



MRI Segmentation Using Inception-based U-Net Architecture and Up Skip Connections

Farnaz Hoseini^{1*}, Hamed Sepehrzadeh², Amir Hosein Talimian³

^{1,2,3}Assistant Professor, Department of Computer Engineering, Technical and Vocational University (TVU), Tehran, Iran.

ARTICLE INFO

Article Type:

Original Research

Received: 04.27.2023

Revised: 06.22.2023

Accepted: 08.13.2023

Keyword:

Segmentation
U-Net
MRI
Deep Learning
Up Skin Connections
Inception Module

*Corresponding Author:

Farnaz Hoseini

Email: f-hoseini@tvu.ac.ir

ABSTRACT

Medical imaging is a non-invasive technique that has caused significant development in diagnosing and identifying human diseases. Among all medical imaging techniques, magnetic resonance imaging (MRI) is more popular. This method is not harmful to human health and can perform imaging of human brain details with high quality. Correct segmentation of brain tumours in MR images is very important. Traditional methods for segmenting medical images are time-consuming and require high expertise. Deep learning methods for brain tumour segmentation from MR images usually use normal convolution layers, in which they will not have the capability of distinguishing micro-scale and large-scale structures. In this research, a new method based on deep learning for brain tumour segmentation on MR images is presented. The proposed method is a generalization of the famous U-Net architecture, with the difference that the Inception module is used instead of normal convolution layers. Due to convolution kernels with different sizes in parallel, the Inception module can extract small-scale and large-scale features from the image. In the architecture of the proposed model, up-skip connections were used to improve the information flow in the forward propagation stage. In addition, a new pre-processing method based on the image mode was presented in this research, which normalizes the image intensity using the image mode. The proposed method was evaluated on the BraTS 2022 dataset and the accuracy results obtained for the Dice similarity coefficient with a value of 0.91 indicate the improvement of the detection accuracy. The evaluation results show that both hypotheses presented on the effect of high jump connections in improving the flow of information and learning are correct, and the use of the Inception module significantly improved the evaluation criteria of the model.



EXTENDED ABSTRACT

Introduction

MRI is used in different fields of studying and diagnosing brain structures and tissues. However, the analysis of medical images is a complex task, because they contain a large amount of information, and in some cases, destructive effects such as noise make it difficult to extract the desired information from these images. In this study, an automatic method for brain MR image segmentation based on U-Net architecture was proposed, which can segment brain MR images with high accuracy compared to other similar methods. Since the use of kernels with the same size in conventional convolutional neural networks (CNN) cannot consider the desired areas with different sizes. In this research, a new deep learning architecture based on U-Net architecture was presented, in which Inception blocks are used instead of simple convolutional layers with kernels of the same size. Since each Inception module includes parallel convolutional layers with different kernel sizes, different structures with different scales can be recognized. In short, the proposed model for brain MR image segmentation uses a U-net architecture based on Inception. Implementation and evaluation of the proposed method were carried out on BraTS 2020 database images. The innovation of the proposed method compared to the usual U-net architecture is the use of the Inception module and the addition of Up skip connections to improve the training process and increase the accuracy of the model. Furthermore, in this research, a new preprocessing method based on the stable intensity normalization method based on the image mode was proposed. The proposed preprocessing method normalizes the image intensity by subtracting the image mode (for example, the grey value of the highest column in the intensity histogram).

Methodology

According to the obtained sub-parameters, the structure of the final proposed U-Net model included 9 Inception blocks, each block containing 2 layers. Each layer also contains a number of Inception modules, according to the setting of these parameters, the complete structure of the proposed model was obtained according to Figure 1.

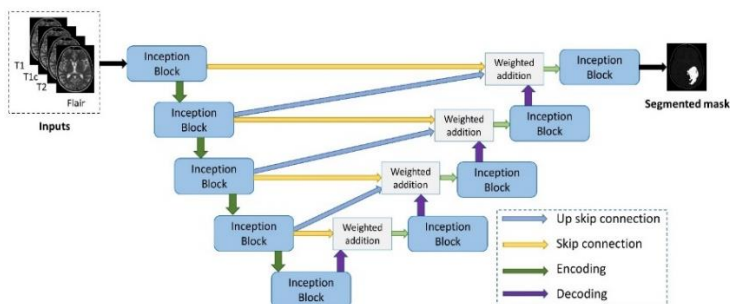


Figure 1 The architecture of the proposed U-Net model.

In the present research, to increase the capacity of the model to segment and optimize its performance, the Inception module was used in the U-Net architecture. The Inception

module demonstrated that it can improve the extraction of visual features with limited computing capacity. In an Inception module, kernels 3×3 , 5×5 , 7×7 were placed in parallel and operated on a common input. In the end, the features obtained by all three kernels were combined and added to the output of 1×1 kernel. The problem with the typical Inception architecture is that the number of learnable network weights increases significantly as the training process begins. Therefore, in this research, an alternative architecture was used for the Inception module. In the alternative architecture, each 5×5 kernel was replaced by two 3×3 kernels. On the other hand, since each 7×7 kernel is equivalent to three 3×3 kernels, three 3×3 kernels were used instead of one 7×7 kernel. Therefore, the alternative architecture is shown in Figure 2. In addition, batch normalization was used in order to increase the speed of convergence and solve the problems related to the training of deep models.

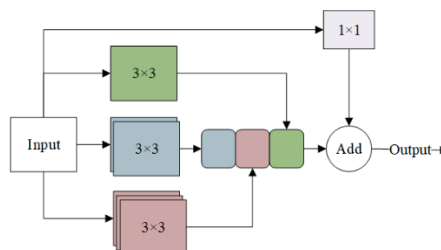
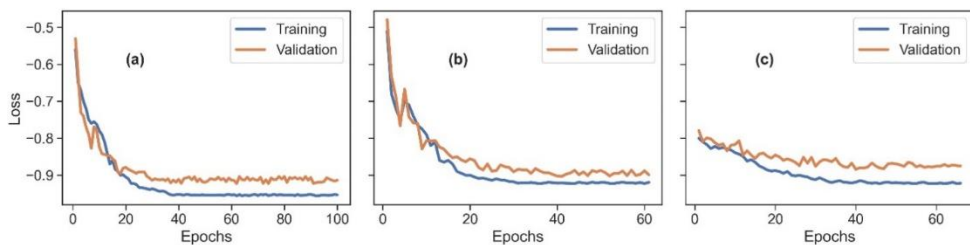


Figure 2. Alternative architecture for the Inception module in this research.

Results and discussion

The proposed optimized U-Net model with 9 Inception modules once with up skip connections (Inception-based U-Net + up skip connections), once with common skip connections (Inception-based U-Net + skip connections) and once without any connections Inception-based U-Net was trained with optimized sub-parameters. The performance comparison results of these optimizer models are shown in Figure 3. As can be observed, the proposed U-Net model with high hop connections (section c in Figure 3) achieved the lowest training and validation error. Although the proposed U-Net model with common up skip connections (section b in Figure 3) achieved a lower training error than the proposed U-Net model without up skip connections (section a in Figure 3), its validation error was higher.



(a) Inception-based U-Net + up skip connections, (b) Inception-based U-Net + skip connections, (c) Inception-based U-Net

Figure 3. Learning graph of proposed U-Net models.

In Workgroup networks, the number of systems (computers) is limited. This is due to limitations in Windows client versions. A view of the computer site space after hardware setup is shown in Figure 3 from different angles. To make the comparisons more understandable, the average values of the evaluation criteria were calculated on the experimental data for the four mentioned segmentation methods as reported in Table 2. This indicates the necessity of using the Inception module, because the Inception module, as mentioned previously, creates the ability to extract small-scale and large-scale structures by using convolutional layers with different kernel sizes and in parallel.

Table 1. Average results of evaluation criteria by different U-Net models proposed for segmentation of MR images from BraTS 2020 dataset.

Models	Dice	Sensitivity	PPV	Jaccard
U-Net	0.864	0.851	0.832	0.884
Inception-based U-Net	0.878	0.859	0.856	0.894
Inception-based U-Net + skip connections	0.893	0.874	0.887	0.904
Inception-based U-Net + up skip connections	0.913	0.894	0.916	0.924

Conclusion

In this research, a new method of brain MR image segmentation based on U-Net model with an Inception module and up skip connections (Inception-based U-Net + up skip connections) was proposed. Modifications to the regular U-Net model included replacing the simple convolution layers with kernels of the same size as the Inception module and adding high-hop connections. In addition to the above, a proposed preprocessing method based on the N4ITK stable intensity normalization method was proposed in this research to increase the comparability of MR images. The proposed method for brain MR image segmentation was evaluated on the BraTS 2020 dataset and the results showed that replacing the conventional convolution layers with the Inception module improves the performance of the network. Moreover, the results obtained for the evaluation criteria confirmed that the hypothesis stated on the effect of jump connections and high jump connections on improving the performance of the predictive model was correct. The use of the proposed pre-processing method based on the N4ITK stable intensity normalization method in this research also improved the performance of the proposed model, because when this pre-processing method is not used, the network performance showed an average drop of 4.1 percent in the evaluation criteria.



بخش‌بندی تصاویر MR با استفاده از معماری U-Net مبتنی بر Inception و اتصالات پرشی بالا

فرناز حسینی^{۱*}، حامد سپهرزاده^۲، امیرحسین تعلیمیان^۳

۱ و ۲-۳- استادیار، گروه مهندسی کامپیوتر، دانشگاه فنی و حرفه‌ای، تهران، ایران.

چکیده

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

دریافت مقاله: ۱۴۰۲/۰۲/۰۷

بازنگری مقاله: ۱۴۰۲/۰۴/۰۱

پذیرش مقاله: ۱۴۰۲/۰۵/۲۲

کلید واژگان:

بخش‌بندی
U-Net
MRI
یادگیری عمیق
اتصالات پرشی بالا
ماژول Inception

*نویسنده مسئول: فرناز حسینی

پست الکترونیکی:

f-hoseini@tvu.ac.ir

تصویربرداری پزشکی یک تکنیک غیرتهاجمی می‌باشد که باعث توسعه قابل توجه در تشخیص و شناسایی بیماری‌های انسان شده است. در بین تمام تکنیک‌های تصویربرداری پزشکی، روش تصویربرداری تشدید مغناطیسی (MRI) دارای محبوبیت بیشتری است. این روش، برای سلامت انسان مضر نیست و می‌تواند تصویربرداری از جزئیات مغز انسان را با کیفیت بالایی انجام دهد. بخش‌بندی صحیح تومور مغزی در تصاویر MR دارای اهمیت بالایی می‌باشد. روش‌های سنتی برای بخش‌بندی تصاویر پزشکی زمان‌بر بوده و نیازمند تخصص بالایی هستند. روش‌های یادگیری عمیق نیز برای بخش‌بندی تومور مغزی از روی تصاویر MR معمولاً از لایه‌های کانولوشن عادی استفاده می‌کنند که در این صورت توانایی تشخیص ساختارهای ریزمقیاس و بزرگمقیاس را نخواهند داشت. در این پژوهش یک روش نوین مبتنی بر یادگیری عمیق برای بخش‌بندی تومور مغزی بر روی تصاویر MR ارائه شده است. روش پیشنهادی تعمیمی از معماری معروف U-Net می‌باشد با این تفاوت که به جای لایه‌های کانولوشن عادی از ماژول Inception استفاده شده است. ماژول Inception به علت استفاده از کرنل‌های کانولوشن با اندازه‌های مختلف به موازات یکدیگر می‌تواند ویژگی‌های ریزمقیاس و بزرگمقیاس را از تصویر استخراج کند. در معماری مدل پیشنهادی برای بهبود جریان اطلاعات در مرحله انتشار رو به جلو از اتصالات پرشی بالا استفاده می‌شود. همچنین یک روش پیش‌پردازش نوین بر پایه مد تصویر نیز در این پژوهش ارائه شده که شدت تصویر را با استفاده از مد تصویر نرمال‌سازی می‌کند. روش پیشنهادی بر روی مجموعه داده BraTS ۲۰۲۲ مورد ارزیابی قرار گرفته و نتایج دقت به دست آمده برای ضریب تشابه دایس با مقدار ۰.۹۱. بیانگر بهبود دقت تشخیص می‌باشد. نتایج ارزیابی نشان می‌دهد که هم فرضیه ارائه شده در مورد تأثیر اتصالات پرشی بالا در بهبود جریان اطلاعات و یادگیری درست بوده و هم استفاده از ماژول Inception توانسته به طور قابل توجهی معیارهای ارزیابی مدل را بهبود ببخشد.



