



به نام خدا

دانشگاه تهران

دانشکده مهندسی برق و کامپیوتر



درس شبکه‌های عصبی و یادگیری عمیق

تمرین سوم

نام دانشجو	تیمین سلوکی	پرسش ۱
شماره دانشجویی		

فهرست

پرسش ۱	پیاده سازی مدل U-Net	۱
۱-۱	آماده سازی مجموعه داده	۱
۲-۱	پیاده سازی مدل	۲
۳-۱	تقویت داده	۴
۱-۴	بهینه ساز، متریک ها، و تابع هزینه	۱۰
۵-۱	آموزش مدل	۱۱
۶-۱	ارزیابی مدل	۱۶
پرسش ۲	تشخیص موجودات زیر آب	Error! Bookmark not defined.

شکل‌ها

- شکل ۱- عکس MRI مغز و ماسک مربوطه ۱
- شکل ۲- معماری مدل U-net ۳
- شکل ۳- Crop and zoom data augmentation ۵
- شکل ۴- Crop with padding data augmentation ۶
- شکل ۵ - shear data augmentation ۶
- شکل ۶ - adding noise data augmentation ۶
- شکل ۷ - Gaussian filter ۷
- شکل ۸- Contrasts data augmentation ۸
- شکل ۹ - data augmentation in this study ۹
- شکل ۱۰ -- data augmentation in this study ۹
- شکل ۱۱ - Report model 1 **Error! Bookmark not defined.**
- شکل ۱۲ -- Report model 2 **Error! Bookmark not defined.**
- شکل ۱۳ - Report model 3 **Error! Bookmark not defined.**
- شکل ۱۴ - Prediction 1 ۱۷
- شکل ۱۵ - Prediction 2 ۱۷
- شکل ۱۶ - Prediction 3 ۱۷
- شکل ۱۷ - Prediction 4 ۱۷
- شکل ۱۸ - Prediction 5 ۱۸

جدول ها

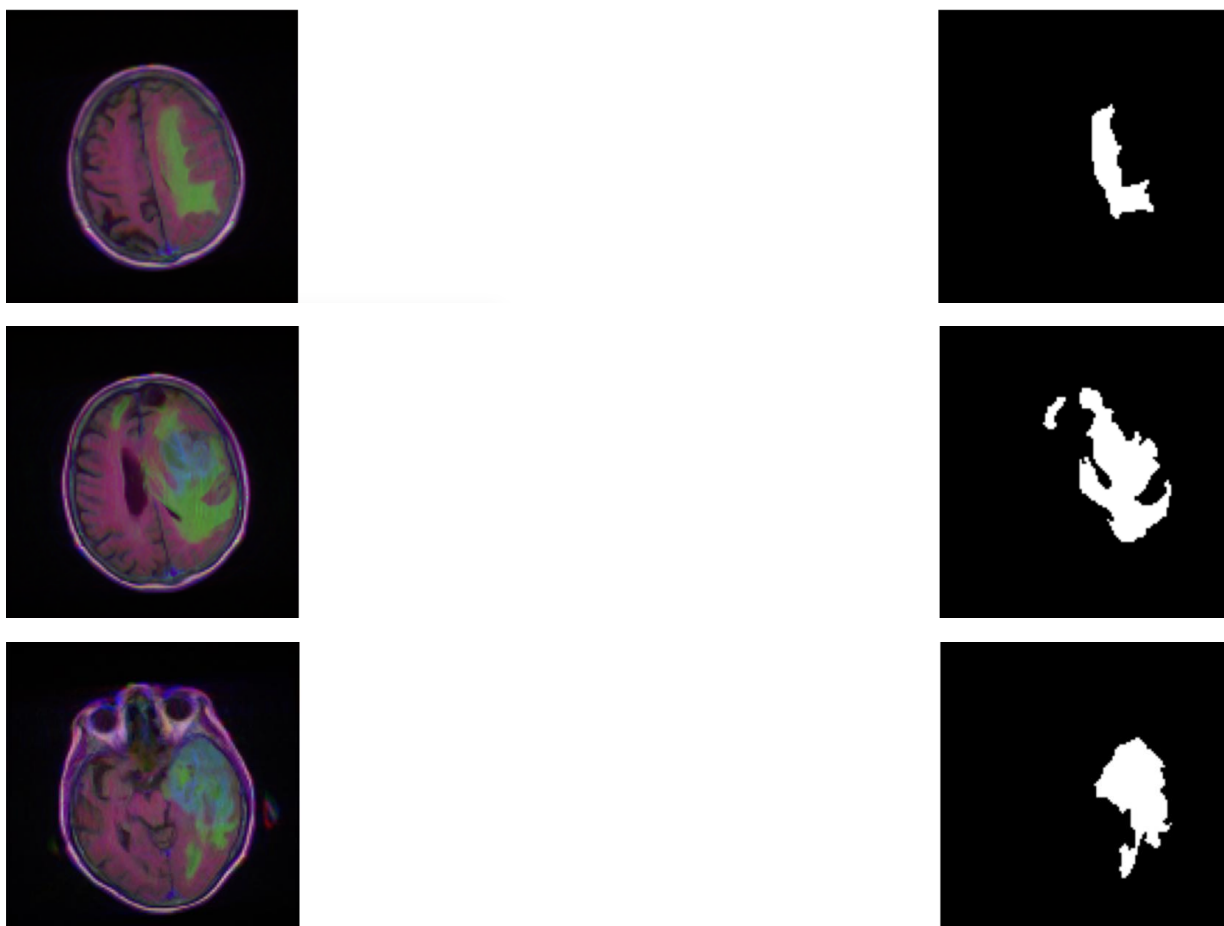
- جدول ۱- عملکرد مدل بر داده ی آموزش ۱۶
- جدول ۲- عملکرد مدل بر داده ی اعتبارسنجی ۱۶
- جدول ۳- عملکرد مدل بر داده ی تست ۱۸

پرسش ۱) پیاده سازی مدل U-Net

۱-۱) آماده سازی مجموعه داده

مجموعه داده ی انتخاب شده، مجموعه داده های تقسیم بندی LGG، شامل تصاویر MR مغز به همراه ماسک های تقسیم بندی ناهنجاری FLAIR دستی است. این تصاویر از آرشیو تصویربرداری سرطان (TCIA) به دست آمده است و مربوط به ۱۱۰ بیمار است. این دیتا مربوط به مجموعه گلیوما درجه پایین تر اطلس ژنوم سرطان (TCGA) با حداقل توالی بازبینی وارونگی ضعیف شده با مایع (FLAIR) و داده های خوشه ژنومی می باشد.

دادگان مربوطه از وبسایت کگل^۱ با استفاده از توکن مربوطه فراخوانی شده و در همان ابتدا عکس ها به سایز ۱۲۸ در ۱۲۸ (طبق مقاله ی پایه) در آمدند. به طور مثال، ۳ نمونه از داده ها به همراه ماسک مربوطه شان را در زیر مشاهده می کنید :



شکل ۱- عکس MRI مغز و ماسک مربوطه

^۱ [Brain MRI segmentation \(kaggle.com\)](https://www.kaggle.com/brain-mri-segmentation)

در مرحله ی بعدی آماده سازی داده، اقدام به نرمالسازی داده شد. در این پروژه نرمالسازی انجام شده MinMaxScaling بود.

پس از آماده سازی داده، دیتا به ۳ بخش آموزش، اعتبارسنجی و آزمون به ترتیب با درصد ۸۰٪، ۱۰٪ و ۱۰٪ تقسیم شد.

۲-۱) پیاده سازی مدل

برای پیاده سازی مدل، از مقاله ی Model Brain Tumor Segmentation using Enhanced U-Net with Empirical Analysis(2022) استفاده شد.

