



دانشگاه صنعتی شریف  
دانشکده‌ی مهندسی برق

پایان‌نامه‌ی کارشناسی ارشد  
بیوالکتریک

عنوان:

# قطعه بندی ساختارهای در ریسک در تصاویر پزشکی با استفاده از روش های یادگیری ماشین

نگارش:

رضا کریمزاده

استاد راهنما:

دکتر عمادالدین فاطمی‌زاده

اسفند ۱۴۰۰

سلام افلا

## چکیده

نگارش پایان‌نامه علاوه بر بخش پژوهش و آماده‌سازی محتوا، مستلزم رعایت نکات فنی و نگارشی دقیقی است که در تهیه‌ی یک پایان‌نامه‌ی موفق بسیار کلیدی و مؤثر است. از آن جایی که بسیاری از نکات فنی مانند قالب کلی صفحات، شکل و اندازه‌ی قلم، صفحات عنوان و غیره در تهیه‌ی پایان‌نامه‌ها یکسان است، با استفاده از نرم‌افزار حروف‌چینی زی‌تک و افزونه‌ی زی‌پرشین یک قالب استاندارد برای تهیه‌ی پایان‌نامه‌ها ارائه گردیده است. این قالب می‌تواند برای تهیه‌ی پایان‌نامه‌های کارشناسی و کارشناسی ارشد و نیز رساله‌ی دکتری مورد استفاده قرار گیرد. این نوشتار به طور مختصر نحوه‌ی استفاده از این قالب را نشان می‌دهد.

**کلیدواژه‌ها:** پایان‌نامه، حروف‌چینی، قالب، زی‌پرشین

# فهرست مطالب

۹	۱ مرور ادبیات
۹	۱-۱ مقدمه
۹	۱-۲ تعاریف و مفاهیم اولیه
۹	۱-۲-۱ تعریف قطعه‌بندی
۱۰	۱-۲-۲ اهداف قطعه‌بندی تصاویر پزشکی
۱۱	۱-۲-۳ چالش‌های قطعه‌بندی تصاویر پزشکی
۱۲	۱-۲-۴ انواع قطعه‌بندی تصاویر پزشکی
۱۴	۱-۲-۵ ارزیابی قطعه‌بندی
۱۵	۱-۲-۶ معیارهای ارزیابی قطعه‌بندی
۱۸	۱-۳ روش‌های قطعه‌بندی
۱۸	۱-۳-۱ روش‌های آستانه‌گذاری
۱۹	۱-۳-۲ روش رشد ناحیه‌ای
۱۹	۱-۳-۳ روش‌های طبقه‌بندی
۲۱	۱-۳-۴ روش‌های خوشه‌بندی
۲۲	۱-۳-۵ مدل‌های شکل‌پذیر
۲۴	۱-۳-۶ مدل‌های مبتنی بر اطلس

۲۴	۱-۳-۷ سایر روش های قطعه بندی
۲۴	۱-۴ قطعه بندی با استفاده از شبکه های عصبی عمیق
۲۵	۱-۴-۱ طبقه بندی پیکسل ها برای قطعه بندی
۲۸	۱-۴-۲ شبکه های تمام کانولوشنی برای قطعه بندی
۳۲	۱-۴-۳ شبکه های تمام کانولوشنی سه بعدی برای قطعه بندی
۳۵	۱-۵ قطعه بندی تومور و ساختارهای در ریسک با یادگیری عمیق
۳۵	۱-۶ جمع بندی

## فهرست شکل‌ها

- ۱-۱ ناهمگنی شدت و میدان در یک تصویر ام‌آرای [۱] . . . . . ۱۲
- ۲-۱ شهود عملکرد معیار Dice [۲] . . . . . ۱۶
- ۳-۱ شهود عملکرد معیار Jaccard [۳] . . . . . ۱۷
- ۴-۱ شهود عملکرد فاصله‌ی Hausdorff [۴] . . . . . ۱۸
- ۵-۱ قطعه‌بندی تصویر ام‌آرای مغزی با روش K-means [۵] . . . . . ۲۱
- ۶-۱ قطعه‌بندی تصویر ام‌آرای مغزی با روش Fuzzy C-means [۶] . . . . . ۲۲
- ۷-۱ قطعه‌بندی با استفاده از مدل‌های شکل‌پذیر [۷] . . . . . ۲۲
- ۸-۱ روش تکامل منحنی و استخراج مجموعه-تراز صفر برای قطعه‌بندی [۸] . . . . . ۲۳
- ۹-۱ روش مبتنی بر اطلس بری قطعه‌بندی تصویر مغزی [۹] . . . . . ۲۴
- ۱۰-۱ لغزاندن یک پنجره روی تصویر ام‌آرای مغزی و استخراج پیچ [۱۰] . . . . . ۲۶
- ۱۱-۱ شبکه‌ی کانولوشنی برای طبقه‌بندی پیکسل‌ها وسطی پیچ‌ها [۱۰] . . . . . ۲۶
- ۱۲-۱ معماری شبکه‌ی کانولوشنی برای طبقه‌بندی بر اساس پیچ ساختارهای در ریسک سر و گردن [۱۱] . . . . . ۲۷
- ۱۳-۱ نتایج بدست آمده در قطعه‌بندی ساختارهای در ریسک سر و گردن (قرمز: خروجی شبکه، سبز: قطعه‌بندی مطلوب) [۱۱] . . . . . ۲۷
- ۱۴-۱ قطعه‌بندی تصاویر با استفاده از رویکرد شبکه‌های تمام کانولوشنی [۱۲] . . . . . ۲۸

- ۱-۱۵ استخراج نقشه‌ی گرمایی وجود یک شی در تصویر با استفاده از ویژگی‌های استخراج شده توسط شبکه‌ی طبقه‌بند [۱۲] . . . . . ۲۹
- ۱-۱۶ تغییر ابعاد ویژگی‌های استخراج شده بعد از لایه‌های ادغام به ابعاد تصویر اصلی [۱۲] ۲۹
- ۱-۱۷ معماری شبکه‌ی شامل کانولوشن و دیکانولوشن برای بهبود دقت قطعه‌بندی [۱۳] ۳۰
- ۱-۱۸ نمایش تفاوت عملکرد کانولوشن و دیکانولوشن [۱۳] . . . . . ۳۰
- ۱-۱۹ معماری شبکه‌ی UNet [۱۴] . . . . . ۳۱
- ۱-۲۰ معماری شبکه‌ی SegNet [۱۵] . . . . . ۳۲
- ۱-۲۱ معماری شبکه‌ی UNet سه‌بعدی [۱۶] . . . . . ۳۳
- ۱-۲۲ یک بلوک با اتصال residual [۱۷] . . . . . ۳۳
- ۱-۲۳ کانولوشن‌های منبسط شده [۱۸] . . . . . ۳۴
- ۱-۲۴ معماری شبکه‌ی سه‌بعدی شامل اتصالات residual و کانولوشن انبساطی [۱۷] . . ۳۴
- ۱-۲۵ معماری شبکه‌ی VNet [۱۹] . . . . . ۳۵

## فهرست جدول‌ها



# فصل ۱

## مرور ادبیات

### ۱-۱ مقدمه

در این فصل سوابق و مطالعات انجام شده برای قطعه‌بندی تصاویر پزشکی بررسی خواهند شد. در ابتدا مفاهیم مقدماتی و تعاریف مرتبط با این حوزه معرفی می‌شوند و سپس روش‌های پیشنهاد شده، برای قطعه‌بندی تصاویر پزشکی مورد مطالعه قرار خواهند گرفت. در نهایت رویکردهای جدید قطعه‌بندی با استفاده از روش‌های یادگیری عمیق معرفی و بررسی می‌شوند.

### ۲-۱ تعاریف و مفاهیم اولیه

در این قسمت تعاریف و مفاهیم اولیه‌ی مرتبط با قطعه‌بندی ارائه می‌گردد. در ابتدا مفهوم قطعه‌بندی معرفی می‌شود و سپس ...

#### ۱-۲-۱ تعریف قطعه‌بندی

همانطور که در فصل دوم اشاره گردید، یک تصویر پزشکی می‌تواند دوبعدی (تصاویر رادیوگرافی) یا سه‌بعدی (تصاویر ام‌آر‌آی و سی‌تی اسکن) که به عناصر سازنده‌ی این تصاویر بر اساس این ابعاد، پیکسل و یا واکسل گفته می‌شود. برای قطعه‌بندی تصاویر تعاریف متنوعی ارائه شده‌است که در زیر به بررسی

هر یک پرداخته می‌شود.

- به جزءبندی یک تصویر به مناطق غیر هم‌پوشان همگن که یک ویژگی یا یک معیار مشترک مانند شدت نوری<sup>۱</sup>، شکل<sup>۲</sup> و یا بافت<sup>۳</sup> همسان دارند قطعه‌بندی گفته می‌شود.
- به روند نسبت دادن یک برجسب با مفهوم بیولوژیکی خاص به پیکسل‌ها و یا واکسل‌ها که این پیکسل‌ها و واکسل‌ها یک ویژگی خاص مشترک دارند و یا متعلق به یک آناتومی مشترک هستند.
- جداسازی ساختارهای مد نظر از یکدیگر و از پس‌زمینه را قطعه‌بندی می‌گویند [۷].

با توجه به تعاریف انجام شده، می‌توان یک تعریف ریاضی برای این عملیات ارایه داد. اگر تصویر را  $I$  بنامیم، آنگاه قطعه‌بندی یعنی مشخص کردن مجموعه‌هایی  $(S_k)$  در تصویر که  $S_k \subset I$  و اجتماع این مجموعه‌ها کل تصویر را بازسازی کند و این مجموعه‌ها هیچ اشتراکی نداشته باشند:

$$\bigcup_{k=1}^K S_k = I \quad (1-1)$$

$$\text{where } S_k \cap S_j = \emptyset \quad \text{for } k \neq j$$

به صورت ایده‌آل قطعه‌بندی به صورتی انجام می‌شود که هر یک از  $S_k$ ‌ها یک ساختار و آناتومی خاص را نشان می‌دهد. شرط دیگر در این فرمول‌بندی متصل بودن هریک از این  $S_k$ ‌هاست که برای ساختارهای آناتومیک یکپارچه این شرط باید برقرار شود. اگر از این شرط چشم‌پوشی کنیم قطعه‌بندی در سطح پیکسل (واکسل) صورت می‌گیرد و در مواردی استفاده می‌شود که ساختارهای مورد نظر از یکدیگر گسسته هستند اما خواص مشابه دیگری دارند [۲۰].

