ترجمه فا

TarjomeFa.Com

ارائه شده توسط:

سایت ترجمه فا

مرجع جدیدترین مقالات ترجمه شده

از نشریات معتبر

یادگیری عمیق و کاربرد های آن در زیست پزشکی

چکیده

پیشرفت های انتزاعی در زیست و تکنولوژی های زیستی، حجم بسیار زیادی از داده های زیستی و فیزیولوژیک را ایجاد کرده است، از جمله تصویر های پزشکی، الکتروانسفالوگرافی، نقشه های ژنوم و توالی های پروتئینی. یادگیری با استفاده از این داده ها منجر به تسهیل درک ما نسبت به سلامت و بیماری های انسان می شود. الگوریتم های مبتنی بر یادگیری عمیق که از شبکه های عصبی مصنوعی توسعه پیدا کرده اند، توانایی بالایی برای استخراج کردن ویژگی و الگوهای یادگیری از داده های پیچیده را از خودشان نشان داده اند. هدف این مقاله فراهم کردن مروری بر روی تکنیک های یادگیری عمیق و بعضی از جدید ترین کاربرد های آن ها در زمینه ی پزشکی می باشد. ما نخست توسعه ی شبکه های عصبی و یادگیری عمیق را ارائه می کنیم. سپس دو بخش اصلی از یادگیری عمیق، یعنی معماری یادگیری عمیق و بهینه سازی مدل را ارائه می کنیم. در نهایت بعضی از نمونه های کاربردی روش یادگیری عمیق ارائه می گردد ، از جمله طبقه بندی تصاویر پزشکی، تحلیل توالی ها، و همچنین طبقه بندی و پیش بینی کردن ساختار پروتئین ها. در نهایت، ما دیدگاه خودمان را در رابطه با جهت آتی در زمینه ی یادگیری عمیق را ارائه می کنیم.

کلمات کلیدی : یادگیری عمیق، داده های بزرگ، انفورماتیک زیستی، انفورماتیک پزشکی ؛ تصویر های پزشکی، توالی با خروجی بالا

مقدمه

یادگیری عمیق یکی از روش های جدید با رشد سریع در زمینه ی یادگیری ماشین می باشد. در این روش تلاش می شود تا استخراج ویژگی ها و اطلاعات مفید از داده هایی با مقیاس بزرگ، با استفاده از شبکه های عصبی عمیق چند لایه ای (DNN ها) ، استخراج شود تا بتوان از داده هایی مانند تصاویر، صداها و متن ها، اطلاعات مفید به دست آورد

[۱]. یادگیری عمیق به صورت عمومی دارای دو ویژگی می باشد : ۱)لایه های چندگانه از واحد های پردازش غیر خطی و ۲) یادگیری با سرپرست یا بدون سرپرست از نمایه ی ویژگی ها در هر لایه [۱].قالب کاری اولیه برای یادگیری عمیق بر اساس شبکه های عصبی مصنوعی (ANN ها) در دهه ی ۱۹۸۰ ایجاد شد [۲] ، در حالی که تاثیر واقعی این روش های یادگیری عمیق در سال ۲۰۰۶ نمایان شد [۳و۴]. از آن زمان تا کنون، یادگیری عمیق در گستره ی زمینه های مختلف، شامل شناسایی خودکار گفتار، شناسایی تصویر، پردازش طبیعی زبان، شناسایی دارو و انفورماتیک زیستی مورد استفاده قرار گرفته است [۵-۷].

در دهه های اخیر، رشد گسترده ای در زمینه ی داده های پزشکی مانند توالی های ژنی، ساختار پروتئین ها، و تصویر های پزشکی ایجاد شده است که به دلیل پیشرفت در تکنولوژی هایی با کارایی بالا بوده است. این طوفان گسترده از داده های بزرگ زیستی، منجر به الزام استفاده از ابزار محاسباتی موثر و کارآمد شده است تا بتوانیم این چنین داده هایی را ذخیره سازی، تحلیل و تفسیر کنیم. قالب های کاری یادگیری عمیق ، مسائل جدیدی را نمایان کرده اند. هدف این مقاله فراهم کردن یک مرور در رابطه با یادگیری عمیق و بعضی از جدید ترین کاربرد های یادگیری عمیق در زمینه ی پزشکی و انفورماتیک زیستی می باشد.ما امید داریم که این مقاله بتواند برای خوانندگان ، مروری مناسب در رابطه با یادگیری عمیق ایجاد کرده و کاربرد های آن برای تحلیل داده های زیستی را نمایان کند.

توسعه ی ANN ها

به عنوان مبنایی برای یادگیری عمیق، شبکه های عصبی مصنوعی یا ANN ها با الهام از روند های زیستی در دهه ی ۱۹۶۰ شکل گرفته اند . در این دهه بود که مشخص شد وقتی گربه ها اشیای مختلف را می دیدند، سلول های مختلف در کورتکس بصری فعال می شدند و بر همین اساس، ایده ی شبکه های عصبی مصنوعی شکل گرفت. این مطالعه ها نشان داد که ارتباط هایی بین چشم و سلول های کورتکس بصری وجود داشت و اطلاعات به صورت لایه به لایه در سیستم بصری پردازش می شد. شبکه های عصبی مصنوعی با متصل کردن نورون های مصنوعی در لایه های شبکه، تلاش کردند تا این روند را تقلید کنند و بتوانند اطلاعات و ویژگی های مفید از اشیا را درک کنند. اما تحقیقات در رابطه با ANN ها بعد از دهه ی ۱۹۶۰ راکد شد، زیرا نتایجی که از ساختار های ضعیف اولیه به دست آمده بود کارایی

چندانی نداشت و ظرفیت محاسباتی کامپیوتر ها در آن زمان پایین بود و نمی توانست عملکرد خوبی را از شبکه های عصبی به دست بیاورد.