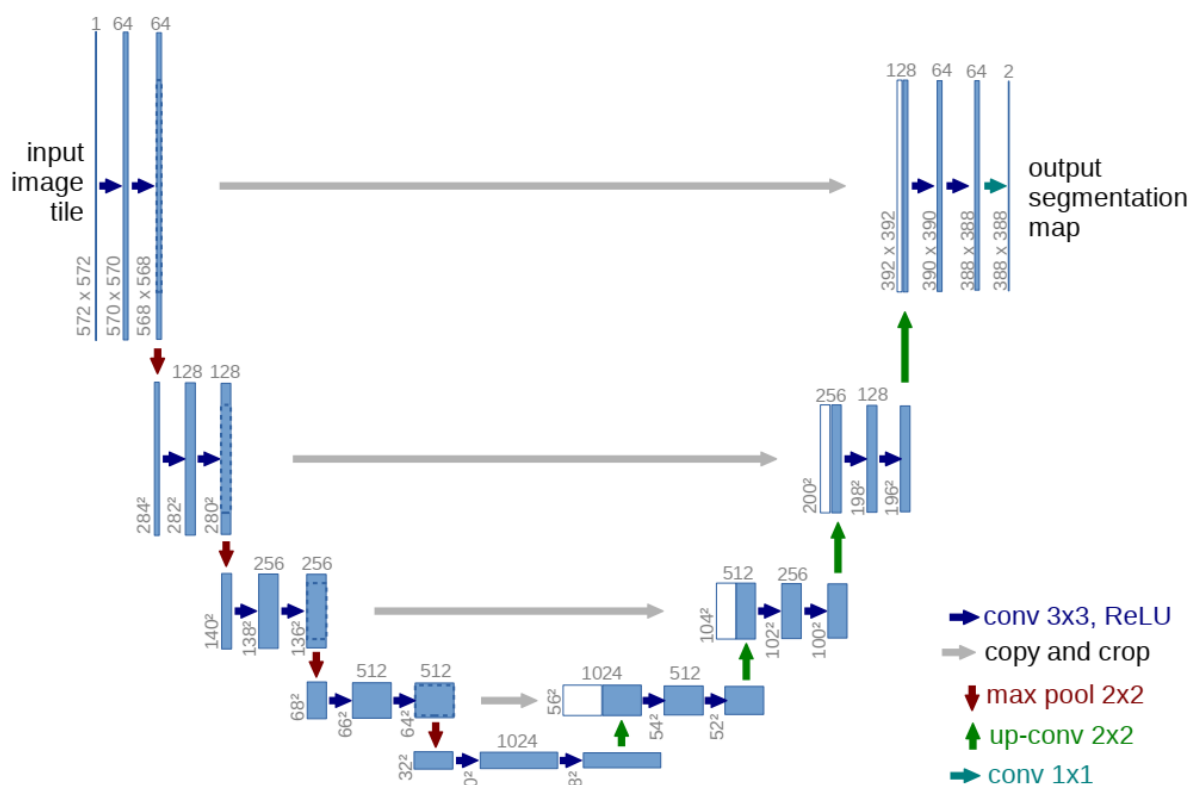
این مقاله بر تقسیم بندی تومورهای مغزی با استفاده از یک مدل پیشرفته U-Net با تجزیه و تحلیل تجربی تمرکز دارد. معماری استفاده شده، معماری اصلاح شده U-Net می باشد. مدل U-Net، در سال ۲۰۱۵ ارائه شد، حال به توضیح مدل می پردازیم:

این مدل ، یک معماری شبکه عصبی کانولوشنی است که در ابتدا برای بخش بندی تصاویر پزشکی ارائه شد.

به صورت کلی، روش کار شبکه های عصبی کانولوشنی برای طبقه بندی تصاویر با downsample کردن تصویر ورودی به یک feature map با رزولوشن پایین تر، آغاز و در انتها به upsampling پایان می یابد.

در بخش بندی تصاویر پزشکی (biomedical image segmentation)، مخصوصا در مواردی چون طبقه بندی سلولی یا طبقه بندی اورگان های بدن، حفظ اطلاعات فضایی (spatial information) بسیار حائز اهمیت است. شبکه ی U-Net برای پاسخ به این چالش ارائه شد: این مدل با ترکیب skip connections، شبکه را قادر می سازد که اطلاعات با رزولوشن بالا را در طول هر مسیر و در هر مرحله ، در طی downsampling و upsampling حفظ کند.

معماری U-Net شامل یک مسیر encoding و در ادامه یک مسیر decoding است. نام گذاری این شبکه به حرف U هم از همین فرایند کاهشی و سپس افزایشی الهام گرفته است.



شکل ۲- معماری مدل U-net

در بخش encoding، متشکل از لایه های کانولوشنی به دنبال لایه های max-pooling می باشد. این لایه ها به تدریج ابعاد فضایی را پایین می آورند و برای جبران این کاهش سایز داده ی ورودی ، تعداد کانال های ویژگی را دو برابر می کند. هر کانولوشن از لایه های کانولوشنال (اغلب با نرمال سازی دسته ای و فعال سازی ReLU) تشکیل شده است که به دنبال آن max-pooling 2×2 با $stride=2$ برای کاهش ابعاد spatial انجام می شود. جزییات معماری شبکه به صورت کامل در عکس بالا شرح داده شده است.

به منظور توضیح بیشتر معماری عکس شبکه می توان گفت که پس از ورود عکس ورودی و خروج از level اول، اندازه ی عکس از $572 \times 572 \times 3$ به $284 \times 284 \times 64$ و در قدم بعد پس از انجام max-pooling با هایپرپارامترهایی که پیش از این ذکر شد، ابعاد به $140 \times 140 \times 256$ رسیده و وارد مرحله ی دوم می شود. در این مرحله هم پس از اعمال دو لایه کانولوشنی با کرنل 3×3 و تابع فعالساز ReLU، ابعاد عکس به $68 \times 68 \times 512$ می رسد. شاهد این هستیم که در encoding تعداد کانال های عکس به منظور جبران کاهش رزولوشن ، دو برابر می شود. (64 به 128 به 512 به 1024). به طور خلاصه، encoding باعث کاهش رزولوشن فضایی اما تقویت چستی feature ها می شود.

بخش decoding، شامل لایه های upsampling و به دنبال آن لایه های کانولوشن است. این لایه ها feature map را به اندازه ورودی اولیه تبدیل می کنند و در عین حال تعداد کانال های ویژگی را کاهش

می‌دهند. این Upsampling می‌تواند با تکنیک‌هایی چون transposed convolutions یا bilinear interpolation انجام شود.

بین هر مرحله ی downsampling و upsampling از Skip connection ها استفاده می‌شود که feature map را از encoder به decoder به هم متصل می‌کنند و اطلاعات فضایی دقیقی را در طول upsampling ارائه می‌دهد. Skip connection نوآوری اصلی این مدل است که به شبکه اجازه می‌دهد اطلاعات فضایی دقیق را حفظ کند و همچنین از vanishing gradient جلوگیری می‌کند و فرآیند آموزش را بهبود می‌دهد.

لایه نهایی U-Net معمولاً از یک لایه کانولوشن 1×1 تشکیل شده است که به دنبال آن یک تابع فعال ساز (به عنوان مثال، sigmoid برای تقسیم بندی باینری یا softmax برای تقسیم بندی چند کلاسه) استفاده می‌شود. خروجی U-Net یک ماسک طبقه بندی (segmentation mask) با پیش بینی های پیکسلی است که در آن هر پیکسل نشان دهنده کلاس یا برجسب پیش بینی است.

از مزایای این شبکه می‌توان گفت که از طریق Skip connection ها حفظ اطلاعات انجام می‌دهد و آموزش کارآمدش جریان گرادینان را در آموزش شبکه، روان می‌کند و همچنین به راحتی می‌توانست با تنظیم معماری این شبکه، تابع loss، دادن و پیکربندی های ورودی/خروجی، شبکه را برای کارهای مختلف تقسیم بندی تصاویر پزشکی تطبیق داد.

معماری پیاده شده در مدل مقاله ی پایه بدین صورت بود که بر روی شبکه ی U-Net دو بعدی که پیش از این توضیح داده شد، پس از تغییر سایز عکس های ورودی به $128 * 128$ ، نرمالسازی MinMax، و استفاده از one hot encoding مدل را پس از تقسیم داده ها به بخش آموزش، تست، اعتبارسنجی با batch size=1 در ۲۳۵ اپاک آموزش داده، در این آموزش از بهینه ساز آدام و تابع هزینه categorical crossentropy استفاده کرد.

