Text recognition and machine learning: For impaired robots and humans

Nadia Gifford[1], Rafiq Ahmad[1]

[1]Department of Computing Science, University of Alberta

Page 31 / WISEST Special Issue / Volume 2 / DOI: 10.29173/aar42

Abstract

As robots and machines become more reliable, developing tools that utilize their potential in manufacturing and beyond is an important step being addressed by many, including the LIMDA team at the University of Alberta. A common and effective means to improve artificial performance is through optical character recognition methods. Within the category of artificial intelligence under classification machine learning, research has focussed on the benefits of convolutional neural networks (CNN) and the improvement provided compared to its parent method, neural networks. Neural networks serious flaw comes from memorization and the lack of learning about what the images contain, while CNN's combat those issues. CNN's are designed to connect information received by the network and begins to closely mimic how humans experience learns. Using the programming language Python and machine learning libraries such as Tensorflow and Keras, different versions of CNN's were tested against datasets containing low-resolution images with handwritten characters. The first two CNN's were trained against the MNIST database against digits 0 through 9. The results from these tests illustrated the benefits of elements like max-pooling and the addition of convolutional layers. Taking that knowledge a final CNN was written to prove the accuracy of the algorithm against alphabet characters. After training and testings were complete the network showed an average 99.34% accuracy and 2.23% to the loss function. Time-consuming training epochs that don't wield higher or more impressive results could also be eliminated. These and similar CNN's have proven to yield positive results and in future research could be implemented into the laboratory to improve safety. Continuing to develop this work will lead to better translators for language, solid text to digital text transformation, and aides for the visual and speech impaired.

Key words:
Text Recognition, Optical Character Recognition (O.C.R), Machine Learning, Neural Networks (N.N), Convolutional Neural Networks (C.N.N)

Cite as: Gifford N., and Ahmad R. 2019. Text recognition and machine learning: for impaired robots and humans. Alberta Academic Review, Vol 2 (2) 31-32, WISEST Special Issue (not peer-reviewed), DOI 10.29173/aar42.

ISSN 2561-5327 (print)
ISSN 2561-5335 (online)

Text Recognition and Machine Learning: For Impaired Robots and Humans

Name	Degree/Position	Year	Supervisor
Nadia Gifford	WISEST Student Researcher	2019	Dr. Rafiq Ahmad

Supported By: Canada, MOTOROLA

Motivation

- For robots to collaborate with humans and react to unwanted situations in production environment.
- To prevent damage and injury with machines using optical text recognition (OCR) programming.

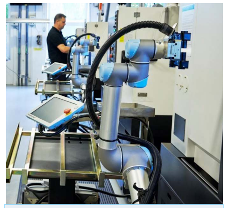

Figure 1: Robot working alongside humans in manufacturing setup (Universal Robots).

Figure 2: Assessing the VF 2TR* in the LIMDA lab (University of Alberta).

Objective

- To assess different machine learning algorithms for text recognition used in manufacturing.
- Implement and test a neural network against a dataset of handwritten examples of characters in the English alphabet.

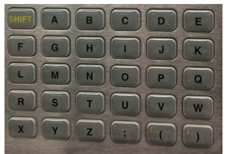

Figure 3: Example of the keys that are desired to be recognized.

*The VF 2TR is a subtractive manufacturing machine with Computer Numerical control (CNC).

Methodology

Figure 4: Step one of Method - Identify the type of machine learning for the task. (COGNUB).

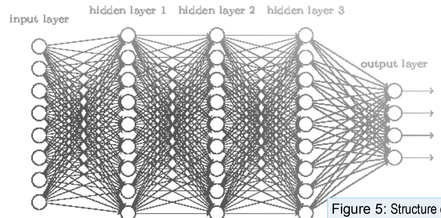

Figure 5: Structure of a deep NN (Nielson).

The form of image classification called Neural Networking (NN) imitates the process of the human brain and uses the information from the input layer to feed forward through layers to determine which answer on the output layer is correct.

Convolutional Neural Networks (CNN's) build on the concepts from neural networks by considering all the information in a specific region of an image and linking the data together to predict if patterns are being formed.