## ۱-۲-۲ اهداف قطعه‌بندی تصاویر پزشکی

در روند قطعه‌بندی تصاویر پزشکی، جزییات مورد نظر در استخراج قطعه‌ها، بسیار به کاربرد کیلینیکی بستگی دارد. هدف قطعه‌بندی بهبود نمایش تصاویر و اجزای مختلف آن برای تشخیص و درمان کارآمدتر است. به طور کلی از مهم‌ترین اهداف قطعه‌بندی می‌توان به موارد زیر اشاره کرد:

<sup>1</sup>Intensity

<sup>2</sup>Shape

<sup>3</sup>Texture

- تشخیص بیماری
- کمی سازی پیشرفت بیماری
- مطالعه‌ی پیشرفت بیماری
- استخراج ویژگی
- تهیه‌ی نقشه‌ی درمان
- موقعیت یابی تومور
- تخمین حجم تومور
- تطبیق تصاویر

به‌طور کلی قطعه‌بندی تصاویر پزشکی سهم عمده‌ای در تحلیل و پردازش تصاویر پزشکی دارد که امکان بررسی مشکلات عملکردی ساختارهای آناتومیکی مختلف را فراهم می‌سازد و تمام فاکتورهای لازم برای تشخیص، کمی‌سازی و مطالعه‌ی پیشرفت بیماری را میسر می‌سازد [۲۱، ۲۲].

### ۱-۲-۳ چالش‌های قطعه‌بندی تصاویر پزشکی

قطعه‌بندی تصاویر پزشکی، با مشکلات زیادی مواجه است که کیفیت قطعه‌بندی نهایی به دلیل وجود برخی اشکالات در تصویر اصلی پدید می‌آید. از جمله‌ی این اشکالات می‌توان به موارد زیر اشاره نمود [۲۳].

#### نویز

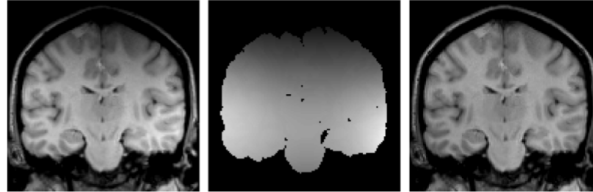
با حضور نویز در تصویر عدم قطعیت بالا می‌رود و مقدار شدت پیکسل‌ها در تصویر تغییر می‌کند که این تغییر باعث دگرگونی ساختار همگن بافت‌ها می‌شود و در نتیجه قطعه‌بندی دشوار می‌گردد [۲۴].

#### ناهمگنی شدت

ناهمگنی شدت<sup>۴</sup> به تغییرات جعلی و هموار در شدت‌های داخل تصویر گفته می‌شود که نمایانگر یک بافت با یک ویژگی مشترک هستند اما به دلیل تغییرات در میدان دستگاه تصویرگر، پدید می‌آیند. این مشکل اغلب در تصاویر ام‌آرای به دلیل تغییر میدان مغناطیسی آن بوجود می‌آیند. شکل ۱-۱ در سمت چپ یک تصویر مغزی دچار این مشکل نشان داده شده است و در شکل وسط، ناهمگونی میدان دستگاه

<sup>4</sup>Intensity In-homogeneity

تصویربرداری استخراج شده و در نهایت در شکل راست با تقسیم تصویر مشکل دار بر ناهمگنی میدان، تصویر ترمیم شده بدست آمده است [۱].



شکل ۱-۱: ناهمگنی شدت و میدان در یک تصویر ام‌آرای [۱]

### اثر حجم جزئی

اثر حجم جزئی<sup>۵</sup> به این دلیل بوجود می‌آید که هر واکسل یک حجم خاص از فضا را با یک مقدار شدت تعریف می‌کند حال اگر دو بافت مختلف در یک واکسل قرار گیرد، مقدار شدت این واکسل ترکیبی از جمع وزن دار شدت‌های این دو بافت است و بنابراین اثر حجم جزئی لبه‌های تصویر را هموارتر و محوتر می‌کند [۲۵].

علاوه بر چالش‌های ذکر شده، چالش‌های دیگری مانند: حرکت بیمار در طی تصویربرداری، شدت‌های یکسان در بافت‌های مختلف، لبه‌های ضعیف در بافت‌های نرم و ... باعث دشوارتر شدن قطعه‌بندی تصاویر پزشکی می‌شوند.

## ۴-۲-۱ انواع قطعه‌بندی تصاویر پزشکی

انواع قطعه‌بندی تصاویر پزشکی را بر اساس بعد تصویر و چالش‌های موجود در قطعه‌بندی، می‌توان به دسته‌های مختلفی تقسیم‌بندی کرد. در زیر به بررسی انواع دسته‌بندی روش‌های قطعه‌بندی پرداخته خواهد شد.

<sup>۵</sup>Partial Volume Effect

### قطعه‌بندی بر اساس بعد تصویر

همانطور که پیش‌تر اشاره گردید، تصاویر پزشکی می‌تواند دوبعدی یا سه‌بعدی باشند بنابراین روش‌های مختلفی بر اساس بعد تصاویر ارائه شده است که می‌تواند در تصاویر دو بعدی یا سه‌بعدی قطعه‌بندی را انجام دهند. همچنین برای تصاویر سه‌بعدی می‌توان الگوریتم قطعه‌بند دوبعدی را در لایه‌های مختلف تصویر اجرا کرد و نتیجه را پشت سر هم قرار داد تا قطعه‌بندی تصویر سه‌بعدی انجام شود. مزیت این روش کاهش هزینه‌های محاسباتی، پیاده‌سازی ساده‌تر و کاهش حافظه‌ی مورد نیاز است. روش‌های سه‌بعدی برای کاربردهای خاص مانند استخراج رویه و سطح ساختارهای سه‌بعدی ارائه شده‌اند که در این حالت نمی‌توان از روش‌های دوبعدی استفاده نمود [۲۶].

### قطعه‌بندی نرم و سخت

قطعه‌بندی سخت<sup>۶</sup> برای یک پیکسل (واکسل) به صورت صفر و یک تعریف می‌شود، یا پیکسل متعلق به کلاس مورد نظر است یا نیست. تعریف ریاضیاتی قطعه‌بندی سخت را می‌توان به صورت فرمول ۲-۱ ارائه داد که در آن تابع عضویت پیکسل  $x_{i,j}$  در موقعیت  $i, j$  تصویر به کلاس  $m$  است. در این نوع قطعه‌بندی امکان حل مشکل اثر حجم جزیی وجود ندارد.

$$m_k(x_{i,j}) = \begin{cases} 1, & x_{i,j} \in S_k \\ 0, & \text{otherwise} \end{cases} \quad (2-1)$$

در دست دیگر، قطعه‌بندی نرم<sup>۷</sup> قرار دارد که امکان طبقه‌بندی پیکسل‌ها به کلاس‌های مختلف را با یک احتمال تعلق نسبت به تمام کلاس‌ها می‌دهد. فرمول‌بندی این نوع قطعه‌بندی را می‌توان در معادلات زیر ملاحظه نمود. که در آن  $K$  تعداد کل کلاس‌های موجود است.

$$\begin{cases} 0 \leq m_k(x_{i,j}) \leq 1 \\ \sum_{k=1}^K m_k(x_{i,j}) = 1 \end{cases} \quad (3-1)$$

با استفاده از قطعه‌بندی نرم، می‌توان اثرات همپوشانی پیکسل‌ها و اثر حجم جزیی را حل نمود. قطعه‌بندی نرم را با یک آستانه‌گذاری می‌توان به قطعه‌بندی سخت تبدیل کرد [۲۷].

<sup>۶</sup>Hard Segmentation

<sup>۷</sup>Soft Segmentation

### قطعه‌بندی گسسته و پیوسته

در قطعه‌بندی گسسته، خروجی نهایی یک نقشه‌ی گسسته، مانند تصویر اولیه است و بر اساس فاصله‌ی فضایی پیکسل‌ها (واکسل‌ها) این قطعه‌بندی صورت می‌گیرد. در سمت دیگر، قطعه‌بندی پیوسته، در یک فضای پیوسته عمل می‌کند و امکان قطعه‌بندی به صورت زیرپیکسل<sup>۸</sup> وجود دارد. به عنوان مثال در مدل‌های شکل‌پذیر قطعه‌بندی نهایی به صورت پیوسته است [۲۸].

### قطعه‌بندی بر اساس تعامل کاربر

انواع روش‌های قطعه‌بندی را بر اساس تعامل کاربر می‌توان به دسته‌های زیر تقسیم‌بندی نمود:

- **روش‌های دستی:** در این گونه روش‌ها، متخصص به صورت دستی و تعیین طبقه‌ی هر یک از پیکسل‌ها (واکسل‌ها) اقدام به قطعه‌بندی می‌کنند که یک عملیات بسیار زمان‌بر و خسته‌کننده برای متخصص و همچنین هزینه‌بر برای استخدام متخصص است.
- **روش‌های نیمه خودکار:** در این نوع روش‌ها، متخصص با ایجاد یک نقطه‌ی اولیه برای شروع الگوریتم، الگوریتم را راه‌اندازی می‌کند و در نهایت خروجی قطعه‌بندی ایجاد می‌شود.
- **روش‌های خودکار:** در این نوع روش‌ها عملیات قطعه‌بندی کاملاً به صورت خودکار انجام می‌شود و نتیجه می‌توان تعداد بسیار زیادی تصویر را با یک دقت خاص، بدون دخالت فرد، بدست آورد. از دیگر مزیت‌های این روش‌ها نسبت به روش‌های قبل سرعت بالاتر است [۲۹].

## ۱-۲-۵ ارزیابی قطعه‌بندی

برای کمی‌سازی عملکرد روش قطعه‌بندی روش‌های ارزیابی بسیار حیاتی هستند. روش‌های ارزیابی را می‌توان به دو دسته‌ی عمده تقسیم‌بندی کرد:

- **مقایسه با قطعه‌بندی متخصص:** این روش پر استفاده‌ترین روش برای ارزیابی الگوریتم‌های قطعه‌بندی است که در آن از قطعه‌بندی شخص متخصص به عنوان استاندارد استفاده می‌شود و در نهایت با معیارهای ارزیابی بین خروجی مدل قطعه‌بند و استاندارد موجود، عملکرد قطعه‌بندی

<sup>۸</sup>Subpixel

کمی سازی می گردد. از جمله مشکلات این روش می توان به، زمان بر و هزینه بر بودن قطعه بندی توسط متخصص و همچنین اختلاف نظر بین متخصصین مختلف اشاره کرد.

• استفاده از فانتوم های فیزیکی و ریاضیاتی: با استفاده از فانتوم های<sup>۹</sup> ریاضیاتی و فیزیکی، اجزای مختلف فانتوم به طور کامل از یکدیگر تفکیک پذیر هستند و یک تصویر سازی دقیق از روند تصویر برداری ایجاد می کنند. بنابراین با قرار دادن قطعه بندی فانتوم به عنوان استاندارد، می توان الگوریتم قطعه بندی را ارزیابی نمود [۳۰].

با توجه به تعریف های فوق ملاحظه گردید که لازم است برای کمی سازی ارزیابی روش قطعه بندی معیارهایی تعریف گردد. در ادامه به معرفی این معیارها پرداخته می شود.

## ۱-۲-۶ معیارهای ارزیابی قطعه بندی

برای کمی سازی عملکرد قطعه بندی معیارهای بسیار زیادی ارایه گردیده است که از جمله مهمترین آنها می توان به موارد زیر اشاره نمود:

- ماتریس درهم آمیختگی<sup>۱۰</sup>
- اختصاص<sup>۱۴</sup>
- دقت<sup>۱۱</sup>
- معیار Dice
- حساسیت<sup>۱۲</sup>
- معیار Jaccard
- صحت<sup>۱۳</sup>
- فاصله ی Hausdorff

معیارهای پر استفاده و شناخته شده در قطعه بندی تصاویر شامل سه معیار آخر (معیار Dice، معیار Jaccard و فاصله ی Hausdorff) هستند که در ادامه به بررسی بیشتر آنها پرداخته می شود.

