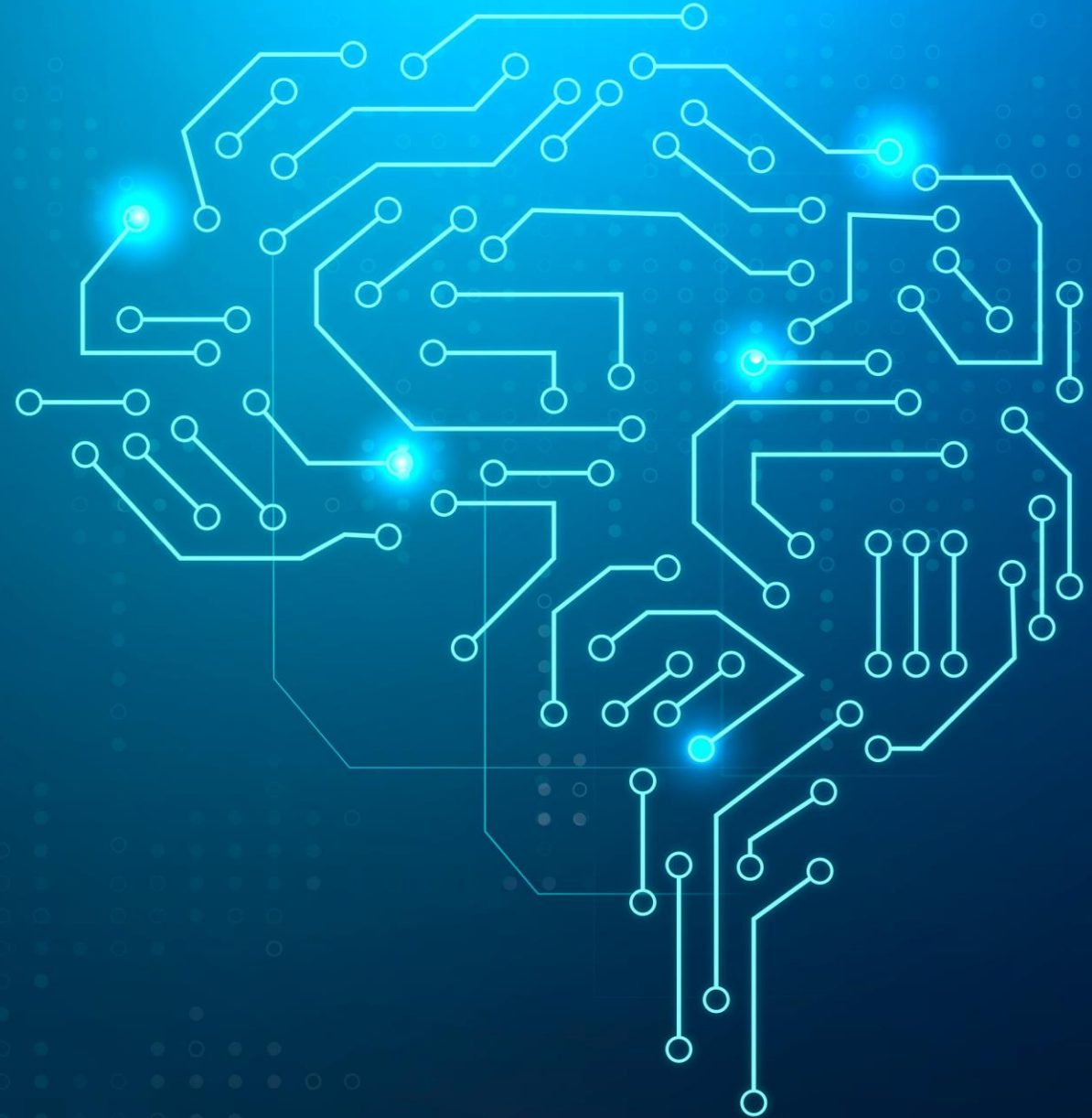




# SEGNET

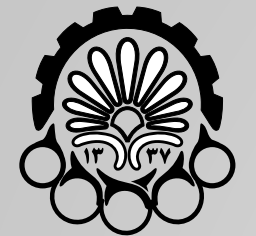
مهدی فیروزبخت  
۴۰۰۱۳۱۰۲۷  
دکتر رضا صفابخش

دپارتمان هوش مصنوعی  
تابستان ۱۴۰۱

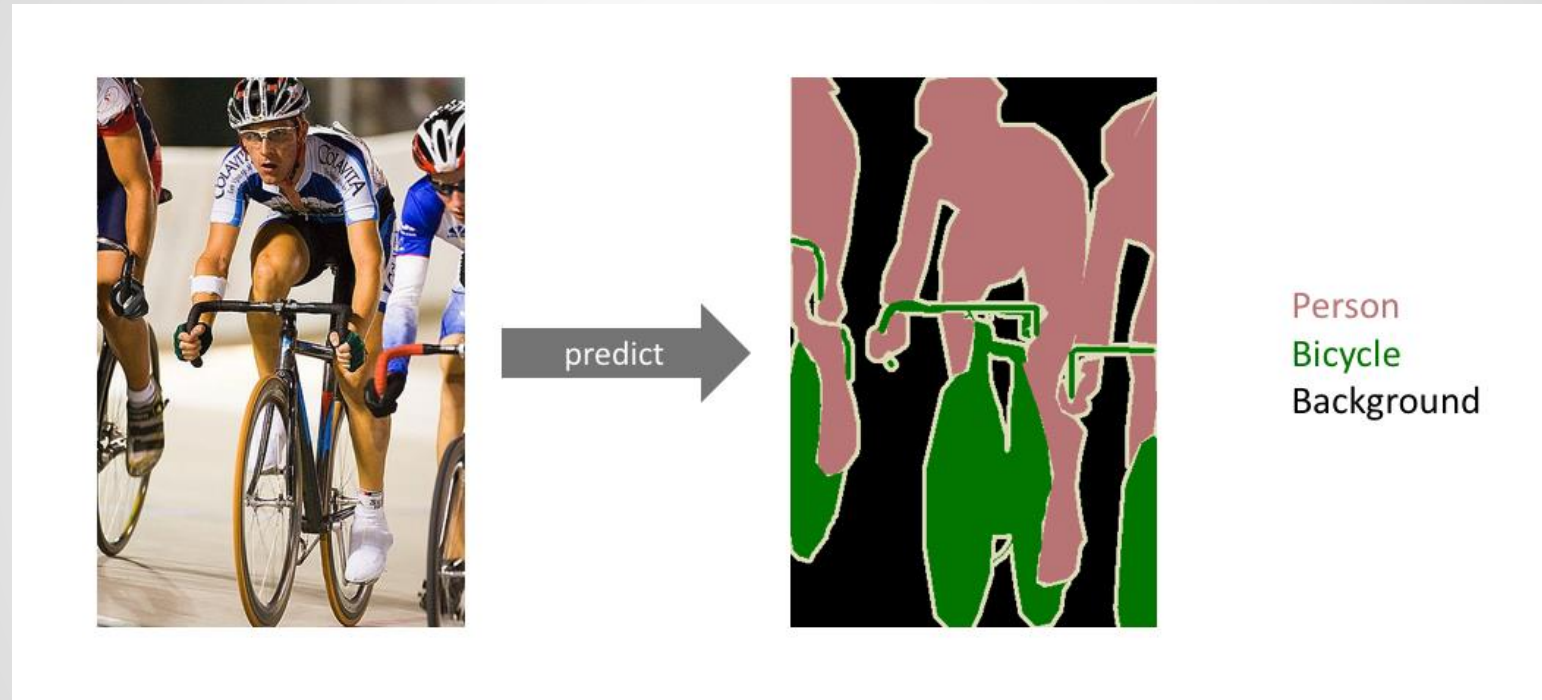


# فهرست

- مقدمه
- تقسیم بندی معنایی
- معماری Segnet
- معماری P-Segnet و NP-Segnet
- معماری FL-Segnet
- معماری M-Segnet
- معماری FA-Segnet
- مراجع و منابع



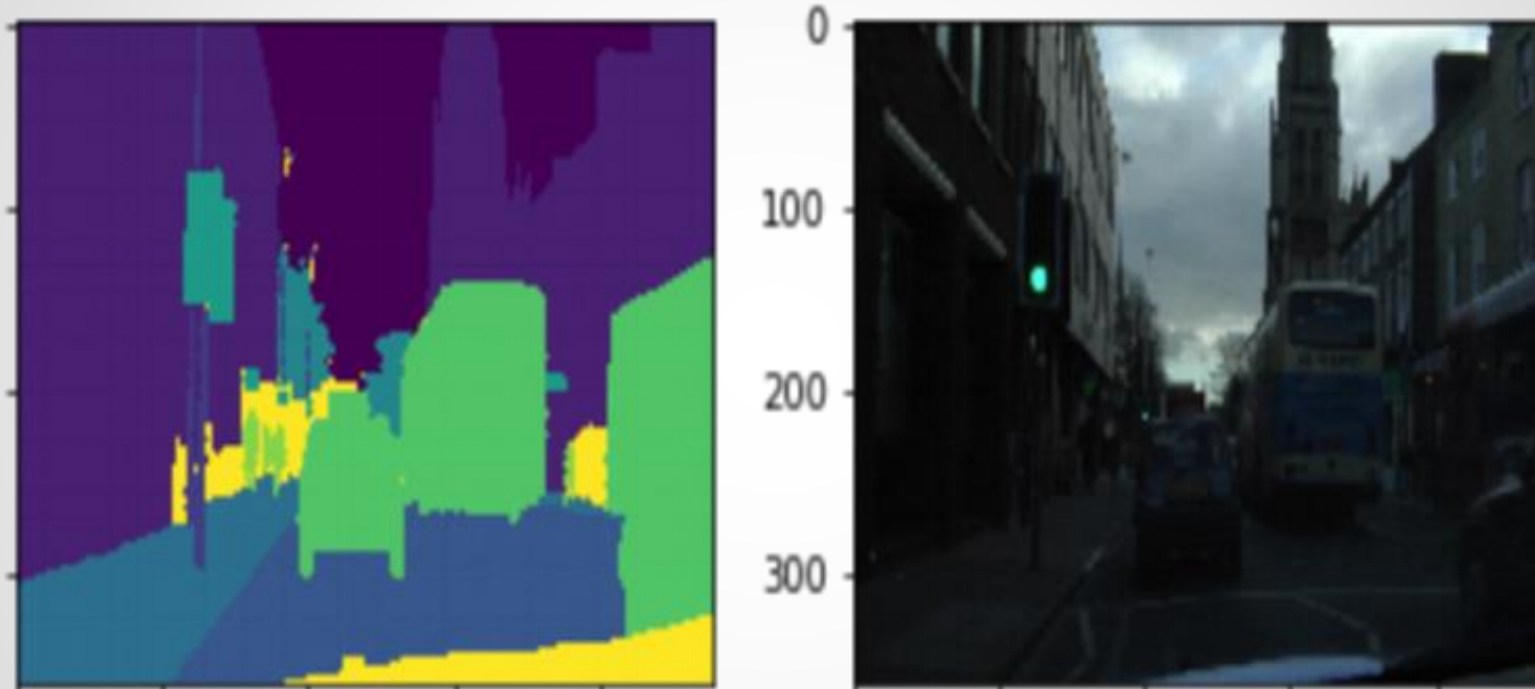
# مقدمه



شبکه عصبی SegNet که توسط الکس کندال ، وی جی بدرینارایانان و روبرتو سیپولا که همگی از دانشگاه کمبریج هستند ، توسعه یافته است، یک شبکه عصبی پیچشی است که برای تقسیم بندی معنایی استفاده شده است.

# تقسیم بندی معنایی

تقسیم بندی تصویر معنایی وظیفه طبقه بندی هر پیکسل در یک تصویر از مجموعه ای از کلاس های از پیش تعریف شده است.

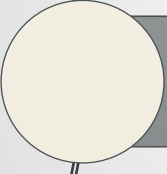


در مثال بالا، پیکسل های متعلق به وسیله نقلیه در کلاس "خودرو" طبقه بندی می شوند، پیکسل های مربوط به جاده ها به عنوان "جاده" برچسب گذاری شده اند.

