

1. Understanding Error Propagation in Deep Learning Neural Network (DNN)

Accelerators and Applications :

فارسی

مشخصات مراجع به طور کامل
نوشته شود.

- مسئله: امروزه *Soft Error* ها زیاد شده و بخاطر انرژی زیادشان باعث ایجاد خرابی در سیستم‌های شبکه عصبی یادگیری عمیق^۱ (*DNN*) می‌شوند. برای رفع این مشکل و افزایش قابلیت اطمینان در گذشته روش‌هایی مانند *TMR* بود که سربار زیادی به همراه داشت. مسئله کاهش سربار و هزینه در سیستم‌های *DNN* با توجه به مسئله قابلیت اطمینان بود.
- راه حل: در این مقاله ویژگی و مشخصه‌های قابلیت اطمینان سیستم‌های *DNN* مورد بررسی قرار گرفته^۲ شده است و مشخص شده است که قابلیت اطمینان سیستم‌های *DNN* وابسته به عواملی مانند نوع داده‌ها، مقدار، استفاده مجدد از داده‌ها و انواع لایه‌ها در طراحی است و سپس با توجه به این عوامل^۳ دو تکنیک کارآمد برای سیستم‌های *DNN* ارائه شده است که یکی از تکنیک‌ها نرم‌افزاری و دیگری سخت‌افزاری است. هدف از ارائه این تکنیک‌ها کاهش سربار و افزایش پوشش^۴ در مسئله قابلیت اطمینان سیستم‌های *DNN* است. این تکنیک‌ها با توجه به توصیف انتشار *Soft Error* ها مکانیزم سخت‌افزاری و نرم‌افزاری برای کاهش هزینه و سربار ارائه می‌دهند.

2. A framework for quantifying the resilience of deep neural networks

- مسئله: امروزه استفاده از شبکه‌های عصبی عمیق در حال رشد است،^۵ همچنین کسری از چرخه‌های محاسباتی که به اجرای آنها اختصاص یافته است نیز در حال زیاد شدن است. این کار سبب شده است که به سمت ساخت سخت‌افزارهای *DNN* برویم. با وجود اهمیتی که امروزه سیستم‌های *DNN* دارند مسئله قابلیت اطمینان در این سیستم‌ها نادیده گرفته شده است. پس مسئله^۶ ی این مقاله کم^۷ توجهی به مسئله قابلیت اطمینان در

سیستم‌های DNN است.

- راه حل: در این مقاله ابتدا با تحریک قابلیت اطمینان الگوریتم‌های ذاتی DNN ها، رابطه‌ای که بین نرخ اشکال و دقت مدل وجود داشته را فهمیده اند سپس یک چارچوب^۱ تزریق اشکال به نام Ares برای یک نوع DNN خاص و سبک وزن ارائه دادند که در ۱۲ درصد سخت‌افزار واقعی مورد تایید قرار گرفته است. همچنین فهمیدند که ساختارهای مختلفی از DNN ها رفتارهای تحمل اشکال متفاوتی دارند. آزمایشات نشان می‌دهد که استفاده از این مدل می‌تواند تا ۵۰ برابر قابلیت اطمینان بالاتری داشته باشند.

3. A Fault-Tolerant Neural Network Architecture

- مسئله: تحقیق‌های زیادی در زمینه ی شتاب دهنده‌های جدید شبکه ی عصبی عمیق مبتنی بر فناوری‌های نوظهور، مانند حافظه دسترسی تصادفی مقاوم (ReRAM)، انجام می‌شود. متأسفانه محدودیتهای فیزیکی در سطح دستگاه که مختص این فناوری‌هاست ممکن است باعث اختلال وزن در حافظه و در نتیجه اختلال کارایی و پایداری شتاب دهنده‌های DNN شود.

فاصله‌گذاری‌ها رعایت شود.

