقت‌های بدست آمده از مدل‌ها .....

# فصل ۱:

## مقدمه

سرطان سینه، به عنوان یکی از رایج ترین سرطان های کشنده در خانم ها شناخته میشود [۱]. ماموگرافی روشی است که به تشخیص زود هنگام سرطان بسیار کمک میکند و باعث میشود نرخ مرگ و میر به دلیل سرطان سینه به طور چشم گیری کاهش یابد [۲]. در بین تمام ناهنجاری ها<sup>۱</sup> در تصاویر سینه که هر کدام میتوانند نشان دهنده وجود سرطان باشند، توده ها مهم ترین یافته ها محسوب میشوند. همچنین اطلاعات ساختاری تومورها میتواند تعیین کننده خوش خیم بودن یا بدخیم بودن سرطان باشد. شکل تومورهای سینه در یکی از دسته بندی های زیر قرار میگیرد: (۱) نامنظم (۲) گرد (۳) بیضی شکل (۴) ابرمانند [۳].

شناسایی با کمک کامپیوتر<sup>۲</sup> که به اختصار سیستم های CAD نیز نام برده میشوند سیستم هایی هستند که در تفسیر تصاویر پزشکی به پزشکان یاری میرسانند. این سیستم ها تصاویر را پردازش کرده و بخش های مشکوک آن، مانند بیماری های احتمالی را مشخص میکنند [۴]. سیستم های CAD میتوانند به متخصصین رادیولوژی کمک کنند تا تومور های سینه را شناسایی و مرز های تومور را مشخص کنند. البته با وجود پیشرفت های چشم گیر سیستم های CAD در سال های اخیر، همچنان تشخیص تومورهای سینه به دلیل پایین بودن نسبت توان سیگنال به توان نویز<sup>۳</sup> و همچنین تنوع شکل تومور ها، یک مسئله چالش برانگیز است. تا به حال چندین مقاله بر روی استخراج ویژگی ها و ترکیب آنها با استفاده از شبکه های کانولوشنی عمیق کار کردند تا عملکرد طبقه بندی تومورهای سینه را افزایش دهند [۵].

علاوه بر کاربردهای ذکر شده در بالا، تعیین کردن نوع مولکول سرطانی در تصاویر ماموگرافی حتی برای متخصصین این زمینه نیز کار بسیار دشواری است. آگاهی از نوع مولکول سرطانی برای تجویز بهترین درمان بسیار حائز اهمیت است [۶]–[۸]. مطالعات اخیر نشان داده است که ارتباط ضعیفی بین شکل بصری تومور و نوع مولکول سرطان آن وجود دارد. در مقاله [۹] با استفاده از شبکه های عصبی کانولوشنی نوع مولکول تومور را از روی بافت استخراج شده از تومور طبقه بندی کردند و به دقت ۶۷٪ دست یافتند. البته واضح است که نمیتوان به طور موثر نوع مولکول تومور را فقط از روی بافت تومور در تصویر مشخص کرد و عوامل دیگری نیز در این موضوع نقش دارند.

در این پروژه ما به سه وظیفه ی مختلف میپردازیم. این سه وظیفه شامل طبقه بندی تومور، طبقه بندی شکل تومور و طبقه بندی نوع تومور می باشد. در ابتدا روش پیشنهادی مقاله اصلی [۱۰] که فقط برای دو وظیفه ی اول روشی پیشنهاد داده بود را ارزیابی میکنیم. در این مقاله برای طبقه بندی تومور از یک معماری بر پایه ی شبکه های مولد رقابتی شرطی و برای طبقه بندی شکل تومور از یک معماری ساده ی همگشتی استفاده شده

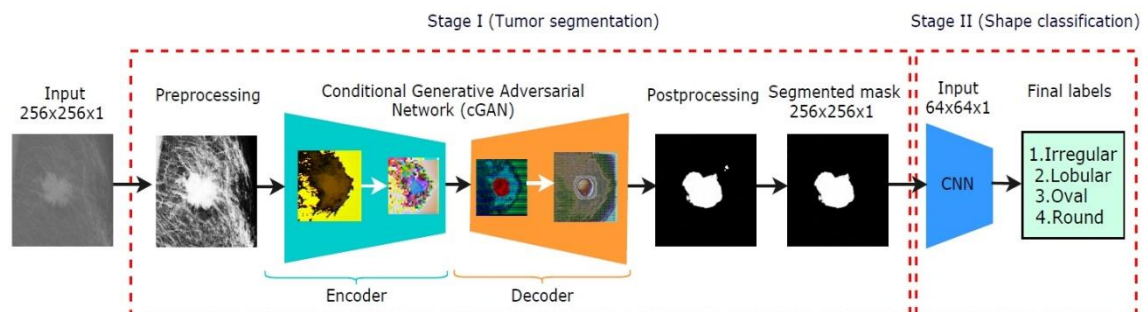
<sup>1</sup> Abnormalities

<sup>2</sup> Computer Aided Diagnosis (CAD)

<sup>3</sup> Signal-to-noise ratio



بود. در مرحله اول، روش پیشنهادی تومور را قطعه بندی و یک ماسک دودویی<sup>۱</sup> تولید میکند. در مرحله دوم، ماسک دودویی ایجاد شده در مرحله قبل به یکی از کلاس های شکل تومور (نامنظم، گرد، بیضی شکل و ابری شکل) طبقه بندی میشود. شکل (۱-۱) شمای کلی خط لوله مقاله اصلی را نشان میدهد.



شکل (۱-۱) خط لوله مقاله اصلی

پس از ارزیابی روش پیشنهادی مقاله اصلی، تلاش بر بهبود نتایج و مقایسه نتایج چند مدل مختلف در هر وظیفه داریم. در این پروژه برای مقایسه cGAN پیشنهاد شده، از دو شبکه دیگر برای قطعه بندی استفاده میکنیم:

- شبکه U-Net
- شبکه Link-Net

همچنین برای ارزیابی مدل پیشنهادی مقاله در قسمت طبقه بندی شکل تومور که وظیفه ای بسیار دشوار برای ماشین است از سه مدل شناخته شده و یک مدل پیشنهادی استفاده میکنیم. مدل های شناخته شده عبارتند از:

- ResNet50
- VGG16
- MobileNet

<sup>1</sup> Binary mask

در آخرین مرحله از پروژه وظیفه جدیدی برای مجموعه دادگان استفاده شده تعریف میکنیم که با گرفتن تصویر کامل ماموگرافی تعیین کند نوع تومور (خوش خیم، بدخیم و نرمال) چیست. برای این وظیفه نیز مدلی ساده پیشنهاد شده و با مدل های معروف زیر مقایسه شده است.

- ResNet50
- VGG19
- DenseNet121

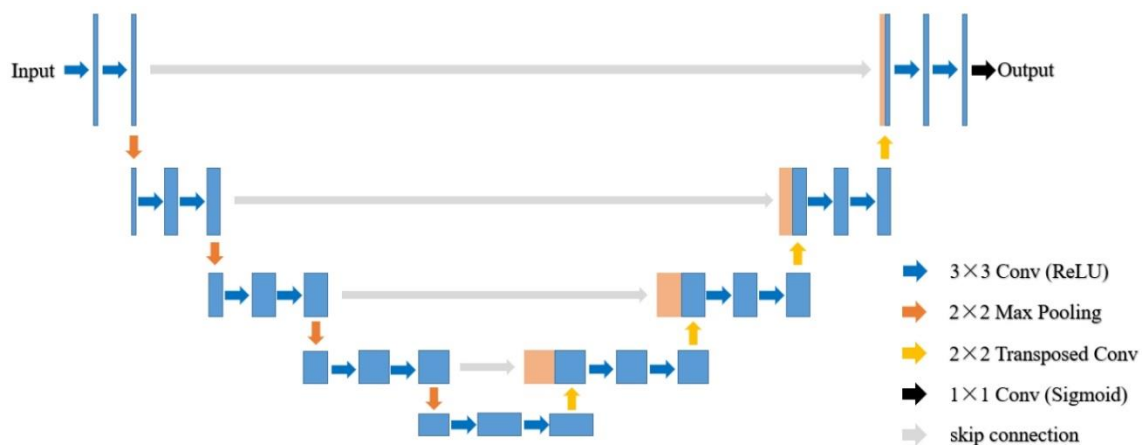
در ادامه این گزارش در بخش دوم مروری بر کارهای انجام شده و ادبیات موضوع خواهیم داشت. در بخش سوم، روش های پیشنهادی را بررسی میکنیم. در فصل چهارم نتایج مدل ها را مقایسه میکنیم تا میزان اثرگذاری آنها بیشتر مشخص شود و در آخرین فصل نتیجه گیری کلی، محدودیت های هر مدل و کارهای آینده را تعریف میکنیم.

## فصل ۲:

### مروری بر ادبیات موضوع

## ۱-۲- قطعه بندی تومور

شبکه های عصبی کانولوشنی میتوانند ویژگی های تصویر ورودی شبکه که شامل اشیاء مختلف در اندازه ها و زوایای مختلف است را به صورت خودکار یاد بگیرند. با افزایش تعداد لایه های شبکه (عمق شبکه) ویژگی های جزئی تر و دقیق تری از تصویر استخراج میشود که باعث می شود قدرت شبکه در حل مسائل مختلف بینایی ماشین از جمله تشخیص شیء، طبقه بندی و قطعه بندی تصویر افزایش یابد. بنابراین روش های بسیاری مبتنی بر یادگیری عمیق معرفی شدند تا مسئله قطعه بندی تصویر را حل کنند [۱۱]. یکی از معماری های شناخته شده برای قطعه بندی معنایی<sup>۱</sup> تصویر، شبکه های کانولوشنی کامل<sup>۲</sup> هستند که به FCN معروف هستند [۱۲]. شبکه های FCN بر پایه لایه های رمزنگار<sup>۳</sup> و رمزگشا<sup>۴</sup> بنا نهاده شده اند که این روش مشکل لایه های کاملاً متصل را در شبکه های عصبی کانولوشنی حل میکند. در این راستا معماری U-Net [۱۳] معرفی شد که لایه های رمزنگار و رمزگشا را به هم متصل می کند دلیل این اتصال حفظ اطلاعات مهم و موثر ویژگی های تصویر ورودی است شکل (۱-۲) معماری شبکه U-Net را نمایش می دهد.



شکل (۱-۲) معماری U-Net

در ادامه معماری Seg-Net [۱۴] که یک نوع معماری FCN است معرفی شد که ایده اصلی این معماری

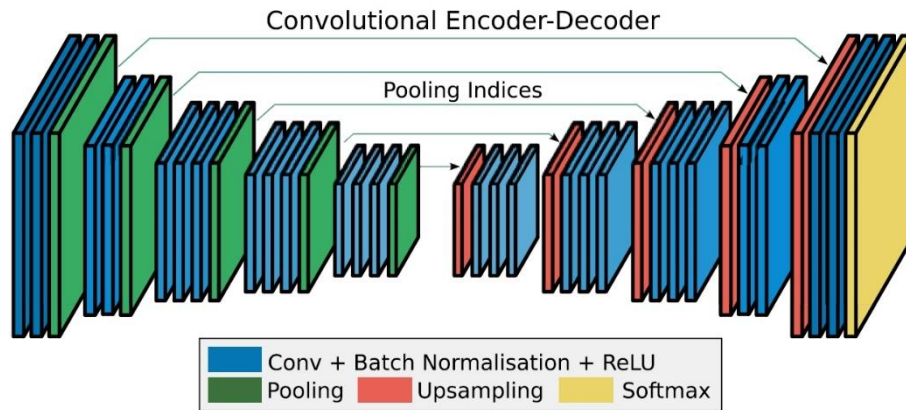
<sup>۱</sup> Semantic Segmentation

<sup>۲</sup> Fully Convolutional Network (FCN)

<sup>۳</sup> Encoder

<sup>۴</sup> Decoder

استفاده از رمزگشاهای سلسله مراتبی است که هر رمزگشا، متناظر یک رمزنگار می باشد. لایه های رمزگشا از لایه ادغام حداکثری<sup>۱</sup> در ورودی سمت رمزنگار استفاده میکند تا خاصیت غیرخطی بودن را روی نقشه ویژگی اعمال کند. شکل (۲-۲) نمایشی کلی از معماری Seg-Net می باشد.



