

### بخش ۳ تمرین ۳

از STDP می توان در SNN (Spiking neural network) برای استخراج ویژگی های بصری که دارای پیچیدگی کمی هستند، در یک فعالیت بدون نظارت استفاده کرد. داشتن لایه های زیاد قدرت تشخیص را افزایش می دهد که این روش به نام Deep learning شناخته می شود.

در طی دهه های اخیر مدل های محاسباتی مختلفی طراحی شدند تا از این پردازش سلسله مراتبی Feed\_forward تقلید کنند. علی رغم موفقیت های محدود بدست آمده در مدل های اولیه، پیشرفت های اخیر در DCNN منجر به مدل های کارآمد شده. اگر چه معماری DCNN (Deep convolutional neural network) به نوعی از سیستم بصری پستانداران الهام گرفته شده است، آن ها کاملاً از مکانیزم های یادگیری و پردازش عصبی در قشر بینایی غافل می شوند. واحد های محاسباتی DCNN ها مقادیر اعشاری را به یکدیگر می فرستند که با سطح فعالیت آن ها در ارتباط است در حالی که نورون های بیولوژیک با ارسال پالس های الکتریکی با یکدیگر در ارتباط هستند.

میانگین نرخ spike در سیستم بصری پستانداران بسیار پایین است لذا نورون ها تنها وقتی بخواهند پیام بسیار مهمی را بفرستند spike می زنند و برخی از اطلاعات می توانند در زمان spike کدگذاری شوند که این کدگذاری موجب محاسبات بسیار سریع و کارآمد در مغز از نظر مصرف انرژی می شوند.

DCNN های امروزی که عملکرد خوبی دارند، با الگوریتم back\_propagation همراه با نظارت آموزش می بینند که ریشه بیولوژیکی ندارند. اگرچه این روش دقت خوبی دارد ولی همگرایی آن کند است. با توجه به اینکه DCNN ها میلیون ها پارامتر آزاد دارند برای جلوگیری از Overfit شدن، به میلیون ها مثال برجسب دار نیاز است. ولی پستانداران مخصوصاً انسان ها، می توانند با تعداد کمی مثال یاد بگیرند در حالی که در بسیاری از اوقات برجسبی در اختیار نیست. این امر را مدیون STDP هستند که یک مکانیزم یادگیری بدون نظارت در قشر بینایی پستانداران است. با توجه به STDP سیناپس هایی که از طریق آن ها spike های presynaptic قبل از spike های postsynaptic رخ می دهند، تقویت می شوند.

امروزه SNN های گوناگونی برای حل وظیفه شناخت اشیا طراحی شده اند که یک گروه از این شبکه ها، نسخه تبدیل شده از Traditional DCNN ها هستند. ایده اصلی جا به جا کردن هر بخش محاسباتی DCNN با یک Spiking neuron است. هدف از ایجاد این شبکه کاهش مصرف انرژی DCNN هاست ولی از مشکلات اجتناب ناپذیر این طور Spike\_rate coding ها، نیاز به spike های زیاد برای هر تصویر و زمان زیاد پردازش هستند. علاوه بر آن، استفاده از الگوریتم Back\_propagation و داشتن همزمان سیناپس های خروجی تحریکی و مهاری در یک نورون از نظر بیولوژیکی امکان پذیر نیست.

یک لایه پیچشی شامل چندین map نورونی می باشد. هر نورون به یک ویژگی بصری تعلق دارد که توسط وزن سیناپس ورودی اش مشخص می شود. نورون های مربوط به یک map مشخص ویژگی بصری یکسانی را در مکان های متفاوت استخراج می کنند. به همین منظور وزن های سیناپسی نورون های متعلق به یک map، همیشه باید یکسان باشند. هر نورون spike های ورودی را از نورون های قرار گرفته در یک پنجره معین از همه map های نورون لایه های قبل دریافت می کند لذا یک ویژگی بصری در یک لایه پیچشی ترکیبی از چند ویژگی ساده تر است که در لایه های قبلی استخراج شده اند. نورون های موجود در همه لایه های پیچشی از نوع non-leaky integrate-and-fire می باشند که spike های ورودی را از نورون های presynaptic دریافت می کنند و هرگاه پتانسیل داخلی آن ها به یک آستانه مشخص شده ای برسد، یک spike منتشر می کنند. در هر گام زمانی مقدار پتانسیل درونی نورون شماره i از طریق فرمول زیر به دست می آید:

$$V_i(t) = V_i(t-1) + \sum_j W_{j,i} S_j(t-1)$$

در این فرمول  $V_i(t)$  بیانگر مقدار پتانسیل داخلی نورون شماره i در لحظه t،  $W_{j,i}$  مقدار وزن سیناپس بین نورون presynaptic شماره j و نورون پیچشی شماره i می باشد. اگر  $S_j(t)=1$  باشد به این معنا است که نورون شماره j در لحظه t یک spike زده است و اگر مقدار آن برابر صفر باشد یعنی در لحظه t هیچ spike ای زده است.