- راه حل: در این مقاله یک معماری شبکه ی عصبی جدید که تحمل پذیر اشکال نیز هست پیشنهاد شده است که مشکل اختلال در وزن‌ها را بدون آموزش مجدد کاهش می‌دهد. در واقع نویسندگان این مقاله یک طبقه بندی کننده منطقی جدید را پیشنهاد کردند که با طراحی مجدد طبقه بندی‌های باینری در کد خروجی تصحیح خطا و الگوریتم آموزش DNN مدرن، پایداری DNN را تقویت می‌کنند. آن‌ها همچنین یک طرح رمزگشایی با طول متغیر را که بهینه سازی شده است برای تقویت دقت در تعداد کمتری از طبقه بندی کننده‌ها ارائه کرده‌اند. نتایج آزمایشات در این مدل نشان می‌دهد که معماری شبکه عصبی تحمل پذیر اشکال می‌تواند به طور موثری تخریب دقت در برابر اختلال در وزن برای شتاب دهنده‌های DNN را با هزینه کم اصلاح کند.

4. CANN: Curable Approximations for High-Performance Deep Neural Network Accelerators

- مسئله: با پیشرفت تکنولوژی موضوع محاسبات تقریبی اهمیت زیادی پیدا کرد. محاسبات تقریبی برای بهبود و تقویت مساحت، کارایی، توان موثر یا انرژی موثر با هزینه ی کاهش کیفیت خروجی بود. به عنوان مثال یادگیری

ماشین یکی از کاربردهایی است که محاسبات تقریبی در آن‌ها نقش مفیدی می‌تواند داشته باشد. حال از دست دادن دقت به خاطر محاسبات به صورت تقریبی در شتاب دهنده‌های سخت‌افزاری شبکه عصبی عمیق ممکن است نتایج مطلوبی نداشته باشد. در این مقاله مسئله‌ی مورد بررسی از دست دادن دقت در شتاب دهنده‌های *DNN* بخاطر تقریب زدن است.

○ راه حل : در این مقاله یک روش جدید برای طراحی شتاب دهنده‌های *DNN* با کارایی بالا ارائه شده است که خطاهای تقریبی از یک قسمت طرح در قسمت بعدی به طور کامل جبران می‌شود. همچنین در این مقاله یک مورد برای تقویت کارایی معماری سخت‌افزاری مبتنی بر آرایه‌ی انقباضی، برای *DNN*‌ها ارائه شده است که معمولاً برای شتاب دادن به الگوریتم‌های یادگیری عمیق استفاده می‌شود.

5. A Fast and Efficient Fault Tree Analysis Using Approximate Computing

○ مسئله : تحلیل درخت اشکال یکی از روش‌های محاسبه‌ی قابلیت اطمینان و تحلیل احتمال خرابی است. روش تحلیل درخت اشکال یک روش عمومی برای محاسبه‌ی قابلیت اطمینان است و شبیه‌سازی آن زمان زیادی می‌برد. بنابراین یکی از مسائلی که نیاز به بررسی و راه حل دارد این است که زمان شبیه‌سازی را با در نظر گرفتن مواردی مانند مساحت، توان مصرفی بهبود داد.

○ راه حل: در این مقاله، یک روش جدید ارائه می‌شود که از محاسبات تقریبی برای کاهش زمان شبیه‌سازی تجزیه و تحلیل درخت اشکال استفاده می‌کند، ضمن اینکه در این روش دقت، مساحت و مصرف توان را به عنوان پارامترهای مؤثر در نظر می‌گیرد. برای این کار یک درخت اشکال با زبان *VHDL* توصیف شده است و سپس در آن مولفه‌های ریاضی زمان بر مثل جمع کننده و ضرب کننده با سخت‌افزارهای تقریبی جایگزین شده‌اند. قابلیت اطمینان آن‌ها از طریق شبیه‌سازی مونت کارلو حساب شده است. و در آخر حدود ۱۵,۲ درصد در زمان شبیه‌سازی، ۶۹,۶ در توان مصرفی و ۹۸,۸ در مساحت بهبود حاصل شده است.

