

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



درس یادگیری عمیق با کاربرد در بینایی ماشین و پردازش صوت

تمرین شماره ۳

نام و نام خانوادگی : مهیار ملکی

شماره دانشجویی : ۸۱۰۱۰۰۴۷۶

خرداد ماه ۱۴۰۱

۳	فهرست اشکال و جداول
۴	مقدمه
۵	سوال اول : تفکیک بافت توموری به وسیله شبکه کانولوشنی
۵	بخش الف – چرایی عدم استفاده از Padding
۵	بخش ب – آموزش مدل U-Net
۹	بخش ج – تأثیر تقویت دادگان
۱۱	بخش د – تغییر به مدل باقیمانده
۱۴	سوال دوم : تشخیص اادات سخن با استفاده از شبکه های بازگشتی
۱۴	بخش الف – پیش پردازش
۱۵	بخش ب – تبدیل متن به بردار
۱۵	بخش پ – پیاده سازی و ارزیابی شبکه ی بازگشتی
۱۷	بخش ت- انواع شبکه های بازگشتی
۱۹	بخش ث – شبکه دو طرفه

فهرست اشکال و جداول

شکل ۱- ساختار شبکه Unet کتابخانه پایتورچ.....	۷
شکل ۲- نمودار خطا مدل Unet ساده.....	۷
شکل ۳- معیارهای ارزیابی مدل Unet ساده.....	۸
شکل ۴- تصاویر خروجی مدل Unet ساده.....	۸
شکل ۵- نمونه تصاویر augment شده.....	۹
شکل ۶- مقایسه نمودار خطا قبل و بعد از تقویت دادگان.....	۱۰
شکل ۷- مقایسه معیارهای ارزیابی قبل و بعد از تقویت دادگان.....	۱۰
شکل ۸- تصاویر خروجی مدل Unet با تقویت دادگان.....	۱۱
شکل ۹- مقایسه نمودار خطا قبل و بعد از اضافه کردن ساختار باقی‌مانده.....	۱۲
شکل ۱۰- مقایسه معیارهای ارزیابی قبل و بعد از اضافه کردن ساختار باقی‌مانده.....	۱۲
شکل ۱۱- تصاویر خروجی شبکه Unet با اعمال تقویت دادگان و ساختار باقی‌مانده.....	۱۲
شکل ۱۲- مقایسه مقادیر (dice و jaccard) آزمون داده‌های ارزیابی روی مدل‌های مختلف.....	۱۳
شکل ۱۳- مقایسه خطای داده‌های ارزیابی مدل‌های مختلف.....	۱۳
شکل ۱۴- نمودار خطای دادگان آموزش و اعتبارسنجی مدل RNN ساده.....	۱۷
شکل ۱۵- معیارهای آزمون دقت شبکه RNN ساده.....	۱۷
شکل ۱۶- مقایسه نمودارهای خطای دادگان آزمون و اعتبارسنجی روی شبکه‌های RNN, LSTM, GRU.....	۱۸
شکل ۱۷- مقایسه معیار دقت دادگان آموزش و اعتبارسنجی در شبکه‌های RNN, LSTM, GRU.....	۱۸
شکل ۱۸- مقایسه معیار F1 micro دادگان آموزش و اعتبارسنجی در شبکه‌های RNN, LSTM, GRU.....	۱۸
شکل ۱۹- مقایسه معیار F1 micro دادگان آموزش و اعتبارسنجی در شبکه‌های RNN, LSTM, GRU.....	۱۸
شکل ۲۰- مقایسه نمودارهای خطای دادگان آزمون و اعتبارسنجی شبکه‌های دوطرفه و یک طرفه GRU.....	۱۹
شکل ۲۱- مقایسه معیار دقت دادگان آزمون و اعتبارسنجی شبکه‌های دوطرفه و یک طرفه GRU.....	۱۹
شکل ۲۲- مقایسه معیار F1 micro دادگان آزمون و اعتبارسنجی شبکه‌های دوطرفه و یک طرفه GRU.....	۲۰
شکل ۲۳- مقایسه معیار F1 macro دادگان آزمون و اعتبارسنجی شبکه‌های دوطرفه و یک طرفه GRU.....	۲۰
شکل ۲۴- مقایسه خطای ساختارهای مختلف RNN روی دادگان ارزیابی.....	۲۱
شکل ۲۵- مقایسه معیارهای دقت ساختارهای مختلف RNN روی دادگان ارزیابی.....	۲۱
جدول ۱- هاپرپارامترهای مورد استفاده برای آموزش شبکه Unet.....	۶
جدول ۲- تعداد داده‌های آموزش و اعتبارسنجی و ارزیابی.....	۱۴
جدول ۳- ساختار شبکه RNN ساده.....	۱۵
جدول ۴- هاپرپارامترهای شبکه RNN ساده.....	۱۶

این تمرین دو بخش متفاوت را شامل می‌شود. در بخش اول کاربرد شبکه های عصبی عمیق در تقطیع^۱ تصاویر و در بخش دوم اهمیت شبکه های عصبی بازگشتی^۲ در مسائلی که در ورودی یا خروجی دنباله داریم و ترتیب اهمیت پیدا می‌کند، بررسی می‌شوند.

مسائل تقطیع کاربردهای صنعتی و تجاری زیادی در خدمات درمانی و تقطیع عکس‌های پزشکی دارند. در سوال اول برای شناسایی تومورهای مغزی، با استفاده از مدل U-Net به تقطیع تصاویر مغزی^۳ و تفکیک بافت تومور می‌پردازیم. همچنین در سوال دوم با استفاده از مدل‌های بازگشتی به مسئله تشخیص ادات^۴ سخن می‌پردازیم.

توضیحات کد:

سوال اول شامل یک فایل پایتون model و سه فایل main برای هر کدام از بخش های سوال می‌باشد. همچنین سوال دوم نیز یک فایل model و چهار فایل main برای هر یک از ساختار ها خواسته شده در سوال را دارا می‌باشد.

برای هر دو سوال با دادن مقادیر دلخواه ورودی به آرگومان WholeDataset می‌توان تعداد ورودی ها را کاهش داد و کد را فقط بروی قسمت کوچکی از داده‌گان اجرا کرد، همچنین با دادن مقدار True کل داده ها به عنوان ورودی در نظر گرفته خواهند شد.

^۱ Segmentation

^۲ Recurrent neural networks

^۳ MRI

^۴ POS tagging

سوال اول : تفکیک بافت توموری به وسیله شبکه کانولوشنی

بخش الف – چرایی عدم استفاده از PADDING

استفاده از padding های متوالی هر یک از مراحل یک شبکه عصبی باعث به وجود آمدن ناحیه هایی در حاشیه های تصویر می شود که به اندازه تصویر ورودی شامل اطلاعات نمی شوند، لذا خروجی شبکه در این نقاط دقت کمتری دارد.

اهمیت این موضوع در کاربردهای پزشکی این می باشد که معمولا تصاویر پزشکی دارای اندازه های بزرگی هستند و تقطیع یک جای آنها توسط شبکه های عمیق و سخت افزار های موجود مشکل می باشد. لذا این تصاویر را به تصاویر کوچکتری تقسیم کرده و به شبکه می دهند و در نهایت خروجی ها را مجددا به هم متصل می کنند. نواحی حاشیه ای که در قسمت قبل گفته شد، در اتصال مجدد این تصاویر مشکل ساز خواهند بود.

بخش ب – آموزش مدل U-NET

در این بخش، شبکه U-net معرفی شده در کتابخانه پایتورچ^۱ پیاده سازی شد. ساختار لایه های این شبکه در شکل ۱ و هایپر پارامترهای مورد استفاده برای آموزش شبکه در جدول ۱ قابل مشاهده است.