۳-۱) تقویت داده

تکنیک های تقویت داده، انواع مختلفی دارند چون تغییرات هندسی مثل چرخش، برش و بزرگنمایی، یا دیگر تغییرات چون اضافه کردن نویز، فیلتر گاوسی، تغییر کانترست عکس، تغییر روشنایی و غیره. در این بخش به توضیح بعضی از این تکنیک ها می‌پردازیم.

❖ Rotation

این تبدیل شامل چرخش تصویر به اندازه ی زاویه دلخواه است. در تجزیه و تحلیل تصاویر پزشکی، افزایش یک مجموعه داده با چرخش های تصادفی در زوایای ۱۰ تا ۱۷۵ درجه یا -۱۵ تا ۱۵ درجه رایج است.

❖ Flips

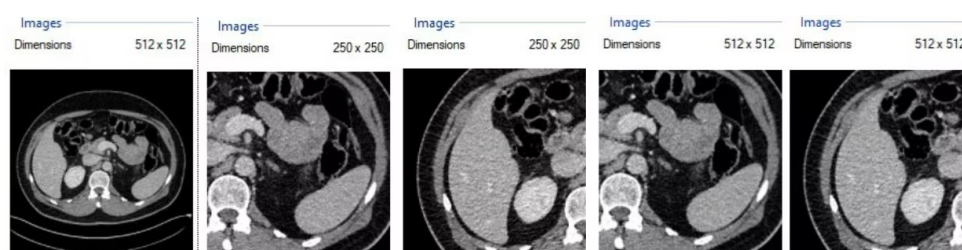
فلیپ تصویر در امتداد یک محور تقارن برای بهبود تنوع مصنوعی تصاویر پزشکی، انجام میگیرد. این فلیپ در راستای افق یا به صورت عمودی انجام میگیرد زیرا که تصاویر را می توان در وضعیت supine یا prone باشند و حاوی تغییرات آناتومیکی (مانند situs inversus) باشد. یک ارگان بدن، هر جایی که در بدن باشد، همیشه همان عضو خواهد بود. برخلاف تصاویر طبیعی که بسته به نوع شیء نمایش داده شده، انعکاس های عمودی (مثلاً یک گربه) هیچ معنای واقعی ای ندارند، زیرا تصویر منعکس شده غیر حقیقی تلقی می شود.

❖ Shift

این تبدیل در امتداد محور x یا y به صورت تصادفی انجام می شود. تصویر تبدیل شده همان جهت تصویر اصلی را حفظ می کند و در جهت تعیین شده حرکت میکند.

❖ Crop and zoom

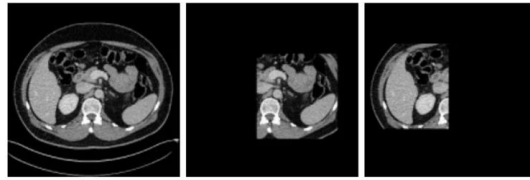
این دو تغییر برخی از قسمت های تصویر را برش می دهند و تنها بخشی از آن را حفظ می کنند. هنگام این تغییر ابعاد تصویر کاهش می یابد و لزوماً سائز تصویر حفظ نمی شود. هنگام زوم کردن البته نسبت تصویر اصلی حفظ می شود. برای انجام این کار، پیکسل هایی که برای بازیابی ابعاد اصلی ایجاد می شوند، دارای یک مقدار درونیابی شده از تابعی هستند که قبل از اعمال زوم روی تصویر اصلی (مثلاً میانگین روشنایی پیکسل های مجاور) تعریف شده است. در نتیجه، کاهش کیفیت را می توان مشاهده کرد.



شکل ۳- Crop and zoom data augmentation

❖ Crop with padding

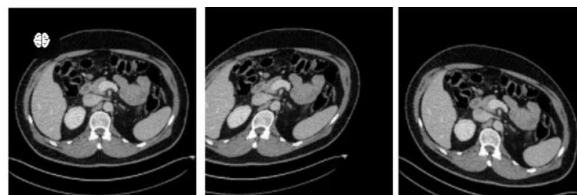
همانند تغییر قبل، برش انجام می شود، اما برای بازیابی/حفظ ابعاد تصویر اصلی، پر کردن با مقدار پیکسل ۰ در اطراف تصویر برش داده شده انجام می شود. مانند شکل زیر:



شکل ۴- Crop with padding data augmentation

❖ Shear

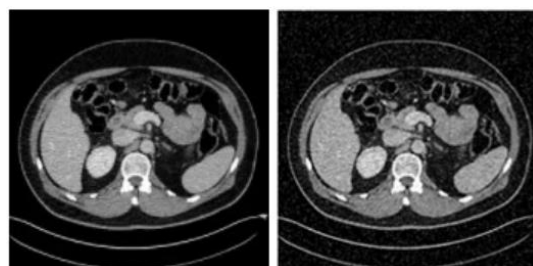
برش افقی (در امتداد محور x) و یا برش عمودی (در امتداد محور y) یک تبدیل افقی است که شامل جابجایی در جهت مخالف بالا و پایین تصویر (Shear افقی) یا سمت راست و چپ تصویر است (Shear عمودی). برخلاف روش های قبلی، در این روش تصویر مخدوش می شود. برش های انجام شده بر روی تصاویر پزشکی برای مثال مقادیر پارامتری از ۰.۱ تا ۰.۳۵ یا با نسبت x/y تا ۱۵٪ دارند. نمونه ای از این تغییر را در زیر مشاهده میکنید: عکس سمت چپ تصویر اصلی و عکس وسط shearing با $x = -0.3$ عکس سمت راست، shearing با $y = 0.3$ می باشد.



شکل ۵ - shear data augmentation

❖ Adding noise

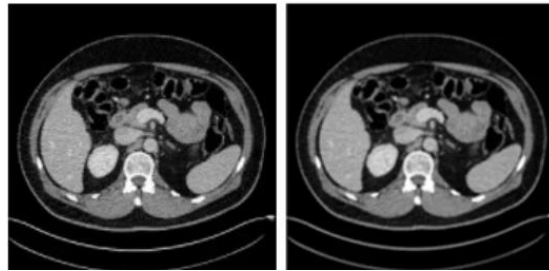
نویز یک تغییر تصادفی در روشنایی تصویر است که کیفیت تصویر اصلی را کاهش می دهد و می تواند چندین نوع باشد: نویز سفید گوسی، نویز پواسون، نویز ضربه ای (نمک و فلفل). در شکل زیر، عکس سمت راست، پس از اضافه کردن نویز salt and pepper بر روی تصویر اصلی سمت چپ می باشد.



شکل ۶- adding noise data augmentation

Gaussian filter ❖

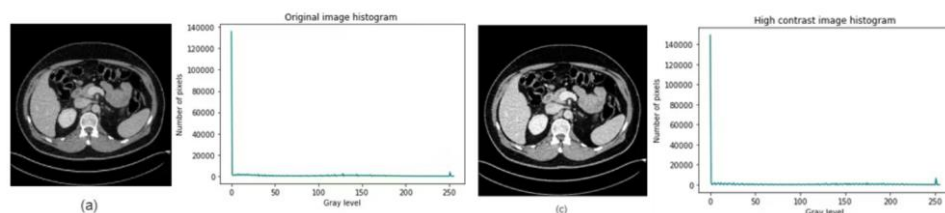
این فیلتر کانولوشن دارای هسته ای به شکل گاوسی است که با اعمال تابع گاوسی روی پیکسل های تصویر، کل تصویر اصلی را اسکن می کند تا آن را محو کند. در این روش، میانگین وزنی پیکسل های کانولوشنی را با دادن وزن بیشتری به پیکسل های نزدیک به مرکز هسته (برخلاف یک فیلتر متوسط که در آن هر پیکسل موجود در هسته وزن یکسانی دارد) ایجاد می کند. در تصویر زیر، عکس سمت چپ، فیلتر شده ی عکس سمت راست با کرنل سایز ۵*۵ است.