مقدمه

تومور مغزی نوعی بیماری است که در اثر رشد غیرعادی سلول‌ها در مغز به وجود می‌آید. این بیماری برای هر فردی در هر سنی ممکن است رخ دهد [۱]. تومورهای مغزی از نظر شکل و اندازه مثل هم نیستند و از این نظر می‌توانند بسیار متنوع باشند. همین‌طور تومورها در یک محل خاص از مغز رخ نمی‌دهند و ممکن است شدت تصاویر متفاوتی داشته باشند [۲]. پزشکان بالینی نقش مهمی در ارزیابی تومورهای مغزی و درمان دارند. هنگامی که یک تومور مغزی از نظر بالینی تشخیص داده می‌شود نیاز به ارزیابی رادیولوژیست برای تعیین موقعیت تومور است. این اطلاعات برای انتخاب بهترین روش درمان از جمله جراحی، شیمی درمانی و پرتودرمانی بسیار مهم و حیاتی است؛ بنابراین ارزیابی تومورهای مغزی با روش‌های تصویربرداری در حال حاضر یکی از مسائل کلیدی در بخش‌های رادیولوژی می‌باشد [۳]. اما همواره پزشکان در تشخیص دقیق مکان تومور مشکل دارند. این مشکل معمولاً ناشی از خستگی زیاد رادیولوژیست، تصاویر با نویز بالا و سایر عوامل محیطی می‌باشد [۴]. بنابراین یک سیستم تشخیص هوشمند به کمک کامپیوتر می‌تواند به رادیولوژیست‌ها در تشخیص دقیق‌تر و زود هنگام کمک کند. استفاده از سیستم‌های تشخیص پزشکی هوشمند در تشخیص بیماری‌های مغزی، به عنوان دستیاری در کنار پزشکان و رادیولوژیست‌ها علاوه بر کمک به آنها، راه را برای شناسایی دقیق و عاری از خطا و تفکیک این بیماری‌ها از سایر بیماری‌های مشابه هموار می‌سازد [۵]. این سیستم‌ها و ابزارهای مشابه معمولاً به رادیولوژیست‌ها و پزشکان در تشخیص مطمئن بیماری کمک می‌کنند که تصویربرداری تشدید مغناطیسی (MRI) [۶] یکی از مهمترین این ابزارها می‌باشد. MRI در زمینه‌های متفاوتی از مطالعه و تشخیص ساختارها و بافت‌های مغز استفاده می‌شود؛ اما تحلیل تصاویر پزشکی کاری پیچیده است، زیرا آنها هم حجم زیادی از اطلاعات را در بر می‌گیرند و هم در برخی موارد آثار مخربی از قبیل نویز باعث مشکل شدن استخراج اطلاعات مورد نظر از این تصاویر می‌گردد. استفاده از روش‌های مبتنی بر کامپیوتر می‌تواند به طور قابل توجهی به کاربران در آنالیز و بررسی تصاویر MR برای بخش‌بندی تومور مغزی کمک کند [۷]. روش‌های مختلفی جهت بخش‌بندی تومور مغزی در تصاویر MR پیشنهاد شده‌اند [۸]؛ اما همچنان دستیابی به دقت بالاتر از نظر معیارهای ارزیابی ارائه شده توسط این روش‌ها حائز اهمیت است. تکنیک‌های محاسباتی و هوش مصنوعی می‌توانند به طور قابل توجهی به ساخت مدل‌های قابل تفسیر، آموزش‌دیده بر روی تصاویر پزشکی موجود برای شناسایی خودکار الگوها، استخراج برجسب‌ها و تشخیص موارد جدید کمک کنند. به طور کلی، مدل‌های هوش مصنوعی جایگزین نیستند، بلکه به عنوان یک روش مکمل، کمکی تخصصی برای افزایش دقت پزشکی هستند. مدل‌های محاسباتی مشتق شده می‌توانند تحقیقات علمی را بهینه کرده و عملکرد بالینی را تسریع بخشند. تقسیم‌بندی تصویر پزشکی فرآیند شناسایی، شناسایی و برجسب‌گذاری بخش‌های آناتومیک، نواحی مرتبط فیزیولوژیکی یا سازمان‌دهی شبکه‌ای ذاتی مغز است. نقشه‌های تقسیم‌بندی مغز، تصمیم‌گیری بالینی را تسهیل می‌کند و یک مرحله اصلی در تشخیص به کمک رایانه ارائه می‌دهد. در تقسیم‌بندی تصویر با یادگیری ماشین کلاسیک، ماسک‌های تقسیم‌بندی، به طور دستی مشتق شده و توسط متخصص برای آموزش یک طبقه‌بندی‌کننده یادگیری ماشین ارسال می‌شوند. به عنوان مثال، یک رویکرد یادگیری آماری برای مدل‌سازی بخش‌بندی ساختار آناتومیک مغز و توسعه یک اطلس مغزی احتمالی. این رویکرد زمان‌بر تلقی می‌شود و به متخصصان این حوزه نیاز دارد تا توضیحات دستی اولیه را به دست آورند. علاوه بر این، زمانی که مرز بین بافت‌های بدخیم و خوش‌خیم مبهم باشد، عملکرد نسبتاً ضعیف است. در این مطالعه روشی خودکار برای بخش‌بندی تصاویر MR مغزی مبتنی بر معماری U-Net پیشنهاد شده است که می‌تواند با دقت بالایی بخش‌بندی تصاویر MR مغزی را نسبت به سایر روش‌های مشابه انجام دهد. از آنجایی که استفاده از کرنل‌ها^۱ با اندازه یکسان در شبکه‌های عصبی کانولوشنی (CNN)^۲ [۹] معمولی

¹ Kernes² Convolutional Neural Network (CNN)

نمی‌تواند نواحی مورد نظر با اندازه‌های مختلف را در نظر بگیرد. در این پژوهش یک معماری یادگیری عمیق جدید مبتنی بر معماری U-Net ارائه می‌شود که در آن به جای استفاده از لایه‌های کانولوشنی ساده با کرنل‌های هم‌اندازه، از بلوک‌های Inception استفاده شده است. از آنجایی که هر ماژول Inception شامل لایه‌های کانولوشنی موازی هم‌اندازه‌ی کرنل‌های متفاوت می‌باشد، لذا می‌توان ساختارهای مختلف با مقیاس‌های متفاوت را تشخیص داد. در معماری پیشنهادی، کرنل‌های کوچک ساختارهای ریزمقیاس و جزئیات را استخراج کرده و کرنل‌های بزرگ‌تر ساختار کلی تصویر را استخراج می‌کنند. از طرفی برای بهبود جریان اطلاعات در مدل از اتصالات پرشی بالا^۱ استفاده می‌شود که در آن علاوه بر استفاده از ویژگی‌های به‌دست آمده از مرحله رمزگذاری^۲ هم‌تراز با مرحله رمزگشایی^۳، از ویژگی‌های به‌دست آمده از یک مرحله پایین‌تر در بخش رمزگذاری نیز استفاده می‌شود. به طور خلاصه مدل پیشنهادی برای بخش‌بندی تصاویر MR مغزی از یک معماری U-net مبتنی بر Inception استفاده می‌کند. اجرا و ارزیابی روش پیشنهادی بر روی تصاویر بانک داده BraTS ۲۰۲۰^۴ انجام شده است. نوآوری روش پیشنهادی نسبت به معماری U-net معمولی استفاده از ماژول Inception و همچنین افزودن اتصالات پرشی بالا یا Up skip connection‌ها برای بهبود فرآیند آموزش و افزایش دقت مدل است. همچنین در این پژوهش یک روش پیش‌پردازش نوین مبتنی بر روش نرمال‌سازی شدت استوار نیز پیشنهاد شده است که بر مبنای مد^۵ تصویر می‌باشد. روش پیش‌پردازش پیشنهادی شدت تصویر^۶ را با کم کردن مد^۷ تصویر (برای مثال: مقدار خاکستری مرتفع‌ترین ستون در نمودار هیستوگرام شدت) نرمال‌سازی می‌کند.

بخش‌بندی ادامه این مقاله به این ترتیب است که در بخش دوم مطالعات انجام شده در حوزه بخش‌بندی تومور مغزی مبتنی بر مدل‌های یادگیری عمیق مربوط سال‌های اخیر بررسی شده است. در بخش سوم روش پیشنهادی و جزئیات معماری U-Net پیشنهادی ارائه شده است. در بخش چهارم تست عملکرد روش پیشنهادی، معرفی بانک داده، محیط شبیه‌سازی و نتایج مربوط به نمودارهای خروجی بررسی شده؛ و نتیجه‌گیری کلی و پیشنهادهای آتی در بخش پنجم ارائه می‌گردند

پیشینه تحقیق

در پژوهش ویجای^۷ و همکاران [۱۰] بر اساس تکنیک‌های یادگیری عمیق^۸ موفقیت آمیز، یک شبکه U-Net که در آن اتصالات باقیمانده با ترکیبی از ادغام هرم فضایی (SPP)^۹ متصل بهم برای بخش‌بندی تومور مغزی پیشنهاد شد. U-Net راه حلی است که معمولاً برای بخش‌بندی معنایی^{۱۰} استفاده می‌شود و از یک روش نمونه برداری پایین برای بخش‌بندی تومورها استفاده می‌کند. U-Net‌ها برای انتقال اطلاعات در طول نمونه برداری به اتصالات باقیمانده متکی هستند. در روش پیشنهادی در این پژوهش اطلاعات از بلوک‌های مختلفی نمونه برداری شده تا دامنه بازسازی را افزایش دهد. این افزایش با ترکیب ویژگی‌های محلی با وابستگی‌های سراسری متناظر آنها فراهم شده است. در شبکه U-Net پیشنهادی از اتصالات پرشی بالا برای افزایش پارامترهای آموزشی در مدل استفاده شده که در نتیجه زمان آموزش و

¹ Up Skip Connections

² Encoding

³ Decoding

⁴ Brain Tumor Segmentation (BraTS 2020) <https://www.kaggle.com/datasets/awsaf49/brats2020-training-data>

⁵ Mode

⁶ Image Intensity

⁷ Vijay

⁸ Deep Learning

⁹ Spatial Pyramid Pooling-powered

¹⁰ Semantic Segmentation

پنجیدگی مدل را افزایش می‌دهد. از سوی دیگر، رویکرد پیشنهادی بدون تغییر تعداد پارامترهای قابل آموزش در ابعاد بزرگتر مانند $160 \times 192 \times 192$ مورد استفاده قرار گرفته است. روش پیشنهادی بر روی مجموعه^۲ BraTS ۲۰۲۱ مورد ارزیابی قرار گرفته و نتایج دقت^۱ به دست آمده برای متریک ضریب تشابه دایس^۲ که به اختصار ضریب دایس^۳ نامیده می‌شود، برابر با ۰.۸۸ بوده است.

در مطالعه هنری^۴ و همکاران [۱۱] به منظور خودکار و استاندارد کردن بخش‌بندی تومور مغزی، چندین مدل شبکه عصبی کانولوشنی (CNN) مبتنی بر معماری U-Net ارائه شده است. معماری منتخب از بین مدل‌های پیشنهادی یک معماری با نظارت عمیق و میانگین‌گیری وزن تصادفی بود. در مدل پیشنهادی از دو گروه مستقل از CNNها با دو خط لوله^۵ آموزشی متفاوت برای آموزش مدل استفاده شده و هر کدام یک نقشه بخش‌بندی تومور مغزی مجزا تولید کرده‌اند. سپس این دو خروجی با در نظر گرفتن عملکرد هر گروه برای مناطق فرعی تومور، برچسب مربوط به هر بیمار را ادغام می‌کند. عملکرد مدل پیشنهادی بر روی مجموعه^۶ داده BraTS ۲۰۲۰ مورد ارزیابی قرار گرفته و دقت به دست آمده برای مدل پیشنهادی بر روی این مجموعه برای ضریب دایس برابر با ۰.۸۱ بوده است.