با وجود پیشرفت های انجام شده در ظرفیت کامپیوتر ها و روش های محاسبه ی آن ها، شبکه های عصبی با استفاده از روش های پس انتشار موثر (BP) توانستند مطالعات مختلف در رابطه با شناسایی الگو را با سهولت انجام دهند. در یک شبکه ی عصبی با تکنیک پس انتشار، طبقه بندی نخست توسط مدل های ANN پردازش میشد و سپس وزن های هر لایه با ارزیابی تفاوت بین مقادیر پیش بینی شده و طبقه بندی های صحیح، تنظیم می شد. با وجود این که روش BP به این شبکه کمک می کرد تا از طریق نزول گرادیان مقدار خطا را به حداقل برساند، به نظر می رسید که این روش تنها برای یک سری از ANN ها قابلیت استفاده را داشت. از طریق بهبود گرادیان های شدید تر با BP ، روش های یادگیری مختلف مانند اندازه ی حرکت، نرخ یادگیری تطبیقی، روش حداقل مربعات و روش شبه نیوتونی و گرادیان ترکیبی (CG) شکل گرفت . اما، به دلیل پیچیدگی شبکه های عصبی، دیگر روش های ساده مانند ماشین بردار پشتیبان (SVM ها) ، جنگل تصادفی ، و الگوریتم نزدیک ترین همسایه ها در فاصله ی k (k-NN) به صورت تدریجی جای ANN ها را از نظر محبوبیت گرفتند (شکل ۱).

توسعه ی یادگیری عمیق

یک شبکه ی عصبی با لایه های پنهان بیشتر، ظرفیت بسیار بیشتری برای استخراج داده ها فراهم می کند. اما یک ANN معمولا به حالت بهینه ی محلی می رسد و وقتی که ساختار های عمیق و پیچیده ای دارد، با نشر گرادیان ر به رو می شود. گرادیانی که به صورت پس انتشار به صورت بازگشت می کند، در راستای لایه های مختلف کم کم مقدارش کم می شود و در نتیجه اصلاح کمی بر روی وزن ها در لایه های نزدیک به ورودی ایجاد می شود. در نتیجه، یک شبکه ی خود رمز نگار عمیق با تمرین از قبل مختص هر لایه (AE) در این زمینه ارائه شده است که منجر به شکل گیری توسعه ای جدید از ANN ها می شود (شکل ۱). در این شبکه، هر لایه با حداقل کردن تفاوت بین داده های اصلی و داده های بازسازی شده، تمرین میابد. این تمرین در هر لایه منجر می شود که مانع نشر گرادیان از بین برود و همچنین انتخاب وزن بندی بهتری برای شبکه های عصبی عمیق (DNN ها) فراهم می کند، در نتیجه داده های

بازسازی شده دیگر در حالت بهینه ی محلی گیر نمی کنند زیرا این حالت بهینه ی محلی معمولا در اثر انتخاب اتفاقی وزن های اولیه ایجاد می شود. به هعاوه، استفاده از واحد های پردازش گرافیکی (GPU ها) باعث شده که محققان مجدد به مبحث یادگیری عمیق علاقه مند شوند.

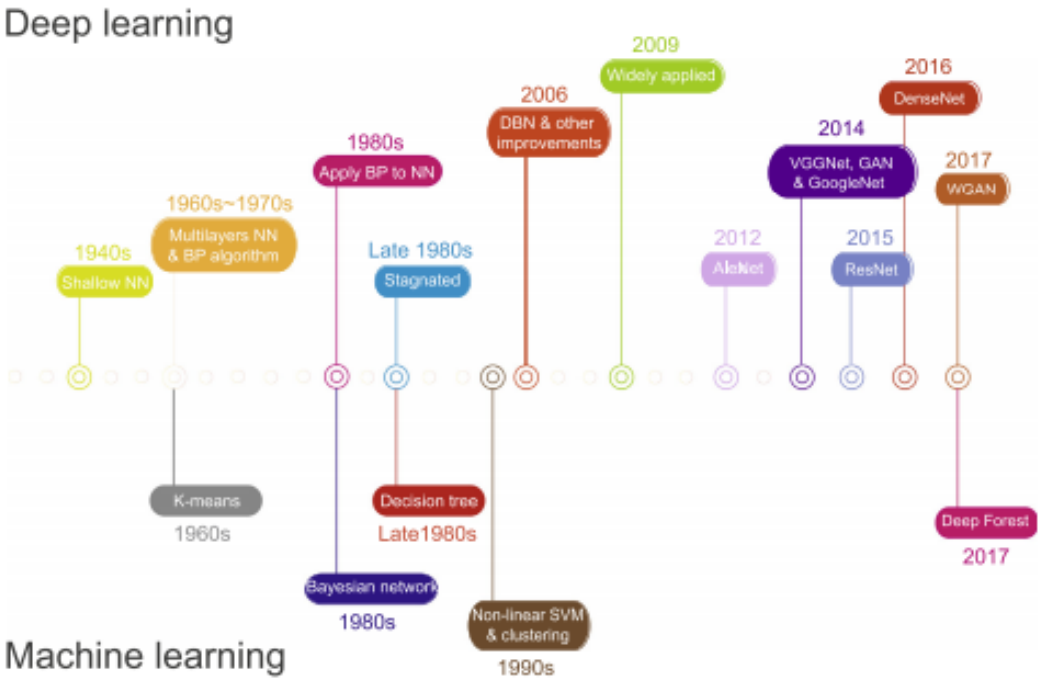
با تمرکز و توجه بیشتر به این روش و تلاش بیشتر، یادگیری عمیق در سالهای اخیر پیشرفت زیادی داشته است و به صورت گسترده در صنایع مختلف مورد استفاده قرار گرفته است. به عنوان مثال، شبکه های باور عمیق (DBN ها) و پشته های ماشین های محدود بلتزمن (RBM) ، در شناسایی تصویر و صدا و پردازش طبیعی زبان مورد استفاده قرار گرفته اند. شبکه های عصبی پیچشی (CNN ها) که بیشتر برای تقلید از درک حیوانات نسبت به اشیا شکل گرفته اند، به صورت گسترده در زمینه ی شناسایی تصویر، بخش بندی تصویر، شناسایی ویدئو و پردازش زبان طبیعی مورد استفاده قرار گرفته اند. شبکه های عصبی بازگشتی (RNN ها) هم دسته ای دیگر از شبکه های عصبی هستند که رفتار پویایی از خودشان نشان می دهند و در این شبکه ها، نورون های عصبی با گام های زمانی با هم در ارتباط هستند. این RNN ها مهم ترین ابزار برای کار با داده های متوالی هستند و در پردازش زبان طبیعی و شناسایی دست خط، به صورت رایج مورد استفاده قرار می گیرند. بعد از این موارد، انواع مختلف از AE ها شامل AE های پشته ای (SAE ها) ، AE های رفع نویز و دیگر نمونه ها نیز در شبکه های عمیق با تمرین از قبل، رواج پیدا کرده اند.

