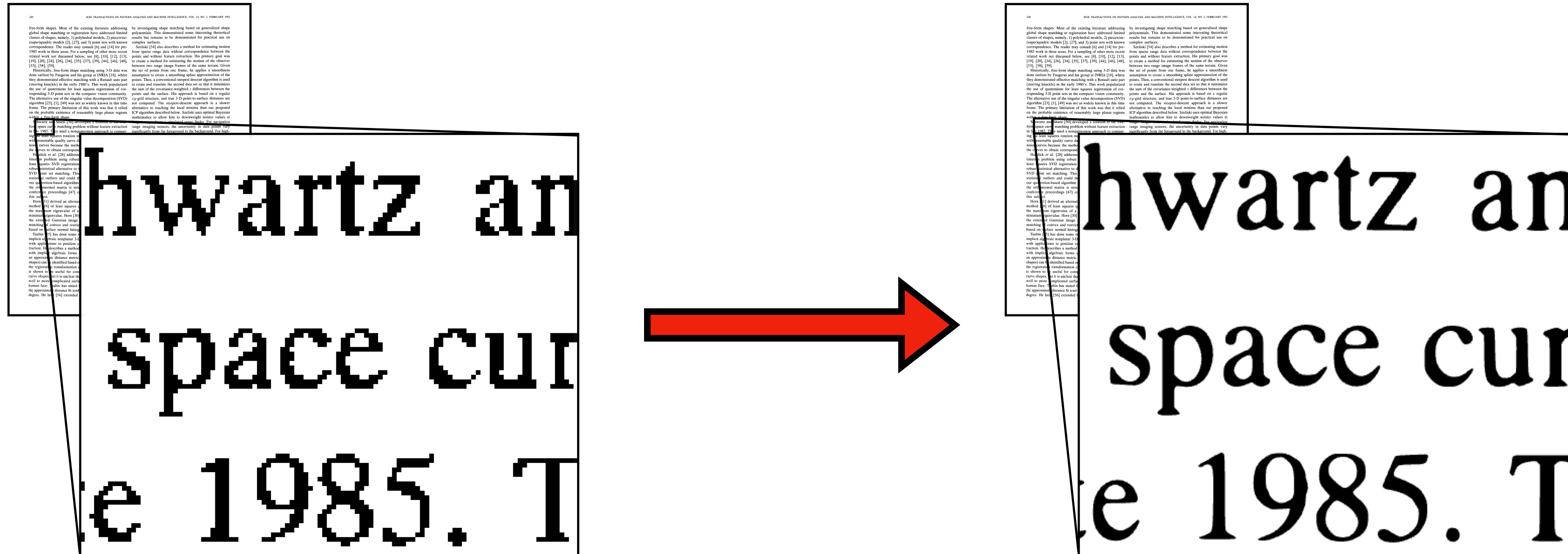


Task: Vectorization of scanned document



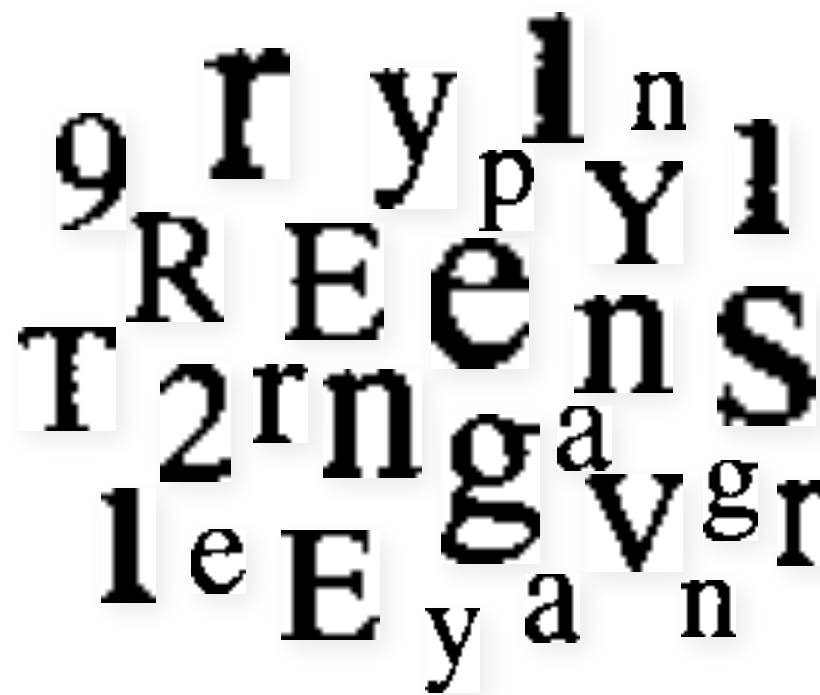
- Improved reading experience and photocopying quality; reduced file size
- **Difficulty:** Low scanning image resolution and severe loss of character details
- **Observation:** The same character will appear multiple times in the document

Algorithm Overview

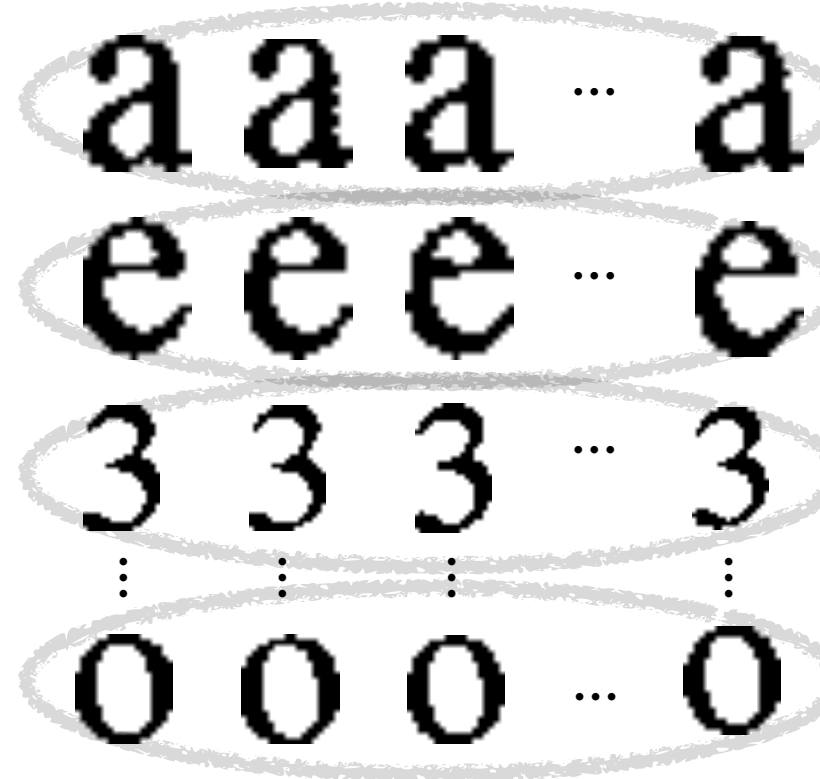
Free-form shapes. Most of the existing literature addressing global shape matching or registration has addressed broad classes of shapes, namely, 2D planar shapes, 2D piecewise-linear shapes [2], [7], and 3D point sets with known correspondences. The reader may consult [1] for a survey of the state-of-the-art in this area. For a sampling of other recent related work, we directed the reader to [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100].

