Handwritten Digit String Recognition using Deep Autoencoder based Segmentation and ResNet based Recognition Approach

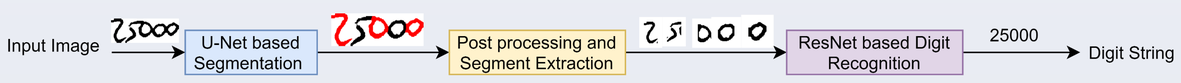
[1]Anuran Chakraborty, [1]Rajonya De, [1]Samir Malakar, [2]Friedhelm Schwenker, and [1]Ram Sarkar

[1]Jadavpur University, [2]Institute of Neural Information Processing, University of Ulm, Ulm, Germany

Introduction

- Segmentation based methods [1] isolate a digit string image into pieces of single digits using various algorithms and then perform digit recognition on these pieces. However, segmentation algorithms may not perform well in case of complex and erratic handwriting.
- Non-segmentation based deep learning methods for digit recognition make use of Recurrent Neural Network (RNN) [2] however they are expensive to train.
- We propose a deep autoencoder based segmentation approach to recognize digit strings.

Proposed Method



The proposed method consists of **four** steps:

- **Segmentation**

1. A U-Net is trained which takes a binarized image as input and gives a segmented image in which the alternate digits of the string have alternate colors (red and black).

- **Post-processing**

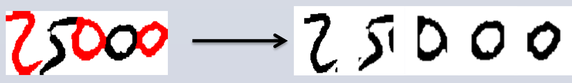
1. The segmentation step may segment a single digit into multiple segments and to handle such cases some rule based post processing methods are applied.
   i. Extract all connected components from the image, discarding all components having size less than α (=10) pixels.
   ii. Categorize the components into small and large based on a threshold.
   iii. Join each small component to form a large component based on the rules illustrated.
   - If a smaller component is connected to only one larger component, then pixels of the smaller component is substituted with the pixel color of a larger component.

   - If a smaller component is in connection with two larger components, then the color of the left-most connected component is substituted with the color of this smaller component.

- **Extraction of segmented components**

1. Take a vertical line passing through the middle of the component and traverse along it from top to bottom, including any component in our path, into the current window.

2. Assuming that all are part of the same digit since it is highly unlikely that two different digits will be written on top/below of one another. The final extraction is done by using a depth first search algorithm to find connected components.



- **Recognition**

1. A CNN based ResNet model is used to perform the recognition of digits which are extracted by segmentation.
2. The idea is to use so-called "identity shortcut connections" that skip one or more layers which handles the problem of vanishing gradients.

Results

- For experiment, the Computer Vision Lab Handwritten Digit String (CVL HDS) database has been used which was opened through the ICDAR 2013 Competition on Handwritten Digit Recognition (HDRC 2013) [3].
- The metrics used are Normalised Levenshtein distance (NLD) and Average NLD (ANLD).
- The table shows a performance comparison of the proposed method with some state-of-the-art methods.

Method	Precision Rate (Top-1)	Average NLD (Top-1)
Beijing [4]	0.8529	0.03
Shanghai [4]	0.4893	0.16
CRNN [5]	0.2601	-
Densenet+CTC [6]	0.4269	-
Proposed Method	0.6150	0.10

Discussion and conclusion

- Our method performs better than most of the state-of-the-art deep learning based methods.
- One advantage of our method is it performs well even with a small amount of data compared to most other deep learning based methods.
- Handwritten digit strings, which are commonly found in bank cheques, postal documents, filled-in forms, etc., are mostly written in cursive.
- For segmenting the cursive digit strings, we have applied a deep autoencoder (U-Net), and then recognized the isolated digits using ResNet.
- In the future, we plan to apply this approach when the entire digit string is written at one go as our segmentation approach fails to isolate such digit srting.

References

[1] R. Saabni,"Recognizing handwritten single digits and digit strings using deep architecture of neural networks," in 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR).IEEE, Sep.2016.

[2] H. Zhan, Q. Wang, and Y. Lu, "Handwritten digit string recognition by combination of residual network and RNN-CTC," in Neural Information Processing. Springer International Publishing, 2017, pp. 583–591.

[3] M. Diem, S. Fiel, A. Garz, M. Keglevic, F. Kleber, and R. Sablatnig,"ICDAR 2013 competition on handwritten digit recognition (HDRC2013)," in 2013 12th International Conference on Document Analysis and Recognition. IEEE, Aug. 2013.

[4] M. Diem, S. Fiel, F. Kleber, R. Sablatnig, J. M. Saavedra, D. Contreras, J. M. Barrios, and L. S. Oliveira, "ICFHR 2014 competition on handwritten digit string recognition in challenging datasets (HDSRC2014)," in 2014 14th International Conference on Frontiers in Handwriting Recognition, IEEE, Sep. 2014.

[5] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 11, pp. 2298–2304, Nov. 2017.

[6] H. Zhan, S. Lyu, and Y. Lu, "Handwritten digit string recognition using convolutional neural network," in 2018 24th International Conference on Pattern Recognition (ICPR), IEEE, Aug. 2018.