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

Anuran Chakraborty, Rajonya De, Samir Malakar, Friedhelm Schwenker, and Ram Sarkar

Agenda

- Introduction
- Proposed Method
- Results
- Conclusion
- References

Introduction

- Handwritten Digit String Recognition (HDSR) is an active research area in the domain of document image processing.

- Segmentation based methods [1] isolate a digit string image into pieces of single digits and then perform digit recognition on these pieces.

- 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 these methods are expensive to train.

Proposed Method

The proposed method consists of four steps:

- ▶ Segmentation
- ▶ Post-processing
- ▶ Extraction of segmented components
- ▶ Recognition

Proposed Method: Segmentation

Proposed Method: Segmentation

▶ In this step a U-Net is trained for which

  ➢ **Input** : Binarized image of the string

  ➢ **Output:** Image of segmented string

▶ The segmented image is an image in which the alternate digits of the string have alternate colors (red and black).

Proposed Method: Post-processing

Need for Post-processing

- The segmentation step may segment a single digit into multiple segments.
- To handle this some rule-based post processing methods are applied.

Post-Processing Steps

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.

Post-Processing Steps

▶ 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*.

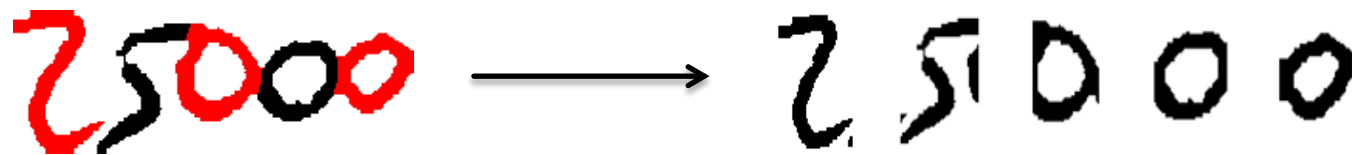# Proposed Method: Extraction of Segmented Components

Proposed Method: Extraction of Segmented Components

- 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.

- Assume 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.

Proposed Method: Recognition

Proposed Method: Recognition

- A CNN based ResNet model is used to perform the recognition of digits which are extracted by segmentation.

- The core idea of ResNet is the introduction of so-called "identity shortcut connections" that skip one or more layers. This handles the problem of vanishing gradients while training deep networks.

Overview of the Proposed Method
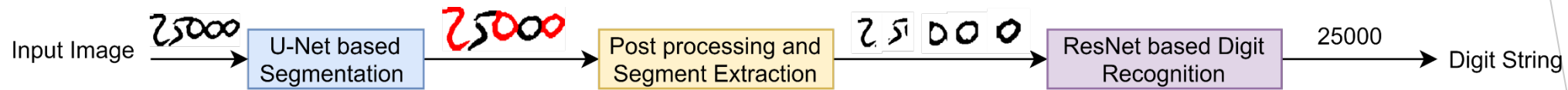
Results

- **Dataset Used** : *Computer Vision Lab Handwritten Digit String (CVL HDS) opened through ICDAR 2013 Competition on Handwritten Digit Recognition (HDRC 2013)* [3].

- **Performance Metrics Used** : *Average Normalised Levenshtein distance (ANLD)* and *Precision Rate*

Results: Comparison

▶ The table shows a performance comparison of the proposed method with other state-of-the-art methods on this dataset.

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

▶ The proposed method performs better than most state-of-the-art deep learning based methods.

Conclusion

- In the era of deep leaning, handwritten isolated digit recognition is considered as an almost saturated research problem.

- But 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 we have recognized the isolated digits using a standard deep learning model (ResNet).

- In the future, we would put forward our effort on the segmentation approach as it is seen that when the entire digit string is written at one go, our segmentation approach fails to isolate some digits.

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.

Thank You