Archive of SID

3rd National Conference on Electrical & Computer Engineering
Distributed systems & Smart Grids
سومین کنفرانس ملی برق و کامپیوتر
سیستم های توزیع شده و شبکه های هوشمند
۱۹ آذر ماه ۹۵

مروری بر برخی از روش های محبوب یادگیری عمیق

سید حسین حسن پور متی کلایی[1]*، رضا سعادتی[2]
[1]دانشگاه آزاد اسلامی واحد علوم و تحقیقات آیت اله آملی، St.h.hasanpour@iauamol.ac.ir
[2]دانشگاه علم و صنعت ایران، Rsaadati@iust.ac.ir

چکیده – یادگیری عمیق در سال های اخیر بسیار مورد توجه قرار گرفته و در مجامع مرتبط با یادگیری ماشین و بخصوص بینایی کامپیوتر بسیار از آن یاد شده است. یادگیری عمیق اشاره به زیرمجموعه ای از الگوریتم ها و روش های مبتنی بر یادگیری ماشین دارد که بدنبال مدلسازی انتزاعات سطح بالا در داده ها میباشند.به واسطه همین خصوصیت ، این نوع از روش ها به موفقیت های چشم گیری در حوزه یادگیری ماشین دست یافته اند و در برخی زمینه ها همانند دسته بندی تصاویر از قدرت انسان نیز پیشی گرفته اند. در این مقاله ما بصورت مختصر به معرفی گونه های مهمی از روش های مبتنی بر یادگیری عمیق میپردازیم و سعی میکنیم دیدی کلی از این روش ها را ارائه کنیم.

کلید واژه- یادگیری عمیق ، یادگیری ماشین، معماری ، هوش مصنوعی

۱- مقدمه

الگوریتم های یادگیری عمیق زیرمجموعه ای از الگوریتم های یادگیری ماشین هستند که هدف آنها کشف چندین سطح از بازنمودهای های توزیع شده از داده ورودی است.یادگیری عمیق همچنین بخشی از روش های یادگیری ماشین است که مبتنی بر یادگیری بازنمایی اند. تحقیقات در این زمینه بدنبال آن هستند که بازنمایی های بهتری ایجاد کرده و مدل هایی ایجاد نمایند تا این بازنمایی ها را از داده های عظیم برچسب گذاری نشده فرا گیرند.اخیرا الگوریتم های یادگیری عمیق زیادی برای حل مسائل هوش مصنوعی سنتی ارائه شده اند که هرکدام نقاط مثبت خاصی دارند. در ادامه ما مروری بر تعدادی از معروف ترین الگوریتم های این حوزه خواهیم داشت و سعی خواهیم کرد بصورت مختصر به معرفی هر روش پرداخته و دیدی کلی از آن ارائه کنیم .

۲-روش ها و پیشرفت های اخیر

طی سال های اخیر، یادگیری عمیق بصورت گسترده در حوزه بینایی کامپیوتر مورد مطالعه قرار گرفته است و به همین دلیل، تعداد زیادی از روش های مرتبط با آن بوجود آمده است. بطور کلی، این روش ها بر اساس روش پایه ای که از آن مشتق شده اند به ۴ دسته مختلف تقسیم میشوند که عبارتند از: شبکه عصبی کانولوشن ،ماشین بولتزمن محدود شده (RBMS) ، اتوانکودرها و در آخر کدگذاری تنک [۱].

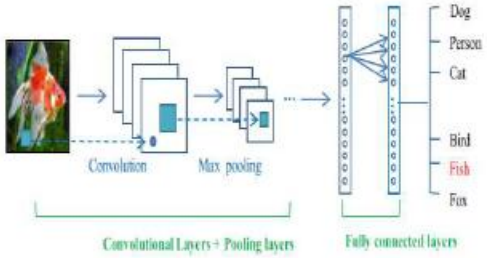
۲-۱-شبکه های عصبی کانولوشن (CNN)