شکل (۲-۲) معماری Seg-Net

با توجه به عملکرد چشمگیری که شبکه های عصبی عمیق بر روی مسئله قطعه بندی معنایی داشتند، از این روش ها در کاربردهای پزشکی نیز استفاده شد [۱۵]. برای مثال، مدل SLSDeep [۱۶] برای قطعه بندی بیماری های پوستی معرفی شد که ایده اصلی آن بزرگنمایی نقشه های ویژگی ورودی از قسمت رمزنگار مدل بود تا نواحی بیماری پوستی به صورت واضح تری قابل مشاهده باشند و همچنین قطعه بندی خودکار آن توسط مدل ساده تر باشد. کاربرد دیگری از شبکه های عمیق در قطعه بندی بیماری قرنیه چشم است که روش [۱۷] به این موضوع پرداخته است. ایده اصلی این مقاله استفاده از شبکه های عمیق چند مقیاسی است. برای آموزش یک شبکه عمیق میتوان به دو صورت عمل کرد. در حالت اول وزن های شبکه از پایه آموزش میبندد [۱۸] اما حالت دومی نیز وجود دارد که از وزن های یک شبکه آموزش دیده استفاده کرد. مانند استفاده از وزن های شبکه VGG [۱۹] که بر روی مجموعه آموزشی معروف ImageNet [۲۰] آموزش دیده اند.

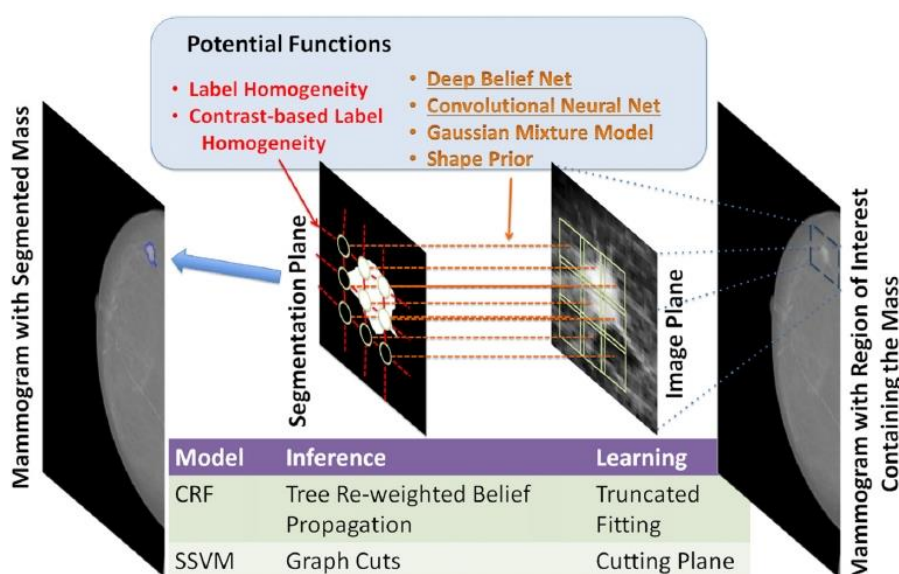
در بحث قطعه بندی تومورهای سینه کارهای بسیاری انجام شده است. یکی از ایده ها در این زمینه استفاده از الگوریتم های رشد منطقه<sup>۲</sup> خودکار بود که حد آستانه این الگوریتم توسط شبکه های عصبی مصنوعی و شبکه های عصبی سلولی<sup>۳</sup> یا به اختصار CeNN بدست می آید و برای قطعه بندی و طبقه بندی تومور بکار می رود [۲۱]. در روش مقاله [۲۲] به منظور کاهش پیچیدگی محاسباتی و افزایش توان مدل از یک

<sup>1</sup> Max Pooling

<sup>2</sup> Region growing

<sup>3</sup> Cellular Neural Network (CeNN)

CeNN غیرخطی و کمی سازی شده استفاده می شود. و در انتهای مدل پیشنهادی از یک مدل پرسپترون چندلایه به عنوان طبقه بند خوش خیم یا بدخیم بودن تومور استفاده می کند. با توجه به اینکه تمرکز اصلی این پروژه قطعه بندی تومور در تصاویر ماموگرافی است به بررسی بعضی از روش های جدید در این زمینه می پردازیم. در مقاله [۲۳] از ماشین بردارهای پشتیبان ساخت یافته<sup>۱</sup> (SSVM) به همراه فیلدهای تصادفی شرطی<sup>۲</sup> استفاده شد که هر دو روش سعی در کمینه کردن مقدار خطای مدل دارند که خطا تابعی از احتمال پیکسل های تصویر بدست آمده از شبکه های عصبی کانولوشنی، شبکه باور عمیق<sup>۳</sup>، مدل مخلوط گاوسی<sup>۴</sup> و شکل قبلی تومور است. معماری این مدل در شکل (۲-۳) نمایش داده شده است.



شکل (۲-۳) معماری مدل معرفی شده در مرجع ۲۳

در مقالات [۲۴]، [۲۵] به حل مشکل مشابهی با مقاله قبلی پرداخته شد تنها تفاوت این روش با روش قبلی نوع بهینه سازی تابع ضرر است که بر پایه ی مشتق شعاعی<sup>۵</sup> است. همچنین نوع دیگری از معیارهای سنجش به نام منظم بودن مقادیر پیکسل سطح خاکستری درون و بیرون تومور در مقاله [۲۵] معرفی شد. یک شبکه FCN با یک لایه CRF در مقاله [۲۶] الحاق شد تا میزان فشردگی خروجی قطعه بندی شده را

<sup>1</sup> Structured Support Vector Machine (SSVM)

<sup>2</sup> Conditional Random Fields (CRF)

<sup>3</sup> Deep Belief Networks

<sup>4</sup> Gaussian Mixture Model (GMM)

<sup>5</sup> Radial Derivative

در موقعیت پیکسل محاسبه کند. این مدل به صورت انتها به انتها<sup>۱</sup> آموزش داده میشود زیرا هم شبکه FCN و لایه CRF توانایی تبدیل داده ها در قسمت روبه جلو و عقب شبکه را دارند. در بعضی مقالات، نویسندگان چند هدف در کنار هم را دنبال میکنند. برای مثال در [۲۷] یک سیستم CAD معرفی شد که از سه مرحله عمیق برای تشخیص، قطعه بندی و طبقه بندی تومورها در تصاویر ماموگرافی استفاده میکرد. همچنین مدل معروف YOLO<sup>۲</sup> برای مشخص کردن مکان تومورها در تصاویر مورد استفاده قرار گرفت [۲۸]. در این مقاله پس از مشخص شدن مکان تومور از یک شبکه عصبی با وضوح کامل<sup>۳</sup> (FrCN) برای قطعه بندی دقیق تومور استفاده شد. روش [۲۹] یکی از نخستین مقالاتی است که از GAN برای کاربرد پزشکی استفاده کرد. در این مقاله تصاویر سه بعدی کبد در تصاویر پرتونگاری کامپیوتری<sup>۴</sup> قطعه بندی میشوند. همچنین [۳۰] از cGAN برای قطعه بندی تصاویر دوبعدی ماموگرافی استفاده میکند که در این پروژه سعی شده است تا نتایج آن را بهبود دهیم.

## ۲-۲- طبقه بندی شکل تومور

در سال های اخیر مقالات زیادی حول طبقه بندی شکل در تصاویر دو بعدی و سه بعدی بر پایه شبکه های عصبی عمیق معرفی شده اند [۳۱]. برای مثال یک شبکه عصبی کانولوشنی [۳۲] که از گراف موجک طیفی<sup>۵</sup> برای پیوستگی دسته ای ویژگی ها برای طبقه بندی شکل استفاده میکند معرفی شد. همچنین مدل PointNet [۳۳] به طور مستقیم از توده نقاط<sup>۶</sup> برای مسائل طبقه بندی اشیاء و قطعه بندی معنایی محلی و سراسری استفاده میکند. توده نقاط مجموعه سه بعدی از نقاط است که خصوصیات سطح خارجی یک شیء را توصیف میکنند. شکل (۲-۴) معماری کلی مدل PointNet را نمایش می دهد.

<sup>1</sup> End-to-end

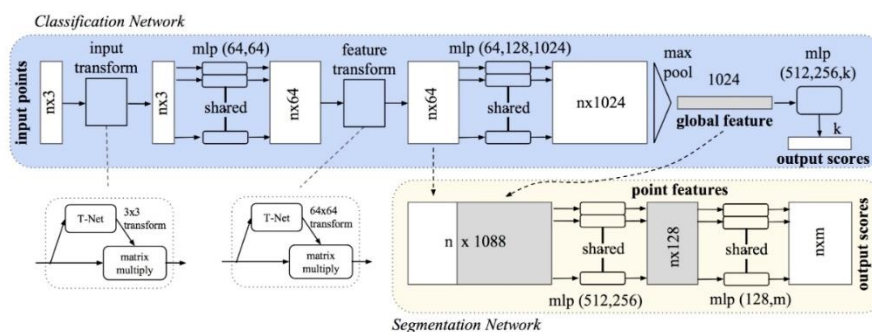
<sup>2</sup> You Look Only Once (YOLO)

<sup>3</sup> Full Resolution Convolutional Network (FrCN)

<sup>4</sup> Computerized Tomography (CT)

<sup>5</sup> Spectral Graph Wavelets

<sup>6</sup> Point Cloud



شکل (۴-۲) معماری مدل Point-Net

تا به حال تعداد زیادی روش برای طبقه بندی شکل در کاربرد پزشکی معرفی شده است [۱۵]. در مقاله [۳۴] از توصیفگرهای حوزه فوریه در کنار شبکه های عصبی کانولوشنی در تصاویر CT استفاده شده است. در مورد تصاویر سینه نیز مدلی خودکار برای توصیف بافت تومور استفاده شد که با استفاده از خروجی توصیفگر بافت شکل تومور را طبقه بندی میکند [۳۵]. بعلاوه در [۳۶] نویسندگان تلاش کردند که یک مدل چند وظیفه ای بر پایه ناحیه را برای طبقه بندی شکل تومور به سه دسته بکار گیرند. مقالات [۳۰]، [۳۷] به طور تخصصی بر روی طبقه اشکال تومور در تصاویر ماموگرافی سینه کار کردند. این دو روش از مجموعه دادگان عمومی CBIS-DDSM برای آموزش و تست مدل خود استفاده کردند.

## ۳-۲- طبقه بندی نوع تومور

تصاویر ماموگرافی به طور کلی دو دسته هستند. تصاویر حاوی توده یا تصاویر نرمال. توده ها نیز به دو دسته خوش خیم و بدخیم تقسیم میشوند. توده خوشخیم از لحاظ تصویری حاشیه هایی بسیار ملایم و یکنواخت دارد در حالی که توده های بدخیم حاشیه هایی تیره رنگ و برجسته دارند و با گذشت زمان این حاشیه ها به صورت تیز و سوزنی شکل در می آیند. Subramaniam و همکارانش در روشی مبتنی بر شبکه های عصبی RBF برای تشخیص سرطان پستان ارایه دادند. آنها در این روش با استفاده از خصوصیات آماری تصاویر ماموگرافی به تحلیل و بررسی آنها پرداختند و توانستند در شناسایی توده ها تا حدودی موفق باشند و روشی دیگر برای تشخیص سرطان پستان با استفاده از آنالیز تصویری ارائه کردند [۳۸]. در [۳۹] نویسندگان الگوریتم جدیدی برای تشخیص توده های سرطانی در تصاویر با استفاده از ویولت دو بعدی



معرفی کردند. Chandra و همکارانش [۴۰] تشخیص سرطان پستان را با استفاده از شبکه های عصبی مصنوعی با تکنیک های یادگیری شدید انجام دادند شبکه عصبی مصنوعی مدل ساده شده ای از سیستم عصبی طبیعی بوده و همانند مغز با پردازش روی داده های تجربی، قابلیت یادگیری دارد. در واقع شبکه ها با انجام محاسبات روی داده های عددی یا مثال ها، قوانین کلی را فرا میگیرند و به همین دلیل به آنها سیستم های هوشمند گفته می شود. یک الگوریتم فازی عصبی برای تشخیص میکروکلسیفیکیشن های خوش خیم و بدخیم در [۴۱] ارائه شد. Jimenez و همکاران طبقه بندی عصبی مصنوعی کانولوشن را برای طبقه بندی توده ها و بافت طبیعی پستان مورد استفاده قرار دادند. بدین صورت که در ابتدا، از منطقه مورد علاقه<sup>۱</sup> تصویر گرفتند و به طور متوسط در زیرمجموعه قرار دادند. سپس آمار اختلاف سطح خاکستری و ویژگی وابستگی خاکستری فضایی از مناطق مختلف را محاسبه کردند و در انتها ویژگی های محاسبه شده را به عنوان ورودی به طبقه بندی کانولوشن دادند [۴۲]. در مقاله ای [۴۳] نویسندگان به کمک شبکه عصبی کانولوشن توانستند با دقت ۹۸٪ تومور و خوشخیم یا بدخیم بودن آن را تشخیص دهند.