در مقایسه این شبکه با شبکه پیاده شده در مقاله اصلی به تفاوت های زیر می توان اشاره کرد:

- اندازه تصاویر ورودی و خروجی:
- در مقاله یونت اصلی اندازه تصاویر ورودی ۵۷۲ و خروجی ۳۸۸ می باشد، ولی در مدل یونت پایتورچ تصاویر ورودی و خروجی هم اندازه و ۲۵۶ در ۲۵۶ پیکسل هستند.
- اضافه کردن پیکسل به مرز تصاویر^۲:
- در مدل اصلی به مرز تصاویر پیکسلی اضافه نمی شود لذا با اعمال متوالی عملیات pooling اندازه تصویر در شبکه کاهش می یابد، ولی در مدل پایتورچ پدینگ صورت گرفته و اندازه تصاویر حفظ می شود.

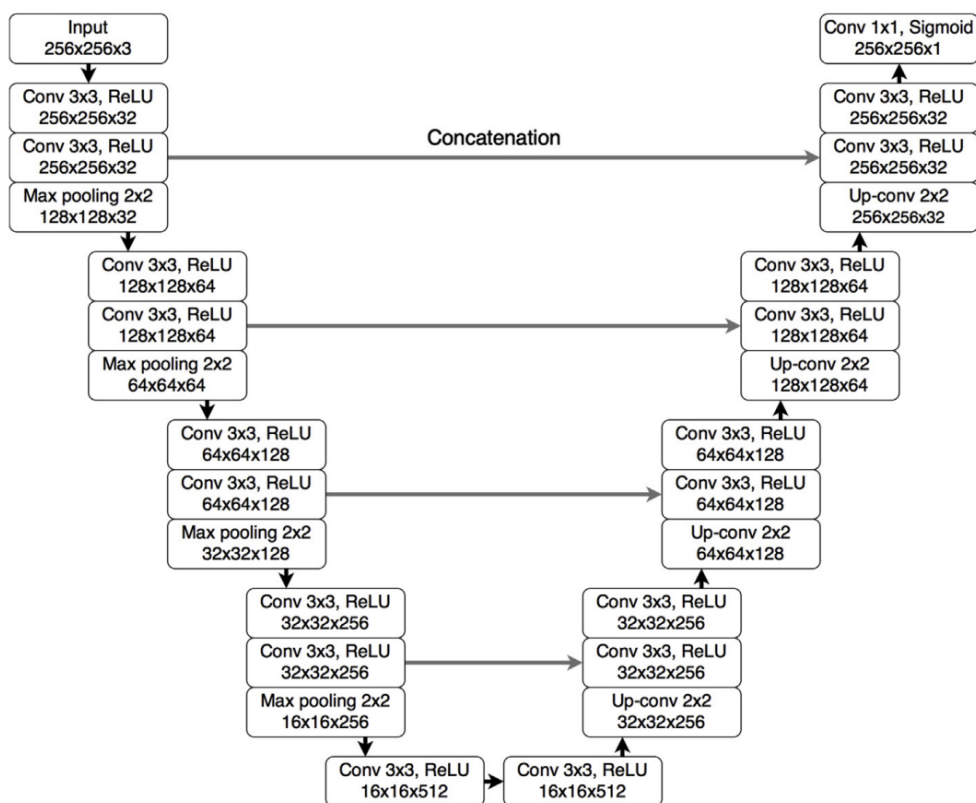
^۱ Pytorch
^۲ Padding

- یکسان سازی دسته ای^۱:
در مدل اصلی از این روش استفاده نمی‌شود، ولی در مدل پایتورچ در هر بلاک عملیات یکسان سازی دسته‌ای انجام می‌شود.
- توابع سافتمکس^۲ و سیگموید^۳:
در مدل اصلی پس از آخرین لایه یک تابع سافتمکس اعمال می‌شود ولی در مدل پایتورچ به جای آن از یک تابع سیگموید استفاده می‌شود.

جدول ۱- هایپرپارامترهای مورد استفاده برای آموزش شبکه Unet

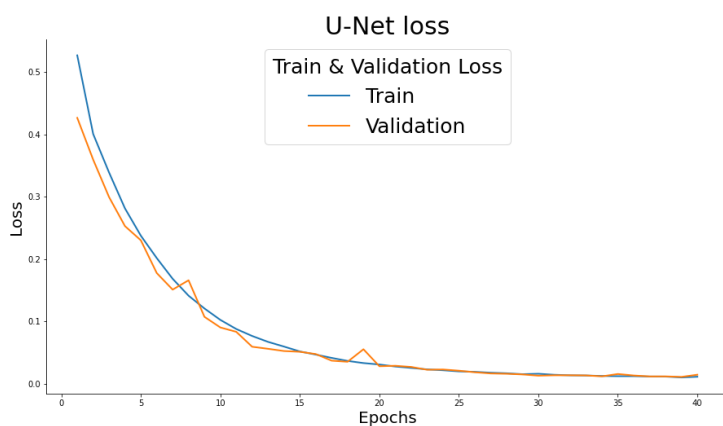
• Batch Size	16
• Epochs	40
• Optimizer	ADAM
• Learning Rate	0.0001
• Weight Decay	0.001
• Loss Function	Binary cross entropy
• Scheduler	LR decreases by factor of 0.1 when LR stopped improving for 2 epochs

^۱ Batch normalization
^۲ Softmax
^۳ Sigmoid



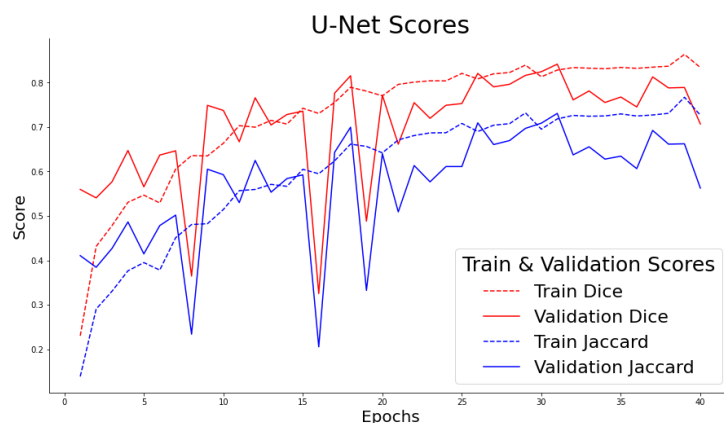
شکل ۱- ساختار شبکه U-Net کتابخانه پایتون

نتایج حاصل از آموزش شبکه در شکل‌های ۲ و ۳ قابل مشاهده هستند. با توجه به نمودار خطی^۱ در شکل ۲ مشاهده می‌شود که دو منحنی آموزش و اعتبارسنجی پس از ۴۰ دوره به خوبی همگرا شده‌اند و دچار فرارازش^۲ نیز نشده‌ایم. همچنین چنانچه در شکل ۳ قابل مشاهده است، معیارهای Dice و Jaccard پس از ۴۰ دوره به ترتیب به مقادیر تقریبی ۰.۸ و ۰.۷ رسیده‌اند. البته برای منحنی‌های اعتبارسنجی نویز زیادی مشاهده می‌شود.



