

RSTDTP (Reward-modulated STDP)

فرآیند STDP شامل LTP و LTD سیناپس‌ها می‌باشد. در صورتی که نورون پیش‌سیناپسی زودتر اسپایک بزند سیناپس بین دو نورون تقویت شده (LTP) و اگر نورون پس‌سیناپسی زودتر از نورون پس‌سیناپسی اسپایک بزند، سیناپس بین دو نورون ضعیف‌تر می‌شود. (LTD)

در RSTDTP اجزای LTD و LTP از STDP توسط غلظت دوپامین تنظیم می‌شوند. در واقع نوعی یادگیری تقویتی داریم که افزایش غلظت دوپامین به معنی دریافت پاداش است و در طی آن STDP اعمال می‌شود و کاهش غلظت دوپامین به معنای مجازات است و در طی آن Anti-STDTP اعمال می‌شود.

از مزایای این روش نسبت به STDP این است که:

۱. می‌تواند بازه زمانی که STDP در آن اتفاق می‌افتد (STDP's window) را تغییر دهد.

۲. قطبیت STDP را تغییر دهد یعنی در بازه‌ای که سابقاً LTD اتفاق می‌فتاده LTP اتفاق بیافتد و برعکس.

کاربرد:

می‌توان از یادگیری تقویتی برای آموزش شبکه‌های عصبی اسپایکی (SNN) برای تشخیص اشیاء مختلف در تصاویر استفاده کرد. در صورت عدم استفاده از یادگیری تقویتی به یک کلاس‌بند خارجی در آخرین لایه شبکه نیاز است و همچنین یادگیری بدون نظارت می‌باشد. اما در این روش یادگیری با نظارت خواهد بود.

در این روش از یک شبکه‌ی عصبی اسپایکی کانولوشنی استفاده می‌شود همراه با یک کدگذاری زمانی بدین ترتیب که نورون‌هایی که با شدت بیشتری فعال می‌شوند زودتر شلیک می‌کنند و نورون‌هایی که با شدت کمتری فعال شده‌اند یا دیرتر شلیک می‌کنند یا اصلاً شلیک نمی‌کنند. پیچیدگی و ویژگی‌های استخراج شده در طول لایه‌ها از آشکارسازهای لبه در لایه اول تا نمونه‌های اولیه‌ی شی (مثلاً چشم در صورت انسان یا چرخ دوچرخه) در بالاترین لایه‌ها افزایش می‌یابد. در بالاترین لایه‌ها هر نورون به یک کلاس اختصاص داده شده است و کلاس اولین نورونی که شلیک کند به عنوان پیش‌بینی در نظر گرفته می‌شود اگر این پیش‌بینی درست باشد به نورون پاداش داده می‌شود. (STDP اعمال می‌شود) در غیر این صورت anti-STDTP اعمال می‌شود.

رویکرد RSTDP ویژگی‌های بصری متمایز کننده را استخراج می‌کند در حالی که STDP هر ویژگی را که به طور مداوم تکرار می‌شود استخراج می‌کند. در نتیجه RSTDP نسبت به STDP عملکرد بهتری دارد. بعلاوه R-STDP برای یادگیری آنلاین مناسب تر است و می‌تواند با تغییرات فاحشی از جمله جایگشت های برجسته سازگار شود. همچنین در این روش استخراج ویژگی و کلاس بندی با استفاده از حداکثر یک اسپایک در هر نورون انجام می‌شود که در نتیجه شبکه ای که از این رویکرد استفاده کند از نظر سخت افزاری سازگار است و در از نظر مصرف انرژی بهینه است.