<sup>1</sup> Region of Interest (ROI)

## فصل ۳:

### روش انتخابی حل مسئله

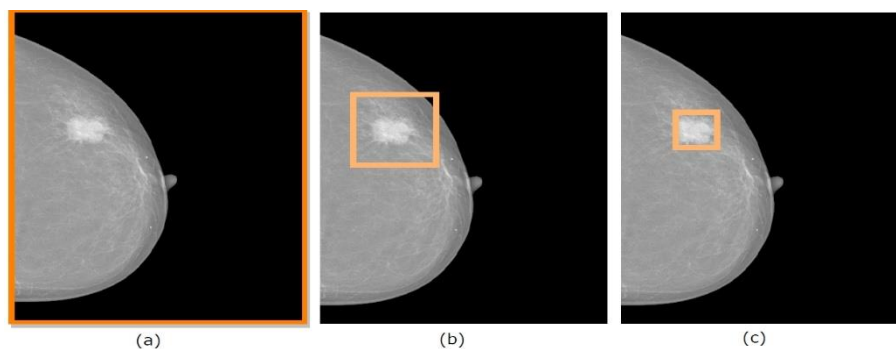
## ۳-۱- قطعه بندی تومور

در این قسمت معماری مدل هایی که برای قطعه بندی تومور پیاده سازی شده اند به صورت دقیق توضیح داده خواهند شد.

### ۳-۱-۱- مدل پیشنهادی cGAN

سیستم CAD معرفی شده که در شکل (۱-۱) نمایش داده شده است که شامل دو مرحله اصلی قطعه بندی و طبقه بندی است. در مرحله اول، چند مرحله پیش پردازشی بر روی تصاویر ماموگرافی به منظور حذف نویز اعمال میشود. مراحل پیش پردازش شامل موارد زیر است:

- اعمال فیلتر گاوسی با استفاده از انحراف معیار  $\sigma = 0.5$
  - بهبود کنتراست تصویر با استفاده از تکنیک یکسان سازی هیستوگرام<sup>۱</sup>
  - مقیاس گذاری بر روی مقادیر پیکسل ها به طوری که تمامی مقادیر بین ۰ و ۱ باشند.
  - برش تصویر و تبدیل تمامی تصاویر به اندازه  $256 \times 256$  پیکسل
- همچنین در این مقاله از قاب های مختلف ناحیه مورد نظر<sup>۲</sup> (ROI) استفاده میشود. این قاب ها به صورت تصاویر کامل ماموگرافی، ROI عریض و ROI تنگ<sup>۳</sup> است. در شکل (۱-۳) نمونه ای از انواع قاب های استفاده شده را مشاهده میکنید.



شکل (۱-۳) (a) ماموگرافی کامل (b) ماموگرافی عریض (c) ماموگرافی تنگ

<sup>1</sup> Histogram Equalization

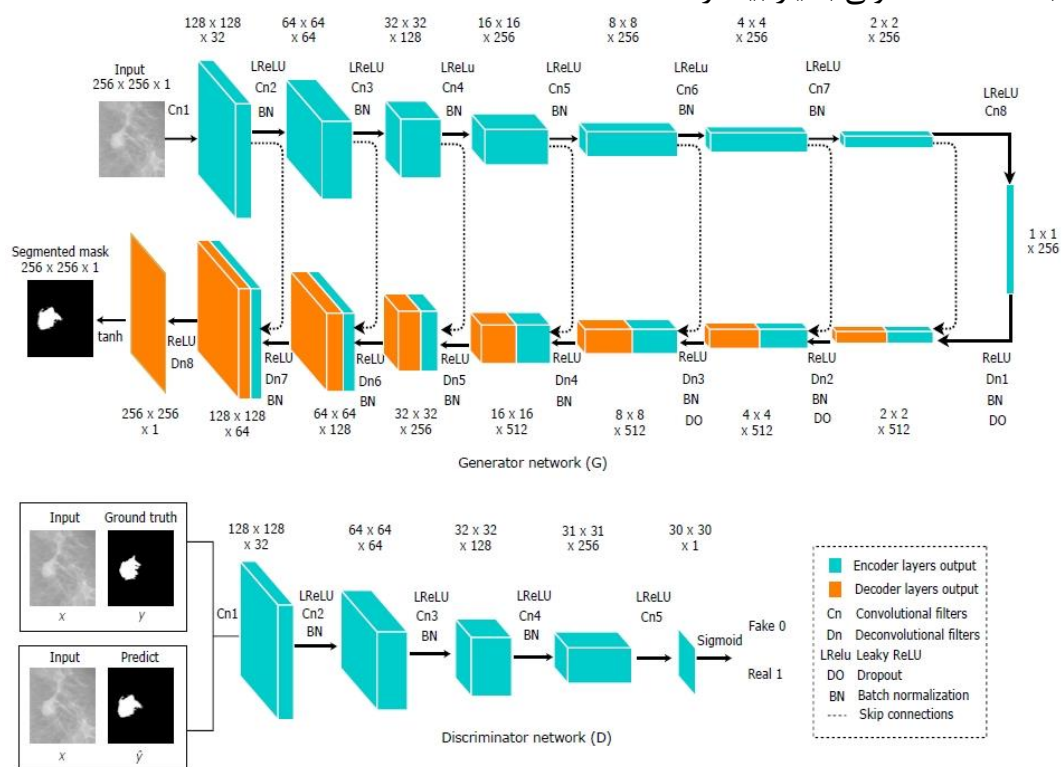
<sup>2</sup> Region of Interest (ROI)

<sup>3</sup> Tight ROI

پس از مرحله پیش پردازش، تصاویر به یک cGAN داده می شوند تا ماسک دودویی تومور سینه را تولید کند. سپس ماسک دودویی تومور نیز پردازش می شود تا نقاط کوچکی که به اشتباه در ماسک بوجود آمده است حذف گردند. پردازش های ذکر شده شامل موارد زیر هستند:

- اعمال عملگر closing با اندازه هسته  $3 \times 3$
- اعمال عملگر erosion با اندازه هسته  $2 \times 2$
- اعمال عملگر dilation با اندازه هسته  $3 \times 3$

cGAN معرفی شده در این مقاله همانند سایر مدل های GAN از دو شبکه تشکیل شده است. شبکه مولد<sup>۱</sup> (G) و شبکه تفکیک کننده<sup>۲</sup> (D). شکل (۲-۳) معماری معرفی شده در مقاله را به همراه جزئیات لایه ها و فیلترها نمایش میدهد. ادغام دو شبکه G و D منجر به یک روند آموزشی پایدار میشود که نیاز به تعداد بالای نمونه ها در مجموعه آموزشی ندارد. همینطور استفاده از GAN شرطی که ماسک اصلی تومور را نیز در ورودی دریافت میکند عملکرد بهتری نسبت به GAN معمولی دارد و همچنین نیاز به داده های آموزشی برای شبکه GAN معمولی بسیار بیشتر است.

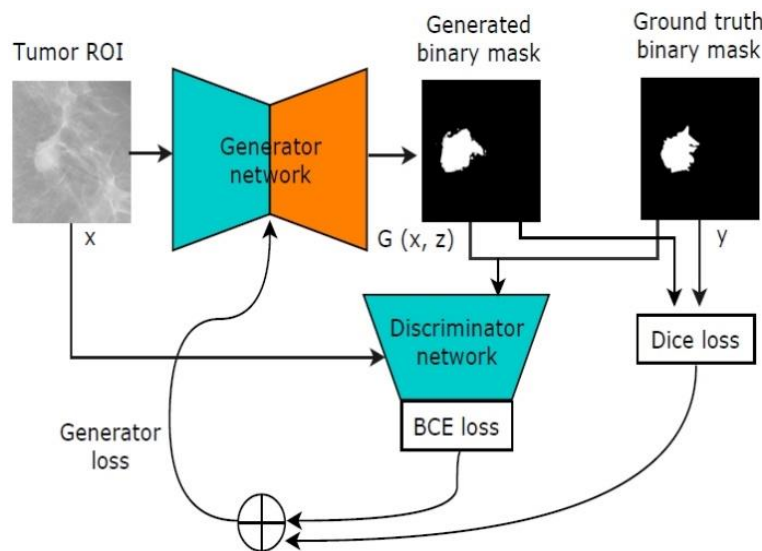


شکل (۲-۳) جزئیات معماری cGAN

<sup>1</sup> Generator Network

<sup>2</sup> Discriminator Network

بهینه سازی هر دو شبکه G و D به صورت همزمان انجام میشود. به عبارتی دیگر در هر قدم بهینه سازی، سعی در بهبود عملکرد هر دو شبکه داریم. بهبود شبکه G به معنی تولید ماسک دودویی دقیق تری است و بهبود شبکه D به معنی تفکیک بهتر بین ماسک واقعی و ساختگی است. در این مقاله تابع ضرر جدیدی نیز ارائه شده بود که ترکیبی از خطای دسته بندی دودویی و Dice Loss است. شکل (۳-۳) چگونگی کارکردن این تابع ضرر را نمایش می دهد.



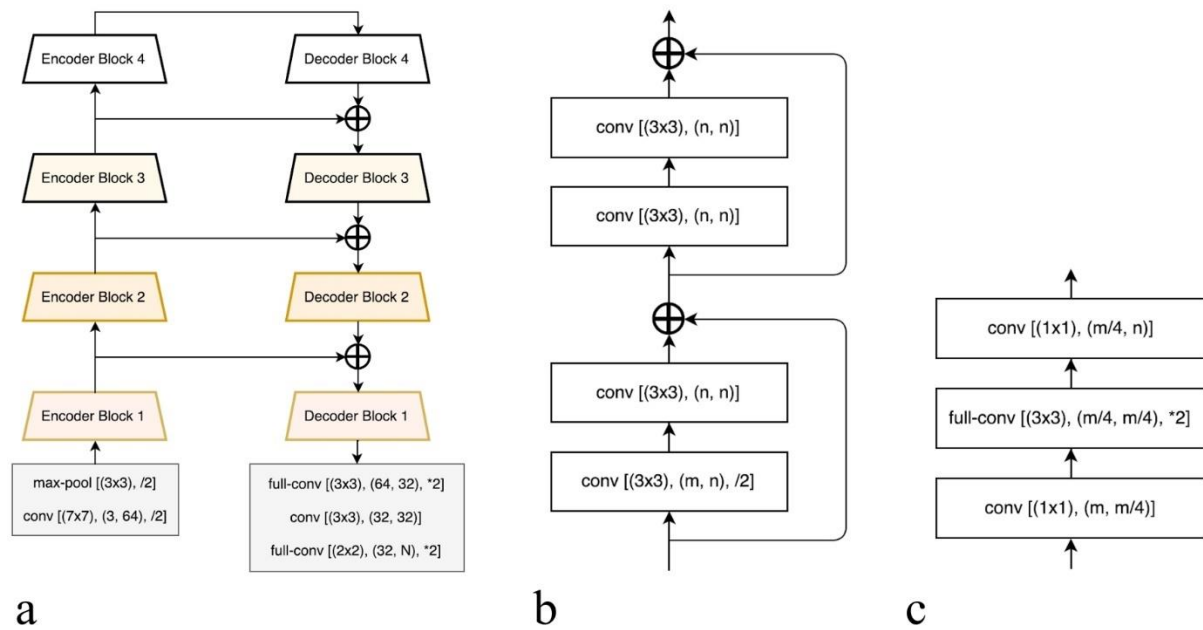
شکل (۳-۳) نحوه استفاده دو تابع ضرر Dice و BCE

برای انتخاب بهترین بهینه ساز، بهینه سازهای SGD، AdaGrad، Adadelta، RMSProp و Adam مورد بررسی قرار گرفتند که بهینه ساز Adam با  $\beta_1 = 0.5$  و  $\beta_2 = 0.999$  و نرخ آموزش برابر 0.0002 و اندازه دسته<sup>۱</sup> برابر ۸ بهترین نتایج را دارد.