## استفاده از تقسیم بندی معنایی



تصاویر پزشکی



سیستم های خودمختار

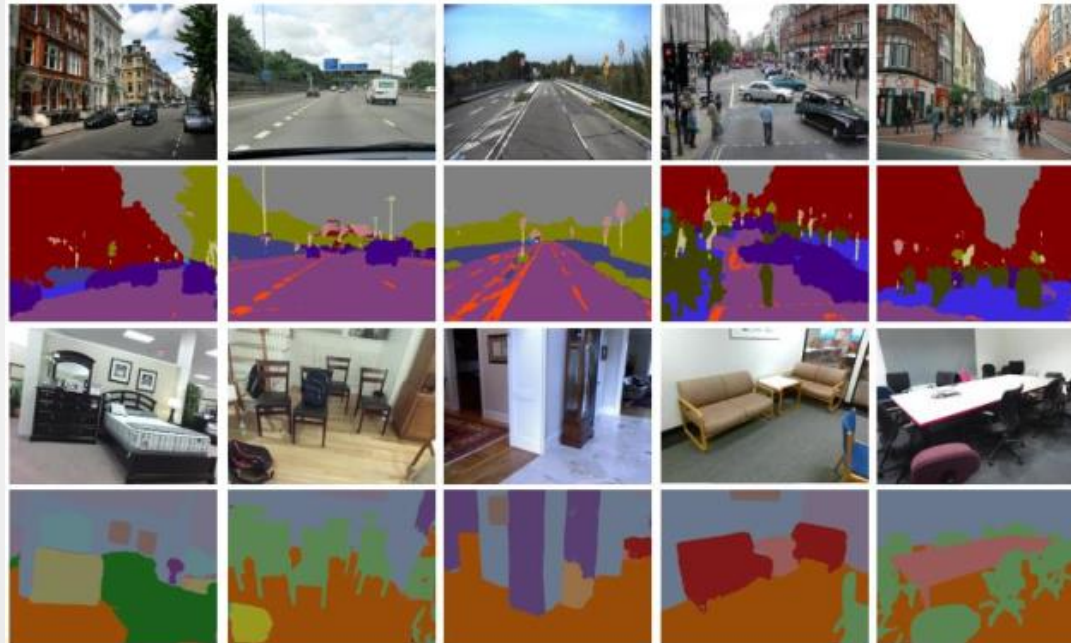


تجزیه و تحلیل تصویر جغرافیایی

# معماري SEGNET

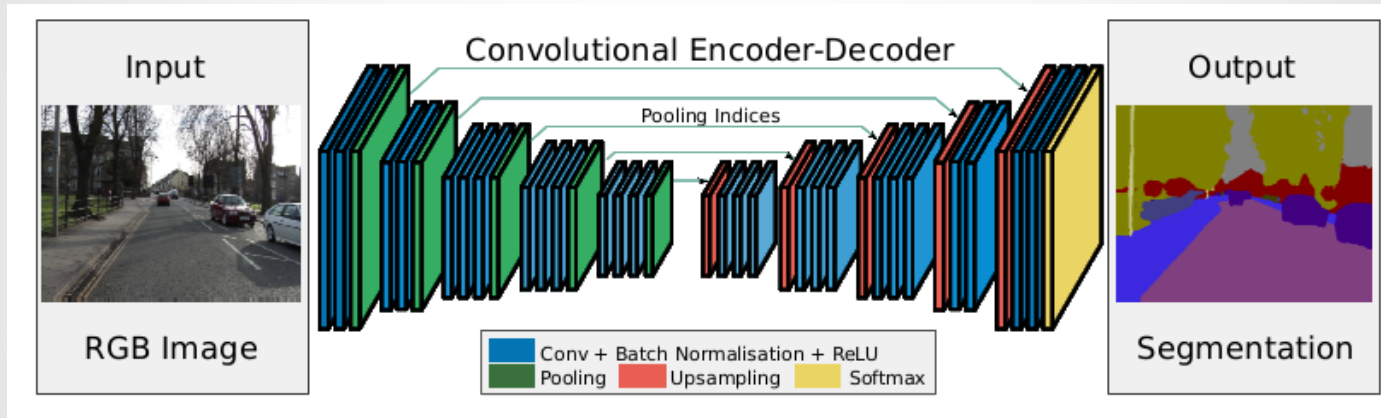


## معرفی معماری



Segnet ، طراحی شده است تا یک معماری کارآمد برای تقسیم بندی معنایی باشد. انگیزه اصلی آن برنامه های درک صحنه جاده است که به توانایی مدلسازی ظاهر (جاده، ساختمان)، تصویر (ماشینها، عابران پیاده) و درک رابطه فضایی (زمینه) بین طبقات مختلف مانند جاده و پیاده روی نیاز دارد.

# معرفی شبکه



SegNet دارای یک شبکه رمزگذار و یک شبکه رمزگشای متناظر است که به دنبال آن یک لایه طبقه‌بندی پیکسلی نهایی قرار دارد.

## ارزیابی شبکه

- دقت جهانی که درصد پیکسل‌هایی را که به درستی در مجموعه داده طبقه‌بندی شده‌اند، اندازه‌گیری می‌کند.
- دقت متوسط کلاس میانگین دقت پیش‌بینی‌کننده در تمام کلاس‌ها
- mIoU روی همه کلاس‌ها است. اگر مساحت پیش‌بینی شده و مساحت کلاس‌بندی شده درست را جعبه در نظر بگیریم ، مساحت مشترک دو جعبه را در صورت کسر و میزان کل مساحتی که این دو جعبه دارند در مخرج کسر قرار می دهیم و حاصل عددی بین ۰ و ۱ خواهد بود.



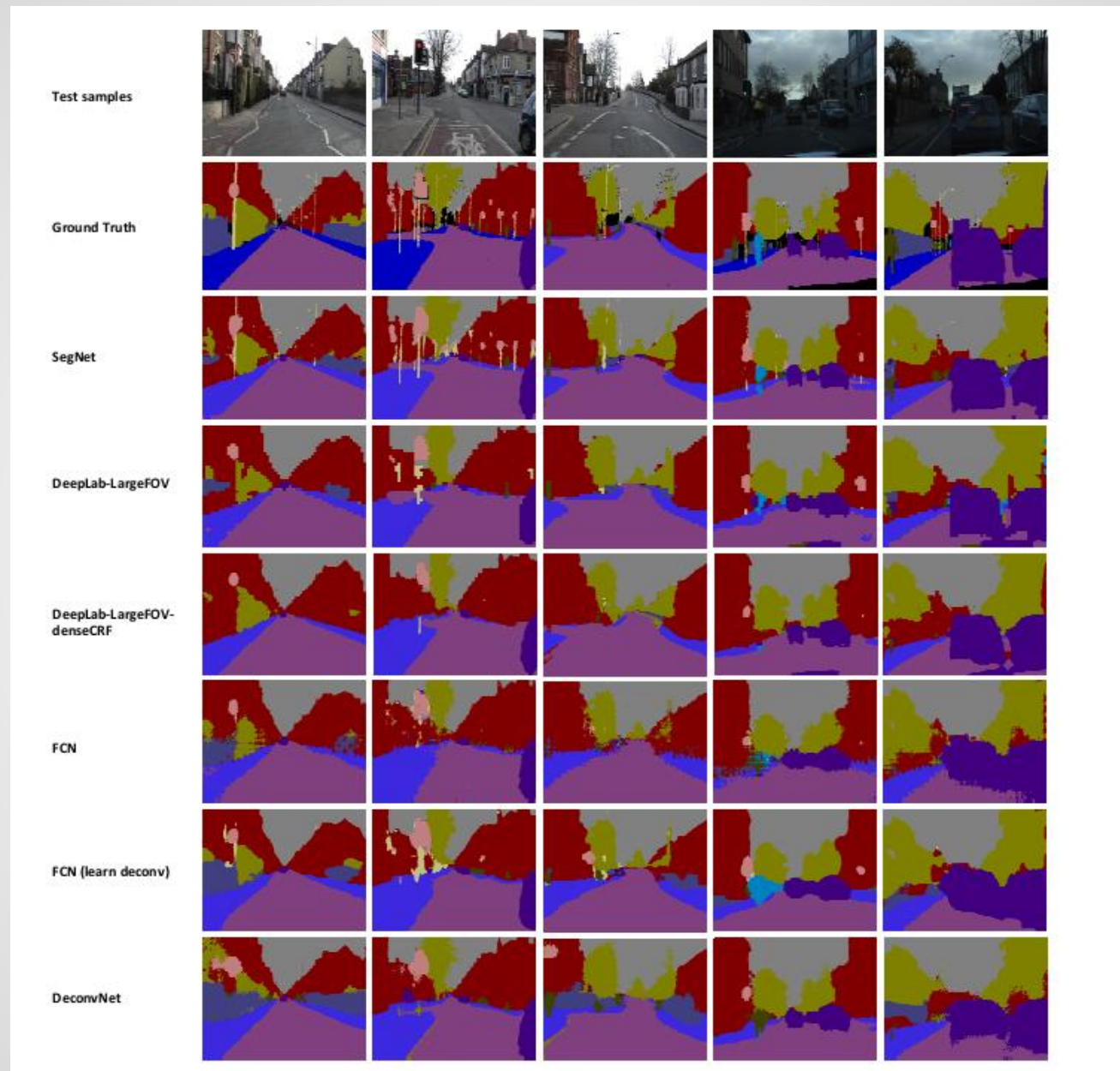
## مقایسه FCN و SEGNET

Variant	Params (M)	Storage multiplier	Infer time (ms)	Median frequency balancing						Natural frequency balancing							
				Test				Train		Test				Train			
				G	C	mIoU	BF	G	C	mIoU	G	C	mIoU	BF	G	C	mIoU
Fixed upsampling																	
Bilinear-Interpolation	0.625	0	24.2	77.9	61.1	43.3	20.83	89.1	90.2	82.7	82.7	52.5	43.8	23.08	93.5	74.1	59.9
Upsampling using max-pooling indices																	
SegNet-Basic	1.425	1	52.6	82.7	62.0	47.7	35.78	94.7	96.2	92.7	84.0	54.6	46.3	36.67	96.1	83.9	73.3
SegNet-Basic-EncoderAddition	1.425	64	53.0	83.4	<b>63.6</b>	48.5	35.92	94.3	95.8	92.0	<b>84.2</b>	56.5	<b>47.7</b>	36.27	95.3	80.9	68.9
SegNet-Basic-SingleChannelDecoder	0.625	1	33.1	81.2	60.7	46.1	31.62	93.2	94.8	90.3	83.5	53.9	45.2	32.45	92.6	68.4	52.8
Learning to upsample (bilinear initialisation)																	
FCN-Basic	0.65	11	24.2	81.7	62.4	47.3	<b>38.11</b>	92.8	93.6	88.1	83.9	55.6	45.0	<b>37.33</b>	92.0	66.8	50.7
FCN-Basic-NoAddition	0.65	n/a	23.8	80.5	58.6	44.1	31.96	92.5	93.0	87.2	82.3	53.9	44.2	29.43	93.1	72.8	57.6
FCN-Basic-NoDimReduction	1.625	64	44.8	<b>84.1</b>	63.4	<b>50.1</b>	37.37	<b>95.1</b>	<b>96.5</b>	<b>93.2</b>	83.5	<b>57.3</b>	47.0	37.13	<b>97.2</b>	<b>91.7</b>	<b>84.8</b>
FCN-Basic-NoAddition-NoDimReduction	1.625	0	43.9	80.5	61.6	45.9	30.47	92.5	94.6	89.9	83.7	54.8	45.5	33.17	95.0	80.2	67.8