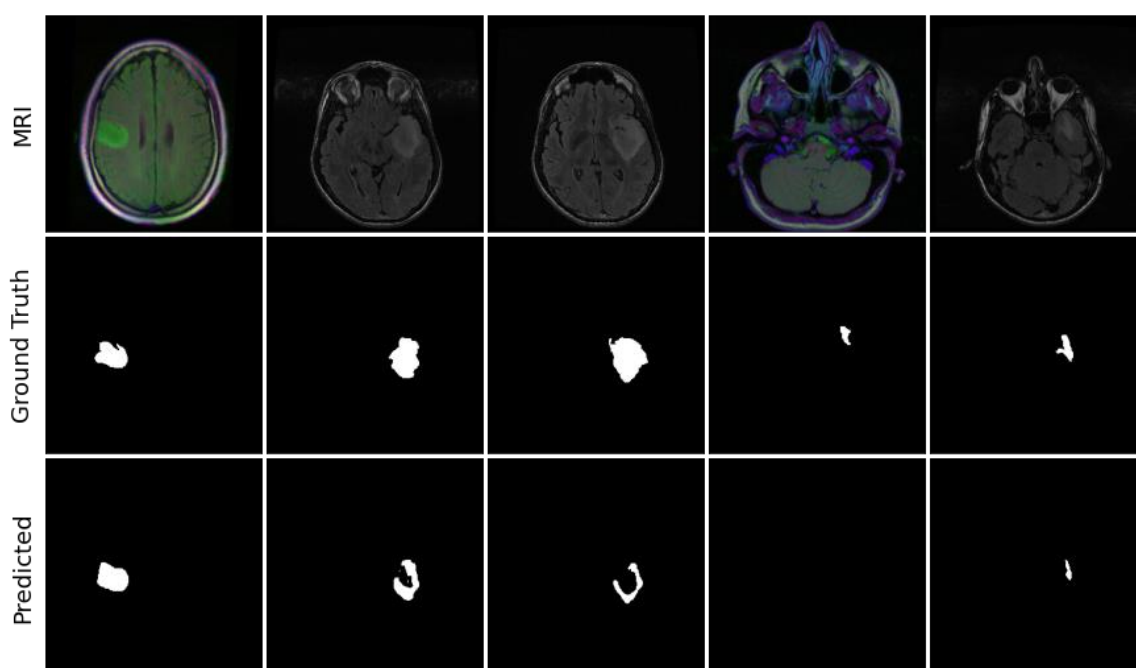
شکل ۲- نمودار خطی مدل U-Net ساده

^۱ Loss
^۲ Overfit



شکل ۳- معیارهای ارزیابی مدل U-Net ساده

در شکل ۴ خروجی مدل روی تعدادی از دادگان ارزیابی مشاهده می‌شود. به نظر می‌رسد وجود بافت‌های تومور با دقت نسبتاً خوبی پیش‌بینی شده است، اما در مواردی تمام بافت تومور به خوبی جدا نشده‌اند.

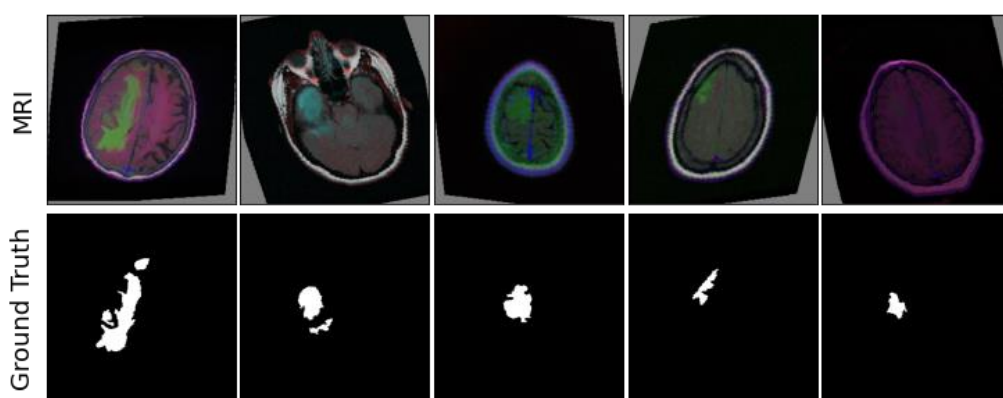


شکل ۴- تصاویر خروجی مدل U-Net ساده

بخش ج – تأثیر تقویت دادگان

روش های تقویت دادگان یکی از راه های تنظیم کردن^۱ مدل می باشد که با افزایش داده ها این امر را امکان پذیر می کند. در این بخش با استفاده از روش های تقویت دادگان زیر به بهبود مدل می پردازیم:

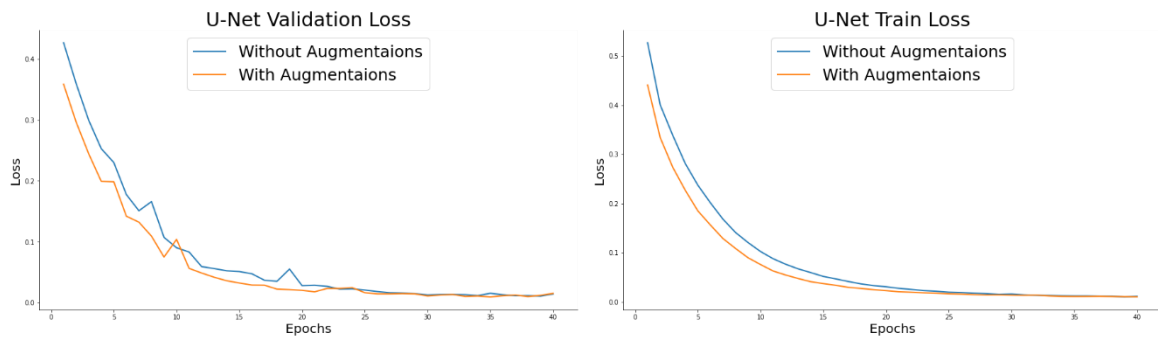
- قرینه کردن افقی^۲: ساختار چپ و راست مغز تا حد زیادی قرینه هستند، لذا می توان از این روش برای تقویت دادگان استفاده کرد تا شبکه، جداسازی تومورهایی در یک طرف هستند را به طور قرینه در سمت دیگر نیز یاد بگیرد. البته ممکن است برخی بافت های توموری مختص یک سمت باشند. لازم به ذکر است که این عمل به صورت عمودی امکان پذیر نیست زیرا در راستای عمودی قرینگی وجود ندارد.
- چرخش^۳: همانند قسمت قبل چرخش نیز می تواند تشخیص بافت های توموری توسط مدل را در نقاط مختلف مغز گسترش دهد.
- تغییر مقیاس^۴: از آنجایی که بافت های توموری از نظر اندازه بسیار متفاوت هستند لذا این تغییر مقیاس نیز می تواند در اینجا مفید واقع شود.
- تبدیل برشی^۵: در برخی از تصاویر MRI ممکن است که سر تکان خورده و تصاویر کمی کج شوند، لذا این تبدیل می تواند مفید باشد.



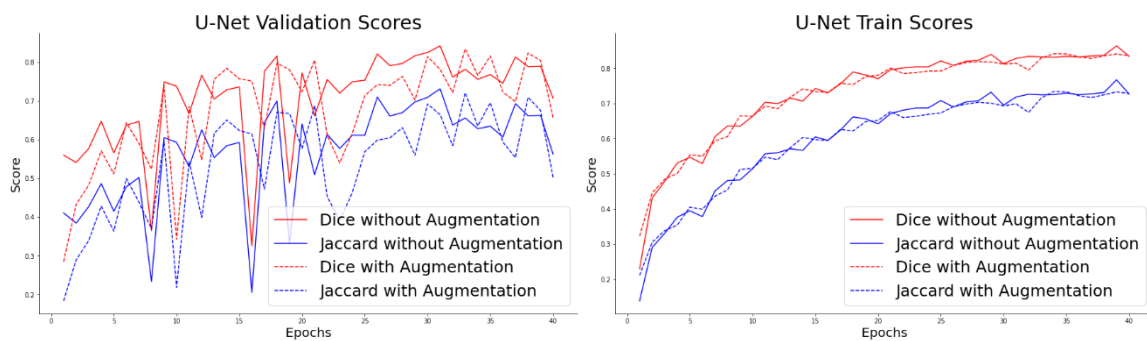
شکل ۵- نمونه تصاویر augment شده

^۱ Regularization
^۲ Horizontal flip
^۳ Rotation
^۴ Scaling
^۵ Shearing

نتایج : چنانچه در شکل ۶ قابل مشاهده است، با تقویت دادگان مدل ما با سرعت کمی بیشتر همگرا شده است. همچنین در شکل ۷ قابل مشاهده است که معیارهای ارزیابی Dice و Jaccard تفاوت چندانی با حالت قبل از تقویت دادگان نکرده اند.