## ۲-۱-۳- مدل Link-Net

مدل Link-Net در مقاله [۴۴] معرفی شد. این مقاله در کنار عملکرد خوبی که دارد مدلی است که به تعداد خیلی زیادی پارامتر برای یادگیری نیاز ندارد. این مدل برای تصاویری با اندازه  $3 \times 360 \times 640$  حدوداً ۱۱.۵ میلیون پارامتر قابل آموزش خواهد داشت. معماری مدل Link-Net در شکل (۴-۳) مشاهده می شود.

<sup>1</sup> Batch Size



شکل (۳-۴) (a) معماری Link-Net (b) ساختار بلوک رمزنگار (c) ساختار بلوک رمزگشا

در معماری شکل (۳-۴) منظور از conv عملگر convolution و full-conv همان full convolution معرفی شده در مقاله [۱۲] است. در بین لایه های همگشتی از نرمالسازی دسته ای و تابع فعال سازی ReLU برای غیرخطی سازی استفاده شده است. در هر بلوک دو متغیر  $m$  و  $n$  وجود دارد که برای هر بلوک تفاوت دارد. مقادیر استاندارد آن ها را در جدول (۳-۱) مشاهده می کنید.

جدول (۳-۱) نقشه ویژگی های ورودی و خروجی

Block	Encoder		Decoder	
	m	n	m	n
1.	64	64	64	64
2.	64	128	128	64
3.	128	256	256	128
4.	256	512	512	256

در این پروژه با توجه به اینکه تعداد داده ها برای آموزش شبکه از ابتدا<sup>۱</sup> میسر نبود از وزن های پیش آموزش دیده این مدل استفاده کردیم و قسمت رمزنگار این مدل را ثابت<sup>۲</sup> کردیم و تنها قسمت رمزگشای این مدل

<sup>۱</sup> Scratch

<sup>۲</sup> Freeze

را آموزش داده ایم.

### ۳-۱-۳- مدل U-Net

مدل U-Net [۱۳] یکی از معروف ترین معماری ها برای وظیفه قطعه بندی تصاویر است. مخصوصا در داده های پزشکی در بسیاری از موارد عملکرد قابل توجهی از خود نشان داده است. شبکه U-Net از یک مسیر منقبض شده و یک مسیر گسترشی تشکیل شده است که آن طرحی به شکل U می بخشد. در مقاله اصلی U-Net مسیر منقبض شده یک شبکه معمولی کانولوشنی است که از استعمال مکرر کانولوشن ها تشکیل شده است و هر یک با استفاده از یک ReLU و یک عملگر پولینگ حداکثر دنبال میشود. در طول انقباض، ضمن افزایش اطلاعات ویژگی، اطلاعات مکانی کاهش می یابد. مسیر گسترش اطلاعات ویژگی و مکانی را از طریق دنباله ای از کانولوشن ها ترکیب می کند و با ویژگی های با وضوح بالا از مسیر انقباضی جمع میکند. شکل (۱-۲) مدل توضیح داده شده را نمایش می دهد. در این پروژه همانند معماری Link-Net مسیر منقبض شده معماری را ثابت کردیم و همینطور به جای استفاده از لایه های مکرر کانولوشنی در مسیر منقبض شده از معماری ResNet34 [۴۵] به عنوان معماری پایه ای<sup>۱</sup> استفاده کردیم. این کار باعث می شود تعداد پارامترهای شبکه افزایش زیادی داشته باشد اما بازنمایی بهتری در مسیر منقبض شده خواهیم داشت و همچنین با توجه به اینکه از وزن های پیش آموخته ثابت استفاده می کنیم مشکل تعداد پارامترها حل خواهد شد.

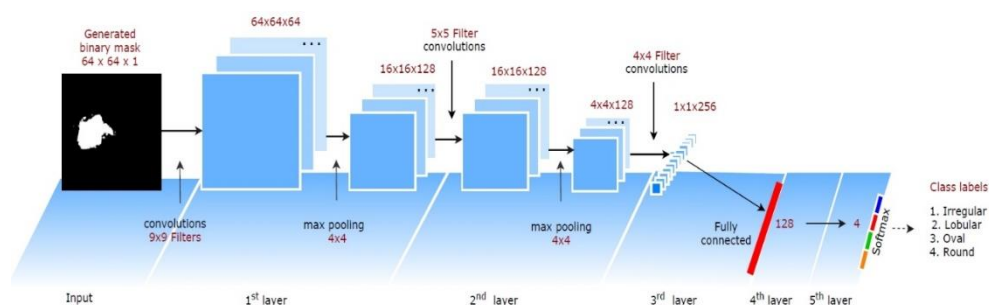
### ۳-۲- طبقه بندی شکل تومور

#### ۳-۲-۱- مدل پیشنهادی مقاله

در مقاله یک مدل بر پایه شبکه های عصبی کانولوشنی (CNN) برای طبقه بندی چند کلاسه معرفی شده است تا شکل تومورهای سینه را کلاسیبندی کند. در این قسمت، مدل با گرفتن تصویر ورودی که ماسک دودویی تولید شده از مدل قطعه بند است، نوع شکل تومور را در یکی از دسته های نامنظم، گرد، بیضی شکل

<sup>1</sup> Backbone

و ابرمانند قرار می گیرد مشخص میکند. تفاوت عمده روش پیشنهادی در این قسمت با سایر روش در این است که سایر روش ها نظیر [۳۵]–[۳۷] تلاش می کنند به صورت مستقیم از روی شدت، بافت، مرز و ... در تصویر خود تومور، شکل آن را مشخص کنند که باعث افزایش هزینه محاسباتی میشود. در این مقاله شکل تومور به راحتی با استخراج ویژگی های ساختاری ماسک دودویی تومور مشخص میشود. شکل (۳–۵) مدل پیشنهادی طبقه بندی شکل تومور را نمایش میدهد.



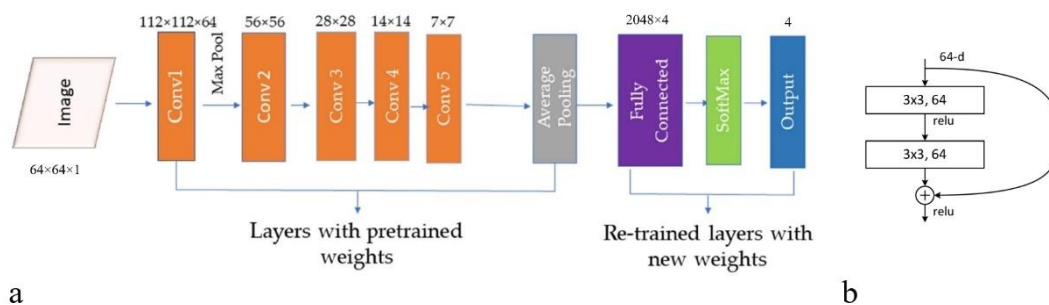
شکل (۳–۵) معماری CNN پیشنهادی مقاله

## ۲-۲-۳- مدل ResNet50

اگرچه انتظار می رود شبکه های عمیق تر نتایج بهتری داشته باشند اما در عمل آموزش یک شبکه عمیق با چالش هایی همراه است. یکی از این چالش ها محو یا انفجار گرادیان است. این مثل تا حدود زیادی توسط مقدار دهی اولیه به صورت نرمال و نرمالیزه کردن لایه های میانی بر طرف شده است. همچنین تخریب در شبکه های عمیق به این معنی است که با افزایش عمق شبکه، دقت شبکه به شکل غیر منتظره ای اشباع می شود و سپس به سرعت کاهش می یابد. در [۴۵] برای حل مسئله تخریب، یک شبکه عمیق مبتنی بر یادگیری باقیمانده ارائه شد. هر بلوک از یک شبکه عمیق دارای چندین لایه متصل به هم است که در تلاش برای تخمین زدن یک نگاشت هستند. رابطه  $F(x) + x$  یک شبکه پیش خور<sup>۱</sup> دارای یم اتصال میان بر قابل پیاده سازی است. در چنین شبکه ای اتصال های میان بر عمل نگاشت همانی را انجام می دهند. همچنین این اتصالات هیچ پارامتری به شبکه اضافه نمیکنند. شکل (۳–۶) معماری شبکه و ساختار بلوک های مدل Res-Net50 مورد استفاده در این پروژه را نمایش میدهد. این مدل به صورت پیش آموخته استفاده شده است و فقط لایه های کاملاً متصل آن آموزش دیده اند.

<sup>1</sup> Feed Forward





ResNet معماری مدل (۳-۶) شکل

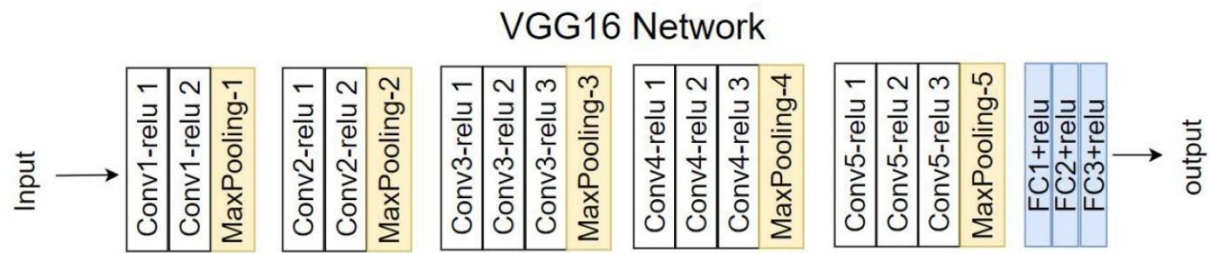
جدول (۲-۳) جزئیات معماری Res-Net با عمق های ۱۸ لایه و ۵۰ لایه را نمایش میدهد.

جدول (۳-۲) معماری Res-Net با عمق های مختلف

layer name	output size	18-layer	50-layer
conv1	$112 \times 112$	$7 \times 7, 64, \text{stride } 2$	
conv2_x	$56 \times 56$	$3 \times 3 \text{ max pool, stride } 2$	
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	$28 \times 28$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
conv4_x	$14 \times 14$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
conv5_x	$7 \times 7$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	$1 \times 1$	average pool, 1000-d fc, softmax	
FLOPs		$1.8 \times 10^9$	$3.8 \times 10^9$

### ۳-۲-۳- مدل VGG16

مدل اصلی VGG16 [۱۹] در مجموعه داده ی ImageNet به دقت 92.7% دست می یابد که با جایگزینی فیلترهای  $3 \times 3$  به جای فیلترهای بزرگ استفاده شده در شبکه AlexNet دقت را افزایش میدهد. معماری شبکه VGG16 در شکل (۷-۳) آمده است. در این پروژه از دو نسخه مختلف VGG استفاده شده است. در قسمت طبقه بندی شکل تومور از مدل VGG16 و در قسمت طبقه بندی نوع تومور از VGG19 که دارای ظرفیت یادگیری بالاتری است استفاده کرده ایم.



شکل (۷-۳) معماری VGG16

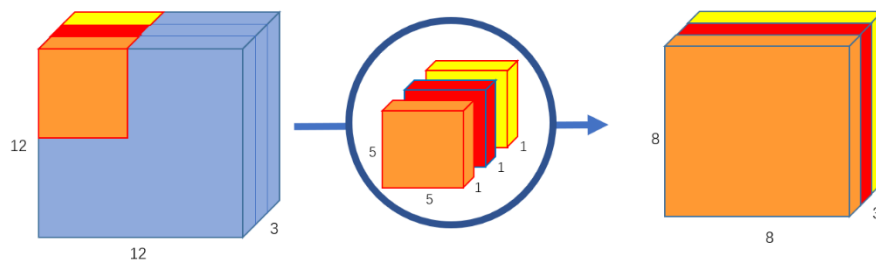
## ۴-۲-۳- مدل MobileNet

معماری مدل های MobileNet [۴۶]، معماری هایی کوچک با تاخیر و تعداد پارامترهای کم هستند. این معماری ها با توجه به ویژگی هایشان مناسب کاربردهای نهفته هستند. این شبکه توسط محققان گوگل با هدف طراحی شبکه های کارآمد، سبک، سریع و با دقت قابل قبول مطرح شده است. در شبکه عصبی MobileNet، عنصر اصلی یک نوع کانولوشن جدید به نام depth-wise separable convolution است. جدول (۳-۳) مقایسه ای بین MobileNet و سایر معماری های شناخته شده از نظر دقت بر روی مجموعه داده های ImageNet، تعداد پارامتر، سایز و عمق معماری است.