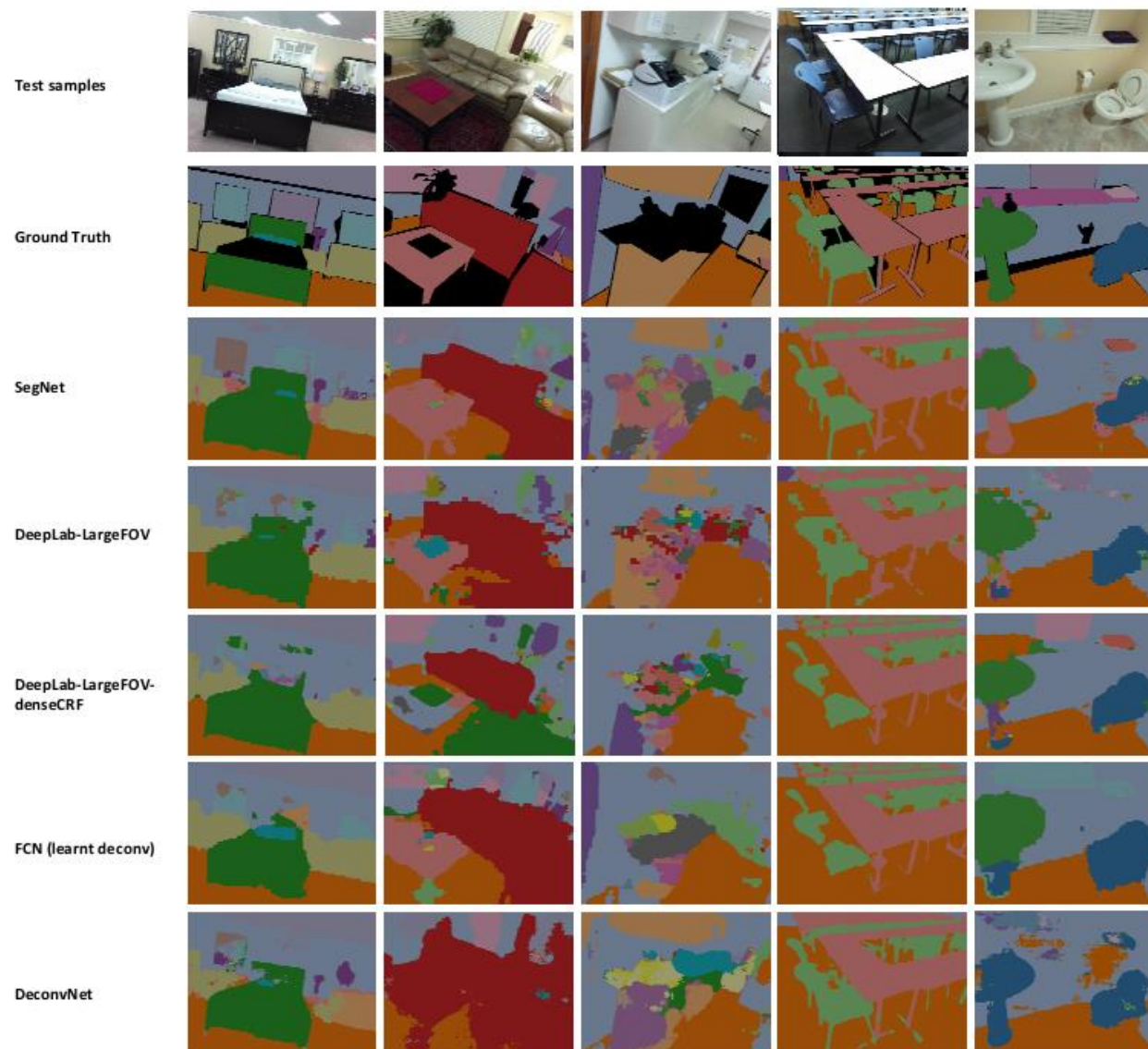


## مقایسه SEGNET و سایر معماری ها

Method	Building	Tree	Sky	Car	Sign-Symbol	Road	Pedestrian	Fence	Column-Pole	Side-walk	Bicyclist	Class avg.	Global avg.	mIoU	BF
SfM+Appearance [28]	46.2	61.9	89.7	68.6	42.9	89.5	53.6	46.6	0.7	60.5	22.5	53.0	69.1	n/a*	
Boosting [29]	61.9	67.3	91.1	71.1	58.5	92.9	49.5	37.6	25.8	77.8	24.7	59.8	76.4	n/a*	
Dense Depth Maps [32]	85.3	57.3	95.4	69.2	46.5	<b>98.5</b>	23.8	44.3	22.0	38.1	28.7	55.4	82.1	n/a*	
Structured Random Forests [31]	n/a											51.4	72.5	n/a*	
Neural Decision Forests [64]	n/a											56.1	82.1	n/a*	
Local Label Descriptors [65]	80.7	61.5	88.8	16.4	n/a	98.0	1.09	0.05	4.13	12.4	0.07	36.3	73.6	n/a*	
Super Parsing [33]	87.0	67.1	96.9	62.7	30.1	95.9	14.7	17.9	1.7	70.0	19.4	51.2	83.3	n/a*	
SegNet (3.5K dataset training - 140K)	<b>89.6</b>	<b>83.4</b>	96.1	<b>87.7</b>	52.7	96.4	<b>62.2</b>	<b>53.45</b>	<b>32.1</b>	<b>93.3</b>	<b>36.5</b>	<b>71.20</b>	<b>90.40</b>	60.10	46.84
CRF based approaches															
Boosting + pairwise CRF [29]	70.7	70.8	94.7	74.4	55.9	94.1	45.7	37.2	13.0	79.3	23.1	59.9	79.8	n/a*	
Boosting+Higher order [29]	84.5	72.6	<b>97.5</b>	72.7	34.1	95.3	34.2	45.7	8.1	77.6	28.5	59.2	83.8	n/a*	
Boosting+Detectors+CRF [30]	81.5	76.6	96.2	78.7	40.2	93.9	43.0	47.6	14.3	81.5	33.9	62.5	83.8	n/a*	









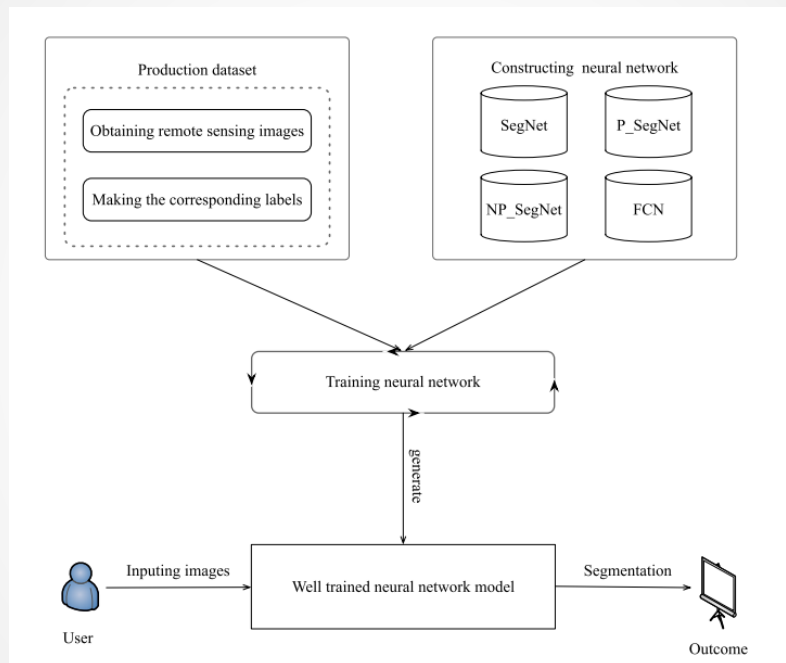
## مقایسه SEGNET در تصاویر داخلی

Network/Iterations	40K				80K				>80K				Max iter
	G	C	mIoU	BF	G	C	mIoU	BF	G	C	mIoU	BF	
SegNet	88.81	59.93	50.02	35.78	89.68	69.82	57.18	42.08	90.40	71.20	60.10	46.84	140K
DeepLab-LargeFOV [3]	85.95	60.41	50.18	26.25	87.76	62.57	53.34	32.04	88.20	62.53	53.88	32.77	140K
DeepLab-LargeFOV-denseCRF [3]	not computed								89.71	60.67	54.74	40.79	140K
FCN	81.97	54.38	46.59	22.86	82.71	56.22	47.95	24.76	83.27	59.56	49.83	27.99	200K
FCN (learnt deconv) [2]	83.21	56.05	48.68	27.40	83.71	59.64	50.80	31.01	83.14	64.21	51.96	33.18	160K
DeconvNet [4]	85.26	46.40	39.69	27.36	85.19	54.08	43.74	29.33	89.58	70.24	59.77	52.23	260K



# معماری P-SEGNET و NP-SEGNET

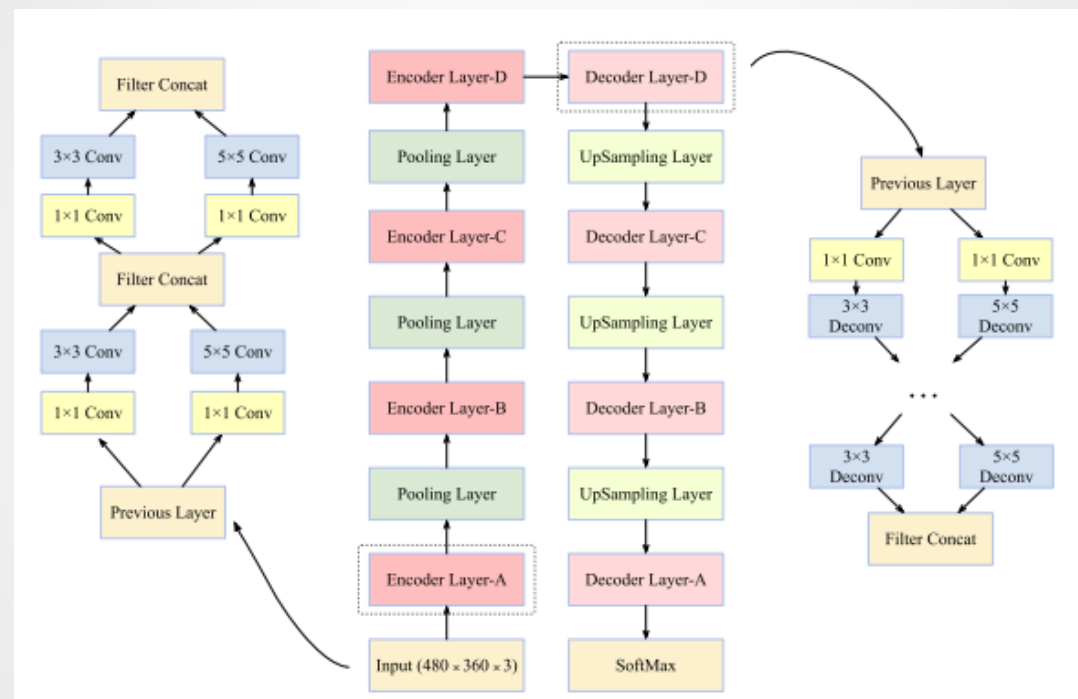
# معرفی معماری



انگیزه اصلی ساخت این معماری ، بهبود عملکرد در تشخیص ابرها بوده است. مدل شبکه عصبی موجود، SegNet، مبنایی برای ارزیابی عملکرد پردازش تشخیص ابر ارائه می‌دهد. با اصلاح معماری آن، می‌توان عواملی را که بر دقت تشخیص ابر و میزان تبعیض اشیاء بسیار مشابه تأثیر می‌گذارند، بررسی کرد. در تشخیص ابر، اشیاء بسیار مشابه شامل برف، بزرگراه‌ها، اشیاء برجسته و غیره است.



