OFFLINE HANDWRITING CHARACTER RECOGNITION

Lan Vu

2006

Department of Information Technology and Management

University of Economics HCMC

OUTLINE

- Introduction
- Processes of the Offline Handwriting Recognition
- Applications
- Discussion

INTRODUCTION

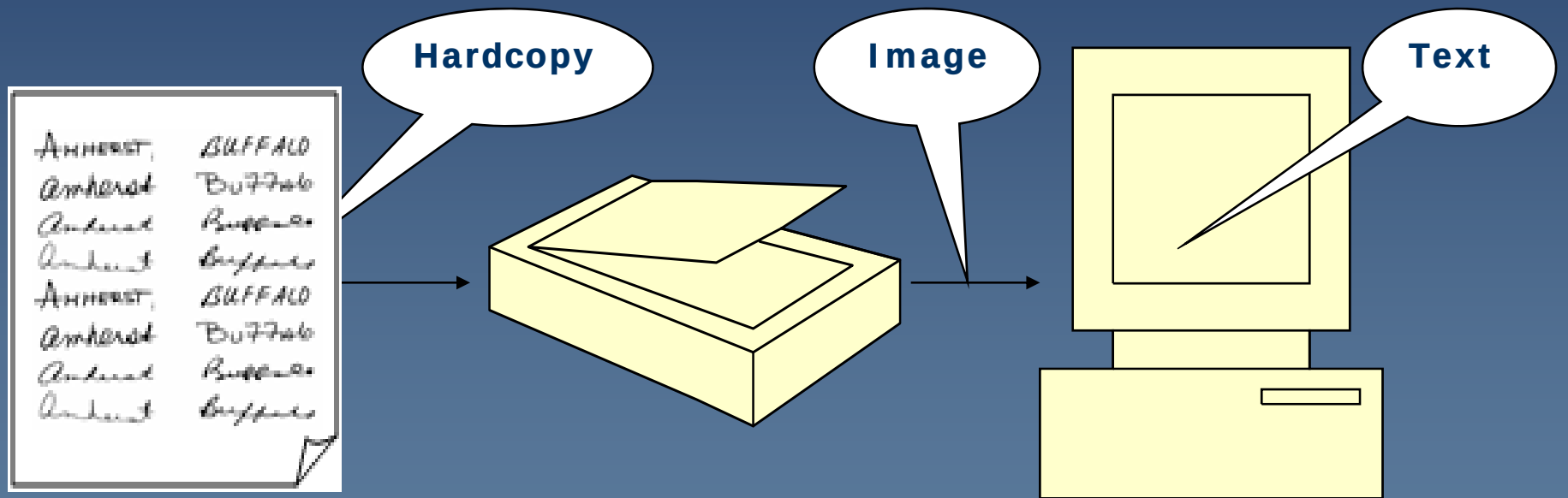
What is Handwriting Character Recognition (HCR)?

- A method of converting visually readable characters into computer readable characters.
- HCR programs are used with scanners to read text in hardcopy into the computer.

Why need HCR?

- People can read characters with eyes
- Computers can’t
- Need a method that allows computer on reading document

What does HCR do ?



What does HCR do ?

- Computer use binary codes (ASCII) to represent characters.
- Scan documents to get its images into computer
- HCR applications will convert the image to ASCII (or to text).
- A word processing program is used to correct the mistakes

The HCR Challenges

- HCR has been the subject of intensive research during the last decades:
  - A very challenging scientific problem
  - Provides a solution for processing large volumes of data automatically.
  - Have numerous applications such as address and zip code recognition, writer identification, etc.
- HRC is one of the benchmark problems of AI research. Recognizing characters seems to be simple with a person but extremely difficult for computer programs to do it.