<sup>9</sup>Phantom

<sup>10</sup>Confusion Matrix

<sup>11</sup>Accuracy

<sup>12</sup>Sensitivity

<sup>13</sup>Precision

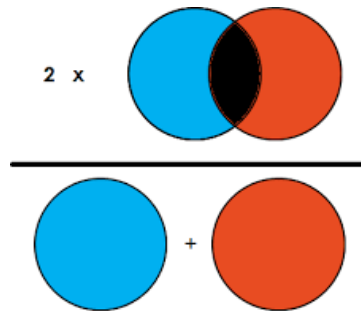
<sup>14</sup>Specificity

### معیار Dice

معیار Dice یک نشانگر آماری برای سنجش شباهت میان دو نمونه است. این معیار با نام F1-Score نیز شناخته می‌شود. تعریف ریاضیاتی این معیار به صورت زیر است که A و B در بحث قطعه‌بندی خروجی الگوریتم و استاندارد مورد نظر هستند.

$$DSC(A, B) = \frac{2 * |A \cap B|}{|A| + |B|} \quad (۴-۱)$$

شکل ۱-۲ شهود عملکرد این معیار را نشان می‌دهد. در واقع هرچقدر همپوشانی دو ناحیه بیشتر باشد این معیار به یک نزدیک‌تر می‌شود و اگر دو ناحیه به طور کامل از یکدیگر جدا باشند مقدار این معیار برابر صفر می‌گردد.



شکل ۱-۲: شهود عملکرد معیار Dice [۲]

### معیار Jaccard

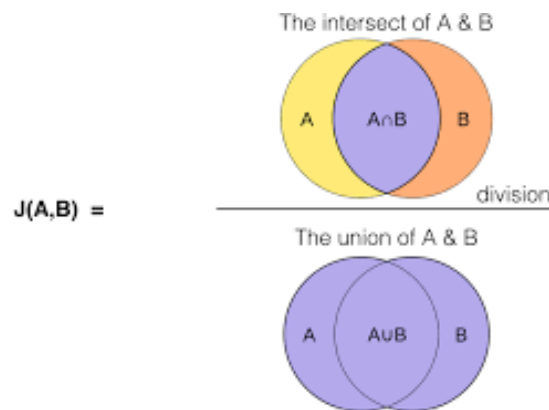
معیار Jaccard نیز مانند معیار Dice یک نشانگر آماری است که میزان شباهت و تفاوت را میان مجموعه‌ی نمونه‌ها نشان می‌دهد. این معیار با نام اشتراک بر اجتماع<sup>۱۵</sup> نیز شناخته می‌شود. فرمول‌بندی این معیار به صورت زیر است.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (۵-۱)$$

برای شهود بیشتر، شکل ۱-۳ را در نظر بگیرید. هرچه میزان هم‌پوشانی دو ناحیه بیشتر شود این معیار به یک نزدیک‌تر می‌شود و اگر این دو ناحیه از هم جدا باشند، مقدار این معیار برابر با صفر خواهد شد.

<sup>15</sup>Intersection over Union





شکل ۱-۳: شهود عملکرد معیار Jaccard [۲]

### فاصله‌ی Hausdorff

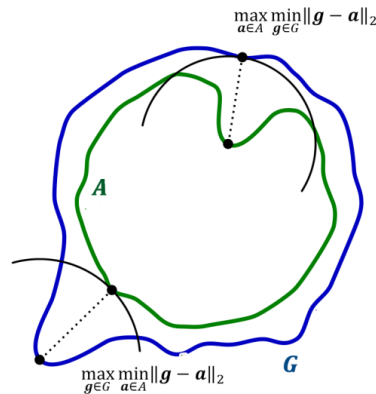
فاصله‌ی Hausdorff در واقع معیاری است که مشخص می‌کند دو مجموعه داده چه مقدار از یکدیگر دوراند. اگر نقاط متناظر در دو مجموعه داده نزدیک به یکدیگر باشند، این فاصله به صفر نزدیک می‌شود و هر چه این نقاط متناظر فاصله‌ی بیشتری بگیرند، مقدار این معیار بزرگتر خواهد شد. تعریف ریاضیاتی این فاصله به صورت زیر است:

$$HD(G, A) = \max(hd(G, A), hd(A, G)) \quad (۱-۶)$$

$$hd(G, A) = \max_{g \in G} \min_{a \in A} \|g - a\|_2$$

$$hd(A, G) = \max_{a \in A} \min_{g \in G} \|g - a\|_2$$

که در آن،  $A$  خروجی الگوریتم قطعه‌بندی و  $G$  استاندارد مطلوب و  $\|\cdot\|_2$  نشان دهنده‌ی نرم اقلیدسی است. در فاصله Hausdorff به صورت شهودی مطابق شکل ۱-۴ بیشینه‌ی کمینه فاصله‌های لبه‌های خروجی الگوریتم با لبه‌های استاندارد مطلوب محاسبه می‌شود.



شکل ۱-۴: شهود عملکرد فاصله‌ی Hausdorff [۴]

### ۳-۱ روش‌های قطعه‌بندی

با نگاه در طول زمان، روش‌ها و تکنیک‌های پیشنهاد شده‌ی قطعه‌بندی با گذشت زمان پیشرفت چشمگیری داشته‌اند. روش‌ها و الگوریتم‌های معرفی شده در این حوزه بسیار فراوان هستند. برای بررسی این روش‌ها در این مطالعه، بر اساس تمرکز روش ارایه شده بر یک ویژگی غالب در تصویر یا یک زمینه‌ی خاص، دسته‌بندی صورت گرفته است و سپس در هر دسته این روش‌ها معرفی شده‌اند. در ادامه به بررسی روش‌های موجود در زمینه‌ی قطعه‌بندی پرداخته می‌شود.

#### ۱-۳-۱ روش‌های آستانه‌گذاری

آستانه‌گذاری<sup>۱۶</sup> یکی از پر استفاده‌ترین روش‌های قطعه‌بندی است که بر اساس شدت پیکسل‌ها عمل می‌کند. با اعمال یک آستانه به شدت پیکسل‌های تصویر می‌توان پیش‌زمینه را از پس‌زمینه تفکیک داد و یک قطعه‌بندی دو کلاس صورت داد. معادله‌ی ۷-۱ توصیف ریاضیاتی قطعه‌بندی ( $S$ ) یک تصویر دوبعدی  $X$  را در مختصات  $i, j$  با آستانه‌ی  $T$  نشان می‌دهد.

$$S(i, j) = \begin{cases} 1 & X(i, j) \geq T \\ 0 & X(i, j) < T \end{cases} \quad (7-1)$$

<sup>16</sup>Thresholding

روش آستانه‌گذاری را می‌توان با تعیین چندین آستانه، به یک قطعه‌بند چند کلاس تعمیم داد. نکته‌ی مهم در این روش، پیدا کردن آستانه بهینه است. برای این منظور روش‌های زیادی ارائه گردیده است که از جمله این روش‌ها می‌توان به موارد زیر اشاره نمود:

- **تیوری تصمیم بیز<sup>۱۷</sup>**: در این روش، در ابتدا با متدهای تخمین تابع توزیع (مانند مدل مخلوط گاوسی<sup>۱۸</sup>)، توزیع شدت پیکسل‌ها استخراج می‌شود. در نهایت روش بیز با کمینه کردن مجموع مقدار مثبت و منفی کاذب<sup>۱۹</sup> مقدار آستانه‌ی بهینه را استخراج می‌کند [۳۱].
- **روش آستانه‌گذاری اوتسو<sup>۲۰</sup>**: در این روش با بیشینه‌سازی یک تابع هزینه، مقدار آستانه‌ی بهینه بدست می‌آید. این تابع هزینه از دو قسمت واریانس بین کلاس‌ها و واریانس درون کلاسی تشکیل شده است که با بیشینه کردن تابع هزینه، واریانس بین کلاس‌ها بیشینه و واریانس درون کلاس‌ها کمینه می‌شود و در نهایت آستانه‌ی مطلوب بدست می‌آید [۳۲].

### ۱-۳-۲ روش رشد ناحیه‌ای

رشد ناحیه‌ای یک روش قطعه‌بندی تعاملی است که یک هسته‌ی اولیه توسط کاربر برای قطعه‌بندی انتخاب می‌شود. سپس با مقایسه‌ی شدت روشنایی هسته‌ی اولیه با پیکسل‌های کناری، اگر این اختلاف کوچکتر از یک آستانه بود این ناحیه توسعه پیدا می‌کند تا در نهایت به پیکسل‌های با شدت‌های متفاوت برسد. مشکل این روش قطعه‌بندی نیمه-خودکار بودن آن است و لازم است هر بار توسط یک متخصص برای هر ناحیه، مقدار هی اولیه صورت گیرد. همچنین قبل از قطعه‌بندی با این روش باید پیش پردازش‌هایی برای بهبود و هم‌نواخت کردن شدت پیکسل‌های متعلق به یک بافت صورت گیرد [۳۳].

### ۱-۳-۳ روش‌های طبقه‌بندی

روش‌های مبتنی بر طبقه‌بندی برای قطعه‌بندی تصاویر، معمولاً با استخراج نقشه‌ی ویژگی<sup>۲۱</sup> از تصویر و دادن این ویژگی‌ها به یک طبقه‌بند برای تصمیم‌گیری نهایی صورت می‌گیرد. این نقشه‌ی ویژگی می‌تواند

<sup>17</sup>Bayes Decision Theorem

<sup>18</sup>Gaussian Mixture Models (GMM)

<sup>19</sup>False Positive and False Negative

<sup>20</sup>Otsu's Thresholding

<sup>21</sup>Feature Map

شامل ویژگی‌های سطح پایین مثل شدت روشنایی و یا پاسخ تصویر نسبت به یک فیلتر تا ویژگی‌های سطح بالاتر استخراج شده توسط الگوریتم‌های دیگر باشد<sup>۲۲</sup>. روش‌های طبقه‌بندی معمولاً از مدل‌های شناخت الگو بهره می‌برند و به نحوی می‌توان این روش‌ها را یادگیری با سرپرست نامید که نتیجه‌ی مطلوب و استاندارد مورد نظر در دسترس است [۳۴]. روش‌های طبقه‌بندی را می‌توان به تکنیک‌های زیر تقسیم‌بندی نمود.

### روش بیشینه‌کردن شباهت

در روش بیشینه‌کردن شباهت<sup>۲۳</sup> هدف ارایه دادن یک طبقه‌بند بر اساس تابع احتمال با کمترین خطای ممکن است. در این تکنیک، نقشه‌ی ویژگی استخراج شده باید خاصیت احتمالاتی ویژگی‌ها را داشته باشد تا این الگوریتم به درستی عمل کند [۳۵].