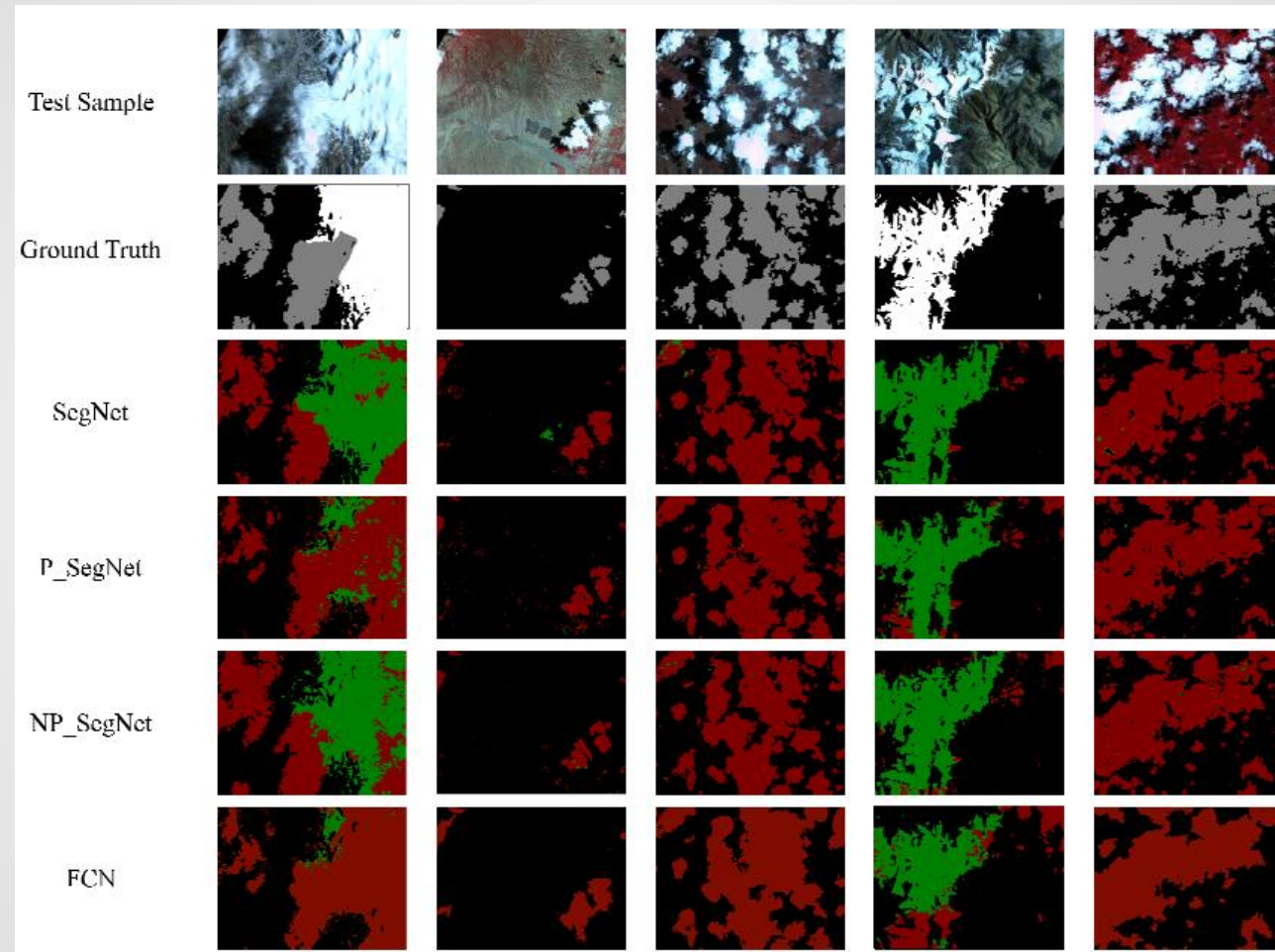
## معرفی شبکه



به منظور بهبود دقت تشخیص ابر، ساختار SegNet اصلاح شد. همانطور که در تصویر بالا نشان داده شده است، در این مقاله مدل اولیه GoogLeNet با استفاده از یک بانک فیلتر برای تولید مجموعه ای از نقشه های ویژگی در یک لایه خاص از شبکه تقلید شده است.

## ارزیابی شبکه

برای آزمایش روی این معماری از ۶۷۰ تصویر آموزشی و ۳۴۲ تصویر RGB آزمایشی و برچسب های تصویر مربوط به آنها استفاده شده است. برای این مجموعه داده، انتظار می رفت که شبکه عصبی سه کلاس را تقسیم کند: ابر، برف و پس زمینه. از دست دادن آنتروپی متقاطع یا (Cross entropy) به عنوان تابع هدف در شبکه استفاده شد.

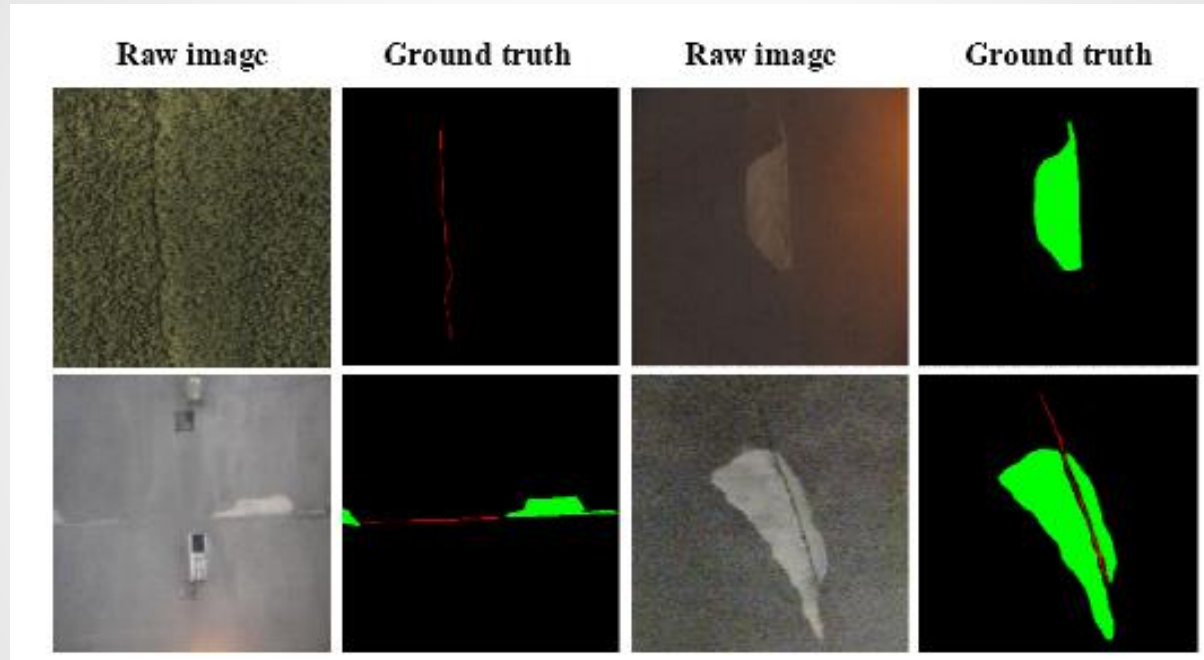




Network / Iteration	50K				100K				500K				
	Cloud	Snow	Background	Class avg.	Cloud	Snow	Background	Class avg.	Cloud	Snow	Background	Class avg.	Macro-F1
SegNet	82.6	75.5	86.0	81.4	88.6	79.5	89.5	85.9	89.9	88.7	90.6	89.7	0.707
P_SegNet	81.3	49.3	92.1	74.2	90.7	55.3	90.8	78.9	93.4	58.7	92.1	81.0	0.586
NP_SegNet	80.5	70.6	90.5	80.5	91.2	70.3	92.4	84.6	92.8	80.8	93.5	89.0	0.712
FCN	75.9	19.3	80.9	58.7	80.4	19.5	87.6	62.5	81.7	20.7	85.5	62.63	-

# معماری FL-SEGNET

## معرفی معماری

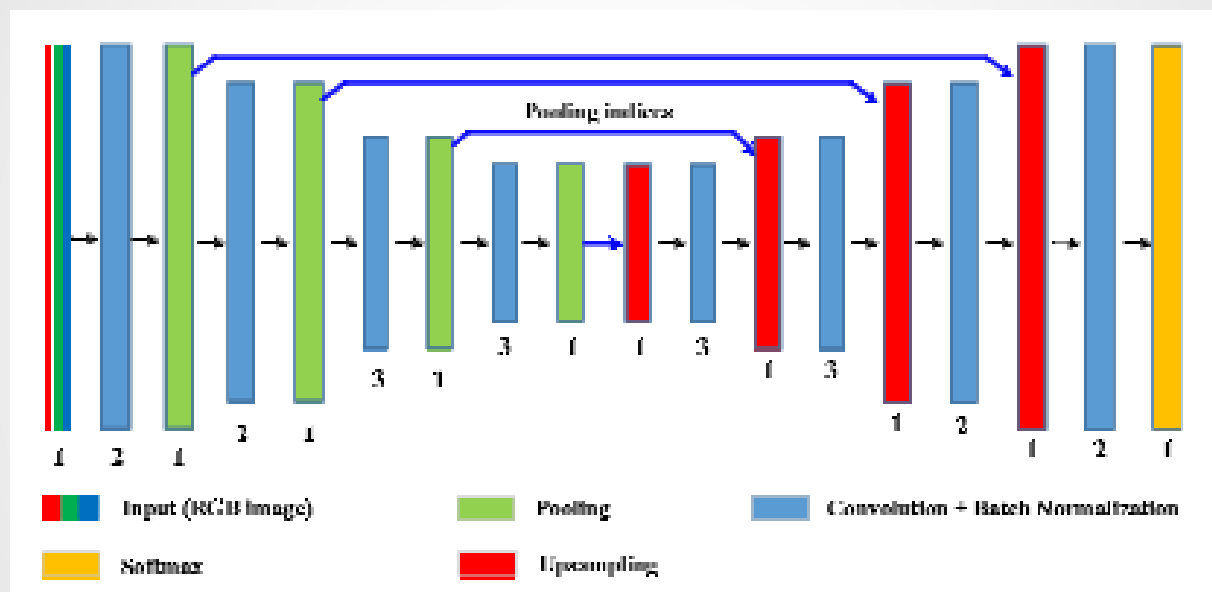


تشخیص آسیب های ساختاری پوشش تونل برای اطمینان از عملیات ایمن تونل حیاتی است. با این حال، تشخیص نقص چندگانه به دلیل عدم تعادل اندازه بین ترک ها، پوسته شدن و پس زمینه، کار چالش برانگیزی است. برای تشخیص دقیق نقص چندگانه در سطح پیکسل تنها با استفاده از شبکه های یک مرحله ای، روش جدیدی پیشنهاد شد که SegNet اصلی را با یک تابع از دست دادن کانونی یکپارچه می کند و به عنوان روش FL-SegNet نامیده می شود.





## معرفی شبکه



در این مطالعه ، به منظور پرداختن به این تصاویر، از تابع از دست دادن کانونی به جای تابع آنتروپی متقابل سنتی (CE) به منظور تمرکز بر یادگیری مثال‌های سخت و کاهش وزن منفی‌های متعدد استفاده شد.