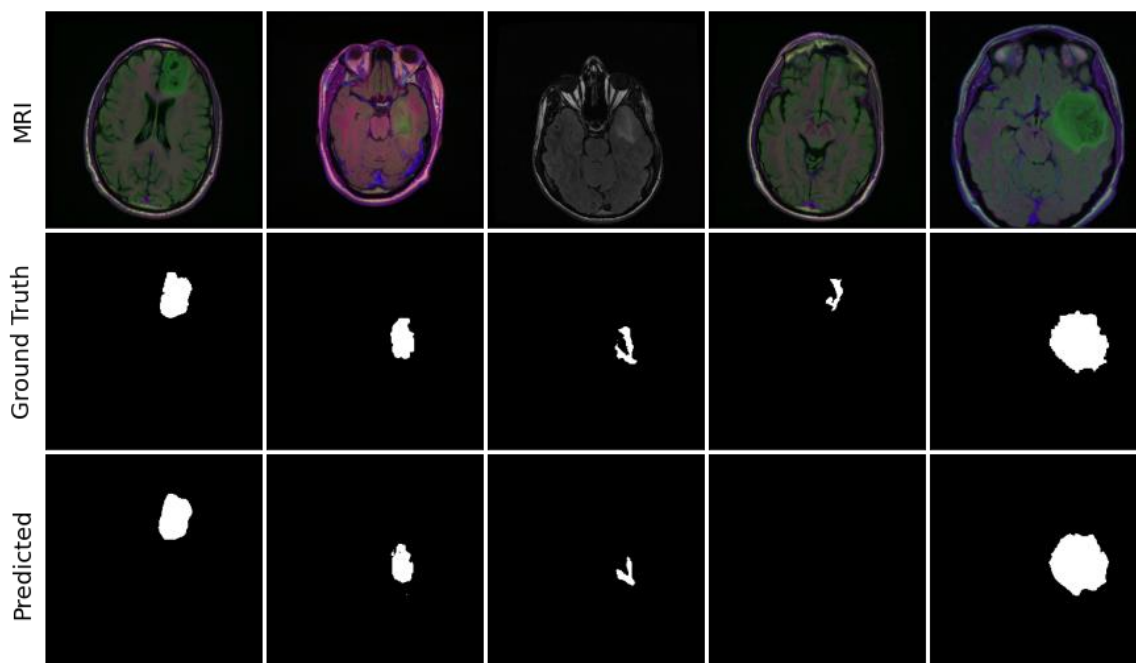


شکل ۶- مقایسه نمودار خطا قبل و بعد از تقویت دادگان



شکل ۷- مقایسه معیارهای ارزیابی قبل و بعد از تقویت دادگان

در شکل ۸ همانند قبل نتایج پیشبینی شبکه روی پنج عدد از تصاویر دادگان ارزیابی آورده شده است. همانطور که قابل مشاهده است به جز یک مورد تمام بافت های مغزی به خوبی تشخیص داده شده اند.



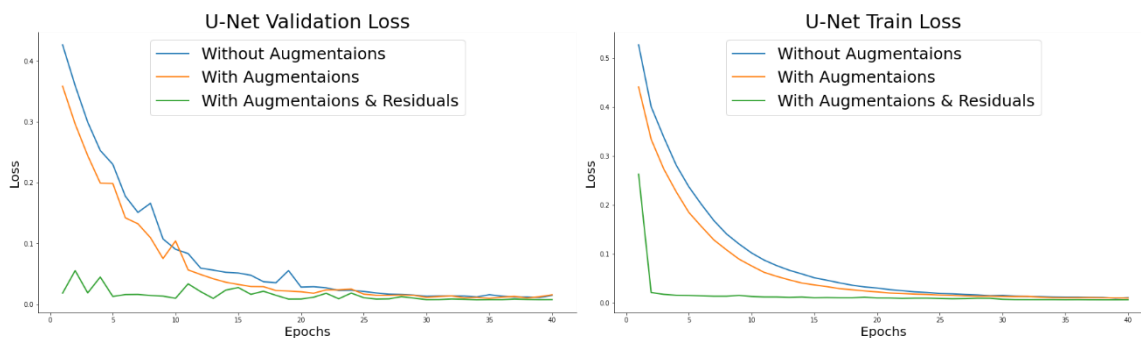
شکل ۸- تصاویر خروجی مدل Unet با تقویت دادگان

بخش د – تغییر به مدل باقیمانده

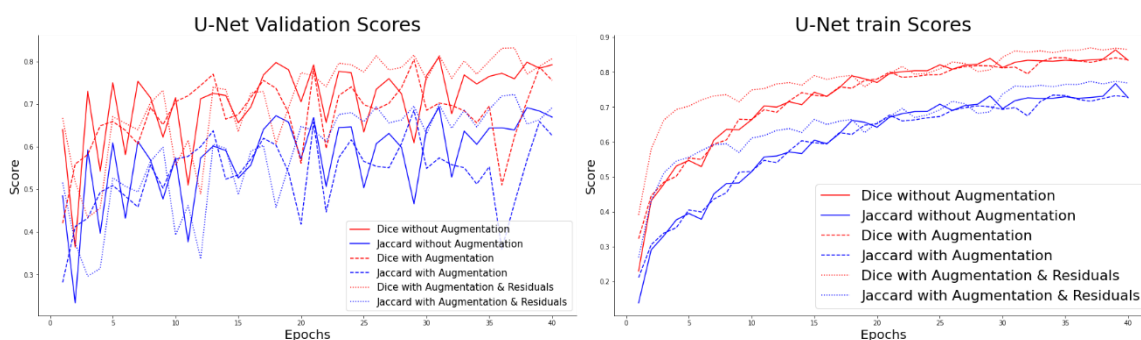
در این بخش با اضافه کردن ساختار باقی‌مانده^۱ به شبکه، تلاش می‌کنیم تا عملکرد آن را بهبود دهیم. این ساختارها به بهبود جریان گرادیان در شبکه کمک کرده و در اینجا باعث شده‌اند که شبکه با سرعت بیشتری به سمت نقطه بهینه همگرا شود. این افزایش سرعت همگرایی باعث می‌شود تا در اینجا با تعداد دوره‌های یکسان، نسبت به مدل بدون ساختار باقی‌مانده به نتایج بهتری برسیم.

همانطور که گفته شد افزایش سرعت همگرایی آموزش شبکه در شکل ۹ قابل مشاهده است. این افزایش سرعت چشمگیر باعث شده است که بعد از تنها یک اپاک مقدار خطا نزدیک به کمینه برسد. همچنین چنانچه در شکل ۱۰ مشاهده می‌شود، افزایش سرعت باعث شده است تا بعد از ۲۰ دوره به مقادیر بهتری برای معیارهای Dice و Jaccard نسبت به قبل برسیم. با توجه به شکل ۱۱ این بهبود و افزایش مقدار معیارهای ارزیابی، در تصاویر خروجی شبکه بروی دادگان ارزیابی نیز قابل مشاهده است زیرا تمام بافت‌های توموری با دقت خوبی در این نمونه‌ها پیش‌بینی و جدا شده‌اند.