با وجود این که کاربرد های یادگیری عمیق بیشتر متمرکز بر روی شناسایی تصویر ، صدا و تحلیل آن ها و پردازش زبان طبیعی بوده است، اما جدیدا کاربرد های نوینی برای این شبکه ها در علوم زیستی ایجاد شده است که در ادامه به صورت مفصل به آن ها خواهیم پرداخت.

توصیفی خلاصه در رابطه با یادگیری عمیق

با وجود این که فرضیات زیر لایه ای و نظریات مختلفی وجود دارد، اما ایده ی اصلی و روند های استخراج ویژگی در بیشتر معماری های NN عمیق (DNN ها) مشابه هم هستند. در انتقال اطلاعات رو به جلو، این شبکه بر اساس ورودی ها به لایه ی اول فعال می شود سپس این لایه ، این فعال سازی را از طریق ارتباطات وزن بندی شده، به لایه

ی نهایی منتقل می کند و در نتیجه پیش بینی یا نتایج بازسازی را ارائه می کند. در روند انتقال اطلاعات رو به عقب، وزن بندی این اتصالات با حداقل کردن تفاوت بین داده های واقعی و داده های پیش بینی شده، به حداقل می رسد.



شکل ۱ زمان بندی توسعه ی شبکه های عصبی عمیق و الگوریتم های رایج یادگیری ماشین

توسعه ی یادگیری عمیق و شبکه های عصبی در قسمت بالایی تصویر نشان داده شده است و چندین الگوریتم رایج در زمینه ی یادگیری ماشین نیز در قسمت پایین تصویر نشان داده شده است. NN : شبکه های عصبی ؛ BP : شبکه های پس انتشار ؛ DBN : شبکه ی باور عمیق ؛ SVM : ماشین بردار پشتیبان ؛ AE : خود رمز نگار ؛ VAE : AE های متغیر ؛ GAN : شبکه های مولد مجادله ای ؛ WGAN : شبکه های واسرشتین GAN

مفاهیم اولیه

توابع فعال سازی

توابع فعال سازی ، لایه های غیر خطی در شبکه های یادگیری عمیق را ایجاد می کنند و ترکیب این توابع با دیگر لایه ها برای شبیه سازی کردن تبدیل های غیر خطی از ورودی به سمت خروجی مورد استفاده قرار می گیرد. ازین رو، استخراج بهتر ویژگی ها مبتنی بر انتخاب توابع فعال سازی مناسب می باشد. در این قسمت، ما توابع فعال سازی

که به صورت رایج مورد استفاده قرار می گیرند را نشان می دهیم که هر کدام از این توابع با استفاده از g نشان داده شده اند.

تابع سیگموید : $g(a) = \frac{1}{1+e^{-a}}$ که a ورودی از لایه ی اولیه می باشد. یک تابع سیگموید متغیر را بر روی بازه ی ۰ تا ۱ نگاشت می کند و به صورت رایج برای ایجاد کردن یک توزیع برنولی، مورد استفاده قرار می گیرد. مصلا به صورت زیر :

$$\tilde{g} = \begin{cases} 0 & \text{if} \quad g(a) \leqslant 0.5 \\ 1 & \text{if} \quad g(a) > 0.5 \end{cases},$$

تانژانت هایپربولیک : $g(a) = \tan\text{h}(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}}$ ، در این قسمت مشتق g به صورت $g' = 1 - g^2$ محاسبه می شود در نتیجه استفاده از این تابع در الگوریتم های پس انتشار ساده می شود.

تابع سافتمکس : $g(a) = \frac{e^{a_i}}{\sum_j e^{a_j}}$ ، a_n را می توان به عنوان توزیع احتمال روی دسته ها در نظر گرفت در نتیجه خروجی این تابع در لایه ی نهایی مورد استفاده قرار می گیرد.

تابع واحد خطی یک سو ساز (ReLU) $g(a) = \max(0, a)$. این تابع فعال ساز و انواع مختلف از این تابع عملکرد برتری را در بسیاری از موارد از خودش نشان می دهد و تا کنون این تابع رایج ترین تابع در یادگیری عمیق بوده است. این تابع می تواند مشکل نشر گرادیان را هم حل کند.

سافتپلاس : $g(a) = \log(1 + e^a)$. این تابع یکی از انواع ReLU می باشد که یک تخمین روان از ReLU (در این مقاله، لگاریتم همیشه لگاریتم طبیعی می باشد) را ارائه می کند.

یک سو سازی مقدار مطلق : $g(a) = |a|$. این تابع زمانی که لایه های کلی مقادیر میانگین در CNN را استفاده می کنند مفید می باشد، ازین مانع از بین رفتن ویژگی های منفی و مثبت می شود.

مکس اون : $g_i(x) = \max_i (b_i + w_i \cdot x)$ ماتریس وزنی در این تابع آرایشی سه بعدی دارد که ردیف ارایه ی سوم در این تابع مطابق با وزن های اتصال به لایه ی بعدی می باشد.

هدف بهینه سازی

یک هدف بهینه سازی معمولا متشکل از یک تابع ضرر و یک عبارت تنظیم می باشد. تابع ضرر در واقع تفاوت بین خروجی های شبکه را اندازه می گیرد که مبتنی بر پارامتر های مدل $f(x|\theta)$ (θ) و y واقعی مورد انتظار می باشد، یعنی طبقه بندی های صحیح در رابطه با وظیفه ی طبقه بندی و یا سطح صحیح در رابطه با وظیفه ی پیش بینی . اما، یک الگوریتم یادگیری خوب نه تنها در زمینه ی تمرین داده ها، بلکه در قسمت تست داده ها نیز باید عملکرد خوبی را داشته باشد. مجموعه ای از برنامه های راهبردی در این قسمت طراحی شده است تا خطای تست به حداقل برسد که این برنامه ها با نام تکنیک های تنظیم شناخته می شوند. بعضی از این عبارت های تنظیم جریمه هایی را برای پارامتر ها در نظر می گیرند تا مانع شکل گیری مدل های پیچیده بشوند. در این قسمت ، ما به صورت خلاصه توابع ضرر مورد استفاده $L(f(x|\theta), y)$ و عبارت های تنظیم $\Omega(\theta)$ را ارائه می کنیم. هدف بهینه سازی در این قسمت به صورت زیر تعریف می شود :

$$\tilde{L}(X, y, \theta) = L(f(x|\theta), y) + \alpha\Omega(\theta) \qquad (1)$$

که α نشان دهنده ی تعادل این دو بخش می باشد و در عمل، تابع ضرر معمولا بر روی نمونه های تمرینی اتفاقی محاسبه می شود نه بر روی توزیع های داده ها، به دلیل این که توزیع بر روی داده ها معمولا مجهول می باشد.

