**Software Requirements Specification**

**for**

**Handwritten digit recognition using Convolutional Neural Network**

**Version 1.0 approved**

**Prepared by**

**Valencia Dias 7368**

**Nivea Dabre 7364**

**Fr. Conceicao Rodrigues college of Engineering**

**17.09.2017**

**Table of Contents**

**Table of Contents ii**

**Revision History ii**

**1. Introduction 1**

1.1 Purpose 1

1.2 Document Conventions 1

1.3 Intended Audience and Reading Suggestions 1

1.4 Product Scope 1

1.5 References 1

**2. Overall Description 2**

2.1 Product Perspective 2

2.2 Product Functions 2

2.3 User Classes and Characteristics 2

2.4 Operating Environment 2

2.5 Design and Implementation Constraints 2

2.6 User Documentation 2

2.7 Assumptions and Dependencies 3

**3. External Interface Requirements 3**

3.1 User Interfaces 3

3.2 Hardware Interfaces 3

3.3 Software Interfaces 3

3.4 Communications Interfaces 3

**4. System Features 4**

4.1 System Feature 1 4

4.2 System Feature 2 (and so on) 4

**5. Other Nonfunctional Requirements 4**

5.1 Performance Requirements 4

5.2 Safety Requirements 5

5.3 Security Requirements 5

5.4 Software Quality Attributes 5

5.5 Business Rules 5

**6. Other Requirements 5**

**Appendix A: Glossary 5**

**Appendix B: Analysis Models 5**

**Appendix C: To Be Determined List 6**

**Revision History**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Date** | **Reason For Changes** | **Version** |
|  |  |  |  |
|  |  |  |  |

* **Introduction**
* **Purpose**

Handwritten digit recognition has recently been of very interest among the researchers because of the evolution of various Machine Learning, Deep Learning and Computer Vision algorithms. In this report, I compare the results of some of the most widely used Machine Learning Algorithms like SVM, KNN & RFC andwith Deep Learning algorithm like multilayer CNN using Keras with Theano and Tensorflow. Using these, I was able to get the accuracy of 98.70% using CNN (Keras+Theano) as compared to 97.91% using SVM, 96.67% using KNN, 96.89% using RFC

* **Document Conventions**

Font face Times new roman

Font size 12

* **Intended Audience and Reading Suggestions**

Some of the research areas include signature verification, bank check processing, postal address interpretation from envelopes etc

* **Product Scope**

Handwritten digit recognition is the ability of a computer system to recognize the handwritten inputs like digits, characters etc. from a wide variety of sources like emails, papers, images, letters etc. This has been a topic of research for decades. Some of the research areas include signature verification, bank check processing, postal address interpretation from envelopes etc

* **References**

https://www.youtube.com/watch?v=7kpYpmw5FfE

https://github.com/anujdutt9/Handwritten-Digit-Recognition-using-Deep-Learning

http://ufldl.stanford.edu/wiki/index.php/Using\_the\_MNIST\_Dataset

https://keras.io/

http://deeplearning.net/software/theano/

* **Overall Description**
* **Product Perspective**

A Convolutional Neural Network (CNN) is a type of feedforward Artificial Neural Network in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. Convolutional Neural Networks consist of neurons that have learnable weights and biases. Each neuron receives some input, performs a dot product and optionally follows it with a non-linearity.The whole Convolutional Neural Network expresses a differentiable scorefunction that is further followed by a Softmax function. The data input into the Convolutional Neural Network is arranges in the form of its width, height and depth as shown in figure below.Should recognise the hand written digits using deep Learning concept

* **User Classes and Characteristics**

<Identify the various user classes that you anticipate will use this product. User classes may be differentiated based on frequency of use, subset of product functions used, technical expertise, security or privilege levels, educational level, or experience. Describe the pertinent characteristics of each user class. Certain requirements may pertain only to certain user classes. Distinguish the most important user classes for this product from those who are less important to satisfy.>

* **Operating Environment**

**Tools**

Keras

Theano/TensorFlow

Python

**Technology**

Deep Learning (Convolutional Neural Network)

* **External Interface Requirements**
* **Software Interfaces**

MNIST Dataset

The MNIST dataset, a subset of a larger set NIST, is a database of 70,000 handwritten digits, divided into 60,000 training examples and 10,000[2] testing samples. The images in the MNIST dataset are present in form of an array consisting of 28x28 values representing an image along with their labels. This is also the same in case of the testing images. The data is stored in four files[2][12]:

1. train-images-idx3-ubyte: training set images[2][12]

2. train-labels-idx1-ubyte:training set labels[2][12]

3. t10k-images-idx3-ubyte: test set images[2][12]

4. t10k-labels-idx1-ubyte: test set labels[2][12]

b.Keras Deep learning python library

c.Theano python library for efficient management of mathematical computations and multi-dimensional arrays

* **System Features**

The CNN for Handwritten Digit Recognition works in three main phases. 1. Phase1 - Input MNIST Data1: The first phase is to input the MNIST data. The MNIST data is provided as 784-d array of pixels. So firstly we convert it to grayscale images using 28x28 matrix of pixels.