در بررسی ژائو^۵ و همکاران [۱۲] یک روش جدید بخش‌بندی تومور مغزی با استفاده از ادغام شبکه‌های عصبی کانولوشنی تماماً متصل (FCNN)^۷ و میدان‌های تصادفی شرطی (CRF)^۸ در یک چارچوب منسجم برای به دست آوردن نتایج بخش‌بندی با ثبات فضایی ارائه شده است. این مدل بخش‌بندی مبتنی بر یادگیری عمیق با استفاده از تکه‌های تصویر دوبعدی در سه مرحله^۹ پیش رو مورد آزمایش قرار گرفت: (۱) آموزش FCNN با استفاده از تکه‌های تصویر؛ (۲) آموزش CRFها به عنوان شبکه‌های عصبی بازگشتی (CRF-RNN)^۸ با استفاده از تکه‌های تصویر و با پارامترهای FCNN ثابت؛ و (۳) میزان‌سازی دقیق FCNNها و CRF-RNN با استفاده از تکه‌های تصویر. نتایج تجربی نشان داد که این روش می‌تواند یک مدل قطعه‌بندی برای چهار مدالیته مورد استفاده در این بانک تصویر ایجاد کند و عملکرد رقابتی مناسبی همانند نمونه‌های ساخته شده در این زمینه داشته باشد. این روش بر اساس داده‌های تصویربرداری ارائه شده توسط بانک تصویر BraTS ۲۰۱۶ مورد ارزیابی قرار گرفته و نتایج دقت به دست آمده برای ضریب دایس برابر با ۰.۸۰ بوده است.

در پژوهش هوایی^۹ و همکاران [۱۳] با استفاده از DNNs^{۱۰} یک مدل عمیق جهت بخش‌بندی تومور مغزی ارائه شده است. در این مدل با استفاده از چندین معماری DCNN، ادغام پیشرفت‌های اخیر در زمینه‌هایی مانند لایه‌های کانولوشنی، حداکثر توزیع واحدهای Maxout و تنظیم پارامترهای حذف تصادفی (Dropout) مورد آزمایش قرار گرفت. محققین در این مطالعه نتایج اولیه برای بهترین شبکه در حال کار خود را بر روی مجموعه^۶ داده‌های BraTS ۲۰۱۳، leaderboard و challenge ارائه نمودند. در حالی که محققین در این پژوهش با معماری خود نتوانستند به بهترین نتایج شرکت‌کنندگان چالش BraTS ۲۰۱۳ برسند، ولی نتایج شان امیدوار کننده بود. روش پیشنهادی بر روی مجموعه^۶ داده‌های BraTS ۲۰۱۳ مورد ارزیابی قرار گرفته و نتایج دقت به دست آمده برای ضریب دایس برابر با ۰.۸۸ بوده است.

¹ Accuracy

² Dice Similarity Coefficient Metric

³ Dice Score

⁴ Henry

⁵ Zhao

⁶ Fully Convolutional Neural Network

⁷ Conditional Random Fields

⁸ Recurrent Neural Networks

⁹ Havaei

¹⁰ Deep Neural Networks

در بررسی میلتری^۱ و همکاران [۱۴] بر اساس شبکه عصبی کانولوشنی (CNN) مبتنی بر معماری U-Net در تصویربرداری پزشکی با تمرکز بر بخش‌بندی پاتولوژی مغز ارائه شده است. این روش بر اساس رای دهی به نام Hough استوار بود. Hough یک استراتژی است که امکان محلی سازی کاملاً خودکار و بخش‌بندی آناتومی مورد نظر را فراهم می‌سازد. این رویکرد نه تنها از نتایج بخش‌بندی CNN استفاده می‌کند بلکه با کشف ویژگی‌های تولید شده توسط عمیق‌ترین بخش مدل U-Net، رای دهی را نیز اجرا می‌نماید. محققین در این مطالعه نشان دادند که این روش بخش‌بندی مبتنی بر یادگیری توانمند، چند ناحیه‌ای و انعطاف پذیر بوده و می‌تواند به راحتی با روش‌های مختلف سازگاری یابد. آنها همگام با تلاش برای نشان دادن توانایی‌ها و رفتار CNN‌ها، یک بررسی سیستماتیک از عملکرد معماری شبکه انجام دادند. در نهایت تأثیر مقدار مختلف داده‌های آموزشی و نیز بعدپذیری داده‌های مختلف (2D، 2.5D و 3D) بر نتایج نهایی را بررسی نمودند. نتایج مربوط به بخش‌بندی تصاویر MR نشان دهنده ۲۶ منطقه از تومورها از نوع میان مغز^۲ بودند. روش پیشنهادی بر روی یک مجموعه از داده‌های سونوگرافی تصاویر MR^۳ مورد ارزیابی قرار گرفته و نتایج دقت به دست آمده برای ضریب دایس برابر با ۰.۸۵ بوده است.

در مطالعه پیرا^۴ و همکاران [۱۵] یک روش بخش‌بندی اتوماتیک بر اساس شبکه عصبی کانولوشنی (CNN)، برای کاوش کرنل‌های کوچک ۳ × ۳ پیشنهاد شده است. استفاده از کرنل‌های کوچک با توجه به تعداد کمتر وزن‌ها در شبکه، امکان طراحی یک معماری عمیق‌تر، علاوه بر داشتن اثر مثبت در برابر برازش بیش از حد را فراهم می‌کند. در این مطالعه استفاده از نرمال‌سازی به عنوان یک مرحله قبل از پردازش مورد بررسی قرار گرفت که اگر چه در روش‌های بخش‌بندی مبتنی بر CNN رایج نیست، اما به همراه تقویت داده‌ها برای بخش‌بندی تومور مغزی در تصاویر MR توانست مؤثر واقع گردد. روش پیشنهادی بر روی دو مجموعه داده 2013 BraTS و 2015 BraTS مورد ارزیابی قرار گرفته و نتایج دقت به دست آمده برای ضریب دایس به ترتیب برابر با ۰.۸۸ و ۰.۷۸ بوده است.

دووراک^۵ و همکاران [۱۶] یک روش پیش‌بینی با ساختار محلی برای انجام وظایف بخش‌بندی سه بعدی مورد بررسی قرار گرفت. این مطالعه ارزیابی سیستماتیک پارامترهای مختلفی را انجام داده که به حاشیه‌نویسی انبوه ساختارهای آناتومیکال مربوط می‌شود. محققین در این مطالعه شبکه عصبی کانولوشنی (CNN) را به عنوان الگوریتم یادگیری انتخاب نمودند، زیرا مشخص شد که این شبکه برای رسیدگی به همبستگی بین ویژگی‌ها بسیار مناسب است. آنها رویکرد خود را بر روی بانک داده 2014 BraTS ارزیابی نمودند که نتایج دقت به دست آمده برای ضریب دایس برابر با ۰.۸۴ بوده است.

زیکیس^۶ و همکاران [۱۷] امکان استفاده مستقیم از مطالعه شبکه عصبی کانولوشنی (CNN) را برای بخش‌بندی بافت تومور مغزی مورد بررسی قرار دادند. در این مطالعه به عنوان ورودی شبکه، اطلاعات شدت چند کانالی از یک بسته کوچک در اطراف هر نقطه با برچسب گذاری، مورد استفاده قرار گرفت. پیش‌پردازش شدت استاندارد تنها برای داده‌های ورودی جهت در نظر گرفتن تفاوت‌های اسکنر اعمال شد. هیچ پیش‌پردازشی برای خروجی CNN اعمال نشد. آنها نتایج اولیه نویدبخشی در مورد داده‌های آموزشی حاصل از بانک تصاویر 2013 BraTS را گزارش کردند. ادامه کار آنها برای ارائه نهایی شامل تغییرات معماری، تنظیمات دقیق پارامترها و آموزش در مورد بانک اطلاعاتی آموزش BraTS

¹ Milletari

² Midbrain

³ MRI Sonography

⁴ Pereira

⁵ Dvořák

⁶ Zikic

۲۰۱۴ خواهد بود که در پژوهش حاضر نتایجی از ادامه این کار ارائه نشده است. نتایج ارزیابی بر روی بانک داده BraTS ۲۰۱۳ دقت به دست آمده برای ضریب دایس را برابر با ۰.۸۳ نشان داده است.

این مرور مطالعاتی شامل مقالاتی است که بین سال‌های ۲۰۱۴ و ۲۰۲۳ منتشر شده‌اند. به طور خاص، ما بر روی مقالاتی تمرکز کردیم که رویکردهای طبقه‌بندی و تقسیم‌بندی تومور مغز را با استفاده از CNN توسعه دادند. در تمام این مطالعات از پایگاه داده استاندارد BraTS مربوط به سال‌های مختلف استفاده شده است. CNNها عملکرد پیشرفته‌ای را در وظایف بینایی کامپیوتری، مانند تقسیم‌بندی و طبقه‌بندی تومور مغزی در چند سال گذشته نشان داده‌اند، ما به سراغ این شبکه‌ها رفتیم؛ اما با این حال، CNNهای مورد استفاده در مطالعات بررسی شده به دلیل اندازهٔ هستهٔ کوچک خود اکثراً نتوانسته‌اند به طور مؤثر اطلاعات دوربرد یا وابستگی‌ها را ضبط کنند. وابستگی‌های دوربرد آنهایی هستند که خروجی مورد نظر به دنباله‌های تصویر ارائه شده در زمان‌های دور بستگی دارد. به دلیل شباهت اندام‌های انسان، بسیاری از نمایش‌های بصری در تصاویر پزشکی به ترتیب سازماندهی می‌شوند. تخریب این توالی‌ها به طور قابل توجهی بر عملکرد یک مدل CNN تأثیر می‌گذارد. به این دلیل است که وابستگی‌های بین توالی‌های تصویر پزشکی (مانند مدالیته، برش و پیچ) حاوی اطلاعات مهمی هستند. در مطالعه حاضر این وابستگی‌های دوربرد را به‌طور مؤثر با تکنیک‌هایی که می‌توانند روابط توالی را پردازش کنند، مدیریت کردیم. در این مطالعه جای CNN معمولی، یک مدل عمیق مبتنی بر معماری U-Net پیشنهاد شده است. معماری پیشنهادی در این مطالعه به معماری U-Net مورد استفاده در مرجع [۱۰] نزدیک است با این تفاوت که U-Net مورد استفاده در این پژوهش به جای استفاده از لایه‌های کانولوشنی ساده با کرنل‌های هم‌اندازه، از بلوک‌های Inception استفاده می‌کند تا با استفاده از لایه‌های کانولوشنی موازی هم در هر ماژول Inception با اندازه‌ی کرنل‌های متفاوت، بتوان وابستگی‌های دوربرد و ساختارهای مختلف با مقیاس‌های متفاوت را تشخیص داد. همچنین در U-Net از اتصالات پرشی بالا یا Up skip connectionها برای بهبود فرآیند آموزش و افزایش دقت استفاده شده است. روش پیش‌پردازش نرمال‌سازی شدت استوار نیز نوآوری دیگر پیشنهاد شده توسط این پژوهش نسبت به پژوهش‌های قبلی می‌باشد. این روش نرمال‌سازی تصویر می‌باشد که شدت تصویر را با کم کردن مد تصویر نرمال می‌کند.