6. Deep neural network accelerator based on FPGA

○ مسئله : در برخی از برنامه‌های کاربردی قابل بازپیکربندی، پیاده‌سازی سخت‌افزار خاص *DNN* در شرایطی که یک کامپیوتر عمومی به دلیل محدودیت در سرعت، هزینه تراشه و مصرف انرژی مناسب نیست، ضروری است.

حال مسئله در این مقاله ارائه یک بستر مناسب برای پیاده سازی سخت افزار شبکه عصبی عمیق با در نظر گرفتن مسئله قابلیت بازپیکربندی آن است.

- راه حل : در این مقاله، یک معماری کارآمد برای شبکه های عصبی عمیق بر روی سیستم محاسباتی قابل بازپیکربندی مانند FPGA پیشنهاد می شود. معماری شبکه عصبی پیشنهادی تنها یک لایه محاسبه فیزیکی واحد برای انجام کل محاسبات شبکه های عصبی عمیق با ویژگی های تعداد لایه های قابل تنظیم، تعداد سلول های عصبی در هر لایه و تعداد ورودی ها دارد. ورودی ها، وزن ها هستند و خروجی های شبکه با فرمت عدد شناور با دقت ۱۶ بیتی نمایش داده می شوند.

7. Deep Neural Network Model and FPGA Accelerator Co-Design: Opportunities and Challenges

- مسئله : با رشد انفجاری الگوریتم های مختلف شبکه عصبی، پیاده سازی آنها با کارایی بالا بر روی پلتفرم های سخت افزاری، مانند GPU و FPGA نیز بسیار مهم می شود. در مقایسه با GPU های پرکاربرد، طراحی و بهینه سازی با FPGA حتی با کمک ابزارهای سطح بالا سنتز سخت تر به نظر می رسد. ولی با این حال مسئله ی سرعت و توان مصرفی یا انرژی مصرفی جزء موارد مهمی است که باید در پیاده سازی به آن ها توجه شود که هر دو عامل در برنامه های یادگیری ماشین مهم هستند.
- راه حل : مطالعات اخیر نشان داده اند که FPGA ها می توانند از GPU ها در راندمان انرژی و سرعت بهتر باشند. در این مقاله، روی شبیه سازی DNN ها و روش طراحی توام سخت افزاری نرم افزاری برای بهبود کارایی DNN روی FPGA بحث شده است.

8. Dependability analysis of safety critical systems: Issues and challenges

- مسئله : سیستم های بحرانی-ایمنی که به تدریج در حوزه هایی مانند انرژی هسته ای، حمل و نقل، پزشکی و سیستم های اطلاعاتی مورد استفاده قرار می گیرند، غالباً یک فرایند رسمی صدور گواهی نامه اتکاپذیری دارند. هدف از فرایند اتکاپذیری این است که این سیستم ها خدمات مورد انتظار را به کاربران خود تحویل دهند. برای تجزیه و تحلیل سیستم های بحرانی، فرایند تجزیه و تحلیل اتکاپذیری پذیرفته شده جهانی پذیرفته شده وجود

ندارد که به انتخاب معیارها، تکنیک‌ها و روش‌ها برای ارزیابی اتکاپذیری چنین سیستم‌های مهمی کمک می‌کند. در این مقاله به بررسی این مسئله پرداخته شده است.

○ راه حل: در این مقاله، یک بررسی جامع به منظور بررسی معیارهای مختلف، تهدیدها، ابزارها، تکنیک‌ها و روش‌های مختلف برای اطمینان از اتکاپذیری سیستم‌های بحرانی مبتنی بر رایانه ارائه شده است. محدودیت‌های این عناصر با توجه به کاربرد آنها در سیستم‌های SC تجزیه و تحلیل می‌شود. هدف از این مقاله تفسیر یک مفهوم بازبینی دقیق، مربوط به طیف گسترده‌ای از امور است. بنابراین، این کار به دانشگاهیان، محققان و دست اندرکاران کمک می‌کند تا این موضوع را عملی کنند و کاستی‌های پژوهش‌های موجود را مورد تجزیه و تحلیل قرار دهند.