扫描文档图像

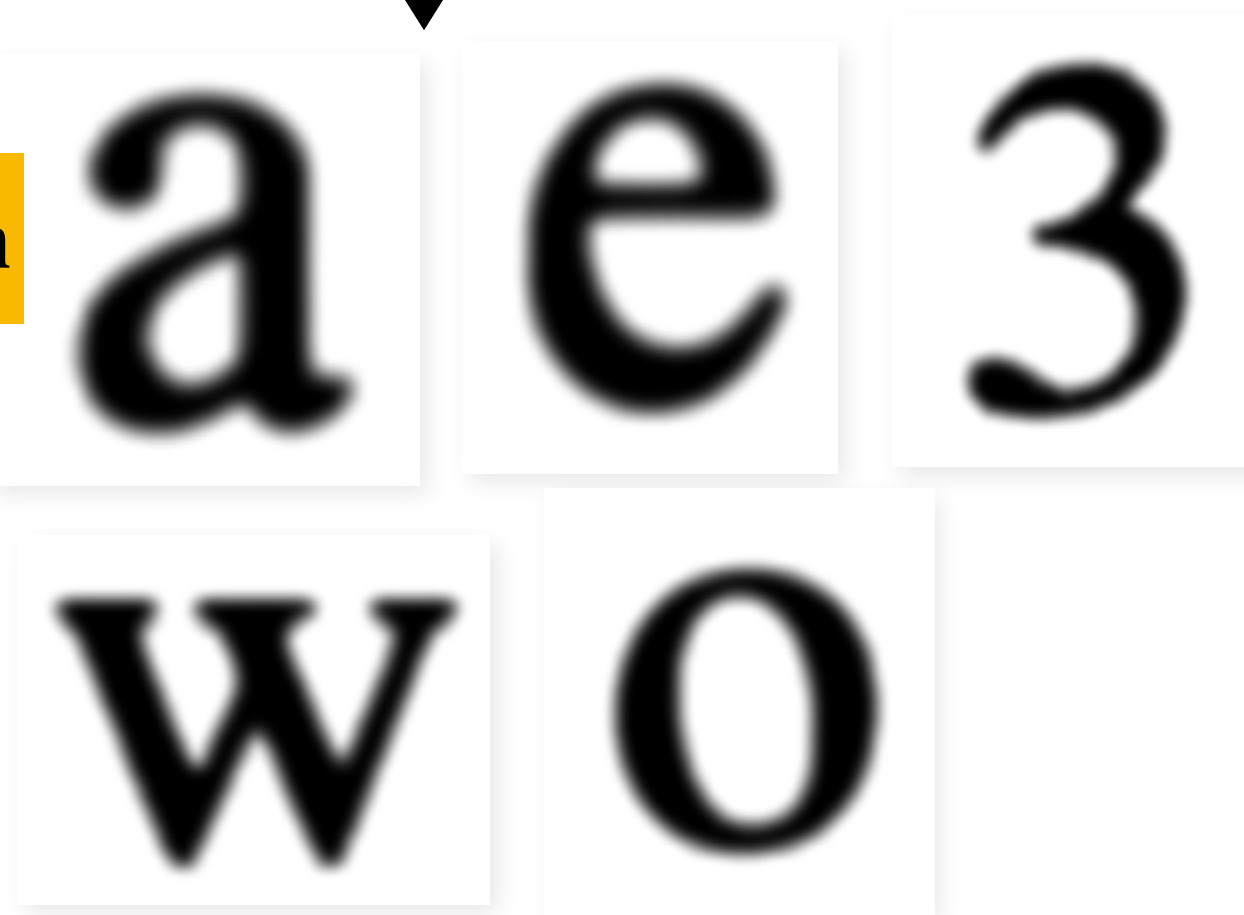
1. Extract



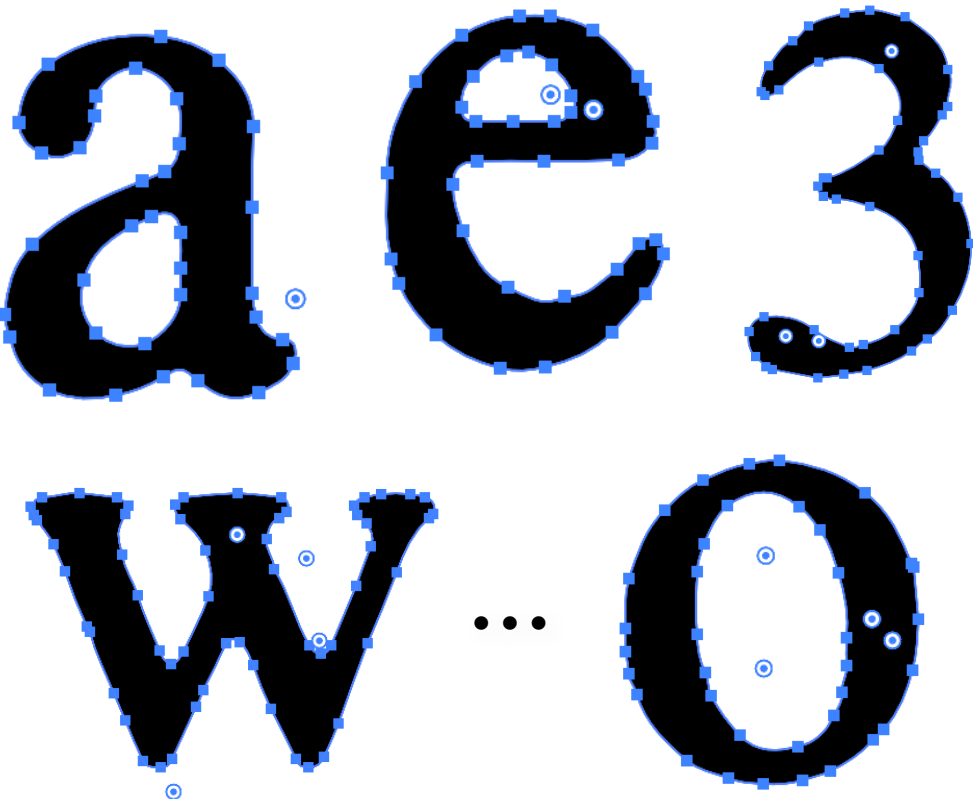
2. Cluster



3. Merge



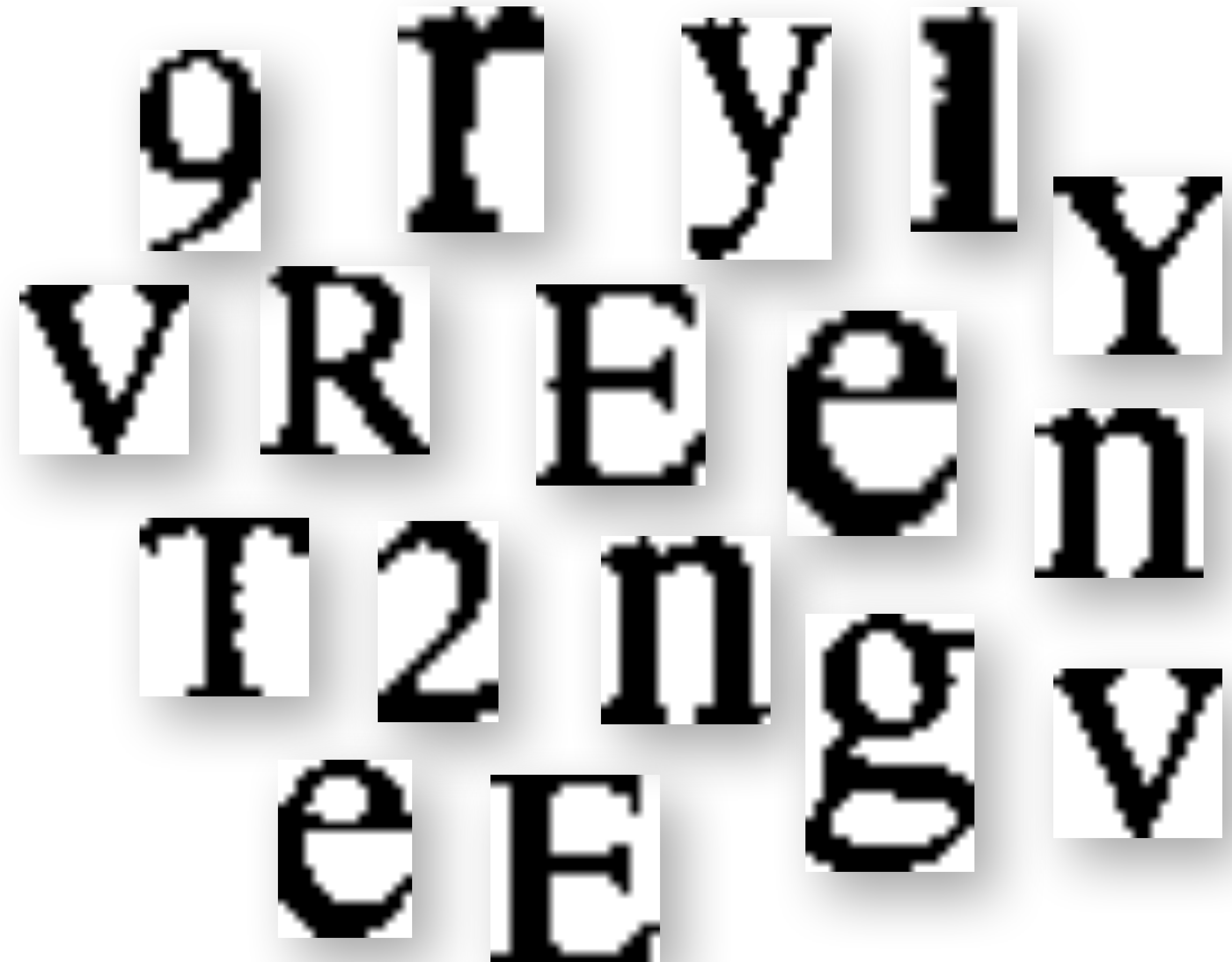
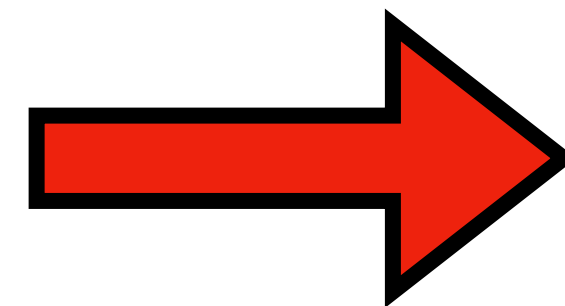
4. Vectorization