تابع ضرر

بیشتر DNN ها از آنتروپی های مشترک بین داده های تمرینی و توزیع مدل به عنوان تابع ضرر استفاده می کنند. رایج ترین آنتروپی مشترک مورد استفاده، احتمال لگاریتمی شرطی می باشد : $L(f(x|\theta), y) = -\log P(f = y|x, \theta)$. این یک مجموعه از توابع ضرر متناظر با توزیع y با در نظر داشتن مقدار ورودی x می باشد. در این قسمت، ما توابع ضرری که به صورت رایج مورد استفاده قرار می گیرند و از همین الگو تبعیت می کنند را ارائه می کنیم.

فرض می کنیم که y حالتی پیوسته دارد و بر روی متغیر مشخص x ، دارای توزیع گاوسی می باشد. تابع ضرر به صورت زیر خواهد بود :

$$L(f(x|\theta),y) = -\log\left[\sqrt{\frac{1}{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(y-f)^2\right)\right]$$
$$= \frac{1}{2\sigma^2}(y-f)^2 + \frac{1}{2}\log(2\pi\sigma^2) \qquad (2)$$

که این تابع به همین صورت به عنوان خطای مربع نیز توصیف می شود. خطای مربع رایج ترین تابع ضرر مورد استفاده در دهه ی ۱۹۸۰ بوده است. اما این روش معمولا پارامتر های جریمه ی شدیدی را برای لایه های نهایی ایجاد می کند در نتیجه نرخ همگرایی شبکه را کند می کند.

در صورتی که y از توزیع برنولی تبعیت کند، سپس تابع ضرر به صورت زیر خواهد بود :

$$L(f(x|\theta),y) = -y\log f(x|\theta) - (1-y)\log(1-f(x|\theta)) \qquad (3)$$

که y یک متغیر گسسته بوده و تنها دارای دو مقدار می باشد، مثلا $y \in \{1,2,\ldots,k\}$ و ما می توانیم از مقدار سافت مکس (قسمت توابع فعال سازی رایج را مطالعه کنید) به عنوان احتمال بر روی دسته های مد نظر، استفاده کنیم. سپس تابع ضرر به صورت زیر خواهد بود :

$$L(f(x|\theta),y) = -\log\left(\frac{e^{a_y}}{\sum_j e^{a_j}}\right) = -a_y + \log\left(\sum_j e^{a_j}\right) \qquad (4)$$

عبارت تنظیم

پارامتر تنظیم L^2 رایج ترین پارامتر تنظیم می باشد و منجر به محدب شدن هدف بهینه سزای می شود و در نتیجه یک راه حل راحت برای شکل گیری مینیموم با استفاده از ماتریس هسیان را ایجاد می کند. این پارامتر تنظیم L^2 را می توان به صورت زیر تعریف کرد :

$$\Omega(\theta) = \frac{1}{2}\|\omega\|^2 \qquad (5)$$

که Ω نشان دهنده ی وزن های واحد های اتصال در شبکه می باشد (مشابه با قسمت بعدی).

در مقایسه با پارامتر L^2 ، پارامتر تنظیم L^1 منجر به شکل گیری راه حل هایی با پراکندگی بیشتر از ω می شود و در نتیجه گروه کوچکی از ویژگی ها را یاد می گیرد این پارامتر را می توان به صورت زیر نشان داد :

$$\Omega(\theta) = \|\omega\|_1 = \sum_i |\omega_i| \tag{6}$$

پارامتر تنظیم فروبینوس بر اساس ضرب داخلی به دست می آید و قابلیت تجزیه ی بلوکی را ندارد، ازین رو محاسبه ی آن راحت تر می باشد. این پارامتر تنظیم را می توان به صورت زیر تعریف کرد :

$$\omega(\theta) = \sqrt{\sum_i \sum_j |\omega_{ij}|^2} = \sqrt{\sum_{i=1}^{rank(\omega)} \sigma_i^2} \tag{7}$$

که σ_i i امین مقدار منفرد بزرگ می باشد. این پارامتر تنظیم دارای تابعی مشابه با نرم هسته ای از نظر تنظیم می باشد.

تابع نرم هسته ای نیز به صورت گسترده به عنوان تابع تنظیم در سال های پیش مورد استفاده قرار گرفته است. این پارامتر در واقع مجموع مقادیر منفرد از ω را اندازه گیری می کند و می توان آن را به صورت زیر نشان داد :

$$\Omega(\theta) = \|\omega\|_* = \sum_{i=1}^{rank(\omega)} \sigma_i \tag{8}$$

روش های بهینه سازی

یک وظیفه ی یادگیری به صورت یک مسئله ی بهینه سازی تبدیل می شود تا بتوان حداقل تابع هدف را با انتخاب کردن ابر پارامتر های مناسب، به دست آورد. روند اصلی روش های مختلف بهینه سازی مشابه هم هستند. نخست، خروجی $f = f(x|\theta_0)$ و هدف بهینه سازی $\tilde{L}$ در مدل با استفاده از پارامتر های اولیه ی θ_0 محاسبه می شوند. پارامتر های شبکه θ سپس تنظیم می شوند تا مقدار تابع هدف از لایه ی نهایی به لایه ی اولیه، کاهش پیدا کند. این روند تا زمانی که مدل مناسب و خطای تناسب کم به دست بیاید تکرار می شود، یعنی مقادیر تابع ضرر به دست بیاید.