### نزدیک‌ترین همسایگی

روش نزدیک‌ترین همسایگی<sup>۲۴</sup> یک روش غیر پارامتری<sup>۲۵</sup> است که با هدف طبقه‌بندی ارایه شده است. در این روش با بررسی ویژگی یک نقطه در نقشه‌ی ویژگی و مقایسه با نقاط نزدیک در نقشه‌ی ویژگی برچسب مطلوب اعمال می‌گردد.

### شبکه‌های عصبی

استفاده از شبکه‌های عصبی برای طبقه‌بندی یک روش پر استفاده است که در این تکنیک ابتدا ویژگی‌های مورد نظر از تصویر استخراج می‌گردد و با رویکرد یادگیری با سرپرست در صورت وجود خروجی مطلوب و یا یادگیری بدون سرپرست در صورت نبود استاندارد مورد نظر، شبکه نسبت به طبقه‌بندی یادگیری را صورت می‌دهد. توجه شود که در این نوع روش‌ها معمولاً ویژگی‌ها از پیش استخراج شده‌اند و عملکرد شبکه‌ی عصبی فقط طبقه‌بندی است [۳۶].

<sup>۲۲</sup> مانند الگوریتم‌های SURF SIFT و ...

<sup>۲۳</sup> Maximum Likelihood

<sup>۲۴</sup> Nearest neighbor

<sup>۲۵</sup> Non-Parametric

### درخت تصمیم

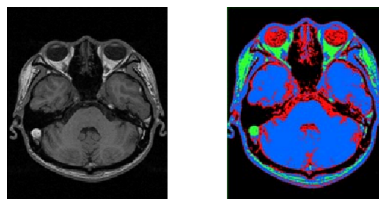
درخت تصمیم<sup>۲۶</sup> یک تکنیک ساده و پر استفاده برای طبقه‌بندی است. طبقه‌بند درخت تصمیم مجموعه‌ای از پرسش‌ها و ارزیابی‌ها را در یک شکل درختی ساختار می‌دهد و با بررسی ویژگی‌ها در هر شاخه در نهایت به یک تصمیم بله یا خیر می‌رسد [۳۷].

### ۱-۳-۴ روش‌های خوشه‌بندی

روش‌های خوشه‌بندی مانند روش‌های طبقه‌بندی یک هدف را دنبال می‌کنند اما در این رویکرد بر عکس طبقه‌بندی از یادگیری بدون سرپرست استفاده می‌شود و داده‌ها بدون حضور استاندارد مطلوب، قطعه‌بندی می‌شوند. دو نوع روش خوشه‌بندی متداول در حوزه‌ی قطعه‌بندی بسیار استفاده می‌شوند که در ادامه به معرفی هر یک پرداخته می‌شود.

#### K-means

در تکنیک K-means، ابتدا تعداد خوشه‌ی مدنظر توسط کاربر مشخص می‌شود و سپس به ازای هر خوشه یک مرکز خوشه به صورت تصادفی در فضای ویژگی قرار می‌گیرد. با محاسبه‌ی فاصله‌ی هر ویژگی نسبت به هر خوشه و سپس نسبت دادن نزدیکترین مرکز خوشه به آن ویژگی، مرکز خوشه‌ها به‌روز رسانی می‌شوند. این روند آنقدر تکرار می‌شود تا در نهایت مرکز خوشه‌ها ثابت شوند و دیگر تغییری نداشته باشند [۳۸]. شکل ۱-۵ در سمت چپ یک تصویر ام‌آرآی را نشان می‌دهد که در سمت راست با روش K-means قطعه‌بندی شده است. نکته‌ی قابل توجه در این قطعه‌بندی، طبقه‌بندی به صورت سخت و قاطع است که هر پیکسل تنها متعلق به یک کلاس خاص است.

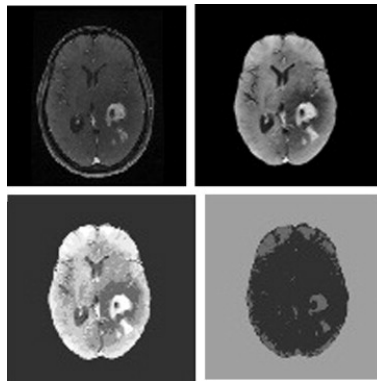


شکل ۱-۵: قطعه‌بندی تصویر ام‌آرآی مغزی با روش K-means [۵]

<sup>26</sup>Decision Tree

## Fuzzy C-means

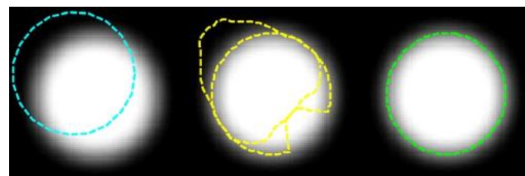
این تکنیک بر اساس تیوری فازی بنا شده است که برعکس K-means که به صورت قطعه‌بندی سخت عمل می‌کند، در اینجا شاهد یک قطعه‌بندی نرم هستیم که امکان محاسبه‌ی احتمال تعلق به هر کلاس را بوجود می‌آورد [۳۹]. شکل ۱-۶ در سمت چپ-بالا یک تصویر ام‌آرآی مغزی را نشان می‌دهد که توسط روش Fuzzy C-means نقشه‌ی احتمال تعلق هر پیکسل به هر کلاس نشان داده شده است.



شکل ۱-۶: قطعه‌بندی تصویر ام‌آرآی مغزی با روش Fuzzy C-means [۶]

## ۵-۳-۱ مدل‌های شکل‌پذیر

یک مدل شکل‌پذیر می‌تواند یک منحنی و یا یک رویه باشد که امکان تغییر آن در طول زمان وجود دارد. تغییرات این نوع مدل‌ها بر اساس تعریف تابع هزینه‌ی خود مدل و اطلاعات درون تصویر صورت می‌گیرد. این نیروهای داخلی و خارجی در نهایت خم شکل‌پذیر را به سمت لبه‌های تصویر و یا یک ویژگی مطلوب در تصویر سوق می‌دهند. این نوع مدل‌ها در دسته‌ی نیمه-خودکار قرار می‌گیرند و لازم است، یک خم اولیه برای شروع الگوریتم بر روی تصویر تعریف کند. شکل ۱-۷ تکامل یک خم اولیه طی چندین تکرار برای قطعه‌بندی را نشان می‌دهد.



شکل ۱-۷: قطعه‌بندی با استفاده از مدل‌های شکل‌پذیر [۷]

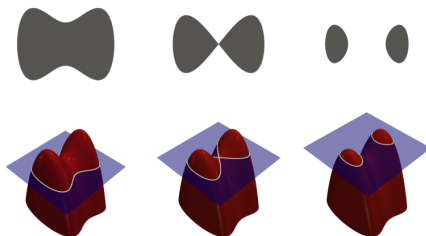
مدل‌های شکل‌پذیر را می‌توان به دو دسته‌ی عمده‌ی مدل‌های شکل‌پذیر پارامتری (صریح) <sup>۲۷</sup> و مدل‌های شکل‌پذیر هندسی (ضمنی) <sup>۲۸</sup> تقسیم نمود.

### مدل‌های شکل‌پذیر پارامتری

به این نوع مدل‌ها، کانتورهای فعال یا مدل‌های snake نیز گفته می‌شود. ایده‌ی اصلی در طراحی این مدل‌ها یافتن مرزها با تطابق مدل شکل‌پذیر و منحنی‌های درون تصویر با استفاده از نیروهای داخلی و خارجی است که با بهینه کردن یک تابع هزینه بدست می‌آیند. نیروهای داخلی شامل قابلیت ارتجاع <sup>۲۹</sup> و سختی <sup>۳۰</sup> مدل است و نیروهای خارجی باعث کشش کانتور به سمت لبه‌های تصویر می‌شود [۴۰].

### مدل‌های شکل‌پذیر هندسی

مدل‌های شکل‌پذیر هندسی و یا مجموعه-تراز <sup>۳۱</sup> از تئوری تکامل منحنی و مجموعه-تراز به‌جای پارامتری کردن منحنی استفاده می‌کنند. عملکرد این مدل‌ها بر اساس یافتن مجموعه-تراز صفر در یک رویه با بعد بالاتر است که قطعه‌بندی نهایی را توصیف می‌کند [۴۱]. شکل ۸-۱ روند تکامل منحنی و استخراج مجموعه-تراز صفر برای قطعه‌بندی یک ساختار دوبعدی را نشان می‌دهد.



شکل ۸-۱: روش تکامل منحنی و استخراج مجموعه-تراز صفر برای قطعه‌بندی [۸]

<sup>27</sup>Parametric (explicit) deformable model

<sup>28</sup>Geometric (implicit) deformable model

<sup>29</sup>Elasticity

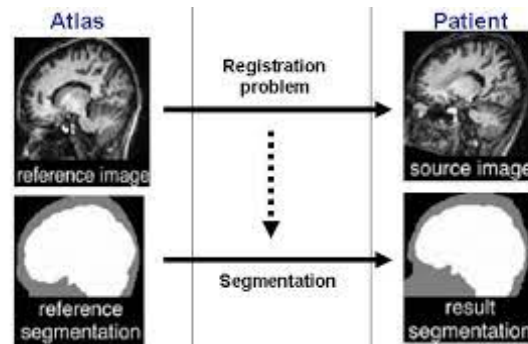
<sup>30</sup>Rigidity

<sup>31</sup>Level Set

### ۱-۳-۶ مدل‌های مبتنی بر اطلس

روش‌های مبتنی بر اطلس زمانیکه یک اطلس و یا قالب استاندارد در دسترس باشد یکی از قدرتمندترین ابزارها برای قطعه‌بندی تصاویر پزشکی هستند. اطلس با گردآوری اطلاعات مربوط به آناتومی که نیاز به قطعه‌بندی دارد، تولید می‌شود و سپس به عنوان مرجع برای قطعه‌بندی استفاده می‌گردد.

به اینگونه روش‌ها به عنوان یک مساله‌ی انطباق تصویر نگاه می‌شود که در ابتدا یک نگاشت یک به یک که تصویر اطلس از پیش قطعه‌بندی شده را بر روی تصویر هدف منطبق می‌کند، پیدا می‌شود و سپس قطعه‌بندی اطلس با همین نگاشت تبدیل داده می‌شود تا قطعه‌بندی هدف حاصل شود. این نوع روش‌ها برای قطعه‌بندی ساختارهایی بهتر نتیجه می‌دهند که تغییر شکل ندارند مانند استخوان‌ها [۴۲]. شکل ۱-۹ روند کلی قطعه‌بندی بر اساس اطلس را نشان می‌دهد.



شکل ۱-۹: روش مبتنی بر اطلس برای قطعه‌بندی تصویر مغزی [۹]

### ۱-۳-۷ سایر روش‌های قطعه‌بندی

علاوه بر آنچه اشاره گردید، روش‌های دیگری نیز برای قطعه‌بندی مانند روش‌های مبتنی بر گراف، روش‌ها هیبرید (استفاده توأم از چند روش قطعه‌بندی)، روش‌های مبنی بر معادلات دیفرانسیل جزئی و ... نیز ارائه شده است [۴۳] که به دلیل کمتر استفاده شدن، از توضیح آن‌ها خودداری می‌شود.