5. Write

矢量文档

1. Character extraction based on connected domain



Scan document images

2. Clustering based on character similarity

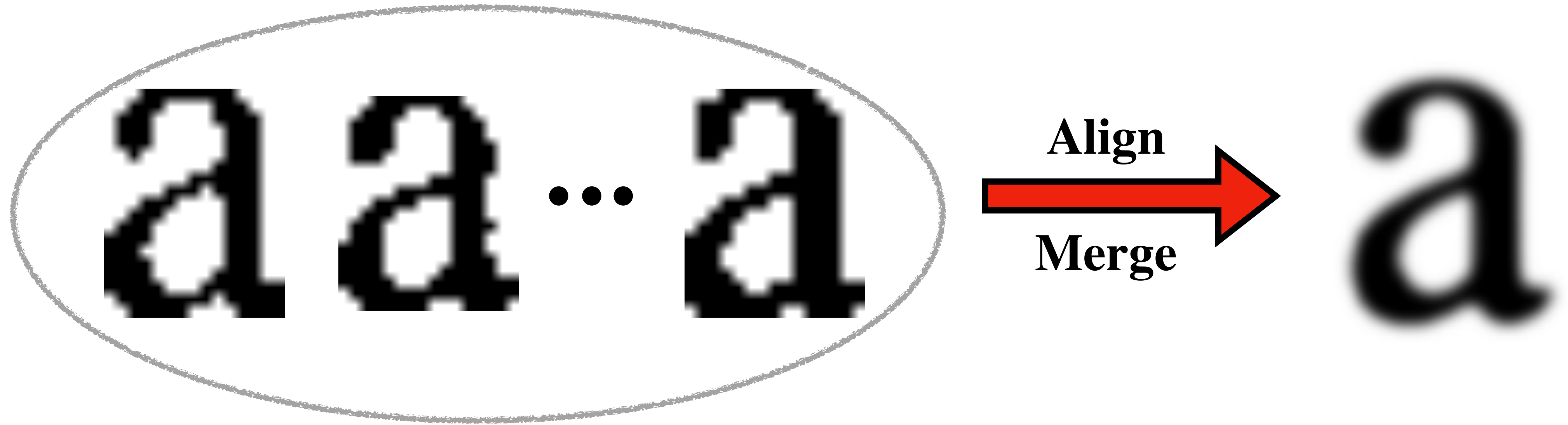
a a ... a

e e ... e

t t ... t

• • • • •

3. Merge-based super-resolution



4. Sketching based on Potrace

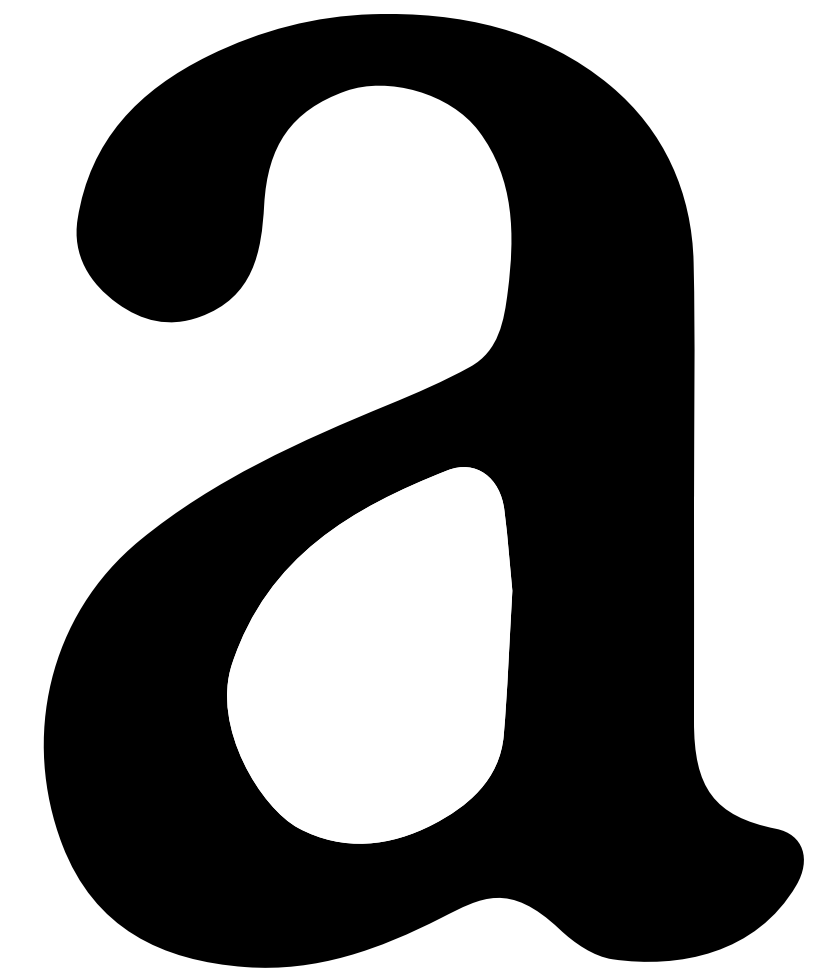


Merged Character

Binarize

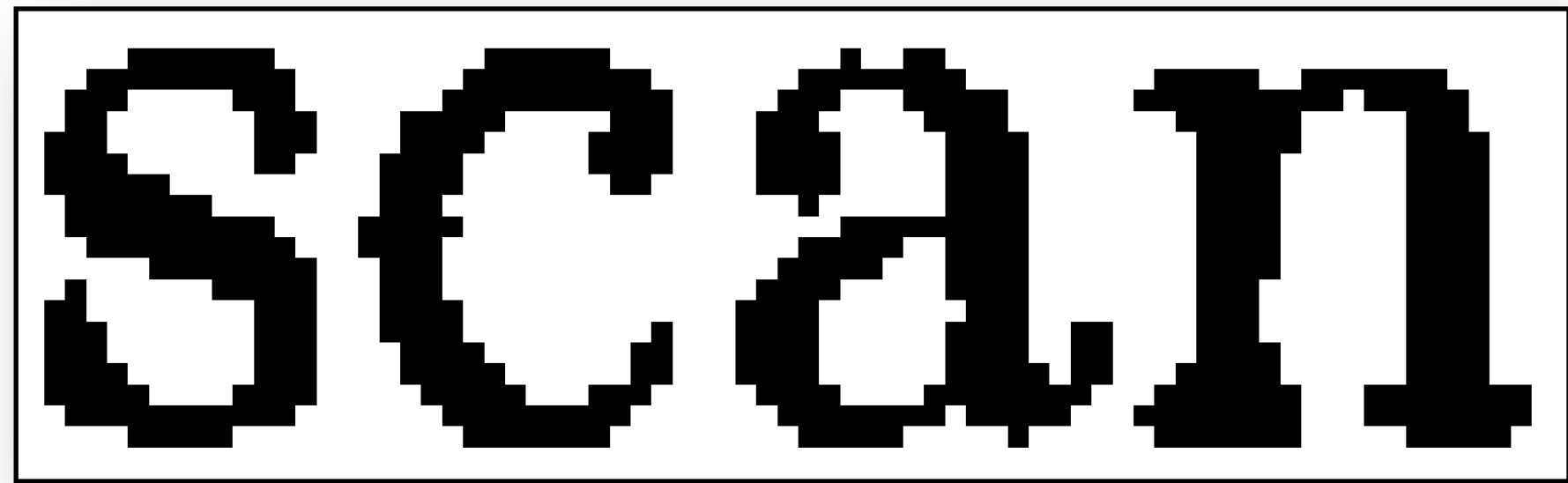


Potrace