## تابع از دست دادن کانونی

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t)$$

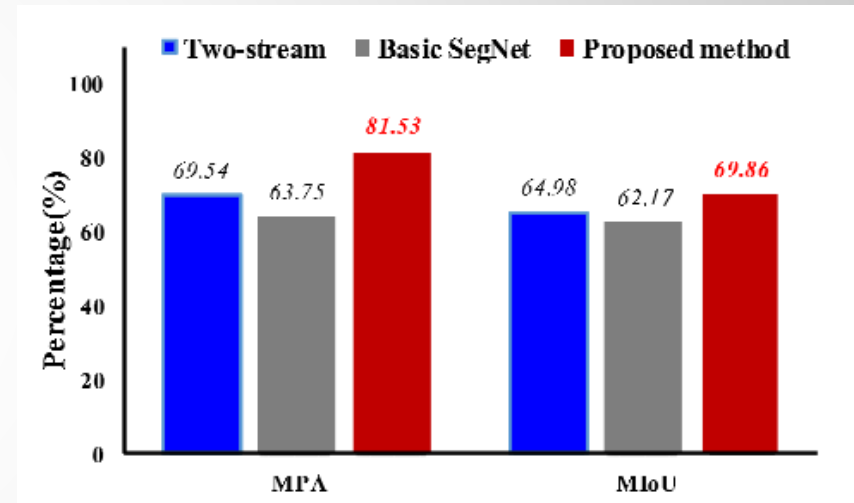
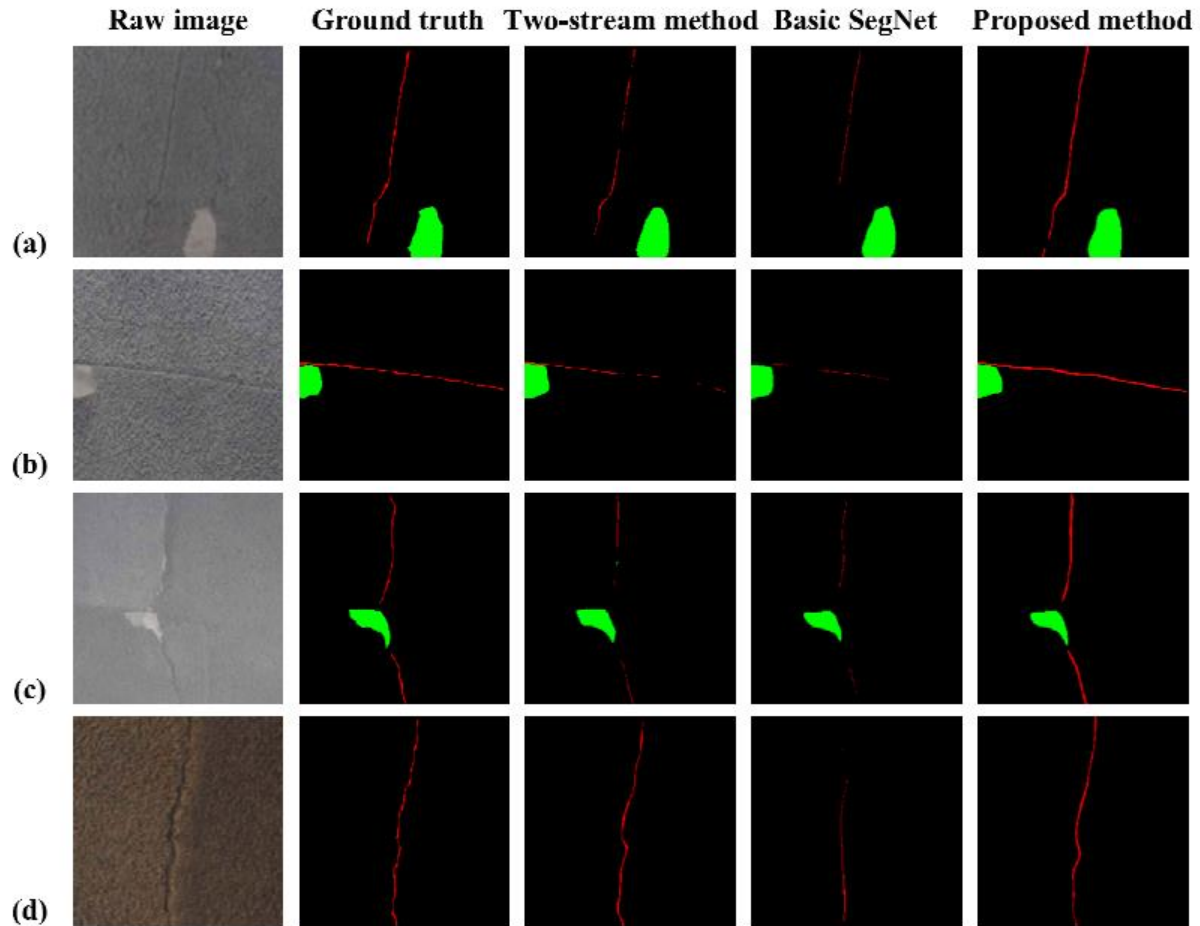
$$p_t = \begin{cases} p & \text{if } y = 1 \\ 1 - p & \text{otherwise.} \end{cases}$$

که در آن  $y$  کلاس داده مرجع است و  $p$  نشان دهنده برچسب پیش بینی با کلاس داده مرجع ۱ است.  $\gamma(1 - p_t)$  به عنوان یک عامل تعدیل کننده برای کاهش اتلاف نسبی برای نمونه های طبقه بندی شده خوب ( $p_t > 0.5$ ) استفاده می شود، و بیشتر بر روی نمونه های طبقه بندی اشتباه تمرکز می شود.  $\gamma$  افزایش می یابد اثر عامل تعدیل نیز افزایش می یابد. به این ترتیب، مشارکت ترک ها، پوسته شدن و پس زمینه ها را می توان در عملکرد از دست دادن متعادل کرد، که نه تنها کارایی را بهبود می بخشد، بلکه دقت شناسایی را نیز افزایش می دهد. در این فرمول آلفا نیز یک ضریب بین ۰ تا ۱ است.

## ارزیابی شبکه

مجموعه داده های مورد استفاده در این مطالعه از تونل نانشییی، که یک تونل جداگانه در بزرگراه داگوانگ در استان جیانگشی است، گرفته شده است. طول کل تونل ۹۲۴ متر است. کل مجموعه داده ها به دو بخش به شرح زیر تقسیم شدند: مجموعه های آموزشی (۲۲۶ تصویر) و مجموعه های تست (۷۲ تصویر). علاوه بر این، برای اثبات عدم تطبیق بیش از حد مدل، تصاویر آموزشی به طور تصادفی به پنج قسمت (چهار قسمت برای آموزش، و یک قسمت برای اعتبارسنجی) برای آزمایش های اعتبارسنجی متقابل پنج برابری این مطالعه تقسیم شدند. برای ارزیابی دقت انواع مختلف روش ها برای تقسیم بندی معنایی، دو معیار ارزیابی به شرح زیر اتخاذ شد: میانگین دقت پیکسل (MPA) و میانگین تقاطع روی اتحاد (MIoU).

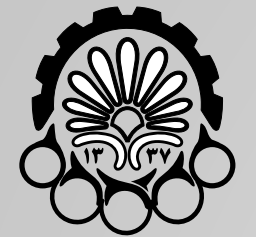
Step	Accuracy	Loss	Val acc	Val loss
1	0.9952	0.0034	0.9867	0.0768
2	0.9950	0.0040	0.9886	0.0450
3	0.9950	0.0039	0.9871	0.0516
4	0.9951	0.0037	0.9850	0.0641
5	0.9952	0.0036	0.9875	0.0449
Mean	0.9951	0.0037	0.9870	0.0565



## چالش های شبکه

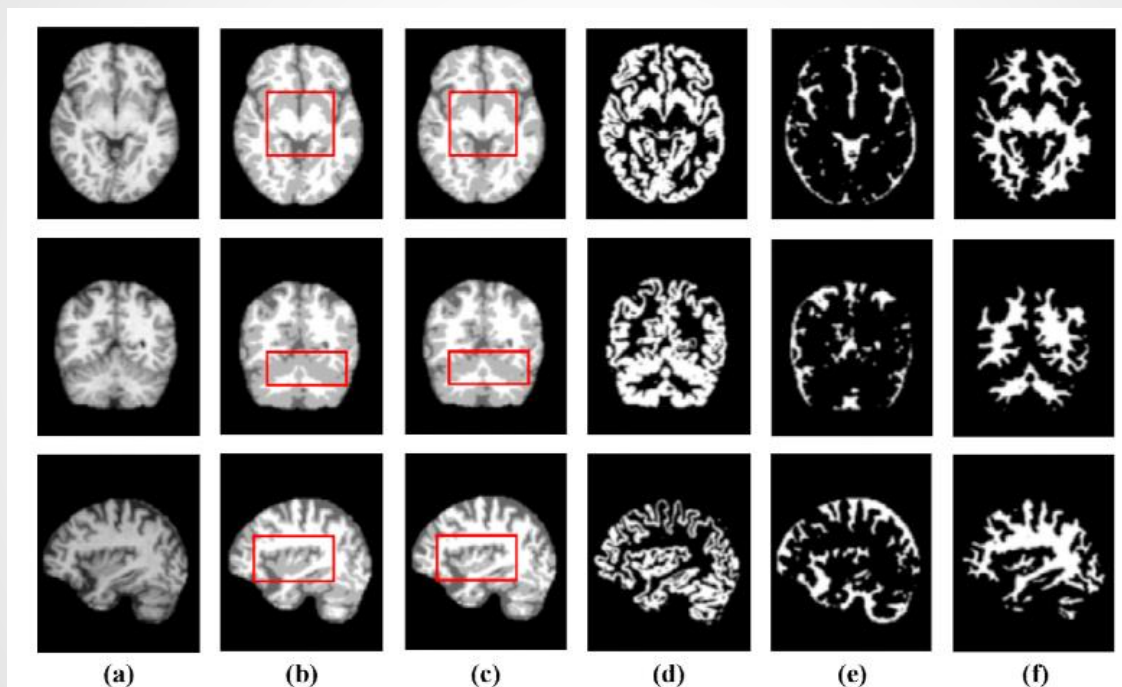
- تداخل پس زمین
- شرایط نوری ناهموار

Method	Background Interference		Uneven Lighting	
	MPA(%)	MIoU(%)	MPA(%)	MIoU(%)
Two-stream	61.52	58.23	66.48	62.73
Basic SegNet	63.77	62.32	56.89	55.64
Proposed method	<b>83.35</b>	<b>75.06</b>	<b>75.67</b>	<b>69.15</b>



# معماری M-SEGNET

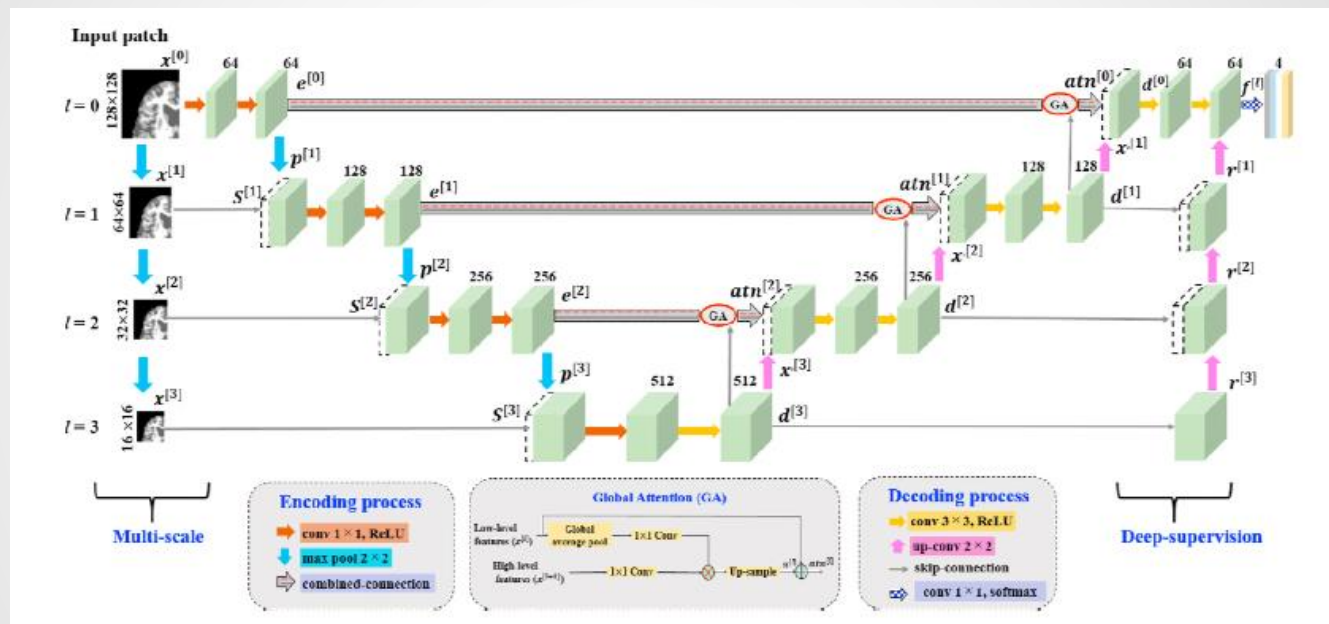
## معرفی معماری



در این مقاله یک معماری جدید M-SegNet با توجه جهانی برای بخش بندی MRI مغز پیشنهاد شده است. معماری M-SegNet پیشنهادی شامل ورودی چند مقیاسی و نظارت عمیق است، به ویژه با مکانیزم توجه جهانی جدید، اندازه‌های مختلف هسته‌های پیچشی در بلوک‌های رمزگذار-رمزگشا، و اتصالات ترکیبی برای تقسیم‌بندی بافت های مغز در MRI ورودی نمونه برداری شده استفاده شده است.