یکی از مهمترین روش های یادگیری عمیق هستند که در آنها چندین لایه با روشی قدرتمند آموزش میبینند[۲]. این روش بسیار کارآمد بوده و یکی از رایجترین روش ها در کاربردهای مختلف بینایی کامپیوتر است. تصویر کلی یک معماری شبکه عصبی کانولوشن در شکل ۲ نمایش داده شده است . بطور کلی، یک شبکه CNN از سه لایه اصلی تشکیل میشود که عبارتند از : لایه کانولوشن، لایه Pooling و لایه تماما متصل. لایه های مختلف وظایف مختلفی را انجام میدهد. در شکل ۲ یک معماری کلی از شبکه عصبی کانولوشن برای دسته بندی تصاویر [۳] بصورت لایه به لایه نمایش داده شده است .در هر شبکه عصبی کانولوشن دو مرحله برای آموزش وجود دارد. مرحله انتشار به جلو و مرحله پس انتشار. در مرحله اول تصویر ورودی به شبکه تغذیه میشود و این عمل چیزی جز ضرب نقطه ای بین ورودی و پارامترهای هر نورون و نهایتا اعمال عملیات کانولوشن در هر لایه نیست. سپس خروجی شبکه محاسبه میشود. در این جا به منظور تنظیم پارامترهای شبکه و یا به عبارت دیگر همان آموزش شبکه، از نتیجه خروجی جهت محاسبه میزان خطای شبکه استفاده میشود. برای اینکار خروجی شبکه را با استفاده از یک تابع خطا با پاسخ صحیح مقایسه کرده و اینطور میزان خطا محاسبه میشود. در مرحله بعدی بر اساس میزان خطای محاسبه شده مرحله پس انتشار آغاز میشود. در این مرحله گرادیانت هر پارامتر با توجه به قاعده زنجیری

www.SID.ir

Archive of SID

محاسبه میشود و تمامی پارامترها با توجه به تاثیری که بر خطای ایجاد شده در شبکه دارند تغییر پیدا میکنند. بعد از بروز آوری شدن پارامترها مرحله بعدی فاز انتشار به جلو شروع میشود. بعد از تکرار تعداد مناسبی از این مراحل آموزش شبکه پایان میابد.



شکل ۲- لایه های مختلف شبکه کانولوشن [۱]

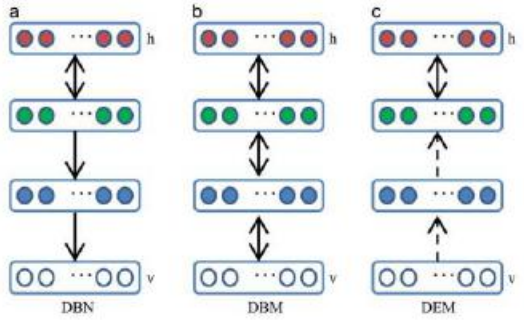
۲-۲- ماشین بولتزمن محدود(RBM)

یک ماشین بولتزمن محدود شده (RBMs) یک شبکه عصبی مولد تصادفی است که توسط هینتون و همکاران در سال ۱۹۸۶ ارائه شد [۲]. یک RBM گونه ای از یک ماشین بولتزمن است که دارای محدودیتی است که در آن واحدهای مشخص و واحدهای پنهان باید تشکیل یک گراف دو قسمتی را دهند. این محدودیت باعث ایجاد الگوریتم های آموزشی بهینه تر .علی الخصوص الگوریتم واگرایی تقابلی مبتنی بر گرادیانت میشود.

هینتون [۳] توضیح مبسوطی در این زمینه داد و یک راه عملی برای آموزش RBMها فراهم کرد. در [۴] مشکلات اصلی آموزش RBM ها و دلایل زیربنایی آنها مورد بحث قرار گرفته و یک الگوریتم جدید که شامل نرخ یادگیری قابل تطبیق و گرادیانت بهبود یافته برای رفع مشکلات مطرحی بود. ارائه شد. یک نمونه مشهور از RBM را میتوان در [۵]یافت. این مدل واحدهای دودویی را به همراه واحدهای خطی تصحیح شده لویزدار به منظور حفظ اطلاعات در مورد شدتهای نسبی در حین گردش اطلاعات در لایه ها تقریب میزند. نه تنها این اصلاح در این مدل بخوبی کار میکند بلکه بصورت گسترده در روش های مختلف مبتنی بر CNN نیز بکار برده میشود[۶, ۷].

با استفاده از RBM ها بعنوان ماجول های یادگیری. ما میتوانیم مدلهای عمیق شبکه باور عمیق(DBN) . ماشین بولتزمن عمیق(DBM) و مدل انرژی عمیق(DEM) را ایجاد کنیم .مقایسه میان این سه مدل در شکل ۳ نشان داده شده است . DBN ها دارای اتصالات بیجهت در دو لایه بالایی که تشکیل یک RBM را داده و دارای اتصالات جهت دار در لایه های پایینی هستند .DBM ها دارای اتصالات بیجهت بین تمامی لایه های شبکه هستند. DEM ها هم دارای واحدهای پنهان قطعی برای لایه های پایین و واحدهای پنهان تصادفی برای لایه بالا هستند[۱].