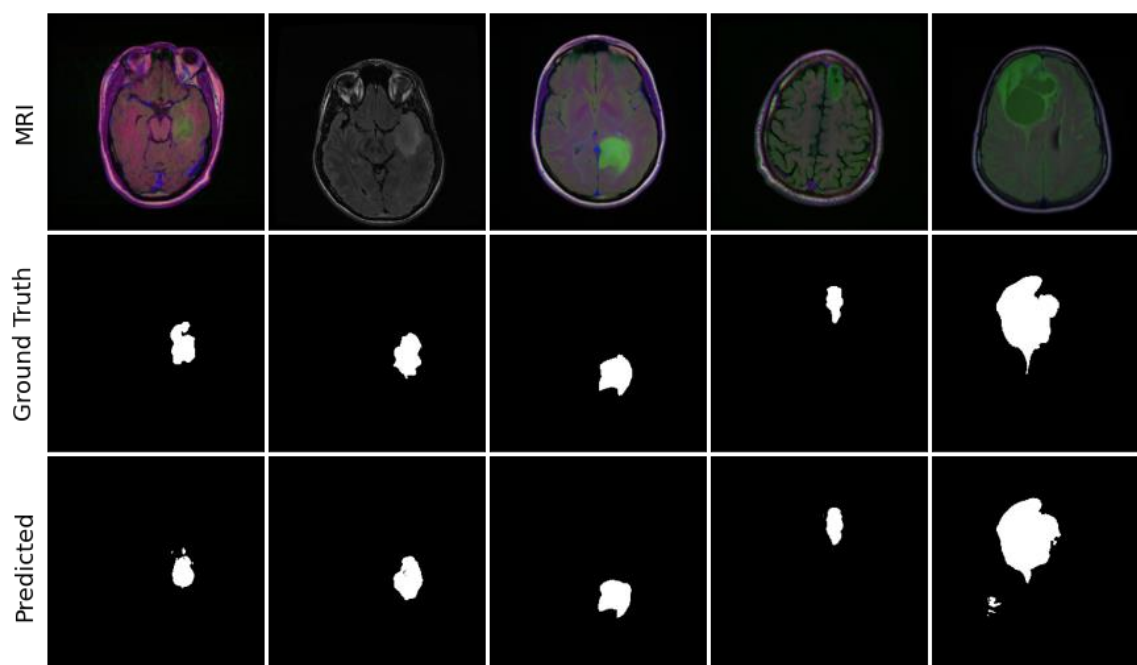
^۱ Residual



شکل ۹- مقایسه نمودار خطا قبل و بعد از اضافه کردن ساختار باقی‌مانده

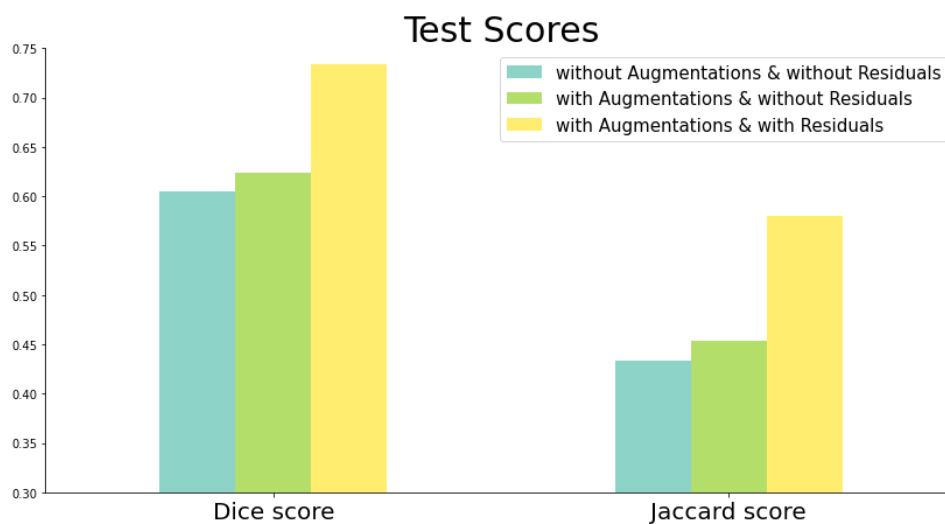


شکل ۱۰- مقایسه معیارهای ارزیابی قبل و بعد از اضافه کردن ساختار باقی‌مانده

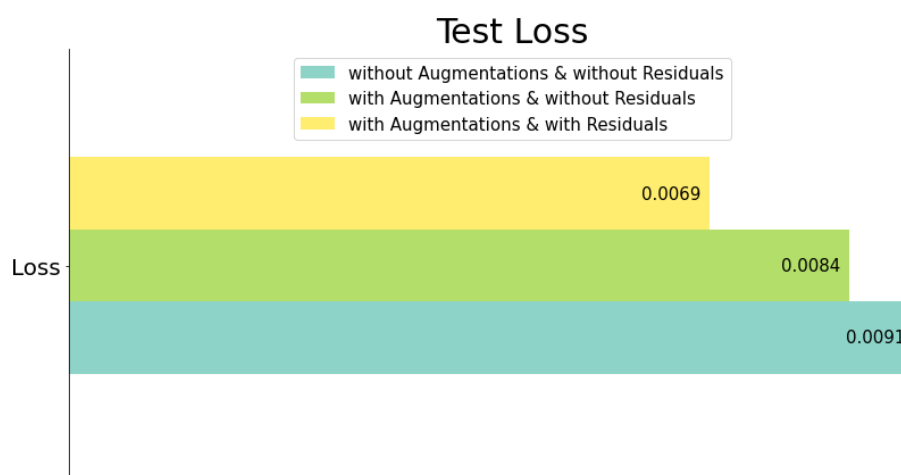


شکل ۱۱- تصاویر خروجی شبکه U-Net با اعمال تقویت دادگان و ساختار باقی‌مانده

در نهایت در شکل های ۱۲ و ۱۳ نتایج نهایی مدل های آموزش دیده، روی دادگان ارزیابی قابل مشاهده است. چنانچه انتظار می رفت نتایج بدست آمده در قسمت های قبلی، به روی دادگان ارزیابی نیز حاصل شده اند.



شکل ۱۲- مقایسه مقادیر (dice و jaccard) آزمون داده های ارزیابی روی مدل های مختلف



شکل ۱۳- مقایسه خطای داده های ارزیابی مدل های مختلف

سوال دوم : تشخیص ادات سخن با استفاده از شبکه های بازگشتی

بخش الف – پیش پردازش

ابتدا پیکره فارسی بیژن خان را بارگزاری کرده و همزمان پیش پردازش های لازم را با استفاده از کتابخانه پارسى نرم و مواردی دیگر اعمال می‌کنیم. لازم به ذکر است که هنگام بارگزاری فایل **Glove** نیز این پیش پردازش ها را اعمال می‌کنیم. پیش‌پردازش ها:

- alphabet_correction پارسى نرم
- semi_space_correction پارسى نرم
- remove_comma_between_numbers پارسى نرم
- remove_repeated_punctuation پارسى نرم
- جايگزارى حروفى مثل « آ، اِ، اُ، اَ، اٌ، اٍ، وُ » با حرف « ا » يا « و »
- جايگزارى تمام اعداد با مقدار صفر
- همچنين بردار متناظر با لغاتى كه آخر آنها « ها » يا « هاى » وجود دارد را برابر با همان لغات بدون اين مقادير در نظر گرفتيم

داده های آموزش، اعتبارسنجی و ارزیابی را به ترتیب با نسبت های ۰.۷ و ۰.۱۵ و ۰.۱۵ تقسیم می کنیم. تعداد داده های هر کدام در جدول ۲ قابل مشاهده است.

جدول ۲- تعداد داده های آموزش و اعتبارسنجی و ارزیابی

۶۲۵۳۸	تعداد داده های آموزش
۱۳۴۰۱	تعداد داده های اعتبارسنجی
۱۳۴۰۲	تعداد داده های ارزیابی

بخش ب - تبدیل متن به بردار

در این بخش از بردار تعبیه GloVe معرفی شده در تمرین استفاده می‌کنیم. این بردار تعبیه شامل ۴۹۹۰۳ لغت همراه با بردارهای ۳۰۰ تایی متناظر آنها می‌باشد. همچنین یک برچسب با نام <unk> نیز برای لغاتی که در آن وجود ندارند نیز در نظر گرفته شده است.