9. *A machine learning based hard fault recuperation model for approximate hardware accelerators*

○ مسئله: این مسئله که همواره دنبال کارایی بالاتر و بهره‌وری انرژی بهتر هستیم منجر به ساخت SoC¹ ناهمگن شده است که شامل چندین شتاب دهنده سخت‌افزاری اختصاصی است. این شتاب دهنده‌ها از موازی بودن ذاتی کارها استفاده می‌کنند و اکثراً تحمل عدم دقت در خروجی‌های خود را دارند، به عنوان مثال در برنامه‌های پردازش سیگنال تصویر و دیجیتال عدم دقت تا حدی قابل چشم‌پوشی است. حال به دلیل پیشرفت تکنولوژی و محدودیت‌های توان، اشکال‌های دائمی در حال افزایش هستند و منجر به خروجی‌های نادرست می‌شوند که این به یک مسئله جدی تبدیل شده است.

○ راه حل: در این مقاله، یک روش ترمیم/تعمیر اشکال که کلی و کم هزینه باشد، ارائه شده است که از تکنیک‌های نظارت بر یادگیری ماشین استفاده می‌کند تا اثر اشکال‌های دائمی را در شتاب دهنده‌های سخت‌افزاری بهبود بخشد. مدل پیشنهادی به هیچ گونه اطلاعاتی در مورد شتاب دهنده احتیاج ندارد و مقیاس پذیری بالایی با مقدار کمی سربار برای مساحت دارد. درنهایت روش پیشنهاد شده در این مقاله ۵۰٪ دقت را بهبود می‌بخشد و میانگین نرخ خطای کلی را ۹۰٪ با سربار مساحت حدود ۵٪ کاهش می‌دهد.

10. A Uniform Modeling Methodology for Benchmarking DNN Accelerators

- مسئله : امروزه با توجه به کاربرد زیاد شتاب دهنده‌ها، از آن‌ها برای پردازش بارهای کاری DNN با استفاده از روش‌هایی که انرژی در آن‌ها کارآمد باشد، تولید شده و همچنان در حال توسعه هستند. فضای طراحی برای شتاب دهنده‌های DNN می‌تواند بسیار بزرگ باشد زیرا می‌توانند از داده‌های مختلف، استراتژی‌های نگاشت داده‌ها، مدارها و فناوری‌های دستگاه استفاده کنند. برای کشف فضای طراحی برای توسعه شتاب دهنده‌های DNN، مهم است که هزینه انرژی مرتبط با یک شتاب دهنده را تخمین بزنید. تخمین هزینه انرژی در یک شتاب دهنده یکی از مسائلی است که باید به آن پرداخته شود.
- راه حل: در این مقاله یک چارچوب مدل سازی یکنواخت، به نام Eva-DNN، برای تخمین انرژی پویای مصرف شده توسط یک شتاب دهنده ی DNN معرفی شده است. در این مقاله در سطوح مختلف حافظه و واحدهای عملیاتی، تعداد دسترسی‌ها و هزینه انرژی مرتبط با آن، مدل می‌شود. برای مدل سازی هزینه انرژی یک واحد عملیاتی، از یک معیار^۱ در سطح دستگاه استفاده می‌شود. Eva-DNN می‌تواند سهم انرژی را در فناوری دستگاه، مدارها، معماری، استراتژی نگاشت داده‌ها و شبکه را دقیق تر مدل کند. نتایج حاصل شده از پیاده سازی این مدل در ۳ معماری مختلف شبکه عصبی، بهبود تخمین هزینه انرژی در حدود ۴ تا ۸ درصد است.