شکل ۳- مقایسه سه مدل عمیق DBNوDBM و DEM [۱]



شبکه باور عمیق یا همان (DBN) Deep Belief Network توسط هینتون ارائه شد[۸] . این شبکه پیشرفت قابل ملاحظه ای در یادگیری عمیق بود. DBN یک مدل مولد احتمالی است که توزیع احتمال مشترک بر روی داده های قابل مشاهده و برچسب ها را فراهم میکند. یک DBN ابتدا از یک استراتژی یادگیری حریصانه لایه به لایه بهینه برای مقدار دهی اولیه (پارامترهای) شبکه عمیق استفاده کرده و سپس تمامی وزنها را بطور مشترک با خروجی های مورد انتظار بدقت تنظیم میکند.

رویه یادگیری حریصانه دو فایده دارد. ۱) مقدار دهی اولیه مناسبی را برای شبکه ایجاد میکند و با اینکار به سختی ای که در انتخاب پارامتر وجود داشته و ممکن است باعث انتخاب نقطه بهینه محلی ضعیف شود تا حدی پاسخ میدهد ۲)رویه یادگیری بدون سرپرست بوده و به هیچ برچسب دسته ای نیاز ندارد بنابر این ضرورت داده برچسب دار برای آموزش را از میان برمیدارد. ایجاد یک مدل DBN اما از لحاظ محاسباتی کار بسیار پرهزینه ای است چراکه نیازمند آموزش چندیدن RBM بوده و مشخص هم نیست که چگونه

www.SID.ir

Archive of SID

میتوان حداکثر احتمال آموزش برای بهینه سازی مدل را تقریب زد [۱] [۸] .

DBN ها با موفقیت توجه محققان را به یادگیری عمیق معطوف کردند و در نتیجه، گونه های زیادی از آنها تولید شدند [۱۲-۹].

۲-۳ ماشین بولتزمن عمیق (DBMs)

ماشین بولتزمن عمیق یا به اختصار DBM , توسط سالاخودینوف و همکاران [۱۴] ارائه شد و یکی دیگر از الگوریتم های یادگیری عمیق است که در آن، واحدهای پردازشی در لایه هایی قرار میگیرنددر قیاس با DBN ها که دو لایه بالایی تشکیل یک مدل گرافی بی جهت و لایه های پایینی تشکیل یک مدل مولد جهت دار را میدادد، DBM دارای اتصالاتی در سراسر ساختار خود است . همانند DBM , RBM هم یکی از زیرمجموعه های خانواده بولتزمن است. تفاوت انها در این است که DBM ها دارای چندین لایه حاوی واحدهای پنهان بوده که واحدهای موجود در لایه های با شماره فرد بصورت شرطی مستقل از لایه های با شماره زوج هستند و بلعکس.

با وجود واحدهای مشخص محاسبه توزیع خلفی بر روی واحدهای پنهان ،دیگر قابل پیگیری نیست , که این نتیجه تعامل بین واحدهای پنهان است. در زمان آموزش شبکه، یک DBM بصورت مشترک تمامی لایه های یک مدل بدون سرپرست خاص را آموزش داده و بجای بیشینه سازی احتمال بصورت مستقیم، از الگوریتم مبتنی بر احتمال بیشینه تصادفی یا به اختصار SML برای بیشینه سازی مرزپایینی در احتمال استفاده میکند. به عبارت بهتر این به معنای اعمال تنها یک و یا چند بروز آوری با استفاده از روش Markov chain Monte Carlo (MCMC) بین هر بروز آوری پارامتر است. به منظور جلوگیری از رسیدن به بهینه محلی ضعیف که بسیاری از واحدهای پنهان را بطور غیرفعال رها میکند، یک استراتژی آموزش مبتنی بر لایه حریصانه نیز به لایه ها در زمان پیش آموزش شبکه DBM که بسیار شبیه همان چیزی است که در DBN[۹] انجام میشود، اضافه میشود [۱] .