جدول (۳-۳) مشخصات مدل های مختلف

Model	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth	Size (MB)
MobileNet	0.704	0.895	4,253,864	88	16
Xception	<b>0.790</b>	<b>0.945</b>	22,910,480	126	88
VGG16	0.713	0.901	138,357,544	23	528
VGG19	0.713	0.900	143,667,240	26	549
ResNet50	0.749	0.921	25,636,712	-	98
ResNet101	0.764	0.928	44,707,176	-	171
InceptionV3	0.779	0.937	23,851,784	159	92

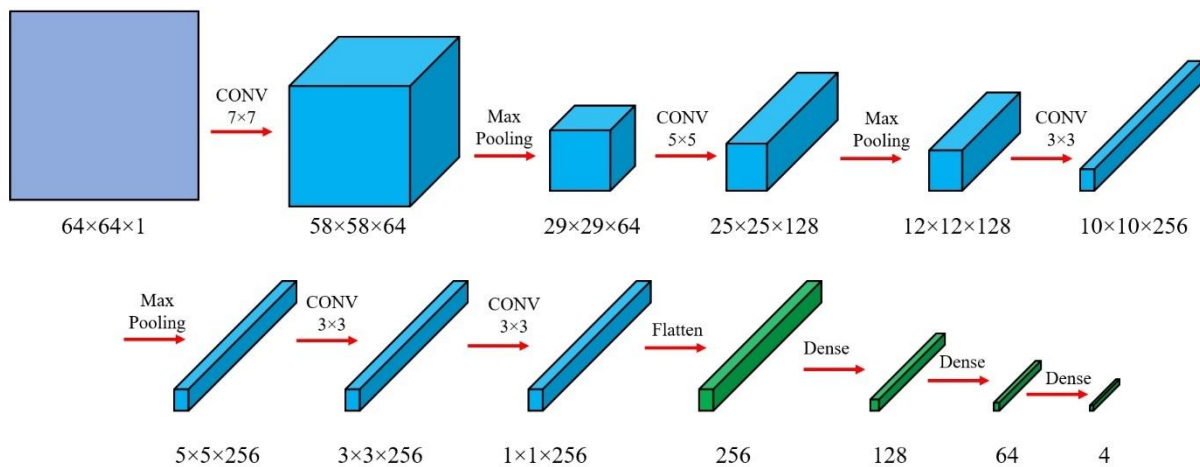
در بین این معماری ها MobileNet کمترین پارامتر و سایز را دارد در حالی شاخص دقت در آن قابل قبول و تقریباً در سطح سایر معماری های سنگین است. شکل (۳-۸) چگونگی عملکرد کانولوشن depth-wise در معماری مدل MobileNet را نمایش میدهد.



شکل (۳-۸) عملکرد کانولوشن depth-wise

### ۳-۲-۵- مدل پیشنهادی جدید

روش پیشنهادی ما برای طبقه بندی شکل تومور در واقع نسخه بهبود یافته مدل پیشنهادی مقاله است. این مدل از اندازه هسته های کوچکتری استفاده میکند و عمق لایه های کانولوشنی و کاملاً متصل آن افزایش یافته است. دلیل این دو تغییر افزایش ظرفیت مدل قبلی است. معماری این مدل در شکل (۳-۹) آمده است.

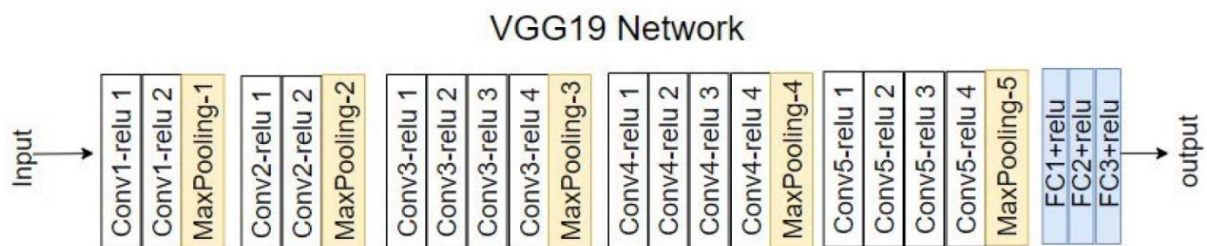


شکل (۳-۹) مدل پیشنهادی طبقه بندی شکل تومور

### ۳-۳- طبقه بندی نوع تومور

#### ۳-۳-۱- مدل VGG19

این مدل در قسمت (۳-۲-۳) معرفی شد. اما در این قسمت از نسخه ی ۱۹ لایه آن استفاده می کنیم که شکل (۳-۱۰) نمایش جزئی آن است. نکته قابل توجه این است که در سری مدل های طبقه بندی نوع تومور در خروجی شبکه ۳ نرون داریم اما در قسمت قبل مدل های طبقه بند شکل تومور ۴ نرون خروجی داشتند.



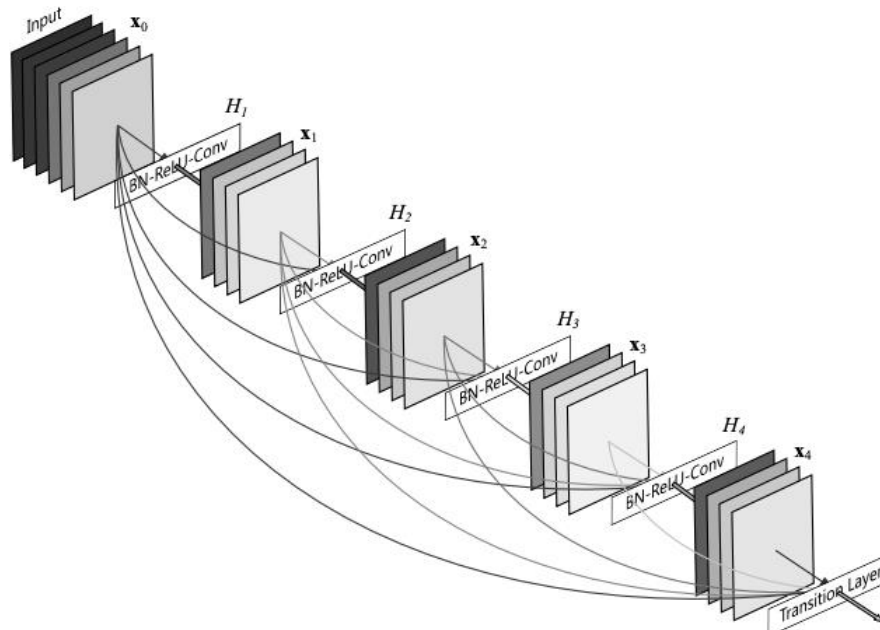
شکل (۳-۱۰) مدل VGG19

#### ۳-۳-۲- مدل ResNet50

این مدل نیز در قسمت (۳-۲-۲) توضیح داده شد. در این قسمت نیز از حالت پیش آموخته شبکه با تغییر لایه های کاملاً متصل استفاده می شود.

#### ۳-۳-۳- مدل DenseNet121

مدل DenseNet [۴۷] دارای بلوک های بسیار باریک و فشرده ای هستند. منظور از فشردگی این است که به طور مثال ۱۲ فیلتر در یک لایه قرار دارد. این شبکه میتواند الگوهای متنوع تری را آموزش ببیند زیرا طبقه بندی DenseNet از ویژگی های تمام سطوح پیچیدگی استفاده می کند. ساختار کلی معماری DenseNet در شکل (۳-۱۱) آورده شده است. همچنین جدول (۳-۴) جزئیات بیشتری از بلوک های مدل DenseNet121 را نمایش می دهد.



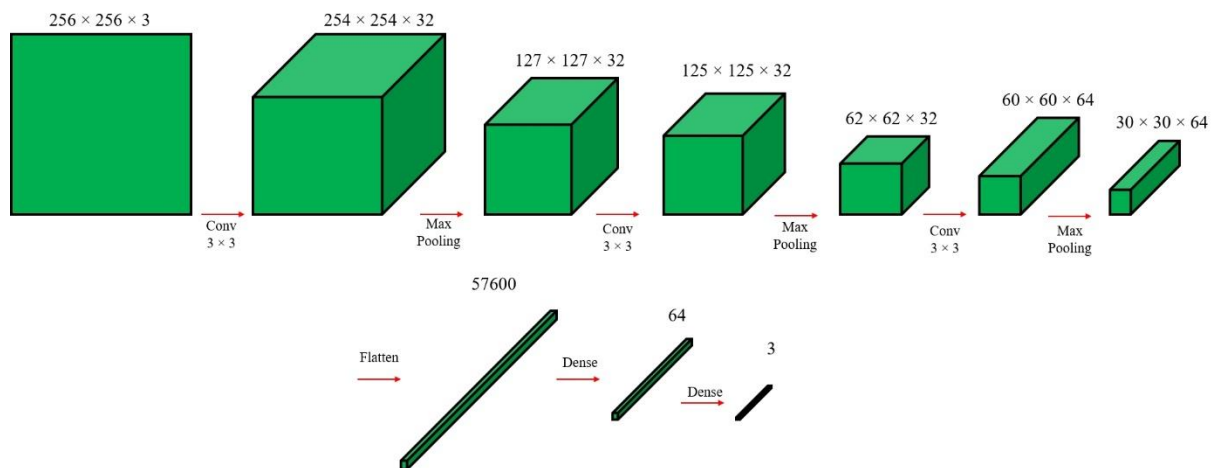
شکل (۳-۱۱) مدل شبکه DenseNet

جدول (۳-۴) جزئیات معماری DenseNet121

Layers	Output Size	DenseNet-121
Convolution	$112 \times 112$	
Pooling	$56 \times 56$	
Dense Block (1)	$56 \times 56$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	$56 \times 56$ $28 \times 28$	
Dense Block (2)	$28 \times 28$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	$28 \times 28$ $14 \times 14$	
Dense Block (3)	$14 \times 14$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$
Transition Layer (3)	$14 \times 14$ $7 \times 7$	
Dense Block (4)	$7 \times 7$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$
Classification Layer	$1 \times 1$	$7 \times 7$ global average pool 1000D fully-connected, softmax

### ۴-۳-۳- مدل پیشنهادی جدید

در این قسمت نیز از مدلی پیشنهادی استفاده می کنیم تا عملکرد آن را در حالی از ابتدا با وزن های تصادفی آموزش می بیند با مدل هایی که از یادگیری انتقالی استفاده می کنند مقایسه کنیم. شکل مدل پیشنهادی در شکل (۳-۱۲) نمایش داده شده است. این مدل از لایه های ساده کانولوشنی و چند لایه کاملاً متصل تشکیل شده است. این مدل دارای ۳۷۱۵۲۹۹ پارامتر قابل آموزش است.



شکل (۳-۱۲) مدل پیشنهادی طبقه بندی نوع تومور

## فصل ۴:

### نتایج و تفسیر آنها

## ۴-۱- مجموعه دادگان

در این پروژه از دو مجموعه داده عمومی برای آموزش و ارزیابی مدل ها استفاده شده است.

### ۴-۱-۱- مجموعه دادگان INbreast

این دیتاست به صورت عمومی در اختیار همگان قرار دارد که این دیتاست ۱۱۵ مورد (در مجموع ۴۱۰ تصویر ماموگرافی) را شامل میشود. این دیتاست شامل توده ها، کلسیفیکاسیون<sup>۱</sup>، عدم تقارن و اعوجاج است. برای تست کردن مدل های قطعه بندی از ۱۰۶ تصویر دارای تومور به همراه ماسک دودویی ناحیه تومور استفاده شده است. در هنگام برش تصویر به قاب های عریض و تنگ هر تومور در یک تصویر قرار می گیرد پس در حالتی که قاب های عریض و تنگ را به شبکه آموزش می دهیم ۹۲ داده آموزشی خواهیم داشت و ۲۳ داده آزمون داریم. اما در حالتی که کل تصویر را به شبکه آموزش می دهیم ۸۵ داده آموزشی و ۲۱ داده آزمون در اختیار داریم. در قسمت انتهایی پروژه نیز برای آموزش مدل طبقه بندی نوع تومور از این مدل استفاده می کنیم که همه ی ۴۱۰ تصویر قابل استفاده هستند و در یکی از دسته های نرمال، خوش خیم و بدخیم قرار می گیرند.