مدل بردار سراسری یا GloVe^۱ یک الگوریتم آموزش بدون نظارت^۲ است که برای به دست آوردن بردارهای متناظر لغات استفاده می‌شود. این مدل توسط پنینگتون و همکاران به منظور ترکیب مدل تجزیه ماتریس مبتنی بر شمارش^۳ و مدل skip-gram مبتنی بر محتوا^۴ ارائه گردیده است.

بخش پ - پیاده‌سازی و ارزیابی شبکه‌ی بازگشتی

در این بخش یک مدل RNN ساده ساخته و این شبکه را آموزش می‌دهیم. ساختار این شبکه در جدول ۳ قابل مشاهده است. همچنین هاپر پارامترهای مورد استفاده نیز در جدول ۴ ذکر شده است.

جدول ۳- ساختار شبکه RNN ساده

Input dimension	300
Hidden state features	64
Output dimension	32

طبق جدول ۳ ابعاد ورودی، به اندازه بردارهای GloVe یعنی ۳۰۰ می‌باشد، همچنین ابعاد خروجی نیز به تعداد کلاس‌ها یعنی همان تگ‌هایی که نشان دهنده ادات سخن هستند، می‌باشد.

در مورد هاپر پارامترهای شبکه نیز پس از آزمون و خطا دیده شد که مقادیر دقت و خطا پس از ۲۰ دوره تا حد خوبی همگرا می‌شوند. بهینه ساز ADAM نیز انتخاب اولیه خوبی برای تمام شبکه بوده و از SGD عملکردی بهتری داد. مقدار سرعت آموزش نیز 0.001 در نظر گرفته شده زیرا مقادیر کمتر از آن باعث

^۱ Global vectors

^۲ Unsupervised

^۳ Count-based matrix factorization

^۴ Context-based skip gram

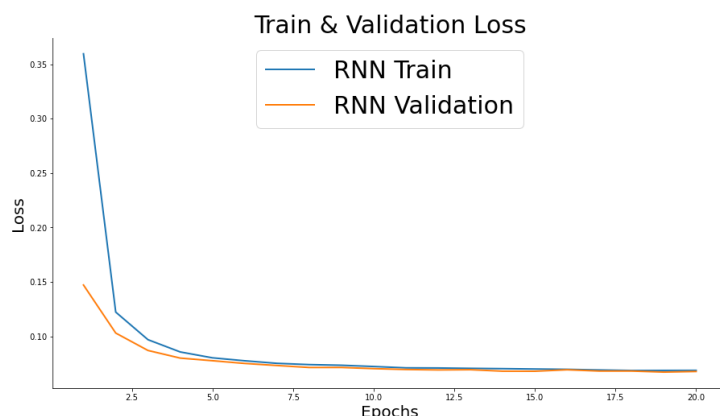
کاهش سرعت همگرایی شبکه و مقادیر بیشتر باعث نویزی شدن نتایج و همگرا نشدن آنها می‌شود. همچنین چون مسئله به نوعی یک طبقه بندی می‌باشد از تابع محاسبه خطای cross entropy استفاده می‌کنیم.

جدول ۴- هایپرپارامترهای شبکه RNN ساده

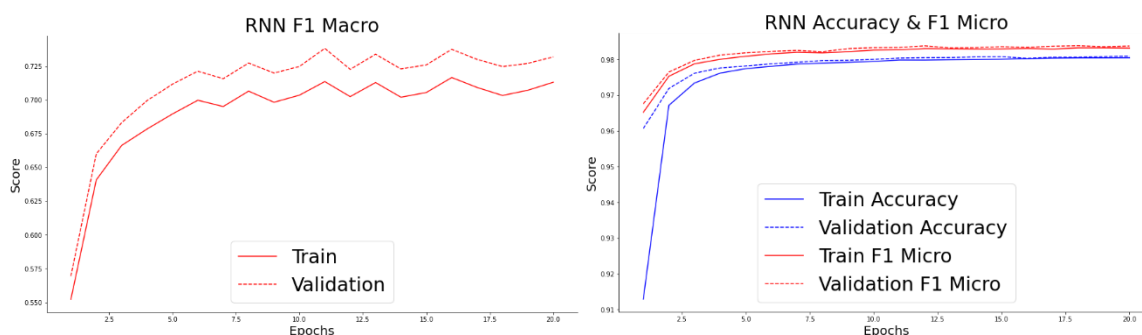
• Batch Size	64
• Epochs	20
• Optimizer	ADAM
• Learning Rate	0.001
• Weight Decay	0.0001
• Loss Function	Cross entropy
• Scheduler	LR decreases by factor of 0.1 when LR stopped improving for 2 epochs

نتایج حاصل از شبکه در شکل‌های ۱۴ و ۱۵ رسم شده‌است. چنانچه در نمودار خطای شکل ۱۴ قابل مشاهده است، پس از ۲۰ اپیاک منحنی خطا تقریباً همگرا شده و به نقطه بهینه رسیده است. همچنین در آزمون های دقت شکل ۱۵ مشاهده می‌شود که حتی در دوره اول به دقت بالایی رسیده‌ایم. این امر می‌تواند به دلیل عدم وجود توازن در داده ها باشد، برای مثال ممکن است تعداد کلماتی که نقش اسم را در جمله دارند بسیار بیشتر از بقیه کلمات باشد.

علاوه بر این، نکته قابل ذکر دیگر این است که نمودار دقت و $F1$ micro بسیار به هم شبیه هستند. این امر بدین خاطر است که در $micro$ معیار $f1$ به صورت $global$ روی تمام داده ها و کلاس ها محاسبه می‌شود (مشابه معیار دقت). اما در $macro$ معیار $f1$ برای هر کلاس به صورت جداگانه محاسبه شده و سپس میانگین گرفته می‌شود. به همین دلیل معیار $macro$ مقادیر کمتری دارد.



شکل ۱۴- نمودار خطای دادگان آموزش و اعتبارسنجی مدل RNN ساده



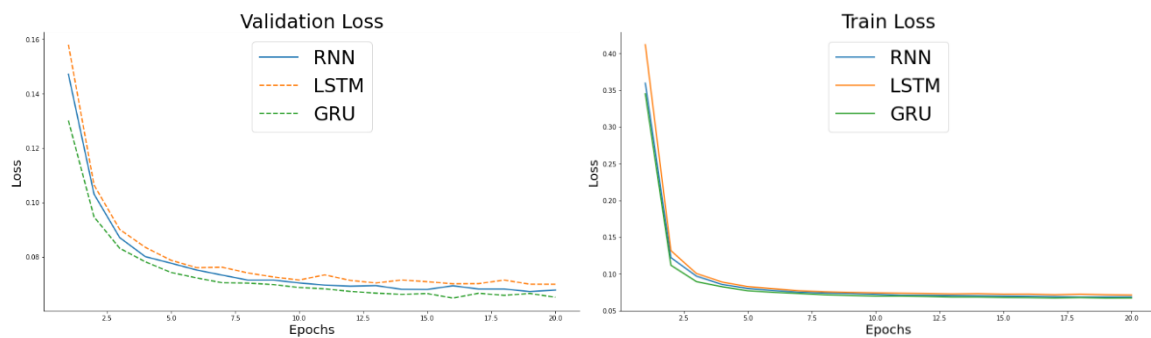
شکل ۱۵- معیارهای آزمون دقت شبکه RNN ساده

بخش ت- انواع شبکه‌های بازگشتی

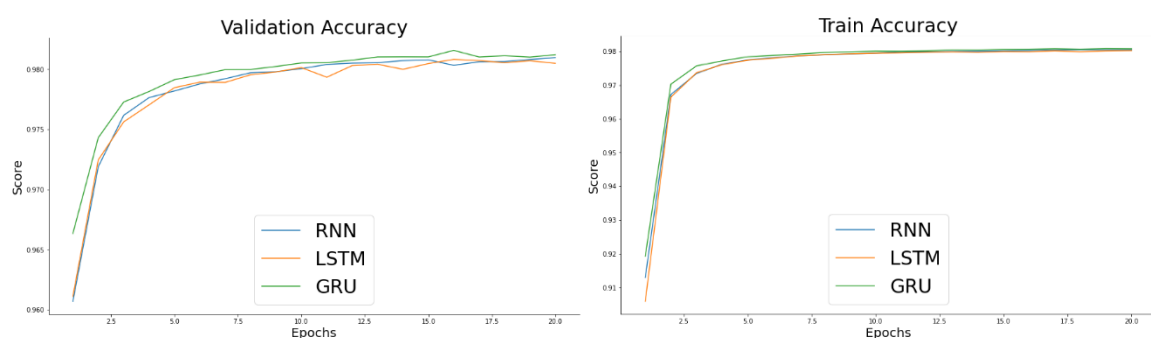
در این قسمت شبکه بازگشتی بخش قبل با کمک ساختارهای LSTM و GRU نیز آموزش دیده و نتایج با هم مقایسه می‌شوند.