آموزش مشترک بهبودهای نویدبخشی را هم در احتمال و هم در کارایی دسته بندی فراگیرنده های ویژگی عمیق به ارمغان اورده است . اما کاستی مهم DBM ها پیچیدگی زمانی استنتاج تقریب است که بسیار بالاتر از DBNهاست. که همین امر بهینه سازی مشترک پارامترهای DBM را برای مجموعه داده های عظیم غیرعملی میکند. به منظور افزایش بهره وری DBM ها , بعضی از محققان الگوریتم استنتاج تقریبی را معرفی کرده اند که از یک مدل تشخیص برای مقداردهی اولیه مقادیر متغییرهای پنهان در تمام لایه ها استفاده میکند و اینطور بطور موثری باعث تسریع استنتاج میشود. روش های زیاد دیگری هم وجود دارند که بدنبال بهبود اثرگذاری DBM ها هستند. بهبودها یا در مرحله پیش آموزش و یا در ابتدای مرحله آموزش میتوانند اتفاق بیفتند.

۲-۴- مدل انرژی عمیق (Deep Energy Model ((DEM)

DEM توسط نیام و همکاران ارئه شده و روشی نوین برای آموزش معماری های عمیق است. برخلاف DBN ها و DBM ها که در خصوصیت داشتن چندین لایه پنهان تصادفی دارای اشتراک هستند، DEM به منظور بهینگی آموزش و استنتاج تنها یک لایه حاوی واحدهای پنهان تصادفی دارد.

این مدل از شبکه های عصبی پیشخور برای مدل کردن از دورنمای انرژی بهره برده و قادر است تمامی لایه ها را در آن واحد آموزش دهد. با ارزیابی کارایی این مدل بر روی تصاویر طبیعی، مشخص میشود که آموزش مشترک چندین لایه باعث بهبود کمی و کیفی نسبت به آموزش لایه به لایه حریصانه میشود. نیام و همکاران از Hybrid Monte Carlo (HMC) برای آموزش این مدل استفاده کرد. گزینه های دیگری همانند واگرایی تقابلی، تطبیق امتیاز و ... نیز وجود دارند.

هرچند RMB ها مثل شبکه های عصبی کانولوشن برای کاربردهای بینایی کامپیوتر مناسب نیستند. اما نمونه های خوبی که از RMB ها جهت استفاده در فعالیتهای مربوط به بینایی کامپیوتر بهره بردند نیز وجود دارد. ماشین بولتزمن شکل، توسط اسلامی و همکاران برای استفاده در کارهای مدلینگ تصاویر شکل دودویی ارائه شد که توزیع احتمال با کیفیتی بر روی شکل شی ها هم به منظور واقعگرایی مثالها از توزیع و هم تعمیم به مثالهای جدید همان کلاس شکل را فرا میگرفت. کای و همکاران مدل CRF و RBM را با هم ترکیب کردند تا هم ساختار محلی و هم سراسری در قطعه بندی صورت را مدل کند . که این امر باعث کاهش پیوسته خطا در برچسب گذاری صورت شد. یک معماری

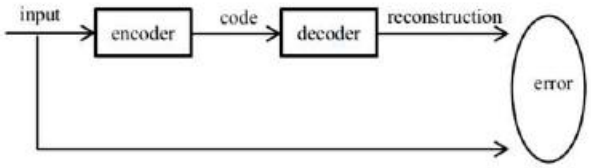
www.SID.ir

Archive of SID

عمیق جدید که برای تشخیص صدا ارائه شد ماجول استخراج ویژگی یک Mean-Covariance RBM را با DBN استاندارد باهم ترکیب کرد. این روش هم با مشکلات ناپهینگی نمایشی GMM ها و هم یک محدودیت مهم در کار قبلی که DBN ها در تشخیص صدا بهره برده بود ،مقابله میکند[۱] .

۲-۵- اتوانکودر

اتوانکودر نوع خاصی از شبکه عصبی مصنوعی است که برای کد کردن بهینه یادگیری مورد استفاده قرار میگیرد[۱] . بجای آموزش شبکه و پیش بینی مقدار هدف Y در ازای ورودی X ، یک اتوانکودر آموزش میبینید تا ورودی X خود را بازسازی کند. بنابر این بردارهای خروجی همان ابعاد بردار ورودی را خواهند داشت. فرآیند کلی یک اتوانکودر در شکل زیر نشان داده شده است. در حین این فرآیند، اتوانکودر با کمینه سازی خطای نوسازی بهینه میشود. کد متناظر همان ویژگی فراگرفته شده است.