### ۴-۱-۲- مجموعه دادگان DDSM

این دیتاست نیز همانند دیتاست INbreast به صورت عمومی منتشر شده است که شامل ۲۶۲۰ تصویر ماموگرافی می باشد. در این پروژه از ۱۴۷۶ تصویر که دارای تومور بودند به همراه ماسک دودویی آنها استفاده کردیم تا مدل طبقه بندی شکل تومور را آموزش دهیم. در این تعداد تصویر به ترتیب ۵۱۶، ۳۸۴، ۴۱۲ و ۱۶۴ تومور در کلاس های نامظنم، ابرمانند، بیضی شکل و گرد قرار دارند. از ۷۷٪ این تعداد تصویر برای آموزش مدل و از باقی تصاویر برای ارزیابی عملکرد مدل آموزش دیده استفاده کردیم.

<sup>1</sup> Calcifications



## ۴-۲- نتایج قطعه بندی تومور

در ابتدا معیارهای ارزیابی برای این قسمت را توضیح می دهیم. همانند مقاله از معیار Dice و معیار IoU که دو معیار معروف در زمینه قطعه بندی تصاویر هستند استفاده می کنیم. رابطه محاسبه این دو معیار در روابط (۱-۴) و (۲-۴) به ترتیب آمده است.

$$Dice\ Coefficient = 2 \times \frac{Y \cap Y_{pred}}{Y + Y_{pred}} \quad (۱-۴)$$

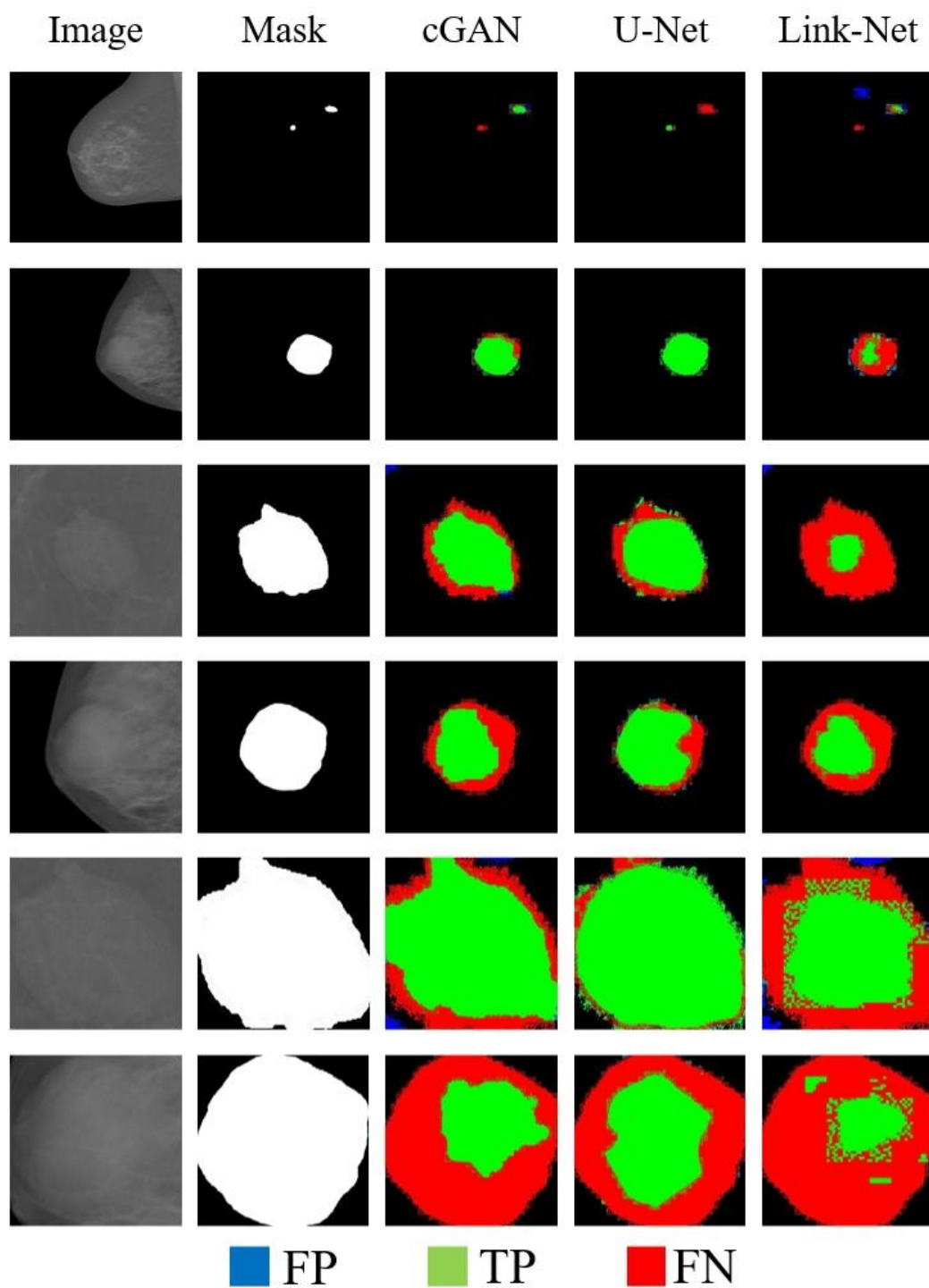
$$Intersection\ over\ Union\ (IoU) = \frac{Y \cap Y_{pred}}{Y \cup Y_{pred}} \quad (۲-۴)$$

در روابط بالا  $Y_{pred}$  نقاطی است که توسط مدل به عنوان تومور تعیین شده است و  $Y$  نقاطی هستند که در ماسک اصلی، تومور هستند. این دو معیار شباهت زیادی به هم دارند و هرچه مقدار عددی این دو معیار نزدیک تر به ۱ باشد نشان از عملکرد خوب مدل دارد. عملکرد هر سه مدل به صورت عددی در جدول (۱-۴) آمده است. در این جدول منظور از  $cGAN^*$  نتایج ارائه شده در مقاله است. نتایج نشان می دهد Link-Net ضعیف ترین عملکرد را دارد اما مقایسه  $cGAN$  پیشنهادی مقاله که برای شبکه G از U-Net استفاده می کند با شبکه U-Net به تنهایی نشان می دهد این دو مدل تفاوت عملکردی چشم گیری ندارند.

جدول (۴-۱) جدول عملکرد مدل های بررسی شده در پروژه

Model	Full Breast		Loose Breast		Tight Breast	
	Avg. Dice	Avg. IoU	Avg. Dice	Avg. IoU	Avg. Dice	Avg. IoU
Link-Net	0.4452	0.3090	0.5528	0.3820	0.5747	0.3337
U-Net	0.5872	0.4482	0.8052	0.7501	0.7474	0.7592
cGAN	0.6063	0.4312	0.7960	0.7516	0.7788	0.7383
cGAN*	0.6638	0.4968	0.8999	0.8181	0.8812	0.7987

همچنین برای مقایسه بصری نتایج، چند نمونه از نتایج هر مدل را در شکل (۴-۱) ملاحظه می کنید.



شکل (۴-۱) نتایج کیفی قطعه بندی

### ۴-۳- نتایج طبقه بندی شکل تومور

برای ارزیابی نتایج هر مدل در این قسمت از معیارهای میانگین صحت<sup>۱</sup>، میانگین فراخوانی<sup>۲</sup> و دقت استفاده می کنیم که روابط مربوط به آن ها در روابط (۳-۴)، (۴-۴) و (۵-۴) به ترتیب آمده است. در مورد معیارهای صحت و فراخوانی برای هر کلاس به صورت جداگانه محاسبات را انجام داده و سپس میانگین می گیریم. همانطور که می دانیم توازنی باید میان صحت و فراخوانی باید برقرار باشد. پس از معیار امتیاز  $F_1$ <sup>۳</sup> که ارتباطی بین دو معیار قبلی برقرار می کند نیز نمایش داده می شود (رابطه (۶-۴)).

$$Precision = \frac{TP}{TP+FP} \quad (۳-۴)$$

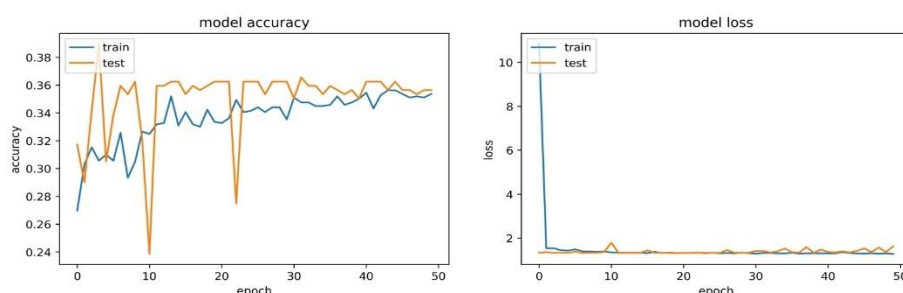
$$Recall = \frac{TP}{TP+FN} \quad (۴-۴)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (۵-۴)$$

$$F1_{Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (۶-۴)$$

حال هر یک از مدل های بررسی شده را به تنهایی بررسی می کنیم و در انتها جدول (۲-۴) را ارائه می دهیم تا یک نمای کلی از عملکرد همه مدل ها داشته باشیم.

مدل پیشنهادی مقاله اصلی، مدلی ساده و در عین حال یکی از بهترین مدل ها در بین مدل های بررسی شده است. تمام مدل ها در این قسمت با ۵۰ epoch اجرا شده اند. مدل پیشنهادی مقاله در از ابتدا و در مراحل ابتدایی همگرا شد. تصویر (۲-۴) نمایش رفتار دقت و تابع ضرر این مدل می باشد.



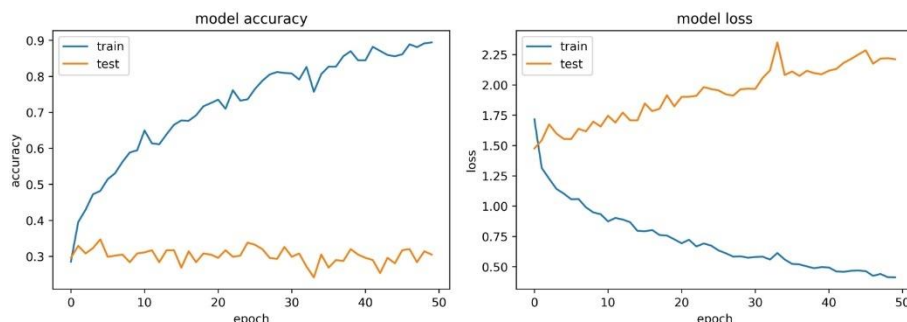
شکل (۲-۴) نمودارهای دقت و ضرر مدل پیشنهادی مقاله

<sup>1</sup> Precision

<sup>2</sup> Recall

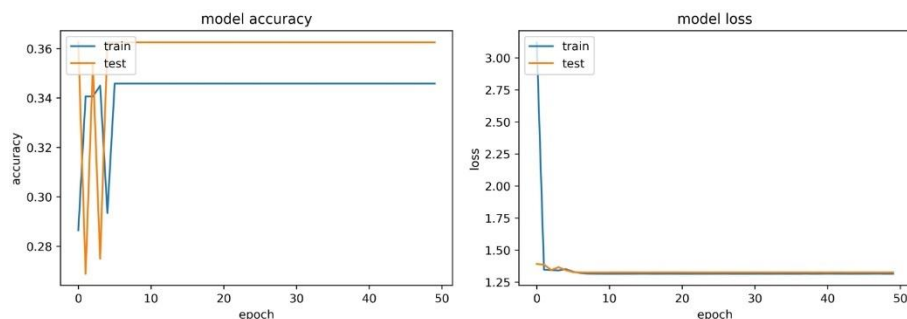
<sup>3</sup> F-Score

مدل ResNet50 برخلاف انتظار نتایج خوبی از خود نشان نداد. نمودارهای این مدل که در شکل (۳-۴) آمده نشان می دهد که این مدل دچار بیش برآزش شده است. البته متعادل نبودن تعداد داده های آموزشی در هر کلاس نیز مزید بر علت شده است زیرا به طور مثال تعداد داده ها با برچسب گرد بسیار کمتر از سایر دسته ها است.