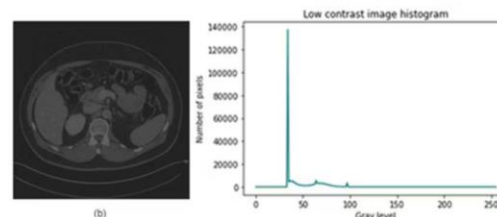


شکل ۷- Gaussian filter

Contrast ❖

کنتراست تصویر زمانی افزایش می یابد که پیکسل های تیره تر تصویر تاریک تر و پیکسل های روشن تر روشن تر شوند: لذا یک تصویر متضاد حاوی مقدار بیشتری سیاه و سفید خواهد بود. افزایش کنتراست را در هیستوگرام تصویر زیر مشاهده می کنید که فاصله بین روشن ترین پیکسل ها و تاریک ترین پیکسل های تصویر بیشتر است. بنابراین هنگامی که کنتراست افزایش می یابد، تعداد پیکسل های سفید و سیاه موجود در تصویر نیز افزایش می یابد، همانطور که می توان با مقایسه هیستوگرام تصاویر عکس سمت a (عکس پایه) و عکس b (کاهش کنتراست) و عکس c (افزایش کنتراست) مشاهده کرد. تغییرات کنتراست انجام شده برای تقویت داده ها از تصاویر پزشکی می تواند از ۰.۸ تا ۱.۲ باشد و اغلب به طور تصادفی در مجموعه داده اعمال می شود.



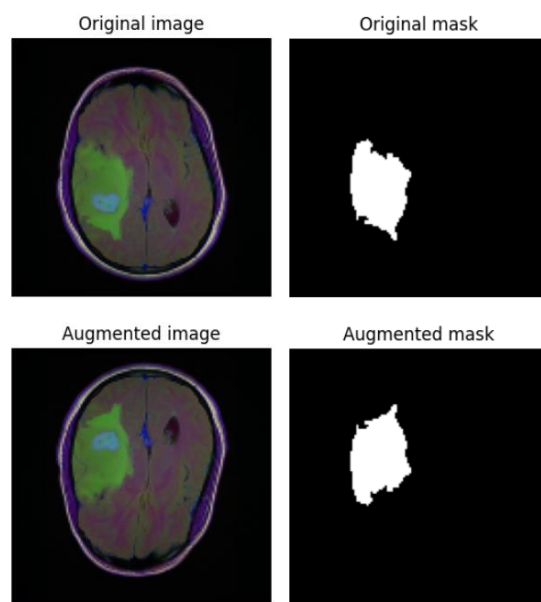


شکل ۸-Contrast data augmentation

❖ Brightness

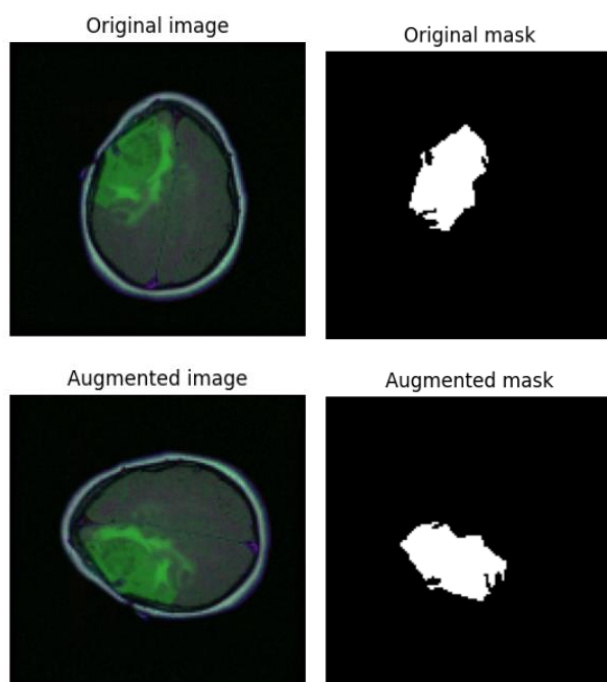
در فضای ادراکی HSV (Hue, Saturation, Value) تصاویر در مقیاس خاکستری مقدار اشباع صفر دارند و هیچ رنگ خاصی ندارند. این تصاویر وابسته به مقدار شاخص های HSV هستند. به منظور افزایش اندازه مجموعه داده در تصویر برداری پزشکی، تغییرات روشنایی متعلق به بازه $[-0.1, 0.1]$ را می توان به صورت تصادفی اعمال کرد.

برای پروژه ی این درس از کتابخانه ی albumentations استفاده شد . کد نوشته شده تضمین می کند که تصویر ورودی و ماسک خروجی مجموعه ای از افزایش ها را با پارامترهای یکسان دریافت خواهند کرد. نکته ی قابل توجه در استفاده از تکنیک های داده افزایی، این است که اگر کیفیت داده های ورودی به قدر کافی خوب نباشد، داده افزایی به جای بهبود آموزش مدل، منجر به تضعیف مدل می شود. نکته ی حائز اهمیت در انتخاب تکنیک های داده افزایی، می بایست به نتیجه ی آگمنت شده ی عکس توجه کرد. از تجربه ی تمرین قبلی درس، چندین تکنیک داده افزایی را همزمان برای تولید داده استفاده کردیم. در کمال تعجب، عملکرد مدل بدون داده های آگمنت شده ۸۵٪ صحت بود و بعد از اضافه کردن داده های آگمنت شده از ترکیب تمامی تکنیک ها، عملکرد مدل به ۶۶٪ صحت افت کرد. نتیجه ای که از تمرین قبل گرفتیم این بود که باید در استفاده ی تکنیک های data augmentation، احتیاط کرد. به همین منظور به طور مثال توصیه نمی شود که از shear با درصد بسیار زیاد استفاده کرد. بنابر این برای پیاده سازی داده افزایی در این پروژه به چند تکنیک اصلی اکتفا کردیم و از ادغام این تکنیک ها پرهیز کردیم. داده افزایی انجام شده در این پروژه، orizontalFlip، VerticalFlip و RandomRotate90. در شکل زیر، نمونه ی یکی از دیتا آگمنتیشن ها را مشاهده می کنید:



شکل ۹- data augmentation in this study

در این شکل، نتیجه ی آگمنت شده ی یک تصویر با VerticalFlip را مشاهده می کنید. تصاویر بالا تصاویر پیش از آگمنت شدن است و تصاویر پایین، آگمنت شده ی دقیقاً تصاویر بالاست. نمونه ی دیگری را مشاهده می کنید پس از RandomRotate90:



شکل ۱۰- data augmentation in this study --

۴-۱) بهینه ساز، متریک ها، و تابع هزینه

در این بخش دو شاخص IoU و Dice Coefficient معرفی می شوند:

یکی از معیارهای رایج مورد استفاده در ارزیابی عملکرد مدل های تقسیم بندی، به ویژه در وظایف بینایی کامپیوتر مانند تشخیص اشیا و segmentation تصویر است. این شاخص، همپوشانی بین دو منطقه یا مجموعه را اندازه گیری می کند:

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{TP}{TP + FP + FN} = \frac{|A \cap B|}{|A \cup B|}$$

شاخص IoU را نمی توان به عنوان تابع هزینه برای آموزش شبکه استفاده کرد زیرا که loss در آموزش شبکه که برای محاسبه ی گرادیان استفاده می شود، و در حالتی که bounding box از تارگت اصلی فاصله دارد و overlap ندارند (که در طی آموزش بسیار رخ می دهد) مقدار IoU برابر با صفر شده فارغ از اینکه box پیش بینی شده به مربع هدف (مثلا بافت سرطانی) نزدیک شده است یا خیر، لذا گرادیان مطلوب برای آموزش شبکه و به روز رسانی وزن های شبکه بدست نمی آید. البته برای آموزش شبکه با IoU، ورژن های جدید این شاخص را می توان استفاده کرد. از محدودیت های دیگر این شاخص به حساسیت آن به بالانس نبودن کلاس ها اشاره می کنیم.