11. Exploratory study of machine learning techniques for supporting failure prediction

- مسئله : پیچیدگی رو به رشد نرم‌افزار تشخیص همه ی اشکال‌ها را دشوار یا حتی غیرممکن می‌کند و چنین اشکال‌های باقی‌مانده در نهایت منجر به خرابی در زمان اجرا می‌شود. پیش بینی خرابی برخط^۱ تکنیکی است که با پیش بینی وقوع خرابی ها بر اساس تجزیه و تحلیل داده‌های گذشته و وضعیت فعلی یک سیستم، سعی در جلوگیری یا کاهش چنین خرابی‌هایی دارد. با توجه به تحولات اخیر فن آوری، الگوریتم‌های یادگیری ماشین توانایی خود در سازگاری و استخراج دانش را در انواع مختلف پیچیده نشان داده اند، و بنابراین برای OFP استفاده شده اند. با همه ی این‌ها هنوز هم بسیار وابسته به مشکل هستند و کارایی آنها می‌تواند تحت تأثیر عوامل مختلفی باشد. مشکلی که در اکثر کارها با استفاده از ML برای OFP وجود دارد این است که آنها فقط روی یک مجموعه کوچک از الگوریتم‌ها و تکنیک‌های پیش بینی تمرکز می‌کنند و هیچ راهی برای مطمئن شدن

از انتخاب درست این الگوریتم‌ها و تکنیک‌ها وجود ندارد.

- راه حل : در این مقاله، یک تحلیل اکتشافی از الگوریتم‌ها و تکنیک‌های مختلف یادگیری ماشین در یک مجموعه داده‌ی حاوی داده‌های خراب ارائه داده می‌شود. نتایج نشان می‌دهد، برای داده‌های یکسان، الگوریتم‌ها و تکنیک‌های مختلف به طور مستقیم بر کارایی پیش بینی تأثیر می‌گذارند و بنابراین باید با دقت انتخاب شوند.

12. BitBlade: Area and energy-efficient precision-scalable neural network accelerator with bitwise summation

- مسئله : شبکه‌های عصبی عمیق (DNN) بسته به کاربردها، نیازمند کارایی مختلف و دارای محدودیت‌های توان هستند. برای به حداکثر رساندن انرژی موثر شتاب دهنده‌های سخت‌افزاری برای برنامه‌های مختلف، شتاب دهنده‌ها باید از پیکربندی‌های مختلف بیتی پشتیبانی کنند. در هنگام طراحی شتاب دهنده‌های قابل بازپیکربندی بیتی، هر المان فرایند باید دارای منطق جمع و شیفت باشد، که مقدار زیادی مساحت و توان را به خود اختصاص می‌دهد. مسئله کاهش مساحت و توان در این شتاب دهنده‌ها است.
- راه حل : در این مقاله یک شتاب دهنده‌ی شبکه عصبی با ویژگی‌های دقت مقیاس پذیر، مساحت و انرژی موثر به نام BitBlade معرفی شده است، که باعث می‌شود سربار کنترل برای متغیر جمع و شیفت با استفاده از روش جمع بندی بیتی کاهش یابد. BitBlade پیشنهادی، هنگامی که در یک فناوری CMOS ۲۸ نانومتری سنتز می‌شود، باعث کاهش ۴۱ درصد در مساحت، و کاهش ۳۶-۴۶ درصد در انرژی نسبت به معماری قبلی می‌شود.

13. Shiraz: Exploiting system reliability and application resilience characteristics to improve large scale system throughput

- مسئله : افزایش چشمگیر نرخ خرابی و شکاف بین سیستم محاسباتی و I/O باعث می‌شود که در آینده استفاده از سیستم‌های محاسباتی موازی به طور چشم‌گیری کم شود. برنامه‌های کاربردی در مقیاس بزرگ به مکانیسم‌های قابلیت اطمینان مانند چک پوینت متکی هستند تا در حضور اشکال، همچنان کار کنند. اگرچه مکانیسم‌های چک پوینت می‌توانند شبیه سازی‌های علمی را به جلو پیش ببرند، اما نوشتن و خواندن در این

کاربردها منجر به سر بار I/O بزرگی می‌شود. کاهش سر بار I/O از مسائلی است که در این مقاله به آن پرداخته شده است.