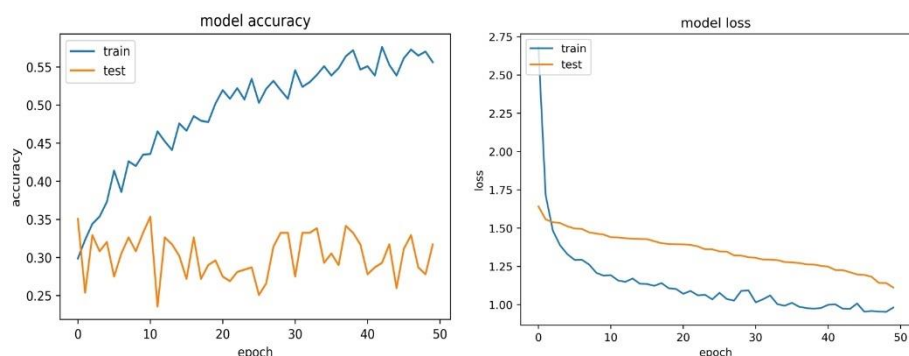
شکل (۳-۴) نمودار دقت و ضرر مدل ResNet

مدل MobileNet عملکردی مشابه روش پیشنهادی مقاله دارد. در این مدل نیز همانند تمام مدل های دیگر مشکل تفکیک کلاس گرد و بیضی و همینطور کلاس های نامنظم و ابرگونه وجود دارد. شکل (۴-۴) نمایش نمودارهای این مدل است.



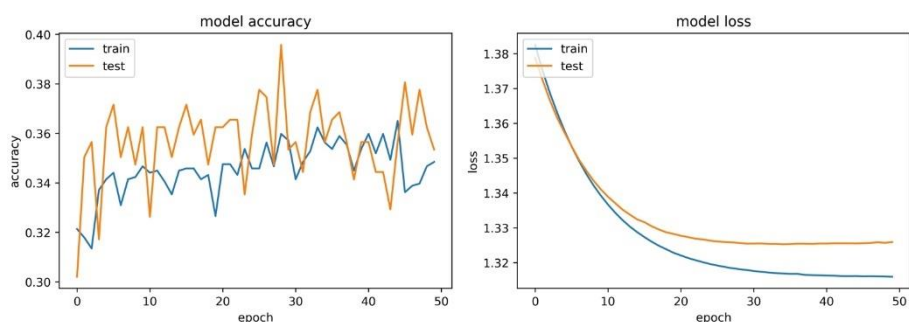
شکل (۴-۴) نمودار دقت و ضرر مدل Mobilenet

مدل VGG16 کمی بهتر از مدل پیشنهادی مقاله و MobileNet عمل کرد و به دقت بالاتری رسید اما همچنان نتایج قابل قبولی نداشته است. البته با توجه به سخت بودن این وظیفه حتی آخرین مدل های این زمینه نیز نتایج خیلی خوبی نمی گیرند. در شکل (۴-۵) نمودارهای این مدل را مشاهده می کنید.



شکل (۴-۵) نمودار دقت و ضرر مدل VGG16

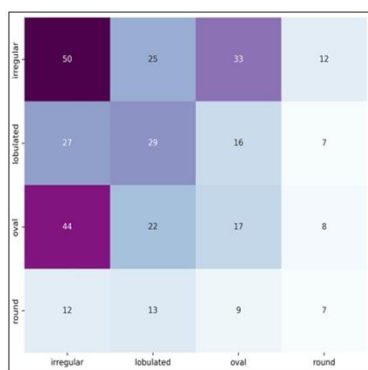
مدل پیشنهادی برخلاف تعداد پارامترهایی که از مدل های دیگر دارد عملکرد خوبی داشت و دقتی نزدیک به ۵۰٪ برای دسته بندی ۴ کلاسه بدست آورد. در این مدل نیز نمودار دقت که در شکل (۴-۶) آمده است نوسانات زیادی دارد اما بهترین مدل در میان سایر مدل هاست که از یادگیری انتقالی استفاده می کنند.



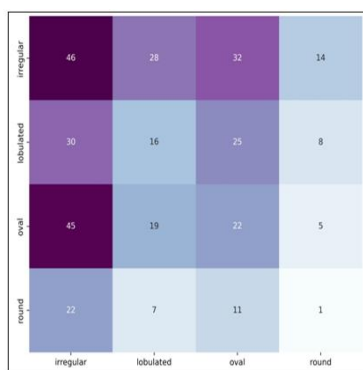
شکل (۴-۶) نمودار دقت و ضرر مدل پیشنهادی جدید

ماتریس همبستگی<sup>۱</sup> همه مدل ها در شکل (۴-۷) آورده شده است. همه ی این ماتریس ها نشان می دهند که تمایز قائل شدن بین دو جفت کلاس گرد، بیضی و نامنظم، ابرگونه به خوبی انجام نمی شود و مدل پیشنهادی سعی کرده است تا حدی آن را بهبود دهد.

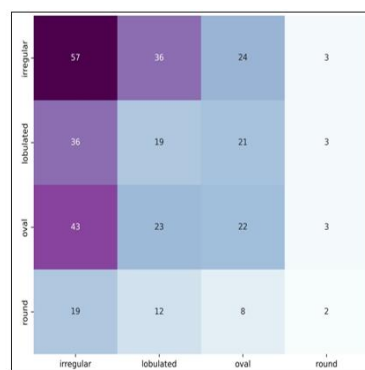
<sup>1</sup> Confusion Matrix



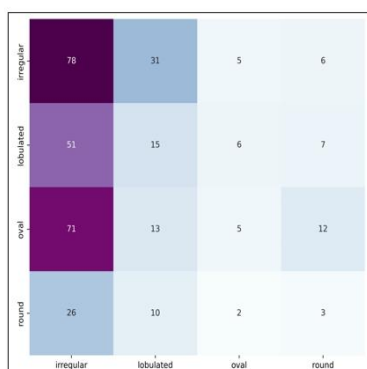
Article Model



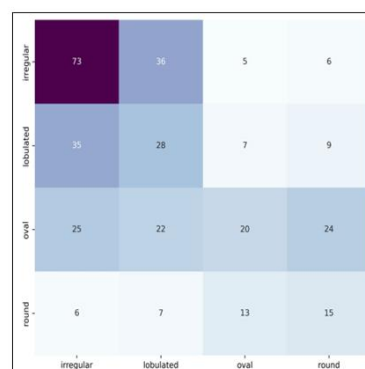
ResNet50



MobileNet



VGG16



Proposed Model

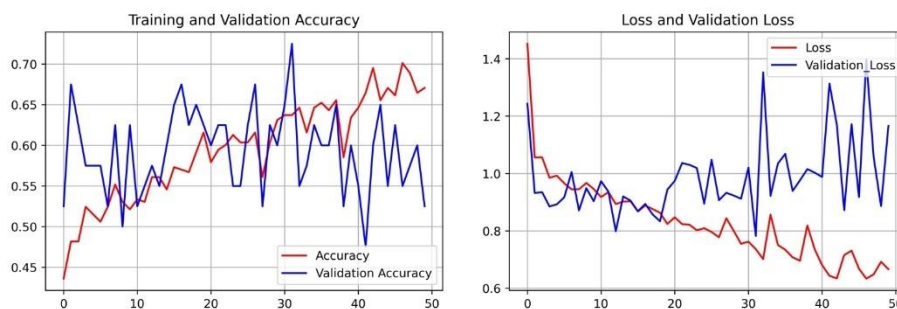
شکل (۴-۷) ماتریس همبستگی مدل های ارزیابی شده

جدول (۴-۲) جدول نتایج مدل های بررسی شده

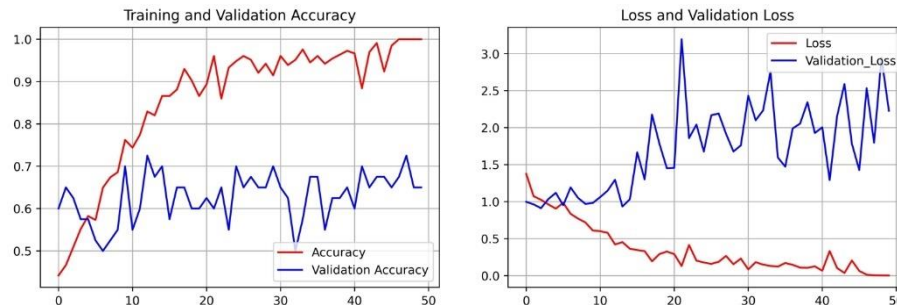
Model	Avg. Precision	Avg. Recall	F1-Score	Accuracy
ResNet50	0.2076	0.2130	0.2103	0.2568
MobileNet	0.2635	0.2515	0.2574	0.3021
VGG16	0.2286	0.2348	0.2317	0.3021
Article Model	0.2836	0.2853	0.2845	0.3112
Proposed Model	0.3765	0.3745	0.3755	0.4018

## ۴-۴- نتایج طبقه بندی نوع تومور

در این قسمت چهار مدل طبقه بندی نوع تومور را بر اساس دقت آن ها بررسی میکنیم. در ابتدا به بررسی عملکرد مدل های DenseNet121 و VGG19 میپردازیم. این دو مدل علی رغم ظرفیت بالای یادگیری و استفاده از وزن های از پیش آموخته شده نتوانستند دقت بالایی بگیرند. نتایج این پروژه نشان می دهد مدل DenseNet121 و VGG19 به ترتیب ۵۹٪ و ۵۷٪ دقت روی مجموعه آزمون دیتاست INbreast بدست آوردند. نمودار دقت و ضرر این دو مدل را در شکل های (۴-۸) و (۴-۹) مشاهده می کنید.

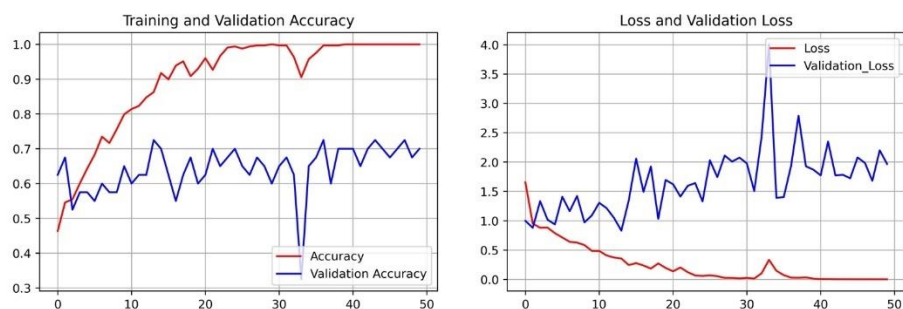


شکل (۴-۸) نمودار دقت و ضرر مدل DenseNet



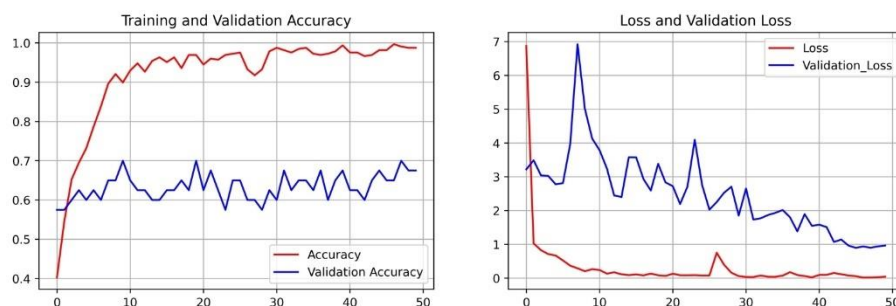
شکل (۴-۹) نمودار دقت و ضرر مدل VGG19

مدل ResNet50 که همانند دو مدل بالا از یادگیری انتقالی استفاده می کند و فقط لایه های کاملاً متصل آن آموزش می بیند عملکرد بهتری داشته است و دقتی برابر ۷۳٪ دارد. شکل (۴-۱۰) مسیر حرکتی آموزش این مدل در طول ۵۰ epoch را نمایش می دهد.



شکل (۴-۱۰) نمودار دقت و ضرر مدل ResNet50

در این پروژه تلاش کردیم با ارائه یک مدل پیشنهادی دقت بالاتری نسبت به سایر مدل ها بدست بیاوریم. اما مدل پیشنهادی ما با ۶۸٪ دقت در رتبه دوم دقت ها بعد از مدل ResNet قرار میگیرد. نمودارهای این مدل در شکل (۴-۱۱) و دقت های مدل ها در جدول (۴-۳) آمده است.