شکل ۳- فرآیند کلی یک اتوانکودر [۱]

بطور کلی، یک لایه ی تنها ،قادر به دریافت ویژگی های متمایز از داده خام نیست . محققان در حال حاضر از اتوانکودر عمیق استفاده میکنند که کد یادگرفته شده از اتوانکودر قبلی را به اتوانکودر بعدی جهت به انجام رساندن کار خود ارسال میکنند.

اتوانکودر عمیق توسط هینتون و همکاران ارائه شد و هنوز بصورت گسترده در مقاله های اخیر مورد مطالعه قرار میگیرد. یک اتوانکودر عمیق اغلب با گونه ای از عملیات پس انتشار خطا مثل روش گرادیان مزدوج آموزش میبینید. هرچند اغلب این مدل کارآمد و موثر است ، اما درصورت وقوع خطا در لایه های اول میتواند بشدت ناکارآمد شود. این مسئله ممکن است باعث شود تا شبکه میانگین داده های آموزشی را بازسازی کند. یک روش مناسب جهت حذف این مشکل پیش آموزش شبکه با وزن های اولیه که راه حل نهایی را تقریب میزنند است[۱] . گونه هایی از اتوانکودرها هم وجود دارند که پیشنهاد کردند نمایش ها را تا حد امکان با توجه به تغییرات در ورودی ثابت نگه داشته شوند[۱] . در جدول زیر ، لیستی شامل گونه های معروف اتوانکودر به همراه خلاصه ای از خصائص و فواید آنها آمده است.

جدول ۱- لیستی شامل گونه های معروف اتوانکودر به همراه خلاصه ای از خصائص آنها

روش	ویژگی ها	نکات مثبت
Saturating autoencoder	افزایش،خطای بازسازی برای ورودی هایی که نزدیک منیفولد داده نباشند	توانایی بازسازی ورودی هایی که نزدیک منیفولد داده اصلی نیستند را محدود میکند.
Convolutional autoencoder	اشتراک وزنها در میان تمامی نواحی در ورودی و حفظ محلی بودن مکانی	از ساختار ۲بعدی تصویر بهره میبرد.
Zero-bias autoencoder	بهره بری از تابع کوچکسازی مناسب به منظور آموزش اتوانکودرها بدون بازتنظیم کننده اضافی	قدرت بیشتر در یادگیری بازنمایی ها بر روی داده هایی با ابعاد بسیار بالای واقعی
Sparse Autoencoder	این کدکننده یک جریمه پراکندگی را به منظور تنک سازی بازنمایی ها اضافه میکند.	۱. باعث جداسازی بیشتر دسته ها از هم میشود.۲ باعث بامعنی شدن بیشتر داده های پیچیده میشود. ۳ با سیستم بینایی هماهنگ است.
Denoising Autoencoder	بازیابی ورودی صحیح از نسخه معیوب	نسبت به نویز مقاوم تر است.
Contractive Autoencoder	افزودن جریمه انقباضی تحلیلی به تابع بازسازی خطا	بصورت بهتری جهات تغییرات محلی دیکته شده توسط داده را دریافت میکند.

