

RSTDTP در واقع یک یادگیری تقویتی است که تفاوت آن reward دادن یا ترشح دوپامین است که افزایش آن باعث تقویت یادگیری STDP میشود و اگر از به مقداری کمتر باشد نوعی punishment رخ میدهد. ترشح دوپامین ممکن است طول فعالیت رو کم و زیاد کند و یا قطبیت را تغییر دهد مثلا جایی که مثبت بوده منفی شود یا برعکس.

$$\begin{aligned}\frac{dc}{dt} &= -\frac{c}{\tau_c} + STDP(\tau)\delta(t - t_{pre/post}), \\ \frac{ds}{dt} &= cd,\end{aligned}$$

در این رابطه وزن سنایسی رو با s و c نشان میدهد که سیناپس اخیرا فعال بوده یا خیر. d میزان دوپامین رو نشان میدهد ، که اگر کمتر از صفر باشد دوپامین کاهش پیدا کرده و STDP به صورت معکوس اجرا میشه و اگر بیشتر از صفر باشد دوپامین زیاد شده و STDP به صورت اصلی خودش اجرا میشه.

c در حالتی زیاد میشه که نوروں اخیرا اسپایک زده باشه و بعد به تدریج سمت صفر میره.

عبارت دوم معادله ی اول وقتی یکی از نوروں های pre یا post فایر کنند، آپدیت میشه.

با استفاده از ترفند reward یا punishment میشه نوروں ها رو کنترل کرد مثلا اگر در به حالت خاص میخوایم فعالیت یک نوروں افزایش پیدا کنه به اون حالت خاص reward میدیم و بقیه رو punish میکنیم.

برای مدل سازی میتوانیم، شبکه های با 4 لایه را در نظر بگیریم، که شامل 2 لایه simple و 2 لایه complex است. مانند

شبکه SNN که از STDP برای یادگیری استفاده کرده بود، اکنون همان شبکه اسپایکی را مدل سازی میکنیم با این تفاوت که

به جای STDP برای یادگیری از RSTDTP استفاده میکنیم.

شبکه جدید با استفاده از RSDTP یادگیری را انجام میدهد و دیگر نیازی به classifier ندارد زیرا با عمل reward یا Punishment تصمیم گیری را اعمال میکند.

این شبکه شامل 4 لایه است:

S1 :

در این لایه با دریافت عکس ورودی، با استفاده از فیلتر gabor ، خطوط با زوایای متفاوت را از عکس استخراج میکنیم.

این ویژگی ها به اسپایک ها تبدیل میشود و نوروں های موجود در هر ناحیه اسپایک میزنند. این لایه اولین لایه ی simple است. در این لایه هر نوروںی که سریعتر اسپایک بزند به لایه بعدی منتقل خواهد شد

: C1

این لایه اول رز ی لایه complex است که به عنوان pooling layer عمل میکند و هر اسپایکی که زودتر بیاید را در هر ناحیه دریافت میکند. این عملیات باعث کاهش نرون ها در لایه s1 میشود.

: S2

این لایه دومین لایه ی simplex است که ویژگی ها را از لایه ی قبلی گرفته و آنها را باهم ترکیب میکند تا در نهایت بتواند اشکال پیچیده تر را انتخاب کند. Learning با استفاده از قاعده ی RSTDTP در این لایه رخ میدهد.

: C2

در این لایه pooling نهایی رخ میدهد. این Pooling باعث میشود که یکی از اسپایک ها در نهایت انتخاب شوند و تصمیم

گیری نهایی انجام شود. اگر این تصمیم درست باشد با استفاده از reward ، تقویت میشود و نتیجه نهایی را مشخص میکند و

در کلاس های مختلف قرار میدهد. اگر تصمیم نادرست باشد، punishment انجام میشود تا دیگر شبکه جلوتر نرود و متوقف شود.