### ۱-۴ قطعه‌بندی با استفاده از شبکه‌های عصبی عمیق

از یک الگوریتم قطعه‌بندی تصاویر پزشکی مناسب انتظار می‌رود موارد زیر را ارضا کند [۲۲].



- دقت بالا
- مقاومت در برابر نویز
- قابلیت اطمینان بالا
- کمترین تعامل با کاربر
- قابلیت تکرار
- سرعت بالا

تمامی الگوریتم‌های کلاسیکی که پیش‌تر بررسی شد، سعی در برآورده کردن موارد بالا داشتند که تا حد قابل قبولی پیش رفته‌اند و توانسته‌اند اعتماد پزشکان و متخصصان را به عنوان یک الگوریتم تشخیص کمکی، استفاده کنند. با ظهور روش‌های یادگیری عمیق و شکست دادن انسان [۴۴] در بسیاری از کارها مانند طبقه‌بندی تصاویر طبیعی [۴۵]، بازی‌های فکری و ... توجه بسیاری برای استفاده از این روش‌ها در طبقه‌بندی تصاویر، جلب شد. با استفاده از متدهای یادگیری عمیق در طبقه‌بندی تصاویر پزشکی، دقت و قابلیت اعتماد بالاتری نسبت به روش‌های کلاسیک بدست آمد و امروزه تقریباً در اکثر نرم‌افزارهای طبقه‌بندی، این روش‌ها جایگزین روش‌های کلاسیک شده‌اند. در ادامه به بررسی مقالات و روش‌ها ارایه شده در این حوزه پرداخته می‌شود.

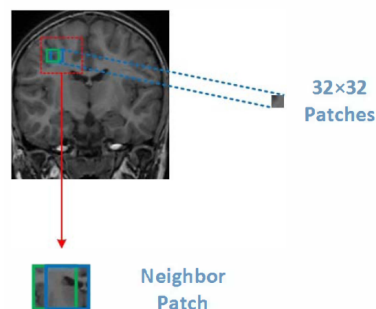
در فصل دوم جزئیات و ابزارهای مربوط به روش‌های یادگیری ماشین و یادگیری عمیق بررسی گردید. در این فصل استفاده‌ی این ابزارها در طبقه‌بندی تصاویر پزشکی و ایده‌هایی که باعث بهبود این روش‌ها شده‌اند بررسی می‌شوند.

## ۱-۴-۱ طبقه‌بندی پیکسل‌ها برای طبقه‌بندی

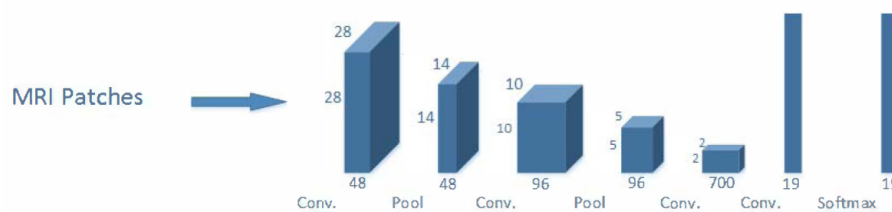
با ظهور و معرفی شبکه‌های عصبی عمیق در سال ۲۰۱۲ برای طبقه‌بندی تصاویر طبیعی [۴۶] و یک جهش بزرگ در دقت طبقه‌بندی، استفاده از این روش‌ها در طبقه‌بندی رونق گرفت. اولین و ساده‌ترین ایده برای ارتباط حوزه‌ی طبقه‌بندی و طبقه‌بندی، همانطور که پیش‌تر اشاره شد، استفاده از شبکه‌های عمیق برای طبقه‌بندی در سطح پیکسل‌هاست. یعنی با استخراج خودکار ویژگی از نواحی اطراف پیکسل با شبکه‌های عصبی عمیق و طبقه‌بندی پیکسل مد نظر به طبقه‌ی مطلوب است.

در [۱۰] برای طبقه‌بندی تصاویر ام‌آرآی از شبکه‌های کانوولوشنی عمیق استفاده شده است که شامل چند لایه پشت سرهم لایه‌های کانوولوشنی و ادغام بیشینه‌گیر است که ویژگی‌ها تصویر را استخراج می‌کند. رویکرد این مطالعه استفاده از روش‌های طبقه‌بندی برای طبقه‌بندی است به این صورت که یک پنجره

با ابعاد  $32 \times 32$  بر روی تصویر لغزانده می‌شود و یک پچ<sup>۳۲</sup> از تصویر استخراج می‌شود، با استفاده از این پچ پیکسل وسط برچسب‌گذاری می‌شود که برای این کار ابتدا توسط شبکه‌ی کانولوشنی ویژگی‌های پچ استخراج می‌شود و سپس این ویژگی‌ها به یک شبکه‌ی تمام متصل برای طبقه‌بندی هر پیکسل داده می‌شود. شکل ۱-۱۰ چگونگی استخراج پچ‌ها از تصویر ام‌آرآی مغزی را نشان می‌دهد و سپس در شکل ۱-۱۱ این پچ‌ها به یک شبکه‌ی کانولوشنی برای استخراج ویژگی و در نهایت طبقه‌بندی توسط یک شبکه‌ی تمام متصل برای قطعه‌بندی ساختارهای مغزی استفاده می‌شود.



شکل ۱-۱۰: لغزاندن یک پنجره روی تصویر ام‌آرآی مغزی و استخراج پچ [۱۰]

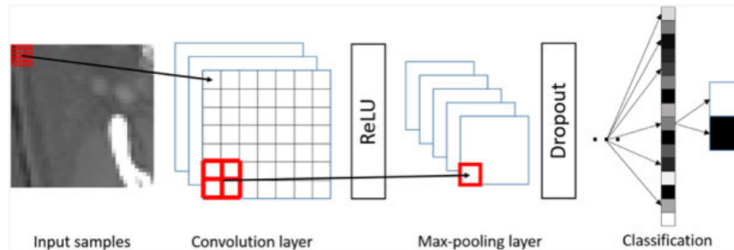


شکل ۱-۱۱: شبکه‌ی کانولوشنی برای طبقه‌بندی پیکسل‌ها وسطی پچ‌ها [۱۰]

ابراگیموف و زینک [۱۱]، ۱۳ ساختار در ریسک را در ۵۰ تصویر سی‌تی اسکن سر و گردن قطعه‌بندی کرده‌اند. رویکرد آن‌ها استفاده از شبکه‌های کانولوشنی عمیق است که شدت‌های یکسان در بافت‌های همسان را استخراج کند و بتواند در تصاویر دیده نشده، این ساختارهای در ریسک را قطعه‌بندی کند. برای آموزش شبکه، پچ‌های مثبت که شامل شدت‌های مثبت و اطراف واکسل ساختار در ریسک هستند و همچنین پچ‌های منفی که شامل واکسل‌های ساختارهای کناری هستند استخراج می‌شود. سپس این پچ‌ها به شبکه‌ی کانولوشنی داده می‌شوند که با ترکیب ویژگی‌های سطح پایین مانند لبه‌ها و نقاط کلیدی با ویژگی‌های سطح بالاتر، به طور موثری ساختارهای در ریسک را قطعه‌بندی نموده است. در شکل زیر

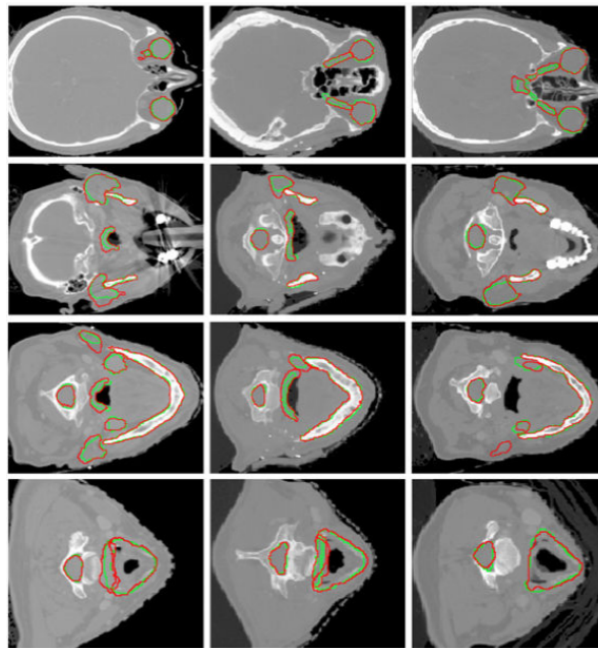
<sup>32</sup>Patch

معماری شبکه‌ی استفاده شده که شامل لایه‌های کانولوشنی، ادغام بیشینه‌گیر و dropout است، نشان داده شده است. در نهایت ویژگی‌های استخراج شده به یک شبکه‌ی تمام متصل برای طبقه‌بندی داده می‌شود.



شکل ۱-۱۲: معماری شبکه‌ی کانولوشنی برای طبقه‌بندی بر اساس پیچ ساختارهای در ریسک سر و گردن [۱۱]

شکل ۱-۱۳: خروجی شبکه (رنگ قرمز) و قطعه‌بندی مطلوب (رنگ سبز) را برای ساختارهای در ریسک سر و گردن مانند چشم‌ها، فک و ... نشان می‌دهد.



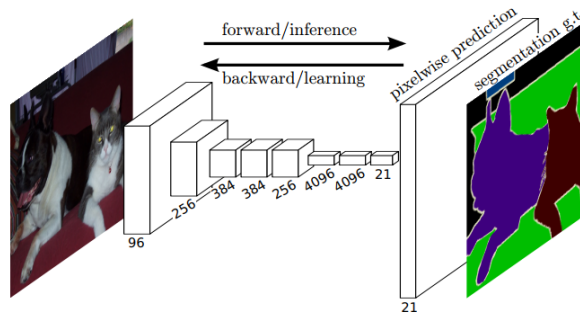
شکل ۱-۱۳: نتایج بدست آمده در قطعه‌بندی ساختارهای در ریسک سر و گردن (قرمز: خروجی شبکه، سبز: قطعه‌بندی مطلوب) [۱۱]

روش‌های قطعه‌بندی بر اساس پیچ، علی‌رغم استفاده‌ی زیاد در مواردی که ابعاد تصاویر برای آموزش

بسیار بزرگ باشد و نیازمند حافظه‌ی زیاد باشد و نیز تعداد دادگان برای آموزش کم باشد (با این روش تعداد دادگان چندین برابر می‌شود) استفاده دارد [۴۷]. اما معایبی از جمله، از بین رفتن اطلاعات گسترده و مربوط به بافت<sup>۳۳</sup> تصویر، ایجاد قطعه‌هایی در جاهایی که مرتبط با بافت هدف برای قطعه‌بندی نیست و همچنین تصمیم‌گیری تنها بر اساس پیکسل‌های کناری است که از استخراج ویژگی‌های سطح بالاتر پیشگیری می‌شود. یکی دیگر از معایب آن به علت وجود شبکه‌ی تمام متصل، داشتن ورودی با ابعاد خاص است زیرا با تغییر ورودی ابعاد ویژگی استخراج شده تغییر می‌کند و در نتیجه باید ابعاد شبکه‌ی تمام متصل تغییر کند (مانند شبکه‌های طبقه‌بند). بنابراین محققان سعی در ارائه‌ی مدل‌هایی برای حل این مشکلات داشتند که در ادامه به بررسی این تکنیک‌ها پرداخته می‌شود [۴۸].