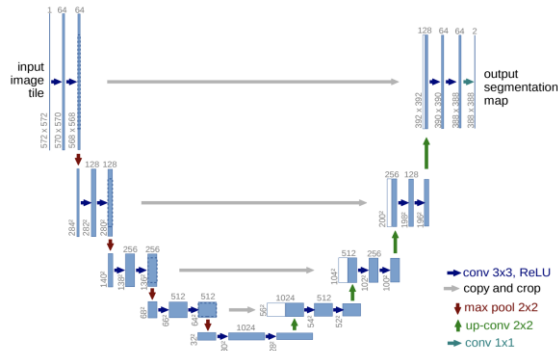
روش پیشنهادی

U-Net یک تکنیک بخش‌بندی معنایی است که برای اولین بار در سال ۲۰۱۵ برای تقسیم‌بندی تصویربرداری پزشکی پیشنهاد شد [۱۸]. این شبکه توسعه یافته یکی از مدل‌های بخش‌بندی یادگیری عمیق قبلی است (یک شبکه کانولوشنالی با پنجره کشویی) که معماری آن نسبتاً ساده بوده و مشابه با همان شبکه عصبی کانولوشنی (CNN) می‌باشد. همان‌طور که در شکل ۱ نشان داده شده معماری U-Net شبیه به حرف U است، که از این رو U-Net نامیده می‌شود. به طور کلی هر U-Net شامل یک رمزگذار^۱ (برای نمونه برداری پایین) و یک رمزگشا^۲ (برای نمونه برداری بالا) با اتصالات پرشی می‌باشد.

¹ Mode

² Encoder

³ Decoder



شکل ۱. معماری مدل U-Net [۱۸].

فلش‌های خاکستری در شکل ۱ نشان‌دهنده اتصالات پرشی هستند که نقشه ویژگی را با رمزگشا پیوند می‌دهد. معماری U-Net دارای یک مسیر گسترش‌سازی در سمت راست و یک مسیر فشرده‌سازی در سمت چپ می‌باشد. مسیر فشرده‌سازی متشکل از دو لایه 3×3 کانولوشنی است. هر یک از این لایه‌های کانولوشنی، یک تابع تابع فعال‌ساز غیرخطی (ReLU^۱) و یک لایه ادغام 2×2 از نوع Max Pooling برای کاهش نمونه‌گیری نگاشت ویژگی دارند.

بهینه‌سازی فرامترهای شبکه

از آنجایی که عملکرد هر مدل یادگیری عمیق تحت تأثیر زیر پارامترهای آن می‌باشد، لذا در این پژوهش با استفاده از داده‌های اعتبارسنجی که ۲۰٪ داده‌های آموزش می‌باشند، ابتدا بهترین زیر پارامترهای مدل عمیق را به دست آورده و سپس از آنها برای آموزش مدل اصلی استفاده شد. جدول ۱ نشان دهنده زیر پارامترهای بهینه و همچنین بازه جستجوی آنها می‌باشد. برای به دست آوردن مقادیر بهینه این زیر پارامترها از یک جستجوی گرید^۲ استفاده شده است. زیر پارامترهایی که بازه جستجویشان بیان نشده است در جستجو دخیل نبوده‌اند. در آموزش مدل‌ها از توقف زود هنگام استفاده شده است به گونه‌ای که اگر خطای مدل روی داده‌های اعتبارسنجی بعد از ۲۰ تکرار کاهش پیدا نکرد، آموزش متوقف می‌شود. این کار برای جلوگیری از بیش‌برازش^۳ انجام می‌شود.

جدول ۱. زیر پارامترهای بهینه مدل‌های U-Net پیشنهادی و بازه جستجوی آنها.

ردیف	زیر پارامتر	مقدار بهینه	بازه جستجو
۱	Inception تعداد بلوک‌های	۹	۱۱, ۷, ۵, ۳
۲	تعداد لایه‌ها در هر بلوک	۲	۳, ۲, ۱
۳	تعداد لایه‌ها در هر لایه به ترتیب بلوک‌ها	۸, ۱۶, ۳۲, ۶۴, ۱۲۸, ۶۴, ۳۲, ۱۶, ۸	۴ - ۲۶۵
۴	وزن‌دهی اولیه کرنل‌ها	he-normal	he-normal, glorot_normal
۵	توابع فعال‌ساز	ReLU	ReLU, sigmoid, tanh
۶	بهینه‌ساز	Adam	Adam, RMSProp, SGD

^۱ Rectified Linear Activation Function

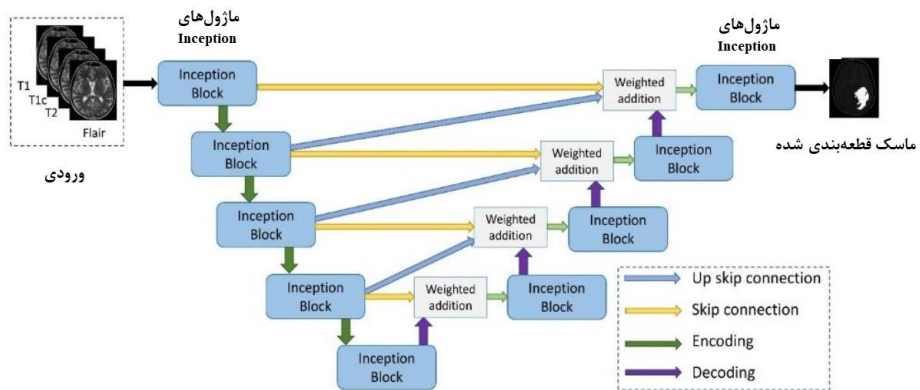
^۲ Grid Search

^۳ Overfitting

ردیف	زیر پارامتر	مقدار بهینه	بازه جستجو
۷	نرخ یادگیری	۰.۰۱	۰.۰۰۱, ۰.۰۰۵, ۰.۰۱
۸	اندازه‌ی پیچ ^۱ (دسته)	۶۴	۱۲۸, ۶۴, ۳۲, ۱۶
۹	پدینگ ^۲	Same	-
۱۰	حداکثر تعداد تکرار ^۳	۱۰۰	-
۱۱	توقف زودهنگام	۲۰	-

معماری U-Net پیشنهادی

با توجه به زیرپارامترهای به دست آمده، ساختار مدل U-Net پیشنهادی نهایی شامل ۹ بلوک Inception شد که هر بلوک شامل ۲ لایه است. هر لایه نیز شامل تعدادی ماژول Inception است که به ترتیب در جدول ۱ گزارش شده‌اند، لذا با توجه به تنظیم این فرآیندها، ساختار کامل مدل پیشنهادی مطابق شکل ۲ به دست آمد.



شکل ۲. معماری مدل U-Net پیشنهادی.

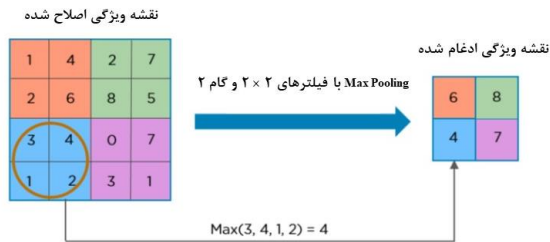
در این شکل، مقادیر فیلتر بر مقادیر موجود در ورودی ضرب شده و حاصل جمع مقادیر به دست آمده به نون لایه بعدی اختصاص داده می‌شوند. با اعمال این فیلتر بر روی کل ورودی، عملیات کانولوشن تمام و یک نقشه ویژگی ساخته می‌شود. از آنجایی که وزن‌های هر فیلتر در طول عمل کانولوشن ثابت است، بنابراین مدل پیشنهادی از اشتراک وزن استفاده می‌کند که این امر سبب کاهش تعداد پارامترها (وزن‌ها) برای آموزش می‌شود. با اعمال فیلترهای متعدد، چندین نقشه ویژگی به دست می‌آید. در ادامه، نقشه‌های ویژگی از یک تابع فعالساز غیرخطی (ReLU) عبور داده می‌شوند که نرخ فعال شدن هر کدام از نرون‌ها را مشخص می‌کند؛ سپس خروجی تابع فعالساز می‌تواند به لایه ادغام (از نوع Max Pooling) داده شود تا این‌که ابعاد ورودی را کاهش دهد. بیشترین مقدار موجود در یک ناحیه را جایگزین کل ناحیه می‌کند. در این حالت مهم‌ترین ویژگی انتخاب می‌شود. استفاده از ادغام سبب می‌شود مدل نسبت به چرخش، تارشدگی، انتقال و تغییر ابعاد مقاوم شود. این امر همچنین ابعاد مسأله را تغییر داده و از بیش برآزش

¹ Batch Size

² Padding

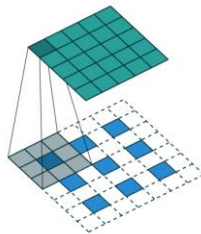
³ Epochs

جولوگیری می‌کند. شکل ۳ نشان دهنده یک مثال از Max Pooling می‌باشد. آموزش مدل U-Net پیشنهادی مانند شبکه‌های عصبی دیگر توسط روش پس انتشار انجام می‌شود.



شکل ۳. عمل ادغام از نوع Max Pooling.

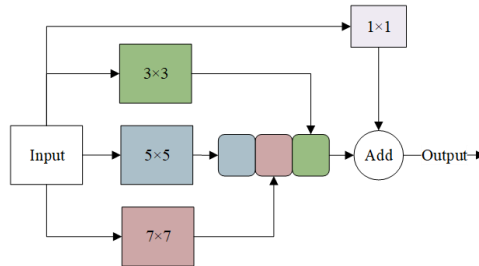
پس از اتمام بخش رمزگذاری، نقشه ویژگی‌های استخراج شده با بخش رمزگشایی داده می‌شود تا این که با استفاده از لایه‌های دیکانولوشن عمل بخش‌بندی را انجام دهد. ابعاد خروجی بخش رمزگشایی برابر با ورودی خواهند بود. عملیات در لایه دیکانولوشن مشابه همان عملیاتی است که در لایه کانولوشن صورت می‌گیرد، با این تفاوت که برخلاف عمل کانولوشن که چندین مقدار را از ورودی توسط فیلترهایی به یک مقدار تبدیل می‌کرد، عمل دیکانولوشن یک مقدار را به چندین مقدار تبدیل می‌کند. نحوه عملکرد این لایه در شکل ۴ نشان داده شده است. همان‌طور که در شکل ۴ قابل مشاهده است، خروجی لایه دیکانولوشن نقشه ویژگی بزرگ شده متراکم می‌باشد.



شکل ۴. عمل دیکانولوشن (مربع‌های آبی ورودی‌های اصلی می‌باشند که با عمل دیکانولوشن ابعاد نهایی افزایش پیدا کرده است).

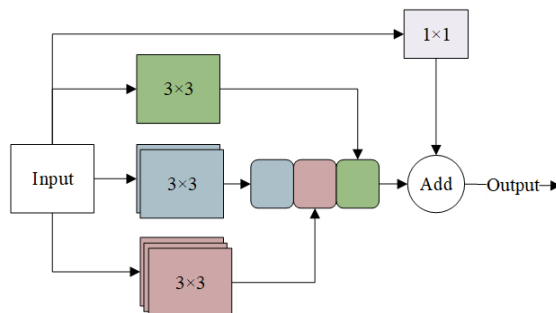
ماژول Inception

در این پژوهش، به منظور افزایش ظرفیت مدل برای بخش‌بندی و بهینه‌سازی عملکرد آن، از ماژول Inception در معماری U-Net استفاده شده است. ماژول Inception نشان داده است که می‌تواند سبب بهبود استخراج ویژگی‌های بصری با ظرفیت محاسباتی محدود شود [۱۹]. شکل ۴ نشان دهنده یک ماژول Inception می‌باشد که در آن کرنل‌های 3×3 ، 5×5 ، و 7×7 به موازات هم قرار گرفته‌اند و بر روی یک ورودی مشترک عمل می‌کنند. در انتها، ویژگی‌های به‌دست آمده توسط هر سه کرنل با یکدیگر ترکیب شده و با خروجی یک کرنل 1×1 جمع می‌شوند.



شکل ۴. نمونه‌ای از ماژول Inception [۱۹].