- راه حل : برای کاهش این چالش، در این مقاله تکنیک جدیدی به نام *Shiraz* ارائه شده است، که این تکنیک نشان می‌دهد چگونه از اختلاف بین چک‌پوینت میان کاربردها و با توجه به دانشی که از ویژگی‌های زمانی خرابی‌ها به دست می‌آید استفاده کنیم تا در پهنای باند و کارایی کلی سیستم بهبود داشته باشیم. ایده اصلی این مقاله این است که برنامه‌هایی با سر بار چک پوینت بالاتر در بازه‌های با قابلیت اطمینان نسبتاً بالا (با میزان خرابی کمتر)، و برنامه‌های دارای سر بار چک پوینت پایین تر در بازه‌هایی با قابلیت اطمینان نسبتاً کم (با میزان خرابی بالاتر) زمان بندی شوند. این کار باعث می‌شود که پهنای باند بهبود یابد و سپس سر بار I/O نیز کم شود.

14. *DPS: Dynamic Precision Scaling for Stochastic Computing-based Deep*

Neural Networks

- مسئله : محاسبه تصادفی یک تکنیک با مزایایی مانند هزینه کم، توان مصرفی کم و مقاومت در برابر خطا است. با این حال، تاکنون شتاب دهنده‌های شبکه عصبی حلقوی^۱ مبتنی بر محاسبات تصادفی بخاطر دقت کم محاسبات تصادفی فقط در شبکه‌های عصبی حلقوی نسبتاً کوچک استفاده می‌شدند. همچنین، معماری‌های قبلی محاسبات تصادفی از قابلیت تغییر دقت به صورت پویا استفاده نمی‌کنند. استفاده از این قابلیت در بازدهی و انعطاف پذیری پیاده سازی شبکه عصبی حلقوی مبتنی بر محاسبات مهم است. پس مسئله ی این مقاله افزایش دقت محاسبات تصادفی در شبکه‌های عصبی حلقوی است.
- راه حل : در این مقاله یک شبکه ی عصبی حلقوی مبتنی بر محاسبات تصادفی (SC-CNN) ارائه شده است که دقت آن به صورت پویا قابل تغییر است و سر بار کمی هم دارد، و نه تنها برای شبکه‌های عصبی عمیق بزرگ کارآمد است، بلکه از شبکه‌های عصبی عمیق مختلف دیگر که نیاز به دقت‌های مختلف دارند، نیز پشتیبانی می‌کند. نتایج آزمایشات نشان می‌دهد که این طراحی نسبت به شبکه‌های عصبی حلقوی که کار ImageNet انجام می‌دهند بسیار کارآمدتر و دقیق تر است.

15. Availability analysis of safety critical systems using advanced fault tree and stochastic Petri net formalisms

- مسئله : تجزیه و تحلیل خرابی‌ها یکی از روش‌های تحلیل در دسترس بودن است. تجزیه و تحلیل خطر پیشرفت قابل توجهی در دو دهه گذشته داشته است. با این حال، حوادث در برخی سیستم‌ها همچنان ادامه دارد. این امر منجر به تعریف مفهوم سیستم‌های ایمنی-بحرانی می‌شود. اگر خرابی منجر به صدمات یا تلفات و خسارت قابل توجه به انسان یا محیط زیست و یا ضرر مهم اقتصادی شود، سیستم می‌تواند به عنوان یک سیستم ایمنی-بحرانی طبقه بندی شود. این موضوع اهمیت مطالعه ی پارامترهای سیستم‌های ایمنی-بحرانی را توجیه می‌کند. پس مسئله ی این مقاله ارزیابی یکی از پارامترهای اتکاپذیری با توجه به اهمیت موضوع سیستم‌های ایمنی-بحرانی است.
- راه حل : پس از بررسی روش‌های قبلی، نویسندگان این مقاله دو فرمول مجزا برای تجزیه و تحلیل سناریوهای خرابی و در دسترس بودن شناسایی کرده اند : شبکه‌های تصادفی عمومی پتری و فرآیندهای مارکوف درخت اشکال. فرمول FTDMP ترکیبی از فرآیند مارکوف و درخت اشکال است که برای غلبه بر محدودیتهای درخت اشکال و در عین حال استفاده از منطق استقرایی است. GSPN یک شبکه پتری با تجزیه و تحلیل احتمالی با استفاده از شبیه سازی مونت کارلو است.

