Synthesis of a Asynchronous Multilayer Perceptron on an FPGA

Fabian Franz

October 2020

Abstract

Chapter 1

Introduction

This project has the claim to design a low-power neural network on an FPGA. To do so, the next sections give a brief introduction to the basic principles of how such a neural network can be modeled.

1.1 Perceptron

A perceptron describes an analog model of a biological human cell in the computer domain. This perceptron can be described with the following graphical and mathematical expressions:

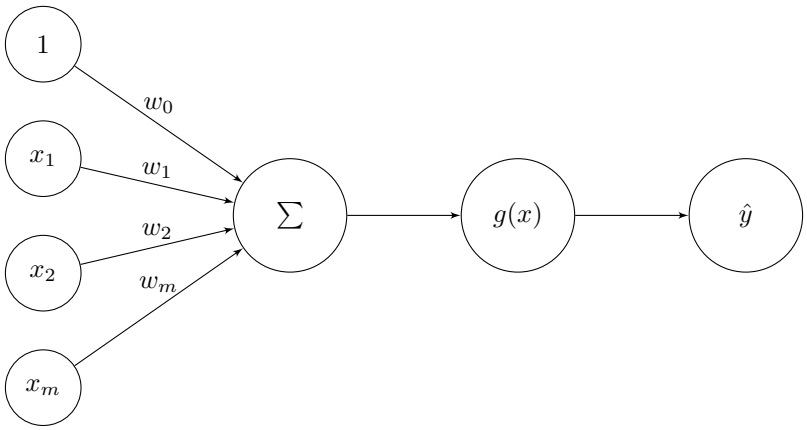

Figure 1.1: Model of perceptron

$$\hat{y} = g\left(w_0 + \sum_{i=1}^{m} x_i w_i\right) \tag{1.1.1}$$

With:

g = Activation function
w_o = Bias

In vector form:

$$\hat{y} = g(W_0 + X^T W) \tag{1.1.2}$$

With:

$$W = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}, \text{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$$

As seen, in a conventional model of a perceptron every input of an layer is multiplied by a weighting. It can be seen, that the implementation of such a perceptron can be handled by given hardware architecture like GPU's or matrix multiplier, because of the possibility to proceed every input matrix and their associated weight matrix independently.

1.2 Activation function

The activation function of a perceptron has the purpose to determines the behavior of the perceptron in response to external stimuli. There are different kinds of activation functions which can be used for different purpuse. The most common used one is the so called signoid function:

1.3 Multilayer perceptron

1.4 Loss optimization

Chapter 2

Method

2.1 Design

2.1.1 Hardware

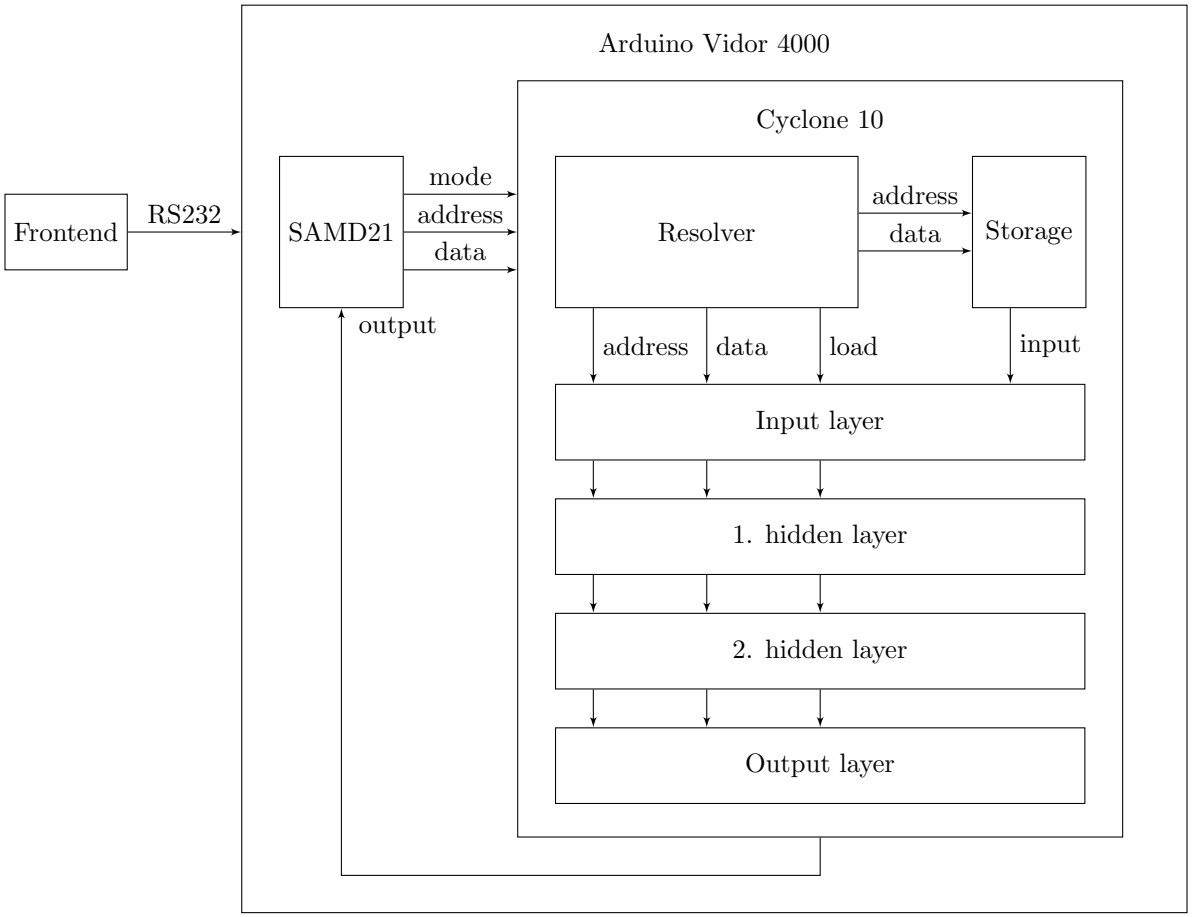

Figure 2.1: Top level hardware model

2.1.2 Software

2.2 Procedure

2.2.1 Iplementation

2.2.2 Validation

Chapter 3

Results

Chapter 4

Discussion