مشکل معماری نشان داده شده در شکل ۴ این است که تعداد وزن‌های قابل یادگیری شبکه با شروع فرآیند آموزش به طور قابل توجهی افزایش پیدا می‌کند. از این رو، در این پژوهش از یک معماری جایگزین برای ماژول Inception استفاده می‌شود. در معماری جایگزین هر کرنل 5×5 توسط دو کرنل 3×3 جایگزین شده است. از آنجایی که هر کرنل 7×7 معادل سه کرنل 3×3 می‌باشد، از این رو به جای کرنل 7×7 از سه کرنل 3×3 استفاده شده است. لذا معماری جایگزین به صورت شکل ۵ می‌باشد. همچنین به منظور افزایش سرعت همگرایی و حل مشکلات مربوط به آموزش مدل‌های عمیق، از نرمال‌سازی دسته‌ای^۱ استفاده شده است.



شکل ۵. معماری جایگزین برای ماژول Inception در این پژوهش.

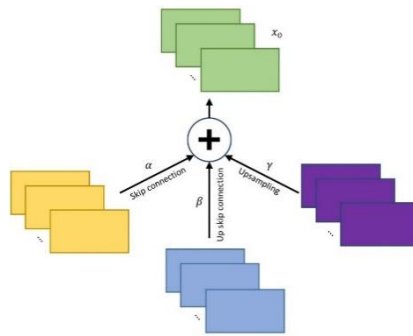
اتصالات پرشی بالا

در اتصالات شبکه‌ی مربوط به این پژوهش برای بهبود مسیر رمزگذاری و رمزگشایی از اتصالات پرشی بالا استفاده شده است. این امر سبب افزایش ظرفیت مدل برای یادگیری ویژگی‌های چند سطحی برای بخش‌بندی دقیق‌تر تومور مغزی می‌شود. در واقع اتصالات پرشی بالا در این پژوهش منجر به ایجاد یک مسیر جدید بین مسیر رمزگذاری و رمزگشایی در زمان انتشار رو به جلو می‌شود که این امر خود سبب استخراج ویژگی‌های سطح پایین‌تری شده و به بازیابی اطلاعات از دست رفته به هنگام رمزگذاری کمک می‌کند. همچنین اتصالات پرشی بالا جریان گرادینان را در فرایند انتشار رو به عقب (پس‌انتشار) نیز بهبود می‌بخشد که به عنوان یک راهنما برای یادگیری لایه‌های قبلی عمل کرده و سبب حل مشکل آموزش مدل می‌شود. همان‌گونه که در شکل ۱ نشان داده شده بود، در معماری U-Net با داشتن N لایه، یک اتصال پرشی متداول خروجی دو لایه i و $N-i+1$ را در یک U-Net معمولی ترکیب می‌کند. این درحالی است که در

¹ Batch Normalization

مدل U-Net پیشنهادی در این پژوهش علاوه بر ترکیب دو لایه i و $N-i+1$ ، دو لایه i و $N-i+2$ نیز ترکیب می‌شوند. این ترکیب‌ها به صورت وزن‌دار صورت می‌گیرند تا اینکه تجمیع نقشه‌های ویژگی سطح پایین با لایه‌های نمونه‌برداری بالا به صورت خودکار انجام شود. برای اطمینان از این‌که ورودی‌های فرایند ترکیب، وزن و اندازه یکسانی داشته باشند اتصالات پرشی بالای پیشنهادی و عملیات نمونه‌برداری بالا در مسیر رمزگشایی از لایه‌های کانولوشنی با کرنل 3×3 و گام‌های 2×2 استفاده می‌کنند تا این‌که به خروجی‌های هم‌اندازه‌ای دست پیدا کنند. این در حالی است که اتصالات پرشی معمولی از کانولوشن 1×1 با گام‌های 1×1 برای تراز کردن اندازه ویژگی‌ها استفاده می‌کنند. عملیات ترکیب وزن‌دار به صورت گرافیکی در شکل ۶ نشان داده شده است که به صورت عبارت (۱) فرموله می‌شود. در این عبارت x_0 خروجی ترکیب وزن‌دار پیشنهادی، x_{skip} ، x_{upskip} و x_{deconv} به ترتیب نقشه‌های ویژگی به دست آمده از اتصالات پرشی متداول، اتصالات پرشی بالا و لایه دیکانولوشن می‌باشند. در انتها، پارامترهای α ، β و γ وزن‌های مربوط به هر کدام از نقشه‌های ویژگی هستند که حین آموزش مدل یاد گرفته می‌شوند.

$$x_0 = \alpha \cdot x_{skip} + \beta \cdot x_{upskip} + \gamma \cdot x_{deconv} \quad (1)$$



شکل ۶. جزئیات اتصالات پرشی بالا در این پژوهش.

روش پیش‌پردازش پیشنهادی (مبتنی بر نرمال‌سازی شدت استوار N4ITK)

از آنجایی که تصاویر MR دارای گستره شدت^۱ متنوعی می‌باشند و به صورت متفاوتی توسط میادین بایاس^۲ تحت تأثیر قرار می‌گیرند، در این پژوهش یک روش نرمال‌سازی شدت استوار^۳ به نام روش N4ITK [۲۰] استفاده شده است. هدف از به‌کارگیری این روش مقایسه‌پذیر کردن تصاویر در کنار تصحیح میدان بایاس^۴ داده‌های MRI توسط روش N4ITK می‌باشد. نوآوری انجام شده در روش N4ITK پیشنهادی، نرمال‌سازی بر مبنای مُد (mode) تصویر است. این روش شدت تصویر را با کم کردن مُد تصویر (برای مثال: مقدار خاکستری مرتفع‌ترین ستون در نمودار هیستوگرام شدت) نرمال‌سازی می‌کند. همچنین در این روش انحراف معیار نیز به ۱ نرمال می‌شود. از آنجایی که نیمی از مغز ماده سفید^۵ می‌باشد، مقدار خاکستری مرتفع‌ترین ستون در هیستوگرام معادل با مقدار خاکستری ماده سفید است، لذا تطابق مقادیر شدت ماده سفید در اسکن‌های MRI و نرمال‌سازی توزیع‌های شدت به خوبی

¹ Intensity

² Bias Fields

³ Robust Intensity Normalization

⁴ Bias Field Correction

⁵ White Matter

می‌تواند تصاویر MR مختلف را مقایسه‌پذیر سازد. با این وجود، محاسبه انحراف معیار بر اساس میانگین شدت لزوماً دارای مفهوم بافتی ثابتی نیست. لذا در این پژوهش از انحراف شدت استوار^۱ به جای انحراف معیار استفاده می‌شود. انحراف استوار بر اساس مقدار خاکستری مرتفع‌ترین ستون هیستوگرام حساب می‌شود که نشان دهنده میزان گسستگی^۲ شدت به مقدار خاکستری ماده سفید می‌باشد. اگر فرض کنیم یک اسکن MRI با وکسل‌های^۳ $\{v_1, \dots, v_N\}$ وجود دارد که هر وکسل دارای شدت $v_k, k = 1, 2, \dots, N$ می‌باشد، انحراف استوار $\bar{\sigma} = \sqrt{\sum_{k=1}^N (\hat{I} - I_k)^2 / N}$ می‌باشد که \hat{I} بیانگر مقدار خاکستری مرتفع‌ترین ستون در هیستوگرام است. فرایند نرمال‌سازی شدت پیشنهادی در این پژوهش به ترتیب مراحل زیر به دست می‌آید:

- ۱- مقادیر شدت را به بین ۰-۲۵۵ تغییر می‌دهیم.
- ۲- هیستوگرام شدت را با بازه^۴ ۲۵۶ حساب می‌کنیم.
- ۳- مقدار خاکستری مرتفع‌ترین ستون هیستوگرام \hat{I} را کم کرده و انحراف استوار را تقسیم می‌کنیم.
- ۴- شدت هر وکسل را به یک ثابت σ ضرب کرده و با ثابت I_0 جمع می‌کنیم، سپس مقادیر شدتی را که زیر ۰ یا بالای ۲۵۵ هستند به ۰ و ۲۵۵ تصحیح می‌کنیم.

در این پژوهش، ثابت‌های σ و I_0 برابر با مقدار خاکستری مرتفع‌ترین ستون هیستوگرام و انحراف استوار تصویر آموزشی HGG ۲۰۲۰ NO.۰۰۰۱ از مجموعه^۵ BraTS ۲۰۲۰ قرار داده شده و توسط روش نرمال‌سازی شدت استوار N4ITK پیشنهادی پیش‌پردازش شده است. همچنین در این پژوهش برای پالس‌های T1c, T1, T2 و Flair مقدار $\sigma = 30, 31, 37$ و مقدار $I_0 = 75, 99, 95$ می‌باشد.

نتایج تجربی

در این بخش به منظور ارزیابی معماری U-Net بهینه ارائه شده، با به‌کارگیری بانک داده BraTS ۲۰۲۰ مدل پیشنهادی مورد آزمایش و بررسی قرار می‌گیرد. ارزیابی مدل پیشنهادی به منظور بخش‌بندی تصاویر MR با چهار مدالیته^۶ استفاده می‌گردد و به منظور مقایسه عملکرد دقت از معیار ارزیابی متریک ضریب تشابه دایس یا به اختصار ضریب دایس استفاده می‌شود.

معرفی بانک داده

در این پژوهش از مجموعه داده‌های چالش بخش‌بندی تومور مغزی چندمده^۶ BraTS ۲۰۲۰ استفاده شده است. در این پژوهش از ۲۲۰ مورد گلیوما با درجه^۷ بالا (HGG^۷) و ۵۴ مورد گلیوما درجه^۸ پایین (LGG^۸) برای مجموعه داده آموزش (Train)، و از ۱۱۰ مورد گلیوما بدون درجه‌بندی برای مجموعه داده^۹ آزمایش (Test) استفاده شد. اسکن مغز هر بیمار دارای مدالیته‌ها یا پالس‌های MRI مختلف می‌باشد که هر کدام از ۱۵۵ بخش تشکیل شده است که در کل ۶۲۰ تصویر برای هر بیمار به وجود می‌آورد. این مدالیته‌ها عبارتند از FLAIR, T1-weighted [T1]-contrast, T1

¹ Robust Intensity Deviation

² Discreteness

³ Voxels

⁴ Bin

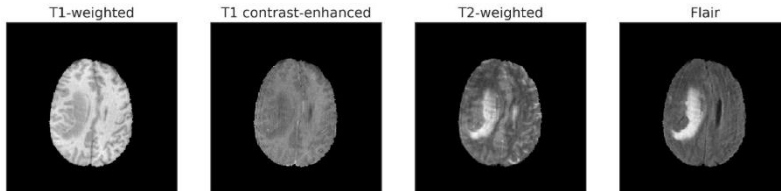
⁵ Modalities

⁶ Multimodal Brain Tumor Segmentation Challenge

⁷ High-grade Glioma

⁸ Low-grade Glioma

enhanced [T1c] و T2-weighted [T2]. یک نمونه از این اسکن‌ها در شکل ۷ قابل مشاهده است که تصویر جمجمه از تمام تصاویر حذف شده است. در بانک داده BraTS ۲۰۲۰ پنج برچسب وجود دارد. این برچسب‌ها شامل نکروسیس^۱ (برچسب ۱)، اِدمَا^۲ (برچسب ۲)، توموری که در حال بهبود نیست (برچسب ۳)، تومور در حال بهبود (برچسب ۴) و هر چیز دیگر (برچسب ۵) می‌باشند.



شکل ۷. نمونه تصویری از بانک داده BraTS ۲۰۲۰.

معیارهای ارزیابی

در این پژوهش عملکرد مدل U-Net پیشنهادی از منظر چهار معیار عملکرد مورد ارزیابی قرار می‌گیرد. این معیارها شامل ضریب تشابه دایس (Dice)، معیار حساسیت (Sensitivity)، مقدار پیش‌بینی مثبت (PPV^۳) و معیار شباهت جاکارد^۴ (Jaccard) [۲۱] می‌باشند.