اما، روش های بهینه سازی مختلف دارای مزیت ها و معایب مختلف از نظر معماری هاو توابع ضرر مختلف می باشند. کاهش گرادیان تصادفی (SGD) و متغیر های آن نیز رایج ترین روش های مورد استفاده هستند که پارامتر ها را بر اساس یک اختلاف متناظر با ماتریس ژاکوبین، به روز رسانی می کند. زمان محاسبه برای هر به روز رسانی حتی با وجود مجموعه داده های تمرینی بزرگ هم خیلی افزایش پیدا نمی کند. روش آداگراد (AdaGrad) پارامتر ها را بر اساس تجمع گرادیان های مربع به روز رسانی می کند در نتیجه در شرایطی که بر روی توابع محدب اعمال شود می تواند همگرایی سریعی داشته باشد اما در شرایط مدل های خاص، عملکرد بدی را از خودش نشان می دهد. روش RMSProp ، یک الگوریتم آداگراد است که برای مسائل بهینه سازی به صورت رایج مورد استفاده قرار می گیرد. یک نوع دیگر از همین الگوریتم از مشتق های مرتبه دوم برای بهبود بهینه سازی استفاده می کند. به عنوان مثال، الگوریتم های حافظه ی محدود برویدن – فلچر – گلدفراب – شان (BFGS) یکی از روش های شبه نیوتونی می باشد که به صورت مکرر تخمین ماتریس معکوس هسیان را اصلاح می کند و ماتریس را ذخیره سازی نمی کند. این BFGS برای کار با مسئله هایی با ابعاد پایین مناسب می باشد ، به خصوص برای روش های پیچه ای (کانولوشن). به علاوه، گرادیان ترکیبی نیز روش های ترکیب و نزول گرادیان را در تصمیم گیری جهت پارامتر ها با هم ترکیب می کند و در نتیجه از محاسبه ی ماتریس معکوس هسیان، اجتناب می کند در حالی که در مدل های RBM معمولا از روش واگرایی مقایسه ای استفاده می شود. با استفاده از GPU ها ، بسیاری از این الگوریتم ها را می توان بسیار سریع تر اجرا کرد. عماری مناسب و توابع هدف کارآمد باید مطابق با داده های مورد استفاده، انتخاب شود. به عنوان یکی از انواع یادگیری ماشین، یادگیری عمیق می تواند با مشکل برازش بیش از حد رو به رو شود ، یعنی این که خطای پایین در داده های تمرینی اما خطای بالا در داده های تست ایجاد کند. علاوه بر عبارت تنظیم، دیگر روش ها نیز برای تنظیم و کاهش خطای تست ضروری هستند.

اضافه کردن نویز به داده های ورودی و یا وزن ها می تواند یکی از تکنیک های مناسب برای تنظیم الگوریتم باشد، مشابه روش رفع نویز در AE. متوقف کردن بهینه سازی پیش از موعد با تنظیم کردن شمار تکرار نیز یکی از روش های رایج برای جلوگلیری از بیش برازش شبکه می باشد. نشر پارامتر درست مشابه با CNN هم می تواند منجر به

شکل گیری تنظیم پارامتری شود. همچنین خارج کردن بعضی از پارامتر ها می تواند باعث شود که واحد ها به صورت مستقل تکامل پیدا کنند و در نتیجه به صورت اتفاقی بخش هایی از واحد ها در ANN در هر تکرار حذف می شود و ازین رو، نتایج بهتری با محاسبه های کمتر به دست می آید.

معماری یادگیری عمیق

AEها

AE ها نسبت به ANN های معمولی متفاوت هستند و ویژگی ها را از داده های دسته بندی نشده به دست می آورند و مقادیر هدف را برابر با ورودی ها در نظر می گیرند. با در نظر داشتن یک بردار ورودی به صورت $\{x^{(1)}, x^{(2)}, x^{(3)}, \ldots\}, x^{(i)} \in R^n,$ ، AE تلاش می کند تا مدل زیر را یاد بگیرد :

$$h_{w,b}(x) = g(Wx + b) \approx x \qquad (9)$$

که W و b پارامتر های مدل و g تابع فعال سازی این مدل می باشد (مطابق با تعاریف ارائه شده در این متن) $h_{W,b}$ هم نشان دهنده ی واحد های پنهان هستند. زمانی که تعداد واحد های پنهان که نشان دهنده ی ابعاد ویژگی ها می باشد، نسبت به ابعاد ورودی پایین تر باشد، AE مطابق با تحلیل بخش اصلی، روند کاهش ابعاد را اجرا می کند. علاوه بر شناسایی الگو، یک AE با یک طبقه بندی کننده در لایه ی نهایی می تواند وظایف طبقه بندی را هم انجام دهد.

RBM ها و DBN ها

RBM ها مدل های گرافیکی مولد هستند که هدفشان یادگیری توزیع داده های تمرینی می باشد. به دلیل این که ما نمی دانیم داده ها از چه نوع توزیعی تبعیت می کنند، ما نمی توانیم به صورت مستقیم پارامتر های این مدل را با استفاده از اصل بیشترین احتمال، محاسبه کنیم. ماشین های بلتزمن (BM ها) از تابع انرژی برای ایجاد کردن توزیع احتمال استتفاده می کنند (معادله های ۱۲ و ۱۳ را در قسمت پایین مشاهده کنید) و سپس پارامتر ها را بهینه سازی می کنند تا زمانی که مدل توزیع واقعی داده ها را پیدا کند. BM های اصلی کارایی عملی زیادی ندارند اما RBM ها به صورت رایج در یادگیری عمیق مورد استفاده قرار می گیرند.

RBM ها، BM ها را بر روی یک گراف دو قسمتی محدود می کنند، یعنی هیچ اتصالی داخل واحد های قابل مشاهده $v = x$ یا لایه های پنهان $\hbar$ وجود ندارد. این محدودیت تضمین می کند که استقلال شرطی لایه های پنهان و قابل مشاهده وجود دارد، یعنی :

$$\begin{aligned} P(h|v) &= \Pi_i p(h_i|v) \\ P(v|h) &= \Pi_j p(v_j|h) \end{aligned} \tag{10}$$

علاوه بر این، بیشتر RBM ها مبتنی بر این فرض کار می کنند که تمام واحد ها در شبکه تنها یکی از دو مقدار احتمالی ۰ یا ۱ را دریافت می کنند یعنی $h_i \in (0,1)$. مشروط نوع تابع فعال ساز، توزیع شرطی واحد های پنهان و قابل مشاهده را می توان به حالت زیر نشان داد :

$$\begin{aligned} p(h_i = 1|v) &= g(W_i v + c_i) \\ p(v_j = 1|h) &= g(W'_j h + b_j) \end{aligned} \tag{11}$$

بر اساس توزیع بلتزمن ، توزیع احتمال بر روی بردار های پنهان و قابل مشاهده به صورت زیر تعریف می شود :

$$P(v,h) = \frac{1}{Z} e^{-E(v,h)} \tag{12}$$

که $Z = \sum e^{-E(v,h)}$ ثابت نرمال کننده و $E(v,h) = -b'v - c'h - h'Wv$ تابع انرژی می باشد.

توزیع احتمال شرطی را هم می توان با ستفاده از انتگرال محاسبه کرد و سپس می توان این پارامتر را با حداقل کردن همگرایی کولباک – لیبلر ، بهینه کرد .