2. Phase2 – Building Network Architecture: In the second phase, we define the models to be used to build a convolutional neural network. Here, we use the Sequential class from Keras to build the network. In this network, we have three layer sets of layers “CONV =>ReLU=> POOL”.

a) First Convolution Layer: In the first layer, we take 20 convolutional filters that go as a sliding window of size 5x5 over all the images of 28x28 matrix size and try to get the pixels with most intensity value.

b) ReLU Function: We know that convolution is a method that uses Back Propagation. So using the ReLU function as the activation function just after the convolutional layer reduces the likelihood of the vanishing gradient and avoids sparsity. This way we don‟t lose the important data and even get rid of redundant data like a lot of 0‟s in the pixels.

c) Pooling Layer: The pooling layer gets the data from the ReLU function and down-samples the steps in the 3D tensor. In short it pools all the pixels obtained from previous layers and again forms a new image matrix of a smaller size. These images are again input into the second set of layers i.e. “CONV =>ReLU=> POOL” and this process goes on till we get to a smallest set of pixels from which we can classify the digit.

3. Phase 3 –Fully Connected Layer: The fully connected layer is used to connect each of the previous layers to the next layers. This layer consists of 500 neurons. Finally, we apply a Softmax Classifierthat returns a list of probabilities for each of the 10 class labels. The class label with the largest probability is chosen as the final classification from the network and shown in the output

**Performance Requirements**

Accuracy should be high

Error rate for training and test images and labels should be low

* **Safety Requirements**

Data backup

Error rate should be less than 2% in order to efficiently recognise the digits

**Appendix A: Glossary**

CNN- Convolutional Neural Network

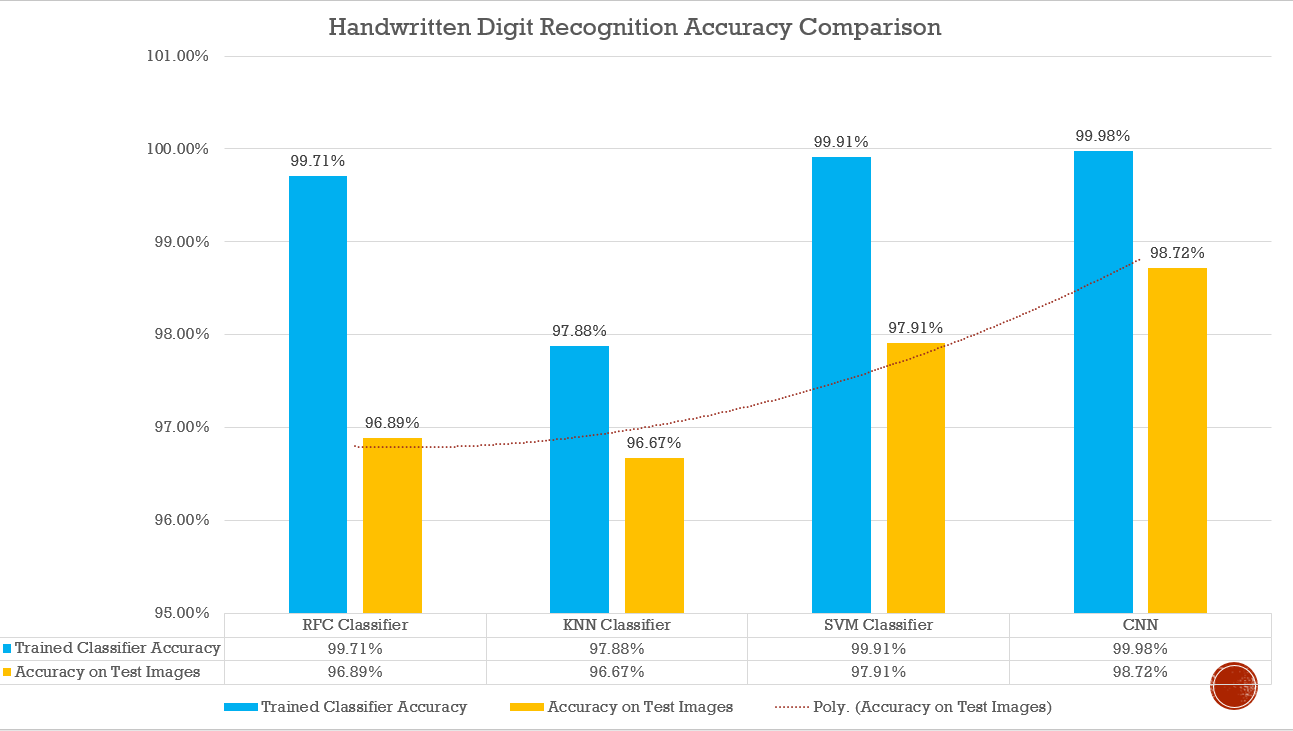
SVM- Supervised vector machine

KNN- K nearest Neighbour

RFC-Random Forest Classifier

MNIST- Modified National Institute of Standards and Technology (DataBase)

**Appendix B: Analysis Models**

****