معیار Dice شباهت بین دو نقشه بخش‌بندی شده را اندازه‌گیری می‌کند که مقدار آن بین صفر و یک است. مقادیر بالا و نزدیک یک نشان دهنده تطابق خوب بین پیش‌بینی‌ها یا واقعیت می‌باشند. با در نظر گرفتن نقشه پیش‌بینی دودویی^۵ به صورت $P \in (0,1)$ که از مدل پیش‌بینی کننده به دست آمده است و بخش‌بندی واقعی $T \in (0,1)$ ، ضریب دایس به صورت عبارت (۲) تعریف می‌شود.

$$Dice(P, T) = \frac{2 \times |P_1 \cap T_1|}{(|P_1| + |T_1|)} \quad (2)$$

در عبارت بالا، $|P_1 \cap T_1|$ بیانگر تعداد عناصری هستند که در هر دو مجموعه وجود دارند و $|P_1|$ و $|T_1|$ اندازهٔ دو مجموعه می‌باشد (به عبارت دیگر، تعداد وکسل‌هایی که $P = 1$ و $T = 1$ باشد).

معیار PPV یا مقدار پیش‌بینی مثبت و معیار Sensitivity برای ارزیابی تعداد بخش‌بندی‌های مثبت استفاده می‌شوند. PPV نسبت تعداد پیش‌بینی‌های مثبت درست به کل پیش‌بینی‌های مثبت بوده و Sensitivity نسبت پیش‌بینی‌های مثبت درست به تعداد پیش‌بینی‌های منفی می‌باشد که مقادیر آنها توسط عبارات (۳) و (۴) به دست می‌آیند.

$$PPV(P, T) = \frac{|P_1 \cap T_1|}{|T_1|} \quad (3)$$

¹ Necrosis

² Edema

³ Positive Predictive Value (PPV)

⁴ Jaccard Similarity

⁵ Binary

$$Sensitivity(P, T) = \frac{|P_1 \cap T_1|}{|T_1|} \quad (۴)$$

معیار شباهت Jaccard از محبوب‌ترین معیارهای ارزیابی در بخش‌بندی تصاویر پزشکی، است که به توسط عبارت (۵) محاسبه می‌شود.

$$Jaccard(P, T) = \frac{|P_1 \cap T_1|}{|P_1 \cup T_1|} \quad (۵)$$

محیط شبیه‌سازی

روش پیشنهادی با استفاده از زبان برنامه‌نویسی پایتون^۱ و کتابخانه تنسورفلو^۲ در محیط ویندوز^۳ و توسط یک رایانه استاندارد دارای پردازنده ۲.۴۰GHz CPU@i7-۵۵۰۰U (TM) Core (R) Intel، کارت گرافیک ۴ Nvidia GeForce GTX ۱۶۶۰ Ti با ۱۵۳۶ هسته و رم ۸ گیگابایت پیاده‌سازی شده است. سخت‌افزار GPU با استفاده از قابلیت هم‌زمانی اجرای بلاک‌ها می‌تواند سرعت اجرا را نسبت به سخت‌افزار CPU افزایش دهد. در صورتی که سایز تصویر کوچک باشد، به دلیل سربار ناشی از انتقال داده‌ها به حافظه پردازنده گرافیکی زمان اجرای محاسبات بر روی سخت‌افزار GPU بیشتر از زمان اجرای محاسبات بر روی سخت‌افزار CPU می‌شود، اما با افزایش تعداد سایز تصویر به دلیل افزایش تعداد بلاک‌ها زمان اجرای الگوریتم رفته رفته کمتر می‌شود. این کاهش به دلیل اجرای هم‌زمان بلاک‌ها بر روی سخت‌افزار GPU می‌باشد. در این مطالعه می‌توان برای موازی‌سازی از نرم‌افزار متلب^۴ مدل scatter/gather استفاده کرد. که یک الگوی متعارف برای انجام عملیات موازی در محیط متلب می‌باشد. در این مدل داده‌ها بر اساس PID^۵ به تمام نقاط مورد نظر در برنامه ارسال شده (scatter) و در نهایت بعد از انجام عملیات به محیط متلب بازگردانده (gather) می‌شوند. سایر کتابخانه‌های مورد استفاده عبارتند از: NumPy برای انجام عملیات ریاضی، Matplotlib برای رسم تصاویر، Seaborn برای رسم نمودارها، SimpleITK برای تصحیح میدان بایاس تصاویر، Scikit-learn برای تقسیم داده‌ها به سه مجموعه آموزش (Train)، آزمایش (Test) و اعتبارسنجی (Validation)، و nibabel برای خواندن تصاویر MR.

خروجی‌های مدل پیشنهادی

قبل از انجام هر کاری، تمام تصاویر با استفاده از روش نرمال‌سازی شدت استوار N4ITK^۶ پیشنهادی پیش‌پردازش می‌شوند. برای این کار ابتدا میدان بایاس تصاویر با استفاده از این روش تصحیح شده و پیش‌پردازش می‌شوند. شکل ۸ نشان دهنده اثر پیش‌پردازش بر روی ۳ نمونه از اسکن‌های T2 می‌باشد که به صورت تصادفی انتخاب شده‌اند. در این شکل ردیف بالا اسکن‌هایی قبل از پیش‌پردازش و ردیف پایین همان تصاویر بعد از پیش‌پردازش را نمایش می‌دهد. همان‌گونه که در شکل ۸ نشان داده شده، شدت و کسل‌ها بعد از اعمال روش پیش‌پردازش پیشنهادی بهبود یافته است که این امر سبب بهبود مقایسه‌پذیری تصاویر مختلف MR می‌شود.

¹ Python

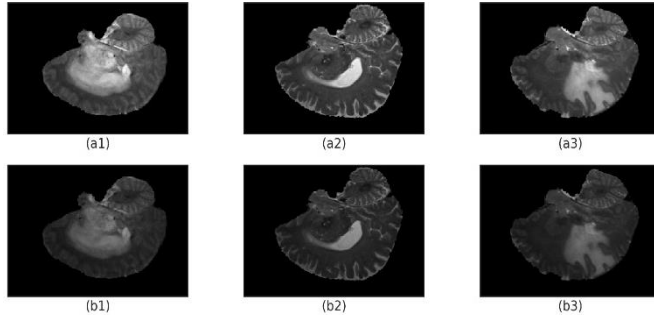
² TensorFlow

³ Windows

⁴ Graphical Processing Unit (GPU)

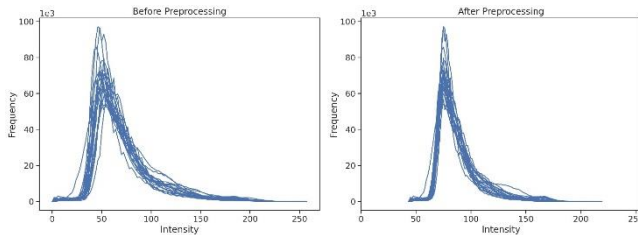
⁵ MATLAB

⁶ Program Identify



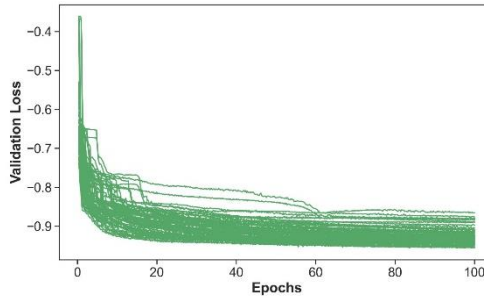
شکل ۸. نمونه‌های تصادفی انتخاب شده از اسکن‌های T2. (a1-a3) قبل از نرمال‌سازی، (b1-b3) بعد از نرمال‌سازی با N4ITK پیشنهادی.

برای انجام بهتر این مقایسه، هیستوگرام شدت ۲۰ اسکن تصادفی از اسکن‌های T2 قبل و بعد از انجام نرمال‌سازی N4ITK پیشنهادی به دست آمده و نتایج توسط نمودارهای شکل ۹ ارائه شده است. همان‌گونه که در شکل ۹ نشان داده شده، عرض هیستوگرام کاهش پیدا کرده است که نشان دهنده تصحیح وکسل‌هایی با شدت بسیار بالا و بسیار پایین می‌باشند.



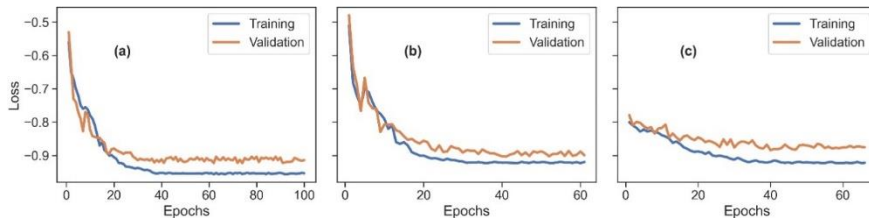
شکل ۹. هیستوگرام شدت در ۲۰ اسکن تصادفی از اسکن‌های T2 قبل و بعد از نرمال‌سازی با N4ITK پیشنهادی.

همان‌طور که پیش‌تر ارائه شد در این پژوهش با استفاده از داده‌های اعتبارسنجی که ۲۰٪ داده‌های آموزش می‌باشند، بهترین زیر پارامترهای مدل عمیق را به دست آورده و سپس از آنها برای آموزش مدل اصلی استفاده کردیم. در جدول ۱ زیر پارامترهای بهینه و همچنین بازه جستجوی آنها ارائه شد. برای به دست آوردن مقادیر بهینه این زیر پارامترها از یک جستجوی گرید استفاده شده است که شکل ۱۰ نشان دهنده نمودار خطا در برابر تکرار بر روی داده‌های اعتبارسنجی در حین این جستجو می‌باشد. زیر پارامترهایی که در جستجوی گرید منجر به کمترین خطای اعتبارسنجی شدند به عنوان زیر پارامترهای بهینه انتخاب شده که مقادیرشان پیش‌تر در جدول ۱ گزارش شد.



شکل ۱۰. خطای اعتبارسنجی در برابر تکرار در حین جستجوی گرید.

مدل U-Net پیشنهادی بهینه شده با ۹ ماژول Inception یکبار با اتصالات پرشی بالا (Inception-based U-Net + skip connections)، یکبار با اتصالات پرشی متداول (Inception-based U-Net + skip connections) و یکبار بدون هیچ‌گونه اتصالات پرشی (Inception-based U-Net) با زیرپارامترهای بهینه‌سازی شده در جدول ۱ آموزش داده شده است. نتایج مربوط به مقایسه عملکرد این مدل‌های بهینه‌ساز با یکدیگر در شکل ۱۱ نشان داده شده است. همان‌گونه که مشاهده می‌شود، مدل U-Net پیشنهادی با اتصالات پرشی بالا (بخش c در شکل ۱۱) به کمترین مقدار خطای آموزش و اعتبارسنجی دست یافته است. مدل U-Net پیشنهادی با اتصالات پرشی متداول (بخش b در شکل ۱۱) هر چند توانسته به خطای آموزش کمتری نسبت به مدل U-Net پیشنهادی بدون اتصالات پرشی (بخش a در شکل ۱۱) دست یابد، اما خطای اعتبارسنجی آن بالاتر است.



(a) مدل U-Net پیشنهادی با اتصالات پرشی بالا (Inception-based U-Net + up skip connections)

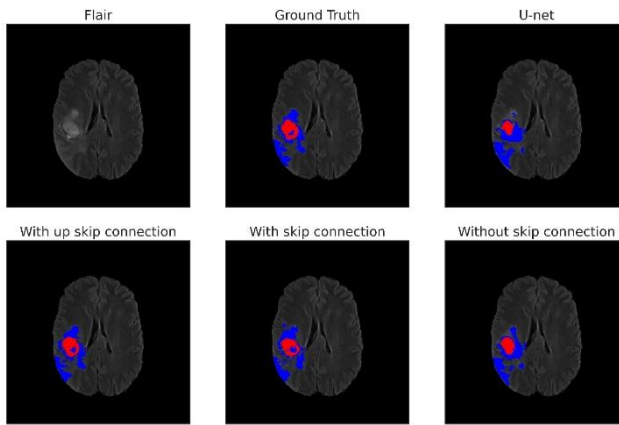
(b) مدل U-Net پیشنهادی با اتصالات پرشی متداول (Inception-based U-Net + skip connections)

(c) مدل U-Net پیشنهادی بدون اتصالات پرشی (Inception-based U-Net)

شکل ۱۱. نمودار یادگیری مدل‌های U-Net پیشنهادی.

