دانشگاه پیام نور

دانشکده فنی مهندسی

پایان نامه برای دریافت درجه کارشناسی ارشد

رشته مهندسی کامپیوتر – گرایش نرم افزار

گروه مهندسی کامپیوتر و فناوری اطلاعات

موازی سازی شبکه های عصبی پیشخور بر روی سیستم های چند پردازنده ای متقارن

سید مانی علایی

استاد راهنما:

دکتر سید علی رضوی ابراهیمی

استاد مشاور:

دکتر محمد هادی معظم

آبان ماه ۱۳۹۱

دانشگاه پیام نور

دانشکده فنی مهندسی

پایان نامه برای دریافت درجه کارشناسی ارشد

رشته مهندسی کامپیوتر – گرایش نرم افزار

گروه مهندسی کامپیوتر و فناوری اطلاعات

موازی سازی شبکه های عصبی پیشخور بر روی سیستم های چند پردازنده ای متقارن

سید مانی علایی

استاد راهنما:

دکتر سید علی رضوی ابراهیمی

استاد مشاور:

دکتر هادی معظم

آبان ماه ۱۳۹۱

بسم الله الرحمن الرحيم

تشکر و قدردانی:

سپاس خدای را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمت های او ندانند و کوشندگان، حق او را گزاردن نتوانند. و سلام و دورد بر محمّد و خاندان پاک او، طاهران معصوم، هم آنان که وجودمان وامدار وجودشان است؛

بدون شک جایگاه و منزلت معلم، اجّل از آن است که در مقام قدردانی از زحمات بی شائبه ی او، با زبان قاصر و دست ناتوان، چیزی بنگاریم. اما از آنجایی که تجلیل از معلم، سپاس از انسانی است که هدف و غایت آفرینش را تامین می کند و سلامت امانت هایی را که به دستش سپرده اند، تضمین؛ بر حسب وظیفه و از باب "**من لم یشکر المنعم من المخلوقین لم یشکر اللّه عزّ و جلّ** " ازپدر و مادر عزیزم این دو معلم بزرگوارم که همواره بر کوتاهی و درشتی من، قلم عفو کشیده و کریمانه از کنار غفلت هایم گذشته اند و در تمام عرصه های زندگی یار و یاوری بی چشم داشت برای من بوده اند؛ از همسر مهربانم که تمام دشواری های این اثر را با مهربانیها و محبت هایش برای من آسان کرد و از استاد با کمالات و شایسته؛ جناب آقای دکتر رضوی که در کمال سعه صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این رساله را بر عهده گرفتند؛ از استاد صبور و مهربان، جناب آقای دکتر معظم ، مدیریت محترم کرسی گروه، که زحمت مشاوره این رساله را کشیدند کمال تشکر و قدردانی را دارم. باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

تقدیم به:

کسانی که دوستشان دارم

چکیده

هدف اصلی کار ارائه شده در این پروژه افزایش سرعت آموزش شبکه‌های عصبی با استفاده از فرآیند موازی سازی می‌باشد. در این پروژه سعی بر آن شده است تا از الگوریتم پس انتشار خطا بر روی شبکه‌های روبه رشد استفاده گردد. امروزه کامپیوترهای موازی بزرگ برای آزمایشگاههای تحقیقاتی و دانشگاه ها در دسترس است. بسیاری از اجراهای شبکه‌های عصبی در کامپیوترهای موازی گزارش شده است. اگر چه تحقیق کاربردهای عصبی در این پروژه نشان می‌دهد که اتصال کوچکی بین توسعه دهندگان کاربرد شبکه های عصبی و محققان فعال در زمینه کاربردهای موازی شبکه عصبی وجود دارد. بسیاری از برنامه‌های آموزشی عصبی اجرایی بر اساس فرضیه‌های غیر واقعی در مورد کاربردهای واقعی آزمایش شده است. به خودی خود برنامه موازی می‌تواند به طور تأثیر گذاری برای یک رتبه بزرگ از کاربردهای عصبی اجرا شود. چندین کاربرد از عصب واقعی در این کار برای آزمایش الگوریتم‌های موازی اجرا شده استفاده شده است. نتایج نشان می‌دهد نفع در این است که از تابع انعطاف پذیر استفاده شود. برای سیستم‌های موازی کوچک، بسیاری از اجراهای موازی می‌توانند به طور موثر آموزش ببینند. اگر چه اهمیت یک کار کاربردی قابل انعطاف برجسته‌تر از تعداد پردازشگرهایی است که افزایش می‌یابد؛ بنابراین برای به دست آمدن منفعت بیشتر از سیستم موازی بزرگ و درجات اصلی چندگانه موازی سازی در الگوریتم آموزشی، می بایست روش ترکیب جنبه‌های الگوریتم آموزشی بر طبق کاربرد عصبی به روز شود. همچنین دو آزمایش بر روی همگرایی پس انتشار شبکه عصبی آموزش دیده گزارش شده است. نتایج نشان می‌دهد که اگر تقسیم بندی مجموعه آموزشی استفاده شود تعداد بیشتری از تکرارها برای همگرایی در طرح موازی نیاز است. اگر چه همیشه مجموع زمان آموزشی مور نیاز با استفاده از کامپیوتر موازی در رابطه با یک کامپیوتر ترتیبی از ساعت به دقیقه کاهش پیدا کرده است.