### ۱-۴-۲ شبکه‌های تمام کانولوشنی برای قطعه‌بندی

ایده‌ی اصلی برای ساخت شبکه‌های تمام کانولوشنی، همانطور که در بالا اشاره شد، قطعه‌بندی بر اساس اطلاعات کل تصویر است. بنابراین با دادن یک تصویر به عنوان ورودی، با همان ابعاد، نقشه‌ی قطعه‌بندی خروجی ایجاد می‌شود [۱۲]. شکل ۱-۱۴ معماری اینگونه شبکه‌ها را نشان می‌دهد که در ابتدا ویژگی‌هایی از کل تصویر توسط لایه‌های کانولوشنی استخراج می‌شود و سپس این ویژگی‌ها با افزایش اندازه و اعمال کانولوشن بر روی ویژگی‌های تغییر اندازه یافته، نقشه‌ی قطعه‌بندی نهایی ایجاد می‌شود.



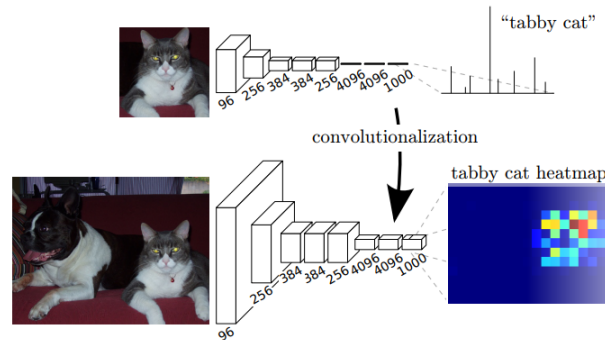
شکل ۱-۱۴: قطعه‌بندی تصاویر با استفاده از رویکرد شبکه‌های تمام کانولوشنی [۱۲]

لانگ و همکاران [۱۲] دریافتند که امکان استخراج نقشه‌ی حرارتی<sup>۳۴</sup> وجود یک شی در یک تصویر را با استفاده از بخش استخراج ویژگی شبکه‌های طبقه‌بند، انجام داد. در شکل ۱-۱۵ وجود یک گربه

<sup>۳۳</sup>Texture

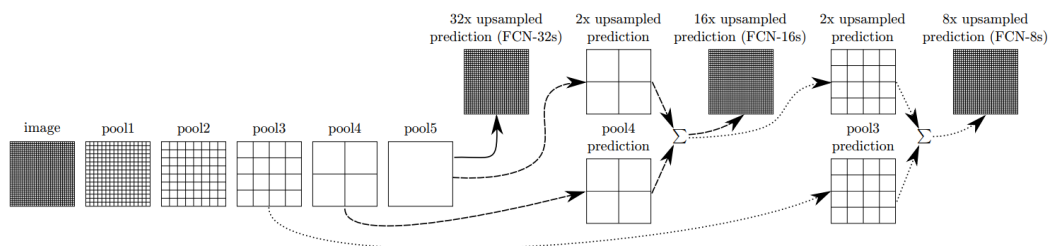
<sup>۳۴</sup>Heatmap

در تصویر توسط شبکه‌های طبقه‌بند پیش‌بینی شده است و با استخراج نقشه‌ی گرمایی از ویژگی‌های استخراج شده، مکان تقریبی آن قابل دستیابی است. بنابراین با این رویکرد و استفاده از معماری شکل بالا، این گروه توانستند قطعه‌بندی را با استفاده از شبکه‌های تمام کانولوشنی صورت دهند.



شکل ۱-۱۵: استخراج نقشه‌ی گرمایی وجود یک شی در تصویر با استفاده از ویژگی‌های استخراج شده توسط شبکه‌ی طبقه‌بند [۱۲]

مشکل موجود در چنین معماری‌هایی بازگرداندن ابعاد ویژگی‌های استخراج شده به اندازه‌ی تصویر اصلی است که برای آن روش‌های زیادی پیشنهاد شده است که در ادامه بررسی خواهند شد. در مقاله‌ی لانگ و همکاران برای حل این مشکل از upsample کردن ویژگی‌های استخراج شده بعد از لایه‌های ادغام و ترکیب ویژگی‌های با ابعاد مختلف با یکدیگر برای رسیدن به ابعاد تصویر اولیه است. روند این کار در شکل ۱-۱۶ نشان داده شده است.

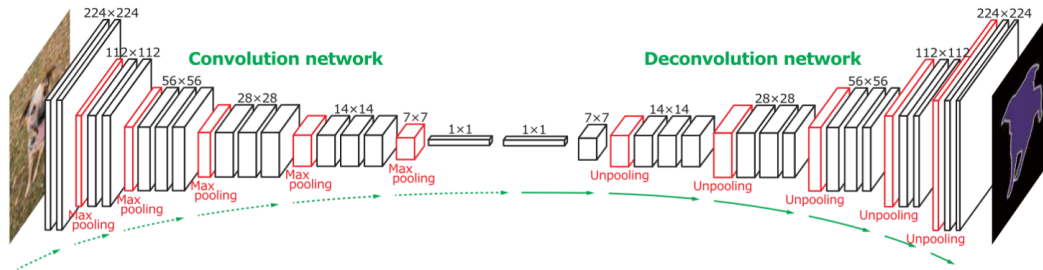


شکل ۱-۱۶: تغییر ابعاد ویژگی‌های استخراج شده بعد از لایه‌های ادغام به ابعاد تصویر اصلی [۱۲]

در مرجع [۱۳] برای حل مشکل upsampling به جای روش‌های معمول درونیابی، فیلترهایی با وزن‌های قابل یادگیری ارائه دادند که باعث بهبود دقت قطعه‌بندی گردید. شبکه‌ی پیشنهادی این گروه طبق شکل ۱-۱۷ از دو بخش کانولوشن و دیکانولوشن<sup>۳۵</sup> تشکیل شده است که در بخش کانولوشن

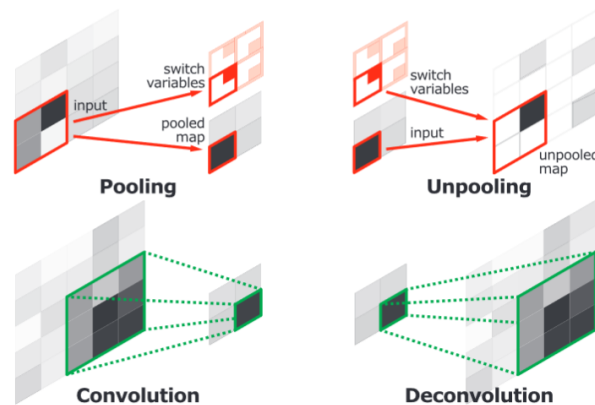
<sup>۳۵</sup>Deconvolution

ویژگی‌های مطلوب توسط لایه‌های کانولوشن و ادغام استخراج می‌شود و سپس بخش دیکانولوشن این ویژگی‌ها را به ابعاد تصویر اصلی تغییر می‌دهد تا نقشه‌ی قطعه‌بندی حاصل شود.



شکل ۱-۱۷: معماری شبکه‌ی شامل کانولوشن و دیکانولوشن برای بهبود دقت قطعه‌بندی [۱۳]

برای تفهیم ایده‌ی لایه‌های دیکانولوشن ارائه شده در این شبکه، شکل ۱-۱۸ را در نظر بگیرید. در سمت بالا-چپ، عملکرد لایه‌ی ادغام بیشینه‌گیر نشان داده شده‌است که مقدار بزرگترین درایه و موقعیت آن را ذخیره می‌کند. در سمت بالا-راست، لایه‌ی برعکس لایه‌ی ادغام نشان داده شده‌است که با دانستن موقعیت بزرگترین درایه در هنگام upsampling آن را در موقعیت پیشین خود قرار می‌دهد. در سمت پایین-چپ، لایه‌ی کانولوشن نشان داده شده‌است که در فصل دوم عملکرد آن بررسی گردید و در نهایت در سمت پایین-راست، عملکرد لایه‌ی دیکانولوشن توضیح داده شده‌است که یک فیلتر با وزن‌های قابل یادگیری در یک مقدار ضرب می‌شود و به ابعاد بزرگتری تغییر اندازه می‌دهد، با لغزاندن این فیلتر و جمع کردن مقادیر همپوشان با یکدیگر، ابعاد ماتریس ویژگی بزرگتر می‌شود.

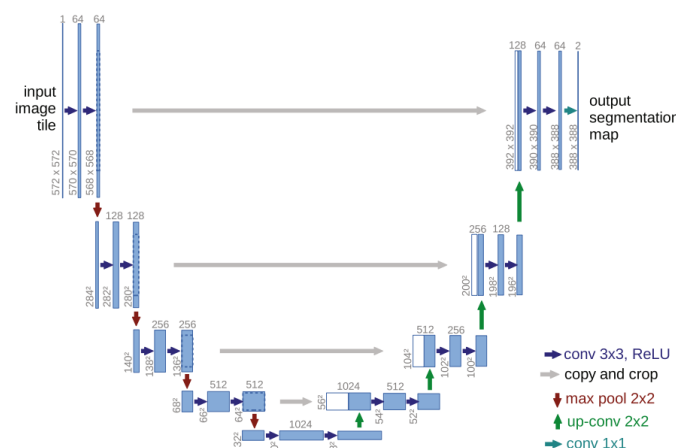


شکل ۱-۱۸: نمایش تفاوت عملکرد کانولوشن و دیکانولوشن [۱۳]

رونبرگر و همکاران [۱۴] برای قطعه‌بندی تصاویر پزشکی یک معماری جدید و پر استفاده به نام UNet ارائه دادند. شکل ۱-۱۹ این معماری را نشان می‌دهد که از دو مسیر انقباضی و انبساطی

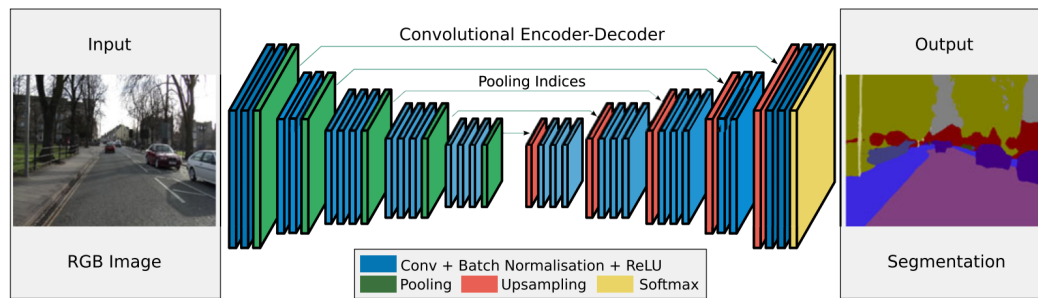
تشکیل شده است. در مسیر انقباضی روند معمول در شبکه‌های کانولوشنی پیاده‌سازی شده است که شامل، لایه‌های کانولوشنی  $3 \times 3$  که بعد از آن‌ها از تابع فعالیت ReLU استفاده شده است و در نهایت یک لایه‌ی ادغام بیشینه‌گیر  $2 \times 2$  برای کاهش ابعاد ماتریس ویژگی به نصف ابعاد اولیه قرار داده شده است. در هر قدم از کاهش ابعاد تعداد کانال‌های ویژگی‌ها دو برابر شده است.