به صورت کلی، با در نظر داشتن معماری کلی این شبکه و پارامتر های بهینه شده، توزیع واحد های قابل مشاهده را می توان به صورت زیر محاسبه کرد :

$$P(v) = \sum_h p(v,h) = \sum_h \frac{e^{-E(v,h)}}{Z} \tag{13}$$

یک DBN را می توان به عنوان مجموعه ای از RBM ها و یا AE ها در نظر گرفت. مشابه با RBM ها ، DBN ها می توانند توزیع نمونه ها و یا نحوه ی طبقه بندی کردن ورودی ها با مشخص بودن دسته بندی ها را یاد بگیرند. اما،

$p(h)$ در فرمول $\tilde{p}(v) = \sum_h p(v,h) = \sum_h p(h)p(v/h)$ در این قسمت توسط یک مدل بهتر بعد از یادگیری وزن های اتصال W توسط یک RBM ، جایگزین می شود.

به علاوه ی استخراج ویژگی، RBM ها همچنین می توانند توزیع داده های دسته بندی نشده را به عنوان یک مدل مولد یاد گرفته و داده های دسته بندی شده را به صورت مدل های متمایز کننده، طبقه بندی کنند (با در نظر داشتن لایه های پنهان به عنوان لایه های نام گذاری). مشابه با AE ها ، RBM ها می توانند پارامتر ها را برای شبکه های پیچیده، از قبل تمرین بدهند.

شبکه های عصبی پیچشی (کانولوشن)

نورون های مصنوعی در شبکه های عصبی پیچشی که نسبت به ساختار های یادگیری عمیق متفاوت هستند ، ویژگی ها را را از بخش کوچکی از تصاویر ورودی دریافت می کنند که این تصاویر، میدان های دریافتی نامیده یم شوند. این نوع از استخراج ویژگی تحت الهام مکانیزم های بصری در ارگانیسم های زنده شکل گرفته است که در این ارگانیسم ها، سلول ها ها در کورتکس بصری نسبت به ناحیه ی کوچکی از میدان بصری ، حساس هستند .

علاوه بر این توابع فعال ساز، دو نوع خاص از لایه ها در CNN ها وجود دارد : لایه های پیچشی و لایه های تجمع (شکل ۲). در لایه ی پیچشی، این تصویر توسط فیلتر های پیچشی مختلف از طریق تغییر میدان دریافتی به صورت گام به گام، کانوالو می شود (شکل 2A). فیلتر های پیچشی پارامتر های مشابه با هم را دارند که بخش بسیار کوچکی از تصویر می باشد که به صورت محسوس منجر به کاهش تعداد ابر پارامتر ها در مدل می شود. یک لایه ی تجمع نیز از ویژگی های ایستای تصاویر استفاده کرده و میانگین، ماکسیموم و دیگر آمار ویژگی ها در قسمت های مختلف از نقشه ی ویژگی ها را دریافت می کند و ازین رو واریانس را کاهش داده و ویژگی های اصلی را ثبت می کند. (شکل 2B).

شبکه های عصبی بازگشتی

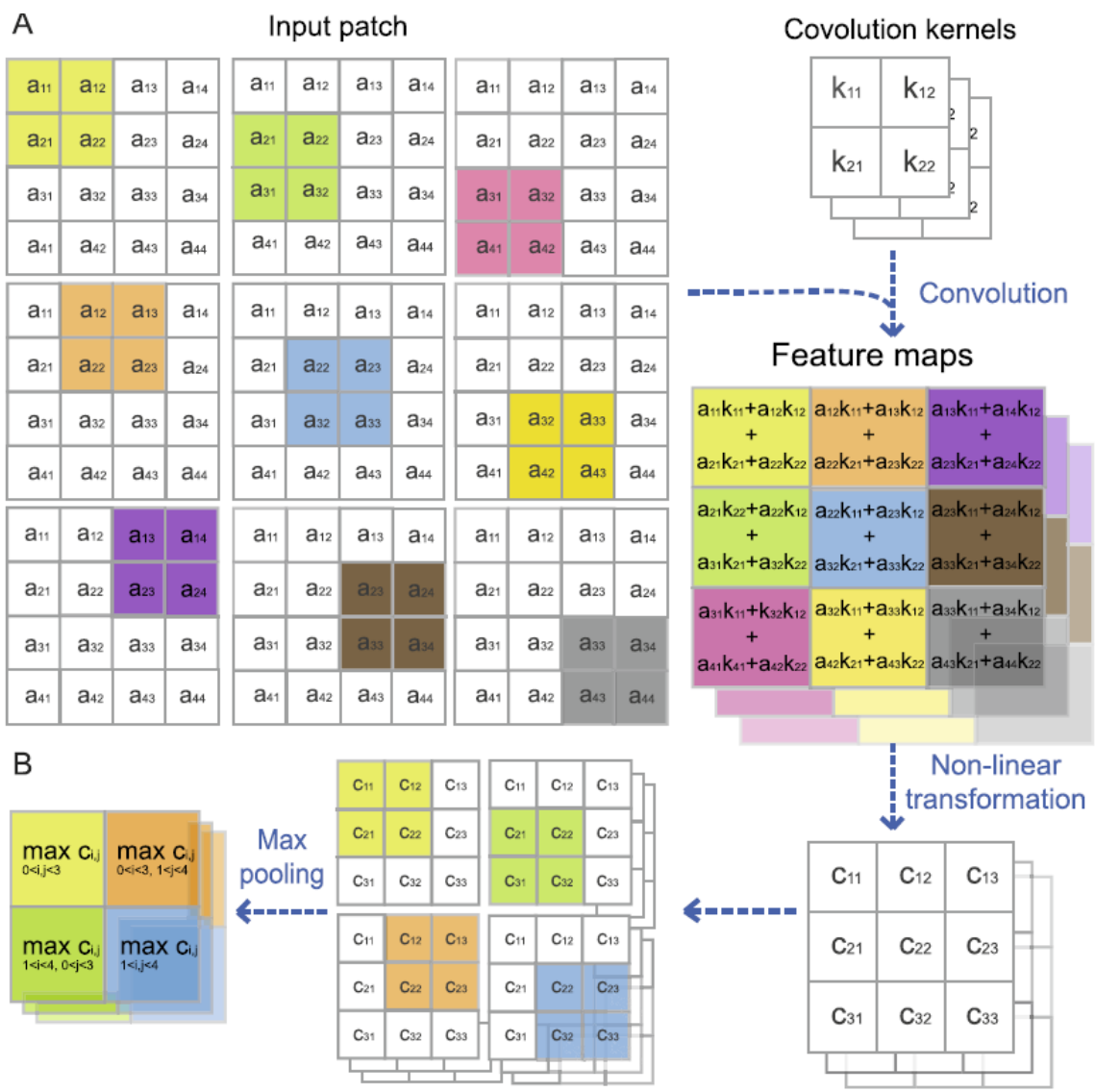
شبکه های عصبی بازگشتی (RNN) نسبت به دیگر روش های یادگیری عمیق در کار با داده های متوالی، عملکرد بهتری دارد. بر اساس ویژگی داده های متوالی، پارامتر ها بر روی گام های زمانی مختلف از مدل های RNN به اشتراک