معیار دیگر، شاخص Dice Coefficient، معیار دیگری است که مثل IoU برای اندازه گیری شباهت یا همپوشانی بین دو نمونه، مجموعه یا تصویر استفاده می شود. این شاخص همسانی بین دو مجموعه یا منطقه را با اندازه گیری همپوشانی بین آنها به صورت کمی نشان می دهد و ارزیابی می کند که یک ماسک تقسیم بندی پیش بینی شده چقدر با ماسک حقیقتی همسو است. ضریب Dice بالاتر (نزدیک تر به ۱) نشان دهنده همپوشانی بهتر و در نتیجه دقت تقسیم بندی بهتر است. ضعف این شاخص این است که با موارد مثبت کاذب (اینجا: فرد/بافت سرطانی که به عنوان بافت سالم برچسب خورده) و منفی کاذب (اینجا: فرد/بافت سالمی که با تومور مغزی برچسب خورده) به طور یکسان رفتار می کند، که ممکن است بسته به کاربرد همیشه برای استفاده و مقایسه مطلوب نباشد. محدودیت دیگر این شاخص این است که در اصل برای تقسیم بندی باینری تعریف شده و برای چندکلاسه نیازمند مقداری اصلاح است.

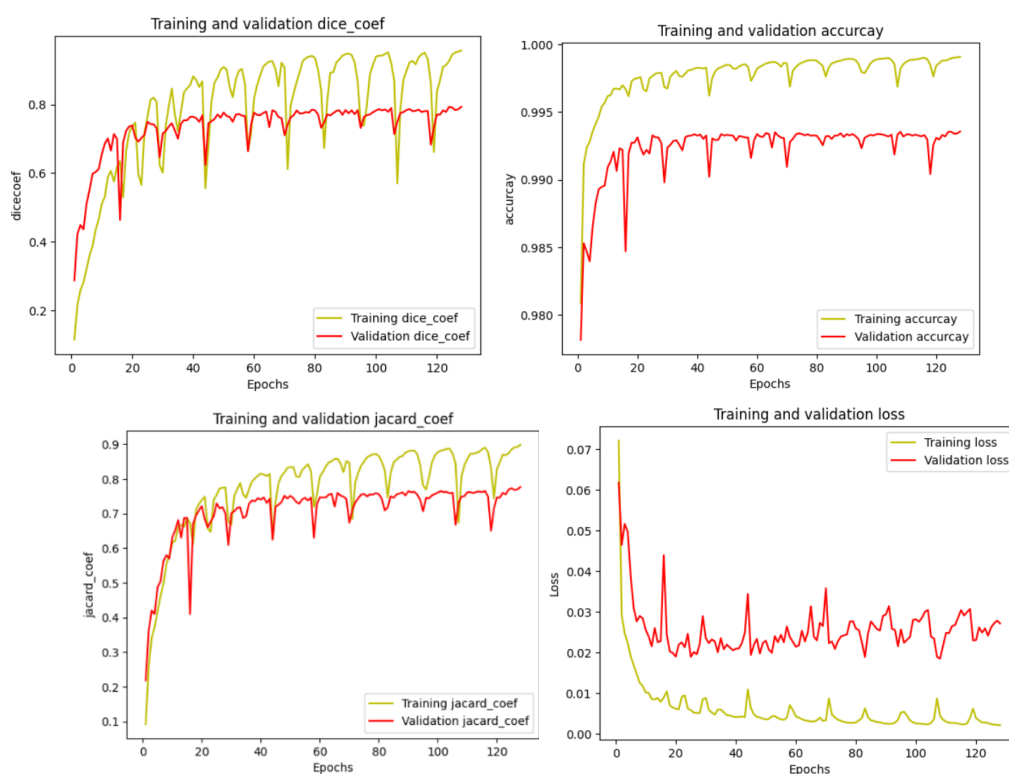
$$\text{Dice Coefficient} = \frac{2TP}{2TP + FP + FN} = \frac{2|A \cap B|}{|A| + |B|}$$

هردوی این شاخص ها در شبکه ی این سوال پیاده شد و نتیجه ی عملکرد، تکامل و مقایسه را در بخش های بعد بدان اشاره کرده ایم.

۵-۱) آموزش مدل

در این بخش پس از تقسیم داده ها به ۳ گروه آموزش، اعتبار سنجی و تست به نسبت ۸-۱-۱ اقدام به آموزش مدل کردیم. بهینه ساز استفاده شده در این فرآیند، آدام بوده و تابع هزینه استفاده شده binary cross entropy است. در فرآیند آموزش، در هر اپاک ۳ شاخص اصلی (accuracy-loss-IoU- Dice Coef) را محاسبه کرده و نهایتاً تکامل آنها را به تصویر کشیدیم. برای آموزش، چندین مدل ترکیب های متفاوت batch size و epoch را امتحان کردیم. تعدادی از گزارش مدل های اجرا شده به همراه هایپرپارامتر های مربوطه را در زیر مشاهده می کنید.

- گزارش مدل با batch_size=16, epochs=100 آموزش دیده با عکس های ۱۲۸×۱۲۸:

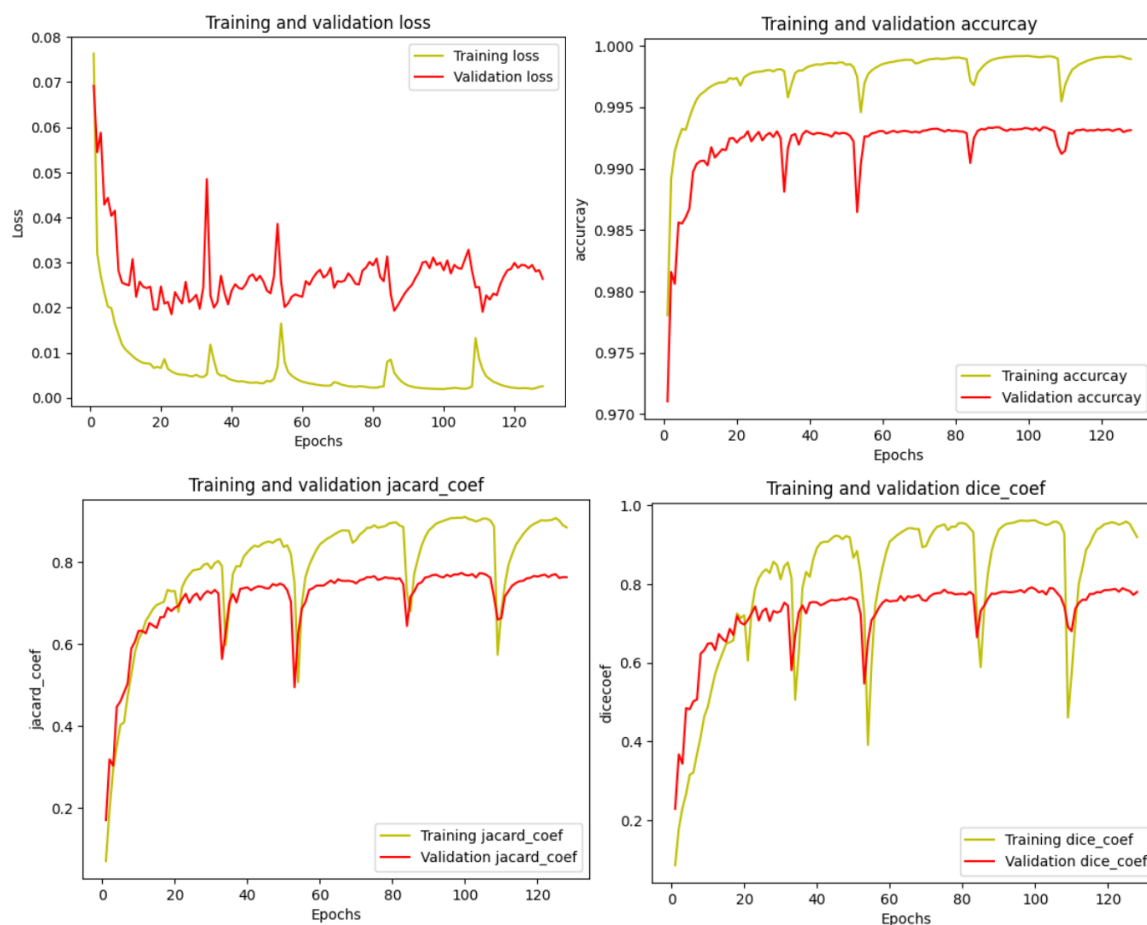


شکل ۱۱- 128*128- batch_size=16, epochs=100

نمودار loss این مدل عملکرد خوبی از خود نشان میدهد. نمودار accuracy، مقداری اختلاف در داده ی آموزش و اعتبارسنجی از خود نشان میدهد. البته این اختلاف در مقادیر بسیار پایین است. نمودار IoU و Dice Coef هر دو عملکرد خوبی نشان می دهند. اما نگرانی ای که در این مدل وجود دارد، احتمال رخداد Overfitting است. برای جلوگیری از این موضوع ۳ تغییر در مدل های بعدی دادیم : (۱) افزایش batch size (۲) کاهش epoch (۳) استفاده از drop out. در ادامه گزارش عملکرد مدل پس از تک تک این تغییرات ارائه شده است:

افزایش batch size:

- گزارش مدل با batch_size=32, epochs=100 آموزش دیده با عکس های 128×128 :

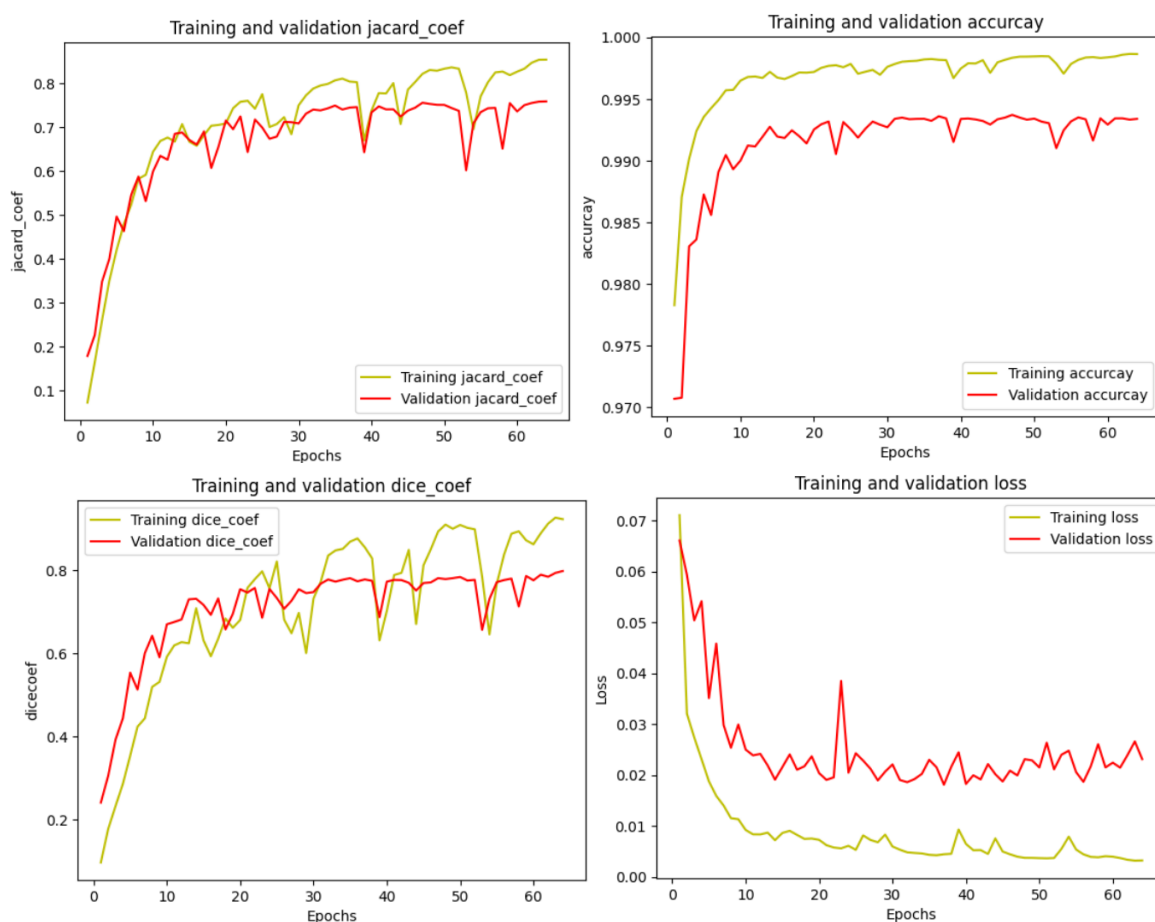


شکل ۱۲-۱ batch_size=32, epochs=100- 128*128

این مدل نسبت به مدل قبلی batch size بیشتری (۱۶ به ۳۲ افزایش) دارد. عملکرد این مدل در پیش بینی خیلی خوب نیست، مسئله ی overfitting نسبت به حالت قبلی تشدید شده است.

کاهش epoch:

- گزارش مدل با batch_size=16, epochs=64 آموزش دیده با عکس های 128×128 :

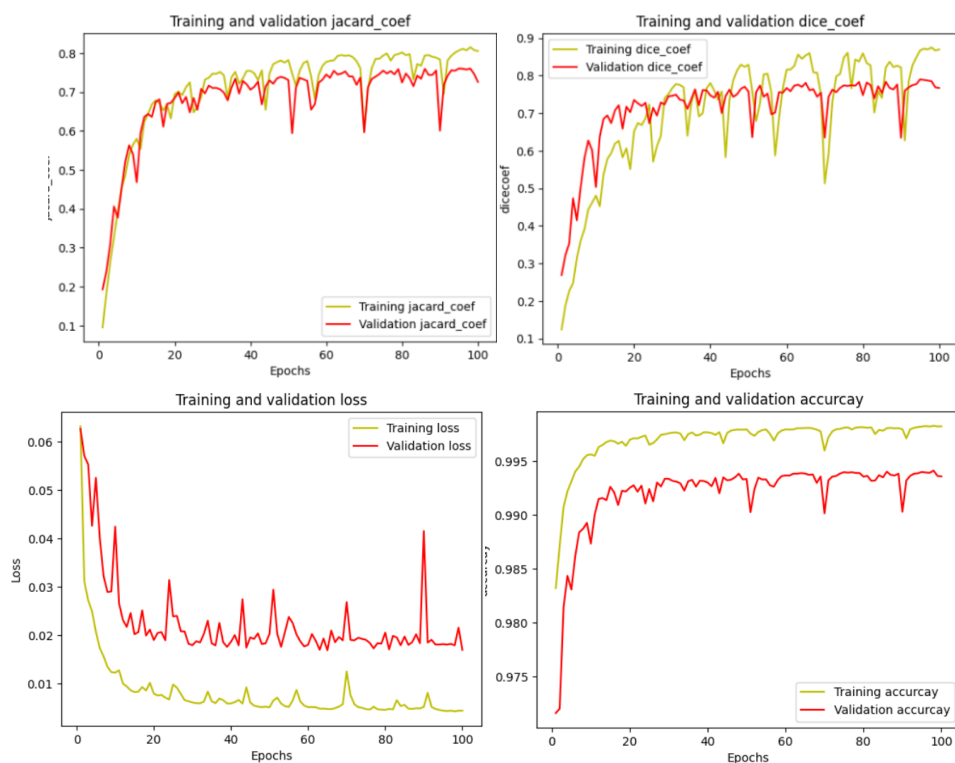


شکل ۱۳- 128*128 batch_size=16, epochs=64

وضعیت overfitting نسبت به حالت قبل، کاهش چشم گیری داشته است. به نظر میرسد که عملکرد این مدل مطلوب است.