چنانچه در شکل‌های ۱۶ تا ۱۹ مشاهده می‌شود، این سه ساختار در ۲۰ دوره ای که آموزش دیده اند تقریباً بر هم منطبق هستند و تفاوت موجود در نمودارها بسیار جزئی می‌باشد. لذا نمی‌توان مقایسه دقیقی بین آنها انجام داد. ولی به هر حال به نظر می‌رسد بهترین نتایج با اختلافی ناچیز توسط ساختار GRU حاصل شده است.

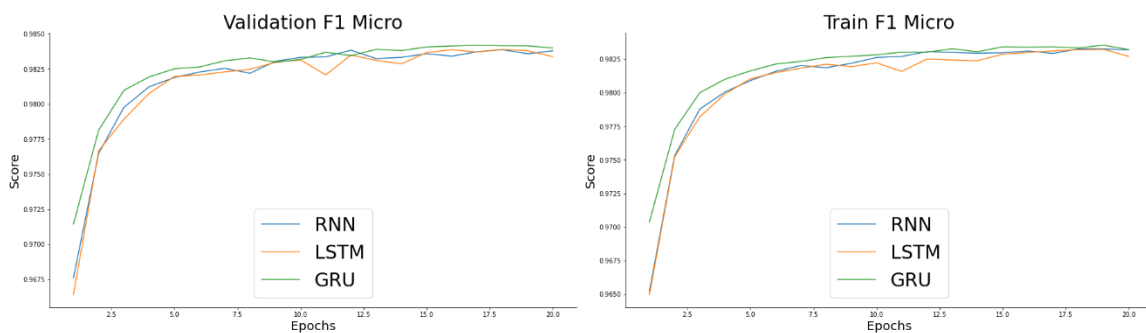
ساختارهای GRU و LSTM با استفاده از ایده سلول‌های حافظه و گیت‌های ورودی و خروجی، مشکلاتی چون محو شدن گرادینان‌ها و مشکل حافظه در دنباله‌های بلند را حل کرده و عملکرد بهتری را نسبت به ساختار ساده RNN نتیجه می‌دهند. (البته ساختار ساده RNN مزیت سادگی و سرعت را دارا می‌باشد) GRU در واقع ساده شده ساختار LSTM می‌باشد و به دلیل ساختار ساده ای که دارد از LSTM در آموزش سریعتر بوده و عملکردی مشابه را در عمل منجر می‌شود.



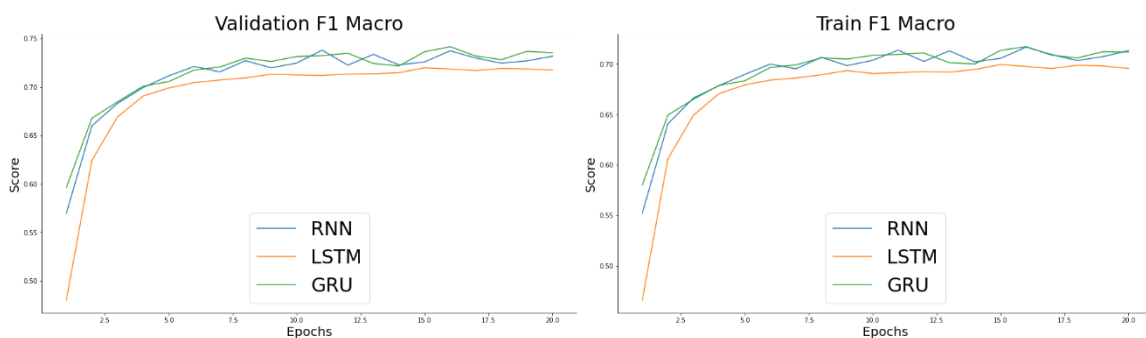
شکل ۱۶- مقایسه نمودارهای خطای دادگان آزمون و اعتبار سنجی روی شبکه های RNN, LSTM, GRU



شکل ۱۷- مقایسه معیار دقت دادگان آموزش و اعتبارسنجی در شبکه های RNN, LSTM, GRU



شکل ۱۸- مقایسه معیار F1 micro دادگان آموزش و اعتبارسنجی در شبکه های RNN, LSTM, GRU

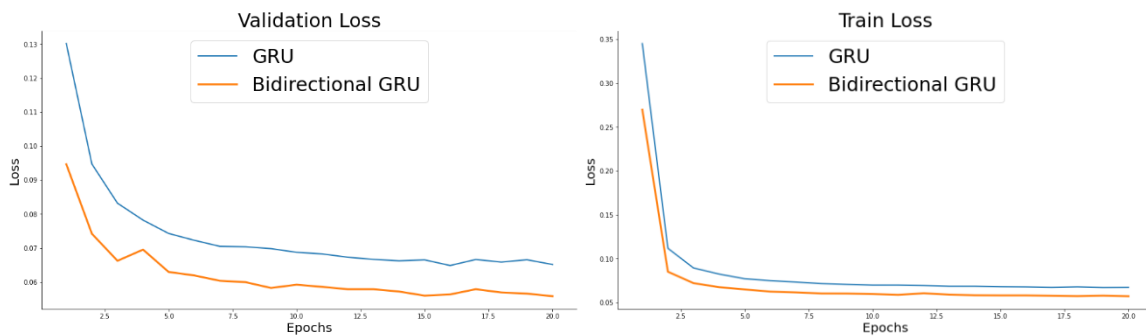


شکل ۱۹- مقایسه معیار F1 micro دادگان آموزش و اعتبارسنجی در شبکه های RNN, LSTM, GRU

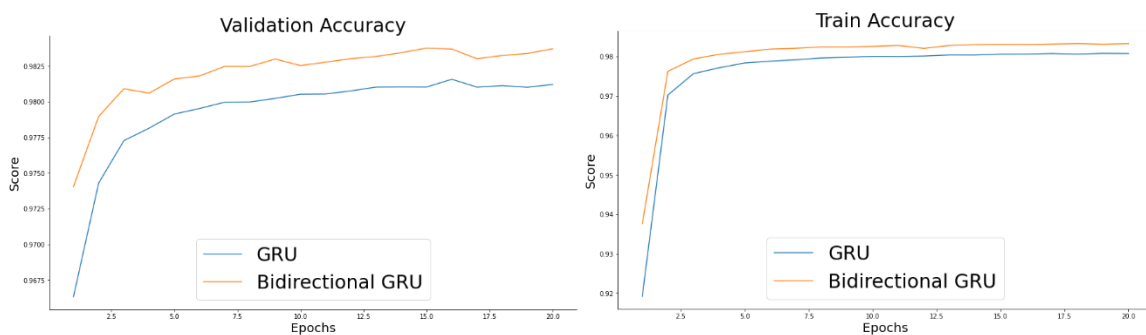
بخش ث – شبکه دو طرفه

در این قسمت بهترین ساختار بخش قبلی (GRU) به صورت دوطرفه پیاده شده و نتایج حاصل از آموزش آن با ساختار GRU معمولی مقایسه می‌شود. در این شبکه ها برای اینکه تصمیم درستی در زمان † داشته باشیم علاوه بر ورودی های قبلی، به ورودی های بعدی نیز توجه می‌شود. البته واضح است که ساختارهای دو طرفه در مسائل **real-time** کاربردی ندارند.