## معرفی شبکه



معماری پیشنهادی از شبکه M-net الهام گرفته شده است، که در ابتدا برای حذف نویز و رنگ آمیزی تصاویر اثر انگشت به طور همزمان استفاده می شد. همانطور که نشان داده شده است، شبکه پیشنهادی دارای یک ساختار مبتنی بر رمزگشا-رمزگشا است که از یک ویژگی کلی M-net، مانند ورودی چند مقیاسی و نظارت عمیق، مکانیزم توجه جهانی جدید، اتصالات ترکیبی، و اندازه‌های مختلف هسته‌های پیچشی تشکیل شده است.



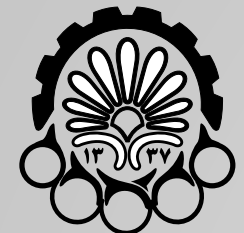
## ارزیابی شبکه

شبکه پیشنهادی بر روی ۲ مجموعه داده مورد بررسی قرار گرفته است : مجموعه‌ای از مطالعات تصویربرداری با دسترسی باز و مجموعه‌های داده مخزن تقسیم‌بندی مغز اینترنتی. در طول آموزش، نزول گرادیان تصادفی را با تکانه ۰/۹۹، نرخ یادگیری ۰/۰۰۱، تقسیم اعتبار ۰/۲، و تعدادی دوره از ۱۰ اتخاذ شده است. برای ارزیابی اثربخشی روش پیشنهادی، از ضریب تشابه دایس، شاخص جاکارد و فاصله هاسدورف برای مقایسه خروجی‌های تقسیم‌بندی با تصویر مرجع استفاده شده است. برای مقایسه نتایج تقسیم‌بندی معماری‌های شبکه مختلف، آزمایش‌هایی را بر روی مدل‌های SegNet، U-net، M-net، U-net++، CE-Net و CFPNet-M تحت شرایط آزمایشی مشابه انجام شده است. در نهایت، روش پیشنهاد شده با توجه به معیارهای ارزیابی مانند DSC و JI برای تقسیم‌بندی MRI مغز به مناطق CSF، GM، و WM بهبود قابل‌توجهی نشان داد و دارای مقادیر متوسط DSC و JI به ترتیب ۰/۹۶ و ۰/۹۲ بود.

$$DSC(g, g') = \frac{2|g \cap g'|}{|g| + |g'|}$$

$$JI(g, g') = \frac{|g \cap g'|}{|g \cup g'|}$$

$$d(g, g') = \max \left\{ \max_{p \in g} \min_{q \in g'} |q - p|, \max_{q \in g'} \min_{p \in g} |p - q| \right\}$$

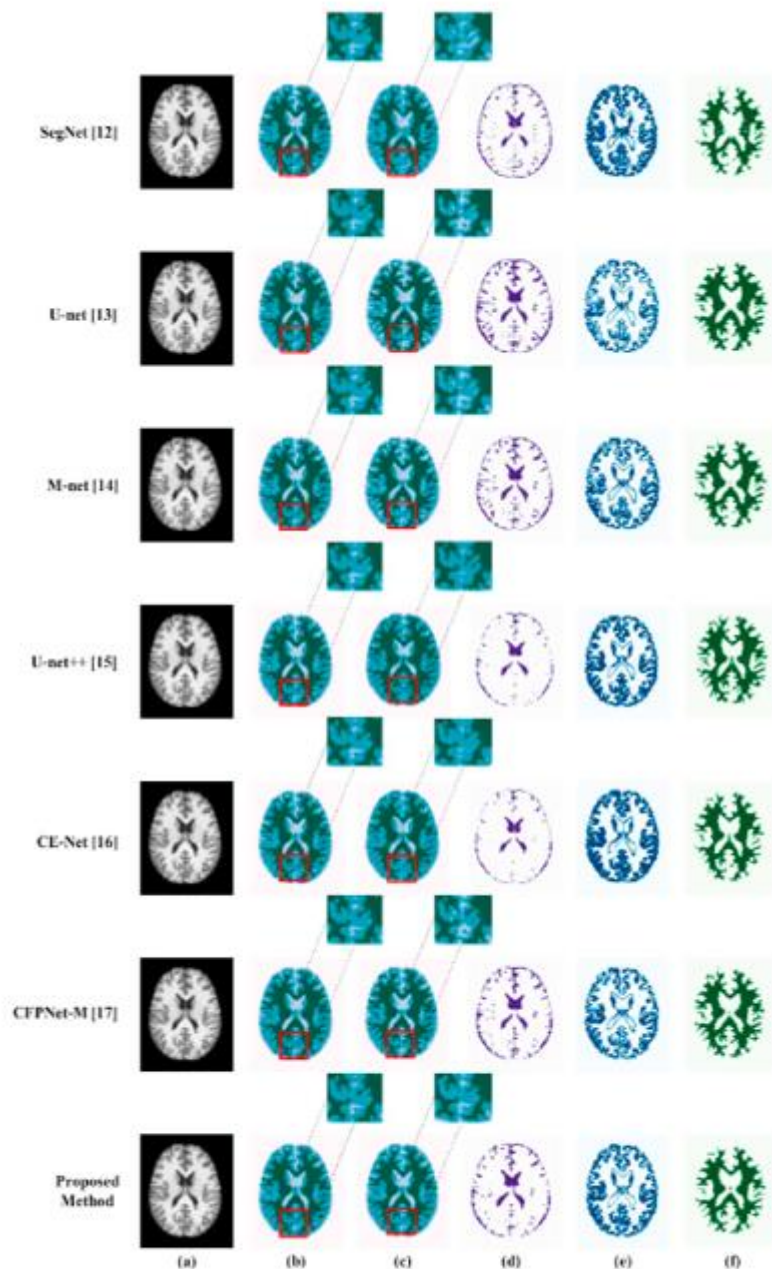
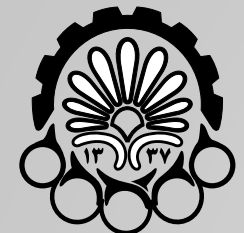


Methods	WM			GM			CSF		
	DSC	JI	HD	DSC	JI	HD	DSC	JI	HD
<b>SegNet [12]</b>	0.89±0.087	0.80±0.096	4.74±0.077	0.86±0.069	0.75±0.089	4.69±0.053	0.85±0.048	0.74±0.064	4.12±0.079
<b>U-net [13]</b>	0.93±0.059	0.87±0.068	4.16±0.064	0.92±0.048	0.85±0.067	4.24±0.046	0.90±0.056	0.82±0.075	3.82±0.039
<b>M-net [14]</b>	0.94±0.068	0.89±0.076	4.02±0.023	0.93±0.055	0.87±0.074	4.11±0.077	0.92±0.044	0.85±0.064	3.79±0.043
<b>U-net + + [15]</b>	0.95±0.075	0.90±0.082	3.78±0.048	0.94±0.035	0.89±0.052	3.84±0.025	0.93±0.039	0.87±0.054	3.56±0.036
<b>CE-Net [16]</b>	0.95±0.039	0.90±0.054	3.65±0.050	0.95±0.042	0.90±0.061	3.57±0.044	0.93±0.043	0.87±0.057	3.21±0.061
<b>CFPNet-M [17]</b>	0.94±0.045	0.89±0.059	4.21±0.023	0.93±0.049	0.87±0.065	4.22±0.077	0.92±0.032	0.85±0.049	3.47±0.043
<b>Proposed</b>	<b>0.96±0.030</b>	<b>0.92±0.053</b>	<b>3.28±0.041</b>	<b>0.96±0.033</b>	<b>0.92±0.048</b>	<b>3.25±0.026</b>	<b>0.95±0.029</b>	<b>0.90±0.042</b>	<b>3.08±0.032</b>
<b>Coronal plane</b>									
<b>SegNet [12]</b>	0.87±0.078	0.77±0.092	5.21±0.023	0.85±0.044	0.74±0.063	5.49±0.053	0.83±0.069	0.71±0.086	5.87±0.084
<b>U-net [13]</b>	0.94±0.065	0.89±0.079	4.88±0.042	0.93±0.057	0.87±0.067	4.95±0.042	0.92±0.063	0.85±0.079	5.34±0.073
<b>M-net [14]</b>	0.94±0.068	0.89±0.082	4.33±0.066	0.92±0.031	0.85±0.043	4.33±0.038	0.92±0.034	0.85±0.051	4.90±0.032
<b>U-net + + [15]</b>	0.94±0.066	0.89±0.073	4.05±0.047	0.93±0.042	0.87±0.057	4.29±0.044	0.93±0.048	0.87±0.065	4.72±0.043
<b>CE-Net [16]</b>	0.95±0.031	0.90±0.046	3.98±0.076	0.94±0.038	0.89±0.050	4.17±0.071	0.93±0.039	0.87±0.049	4.17±0.050
<b>CFPNet-M [17]</b>	0.94±0.055	0.89±0.068	4.29±0.066	0.92±0.046	0.85±0.059	4.59±0.038	0.92±0.041	0.85±0.055	4.45±0.032
<b>Proposed</b>	<b>0.96±0.024</b>	<b>0.92±0.038</b>	<b>3.43±0.046</b>	<b>0.95±0.024</b>	<b>0.90±0.036</b>	<b>3.48±0.066</b>	<b>0.94±0.032</b>	<b>0.89±0.048</b>	<b>3.64±0.036</b>
<b>Sagittal plane</b>									
<b>SegNet [12]</b>	0.88±0.036	0.79±0.054	5.53±0.027	0.85±0.083	0.74±0.095	5.26±0.033	0.84±0.060	0.72±0.077	5.69±0.088
<b>U-net [13]</b>	0.94±0.058	0.89±0.070	5.11±0.030	0.92±0.064	0.85±0.080	5.11±0.026	0.93±0.058	0.87±0.075	5.21±0.079
<b>M-net [14]</b>	0.94±0.038	0.89±0.055	5.34±0.046	0.92±0.053	0.85±0.070	4.67±0.026	0.93±0.049	0.87±0.059	5.04±0.082
<b>U-net + + [15]</b>	0.95±0.050	0.90±0.066	4.46±0.031	0.94±0.038	0.89±0.049	4.32±0.019	0.94±0.053	0.89±0.068	4.56±0.041
<b>CE-Net [16]</b>	0.95±0.043	0.90±0.064	4.13±0.020	0.94±0.045	0.89±0.058	4.25±0.034	0.94±0.051	0.89±0.064	4.28±0.055
<b>CFPNet-M [17]</b>	0.94±0.041	0.89±0.059	5.24±0.052	0.92±0.044	0.85±0.060	4.35±0.037	0.93±0.042	0.87±0.055	5.23±0.049
<b>Proposed</b>	<b>0.95±0.029</b>	<b>0.90±0.047</b>	<b>3.68±0.035</b>	<b>0.95±0.021</b>	<b>0.90±0.035</b>	<b>3.16±0.042</b>	<b>0.95±0.036</b>	<b>0.90±0.047</b>	<b>3.79±0.027</b>