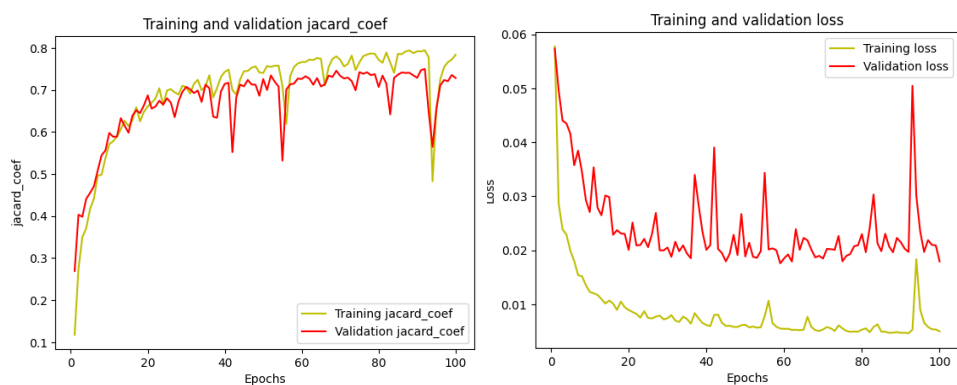
برای بهبود عملکرد مدل و جلوگیری از overfitting تصمیم گرفتیم که لایه های dropout به مدل اضافه کنیم، نتیجه ی چند مدل اجرا شده از این قرار بود:

- گزارش مدل با $\text{batch_size}=16, \text{epochs}=100$ آموزش دیده با عکس های 128×128 با $\text{dropout}=0.1$:



شکل ۱۴- شکل ۱۳ $\text{batch_size}=16, \text{epochs}=100 - \text{dropout}=0.1 - 128 \times 128$

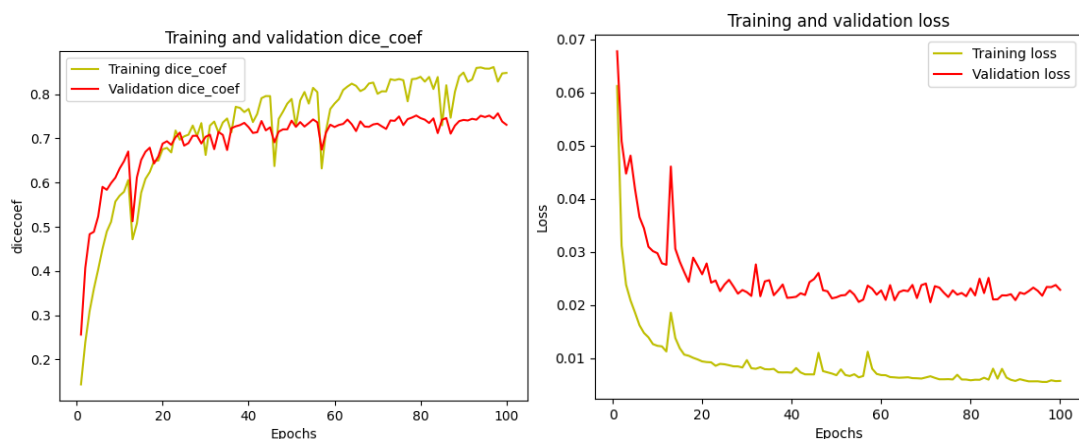
- گزارش مدل با $\text{batch_size}=16, \text{epochs}=100$ آموزش دیده با عکس های 128×128 با $\text{dropout}=0.2$:



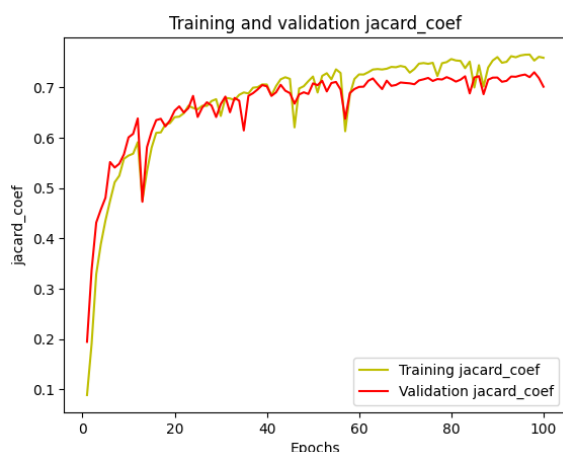
شکل ۱۵- $\text{batch_size}=16, \text{epochs}=100 - 128 \times 128 - \text{dropout}=0.2$

استفاده از dropout مقداری باعث کاهش overfit شده اما باعث نوسانات زیادی در فرآیند آموزش می شود.

- گزارش مدل با $\text{batch_size}=16, \text{epochs}=100$ آموزش دیده با عکس های 64×64 با $\text{dropout}=0.2$:



شکل ۱۶ - $\text{batch_size}=16, \text{epochs}=100$ - 64×64 - $\text{dropout}=0.2$



در دیتای سایز 64×64 ، مدل بنا بر شاخص IoU از خود overfit کمی نشان می دهد. اما دقت اش از مدل آموزش دیده با عکس های 128×128 کمتر است.

همانطور که از نمودار های بالا بر می آید، dropout در این مسئله با این دیتا خیلی باعث بهبود عملکرد مدل نشده است.

در این بخش تاثیر هایپرپارامتر ها بر عملکرد مدل را بررسی کردیم. بهترین مدل تا به این لحظه مدل آموزش دیده شده بر روی دیتای 128×128 با 128 اپاک با $\text{batch size}=16$ بوده است. این بهترین مدل برای پیش بینی مرحله ی بعد استفاده می شود.

۶-۱) ارزیابی مدل

پس از آموزش مدل های متعدد، برترین مدل (مدل ۵ فایل کد) را انتخاب کردیم و برای پیش بینی به کار بردیم. عملکرد مدل بر داده ی آموزش را در زیر مشاهده می کنید:

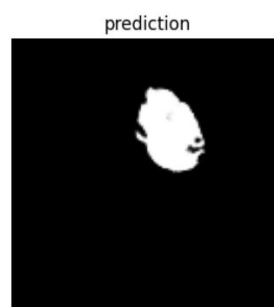
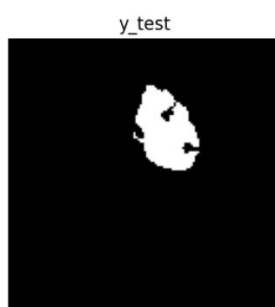
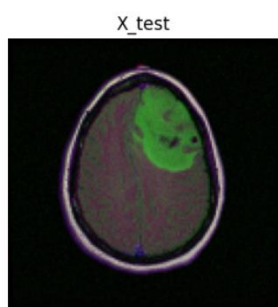
جدول ۱- عملکرد مدل بر داده ی آموزش

METRIC	PERFORMANCE ON TRAIN DATA
Loss	0.0022
Accuracy	0.9991
dice coef	0.9568
jacard (IoU)	0.8986

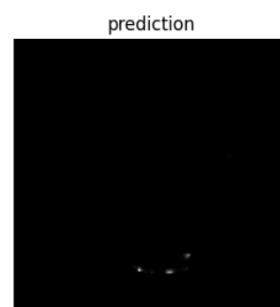
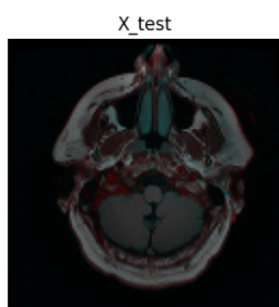
جدول ۲- عملکرد مدل بر داده ی اعتبارسنجی

METRIC	PERFORMANCE ON VALIDATION DATA
Loss	0.0272
Accuracy	0.9936
dice coef	0.7930
jacard (IoU)	0.7766

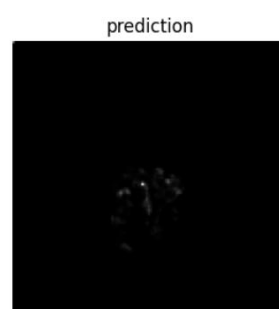
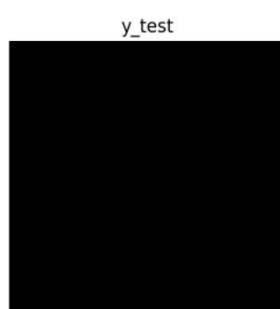
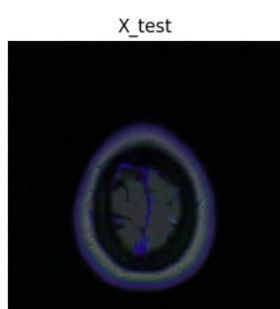
نتیجه ی پیش بینی ۵ کیس به صورت تصادفی انتخاب شده از داده های تست را در زیر مشاهده می کنید:



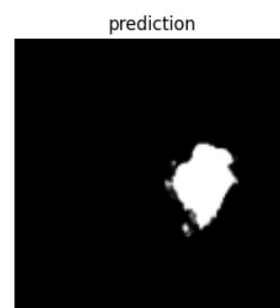
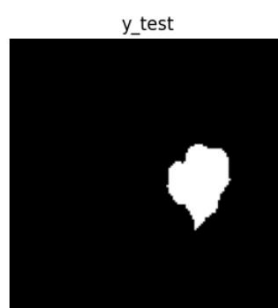
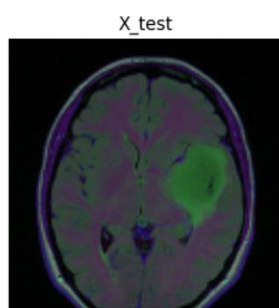
شکل ۱۷ - Prediction ۱



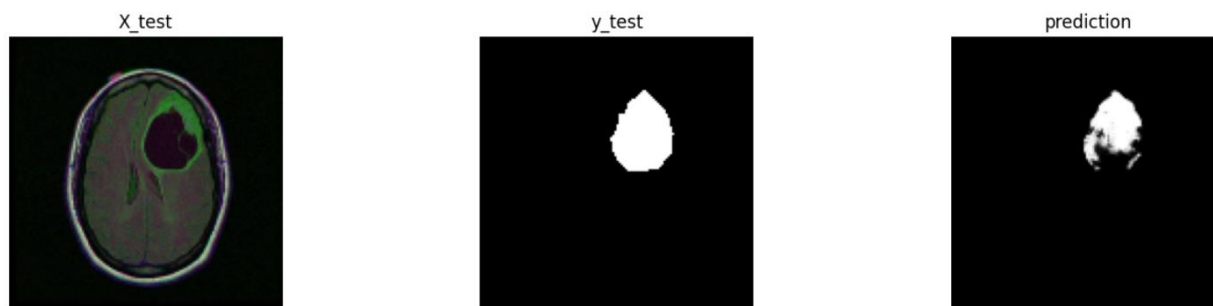
شکل ۱۸ - Prediction 2



شکل ۱۹ - Prediction 3



شکل ۲۰ - Prediction 4



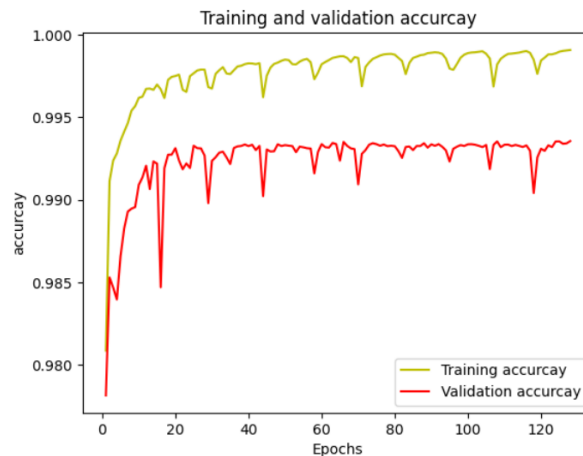
شکل ۵۲۱ - Prediction

همانطور که از عکس ها پیداست، مدل عملکرد بسیار خوبی از خود در تشخیص بافت های سرطانی از غیر سرطانی داشته است. به طور مثال در عکس اول، بافت سرطانی به خوبی پیش بینی شده است و در عکس دوم که بافت مغز سالم بوده و بافت سرطانی ای وجود نداشته، مدل هیچ گونه بافت سرطانی ای هم پیش بینی نکرده است. این عملکرد خوب، خود را در شاخص های بالای مقایسه ای IoU , accuracy و Dice coef هم نشان داده است:

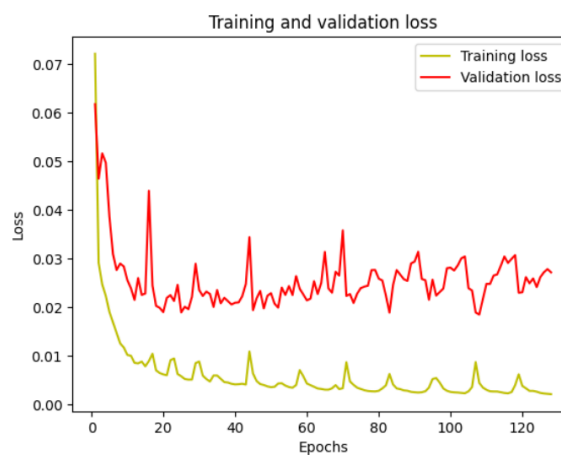
جدول ۳- عملکرد مدل بر داده ی تست

METRIC	PERFORMANCE ON TEST DATA
Loss	0.011173075996339321
Accuracy	0.9974997639656067
dice coef	0.8621870875358582
jacard (IoU)	0.7719519734382629
Mean IoU	0.9 3896

این مدل پس از بررسی مدل های متعدد با هایپرپارامتر های متعدد بدست آمده است. مسئله ی بیش برازش در این مدل نسبت به مدل های دیگر پس از بررسی اثر batch size, epoch و استفاده و عدم استفاده از dropout بدست آمد. قابلیت تعمیم پذیری این مدل بسیار خوب است و در ۵ نمونه ی تصادفی خطای بسیار بسیار کمی از خود نشان داد(تنها در یک مورد که بخشی از بافت سرطانی پیش بینی نشد- شکل ۲۱) و این بسیار خوب است.



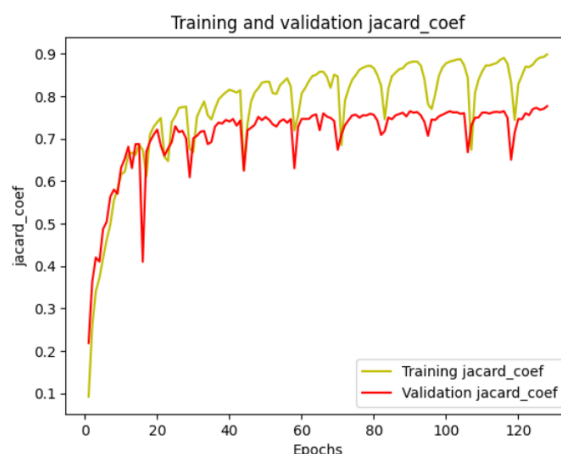
در نمودار accuracy، ابتدا شاهد سیر صعودی accuracy و بعد ثابت شدن آن هستیم. می توان گفت در تک تک تکرار ها، accuracy بالای ۹۸٪ است. اما این شاخص خیلی شاخص قابل اطمینانی برای بررسی عملکرد شبکه ی U-Net نیست. برای بررسی عملکرد شبکه، بهتر است از دو شاخص IoU و Dice Coef کمک بگیریم.



در این نمودار سیر تکاملی خطا در طی ایپاک ها، مشخص است که مدل دارد از داده ها یادمیگیرد و سیر کلی خطا، کاهشی است. ما شاهد یک نرخ کاهشی آهسته و پیوسته هستیم.



در مورد Dice coef، ابتدا شاهد نرخ افزایشی تند در ابتدای تکرار ها و بعد نسبتاً ثابت شدن مقدار شاخص هستیم. البته شاهد مقداری نوسان هستیم اما نوسانات کوچک. این نوسانات می توانند نشان از خطر بیش برآزش باشند. که البته ما راهکار های کاهش این خطر را امتحان کردیم و این مدل، بهترین نتیجه بود.



در این نمودار شاخص IoU، مشاهده می شود که پیوسته این شاخص افزایش یافته است، اگرچه مقداری نوسان دارد. همچنین به این اشاره می کنیم که اینکه چه مقدار ضریب ژاکارد ایده ال است بسته نوع مسئله و دیتای استفاده شده دارد اما به طور کلی، این مقدار بدست آمده در این مسئله مطلوب است.