در مسیر انبساطی، بعد از هر بار upsampling یک کانولوشن  $2 \times 2$  قرار داده شده است که در تصویر با نام 'Up-convolution' نشان داده شده و در طی این عملیات تعداد کانال‌های ویژگی نصف می‌شود. نکته‌ی مهم در این معماری قرار دادن مسیرهای اتصالی از مسیر انقباضی به مسیر انبساطی است. این کار باعث حل شدن مشکل محو شدن گرادیان طی به‌روز رسانی وزن‌ها می‌شود و همچنین اطلاعات و جزئیاتی که به دلیل استفاده از لایه‌های ادغام از بین می‌روند دوباره در این مسیر احیا می‌شود و باعث بهبود دقت قطعه‌بندی می‌گردد. در نهایت بر روی کانال‌های ویژگی بدست آمده یک کانولوشن  $1 \times 1$  برای تغییر عمق کانال‌ها زده می‌شود و نقشه‌های قطعه بندی مطلوب با کمینه کردن تابع هزینه ایجاد می‌شوند.



شکل ۱-۱۹: معماری شبکه‌ی UNet [۱۴]

برای بهبود عملکرد upsampling در معماری UNet همانند آنچه در شکل ۱-۱۸ بالا-راست، مشاهده شد، معماری جدیدی به نام SegNet معرفی گردید [۱۵] که مختصات بزرگترین درایه‌ها را در لایه‌های ادغام مسیر انقباضی ذخیره می‌کند و سپس برای upsample کردن، مقادیر را دقیقاً در همان درایه‌های متناظر قرار می‌دهد و برعکس روش‌های دیگر، مابقی درایه‌ها برابر با صفر قرار داده می‌شوند. شکل ۱-۲۰ معماری این شبکه را نشان می‌دهد که علاوه بر مسیرهای اتصالی، درایه‌های لایه‌های ادغام نیز عبور داده می‌شود.



شکل ۱-۲۰: معماری شبکه‌ی SegNet [۱۵]

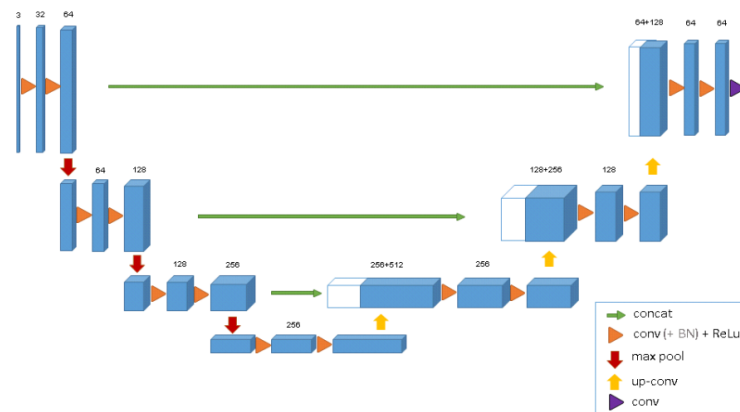
### ۱-۴-۳ شبکه‌های تمام کانولوشنی سه‌بعدی برای قطعه‌بندی

با گسترش استفاده‌ی شبکه‌های کانولوشنی دوبعدی برای پردازش تصاویر دوبعدی و حاصل شدن نتایج خیره‌کننده، استفاده‌ی این نوع شبکه‌ها برای پردازش ویدیو نیز مورد توجه قرار گرفت. در دنباله‌های ویدیویی، علاوه بر ابعاد طول و عرض تصویر بعد زمان نیز اضافه می‌شود و یک رشته‌ی سه‌بعدی برای پردازش پدید می‌آید. بنابراین با توسعه‌ی شبکه‌های کانولوشنی دوبعدی به سه‌بعدی امکان پردازش ویدیو نیز مهیا شد [۴۹، ۵۰].

با این توسعه، پردازش تصاویر سه‌بعدی پزشکی مانند ام‌آر‌آی و سی‌تی اسکن، نیز مورد توجه قرار گرفت که بحث قطعه‌بندی این نوع تصاویر را نیز شامل می‌شود. بنابراین با تعمیم شبکه‌های دوبعدی به سه‌بعدی تمام کانولوشنی، می‌توان مسأله‌ی قطعه‌بندی را با تصویر سه‌بعدی انجام داد. در ادامه به بررسی شبکه‌های سه‌بعدی ارایه شده برای قطعه‌بندی تصاویر سه‌بعدی پرداخته خواهد شد.

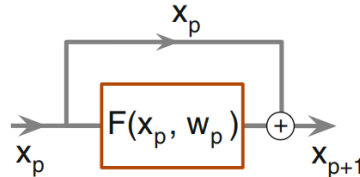
شبکه‌ی UNet سه‌بعدی یا 3D UNet مشابه معماری UNet دوبعدی از مسیرهای انقباضی و انبساطی تشکیل شده‌است با این تفاوت که تمامی محاسبات کانولوشن و ادغام در فضای سه‌بعدی انجام می‌شود [۱۶]. معماری این شبکه در شکل ۱-۲۱ نشان داده شده‌است.





شکل ۱-۲۱: معماری شبکه‌ی UNet سه‌بعدی [۱۶]

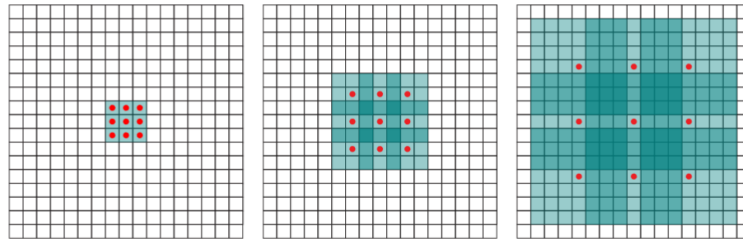
بلوک‌های residual توسط هی و همکاران [۵۱] برای آموزش موثر شبکه، ارایه شدند. ایده‌ی اصلی این بلوک‌ها ایجاد یک تطابق یک به یک برای میانبر کردن لایه‌های شامل پارامتر و تجمیع ورودی و خروجی در انتهای این بلوک است. این بلوک‌ها انتشار اطلاعات را هموارتر می‌کنند و باعث آموزش بهتر شبکه می‌شوند. در شکل ۱-۲۴ یک بلوک دارای اتصالات residual نشان داده شده‌است.



شکل ۱-۲۲: یک بلوک با اتصال residual [۱۷]

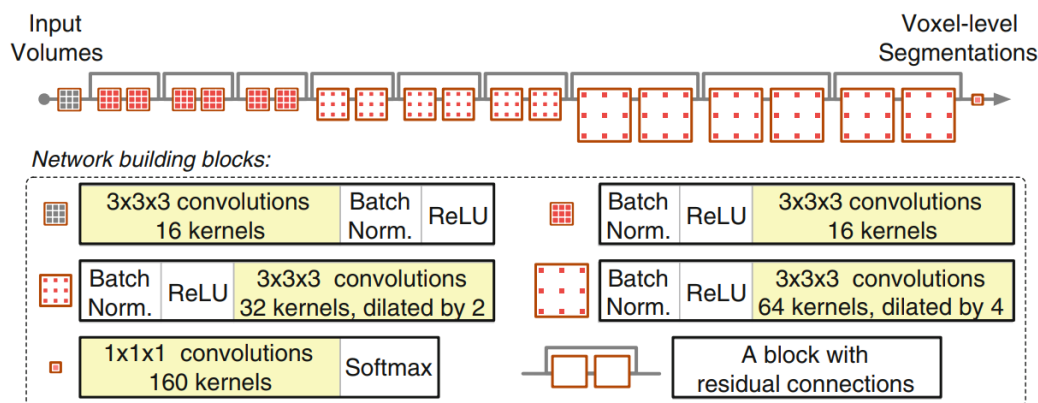
کانوولوشن‌های منبسط شده<sup>۳۶</sup> همانند کانوولوشن‌های عادی عمل می‌کنند با این تفاوت که بین المان‌های سازنده‌ی آن‌ها بر حسب ضریب انبساط صفر قرار می‌گیرد. به این ترتیب بدون افزایش تعداد پارامترهای فیلتر، میدان دید بزرگتر و استخراج ویژگی از ناحیه‌ی بزرگتری صورت می‌پذیرد که باعث استخراج ویژگی‌های عمیق‌تری می‌شود. شکل ۱-۲۳، میدان دید کانوولوشن‌های منبسط شده با ضریب انبساط به ترتیب از چپ به راست ۱، ۲ و ۴ را نشان می‌دهد.

<sup>36</sup>Dilated Convolution



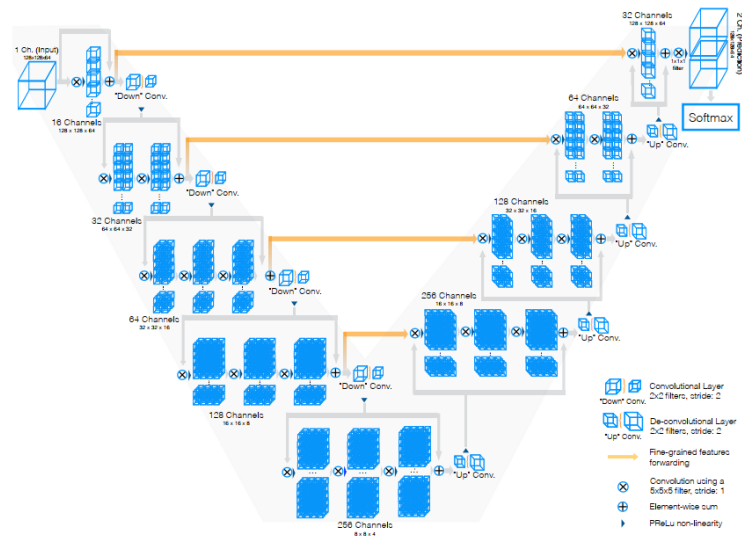
شکل ۱-۲۳: کانولوشن‌های منبسط شده [۱۸]

با استفاده از دو ایده‌ی معرفی شده در بالا (بلوک‌های residual و کانولوشن‌های منبسط شده) لی و همکاران یک شبکه‌ی سه‌بعدی برای قطعه‌بندی تصاویر ام‌آرآی مغزی ارایه کرده‌اند. این شبکه شامل ۲۰ لایه‌ی کانولوشنی است. در ۷ لایه‌ی اول، از کانولوشن‌های  $3 \times 3 \times 3$  استفاده شده‌است که وظیفه‌ی آن‌ها استخراج ویژگی‌های سطح پایین مانند لبه‌ها و گوشه‌هاست. در لایه‌های بعدی ضریب انبساط برای کانولوشن‌های منبسط شده، ۲ و ۴ است. این لایه‌ها وظیفه‌ی استخراج ویژگی‌های سطح بالاتر را دارند. بلوک‌های residual به هر جفت لایه‌ی کانولوشنی برای بهبود عملکرد اضافه گردیده‌است. همچنین از روش Batch-normalization برای کمک به فرآیند تعمیم‌پذیری استفاده شده‌است. شکل زیر معماری این شبکه را نشان می‌دهد.