گذاشته می شوند. با در نظر داشتن کلام به عنوان یک نمونه، این داده ها به صورت زیر می باشند : بعضی از حروف صدادار ممکن است نسبت به دیگر صدا ها بیشتر طول بکشند ؛ تفاوت های موجود موجب می شود که گام های زمانی مطلق معنی نداشته باشند و ازین رو، باید پارامتر های مدل در راستای گام های زمانی مشخص شوند.

علاوه بر اشتراک پارامتر ها، RNN ها نسبت به شبکه های چندلایه با داشتن یک چرخه متفاوت هستند که این چرخه، نشان دهنده ی بازگشت داده ها از لایه های پنهان به لایه های پنهان می باشد. یک شبکه ی بازگشتی متناظر با معادله های زیر میباشد :

$$\begin{aligned} h_{(t)} &= g(b + Uh_{(t-1)} + Wx_{(t)}) \\ o_{(t)} &= c + Vh_{(t)} \end{aligned} \tag{14}$$

که t نشان دهنده ی برچسب زمانی، W و V هم به ترتیب نشان دهنده ی وزن اتصال لایه های پنهان و لایه ی ورودی ، و لایه ی پنهان و لایه ی خروجی می باشد. b و c نیز مقادیر آفست به ترتیب برای لایه های قابل مشاهده و پنهان می باشند. g نیز نشان دهنده ی تابع فعال ساز و U هم نشان دهنده ی وزن اتصال واحد ها در زمان t-1 به لایه های پنهان در زمان t می باشد (شکل ۳).



شکل ۲ نمایش شبکه های عصبی پیچشی

A) در لایه ی پیچشی، میدان (بلوک های مختلف رنگی در جدول) از گام ورودی (که توسط a نشان داده شده است) توسط ماتریس هایی ضرب می شوند (هسته ی کانوالو ، که توسط k نشان داده می شود). B) در لایه ی تجمع ، نتایج کانولوشن به صورت خلاصه نشان داده شده است (بیشترین تجمع به عنوان نمونه در این قسمت در نظر گرفته می شود). a_{ij}, c_{ij}, k_{ij} نشان دهنده ی عدد هایی هستند که در ردیف i و ستون j در ماتریس متناظر می باشند.

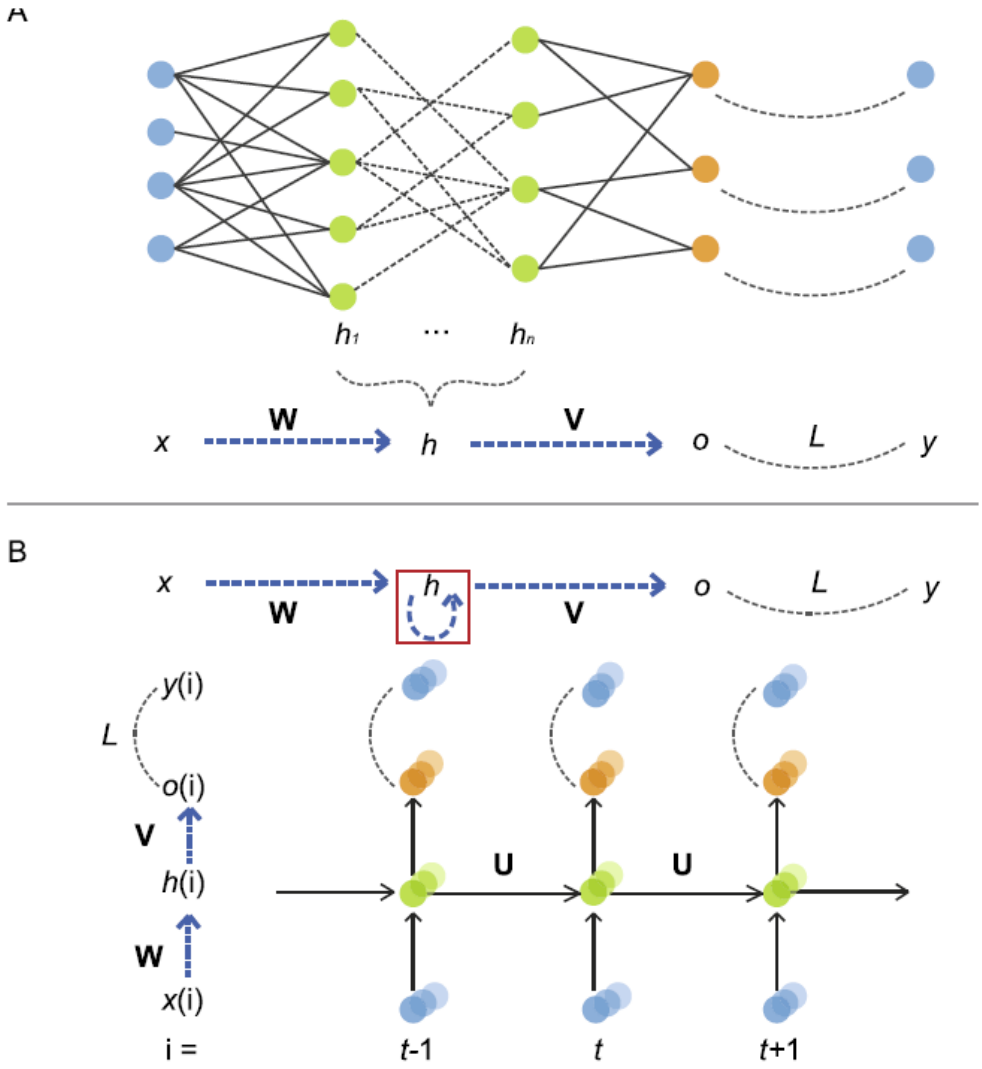
مشابه با دیگر معماری های یادگیری عمیق، RNN ها را می توان با استفاده از روش های پس انتشار آموزش داد. یک متغیر از روش BP با نام پس انتشار زمانی (BPTT) ، یکی از روش های استاندارد بهینه سازی برای RNN ها می باشد و بعضی از روش های جایگزین نیز در این قسمت برای افزایش سرعت بهینه سازی و یا افزایش ظرفیت آن ارائه شده است.