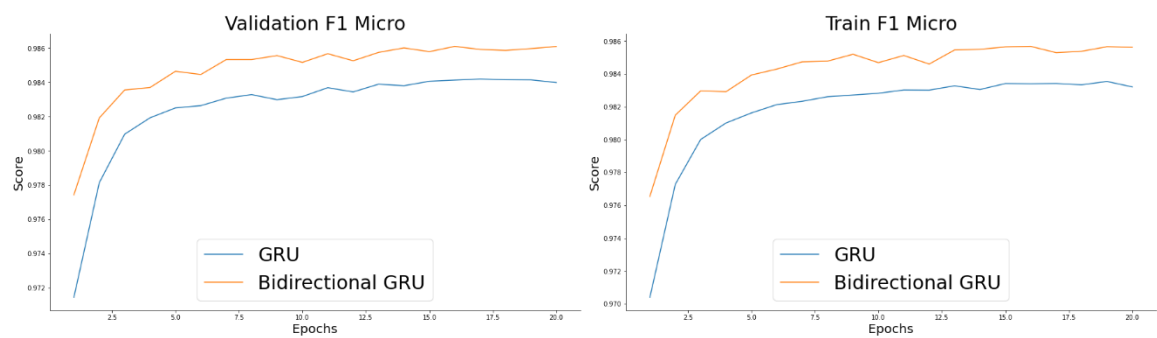
چنانچه در نمودار خطای شکل ۲۰ و نمودار های آزمون دقت شکل های ۲۱ تا ۲۳ قابل مشاهده است استفاده از ساختار دو طرفه باعث بهبود چند درصدی نتایج شده است.



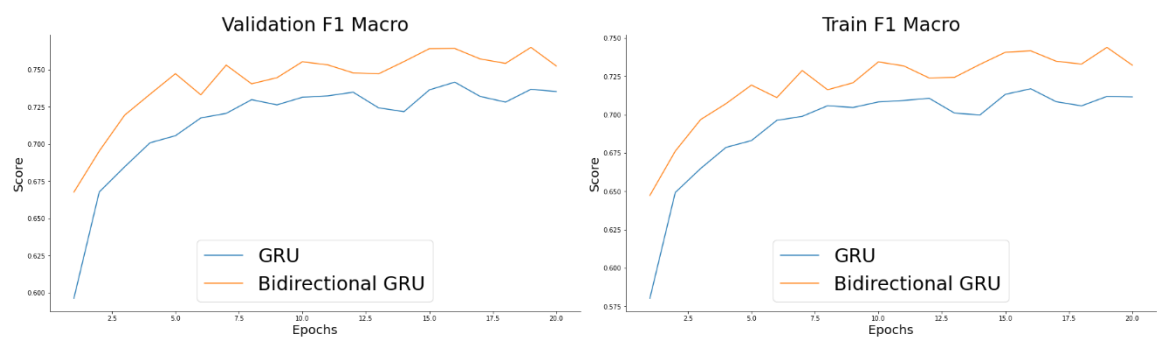
شکل ۲۰- مقایسه نمودارهای خطای دادگان آزمون و اعتبار سنجی شبکه های دوطرفه و یک طرفه GRU



شکل ۲۱- مقایسه معیار دقت دادگان آزمون و اعتبار سنجی شبکه های دوطرفه و یک طرفه GRU



شکل ۲۲- مقایسه معیار F1 micro دادگان آزمون و اعتبار سنجی شبکه های دوطرفه و یک طرفه GRU

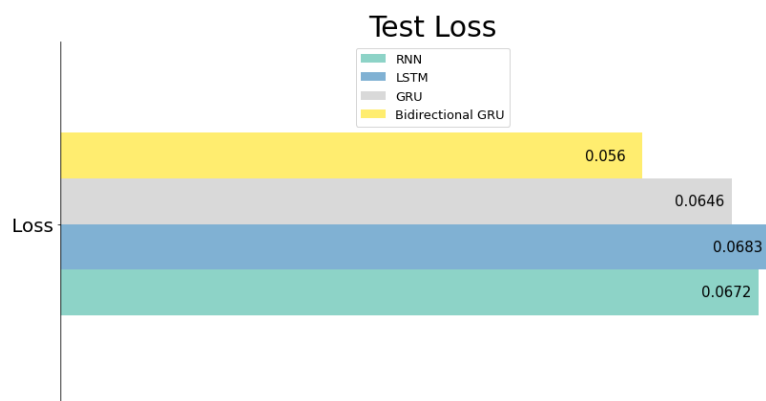


شکل ۲۳- مقایسه معیار F1 macro دادگان آزمون و اعتبار سنجی شبکه های دوطرفه و یک طرفه GRU

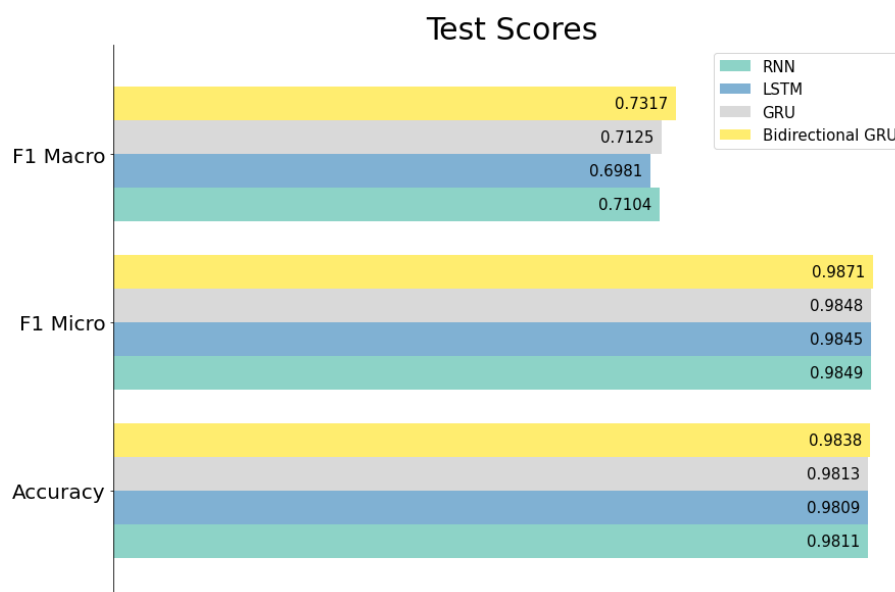
در نهایت نتایج نهایی خطا و معیارهای دقت شبکه آزموده شده GRU با ساختار دوطرفه، بروی دادگان ارزیابی در شکل‌های ۲۳ و ۲۴ قابل مشاهده هستند. همچنین از این شبکه برای تشخیص ادات یک جمله دلخواه استفاده می‌کنیم:

" تشخیص ادات سخن ، یکی از مسائلی است که در پردازش متن به عنوان گام اولیه برای سایر کارها ، کاربرد بسیار دارد . "

(تشخیص , N_SING) (ادات , N_SING) (سخن , N_SING) (, DELM) (یکی , PRO) (از , P)
 (مسائلی , N_PL) (است , V_PRS) (که , CON) (در , P) (پردازش , N_SING) (متن , N_SING)
 (به , P) (عنوان , N_SING) (گام , N_SING) (اولیه , ADJ) (برای , P) (سایر , N_SING)
 (کارها , N_PL) (, DELM) (کاربرد , N_SING) (بسیار , ADJ) (دارد , V_PRS) (, DELM)



شکل ۲۴- مقایسه خطای ساختارهای مختلف RNN روی دادگان ارزیابی



شکل ۲۵- مقایسه معیارهای دقت ساختارهای مختلف RNN روی دادگان ارزیابی