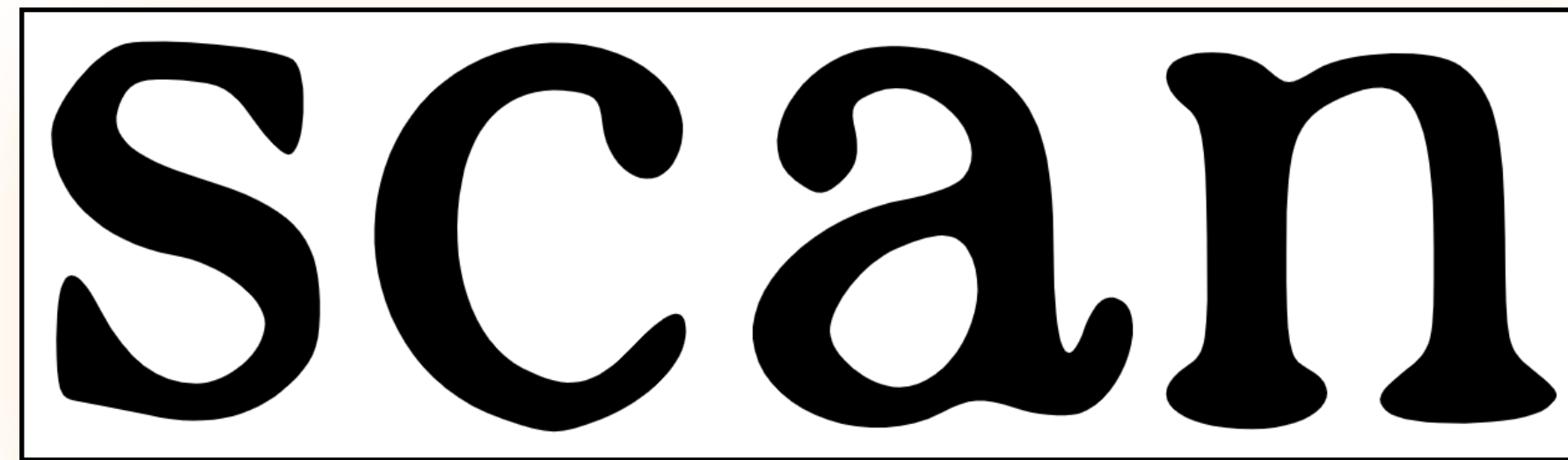


Vector Character

5. Result & Comparison



Input



This method



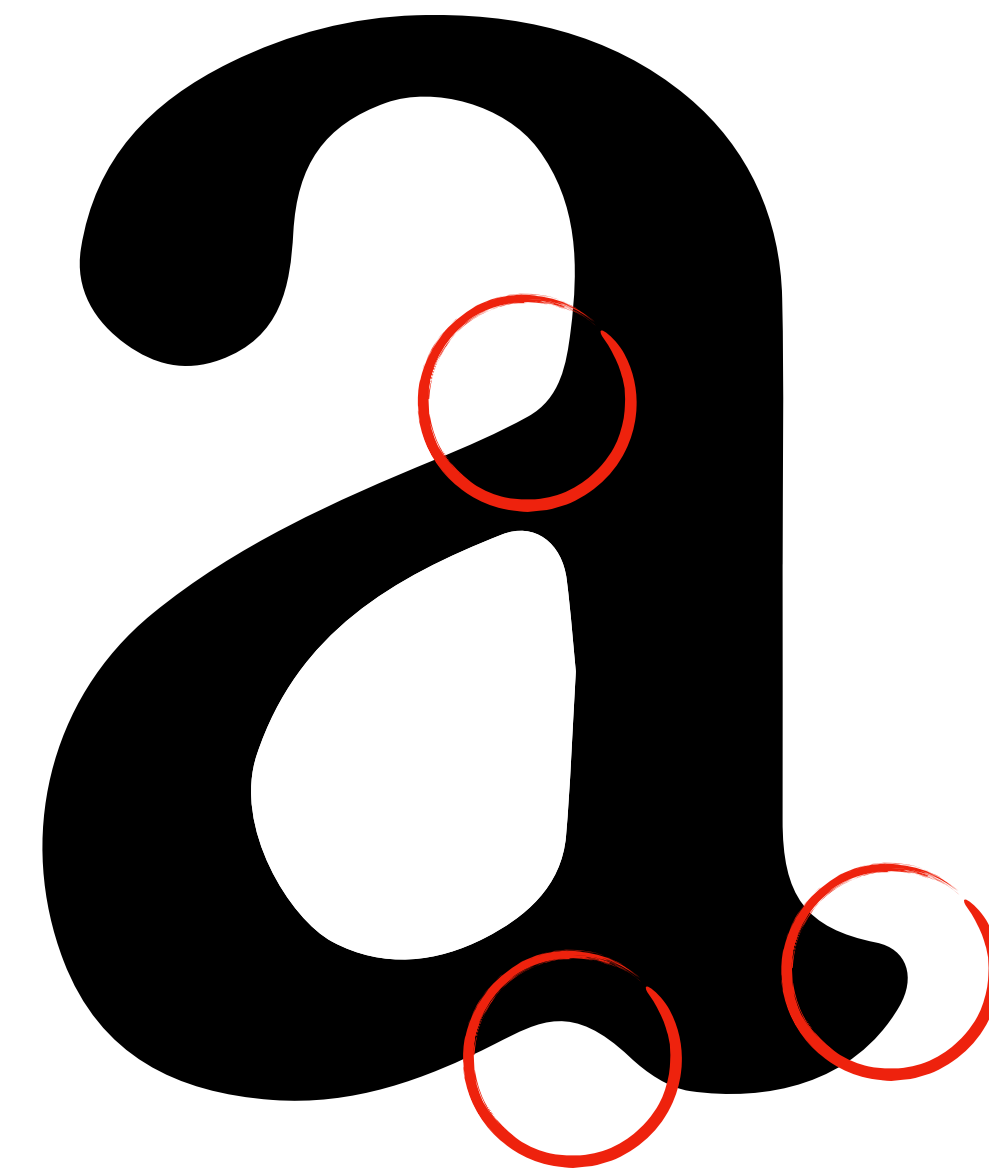
Adobe Illustrator

Remained Issue: Loss of Serif Details



Merged Image

**Binarization
Vectorization**



Over-Smoothed Serifs

Improved Based on Machine Learning



Merged Image

X_i

~~Binarization
Vectorization~~



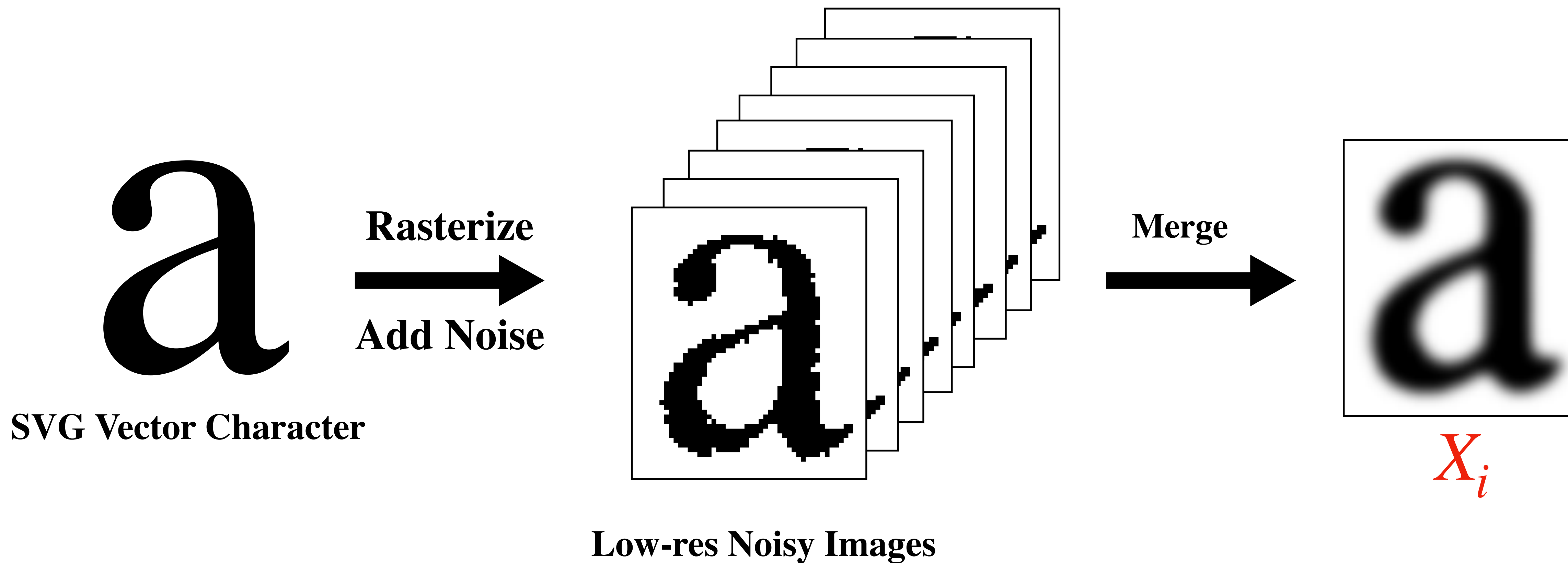
Deep Learning
Model



Target Vector

