

Implementação de redes neurais artificiais em FPGAs para a seleção de partículas subatômicas

Isabella Silva de Campos¹, Juliano Alves de Oliveira² e Renato Giacomini¹

² Departamento de Elétrica, Centro Universitário FEI

unieiscampos@fei.edu.br

Resumo: O trabalho introduz o estudo da implementação de uma rede Neural Artificial (RNA) em uma placa FPGA para a análise de partículas subatômicas, a fim de classificá-las em tempo real. A solução proposta permite um treinamento offline das redes e a classificação final da implementação em “hardware description language” (VHDL).

1. Introdução

Um acelerador de partículas é conhecido por aumentar a velocidade das partículas para analisar seus comportamentos [1]. Em 2013, foi através do maior acelerador de partículas do mundo, o Large Hadron Collider (LHC), localizado na Suíça, que descobriram uma partícula chamada Bóson de Higgs [1], contudo é através da placa FPGAs junto a programação de descrição hardware VHDL (VHSIC Hardware Description Language) que há o processamento dos dados dos elementos. Devido ao grande volume de feixes gerados pelo LCH, o número de eventos é incontável, gerando um grande volume de dados, sendo necessário a compilação dos dados gerados por esses aceleradores, afim de se reter apenas eventos de interesse de estudo. Os sinais são recebidos pela placa FPGA que de forma eficaz, consegue fazer o processamento dos dados, entretanto não é tão ágil, uma vez que são guardados em um banco de dados e não são compilados instantaneamente.

O uso de ferramentas como Field Programmable Gate Arrays (FPGAs), auxiliam em desenvolvimentos de circuitos de sistemas digitais. A placa é um dispositivo lógico onde as conexões das entradas e saídas são determinadas através da linguagem de hardware. A linguagem de descrição de hardware (do inglês, Hardware Description Language -HDL) foi desenvolvida a fim de descrever circuitos integrados muito complexos e o uso do VHDL (Very High Speed Integrated Circuits) tem o objetivo de facilitar o design dos circuitos lógicos [2][3].

As Redes Neurais Artificiais (RNAs) fazem parte do campo de estudo da Inteligência Artificial, e foram inspiradas inicialmente pela unidade básica do cérebro, o neurônio pelo fato que, o cérebro humano é capaz de processar diversas informações simultaneamente, através de uma rede complexa de neurônios [4].

Os neurônios são arquitetados de acordo com a realidade computacional e projetados para cumprir uma tarefa específica. Cada entrada x_i de um neurônio tem um peso w_i respectivo, que após sua combinação, passa por uma “função de ativação”, a qual é responsável por definir o nível de ativação da saída [4], como demonstrado na Figura 1 abaixo.

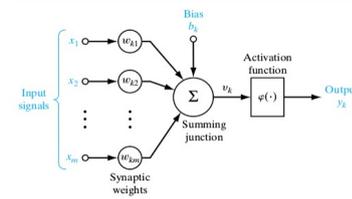


Figura 1 – Representação da arquitetura de uma rede neural artificial [5].

A adoção de redes neurais em sistemas de classificação de dados mostra grandes vantagens em relação a métodos clássicos de classificação [5], pois não são necessárias a criação de parâmetros de forma manual para caracterizar o evento, pelo fato que o próprio processo de treinamento da rede neural nos dados brutos permite a extração dessas características, tornando o desenvolvimento de sistemas de análise mais acelerados e precisos.

Para alcançar os objetivos descritos, foi necessário encontrar a melhor maneira de treinar e implementar a rede desejada em descrição de hardware. As aplicações e desenvolvimentos foram demonstradas nas próximas seções do artigo.

2. Métodos e Materiais

Para iniciar a familiarização com as redes neurais artificiais, foi necessário estudar sobre suas funções de ativações, aprender métodos de implementação e escolher a melhor maneira de trabalhar com elas.

Um estudo sobre a implementação das Redes Neurais artificiais demonstrou que a Multilayer Perceptron (MLP) pode se adaptar em quantas camadas forem desejadas, uma vez que as MLPs usam funções não lineares, facilitando na adaptação do modelo com base de dados variados [4].

Uma implementação de uma RNA, foi estudada, baseado no estudo de Carlos Alberto de Albuquerque [6]. Como primeiros passos na pesquisa, foi descrita uma função conhecida para melhor compreensão das redes. Foi arquitetada uma descrição em VHDL para o comportamento de uma RNA para desempenhar a função de uma porta lógica XOR. A rede resultante possuía duas entradas, duas camadas ocultas e uma saída. O Matlab foi usado para o treinamento da rede e passagem dos dados de ponto fixo para linguagem de hardware.

Através do aplicativo fornecido pelo Mathlab, “HDL Code”, foi possível, computar a rede neural artificial em VHDL e assim, transferir a descrição para o Quartus Prime II da Intel.

O software usado para a criação e compilação do VHDL foi o Quartus II da Intel. Ademais, foi possível simular o circuito resultante, além de criar uma visão RTL dos blocos.