شکل (۴-۱۱) نمودار دقت و ضرر مدل پیشنهادی

جدول (۴-۳) دقت های بدست آمده از مدل ها

Model	Accuracy
DenseNet	59%
VGG19	57%
ResNet	73%
Proposed Method	68%



## فصل ۵:

### جمع‌بندی و پیشنهادها

## ۵-۱- جمع‌بندی

در این پروژه تلاش کردیم تا سه وظیفه اساسی قطعه‌بندی، طبقه‌بندی شکل و طبقه‌بندی نوع تومور را به کمک شبکه‌های عصبی عمیق انجام دهیم. دلیل انجام طبقه‌بندی شکل تومور وجود رابطه بین آن و نوع مولکول سرطانی است که در حال حاضر با آزمایشات سخت پزشکی مشخص می‌شود. در قسمت قطعه‌بندی تومور تصاویر را با سه حالت مختلف برای آموزش به شبکه‌ها دادیم. در حالتی که از قاب عریض استفاده شود میتوان نتایج بهتری برای قطعه‌بندی انتظار داشت. مدل U-Net در هر دو حالتی که به عنوان شبکه G در cGAN و یا به تنهایی برای قطعه‌بندی استفاده شود نتایج یکسانی دارد. در قسمت طبقه‌بندی شکل تومور با توجه به دشوار بودن وظیفه تعریف شده و همینطور متوازن نبودن دیتاست مورد استفاده نتایج مدل‌ها ضعیف‌تر از حد انتظار است. اما در این پروژه با ارائه یک مدل کانولوشنی دقتی بالاتر از مدل‌های از پیش آموزش دیده گرفتیم. مدل پیشنهادی ما بر پایه‌ی مدل مقاله است که تلاش شده تا ظرفیت یادگیری آن با کاهش سایز هسته‌ها و بالابردن عمق شبکه افزایش یابد. در آخرین قسمت پروژه با اضافه کردن وظیفه طبقه‌بندی نوع تومور که زمینه تحقیقاتی بسیار فعالی در این حوزه می‌باشد سعی کردیم تا با استفاده از مدل‌های رایج خوش‌خیم بودن، بدخیم بودن و یا نرمال بودن تصویر پستان را مشخص کنیم. در این بین مدل ResNet50 که لایه‌های کانولوشنی آن از پیش آموزش دیده‌اند به دقت قابل قبولی دست یافته‌است.

## ۵-۲- پیشنهادها

در بررسی مسائل مربوط به تصاویر پزشکی با توجه به حجم عظیم داده که ویژگی‌های مربوط به تک تم پیکسل‌های تصویر هستند، ممکن است استفاده از تعداد طبقه‌بندی‌های بیشتر، اجماع طبقه‌بندی‌ها و رای گیری بر اساس نتیجه تصمیم‌گیری آن‌ها موجب حصول نتیجه مطلوبی مخصوصاً در وظیفه طبقه‌بندی شکل تومور گردد. همچنین آموزش یک مدل که وزن هر کلاس (نسبت تعداد داده‌های موجود در هر کلاس به تعداد کل داده‌ها) را در دیتاست در نظر بگیرد میتواند بسیار مفید باشد.

## مراجع

- [1] R. L. Siegel *et al.*, "Colorectal cancer statistics, 2017," *CA: A Cancer Journal for Clinicians*, vol. 67, no. 3, pp. 177–193, May 2017, doi: 10.3322/CAAC.21395.
- [2] B. Lauby-Secretan *et al.*, "Breast-Cancer Screening — Viewpoint of the IARC Working Group," <http://dx.doi.org/10.1056/NEJMSr1504363>, Jun. 2015, doi: 10.1056/NEJMSr1504363.
- [3] J. Tang, R. M. Rangayyan, J. Xu, I. E. el Naqa, and Y. Yang, "Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 2, pp. 236–251, 2009, doi: 10.1109/TITB.2008.2009441.
- [4] "Computer-aided Diagnosis: The Tipping Point for Digital Pathology." <https://digitalpathologyassociation.org/blog/computer-aided-diagnosis-the-tipping-point-for-digital-pathology/> (accessed May 15, 2022).
- [5] Z. Jiao, X. Gao, Y. Wang, and J. Li, "A parasitic metric learning net for breast mass classification based on mammography," *Pattern Recognition*, vol. 75, 2018, doi: 10.1016/j.patcog.2017.07.008.
- [6] N. Cho, "Molecular subtypes and imaging phenotypes of breast cancer," *Ultrasonography*, vol. 35, no. 4, 2016, doi: 10.14366/usg.16030.
- [7] S. Liu, X. D. Wu, W. J. Xu, Q. Lin, X. J. Liu, and Y. Li, "Is there a correlation between the presence of a spiculated mass on mammogram and luminal a subtype breast cancer?," *Korean Journal of Radiology*, vol. 17, no. 6, 2016, doi: 10.3348/kjr.2016.17.6.846.
- [8] K. Tamaki *et al.*, "Correlation between mammographic findings and corresponding histopathology: Potential predictors for biological characteristics of breast diseases," *Cancer Science*, vol. 102, no. 12, 2011, doi: 10.1111/j.1349-7006.2011.02088.x.
- [9] V. K. Singh *et al.*, "Classification of breast cancer molecular subtypes from their micro-texture in mammograms using a VGGNet-based convolutional neural network," in *Frontiers in Artificial Intelligence and Applications*, 2017, vol. 300, doi: 10.3233/978-1-61499-806-8-76.
- [10] V. K. Singh *et al.*, "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network," *Expert Systems with Applications*, vol. 139, 2020, doi: 10.1016/j.eswa.2019.112855.
- [11] J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, 2015, doi: 10.1016/j.neunet.2014.09.003.
- [12] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015, vol. 07-12-June-2015, doi: 10.1109/CVPR.2015.7298965.
- [13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015, vol. 9351, doi: 10.1007/978-3-319-24574-4\_28.
- [14] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, 2017, doi: 10.1109/TPAMI.2016.2644615.

- [15] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42. 2017. doi: 10.1016/j.media.2017.07.005.
- [16] M. M. K. Sarker *et al.*, “SLSDeep: Skin lesion segmentation based on dilated residual and pyramid pooling networks,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11071 LNCS. doi: 10.1007/978-3-030-00934-2\_3.
- [17] H. Fu, J. Cheng, Y. Xu, D. W. K. Wong, J. Liu, and X. Cao, “Joint Optic Disc and Cup Segmentation Based on Multi-Label Deep Network and Polar Transformation,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, 2018, doi: 10.1109/TMI.2018.2791488.
- [18] N. Tajbakhsh *et al.*, “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, 2016, doi: 10.1109/TMI.2016.2535302.
- [19] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2015.
- [20] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” 2010. doi: 10.1109/cvpr.2009.5206848.
- [21] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, “Benign and malignant breast tumors classification based on region growing and CNN segmentation,” *Expert Systems with Applications*, vol. 42, no. 3, 2015, doi: 10.1016/j.eswa.2014.09.020.
- [22] Z. Liu, C. Zhuo, and X. Xu, “Efficient segmentation method using quantised and non-linear CeNN for breast tumour classification,” *Electronics Letters*, vol. 54, no. 12, 2018, doi: 10.1049/el.2018.1213.
- [23] N. Dhungel, G. Carneiro, and A. P. Bradley, “Deep learning and structured prediction for the segmentation of mass in mammograms,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9349, 2015. doi: 10.1007/978-3-319-24553-9\_74.
- [24] J. S. Cardoso, I. Domingues, and H. P. Oliveira, “Closed shortest path in the original coordinates with an application to breast cancer,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 29, no. 1, 2015, doi: 10.1142/S0218001415550022.
- [25] J. S. Cardoso, N. Marques, N. Dhungel, G. Carneiro, and A. P. Bradley, “Mass segmentation in mammograms: A cross-sensor comparison of deep and tailored features,” in *Proceedings - International Conference on Image Processing, ICIP*, 2018, vol. 2017-September. doi: 10.1109/ICIP.2017.8296579.
- [26] W. Zhu, X. Xiang, T. D. Tran, G. D. Hager, and X. Xie, “Adversarial deep structured nets for mass segmentation from mammograms,” in *Proceedings - International Symposium on Biomedical Imaging*, 2018, vol. 2018-April. doi: 10.1109/ISBI.2018.8363704.
- [27] M. A. Al-antari, M. A. Al-masni, M. T. Choi, S. M. Han, and T. S. Kim, “A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification,” *International Journal of Medical Informatics*, vol. 117, 2018, doi: 10.1016/j.ijmedinf.2018.06.003.
- [28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016-December. doi: 10.1109/CVPR.2016.91.
- [29] D. Yang *et al.*, “Automatic liver segmentation using an adversarial image-to-image network,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in*

- Artificial Intelligence and Lecture Notes in Bioinformatics*), 2017, vol. 10435 LNCS. doi: 10.1007/978-3-319-66179-7\_58.
- [30] V. K. Singh *et al.*, “Conditional generative adversarial and convolutional networks for X-ray breast mass segmentation and shape classification,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11071 LNCS. doi: 10.1007/978-3-030-00934-2\_92.
  - [31] L. Kurnianggoro, Wahyono, and K. H. Jo, “A survey of 2D shape representation: Methods, evaluations, and future research directions,” *Neurocomputing*, vol. 300, 2018, doi: 10.1016/j.neucom.2018.02.093.
  - [32] M. Masoumi and A. ben Hamza, “Spectral shape classification: A deep learning approach,” *Journal of Visual Communication and Image Representation*, vol. 43, 2017, doi: 10.1016/j.jvcir.2017.01.001.
  - [33] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “PointNet: Deep learning on point sets for 3D classification and segmentation,” in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017, vol. 2017-January. doi: 10.1109/CVPR.2017.16.
  - [34] Y. Xie, J. Zhang, Y. Xia, M. Fulham, and Y. Zhang, “Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT,” *Information Fusion*, vol. 42, 2018, doi: 10.1016/j.inffus.2017.10.005.
  - [35] P. Kisilev, E. Walach, S. Hashoul, E. Barkan, B. Ophir, and S. Alpert, “Semantic description of medical image findings: structured learning approach,” 2015. doi: 10.5244/c.29.171.
  - [36] 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, 2017, doi: 10.1109/TPAMI.2016.2577031.
  - [37] S. T. Kim, H. Lee, H. G. Kim, and Y. M. Ro, “ICADx: interpretable computer aided diagnosis of breast masses,” 2018. doi: 10.1117/12.2293570.
  - [38] Y. Chen, Y. Wang, and B. Yang, “Evolving hierarchical RBF neural networks for breast cancer detection,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2006, vol. 4234 LNCS-III. doi: 10.1007/11893295\_16.
  - [39] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, and F. Caselli, “Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing,” *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 7, 2008, doi: 10.1109/TIM.2007.915470.
  - [40] C. Prasetyo, A. Kardiana, and R. Yuliwulandari, “Breast Cancer Diagnosis using Artificial Neural Networks with Extreme Learning Techniques,” *International Journal of Advanced Research in Artificial Intelligence*, vol. 3, no. 7, 2014, doi: 10.14569/ijarai.2014.030703.
  - [41] B. Verma and J. Zakos, “A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 5, no. 1, 2001, doi: 10.1109/4233.908389.
  - [42] G. Jiménez and D. Racocanu, “Deep learning for semantic segmentation versus classification in computational pathology: Application to mitosis analysis in breast cancer grading,” *Frontiers in Bioengineering and Biotechnology*, vol. 7, no. MAY, 2019, doi: 10.3389/fbioe.2019.145.

- [43] A. al Nahid and Y. Kong, "Histopathological breast-image classification using local and frequency domains by convolutional neural network," *Information (Switzerland)*, vol. 9, no. 1, 2018, doi: 10.3390/info9010019.
- [44] A. Chaurasia and E. Culurciello, "LinkNet: Exploiting encoder representations for efficient semantic segmentation," in *2017 IEEE Visual Communications and Image Processing, VCIP 2017*, 2018, vol. 2018-January. doi: 10.1109/VCIP.2017.8305148.
- [45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, vol. 2016-December. doi: 10.1109/CVPR.2016.90.
- [46] G. Andrew and Z. Menglong, "Efficient convolutional neural networks for mobile vision applications," *Mobilenets*, 2017.
- [47] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017, vol. 2017-January. doi: 10.1109/CVPR.2017.243.