عنوان الرسالة : مسرع الأجهزة المستند إلى مجموعة بوابة قابلة للبرمجة الميدانية لوظيفة دالة السوفت ماكس

اسم المشرف : د/ محمد بلال

اسم الطالب : محمد عبدالله مبارك الشهراني

ملخص الرسالة (عربي) :

تعد وظيفة Softmax جزءًا لا يتجزأ من أطر اكتشاف الكائنات استنادًا إلى جميع الشبكات العصبية العميقة أوالضحلة. في حين أن تكوين طبقات التشغيل المختلفة في الشبكة العصبية يمكن أن يكون مختلفًا تمامًا، فإن عملية softmaxثابتة. مع التطورات الحديثة في مناهج اكتشاف الكائنات، خاصة مع إدخال شبكات عصبية تلافيفية عالية الدقة، اقترح الباحثون والمطورون بنى أجهزة مختلفة لتسريع التشغيل الكلي لهذه الخوارزميات كثيفة الحوسبة Xilinx .، أحد بائعي FPGA الرائدين ، قدم مؤخرًا مجموعة أدوات تطوير الشبكة العصبية العميقة لهذا الغرض بالضبط. ومع ذلك، نظرًا للطبيعة المعقدة لأجهزة softmaxالحسابية التي تتضمن وظيفة أسية، فإن هذه الوظيفة متاحة فقط للأجهزة الأكبر حجمًا. بالنسبة للأجهزة الأصغر، لا بد أن يتم تنفيذ هذه العملية في البرامج. في هذه الأطروحة، سوف نستكشف تنفيذ الأجهزة لهذه الوظيفة في ضوء الأعمال السابقة المذكورة في الأدبيات. سيكون الهدف الرئيسي من هذه الدراسة هو التوصل إلى تنفيذ فعال ومنخفض التكلفة ومناسب للأجهزة الصغيرة أيضًا. لهذا الغرض، نعتزم النظر في عمليات التنفيذ الخطية التقريبية والقطعة أيضًا. سيتم تنفيذ المنتج النهائي كمسرع للأجهزة مع واجهة ناقل قياسية في الصناعة بحيث يمكن دمجها بسلاسة مع معالج Xilinx والمنطق القابل للبرمجة. أحد الأمور الأخرى المثيرة للقلق هو تصميم وتطوير واختبار أنظمة الأجهزة الكاملة التي تتطلب الكثير من الجهد والوقت. وبالتالي، هناك اتجاه متزايد لوصف سلوك الدائرة باستخدام أدوات التركيب عالية المستوى لتقليل وقت التطوير وإنفاق المزيد من الجهد على مستوى التصميم الحسابي. توفر مربعات أدوات المبرمج بلغة وصف الأجهزة ((HDL و Vision HDL في®Simulink بواسطة ®Mathworks تجريدًا عالي المستوى لمعالجة الصور وخوارزميات التعلم الآلي للنشر السريع على مجموعة متنوعة من الأجهزة المستهدفة. ونتيجة لذلك، يتم تقصير دورة تطوير تطبيقات أجهزة معالجة الوسائط المتعددة باستخدام هذه الأدوات. وبالتالي، بهذه الروح، سيتم تنفيذ المشروع باستخدام أدوات توليف عالية المستوى متوفرة في بيئة ماتلاب. سيستخدم إعداد الاختبار الشبكات العصبية القياسية مثل Resnet و Squeezenet وما إلى ذلك لتحديد الدقة الوظيفية. سيتم نقل الإعداد بالكامل إلى لوحة تطويرXilinx FPGA ، أي Zedboard التي تحتوي على مكونات الأجهزة الضرورية مثل واجهات USB و Ethernet وHDMI وما إلى ذلك لتنفيذ نظام يعمل بالكامل قادر على معالجة تطبيق التعلم الآلي في الوقت الفعلي.

Thesis Title: FPGA-based Hardware Accelerator for SoftMax Function

2021

SUPERVISOR: Dr. Mohammed Belal

STUDENT NAME: Mohammed Abdullah Mubarak Alshahrani

STUDENT ID: 1900641

Research Summary(English):

Softmax function is an integral part of object detection frameworks based on all deep or shallow neural networks. While the configuration of different operation layers in a neural network can be quite different, softmax operation is fixed. With the recent advances in object detection approaches, especially with the introduction of highly accurate convolutional neural networks, researchers and developers have suggested different hardware architectures to speed up the overall operation of these compute-intensive algorithms. Xilinx, one of the leading FPGA vendors, has recently introduced a deep neural network development kit for exactly this purpose. However, due to the complex nature of softmax arithmetic hardware involving exponential function, this functionality is only available for bigger devices. For smaller devices, this operation is bound to be implemented in software. In this thesis, we will explore the hardware implementation of this function in the light of the earlier works reported in the literature. The main objective of this study will be to come up with an efficient and low-cost implementation that is suitable for smaller devices as well. For this purpose, we intend to look into approximate and piece-wise linear implementations as well. The final product will be implemented as a hardware accelerator with industry standard bus interface so that it can be integrated seamlessly with Xilinx processor and programmable logic. One other matter of concern is the design, development and testing of full hardware systems which requires a lot of effort and time. Thus, there is an increasing trend of describing the circuit behavior using high-level synthesis tools to reduce the development time and spend more effort at the algorithmic design level. Hardware Description Language (HDL) Coder and Vision HDL toolboxes in Simulink® by Mathworks® provide higher level abstraction of image processing and machine learning algorithms for quick deployment on a variety of target hardware. Resultantly, the development cycle of multimedia

processing hardware applications using these tools gets shortened. Thus, in this spirit, the project will be implemented using high-level synthesis tools available in Matlab environment. The test setup will use standard neural networks such as Resnet and Squeezenet etc. to determine the functional accuracy. The entire setup will be ported to a Xilinx FPGA development board i.e. Zedboard which contains the necessary hardware components such as USB, Ethernet and HDMI interfaces etc. to implement a fully working system capable of processing a machine learning application in real-time.