16. Code-dependent and architecture-dependent reliability behaviors

- مسئله : امروزه نیاز به محاسبات و بهره‌وری بالاتر زیاد شده است. تنوع کدهایی که باید اجرا شوند همراه با پیچیدگی معماری، چالش‌های جدیدی را در ارزیابی قابلیت اطمینان سیستم‌ها و برنامه‌های محاسباتی به وجود می‌آورد. مسئله‌ای که این مقاله به آن پرداخته است این است که قابلیت اطمینان را در معماری‌ها و کدهای وابسته ارزیابی کند.
- راه حل : در این مقاله، رفتارهای قابلیت اطمینان شش معماری مختلف، یک پردازنده اینتل، سه GPU از شرکت انویدیا، یک APU از شرکت AMD و یک ARM با اجرای هشت کد مختلف مقایسه شده است. ابتدا هر دو مورد خرابی آرام داده (SCD) و نرخ خطاهای غیر قابل بازیابی تعیین می‌شود. سپس با در نظر گرفتن تفاوت بین مقادیر خراب و پیش بینی شده و همچنین بخشی از خروجی که خراب شده است، خطاهای مشاهده شده

را تحلیل می‌کنیم. از این تحلیل‌ها، ویژگی‌های قابلیت اطمینان را که مربوط به سخت‌افزار پایه و رفتارهای ذاتی کد اجرا شده است، شناسایی می‌کنیم.

17. Accelerating fpga prototyping through predictive model-based HLS design space exploration

- مسئله : حرکت به سمت سیستم‌های محاسباتی ناهمگن مبتنی بر شتاب دهنده‌های سخت‌افزاری پیچیده، باعث شده است که بسیاری از شرکت‌های طراحی VLSI از سنتز سطح بالا استقبال کنند. بالا بردن سطح انتزاع از RTL به سطح رفتاری مزیت‌های زیادی دارد که یکی از مهمترین آنها، این است که وقتی از یک مدار برای هدف خاصی استفاده می‌شود با تغییر تکنولوژی می‌توان از آن مدار با هدف دیگر استفاده کرد. با این حال کاوش در فضای طراحی بسیار وقت گیر است و به راحتی می‌تواند چندین روز برای طراحی‌های پیچیده طول بکشد.
- راه حل : پس به دلیل پیچیدگی در طراحی ASIC های بزرگ، تیم‌هایی که وظیفه ی تست و تایید این مدارها را دارند، از تست شبیه سازی سخت‌افزاری روی fpga و نمونه اولیه سازی برای آزمایش مدار استفاده می‌کنند. در این مقاله، روشی را ارائه داده اند تا به صورت خودکار طراحی های ASIC به عنوان توصیف‌های رفتاری برای HLS به FPGAها براساس مشاهداتی که در مرحله سنتز به دست می‌آید بهینه سازی مجدد شوند. این مقاله یک روش مدل سازی پیش بینی شده را ارائه می‌دهد که به عنوان ورودی نتایج یک ASIC HLS DSE را انجام می‌دهد و به طور خودکار، بدون نیاز به بررسی مجدد در توصیف رفتاری، معماری‌های میکرو معماری بهینه را برای FPGA هدف پیدا می‌کند.

18. Peregrine: A Flexible Hardware Accelerator for LSTM with Limited Synaptic Connection Patterns

- مسئله : تلفیق توانایی یادگیری عمیق در دستگاه‌های تلفن همراه یا IoT یک مبحث تحقیقاتی مهم محسوب می‌شود. بر اساس گزارش گوگل⁶ درصد استفاده از سرور شبکه عصبی حلقوی در مقایسه با حافظه کوتاه مدت (LSTM) ناچیز است. این بدین معناست که برای بهبود کارایی و انرژی کارآمد در موتورهای استنتاج LSTM باید تحقیقاتی به عمل آید. این مقاله نیز به این مسئله می‌پردازد.