کلمات کلیدی: شبکه های عصبی، موازی سازی، محیط عملیاتی موازی

Abstract

The main object of the work present herein is to speed up neural network training using parallel processing. The back propagation trained feed – forward neural network was selected for this, research, since it has attracted most interest among neural network researchers.Today, large parallel computers are becoming more accessible to universities and research labs. Many implementations of neural networks on parallel computers have been report. However, survey of work on neural applications in this thesis indicates that there exists little contact between neural application developers and researchers in the field of parallel implementations of neural networks.Many implemented parallel neural training programs are tested based on unrealistic assumptions about real applications. Using large neural networks give rise to better performance then that obtainable in reality, since real neural networks are usually small. As such, the parallel program can execute efficiently for large of neural applications. Several real neural applications used in this work to test the implemented parallel algorithms. The result shows that it is beneficial to use flexible mapping.For small parallel systems; many parallel implementations can train efficiently. However, the importance of a flexible job assignment is more prominent as the number of processors increases. Thus, to gain the full benefit of a large parallel system the multiple inherent of degrees of parallelism in the training algorithm must be combined. Also the method of combining the parallel aspects of the training algorithm should be adaptable according to the given neural application.Two tests on convergence of back propagation trained neural network are also reported.The results indicate that if training set partitioning is used, in the parallel scheme, a larger number of iterations are needed for convergence. However, still the total training time required, by using a parallel computer, is reduced from hours to minutes.

اصطلاحات :

سلول	عنصر فرآیند سازی
دوره	تکرار رجوع شود
تکرار	چرخه کامل از آموزش الگوش ارائه شده
شبکه های عصبی	شبکه های عصبی مصنوعی
واحد	نورون

سر واژه ها :

ANN	شبکه عصبی مصنوعی
ASIC	کاربرد مدار مجتمع مشخص
ASIC	کاربرد مدار مجتمع مشخص
CPS	اتصالات در هر ثانیه
CPS	اتصالات در هر ثانیه
CPV	واحد مرکزی فرآیند سازی
CPV	واحد مرکزی فرآیند سازی
CUPS	اتصالات به روز شده در هر ثانیه
CVPS	اتصالات به روز شده در هر ثانیه

DRAM	RAM دینامیک (پویا)
DRAM RAM	دینامیک (پویا)
DSP	تک پردازشگر دیجیتالی
DSP	تک پردازشگر دیجیتالی
FLOPS	عملیات نقطه شناور در هر ثانیه
FLOPS	عملیات نقطه شناور در هر ثانیه
FPGA	آرایش مدار ورودی قابل برنامه ریزی
FPGA	آرایش مدار ورودی قابل برنامه ریزی
FPV	نقطه واحد شناور
FPV	نقطه واحد شناور
LCA	آرایش سلول منطقی
LCA	آرایش سلول منطقی
MEM	دستورالعمل جریان چند لایی، جریان اطلاعات چند لایه
MEM	دستورالعمل جریان چند لایی، جریان اطلاعات چند لایه
MLP	شبکه ادراکی چند لایه
MLP	شبکه ادراکی چند لایه
NTH	آزمایشگاه نروژی تکنولوژی
NTH	آزمایشگاه نروژی تکنولوژی

PE	عنصر پردازشی
RAM	دسترسی حافظه به طور تصادفی
RAM	دسترسی حافظه به طور تصادفی
RENNS	کمک کننده شبکه عصبی با قابلیت طراحی مجدد
RENNS	کمک کننده شبکه عصبی با قابلیت طراحی مجدد
SIMD	دستور العمل تک جریان، مجریان اطلاعات چند لایه
SIMD	دستور العمل تک جریان، مجریان اطلاعات چند لایه
SRAM	RAM ایستا (ساکن)
VLSI	مقیاس مجتمع سازی (انتگرال گیری) بسیار بزرگ
VRAM	ویدئو RAM

