

SALÃO DE  
INICIAÇÃO CIENTÍFICA  
**XXIX SIC**  
**UFRGS**  
PROPESQ



múltipla   
**UNIVERSIDADE**  
inovadora  inspiradora

<b>Evento</b>	Salão UFRGS 2017: SIC - XXIX SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
<b>Ano</b>	2017
<b>Local</b>	Campus do Vale
<b>Título</b>	Análise de Confiabilidade de Redes Neurais em SoC
<b>Autor</b>	FABIANO PEREIRA LIBANO
<b>Orientador</b>	FERNANDA GUSMAO DE LIMA KASTENSMIDT

## **Análise de Confiabilidade de Redes Neurais em SoC**

Fabiano Pereira Libano, Fernanda Gusmão de Lima Kastensmidt (Orientadora)

Instituto de Informática - Universidade Federal do Rio Grande do Sul

O projeto de sistemas de heterogêneos, que combinam recursos de hardware e software para a implementação de algoritmos tradicionais, mostra-se atualmente como uma tendência do mercado, e conseqüentemente, uma vertente importante da pesquisa científica. A adoção de tais sistemas justifica-se pelo considerável acréscimo de desempenho com a divisão de workload, visto que muitos algoritmos possuem trechos sequenciais e paralelizáveis. Para que o desenvolvimento de tais sistemas fosse possível, utilizou-se a plataforma híbrida ZedBoard (Xilinx), que através do chip Zynq, apresenta uma porção de hardware reconfigurável (FPGA), além de dois cores ARM, dedicados ao software.

Dentre os algoritmos explorados, destacam-se as Redes Neurais, dada sua natureza altamente paralelizável, e a grande gama de aplicações safety-critical onde as mesmas podem ser utilizadas. Neste projeto, decidiu-se explorar duas redes com finalidades distintas. A primeira, voltada para a classificação de instâncias, com base em features extraídas de imagens, ou seja, reconhecimento de padrões. A segunda, comportando-se como um regressor linear, em outras palavras, um preditor de uma variável Y em função de N entradas.

A partir de então, teve início a descrição do hardware de ambas as redes em VHDL, sendo que um dos maiores obstáculos encontrados, foi a implementação da função de ativação dos neurônios que compõem a rede, por tratar-se de uma expressão exponencial. A solução adotada foi a discretização da curva contínua. É importante destacar que ambas as redes foram testadas com três níveis de discretização, totalizando seis designs diferentes. Como citado anteriormente, o SoC utilizado também possui cores ARM, que ficaram responsáveis pelo estímulo do design implementado na FPGA, bem como pela posterior avaliação de corretude dos resultados. A comunicação e troca de dados entre FPGA e ARM, se deu via interface AXI (padrão Xilinx).

O passo seguinte da pesquisa, consistiu na avaliação da confiabilidade das redes neurais. Tal avaliação baseia-se nos resultados de dois experimentos fundamentalmente diferentes: teste controlado sob radiação usando um acelerador de partículas (realizado no Laboratório de Física da USP), e campanha exaustiva de injeção de falhas na memória de configuração da FPGA. Os dados coletados no experimento com radiação fornecem a probabilidade da ocorrência de erros e defeitos (erros críticos ao funcionamento da aplicação), enquanto que com a injeção de falhas, conseguimos avaliar a probabilidade de uma falha afetar um output. Resultados preliminares, demonstram que há uma predominância de erros em relação à defeitos, e que falhas nas primeiras camadas das redes têm maior probabilidade de afetar o output, devido à característica de conexão total entre neurônios de camadas vizinhas.

No futuro pretende-se avaliar o comportamento de redes consideravelmente maiores, utilizando para tanto, SoCs mais capazes, de modo a verificar que os resultados obtidos até o momento podem ser estendidos, isto é, escaláveis conforme a dimensão e o propósito de aplicação de uma rede neural genérica.