شکل ۱-۲۴: معماری شبکه‌ی سه‌بعدی شامل اتصالات residual و کانولوشن انبساطی [۱۷]

شبکه‌ی VNet نیز مانند UNet سه‌بعدی از کانولوشن و محاسبات سه‌بعدی استفاده می‌کند. تفاوت این دو معماری مطابق شکل ۱-۲۵، اضافه شدن بلوک‌های residual به معماری UNet است. این کار باعث جلوگیری از محو شدن گرادیان در لایه‌ها نخستین و همچنین عمیق‌تر کردن شبکه برای استخراج ویژگی‌های سطح بالا می‌شود [۱۹].



شکل ۱-۲۵: معماری شبکه‌ی VNet [۱۹]

علی‌رغم وجود مزیت‌های فراوان استفاده از شبکه‌های سه‌بعدی در استخراج ویژگی‌های فضایی بهتر، این شبکه‌ها با مشکلاتی از قبیل، هزینه‌ی محاسباتی بالا و حافظه‌ی مورد نیاز بالا برای استفاده و آموزش روبرو هستند [۵۲].

## ۱-۵ قطعه‌بندی تومور و ساختارهای در ریسک با یادگیری عمیق

## ۱-۶ جمع‌بندی

## مراجع

- [1] U. Vovk, F. Pernus, and B. Likar. A review of methods for correction of intensity inhomogeneity in mri. *IEEE transactions on medical imaging*, 26(3):405–421, 2007.
- [2] Metrics to evaluate your semantic segmentation model, <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>. Accessed: 2022-04-12.
- [3] Understand jaccard index, jaccard similarity in minutes, <https://medium.com/data-science-bootcamp/understand-jaccard-index-jaccard-similarity-in-minutes-25a703fbf9d7>. Accessed: 2022-04-12.
- [4] Hausdorff distance, <https://en.wikipedia.org/wiki/hausdorffdistance>. Accessed: 2022-04-12.
- [5] H. Ng, S. Ong, K. Foong, P.-S. Goh, and W. Nowinski. Medical image segmentation using k-means clustering and improved watershed algorithm. In *2006 IEEE southwest symposium on image analysis and interpretation*, pages 61–65. IEEE, 2006.
- [6] S. Madhukumar and N. Santhiyakumari. Evaluation of k-means and fuzzy c-means segmentation on mr images of brain. *The Egyptian Journal of Radiology and Nuclear Medicine*, 46(2):475–479, 2015.
- [7] slides of medical image analysis and processing (miap) course, segmentation chapter, tutor: Emad fatemizadeh, sharif university of technology,. Accessed: 2020.
- [8] Level set, <https://en.wikipedia.org/wiki/level-setmethod>. Accessed: 2022-04-18.

- [9] S. Bourouis, R. Alroobaea, S. Rubaiee, and A. Ahmed. Toward effective medical image analysis using hybrid approaches—review, challenges and applications. *Information*, 11(3):155, 2020.
- [10] Z. Cui, J. Yang, and Y. Qiao. Brain mri segmentation with patch-based cnn approach. In *2016 35th Chinese Control Conference (CCC)*, pages 7026–7031. IEEE, 2016.
- [11] B. Ibragimov and L. Xing. Segmentation of organs-at-risks in head and neck ct images using convolutional neural networks. *Medical physics*, 44(2):547–557, 2017.
- [12] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
- [13] H. Noh, S. Hong, and B. Han. Learning deconvolution network for semantic segmentation. In *Proceedings of the IEEE international conference on computer vision*, pages 1520–1528, 2015.
- [14] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [15] 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*, 39(12):2481–2495, 2017.
- [16] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.
- [17] W. Li, G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso, and T. Vercauteren. On the compactness, efficiency, and representation of 3d convolutional networks: brain parcellation as a pretext task. In *International conference on information processing in medical imaging*, pages 348–360. Springer, 2017.
- [18] F. Yu and V. Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv:1511.07122*, 2015.

- [19] F. Milletari, N. Navab, and S.-A. Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. IEEE, 2016.
- [20] D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(1):315–337, 2000.
- [21] T. Zuva, O. O. Olugbara, S. O. Ojo, and S. M. Ngwira. Image segmentation, available techniques, developments and open issues. *Canadian Journal on Image Processing and Computer Vision*, 2(3):20–29, 2011.
- [22] N. Sharma and L. M. Aggarwal. Automated medical image segmentation techniques. *Journal of medical physics/Association of Medical Physicists of India*, 35(1):3, 2010.
- [23] V. Shrimali, R. Anand, and V. Kumar. Current trends in segmentation of medical ultrasound b-mode images: a review. *IETE technical review*, 26(1):8–17, 2009.
- [24] S. S. Al-Amri, N. V. Kalyankar, and S. D. Khamitkar. A comparative study of removal noise from remote sensing image. *arXiv preprint arXiv:1002.1148*, 2010.
- [25] J. Tohka. Partial volume effect modeling for segmentation and tissue classification of brain magnetic resonance images: A review. *World journal of radiology*, 6(11):855, 2014.
- [26] N. Makris, J. Kaiser, C. Haselgrove, L. J. Seidman, J. Biederman, D. Boriel, E. M. Valera, G. M. Papadimitriou, B. Fischl, V. S. Caviness Jr, et al. Human cerebral cortex: a system for the integration of volume-and surface-based representations. *Neuroimage*, 33(1):139–153, 2006.
- [27] P. Sinha, M. Tuteja, and S. Saxena. Medical image segmentation: hard and soft computing approaches. *SN Applied Sciences*, 2(2):1–8, 2020.
- [28] E. Debreuve, M. Gastaud, M. Barlaud, and G. Aubert. Using the shape gradient for active contour segmentation: from the continuous to the discrete formulation. *Journal of Mathematical Imaging and Vision*, 28(1):47–66, 2007.

- [29] N. Ironside, C.-J. Chen, S. Mutasa, J. L. Sim, D. Ding, S. Marfatiah, D. Roh, S. Mukherjee, K. C. Johnston, A. M. Southerland, et al. Fully automated segmentation algorithm for perihematoma edema volumetry after spontaneous intracerebral hemorrhage. *Stroke*, 51(3):815–823, 2020.
- [30] B. Alfano, M. Comerci, M. Larobina, A. Prinster, J. P. Hornak, S. E. Selvan, U. Amato, M. Quarantelli, G. Tedeschi, A. Brunetti, et al. An mri digital brain phantom for validation of segmentation methods. *Medical image analysis*, 15(3):329–339, 2011.
- [31] Z.-K. Huang and K.-W. Chau. A new image thresholding method based on gaussian mixture model. *Applied mathematics and computation*, 205(2):899–907, 2008.
- [32] N. Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.
- [33] R. Adams and L. Bischof. Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence*, 16(6):641–647, 1994.
- [34] J. C. Bezdek, L. Hall, L. Clarke, et al. Review of mr image segmentation techniques using pattern recognition. *MEDICAL PHYSICS-LANCASTER PA-*, 20:1033–1033, 1993.
- [35] L. Le Cam. Maximum likelihood: an introduction. *International Statistical Review/Revue Internationale de Statistique*, pages 153–171, 1990.
- [36] T. Song, M. M. Jamshidi, R. R. Lee, and M. Huang. A modified probabilistic neural network for partial volume segmentation in brain mr image. *IEEE Transactions on Neural Networks*, 18(5):1424–1432, 2007.
- [37] R. B. Bhatt, G. Sharma, A. Dhall, and S. Chaudhury. Efficient skin region segmentation using low complexity fuzzy decision tree model. In *2009 Annual IEEE India Conference*, pages 1–4. IEEE, 2009.
- [38] M.-N. Wu, C.-C. Lin, and C.-C. Chang. Brain tumor detection using color-based k-means clustering segmentation. In *Third international conference on intelligent information hiding and multimedia signal processing (IIH-MSP 2007)*, volume 2, pages 245–250. IEEE, 2007.

- [39] W. Wiharto and E. Suryani. The comparison of clustering algorithms k-means and fuzzy c-means for segmentation retinal blood vessels. *Acta Informatica Medica*, 28(1):42, 2020.
- [40] L. He, Z. Peng, B. Everding, X. Wang, C. Y. Han, K. L. Weiss, and W. G. Wee. A comparative study of deformable contour methods on medical image segmentation. *Image and vision computing*, 26(2):141–163, 2008.
- [41] Y. Jiang, M. Wang, and H. Xu. A survey for region-based level set image segmentation. In *2012 11th International Symposium on Distributed Computing and Applications to Business, Engineering & Science*, pages 413–416. IEEE, 2012.
- [42] H. Kalinic. Atlas-based image segmentation: A survey. *Croatian Scientific Bibliography*, pages 1–7, 2009.
- [43] S. Masood, M. Sharif, A. Masood, M. Yasmin, and M. Raza. A survey on medical image segmentation. *Current Medical Imaging*, 11(1):3–14, 2015.
- [44] Q. Yu, Y. Yang, F. Liu, Y.-Z. Song, T. Xiang, and T. M. Hospedales. Sketch-a-net: A deep neural network that beats humans. *International journal of computer vision*, 122(3):411–425, 2017.
- [45] W. Wang, Y. Yang, X. Wang, W. Wang, and J. Li. Development of convolutional neural network and its application in image classification: a survey. *Optical Engineering*, 58(4):040901, 2019.
- [46] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.
- [47] T. B. Sekou, M. Hidane, J. Olivier, and H. Cardot. From patch to image segmentation using fully convolutional networks—application to retinal images. *arXiv preprint arXiv:1904.03892*, 2019.
- [48] E. Shelhamer, J. Long, and T. Darrell. Fully convolutional networks for semantic segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(4):640–651, 2016.



- 
- [49] D. Maturana and S. Scherer. Voxnet: A 3d convolutional neural network for real-time object recognition. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 922–928. IEEE, 2015.
  - [50] Q. Chen, Z. Liu, Y. Zhang, K. Fu, Q. Zhao, and H. Du. Rgb-d salient object detection via 3d convolutional neural networks. *arXiv preprint arXiv:2101.10241*, 2021.
  - [51] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
  - [52] T. Lei, R. Wang, Y. Wan, X. Du, H. Meng, and A. K. Nandi. Medical image segmentation using deep learning: A survey. 2020.

## **Abstract**

We present a standard template for typesetting theses in Persian. The template is based on the X<sub>Y</sub>TeX Persian package for the L<sup>A</sup>T<sub>E</sub>X typesetting system. This write-up shows a sample usage of this template.

**Keywords:** Thesis, Typesetting, Template, X<sub>Y</sub>TeX Persian



Sharif University of Technology  
Department of Electrical Engineering

M.Sc. Thesis

# **Organs at Risk (OAR) segmentation using machine learning methods**

By:

**Reza Karimzadeh**

Supervisor:

**Dr. Emad Fatemizadeh**

February 2022