Partindo para uma aplicação mais complexa em termos de classificação, a partir do artigo publicado de em equipe sobre análises e simulações em um transistor [7] foi possível utilizar os dados extraídos através do acelerador de partículas Pelletron, localizado no Laboratório Aberto de Física Nuclear (LAFN). Os dados encontrados, são diferentes tipos de íons, os quais foram separados por pulsos no Matlab. Na Figura 2 abaixo, cada cor representa um íon diferente.

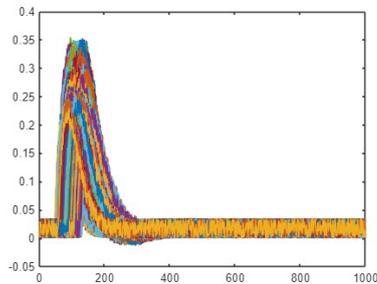


Figura 2 – Pulsos de íons extraídos do Acelerador de Partículas Pelletron

Com isso, foi possível encontrar a melhor maneira de classificação dos dados. Com o auxílio do “Classification Learner”, do Matlab, foi possível chegar em uma porcentagem de 85.2% de eficácia para as Redes Neurais artificiais.

3. Resultados

Enquanto a rede é treinada, o Matlab abre um comando em inglês, chamado de “Neural Network Training Tool” que demonstra a arquitetura da porta lógica XOR em Rede Neural Artificial, incluindo o número de entradas (inputs), camadas ocultas (hidden) e saídas (outputs), apontados na Figura 3.

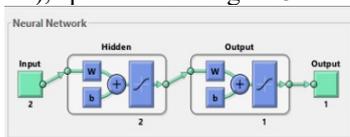


Figura 3 – RNA da Porta XOR

O processo de treinamento pode ser avaliado pela performance de erro. A figura 4 demonstra a melhor performance da rede, a qual foi 0,5054 em 5000 épocas (epoch).

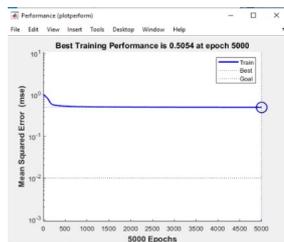


Figura 4 – Performance da porta lógica XOR em RNA

Os resultados das bias e pesos da rede foram convertidos para ponto fixo, facilitando a passagem para a representação em hardware. Após a descrição dos blocos em VHDL, foi possível compilar as waveforms da porta XOR em rede neural artificial além de implementar na placa FPGA. A figura 5 demonstra a waveform que representa a entrada “0011” resultando no número “90” em saída de FPGA.

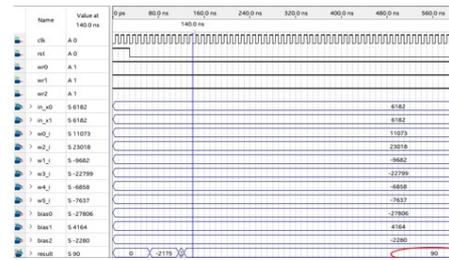


Figura 5 – Waveform que representa a porta lógica XOR

Em seguida, com os dados extraídos de íons pelo Pelletron, foi possível selecionar cinco amostras para criação de uma rede simples e eficaz. Sua arquitetura foi representada na Figura 6 abaixo.

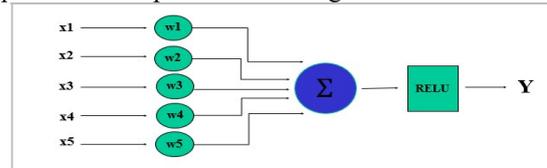


Figura 6 – Arquitetura da rede neural para seleção de partículas

Em seguida, será necessário implementar rede neural em VHDL no Quartus Prime, simulando seu comportamento através da simulação na placa FPGA.

4. Conclusões

Através dos estudos durante um ano, notou-se que é possível implementar redes neurais integrando vertentes oferecidas pelo Mathwork’s Matlab e Quartus Prime da Intel. O Matlab demonstrou ser capaz e assertivo no treinamento das redes, enquanto o Quartus auxiliou na simulação e compilação da descrição em VHDL.

5. Referências

- [1] CERN, “The Higgs Boson,” 29 May 2021. [Online]. Available: <https://home.cern/science/accelerators/large-hadron-collider>.
- [2] R. J. W. N. S. M. G. TOCCI, Digital Systems-Principles and Application, São Paulo: Pearson Education do Brasil, 2007.
- [3] P. D. R. Giacomini, *Apostila Básica de VHDL*.
- [4] S. Haykin, NeuralNetwork: A comprehensive Foundation, Ontário,Canada.
- [5] G. D. S. KOROL, “AN FPGA IMPLEMENTATION FOR CONVOLUTIONAL NEURAL NETWORK,” pp. 13-24, 2019.
- [6] C. d. A. A. Siva, "Contribuição para o estudo do embarque de uma rede neural artificial em field programmable gate array (FPGA)," Natal, 2010.
- [7] J. Oliveira, M. A. Guazzelli e M. A. Assis, “Single event effect: Simulations and analysis on 3N163 PMOS transistor,” *IEEEEXPLORE*, 12-14 Março 2018.

Agradecimentos

À instituição Centro Universitário FEI pelo do suporte financeiro.

¹ Aluno de IC do Centro Universitário FEI (PIBIC-FEI). Projeto com vigência de 08/2021 a 08/2022.