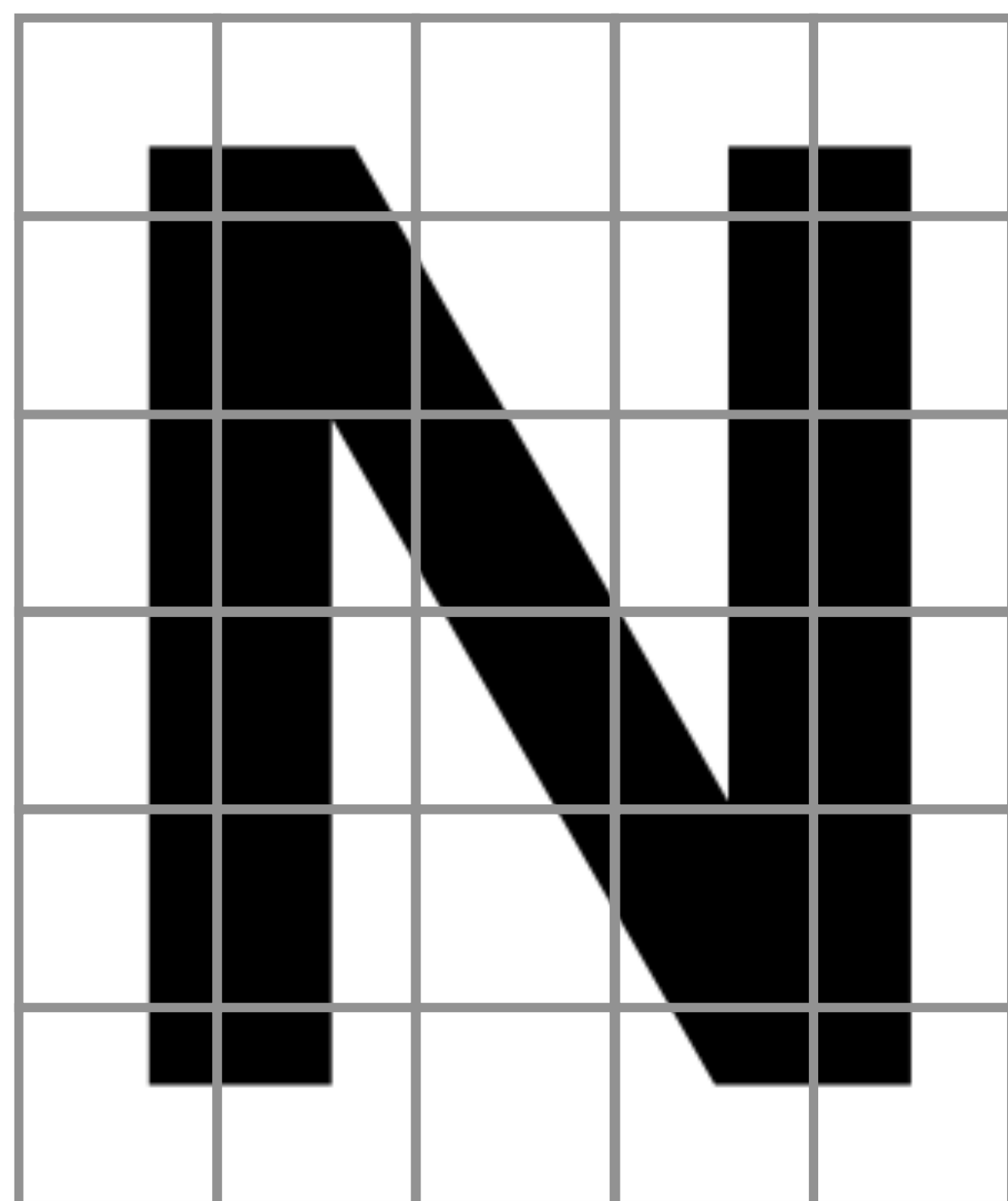
Y_i

Construct Training-Set: Merged Image X_i

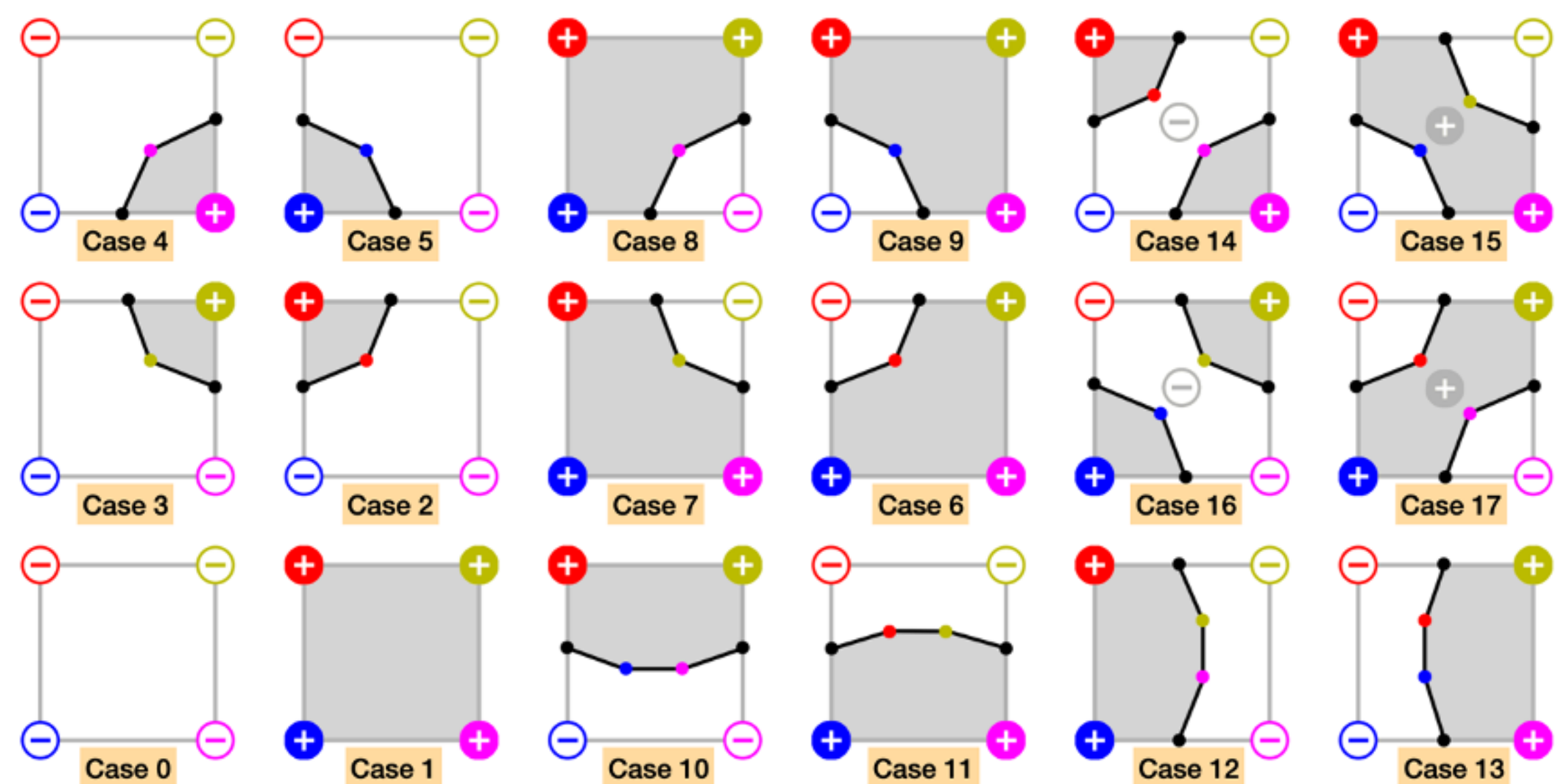


Construct Training-Set: Target Vector Y_i

Y_i is represented in a format resembling the Marching-Square Algorithm

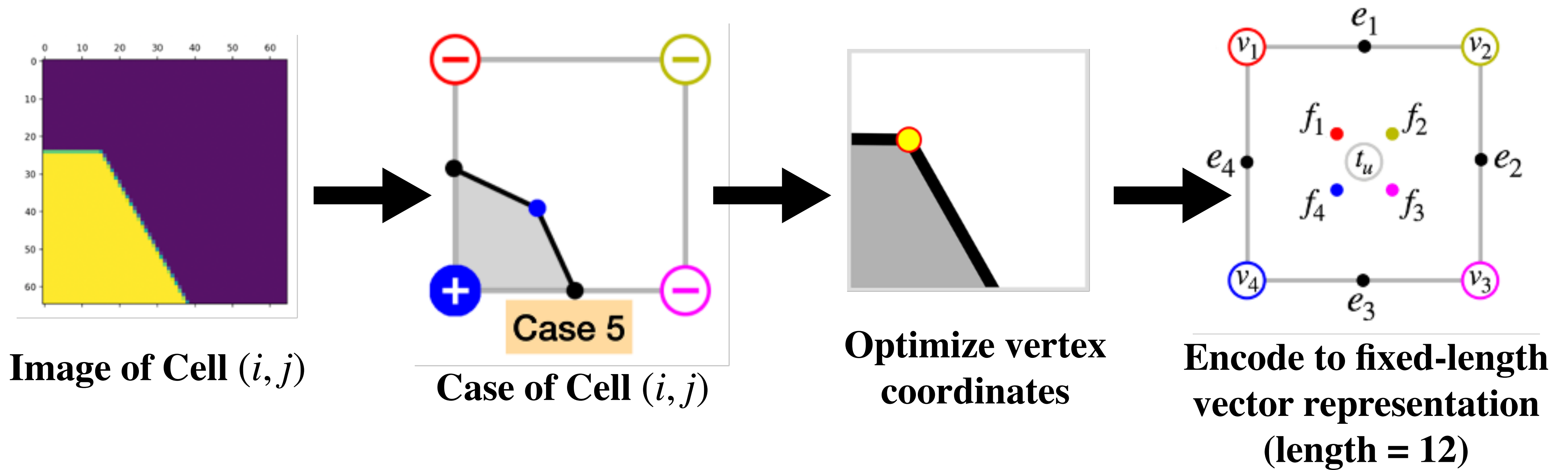


SVG Vector Character
Divide into $H \times W$ Grid



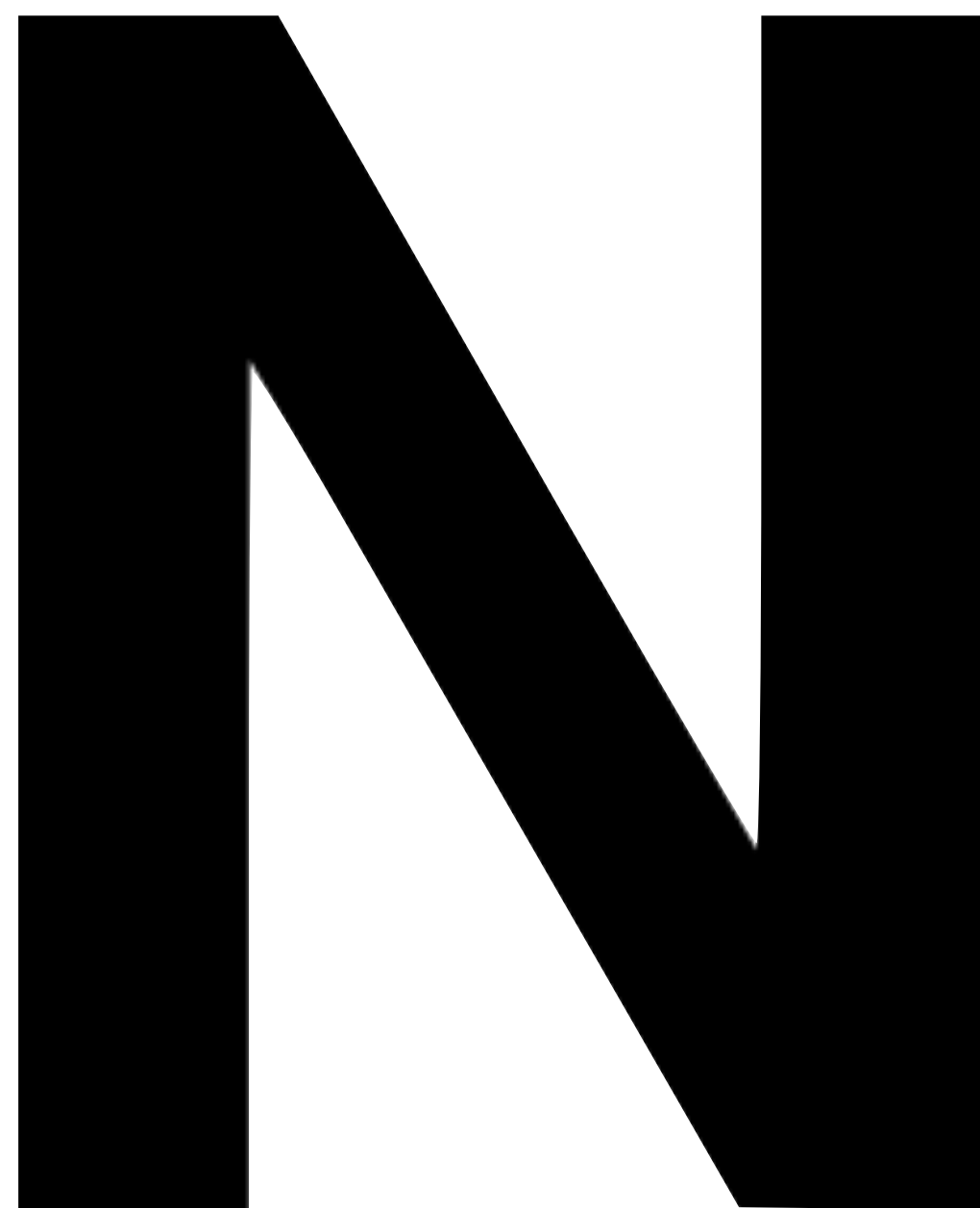
Each grid cell corresponds to a Case above