علائم:

$\propto$	نسبت به
C	عدد کلی عناصر فرآیند سازی (سلولها)
C_X، C_y	عدد کلی عناصر پردازشیدر ابعاد y،x
$C_{x,y}$	فرایند سازی شاخص سلولی 2-D
$D_{p,k}$	مقدار نشانگاه برای الگو p واحد خروجی k
E	آستانه خطا

E_p	میزان خطا برای الگو p
E	میزان خطا برای مجموعه آموزشی کامل
$E_{\%}$	درصد الگوهایی که به طور کامل آموزش ندیده اند
E_{RMSE}	میانگین خطای توان دوم
F	تابع هلالی شکل
E	تعداد عملکرد های نقطه شناور
K_e	نسبت عدد تکرار برای یادگیری هر دوره
L	طول بردار
M	عدد متوسط واحد پردازشگر
N	نیاز به هیچ تکرار آموزشی وجود ندارد
N_i	تعداد واحدهای ورودی
N_n	تعداد واحدهای مخفی
N_o	تعداد واحد های خروجی
N	شاخص تکرار آموزشی
N_i	تعداد بخشهای ورودی تعیین شده برای یک پردازشگر
N_h	تعداد بخشهای مخفی تعیین شده برای یک پردازشگر
N_o	تعداد بخشهای خروجی تعیین شده برای یک پردازشگر
N	شاخص تکرار آموزشی

P	عدد کلی الگوها در مجموعه آموزشی
P_c	تعداد الگوهای آموزشی پردازشگرC
P	شاخص الگو
t_x	زمان استفاده شده برای عملکرد x در ثانیه
T_x	زمان کل استفاده شده برای عملکرد x در ثانیه
$W_{h,ji}$	مقدار وزن، واحد ارتباطی ورودی i و واحدهای مخفی j
$W_{o,kj}$	مقدار وزن، واحد ارتباطی ورودی j و واحدهای مخفیk
$\Delta W_{h,ji}$	تعداد تغییر وزن برای وزن $w_{h,ji}$
$\Delta W_{o,kj}$	تعداد تغییر وزن برای وزن $W_{o,kj}$
x_i	مقدار ورودی به نورون i
$y_{o,k}$	مقدار خروجی از واحد خروجی

حروف یونانی:

$\propto$	تکانه
δ	بیان تکرار آموزشی به عنوان تابعی از μ
θ	مقدار پیشقدر
$\delta_{h,j}$	مقدار خطای دلتا برای واحد مخفی j

$\delta_{o,k}$	مقدار خطای دلتا برای واحد خروجی k
η	نسبت یادگیری

فهرست مطالب :

عناوین **صفحه**

فصل ۱ ..... ۳

۱-۱مقدمه ۴

۱-۲ تعریف مساله و بیان سوال های اصلی تحقیق ۴

۱-۳سابقه و ضرورت انجام تحقیق ۵

۱-۴فرضیه ها ۵

۱-۵هدف ها ۵

۱-۶کاربردها ۶

۱-۷جنبه نوآوری تحقیق ۶

۱-۸روش تحقیق ۶

۱-۹مراحل انجام تحقیق ۷

۱-۱۰ساختار پایان نامه ۷

فصل ۲ ..... ۸

۲-۱ مقدمه ۹

۲-۲ نحوه کار شبکه‌های عصبی ۱۰

۲- ۳ تاریخچه پیدایش شبکه‌های عصبی ۱۳

۲-۴ کاربرد شبکه‌های عصبی ۱۵

۲-۵ متداول‌ترین انواع شبکه‌های عصبی ۱۶

۲-۵-۱ پرسپترون ساده ۱۶

فهرست مطالب :