اگر مقدار  $V_i$  به آستانه تعیین شده برسد، آنگاه نورون یک spike می زند و  $V_i$  از ابتدا مقداردهی می شود.

$$V_i(t) = 0 \text{ and } S_i(t) = 1, \text{ if } V_i(t) \geq V_{thr}.$$

یک مکانیزم مهار در همه لایه های پیچشی وجود دارد به این صورت که هرگاه یک نورون در یک مکان خاص spike بزند، باعث مهار نورون های دیگر که در همان مکان از map های دیگر قرار دارند می شود و به آن ها اجازه spike زدن را نمی دهد تا زمانی که عکس بعدی نشان داده شود. این اتفاق موجب یک کدگذاری مفید می شود زیرا در هر مکان حداکثر یک spike داریم که نشان دهنده وجود یا عدم وجود یک ویژگی خاص در آن مکان است.

لایه های Pooling با انجام عملیات غیرخطی max pooling بر روی مجموعه ای از نورون های همسایه با ویژگی برتر یکسان، به شبکه کمک می کنند تا از پراکندگی دور شود. نورون های Pooling از نوع integrate-and-fire هستند که وزن های سیناپسی ورودی و مقدار آستانه پتانسیل، همه با مقدار یکسانی مقاداردهی شده اند و به همین علت اولین spike ورودی آن ها را فعال می کند و منجر به یک spike خروجی می شود. هر نورون Pooling حداکثر یک بار اجازه spike زدن دارد. در لایه های Pooling هیچ یادگیری رخ نمی دهد بلکه اطلاعات بصری فشرده می شود.

یادگیری فقط در لایه های پیچشی رخ می دهد که باید یافتن ویژگی های بصری از ترکیب کردن ویژگی های ساده تری که در لایه قبل اسنخراج شده اند را یاد بگیرند. یادگیری در یک لایه پیچشی زمانی شروع می شود که یادگیری در لایه پیچشی قبلی تمام شود به عبارت دیگر یادگیری لایه به لایه رخ می دهد. هنگامی که یک تصویر جدید ارائه می شود، نورون های لایه پیچشی با یکدیگر رقابت می کنند و آن هایی که زودتر spike بزنند، موجب رخ دادن STDP می شوند و الگوی ورودی را یاد می گیرند. یک نسخه ساده شده از STDP به صورت فرمول زیر است که طبق آن مقدار تغییر وزن سیناپس را بدست می آورند:

$$\Delta w_{ij} = \begin{cases} a^+ w_{ij} (1 - w_{ij}), & \text{if } t_j - t_i \leq 0, \\ a^- w_{ij} (1 - w_{ij}), & \text{if } t_j - t_i > 0, \end{cases}$$

i و j به شماره نورون های presynaptic و postsynaptic اشاره می کنند.  $a^+$  و  $a^-$  نرخ یادگیری هستند.

همان طور که از این فرمول می توان فهمید، مقدار اختلاف زمانی بین دو spike تاثیری بر روی تغییر وزن ندارد بلکه علامت آن تاثیرگذار است. انتخاب مقادیر بزرگ برای  $a^+$  و  $a^-$  موجب کاهش حافظه یادگیری می شود که سبب می شود تا نورون ها آخرین تصویری که ارائه شده است را یاد بگیرند و تصاویری که قبلاً نمایش داده شده اند را یاد بگیرند و عکس هایی که قبلاً نمایش داده شده اند را یاد نگیرند. اگر  $a^-$  را بزرگ تر از  $a^+$  در نظر بگیریم وزن های سیناپسی به تدریج decay می شوند و نورون ها نمی توانند به آستانه پتانسیل برسند و spike نمی زنند بنابراین بهتر است تا  $a^+$  را کمی بزرگ تر از  $a^-$  در نظر بگیریم زیرا اگر خیلی بزرگ تر در نظر بگیریم، نورون ها تمایل دارند که بیشتر از یک الگو را یاد بگیرند و به همه آن ها پاسخ می دهند.

اخیراً DCNN های با نظارت بر روی بیشتر مسائل شناخت اشیاء به دقت بالایی دست یافته اند. معماری این شبکه ها تا حد زیادی از پردازش سلسله مراتبی عمیق در قشر بینایی الهام گرفته شده است ولی مکانیزم یادگیری و پردازش استفاده شده در DCNN ها با قشر بینایی (جایی که نورون ها با استفاده از spike ها با یکدیگر ارتباط برقرار می کنند و الگو های spike های ورودی را به طور عمده بدون نظارت یاد می گیرند) در تناقض است. به کارگیری چنین مکانیزم هایی در DCNN ها می تواند مصرف انرژی را کاهش دهد و نیاز آن ها به یادگیری با نظارت با میلیون ها تصویر برجسب دار که بسیار پر هزینه است را کاهش دهد.

امیرحسین باباجانی

۹۷۲۲۲۰۰۹