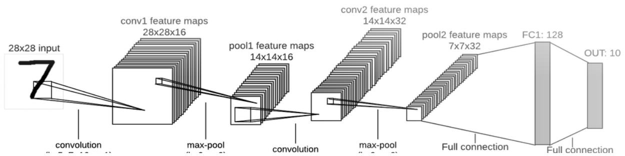

Figure 6: Structure of a CNN (EasyTensorFlow).

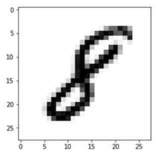

Figure 7: Example of the MNIST digits used to test prototype networks.

Figure 8: Example of the dataset used on final network (Patel).

Results

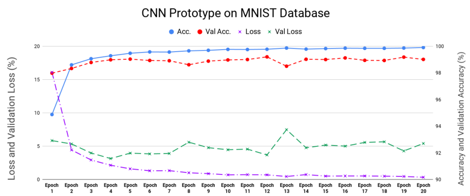

Figure 9: Line graph of the relationship between Loss function and Accuracy on prototype CNN (Sharma).

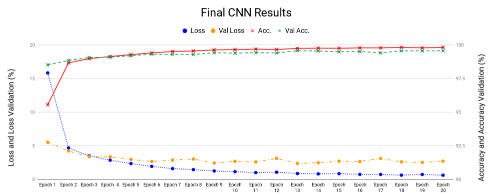

Figure 10: Line graph of the relationship between Loss function and Accuracy on final CNN (Patel).

- Using machine learning libraries like Tensorflow, Keras and Numpy, the slight variations in layer variables between the similar CNN's created a network with an average of 99.3% accuracy and an average 3% loss from the cost function.
- The learning curve of the networks both prove that improvement levels out between 10-15 epochs (training tests), less epochs needed for adequate results means less time wasted running tests.

Future Research

Future testing can determine if this CNN can be useful in the lab.

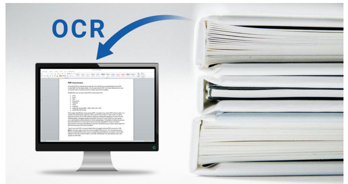

Figure 11: Document transfers to cyberspace (Investintech).

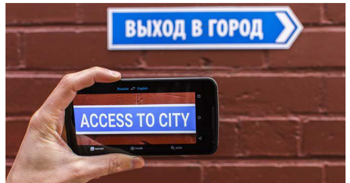

Figure 12: OCR's translating in real time (Verdugo).

Figure 13: Portable voice stick made with OCR to instantly translate text to audio for the visually impaired (Park).

References

- Allibhai, E. (2018, October 16). "Building a Convolutional Neural Network (CNN) in Keras". *Towards Data Science*. Retrieved from https://towardsdatascience.com/building-a-convolutional-neural-network-cnn-in-keras-329fbbadc5f5
- COGNUB. (2019). Retrieved from http://www.cognub.com/index.php/cognitive-platform/
- EasyTensorFlow. (2018). Retrieved from https://www.easy-tensorflow.com/tf-tutorials/convolutional-neural-nets-cnns
- Investintech. (2019). "OCR Software". *Investintech.com.* Retrieved from https://www.investintech.com/resources/articles/ocrprogram/
- Neilsen, M.A. (2015). "Neural Networks and Deep Learning". *Determination Press.* Retrieved from http://static.latexstudio.net/article/2018/0912/neuralnetworksanddeeplearning.pdf
- MNIST. Retrieved from http://yann.lecun.com/exdb/mnist/
- Park. S. (2008). "Voice Stick". Retrieved from https://www.yankodesign.com/2008/08/21/voice-stick/
- Patel. S. (2018). "A-Z Handwritten Alphabets in .csv format". Image data-set from *Kaggle*. Retrieved from https://www.kaggle.com/sachinpatel21/az-handwritten-alphabets-in-csv-format
- Sharma, A. (2017, December 5). "Convolutional Neural Networks in Python with Keras". *DataCamp*. Retrieved from https://www.datacamp.com/community/tutorials/convolutional-neural-networks-python
- Universal Robots. (2017). Retrieved from https://www.youtube.com/watch?v=Psz4sX4NgO0
- Verdugo. D.G. (2016, May 19). "OCR on Android". Retrieved from https://solidgeargroup.com/ocr-on-android

Laboratory of Intelligent Manufacturing, Design and Automation

Department of Mechanical Engineering, University of Alberta

Contact:
rafiq.ahmad@ualberta.ca
https://sites.ualberta.ca/~rafiq1/