شکل ۱۲ نشان دهنده نتایج بصری بخش بندی تصاویر MR مغزی با استفاده از مدل‌های U-Net پیشنهادی می‌باشد که بر روی یک اسکن Flair از مجموعه داده آزمایشی (BraTS ۲۰۲۰) اعمال شده است. در این شکل نواحی قرمز نشان دهنده تومور در حال بهبود و نواحی آبی نشان دهنده ادمای می‌باشند. مدل U-Net پیشنهادی با اتصالات پرشی بالا بهترین بخش بندی را داشته است. علاوه بر این مدل U-Net پیشنهادی با اتصالات پرشی متداول نیز عملکرد بهتری نسبت به مدل U-Net پیشنهادی بدون اتصالات پرشی داشته است. این امر نشان دهنده تأثیر مثبت افزودن اتصالات پرشی به شبکه می‌باشد. همان‌گونه که پیش‌تر اشاره شد، اتصالات پرشی باعث بهبود جریان اطلاعات در مرحله انتشار رو به جلو و بهبود جریان گرادینان خطا در مرحله انتشار رو به عقب می‌شود. همچنین نتایج خروجی در شکل ۱۲ نشان

می‌دهد که مدل U-Net پیشنهادی سبب بازیابی اطلاعات از دست رفته به هنگام رمزگذاری و آموزش بهتر مدل نیز شده است.



به ترتیب از بالا سمت چپ: یک اسکن Flair، نتیجه قطعه بندی صحیح نمایش داده شده در مجموعه داده، قطعه بندی با مدل U-Net معمولی، قطعه بندی با مدل U-Net پیشنهادی با اتصالات پرشی بالا، قطعه بندی با مدل U-Net پیشنهادی با اتصالات پرشی متداول و قطعه بندی با مدل U-Net پیشنهادی بدون اتصالات پرشی

شکل ۱۲. نتایج بخش‌بندی تصویر MR بر روی یک اسکن Flair از مجموعه داده BraTS ۲۰۲۰.

برای قابل درک شدن بیشتر مقایسه‌ها، مقادیر میانگین معیارهای ارزیابی بر روی داده‌های آزمایشی برای چهار روش بخش‌بندی ذکر شده محاسبه گردیده و در جدول ۲ گزارش شده است. بر اساس نتایج گزارش شده در جدول ۲، مدل U-Net پیشنهادی با اتصالات پرشی بالا بهترین عملکرد را در بخش‌بندی تصاویر MR مغزی داشته است. از طرفی تمام روش‌های مبتنی بر ماژول Inception نسبت به U-Net معمولی عملکرد بهتری داشته‌اند. این امر نشان دهنده لزوم استفاده از ماژول Inception می‌باشد، چرا که ماژول Inception همان‌گونه که پیش‌تر اشاره شد توانایی استخراج ساختارهای ریزمقیاس و بزرگ مقیاس را با بهره‌گیری از لایه‌های کانولوشنی با اندازه کرنل‌های متفاوت و موازی ایجاد می‌کند.

جدول ۲. میانگین نتایج معیارهای ارزیابی توسط مدل‌های مختلف U-Net پیشنهادی برای بخش‌بندی تصاویر MR از

مجموعه داده BraTS ۲۰۲۰.

Jaccard	PPV	Sensitivity	Dice	مدل پیشنهادی
۰.۸۸۴	۰.۸۳۲	۰.۸۵۱	۰.۸۴۶	مدل U-Net معمولی (U-Net)
۰.۸۹۴	۰.۸۵۶	۰.۸۵۹	۰.۸۷۸	مدل U-Net پیشنهادی بدون اتصالات پرشی (Inception-based U-Net)
۰.۹۰۴	۰.۸۸۷	۰.۸۷۴	۰.۸۹۳	مدل U-Net پیشنهادی با اتصالات پرشی متداول (Inception-based U-Net + skip connections)
۰.۹۲۴	۰.۹۱۶	۰.۸۹۴	۰.۹۱۳	مدل U-Net پیشنهادی با اتصالات پرشی بالا (Inception-based U-Net + up skip connections)

نتایج خروجی در جدول ۲ نشان می‌دهد که افزودن اتصالات پرشی و اتصالات پرشی بالا سبب بهبود عملکرد شبکه شده است. افزودن ماژول Inception سبب افزایش ضریب Dice به میزان ۱.۶٪، Sensitivity به میزان ۱٪، PPV به میزان ۲.۸٪ و Jaccard به میزان ۱.۱٪ شده است. این در حالی است که معیارهای ارزیابی ذکر شده بعد از اضافه کردن اتصالات پرشی متداول و ماژول Inception نسبت به شبکه U-Net معمولی به میزان ۳.۳٪، ۲.۷٪، ۶.۶٪ و ۲.۳٪ افزایش داشته است. این افزایش نسبت به زمانی که تنها ماژول Inception استفاده شده بود به میزان ۱.۷٪، ۱.۷٪، ۳.۸٪ و ۱.۲٪ می‌باشد. افزودن اتصالات پرشی بالا به همراه ماژول Inception سبب بهبود معیارهای ارزیابی به میزان‌های ۵.۷٪، ۵٪، ۱۰.۱٪ و ۴.۵٪ نسبت به U-Net معمولی شده است. همچنین افزودن اتصالات پرشی بالا منجر به افزایش عملکرد مدل نسبت به U-Net پیشنهادی بدون اتصالات پرشی به میزان ۴.۱٪، ۴.۰٪، ۷.۳٪ و ۳.۳٪ و نسبت به U-Net پیشنهادی با اتصالات پرشی متداول به میزان ۲.۴٪، ۲.۳٪، ۳.۵٪ و ۲.۲٪ می‌باشد. برای راحتی مقایسه، این مقادیر در جدول ۳ گزارش شده‌اند. همان‌گونه که مشاهده می‌شود، PPV در هر مرحله بیشترین بهبود را داشته است. PPV معیار مهمی است، زیرا که میزان پیش‌بینی‌های مثبت درست را نشان می‌دهد. این امر نشان دهنده آن است که بخش‌بندی تصاویر MR مغزی با دقت بیشتری انجام می‌پذیرد که از اهمیت بالایی برخوردار است.

جدول ۳. میزان بهبود معیارهای ارزیابی توسط مدل U-Net پیشنهادی با افزودن اتصالات پرشی و اتصالات پرشی بالا نسبت به یکدیگر و نسبت به U-Net معمولی.

مدل پیشنهادی	معیار ارزیابی (%)	بهبود نسب به مدل U-Net معمولی	بهبود نسب به مدل U-Net پیشنهادی بدون اتصالات پرشی	بهبود نسبت به مدل U-Net پیشنهادی با اتصالات پرشی متداول (+ Inception-based U-Net skip connections)
مدل U-Net پیشنهادی بدون اتصالات پرشی	Dice	۱.۶	-	-
Inception-based U-(Net)	Sensitivity	۱	-	-
	PPV	۲.۸	-	-
	Jaccard	۱.۱	-	-
مدل U-Net پیشنهادی با اتصالات پرشی متداول	Dice	۳.۳	۱.۷	-
	Sensitivity	۲.۷	۱.۷	-
Inception-based U-(Net + skip connections)	ppv	۶.۶	۳.۸	-
	Jaccard	۲.۳	۱.۲	-
مدل U-Net پیشنهادی با اتصالات پرشی بالا	Dice	۵.۷	۴.۱	۲.۴
	Sensitivity	۵	۴	۲.۳
Inception-based U-(Net + up skip connections)	PPV	۱۰.۱	۷.۳	۳.۵
	Jaccard	۴.۵	۳.۳	۲.۲

تأثیر روش پیش‌پردازش پیشنهادی مبتنی بر نرمال‌سازی شدت استوار N4ITK بر روی مدل U-Net پیشنهادی با اتصالات پرشی بالا در جدول ۴ بررسی شده است. در این جدول، عملکرد مدل U-Net پیشنهادی با اتصالات پرشی بالا یک‌بار با روش پیش‌پردازش پیشنهادی و یک‌بار بدون انجام این پیش‌پردازش گزارش شده است. مقادیر موجود در جدول ۴ میانگین معیارهای ارزیابی بر روی مجموعه داده BraTS ۲۰۲۰ را نمایش می‌دهد. نتایج نشان می‌دهد که در صورت عدم استفاده از روش نرمال‌سازی شدت استوار N4ITK عملکرد مدل U-Net پیشنهادی به صورت چشم‌گیری

کاهش می‌یابد. این امر به دلیل آن است که مقایسه‌پذیری تصاویر مختلف MR در صورت انجام ندادن پیش‌پردازش کاهش می‌یابد و در نهایت باعث بدتر شدن عملکرد مدل می‌شود.

جدول ۴. مقایسه عملکرد مدل U-Net پیشنهادی با اتصالات پرشی بالا/با/بدون اعمال روش پیش‌پردازش پیشنهادی (مبتنی بر نرمال‌سازی شدت استوار N4ITK).

روش پیش‌پردازش پیشنهادی	Dice	Sensitivity	PPV	Jaccard
با نرمال‌سازی شدت استوار N4ITK	۰.۹۱۳	۰.۸۹۴	۰.۹۱۶	۰.۹۲۴
بدون نرمال‌سازی شدت استوار N4ITK	۰.۸۸۴	۰.۸۶۰	۰.۸۵۵	۰.۸۹۸

بررسی نتایج به دست آمده در جدول‌های ۱ تا ۴ نشان می‌دهد که مدل U-Net پیشنهادی با اتصالات پرشی بالا کارایی بالایی دارد. جدول ۵ مقایسه‌ای از عملکرد دقت این مدل با سایر کارهای انجام شده در بخش ۲ مربوط به سال‌های اخیر را بر اساس ضریب دایس (Dice) نمایش می‌دهد.

جدول ۵. مقایسه دقت مدل U-net پیشنهادی با اتصالات پرشی بالا با سایر کارهای انجام شده در سال‌های اخیر بر اساس ضریب دایس (Dice).