کاربرد های این روش در زیست پزشکی

به لطف پیشرفت در تکنولوژی ها با کارایی بالا، طوفان داده های پزشکی و زیستی در دهه های اخیر ایجاد شده است که این داده ها شامل داده های مرتبط با تصویر های پزشکی، توالی های ژنی و ساختار های پروتئینی می باشد. بعضی از کاربرد های موفق یادگیری عمیق در زمینه ی پزشکی ، در این قسمت مورد بررسی قرار گرفته و در جدول ۱ نشان داده شده است.

طبقه بندی تصویر های پزشکی و بخش بندی آن ها

یادگیری ماشینی برای تصویر های پزشکی ، یک ابزار قوی در تشخیص و یا ارزیابی بیماری ها بوده است. معمولا، ویزگی های متمایز اشاره به تفسیر تصویر های پزشکی دارد که معمولا پیش از این، به صورت دستی توسط انسان ها برای طبقه بندی (شناسایی کردن آسیب ها یا حالت های غیر عادی) و یا بخش بندی منطقه های مد نظر (بافت ها و ارگان ها) در زمینه های پزشکی مختلف، مورد استفاده قرار می گرفته است. این روند نیازمند مشارکت متخصص ها و پزشکان می باشد. در هر صورت، پیچدگی و ابهام تصویر های پزشکی، دانش محدود برای تفسیر تصاویر پزشکی و الزام مقادیر داده های تفسیری مانع استفاده ی گسترده از یادگیری ماشینی در دامنه ی تصویر های پزشکی شده است. اما باید به این نکته اشاره کرد که روش های یادگیری عمیق ، موفقیت زیادی در انواع وظایف بصری کامپیوتری مانند شناسایی اشیا، مشخص کردن مکان آن ها و یا بخش بندی تصویر های طبیعی به دست آورده است. این موارد ، باعث شده که این روش ها از نظر تفسیر تصویر های پزشکی نیز امید بخش باشند.



شکل ۳ نمایش شبکه ی عصبی بازگشتی

حالت باز شده ی شبکه های رایج عصبی (قسمت بالا) و طرح آن (پایین). B) نمایش شبکه های عصبی بازگشتی (قسمت بالا) و حالت باز شده ی آن ها (پایین). مربع های قرمز در این قسمت نشان دهنده ی یک گام تاخیر زمانی هستند. فلش ها در قسمت B متفاوت از قسمت A ، نشان دهنده ی مجموعه ای از اتصالات می باشد. W و B در این قسمت نشان دهنده ی وزن ماتریس و بردار بایاس می باشد. X و Y نیز به ترتیب نشان دهنده ی ورودی و خروجی شبکه می باشند؛ h نشان دهنده ی واحد های پنهان می باشد ؛ L نشان دهنده ی زوج های تبدیل مانند لایه های با ارتباط متراکم و یا لایه های خروجی می باشد ؛ U نشان دهنده ی تبدیل بین دو نقطه ی همسایه می باشد و t هم نشان دهنده ی نقطه ی زمانی می باشد.

بخش بندی کردن بافت ها و اندام برای ارزیابی کیفی و کمی تصویر های پزشکی موضوع بسیار مهمی می باشد. پیریا و همکارانش از تقویت داده ها، هسته های کانوالو کوچک و یک مرحله ی پزش پردازش استفاده کردند تا بتوانند تقسیم بندی مناسب برای تومور مغزی به دست بیاورند. روش های بخش بندی مبتنی بر CNN اولین جایگاه رد در چالش بخش بندی تومور مغزی (BRAST) در سال ۲۰۱۳ به دست آورد و در سال ۲۰۱۵ نیز جایگاه دوم را به دست آورد. هاوای و همکارانش یک روش بخش بندی تومور مغزی خودکار مبتنی بر DNN ها را در تصویر های رزونانس مغناطیسی (MRI) با روند تمرین دو فازی ارائه کردند که جایگاه دوم را در BRAST سال ۲۰۱۳ به دست آورد. روش آن ها بر روی مجموعه داده های عمومی INbreast و دیتابیس های دیجیتال برای پویش ماموگرافی (DDSM) تست شد و از نظر صحت و کارایی نسبت به جدید ترین روش ها که بر روی DDSM تست شده بود، به دست آورد. کاربرد های دیگر در زمینه ی پزشکی که از معماری های یادگیری عمیق استفاده می کنند نیز در بخش بندی بطن چپ قلبی از داده های MR ، مقطع نگاری کامپیوتری از پانکراس (CT) ، تصویر برداری رزونانس مغناطیسی از غضروف های درشت نی (MRI) و تصویر برداری از پروستات ، و تصویر برداری های MR مغزی از هیپوکامپ بررسی شده است. تمایز بافت ها و یا ارگان ها در تصویر های پزشکی با نام بخش بندی معنایی شناخته می شود که در این زمینه، هر پیکسل از تصویر ، در یک دسته بندی قرار می گیرد. عضله های اسکلتی، اندام ها و چربی در تصویر های CT به خوبی از طریق این بخش بندی های معنایی بر مبنای معماری های DNN شناسایی می شوند. به صورت مشابه، این بخش بندی های معنایی از تصویر های MR هم توانستند نتایج بخش بندی مناسبی را ایجاد کنند.

ترجمه فا
TarjomeFa.Com

این مقاله، از سری مقالات ترجمه شده رایگان سایت ترجمه فا میباشد که با فرمت PDF در اختیار شما عزیزان قرار گرفته است. در صورت تمایل میتوانید با کلیک بر روی دکمه های زیر از سایر مقالات نیز استفاده نمایید:

✔ لیست مقالات ترجمه شده

✔ لیست مقالات ترجمه شده رایگان

✔ لیست جدیدترین مقالات انگلیسی ISI

سایت ترجمه فا ؛ مرجع جدیدترین مقالات ترجمه شده از نشریات معتبر خارجی