Construct Training-Set: Target Vector Y_i

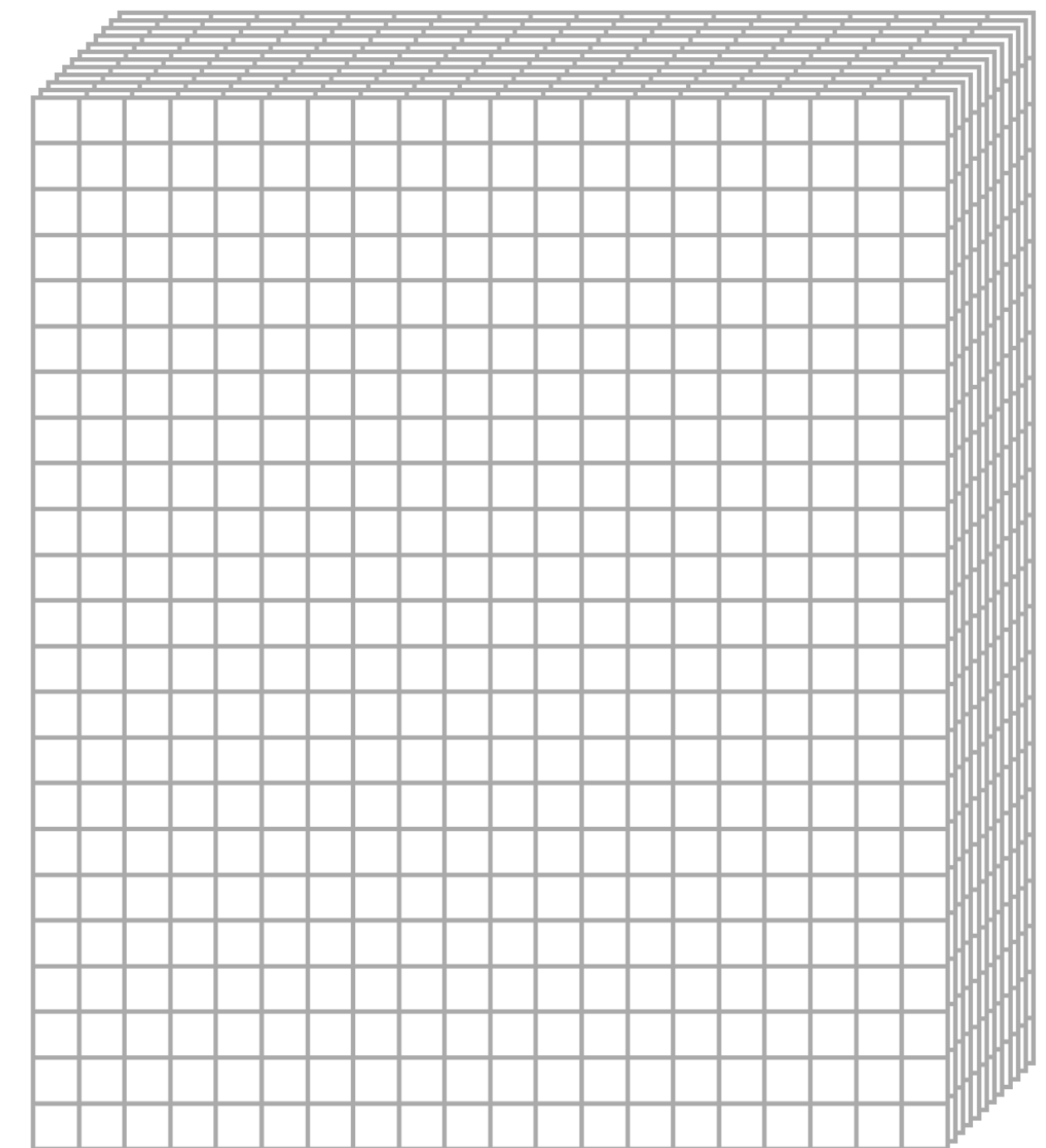
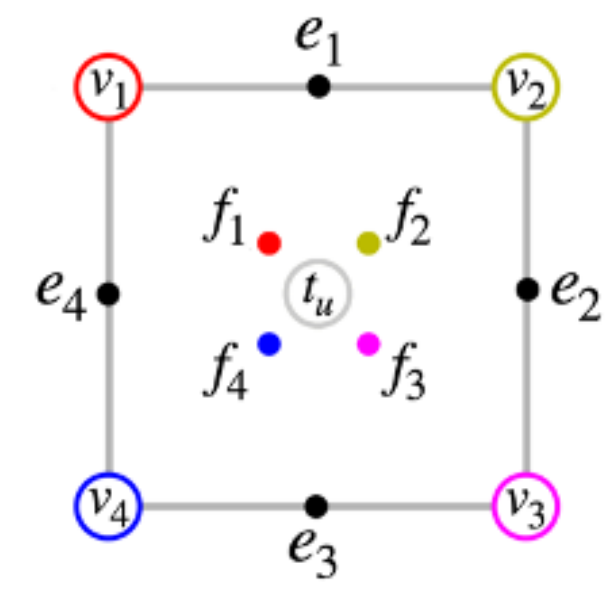
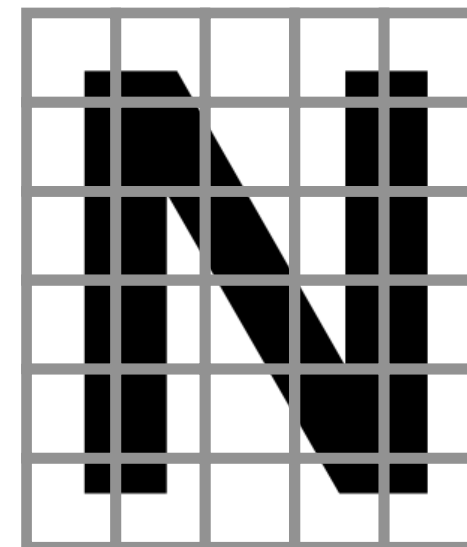


$H \times W$ cells in total, so the entire grid is represented as a $H \times W \times 12$ tensor