عناوین **صفحه**

۲-۵-۲ الگوریتم پس انتشار- چند لایه ۱۷

۲-۵-۳ شبکه‌های پیش خور ۱۹

۲-۵-۴ شبکه‌های باز خور ۱۹

۲-۵-۵ شبکه‌های خطی ۱۹

۲-۵-۶ نقشه‌های ویژگی کوهونن ۱۹

۲-۶ پردازش اطلاعات در شبکه‌های عصبی ۲۰

۲-۷ آموزش شبکه‌های عصبی ۲۱

۲-۸ نتیجه گیری: ۲۴

فصل ۳ ..... ۲۵

۳-۱ مقدمه ۲۶

۳-۲ شبکه عصبی چند لایه ۲۷

۳-۲-۱ توضیحاتی جزئی‌تر از یادگیری الگوریتم BP ۲۸

۳-۲-۲ توجه اساسی به الگوریتم Bp ۳۵

۳-۲-۲-۱ مجموعه آموزشی : ۳۵

۳-۲-۲-۲ اندازه شبکه: ۳۶

۳-۲-۲-۳ سه رویکرد متفاوت از استراتژی‌های به روز رسانی وزن‌ها: ۳۶

۳-۲-۲-۴ مدت به روز رسانی وزن: ۳۶

۳-۲-۲-۵ درجات موازی سازی: ۳۷

فهرست مطالب :

عناوین **صفحه**

۳-۲-۳ تکانه ۳۸

۳-۲-۴ اجرای یادگیری ۳۹

۳-۲-۵ تعمیم دادن ۴۲

۳-۲-۶ پیشرفت شبکه‌های عصبی رو به رشد ۴۳

۳-۲-۶-۱ تکنیک‌های فرایند سازی ۴۴

۳-۲-۷ تأثیر مدت بروزرسانی وزن ۴۴

۳-۳ کاربردهای شبکه عصبی ۴۶

۳-۳-۱ فرایند سازی تصویر ۴۷

۳-۳-۱-۱ تصویرهای ماهواره‌ای ۴۷

۳-۳-۱-۲ تراکم تصویر ۴۷

۳-۳-۲ آیا کاربرد شبکه عصبی نیاز به سخت افزار موازی دارد؟ ۴۸

فصل ۴ ..... ۵۱

۴-۱ مقدمه ۵۲

۴-۲ محاسبه توصیفی برای درجه موازی‌سازی BP ۵۲

۴-۲-۱ آموزش مجموعه‌های موازی ساز ۵۲

۴-۲-۲ کانال بندی ۵۴

۴-۲-۳ گره موازی‌سازی ۵۴

۴-۲-۳-۱ موازی‌سازی نرون (رشته عصبی) ۵۵

فهرست مطالب :

عناوین **صفحه**

۴-۲-۳-۲ موازی‌سازی سیناپس ۵۷

۴-۲-۴ اهمیت هر یک از درجات موازی‌سازی ۵۸

۴-۳ محاسبه کننده‌های موازی برای شبیه‌سازی شبکه‌های عصبی مصنوعی ۵۸

۴-۳-۱ جنبه‌های کلی پردازش موازی ۵۹

۴-۳-۱-۱ تقسیم بندی فلاین: ۶۱

۴-۴ تحقیقی بر اجراهای موازی‌سازی متفاوت ۶۲

۴-۴-۱ اهداف کلی اجراهای کامپیوتر (محاسبه‌گر) ۶۳

۴-۴-۲ ایستگاه کاری به هم پیوسته ۶۶

۴-۴-۳ تحقیق در مورد استفاده از مدل‌های ماشین‌های موازی ۶۶

۴-۴-۳-۱ سیستم‌های هدایت داده‌ها ۶۶

۴-۴-۳-۲ مکان‌یابی شبکه ۶۷

۴-۴-۴ هدف سخت افزار خاص در مقابل عام ۷۱

فصل ۵ ..... ۷۳

۵-۱ مقدمه ۷۴

۵-۲ پردازش موازی چیست؟ ۷۵

۵-۲-۱ کاربردهای پردازش موازی ۷۵

۵-۲-۲ محاسبه موازی ۷۶

۵-۲-۲-۱ زیاد شدن توان عملیاتی : ۷۷

فهرست مطالب :

عناوین **صفحه**

۵-۲-۲-۲ تحمل پذیری در برابر خطا : ۷۷

۵-۳ تقسیم بندی سیستم عامل‌های چند پردازنده‌ای : ۷۸

۵-۴ پردازش موازی در واحد پردازش گرافیکی - GPU ۷۹

۵-۴-۱ کودا چیست؟ ۸۱

۵-۵ تشخیص اعداد دستنوشته: ۸۴

۵-۵-۱ استخراج ویژگی‌ها ۸۵

فصل ۶ ۸۷

الف –پیاده سازی موازی شبکه‌های عصبی پیشخور- اجرا بر روی GPU ۸۸

ب –پیاده سازی سریال شبکه‌های عصبی پیشخور- اجرا بر روی CPU ۱۴۷

ج- ارائه جدول زمان مقایسه ای اجرای شبکه عصبی در حالت سریال و موازی ۱۹۱

د- ارائه نمایش زمان اجرای پروژه در حالت آموزش شبکه به صورت سریال و موازی ۱۹۲

پیشنهادات: ۱۹۴

مراجع: ۱۹۵

کتاب‌ها و مجلات : ۱۹۹

منابع : ۱۹۹

لغت نامه انگلیسی به فارسی ۲۰۱

فهرست فرمول ها :