www.SID.ir

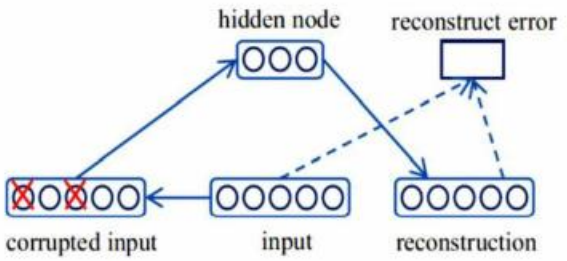
Archive of SID

۲-۶- اتوانکودر تنک

این نوع اتوانکودرها بدنبال استخراج ویژگی های پراکنده از داده خام هستند. پراکندگی نمودها میتواند یا از طریق جریمه بایاس های واحد پنهان و یا مستقیما با جریمه خروجی مقادیر واحد پنهان بدست بیایند. بازنمودهای پراکنده دارای چندین فایده احتمالی اند که عبارتند از : ۱) استفاده از نمودهای با ابعاد زیاد باعث افزایش احتمال انکه دسته های متفاوت براحتی همانند نظریه موجود در SVM , قابل جداشدن از هم باشند میشود. ۲)نمودهای پراکنده برای ما تفسیر ساده ای از داده ورودی پیچیده را در قالب تعدادی بخش فراهم میکنند. ۳) دید بیولوژیکی از نمودهای پراکنده در نواحی دیداری ابتدایی استفاده میکند [۱] . گونه بسیار مشهوری از اتوانکودر تنک یک مدل ۹ لایه ای بصورت محلی متصل با pooling و نرمال سازی شفافیت است. این مدل به سیستم اجازه میدهد تا یک صورت یاب را بدون نیاز به برچسب زنی تصاویر بصورت "حاوی صورت و بدون صورت " آموزش دهد. ویژگی یاب حاصل , در مقابل انتقال , تغییر مقیاس و چرخش خارج از صفحه بسیار مقاوم است.

۲-۷- Denoising autoencoder

به منظور افزایش قدرت این مدل، وینسنت مدلی بنام Denoising AutoEncoder (DAE)را ارئه داد که قادر به بازیابی ورودی صحیح از نسخه خراب بود و به این ترتیب مدل را به دریافت ساختار توزیع ورودی مجبور میکرد[۱]. فرآیند یک DAE در شکل زیر نشان داده شده است .



شکل ۴- Denoising autoencoder [۱]

۲-۸- Contractive autoencoder

Contractive autoencoder یا به اختصار CAE, توسط ریفای و همکاران ارائه شد که دنباله DAE بوده و انگیزه مشابه ای از یادگیری قدرتمند نمودها با آن داشت[۸] در حالی که DAE با تزریق نویز در مجموعه آموزشی باعث قدرتمند کردن تمامی عملیات نگاشت میشد, CAE با اضافه کردن جریمه افتراقی تحلیلی به تابع خطای بازسازی به این مهم دست پیدا میکرد. هرچند تفاوتهای قابل توجه بین DAE و CAE توسط بنجیو و همکاران عنوان شدند , اما الین و همکاران نشان دادند که DAE و یک شکل از CAE بسیار نزدیک به هم هستند. یک DAE با نویز خرابی کوچک را میتوان بصورت گونه ای از CAE که در آن جریمه افتراقی بجای اینکه بروی کدکننده باشد بر روی کل تابع نوسازی بوده ، در نظر گرفت. هم DAE و هم CAE هردو بصورت موفقیت آمیزی در رقابت انتقال یادگیری و بی سرپرست مورد استفاده قرار گرفتند[۱] .

۲-۸- Sparse Coding

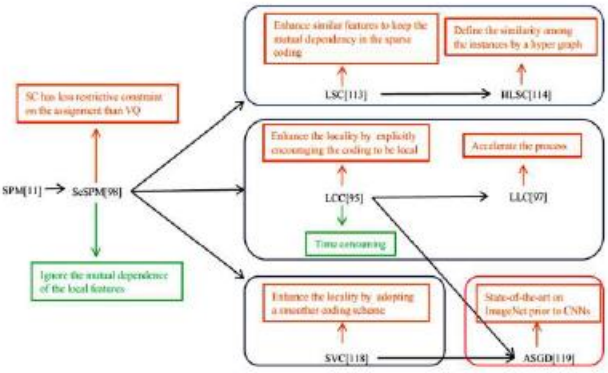
Sparse Coding نوعی روش بدون سرپرست است که به منظور یادگیری مجموعه فراکامل از توابع پایه برای توصیف داده ورودی مورد استفاده قرار میگیرد.این روش مزیتهای متعددی دارد[۱]. بعنوان مثال میتوان به این موارد اشاره کرد : ۱) این روش قادر به بازسازی بهتر توصیفگر با استفاده از چندین پایه و گرفتن(دریافت) ارتباطات بین توصیف گرهای مشابه که دارای پایه های مشترک اند. است . ۲)پراکندگی این اجازه را میدهد تا بازنمود . خصائص برجسته تصاویر را دریافت کند. ۳)این روش همسو با سیستم بینایی بیولوژیکی است که استدلال میکند ویژگی های پراکنده سیگنالها برای یادگیری مفید هستند. ۴) مطالعه آمار تصاویر نشان میدهد که تکه تصاویر سیگنالهای پراکنده هستند. ۵) الگوها با ویژگی های پراکنده از لحاظ خطی بیشتر قابل جداسازی هستند.