ردیف	مرجع	سال چاپ	بانک داده	معیار ارزیابی Dice (%)
۱	[۱۰]	۲۰۲۳	BraTS ۲۰۲۱	۰.۸۸
۲	[۱۱]	۲۰۲۱	BraTS ۲۰۲۰	۰.۸۱
۳	[۱۲]	۲۰۱۸	BraTS ۲۰۱۶	۰.۸۰
۴	[۱۳]	۲۰۱۷	BraTS ۲۰۱۳	۰.۸۸
۵	[۱۴]	۲۰۱۷	MRI Sonography	۰.۸۵
۶	[۱۵]	۲۰۱۶	BraTS ۲۰۱۳	۰.۸۸
			BraTS ۲۰۱۵	۰.۷۸
۷	[۱۶]	۲۰۱۵	BraTS ۲۰۱۴	۰.۸۴
۸	[۱۷]	۲۰۱۴	BraTS ۲۰۱۳	۰.۸۳
۹	#	-	BraTS ۲۰۲۲	۰.۹۱

با توجه به موفقیت شبکه‌های عصبی کانولوشنی در دسته‌بندی و تشخیص اشیا، روشن است که این شبکه‌ها می‌توانند ویژگی‌های سطح بالا و مفیدی را از تصاویر بانک داده‌های بزرگ استخراج کنند. از این رو، استفاده از این شبکه‌ها در زمینه تشخیص پزشکی که یکی از الگوهای مهم یادگیری ماشین است، ضروری به نظر می‌رسد. همچنین پیشینه پژوهش‌های بررسی شده در این بخش نشان داد که در زمینه دسته‌بندی به معیار دقت توجه ویژه‌ای نشان داده شده است. بنابراین، انتخاب و طراحی شبکه‌ای عمیق نیز یک ضرورت به نظر می‌رسد تا بتواند ضمن داشتن دقت مناسب، از سرعت خوبی نیز برخوردار باشد. در این تحقیق، ابتدا یک بانک داده استاندارد انتخاب شده و بعد از فراخوانی تصاویر، به منظور بهبود کیفیت، بر روی تصاویر پیش‌پردازش‌هایی اعمال می‌شود. پس از پیش‌پردازش لیست ویژگی‌های کاربردی تهیه می‌شود. در نهایت از یک شبکه عصبی کانولوشنی عمیق برای فرآیند یادگیری و تشخیص استفاده می‌شود. اولین

یافته‌ای که از این پژوهش حاضر به دست آمد این بود که با عمیق کردن شبکه عصبی کانولوشنی و کنترل بیش برآزش می‌توان به دقت بالاتری رسید. نتایج به دست آمده در جدول ۲ نشان می‌دهد که مدل عمیق پیشنهادی دارای کارایی مناسبی است. جدول ۵ خلاصه‌ای از عملکرد مدل عمیق پیشنهادی را در مقایسه با سایر کارهای مورد مطالعه در بخش ۲ مربوط به این حوزه در سال‌های اخیر را نمایش می‌دهد. به منظور ایجاد یک مقایسه عادلانه، تنها معیار ارزیابی Dice برای مقایسه دقت مورد بررسی قرار گرفت. مقالاتی مورد بررسی در مراجع مورد مقایسه در این جدول همگی از یک مدل مبتنی بر شبکه‌های عصبی کانولوشنی با معماری‌های مختلف بر روی بانک داده BraTS استفاده کردند.

نتیجه‌گیری و پیشنهادهای آتی

تصویربرداری تشدید مغناطیسی (MRI) در زمینه‌های متفاوتی از مطالعه و تشخیص ساختارها و بافت‌های مغز استفاده می‌شود؛ اما تحلیل تصاویر پزشکی کاری پیچیده است، زیرا آنها هم حجم زیادی از اطلاعات را در بر می‌گیرند و هم در برخی موارد آثار مخربی از قبیل نویز باعث مشکل شدن استخراج اطلاعات مورد نظر از این تصاویر می‌شوند. بنابراین، استفاده از روش‌های مبتنی بر کامپیوتر می‌تواند کمک شایانی به پزشکان در آنالیز، بررسی و بخش‌بندی تصاویر MR داشته باشد. در این پژوهش یک روش نوین بخش‌بندی تصاویر MR مغزی مبتنی بر مدل U-Net با ماژول Inception و اتصالات پرشی بالا (Inception-based U-Net + up skip connections) پیشنهاد شد. اصلاحات انجام شده بر روی مدل U-Net معمولی شامل جایگزین کردن لایه‌های کانولوشن ساده با کرنل‌های هم‌اندازه با ماژول Inception، و افزودن اتصالات پرشی بالا بود. هدف از افزودن ماژول‌های Inception به شبکه استخراج ساختارهای ریزمقیاس و بزرگ‌مقیاس در تصاویر MR مغزی بود. از طرفی افزودن اتصالات پرشی بالا سبب بازیابی اطلاعات از دست رفته هنگام رمزگذاری شده و جریان اطلاعات را در مرحله انتشار رو به جلو بهبود بخشید. همچنین با بهبود جریان گرادینان خطا در مرحله پس‌انتشار سبب یادگیری بهتر لایه‌های کم عمق شد. علاوه بر موارد بالا، یک روش پیش‌پردازش پیشنهادی مبتنی بر روش نرمال‌سازی شدت استوار N4ITK در این پژوهش پیشنهاد شد تا این که بتواند مقایسه‌پذیری تصاویر MR را افزایش دهد. روش پیشنهادی برای بخش‌بندی تصاویر MR مغزی بر روی مجموعه داده‌های BraTS ۲۰۲۰ ارزیابی شد و نتایج نشان داد که جایگزین کردن لایه‌های کانولوشن متداول با ماژول Inception سبب بهبود عملکرد شبکه می‌شود. همچنین نتایج به دست آمده برای معیارهای ارزیابی ثابت می‌کنند که فرضیه بیان شده در مورد تأثیر اتصالات پرشی و اتصالات پرشی بالا بر بهبود عملکرد مدل پیش‌بینی کننده درست بوده است. استفاده از روش پیش‌پردازش پیشنهادی مبتنی بر روش نرمال‌سازی شدت استوار N4ITK در این پژوهش نیز سبب بهبود عملکرد مدل پیشنهادی می‌شود، چرا که در صورت عدم استفاده از این روش پیش‌پردازش، عملکرد شبکه به طور متوسط افت ۴.۱ درصدی در معیارهای ارزیابی را از خود نشان داد.

هرچند روش پیشنهادی عملکرد بسیار خوبی در بخش‌بندی تصاویر MR مغزی از خود نشان داد، با این حال پیشنهادی‌های دیگری جهت بهبود عملکرد شبکه و همچنین ارزیابی دقیق‌تر عملکرد آن ارائه می‌شوند که عبارتند از استفاده از الگوریتم‌های بهینه‌سازی تصادفی مانند الگوریتم ژنتیک^۱ و بهینه‌سازی ازدحام ذرات^۲ برای آموزش و بهینه‌سازی زیر پارامترهای شبکه، بررسی عملکرد روش پیشنهادی در بخش‌بندی ترتیبی هسته^۳، تمام^۴ و بهبود^۵ برای

¹ Genetic algorithm

² Particle swarm optimization

³ Core

⁴ Complete

⁵ Enhancing

بخش‌بندی تومور مغزی و بررسی تأثیر استفاده از بلوک‌های Conv LSTM^۱ به‌جای ماژول Inception برای پردازش تصاویر MR مغزی.

References

- [1] Herholz, K., Langen, K.-J., Schiepers, C., & Mountz, J. M. (2012). Brain Tumors. *Seminars in Nuclear Medicine*, 42(6), 356-370. <https://doi.org/10.1053/j.semnuclmed.2012.06.001>
- [2] Ayadi, W., Elhamzi, W., Charfi, I., & Atri, M. (2021). Deep CNN for Brain Tumor Classification. *Neural Processing Letters*, 53(1), 671-700. <https://doi.org/10.1007/s11063-020-10398-2>
- [3] Abd-Ellah, M. K., Awad, A. I., Khalaf, A. A. M., & Hamed, H. F. A. (2019). A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. *Magnetic Resonance Imaging*, 61, 300-318. <https://doi.org/10.1016/j.mri.2019.05.028>
- [4] Huang, J., Shlobin, N. A., Lam, S. K., & DeCuyper, M. (2022). Artificial Intelligence Applications in Pediatric Brain Tumor Imaging: A Systematic Review. *World Neurosurgery*, 157, 99-105. <https://doi.org/10.1016/j.wneu.2021.10.068>
- [5] Das, S., Nayak, G. K., Saba, L., Kalra, M., Suri, J. S., & Saxena, S. (2022). An artificial intelligence framework and its bias for brain tumor segmentation: A narrative review. *Computers in Biology and Medicine*, 143, 105273. <https://doi.org/10.1016/j.combiomed.2022.105273>
- [6] Wald, L. L., McDaniel, P. C., Witzel, T., Stockmann, J. P., & Cooley, C. Z. (2020). Low-cost and portable MRI. *Journal of Magnetic Resonance Imaging*, 52(3), 686-696. <https://doi.org/10.1002/jmri.26942>
- [7] Naglah, A., Khalifa, F., Khaled, R., Abdel Razek, A. A. K., Ghazal, M., Giridharan, G., & El-Baz, A. (2021). Novel MRI-Based CAD System for Early Detection of Thyroid Cancer Using Multi-Input CNN. *Sensors*, 21(11), 3878. <https://doi.org/10.3390/s21113878>
- [8] Yamanakkanavar, N., Choi, J. Y., & Lee, B. (2020). MRI Segmentation and Classification of Human Brain Using Deep Learning for Diagnosis of Alzheimer's Disease: A Survey. *Sensors*, 20(11), 3243. <https://doi.org/10.3390/s20113243>
- [9] Chauhan, R., Ghanshala, K. K., & Joshi, R. C. (2018, December 15-17). *Convolutional Neural Network (CNN) for Image Detection and Recognition* [Paper presentation]. 2018 First International Conference on Secure Cyber Computing and Communication, Jalandhar, India. <https://doi.org/10.1109/ICSCCC.2018.8703316>
- [10] Vijay, S., Guhan, T., Srinivasan, K., Vincent, P. M. D. R., & Chang, C.-Y. (2023). MRI brain tumor segmentation using residual Spatial Pyramid Pooling-powered 3D U-Net. *Frontiers in Public Health*, 11, 1091850. <https://doi.org/10.3389/fpubh.2023.1091850>
- [11] Henry, T., Carré, A., Lerousseau, M., Estienne, T., Robert, C., Paragios, N., & Deutsch, E. (2021). Brain Tumor Segmentation with Self-ensembled, Deeply-Supervised 3D U-Net Neural Networks: A BraTS 2020 Challenge Solution. In A. Crimi & S. Bakas (Eds.), *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (pp. 327-339). Springer International Publishing. https://doi.org/10.1007/978-3-030-72084-1_30

¹ Convolution-Long Short-Term Memory

- [12] Zhao, X., Wu, Y., Song, G., Li, Z., Zhang, Y., & Fan, Y. (2018). A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. *Medical Image Analysis*, 43, 98-111. <https://doi.org/10.1016/j.media.2017.10.002>
- [13] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., & Larochelle, H. (2017). Brain tumor segmentation with Deep Neural Networks. *Medical Image Analysis*, 35(4), 18-31. <https://doi.org/10.1016/j.media.2016.05.004>
- [14] Milletari, F., Ahmadi, S.-A., Kroll, C., Plate, A., Rozanski, V., Maiostre, J., Levin, J., Dietrich, O., Ertl-Wagner, B., Bötzel, K., & Navab, N. (2017). Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Computer Vision and Image Understanding*, 164, 92-102. <https://doi.org/10.1016/j.cviu.2017.04.002>
- [15] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *Institute of Electrical and Electronics Engineers Transactions on Medical Imaging*, 35(5), 1240-1251. <https://doi.org/10.1109/TMI.2016.2538465>
- [16] Dvořák, P., & Menze, B. (2016). Local Structure Prediction with Convolutional Neural Networks for Multimodal Brain Tumor Segmentation. In B. Menze, G. Langs, A. Montillo, M. Kelm, H. Müller, S. Zhang, W. Cai, & D. Metaxas (Eds.), *Medical Computer Vision: Algorithms for Big Data* (pp. 59-71). Springer International Publishing. https://doi.org/10.1007/978-3-319-42016-5_6
- [17] Zikic, D., Ioannou, Y., Brown, M., & Criminisi, A. (2014, September 14-18). *Segmentation of brain tumor tissues with convolutional neural networks*[Paper presentation]. Medical Image Computing and Computer-Assisted Intervention, Boston, Massachusetts. https://www.researchgate.net/publication/303703706_Segmentation_of_Brain_Tumor_Tissues_with_Convolutional_Neural_Networks
- [18] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234-241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28
- [19] Cahall, D. E., Rasool, G., Bouaynaya, N. C., & Fathallah-Shaykh, H. M. (2019). Inception Modules Enhance Brain Tumor Segmentation. *Frontiers in Computational Neuroscience*, 13, 44. <https://doi.org/10.3389/fncom.2019.00044>
- [20] Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. *Institute of Electrical and Electronics Engineers Transactions on Medical Imaging*, 29(6), 1310-1320. <https://doi.org/10.1109/TMI.2010.2046908>
- [21] Bag, S., Kumar, S. K., & Tiwari, M. K. (2019). An efficient recommendation generation using relevant Jaccard similarity. *Information Sciences*, 483, 53-64. <https://doi.org/10.1016/j.ins.2019.01.023>