Construct Training-Set: Target Vector Y_i

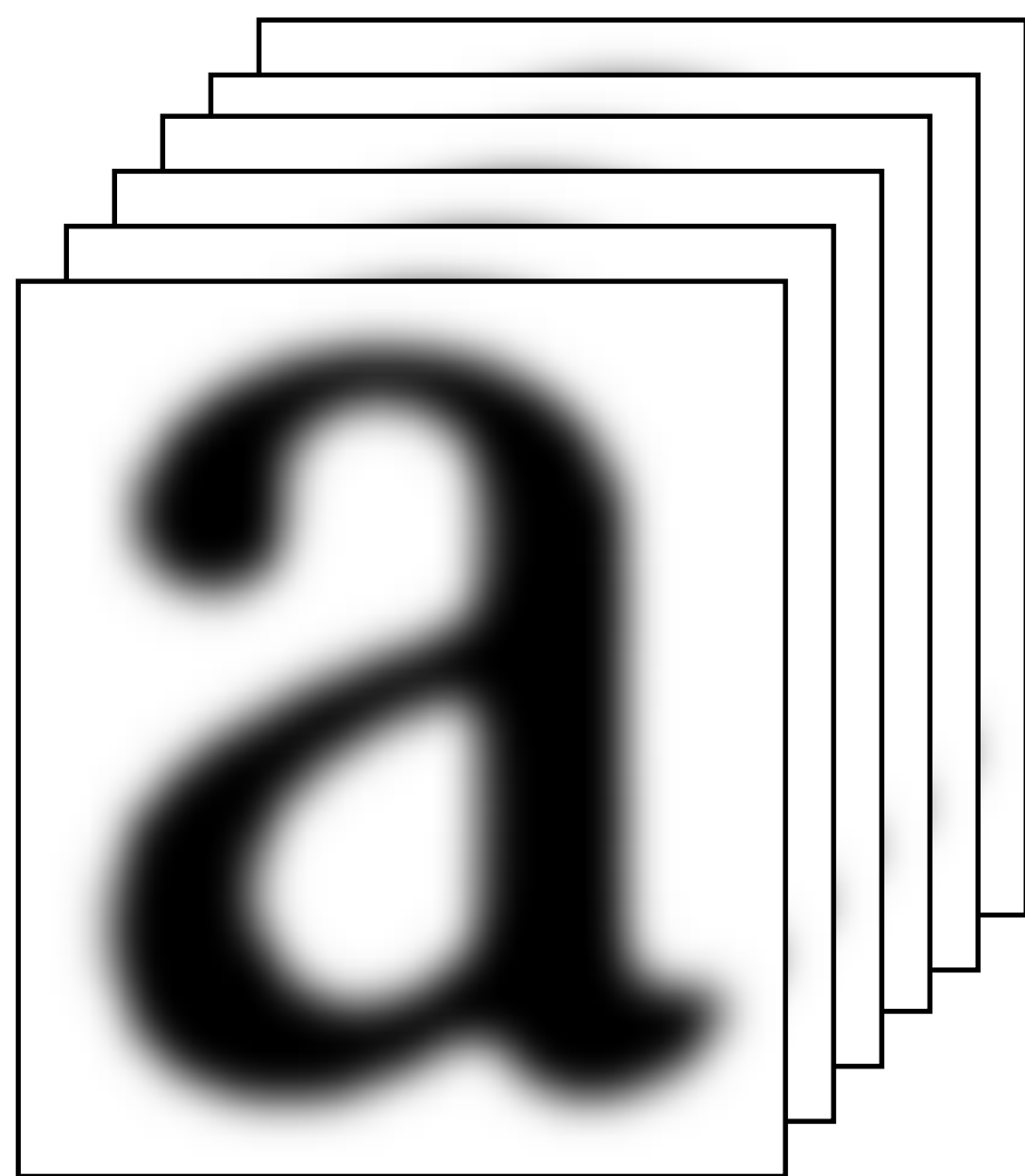


SVG Vector Character



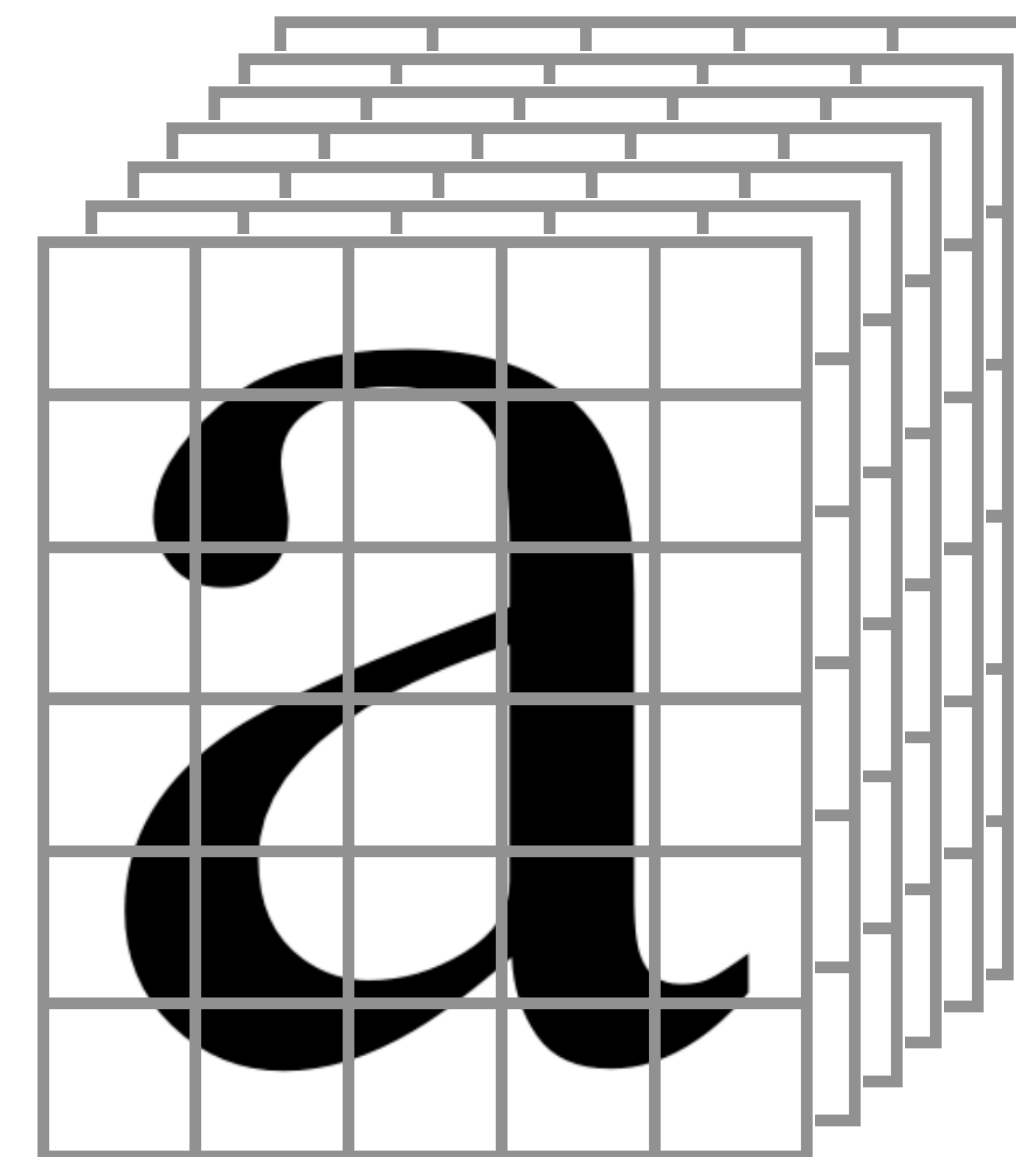
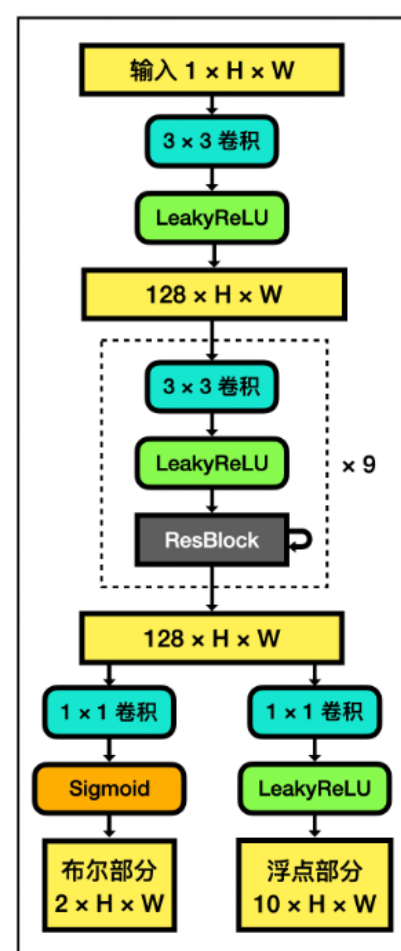
Y_i
(H, W, 12) Tensor

Learn the mapping of $X_i \rightarrow Y_i$ with CNN



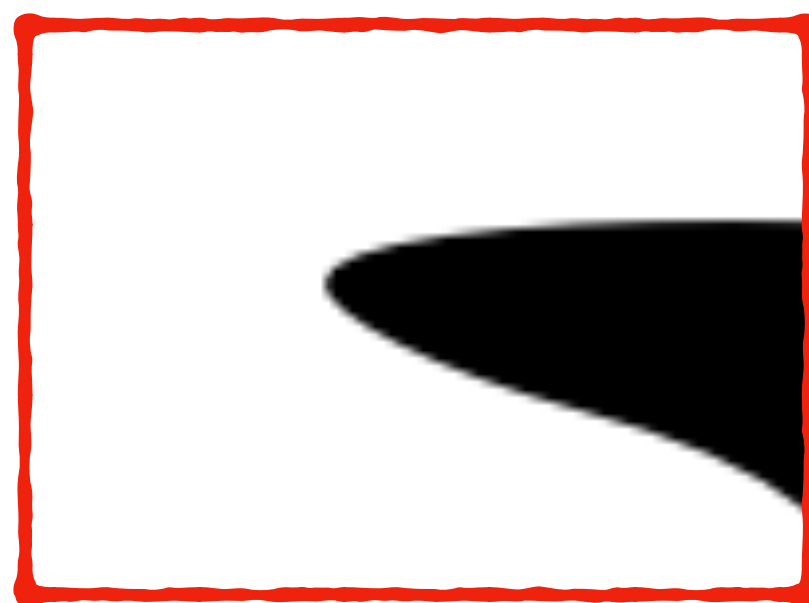
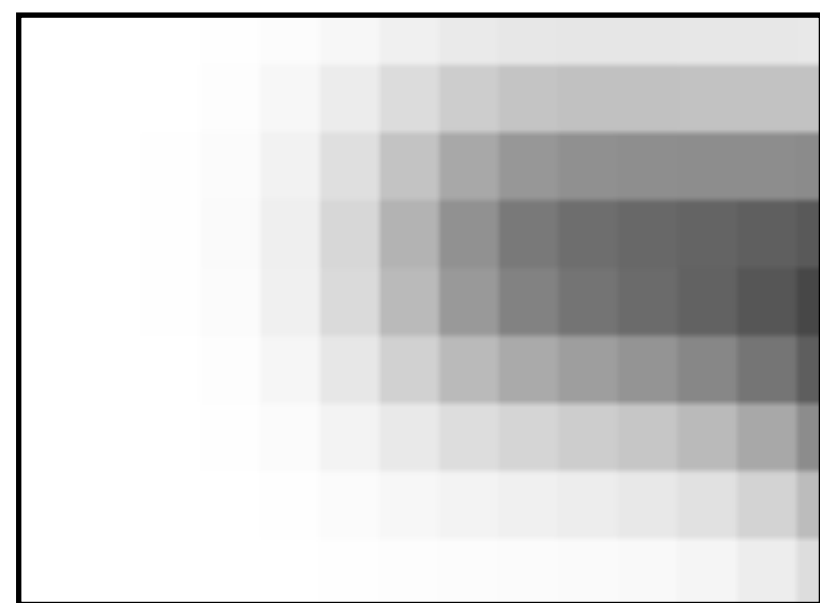
X_1, \dots, X_N
($H \times W \times 1$)