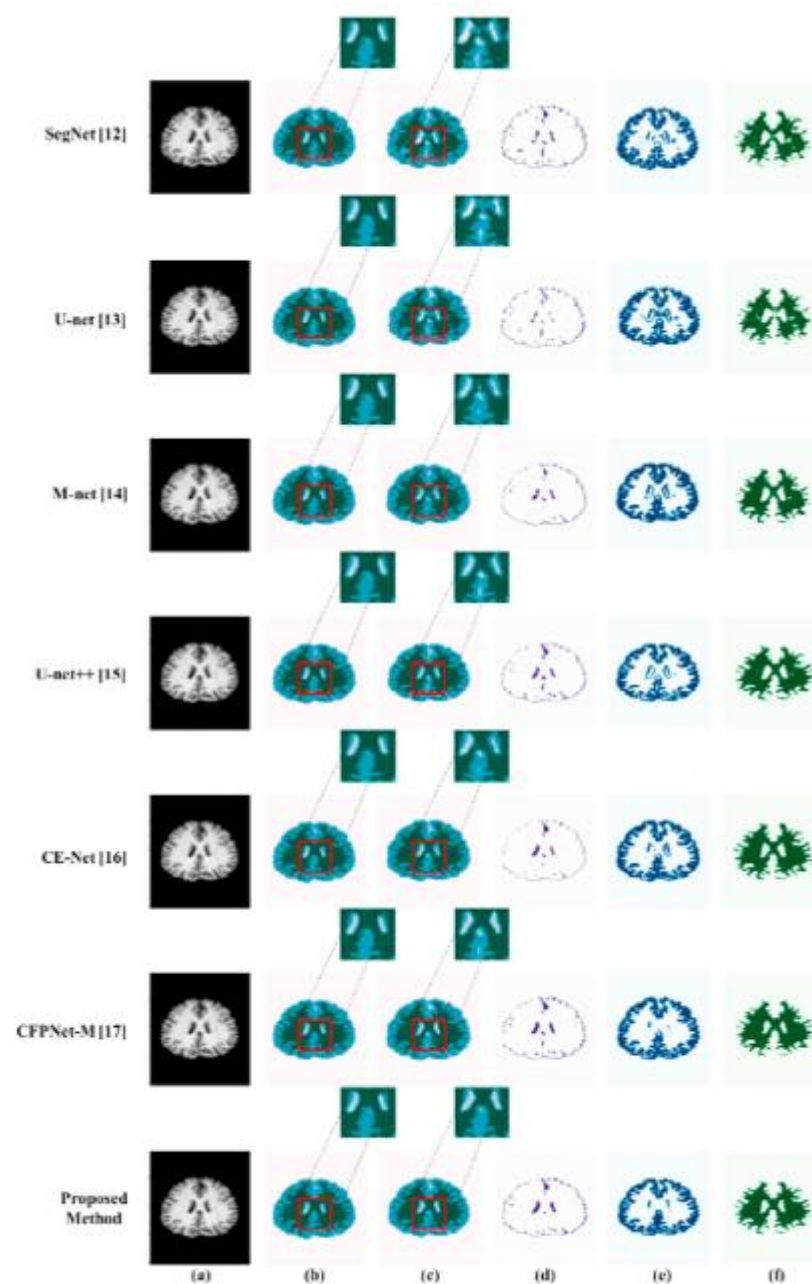


Methods	WM			GM			CSF		
	DSC	JI	HD	DSC	JI	HD	DSC	JI	HD
<b>SegNet [12]</b>	0.72±0.062	0.56±0.075	6.51±0.65	0.75±0.049	0.60±0.058	6.53±0.91	0.68±0.079	0.52±0.091	6.96±0.46
<b>U-net [13]</b>	0.89±0.048	0.80±0.064	5.14±0.51	0.91±0.017	0.83±0.023	4.87±0.51	0.84±0.065	0.72±0.079	5.24±0.31
<b>M-net [14]</b>	0.90±0.043	0.82±0.058	4.76±0.39	0.92±0.053	0.85±0.028	4.45±0.65	0.84±0.039	0.72±0.048	4.84±0.18
<b>U-net + + [15]</b>	0.88±0.055	0.79±0.069	5.37±0.36	0.89±0.037	0.80±0.049	5.17±0.29	0.83±0.058	0.71±0.072	5.34±0.64
<b>CE-Net [16]</b>	0.89±0.052	0.80±0.063	4.98±0.84	0.90±0.068	0.82±0.083	4.95±0.38	0.82±0.037	0.69±0.051	4.74±0.93
<b>CFPNet-M [17]</b>	0.89±0.056	0.80±0.071	4.37±0.55	0.91±0.033	0.83±0.071	4.66±0.67	0.84±0.050	0.72±0.066	4.54±0.59
<b>Proposed</b>	<b>0.90±0.038</b>	<b>0.82±0.049</b>	<b>4.59±0.64</b>	<b>0.92±0.055</b>	<b>0.85±0.028</b>	<b>4.43±0.47</b>	<b>0.84±0.032</b>	<b>0.72±0.055</b>	<b>4.42±0.24</b>
<b>Coronal plane</b>									
<b>SegNet [12]</b>	0.70±0.061	0.54±0.081	6.32±0.82	0.73±0.052	0.57±0.062	6.21±0.84	0.66±0.071	0.49±0.086	6.84±0.75
<b>U-net [13]</b>	0.88±0.052	0.79±0.067	5.45±0.67	0.90±0.044	0.82±0.056	5.17±0.38	0.83±0.032	0.71±0.043	5.54±0.47
<b>M-net [14]</b>	0.89±0.066	0.80±0.078	4.61±0.21	0.91±0.035	0.83±0.046	4.56±0.19	0.84±0.075	0.72±0.091	4.83±0.25
<b>U-net + + [15]</b>	0.88±0.041	0.79±0.063	5.21±0.39	0.91±0.063	0.83±0.077	5.24±0.24	0.82±0.044	0.69±0.067	5.73±0.39
<b>CE-Net [16]</b>	0.89±0.034	0.80±0.055	4.89±0.21	0.90±0.049	0.82±0.068	5.98±0.93	0.83±0.056	0.71±0.072	5.21±0.20
<b>CFPNet-M [17]</b>	0.88±0.042	0.79±0.062	4.48±0.66	0.89±0.038	0.80±0.056	4.65±0.78	0.83±0.059	0.71±0.077	4.78±0.75
<b>Proposed</b>	<b>0.91±0.026</b>	<b>0.83±0.043</b>	<b>4.39±0.42</b>	<b>0.92±0.071</b>	<b>0.85±0.040</b>	<b>4.52±0.36</b>	<b>0.83±0.033</b>	<b>0.71±0.047</b>	<b>4.26±0.52</b>
<b>Sagittal plane</b>									
<b>SegNet [12]</b>	0.71±0.043	0.55±0.059	6.49±0.61	0.74±0.073	0.59±0.089	6.36±0.76	0.65±0.083	0.48±0.095	6.99±0.41
<b>U-net [13]</b>	0.86±0.049	0.75±0.062	5.75±0.37	0.89±0.046	0.80±0.057	5.77±0.21	0.80±0.071	0.67±0.089	5.83±0.15
<b>M-net [14]</b>	0.87±0.036	0.77±0.048	4.89±0.14	0.90±0.062	0.82±0.076	5.42±0.06	0.81±0.060	0.68±0.073	4.98±0.09
<b>U-net + + [15]</b>	0.85±0.033	0.74±0.049	4.57±0.54	0.88±0.067	0.79±0.081	4.96±0.22	0.79±0.049	0.65±0.068	5.60±0.44
<b>CE-Net [16]</b>	0.86±0.054	0.75±0.065	5.34±0.66	0.89±0.051	0.80±0.067	5.86±0.55	0.79±0.039	0.65 p0.052	5.25±0.37
<b>CFPNet-M [17]</b>	0.86±0.046	0.75±0.062	4.76±0.57	0.88±0.043	0.79±0.059	5.72±0.66	0.80±0.041	0.67±0.053	4.56±0.48
<b>Proposed</b>	<b>0.89±0.032</b>	<b>0.80±0.049</b>	<b>4.46±0.52</b>	<b>0.90±0.029</b>	<b>0.82±0.042</b>	<b>5.42±0.31</b>	<b>0.82±0.020</b>	<b>0.69±0.035</b>	<b>4.31±0.32</b>

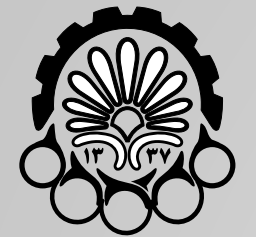




نتیجه بررسی در مجموعه داده OASIS

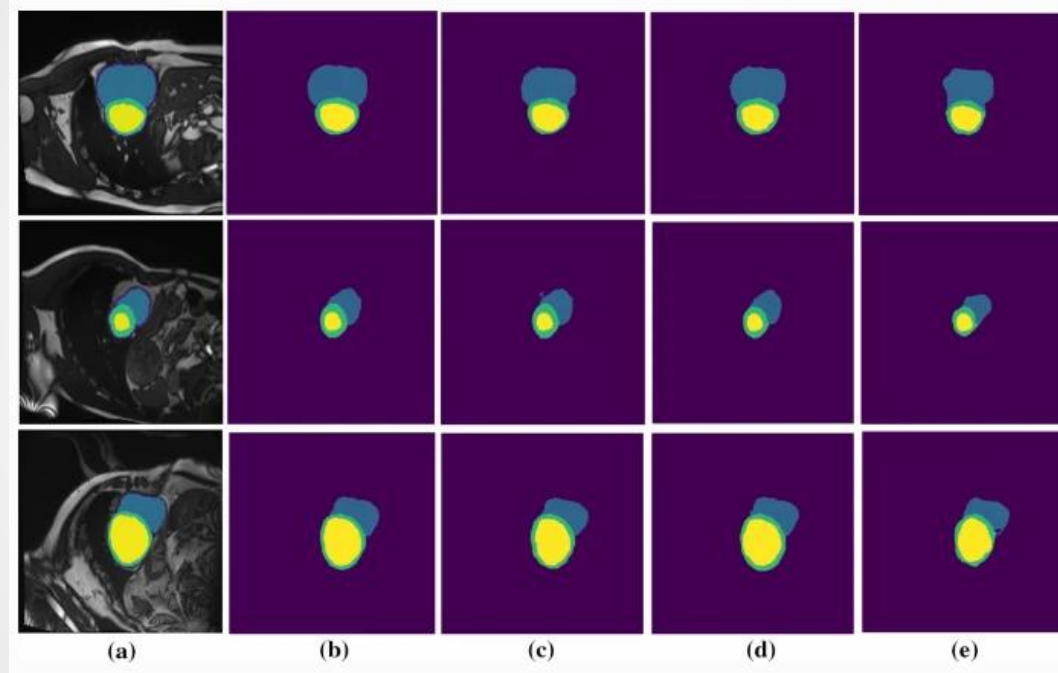


نتیجه بررسی در مجموعه داده IBSR



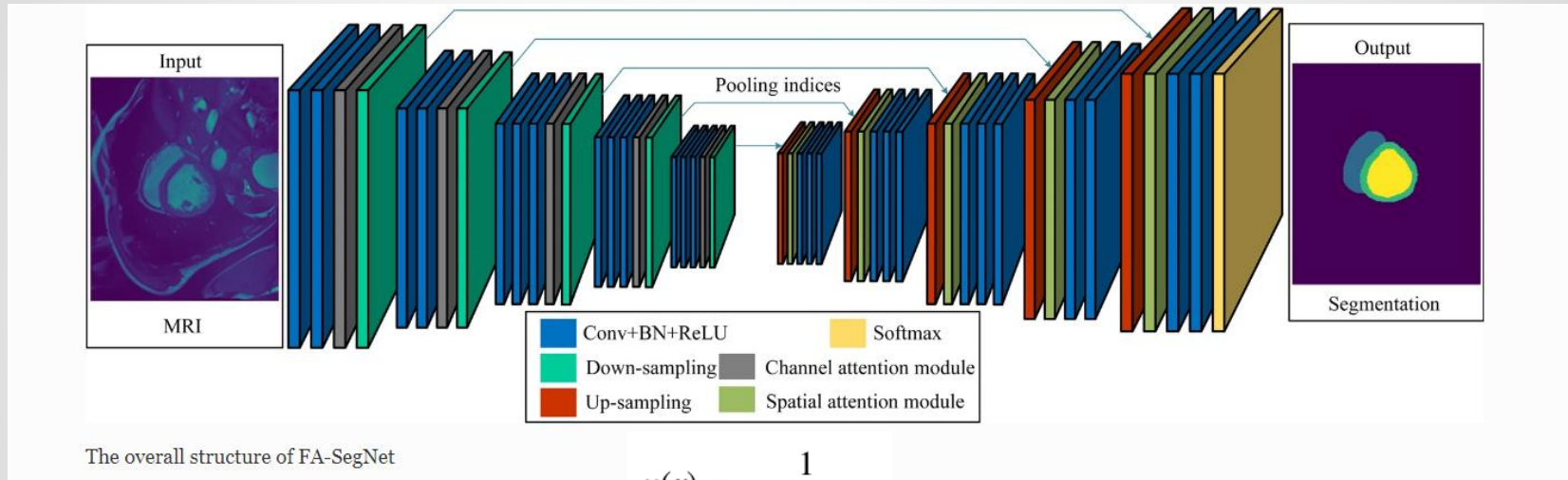
# معماری FA-SEGNET

## معرفی معماری



با هدف تقسیم‌بندی تصاویر پزشکی با تشخیص کم و نویز پس‌زمینه بالا، یک مدل تقسیم‌بندی تصویر شبکه عصبی پیچشی عمیق بر اساس مکانیسم توجه فازی پیشنهاد شده است که FA-SegNet نامیده می‌شود. SegNet را به عنوان چارچوب اصلی می‌گیرد. در واحد فرو نمونه برداری برای استخراج ویژگی تصویر، یک واحد توجه کانال فازی اضافه شده است تا تشخیص مناطق هدف مختلف را تقویت کند. در واحد فرا نمونه‌برداری برای بازیابی اندازه تصویر و ادغام ویژگی‌های چند مقیاسی، یک واحد توجه فضایی فازی اضافه می‌شود تا از دست رفتن جزئیات تصویر را کاهش دهد و حوزه دریافتی را گسترش دهد.