۲-۹- الگوریتم های مطرح Sparse coding

در این بخش ما به توضیح تعدادی از الگوریتم های مطرح Sparse Coding در حوزه بینایی کامپیوتر میپردازیم . در شکل زیر بعضی از این الگوریتم ها به همراه دستاوردها و مشکلات آنها نمایش داده شده اند.

www.SID.ir

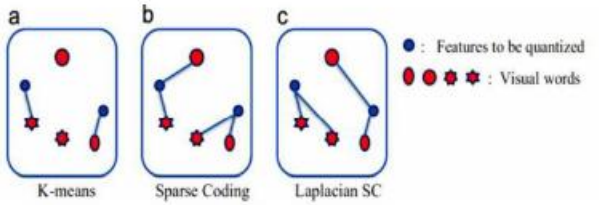
Archive of SID



شکل ۵- دستاوردها و مشکلات تعدادی از الگوریتم های مطرح Sparse Coding [۱]

یک نمونه از الگوریتم های SC , الگوریتم Sparse Coding SPM یا به اختصار ScSPMاست. که بسطی از الگوریتم Spatial(SPM) Pyramid Matchingاست برخلاف SPM که از (VQ) Vector Quantizationبرای نمایش تصویر استفاده میکند. ScSPM از (SC) Sparse coding به همراه multi scale spatial max pooling برای این منظور استفاده میکند. Codebook مربوط به SC یک over-complete basis بوده و هر ویژگی قادر به فعالسازی تعداد کمی از آنها میباشد. در قیاس با SC , VQ خطای بازسازی بسیار کمتری را بواسطه قید نسبتا آزاد تری که بروی عمل انتساب لحاظ میکند دارد. کوتس و همکاران در رابطه با دلایل موفقیت SCنسبت به VQ تحقیقات بیشتری انجام داد و نتایج را در قالب مقاله ای با جزییات فراوان ارائه نمود. یک اشکال ScSPM این است که این روش بصورت جداگانه با ویژگی های محلی برخورد میکند و به این ترتیب وابستگی دوطرفه بین آنها را نادیده میگیرد. که این خود باعث حساسیت بیش از حد به تغییرپذیری ویژگی میشود. این یعنی SCها ممکن است حتی برای ویژگی های مشابه هم نیز بشدت تغییر کنند.

به منظور برطرف سازی این مشکل، گو و همکاران طی مقاله ای روش کدگذاری تنک لاپلاسی یا به اختصار LSC را ارائه کرد که در آن ویژگی های مشابه نه تنها به مراکز دسته های بطور بهینه انتخاب شده انتساب داده میشوند بلکه تضمین میشود که مراکز دسته های انتخابی مشابه یکدیگر هم باشند. در شکل زیر تفاوت بین Sparse Coding , K-means و Laplacian Sparse Coding نمایش داده شده است .



شکل ۶- تفاوت بین Sparse Coding , K-means و Laplacian Sparse Coding [۱]

۲-۱۰- بحث

برای مقایسه بین این ۴ دسته از روش های یادگیری عمیق و بدست آوردن درک درستی از آنها، خلاصه ای از فواید و مشکلات هریک با توجه به خصائص متنوع آنها در جدول شماره ۲ آمده است .

در جدول زیر در کل ۹ خصیصه وجود دارد. که در آن تعمیم پذیری اشاره به این دارد که روش مورد نظر برای رسانه ها (تصویر، صدا و...) و کاربردهای متنوعی مثل تشخیص صدا ، تشخیص تصویر و ... موثر بوده است.

یادگیری بدون سرپرست هم اشاره به توانایی یادگیری خودکار یک مدل عمیق بدون تفسیر نظارتی است. یادگیری ویژگی هم توانایی یادگیری خودکار ویژگی ها بر اساس یک مجموعه داده است . آموزش بلادرنگ و پیش بینی بلادرنگ هم به ترتیب اشاره به بهره وری یادگیری و فرآینده ای استنتاجی هستند.

ادراک بیولوژیکی و توجیه نظری هم نشانگر آن هستند که آیا روش مورد نظر دارای شالوده بیولوژیکی یا مبانی نظری است یا خیر.