CNN
→

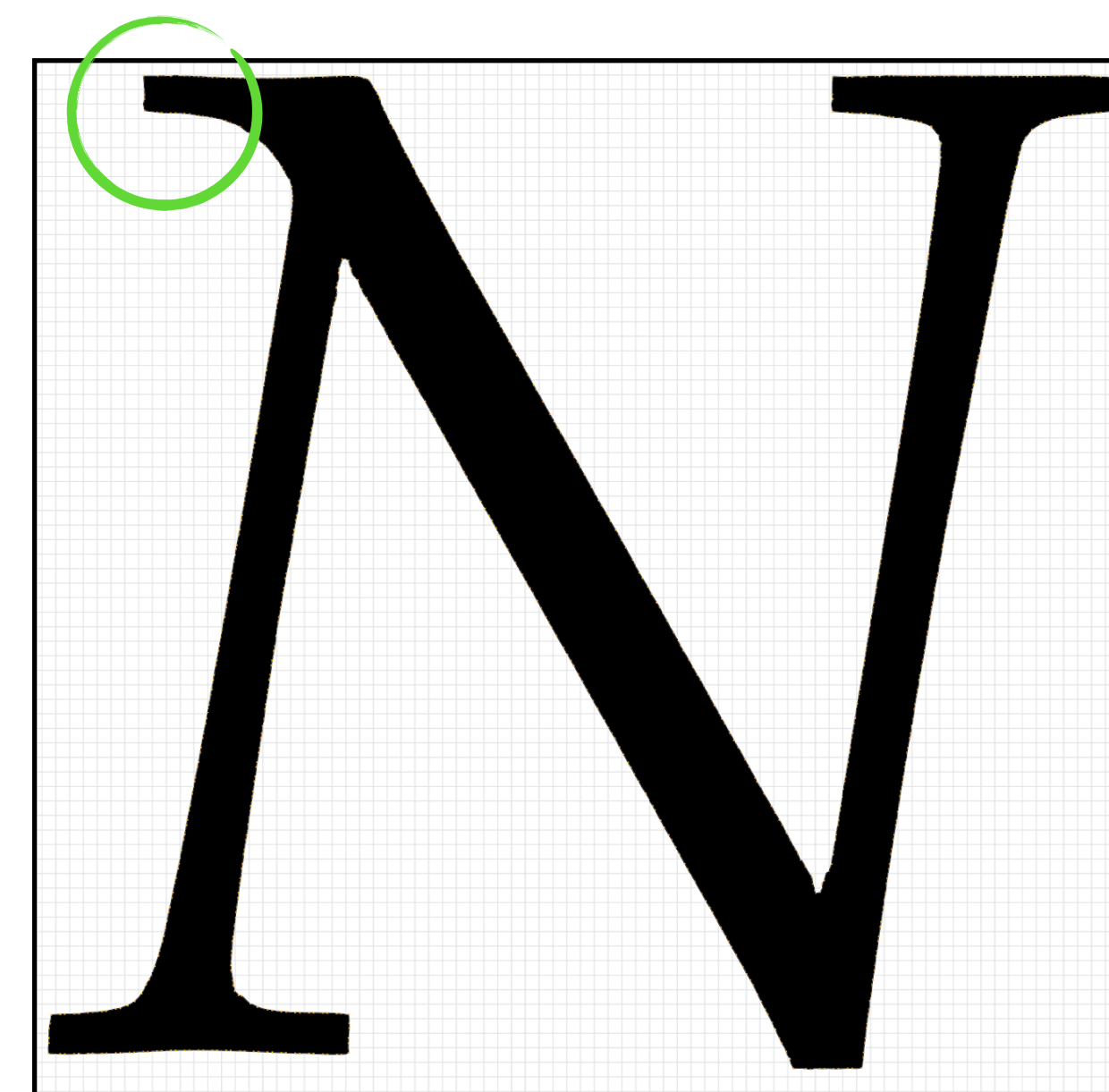
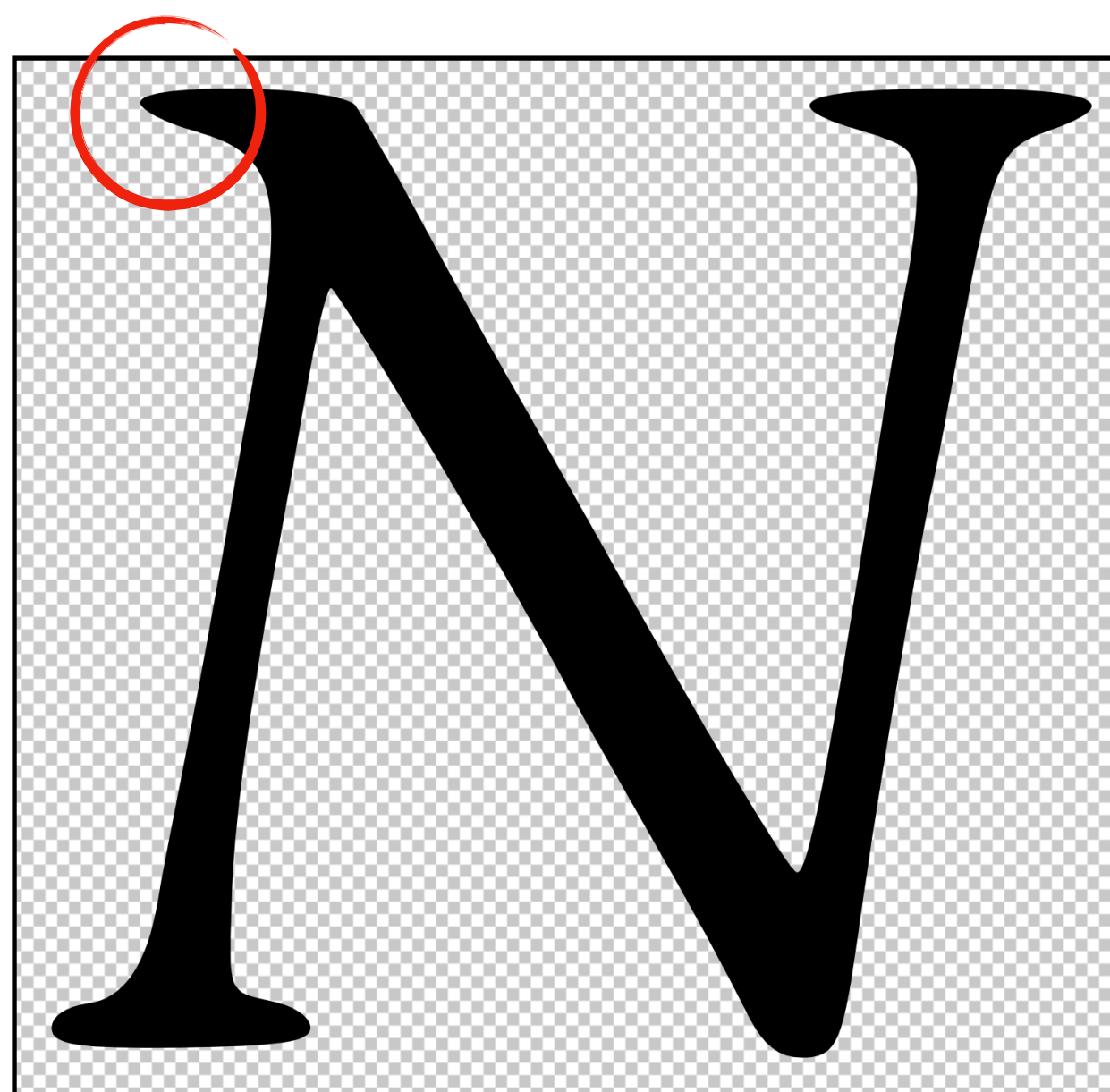
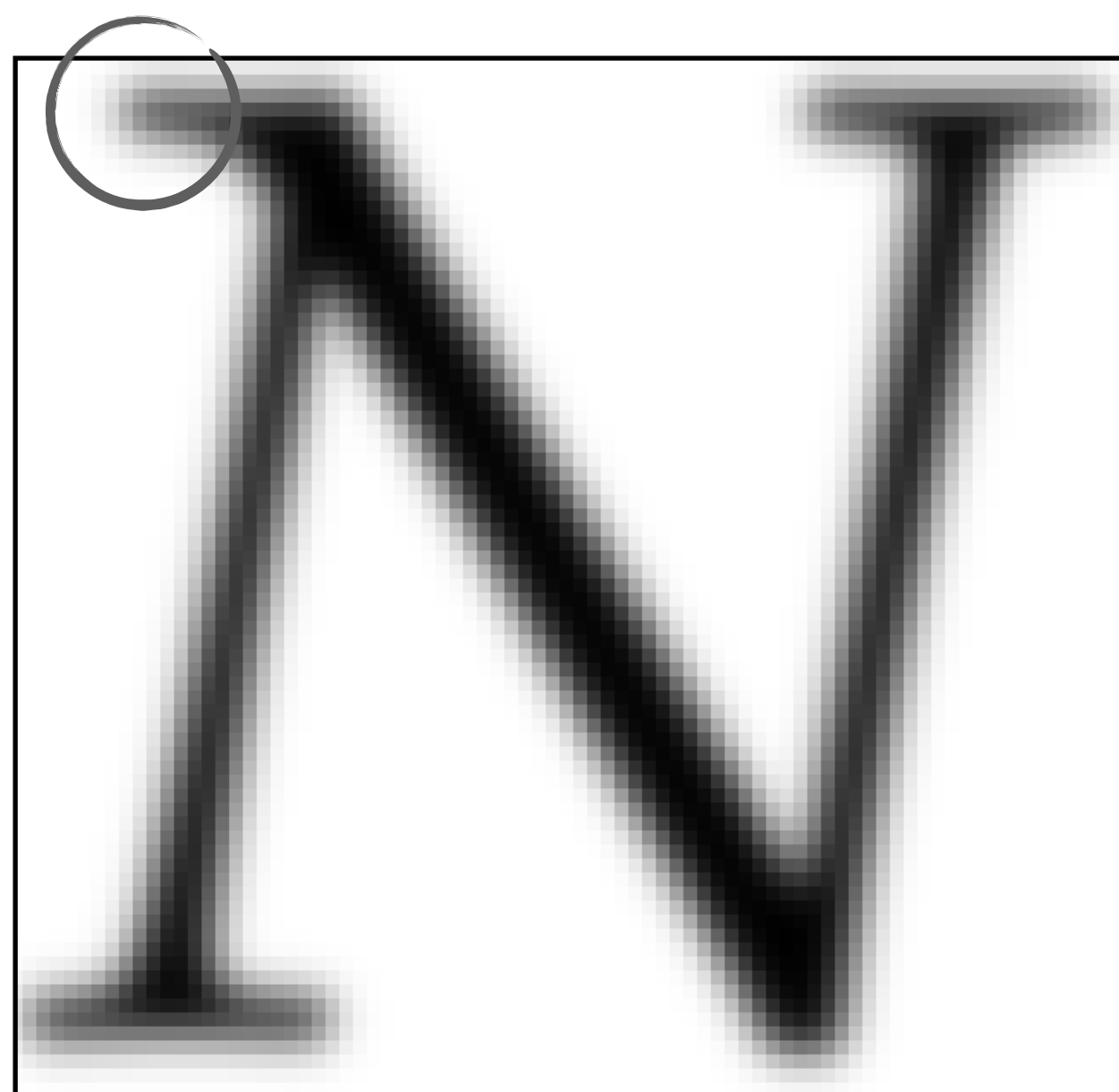


Y_1, \dots, Y_N
($H \times W \times 12$)

Results & Comparison



high-fidelity serif



Input

Previous (Potrace)

CNN

Reference

- [1] PyMuPDF. 2023. An enhanced Python binding for MuPDF, a lightweight PDF, XPS, and E-book viewer, renderer, and toolkit.
- [2] Peter Selinger. 2019. A tool for tracing a bitmap, which means, transforming a bitmap into a smooth, scalable image.
- [3] World Wide Web Consortium. 2011. Scalable Vector Graphics (SVG) 1.1 (Second Edition). <https://www.w3.org/TR/SVG11/>
- [4] William E Lorensen, Harvey E Cline. 1987. Marching cubes: A high resolution 3D surface construction algorithm. SIGGRAPH Comput. Graph. 21, 4 (July 1987), 163–169.
- [5] Rumelhart D E, Hinton G E, Williams R J. 1986. Learning representations by back-propagating errors. Nature, 323, 533–536.
- [6] LeCun Y, Boser B, Denker J S, Henderson D, Howard R E, Hubbard W, Jackel L D. 1989. Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4), 541-551.
- [7] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan et al. 2019. Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems 32 (2019).
- [8] He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- [9] Chen Zhiqin, Zhang Hao. 2021. Neural marching cubes. ACM Trans. Graph. 40, 6, Article 251 (December 2021).
- [10] Kurilin Ilya V et al. 2013. Generation of PDF with vector symbols from scanned document.
- [11] Ncraun. 2013. SmoothScan is a tool to convert scanned text into a vectorized output form.
- [12] Azriel Rosenfeld, John L Pfaltz. 1966. Sequential Operations in Digital Picture Processing. J. ACM 13, 4 (Oct. 1966), 471–494.
- [13] FontForge. 2023. Free (libre) font editor for Windows, Mac OS X and GNU+Linux.
- [14] Maas A L. 2013. Rectifier Nonlinearities Improve Neural Network Acoustic Models.