- راه حل : در این مقاله یک راه حل جامع برای طراحی یک شتاب دهنده ی LSTM با کارایی بالا ارائه شده است. در واقع یک معماری سخت‌افزاری سریع و انعطاف پذیر با نام Peregrine پیشنهاد شده است که این معماری از الگوریتم تا طراحی سخت‌افزار توسط دسته‌ای از نوآوری‌ها پشتیبانی می‌شود. در این معماری ابتدا با محدود کردن الگوهای اتصال سیناپسی در شبکه ی LSTM، ردپای حافظه را به حداقل می‌رسانند. سپس Peregrine رمزگشایی موازی هافمن را با تطبیق کلاک برای انعطاف‌پذیری در طیف گسترده‌ای از سطوح پراکنده در ماتریس‌های وزنی فراهم می‌کند. تمام این ویژگی‌ها در یک معماری سخت‌افزاری جدید گنجانیده شده اند تا بهره وری انرژی را به حداکثر برساند.

ظاهرا اندازه فونت تغییر کرده است.

19. Design of a Safe Convolutional Neural Network Accelerator

- مسئله : امروزه شتاب دهنده‌های یادگیری ماشین با پیشرفت چشمگیر توان و کارایی، نسبت به CPU و GPU پیشرفت کرده اند. در چند سال اخیر، تکنیک‌های یادگیری ماشین برای دسته بندی دقت در کاربردهای مختلف زیاد استفاده می‌شوند. یادگیری ماشین به توان محاسباتی و پهنای باند بالا نیاز دارد. شتاب دهنده‌های یادگیری ماشین به دلیل توان و کارایی موثر در سیستم‌های ایمنی-بحرانی کاربرد زیادی دارند. این سیستم‌ها نیاز به ایمنی بالایی دارند و این مسئله از مواردی است که در این مقاله به آن‌ها پرداخته می‌شود.
- راه حل : در این مقاله یک الگوریتم مبتنی بر بررسی خطا برای تشخیص خطای همزمان مبتنی بر یک شبکه عصبی حلقوی با کیفیت صنعتی تولید شده است که اولویت این الگوریتم تامین نیازهای پوشش با ایمنی بالا و افزایش سطح کارایی و توان است. همچنین یک جستجوگر خوشه‌ای مبتنی بر الگوریتم با محلی سازی خطاهای درشت دانه برای بهبود دسترس پذیری در زمان اجرا توسعه داده شده است.

20. On the Evaluation Measures for Machine Learning Algorithms for Safety-critical System

- مسئله : امروزه توانایی الگوریتم‌های یادگیری ماشین بسیاری از سازنده‌های سیستم‌ها را تحریک کرده که از این الگوریتم‌ها در محصولاتشان استفاده کنند. حال برخی از این سیستم‌ها می‌توانند سیستم‌های بحرانی ایمنی

باشند که آسیب در این سیستم‌ها جان و مال انسان‌ها را در خطر می‌اندازد. بنابراین باید کارایی الگوریتم‌های یادگیری ماشین را قبل از استفاده در محیط عملیاتی از نظر نیاز به ایمنی بسنجیم.

○ راه حل : در این مقاله برای حل مشکل، با توجه به نمودار توزیع تعداد مشاهدات در برابر احتمال پیش‌بینی

شده در حالتی که پیشگویی نادرست است یعنی تشخیص مثبت ~~بوده ولی~~ پیشگویی منفی یا برعکس به جای

اینکه تصمیم گرفته شود و اشتباه باشد تصمیم " I do not Know " به عنوان راه حل ارائه شده است. در

واقع هدف این است که با استفاده از تحلیل روش گذشته به روش جدیدی دست یابد که در آن با توجه به

فرمول‌ها و تحلیل آن‌ها اثبات کند که روش جدید در مسئله ایمنی در سیستم‌های بحرانی-ایمنی سودمندتر

است.

دفعی دانم»