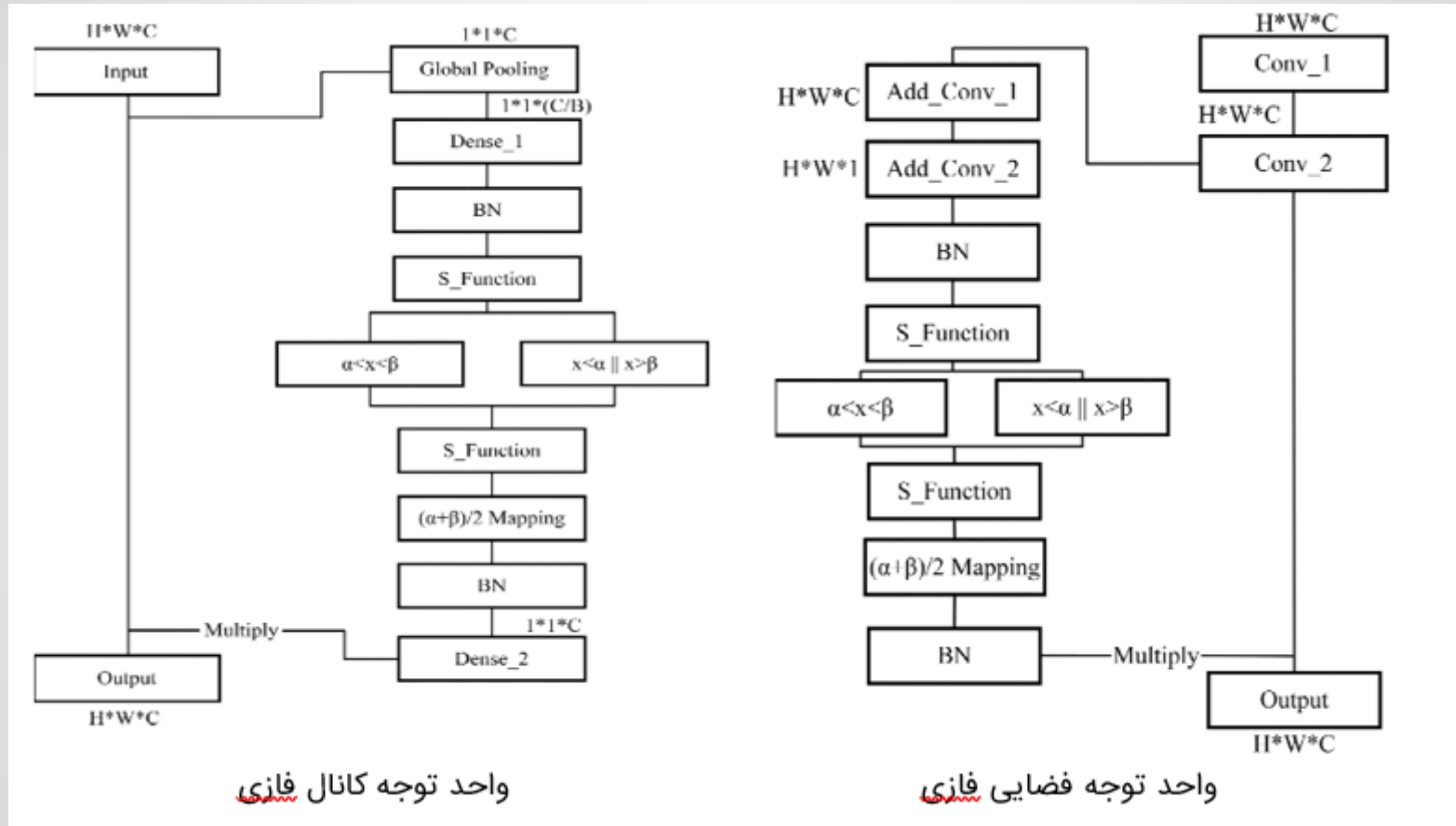
## معرفی شبکه



$$\mu(x) = \frac{1}{1 + e^{ax+b}}.$$

در مرحله فرو نمونه برداری ، با استفاده از تابع عضویت تعریف شده ، مقدار قطعیت هر کانال مورد بررسی قرار میگیرد. مقادیر نزدیک به ۰.۵ قطعیت کمی دارند و مقدار دور از ۰.۵ از قطعیت بالایی برخوردار هستند. طبق آزمایشات انجام شده این کانال ها همان تصاویری را که اهمیت بالایی دارند را ایجاد میکنند. پس در این مکانیزم توجه ، وزن این کانال های با اهمیت افزایش میابد.

در مرحله فرا نمونه برداری ، از همین تابع عضویت استفاده میشود. عدم قطعیت عضویت به عنوان وزن برای ارزیابی اهمیت پیکسل ها استفاده می شود. وزن پیکسل های با عدم قطعیت کم نقش مهمتری ایفا میکند. در واحد فرا نمونه برداری، یک واحد توجه فضایی فازی اضافه می شود تا از دست رفتن ویژگی های جزئیات تصویر را کاهش دهد و حوزه دریافتی را گسترش دهد.



مکانیسم توجه فازی به خوبی اطلاعات زمینه فعال شده با پیچیدگی اصلی را حفظ می‌کند. این می‌تواند اطلاعات ویژگی‌های منطقه و لبه را در وظیفه تقسیم‌بندی بهتر استخراج کند و تداخل اطلاعات مستقل از کار را سرکوب کند.



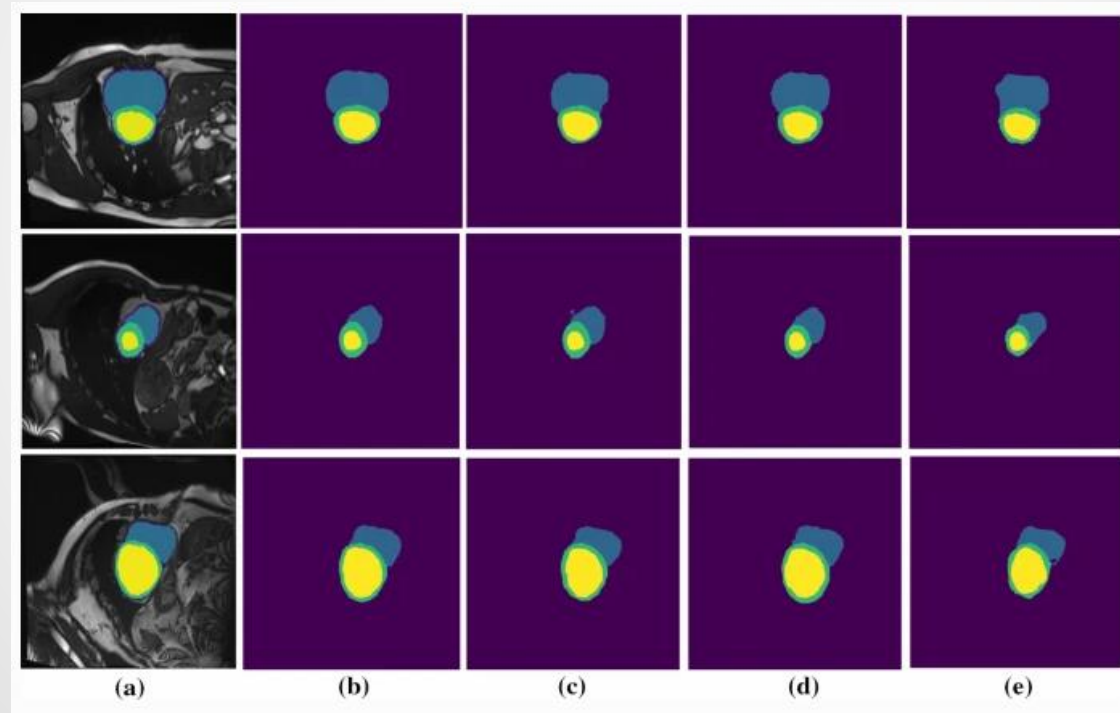
## ارزیابی شبکه

در آموزش FA-SegNet از تلفات آنتروپی متقاطع به عنوان تابع ضرر و از الگوریتم انتشار برگشتی برای یادگیری پارامترهای شبکه استفاده می شود. هنگام ارزیابی عملکرد FA-SegNet پیشنهاد شده در این مقاله، معیارهای ارزیابی تقسیم بندی تصویر عبارتند از: دقت پیکسل (PA)، دقت میانگین پیکسل (MPA)، تقاطع بر روی اتحاد (IOU)



دانشگاه صنعتی امیرکبیر  
( پلی تکنیک تهران )

نتایج کیفی چهار روش مختلف در تصویر زیر نشان داده شده است که در آن (a) تصویر اصلی MRI با علائم تقسیم بندی استاندارد است. (b) نتیجه تقسیم بندی FA-SegNet است که در این مقاله پیشنهاد شده است. (c) نتیجه تقسیم بندی مدل اصلی SegNet است. (d) نتیجه تقسیم بندی مدل U-net است. (e) نتیجه تقسیم بندی شبکه کاملاً پیچشی است.



از نتایج تجربی جدول زیر می توان دریافت که FA-SegNet بهترین نتایج را در PA، MPA و MIoU در مقایسه با سایر الگوریتم ها به دست می آورد. به طور خاص، در مقایسه با روش SegNet اصلی، مدل پیشنهادی ۰.۰۴، ۰.۰۱ و ۰.۰۵ به ترتیب در PA، MPA و MIoU بالاتر است، که ثابت می کند که واحد توجه فازی پیشنهادی تأثیر خوبی در استخراج ویژگی عمیق دارد.

Model	PA	MPA	MIoU
FCN	0.8740	0.8534	0.6340
U-net	0.9372	0.9216	0.8460
SegNet	0.9022	0.9157	0.8165
Proposed	0.9475	0.9247	0.8618

## مراجع و منابع

- [۱]. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation", 2017
- [۲]. "P\_Segnet and NP\_Segnet: New Neural Network Architectures for Cloud Recognition of Remote Sensing Images", 2019
- [۳]. "A Deep-Learning-Based Multiple Defect Detection Method for Tunnel Lining Damages", 2019
- [۴]. "A novel M-SegNet with global attention CNN architecture for automatic segmentation of brain MRI", Yamanakkanavar and Lee, 2021
- [۵]. " An FA-SegNet Image Segmentation Model Based on Fuzzy Attention and Its Application in Cardiac MRI Segmentation." , Yang, R., Yu, J., Yin, J. et al. , 2022

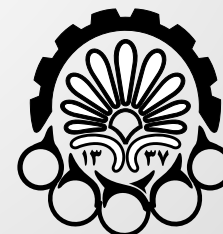
با تشکر فراوان

مهدی فیروزبخت

۴۰۰۱۳۱۰۲۷

[mahdifiruzbakht23@gmail.com](mailto:mahdifiruzbakht23@gmail.com)

[mahdi.firuzbakht@aut.ac.ir](mailto:mahdi.firuzbakht@aut.ac.ir)



دانشگاه صنعتی امیرکبیر  
( پلی تکنیک تهران )