عناوین **صفحه**

شکل۲- ۱ نمایی از یک نرون طبیعی[Laurene V. Fausett] ۱۰

شکل۲- ۲ نمایی از یک پرسپترون ساده [Laurene V. Fausett] ۱۷

شکل۳- ۱ مدل ادراکی [Jim Torresen,1996] ۲۶

شکل۳- ۲ شبکه عصبی رو به رشد با دو لایه وزن [Jim Torresen,1996] ۲۷

شکل۳- ۳ تابع سیگموئید [Jim Torresen,1996] ۲۹

شکل۳- ۴ تعداد اجرای نقطه تغییر برای محاسبه و به روز کردن وزن برای وزن‌های بلوک flbp [Jim Torresen,1996] ..۴۱

شکل۳- ۵ خطای مجموعه آموزشی در مقابل خطای آزمون به عنوان یک تابع عددی از تکرار آموزشی از [۴۹] ۴۳

شکل۳- ۶ خطای کل آموزش برای فرکانس‌های مختلف وزن به روز شده [Jim Torresen,1996] ۴۵

شکل۴- ۲ نقشه وزن ماتریسها برای کانال بندی [Jim Torresen,1996] ۵۳

شکل۴- ۱ آموزش مجموعه موازی سازی برای یادگیری الفبای انگلیسی [Jim Torresen,1996] ۵۳

شکل۴- ۳ کانال بندی برای آموزش الگوها [Jim Torresen,1996] ۵۴

شکل۴- ۴ اصول سه عنصر فرآیندسازی [Jim Torresen,1996] ۵۵

شکل۴- ۵ موازی سازی سیناپس [Jim Torresen,1996] ۵۷

شکل۴- ۶ تعدادی از مکان یابها برای مرتبط ساختن PE ها وجود دارد [۱۴۷] [Jim Torresen,1996] ۶۰

شکل۴- ۷ اجرای شبکه لوله‌ای [۱۳۰] [Jim Torresen,1996] ۶۸

شکل۴- ۸ حلقه اندازه N2 محدود به حلقه درازتر با سایر N1 [۱۱۹] [Jim Torresen,1996] ۷۰

شکل۴- ۹ رابطه بین سرعت و ترکیب بندی مجدد سخت افزار [Jim Torresen,1996] ۷۲

شکل۵- ۱ مقایسه معماری GPU با CPU [IE-GPU] ۸۳

شکل۵- ۲ پنجره توجه ۸۶

شکل۵- ۳ تقسیم بندی عدد دستنوشته ۸۶

شکل۵- ۴ نمایش تعداد برخورد بین خطوط در عدد دستنوشته ۸۷

فهرست فرمول ها :

عناوین **صفحه**

فرمول۲- ۱ ..... ۱۷
فرمول۳- ۲ ..... ۲۸
فرمول۳- ۳ ..... ۲۹
فرمول۳- ۴ ..... ۲۹
فرمول۳- ۵ ..... ۳۰
فرمول۳- ۶ ..... ۳۰
فرمول۳- ۷ ..... ۳۰
فرمول۳- ۸ ..... ۳۱
فرمول۳- ۹ ..... ۳۱
فرمول۳- ۱۰ ..... ۳۱
فرمول۳- ۱۱ ..... ۳۱
فرمول۳- ۱۲ ..... ۳۲
فرمول۳- ۱۳ ..... ۳۲
فرمول۳- ۱۴ ..... ۳۲
فرمول۳- ۱۵ ..... ۳۲
فرمول۳- ۱۶ ..... ۳۳
فرمول۳- ۱۷ ..... ۳۳
فرمول۳- ۱۸ ..... ۳۳
فرمول۳- ۱۹ ..... ۳۳
فرمول۳- ۲۰ ..... ۳۴
فرمول۳- ۲۱ ..... ۳۴
فرمول۳- ۲۲ ..... ۳۴
فرمول۳- ۲۳ ..... ۳۴
فرمول۳- ۲۴ ..... ۳۵
فرمول۳- ۲۵ ..... ۳۵
فرمول۳- ۲۶ ..... ۳۸
فرمول۳- ۲۷ ..... ۳۸
فرمول۳- ۲۸ ..... ۳۸