



AaBbCcDdEeFfGg

## HANDWRITTEN TEXT RECOGNITION

Aditya Gourav 2K11/CO/007

adi.dtu@gmail.com

Aman Gupta 2K11/CO/017

amangupta687@gmail.com

Ashish Yadav 2K11/CO/029

ashuyadav93@gmail.com

Kartik Singal 2K11/CO/052

ka4tik@gmail.com

*Handwriting recognition is the ability of a computer to receive and interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices.*

## ABSTRACT

Offline handwriting recognition—the transcription of images of handwritten text—is an interesting task, in that it combines computer vision with sequence learning. In most systems the two elements are handled separately, with sophisticated pre-processing techniques used to extract the image features and sequential models such as HMMs used to provide the transcriptions. By combining two recent innovations in neural networks—multidimensional recurrent neural networks and connectionist temporal classification—this project introduces a globally trained offline handwriting recogniser that takes raw pixel data as input.

## INTRODUCTION

The image of the written text may be sensed "off line" from a piece of paper by optical scanning (optical character recognition) or intelligent word recognition. Alternatively, the movements of the pen tip may be sensed "on line", for example by a pen-based computer screen surface.

## METHODS

- The three components of our recognition system are:  
Multidimensional recurrent neural networks and  
multidimensional LSTM in particular.
- The connectionist temporal classification output layer.
- The Network hierarchical structure.

### MULTIDIMENSIONAL RECURRENT NEURAL NETWORKS

The basic idea of multidimensional recurrent neural networks (MDRNNs) is to replace the single recurrent connection found in standard recurrent networks with as many connections as there are spatial-temporal dimensions in the data. These connections allow the network to create a flexible internal representation of surrounding context, which is robust to localised distortions.

### MULTIDIMENSIONAL LSTM

Long Short-Term Memory (LSTM) is an RNN architecture designed for data with long range interdependencies. An LSTM layer consists of recurrently connected ‘memory cells’. The activation of each cell is controlled by three multiplicative gate units: the input gate, forget gate and output gate. The gates allows

the cells to store and retrieve information over long periods of time, giving them access to long range context.

### **CONNECTIONIST TEMPORAL CLASSIFICATION**

Connectionist temporal classification (CTC) is an output layer designed for sequence labelling with RNNs. Unlike other neural network output layers it does not require pre-segmented training data, or post processing to transform its outputs into transcriptions. Instead, it trains the network to directly estimate the conditional probabilities of the possible labelling given the input sequences.

### **NETWORK HIERARCHY**

Many computer vision systems use a hierarchical approach to feature extraction, with the features at each level used as input to the next level. This allows complex visual properties to be built up in stages. Typically, such systems use subsampling, with the feature resolution decreased at each stage. They also generally have more features at the higher levels. The basic idea is to progress from a small number of simple local features to a large number of complex global features.

## APPLICATIONS

- Signature Verification
- Postal-Address Interpretation
- Bank-Check Processing
- Writer Recognition

## REFERENCES

- [1] A. Graves. *Supervised Sequence Labelling with Recurrent Neural Networks*. PhD thesis.
- [2] A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber. Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. In *Proceedings of the International Conference on Machine Learning, ICML 2006*, Pittsburgh, USA, 2006.
- [3] A. Graves, S. Fernández, M. Liwicki, H. Bunke, and J. Schmidhuber. Unconstrained online handwriting recognition with recurrent neural networks. In J. Platt, D. Koller, Y. Singer, and S. Roweis, editors, *Advances in Neural Information Processing Systems 20*. MIT Press, Cambridge, MA, 2008.
- [4] A. Graves, S. Fernández, and J. Schmidhuber. Multidimensional recurrent neural networks. In *Proceedings of the 2007 International Conference on Artificial Neural Networks*, Porto, Portugal, September 2007.