تغییرناپذیری اشاره به این مساله دارد که آیا روش مورد نظر در مقابل تبدیلاتی همانند : چرخش، تغییر مقیاس ، و انتقال مقاوم است یا خیر. مجموعه آموزشی کوچک هم اشاره به توانایی یادگیری یک مدل عمیق با استفاده از تنها تعداد کمی نمونه است . این نکته حائز اهمیت است که جدول زیر تنها یافته های کلی فعلی را نمایش داده و صبحتی از فرصتهای آینده و یا نمونه های اختصاصی نمیکند.

www.SID.ir

Archive of SID

جدول۲- خلاصه ای از فواید و مشکلات هریک

خصائص	CNN	RBM	AutoEncoder	Sparse coding
تعمیم	بله	بله	بله	بله
یادگیری بی سرپرست	خیر	بله	بله	بله
یادگیری ویژگی	بله	بله*	بله*	خیر
آموزش بلادرنگ	خیر	خیر	بله	بله
پیش بینی بلادرنگ	بله	بله	بله	بله
درک بیولوژیکی	خیر	خیر	خیر	بله
دارای توجیه نظری	بله*	بله	بله	بله
تغییر ناپذیری	بله*	خیر	خیر	بله
مجموعه آموزشی کوچک	بله*	بله*	بله	بله

توجه: "بله" اشاره به این دارد که دسته دراین خصوصیت بخوبی عمل میکند. در غیر اینصورت با خیر علامت گذاری شده است. "بله*" اشاره به توانایی ابتدایی و یا ضعیف دارد.

۳-نتیجه گیری

در این مقاله یادگیری عمیق بصورت مختصر معرفی شده و سپس طی بخشهای چندگانه به معرفی روش های مهم این شاخه و توضیح زیر شاخه های مهم هریک پرداخته شد. چیستی و نکات مثبت هریک از روش ها به اختصار بیان گردید و سپس در انتها تفاوت ها و قابلیت های پشتیبانی شده توسط هر روش به اختصار بیان گردید. روش های توضیح داده شده در این مقاله هرچند جزو مشهورترین روش های یادگیری عمیق اند اما تنها روش های فعال در این شاخه نیستند. بدیهی است بواسطه اهمیت این شاخه پرداختن به سایر روش های آن از اهمیت علمی بالایی برخوردار بوده و میتواند به عنوان یکی از جهات تحقیقات در این زمینه مورد استفاده واقع شود.

مراجع

۱. Guo, Y., et al., *Deep learning for visual understanding: A review*. Neurocomputing.

۲. Hinton, G. and T. Sejnovski, *Learning and Re-learning in Boltzmann Machines, PDP*. ۱۹۸۶, MTI Press.

۳. Hinton, G., *A practical guide to training restricted Boltzmann machines*. Momentum, ۲۰۱۰. ۹(۱): p. ۹۲۶.

٤. Cho, K., T. Raiko, and A.T. Ihler. *Enhanced gradient and adaptive learning rate for training restricted Boltzmann machines*. in *Proceedings of the ۲۸th International Conference on Machine Learning (ICML-۱۱)*. ۲۰۱۱.

٥. Nair, V. and G.E. Hinton. *Rectified linear units improve restricted boltzmann machines*. in *Proceedings of the ۲۷th International Conference on Machine Learning (ICML-۱۰)*. ۲۰۱۰.

٦. Krizhevsky, A., I. Sutskever, and G.E. Hinton. *Imagenet classification with deep convolutional neural networks*. in *Advances in neural information processing systems*. ۲۰۱۲.

۷. Zeiler, M.D. and R. Fergus, *Stochastic pooling for regularization of deep convolutional neural networks*. arXiv preprint arXiv:۱۳۰۱.۳۵۵۷, ۲۰۱۳.

۸. Bengio, Y., A. Courville, and P. Vincent, *Representation learning: A review and new perspectives*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, ۲۰۱۳. ۳۵(۸): p. ۱۷۹۸-۱۸۲۸.

۹. Lee, H., C. Ekanadham, and A.Y. Ng. *Sparse deep belief net model for visual area V۲*. in *Advances in neural information processing systems*. ۲۰۰۸.

۱۰. Lee, H., et al. *Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations*. in *Proceedings of the ۲٦th Annual International Conference on Machine Learning*. ۲۰۰۹. ACM.

۱۱. Lee, H., et al., *Unsupervised learning of hierarchical representations with convolutional deep belief networks*. Communications of the ACM, ۲۰۱۱. ٥٤(۱۰): p. ۹٥-۱۰۳.

۱۲. Nair, V. and G.E. Hinton. *۳D object recognition with deep belief nets*. in *Advances in Neural Information Processing Systems*. ۲۰۰۹.

www.SID.ir