The HCR Challenges

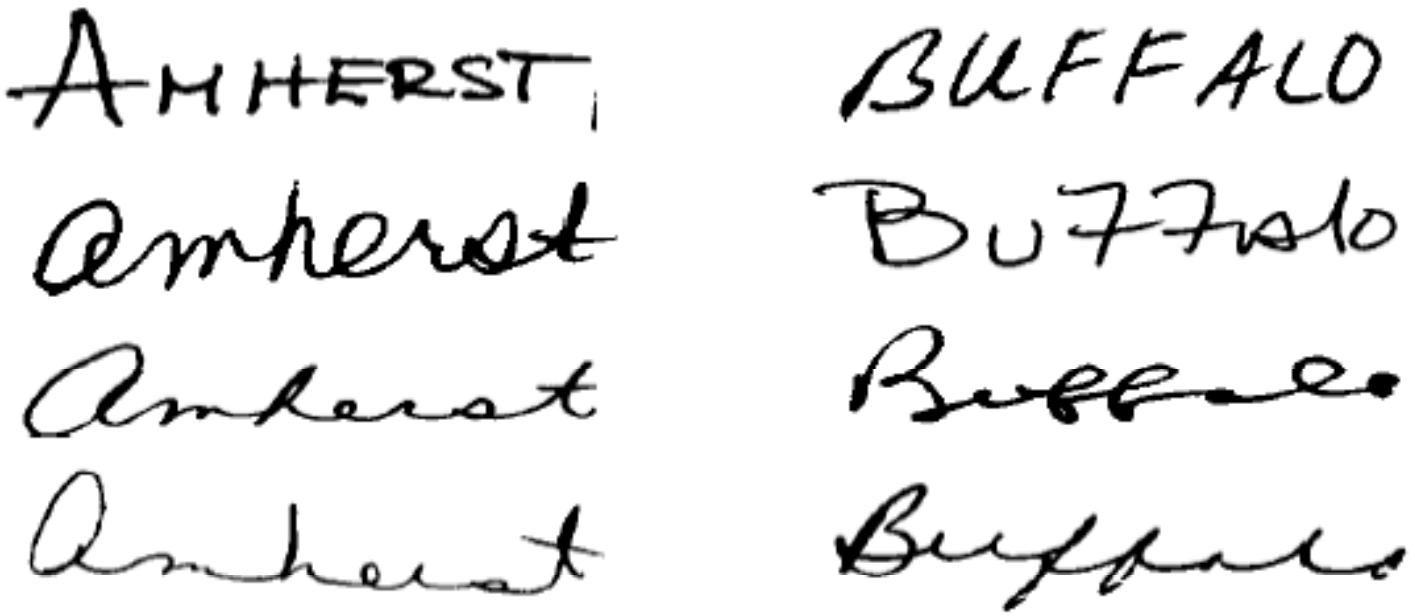


Figure 1.3: Handwritten words in different writing style

Processes of the Handwriting Character Recognition

How does HCR application work?

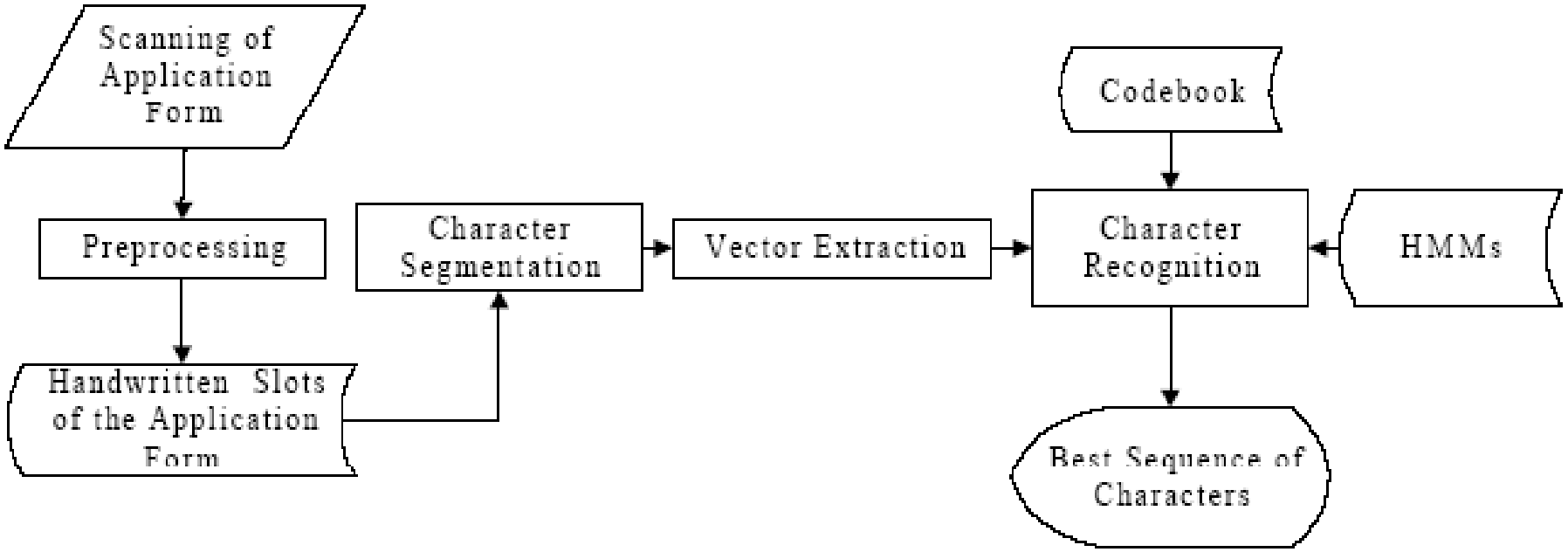
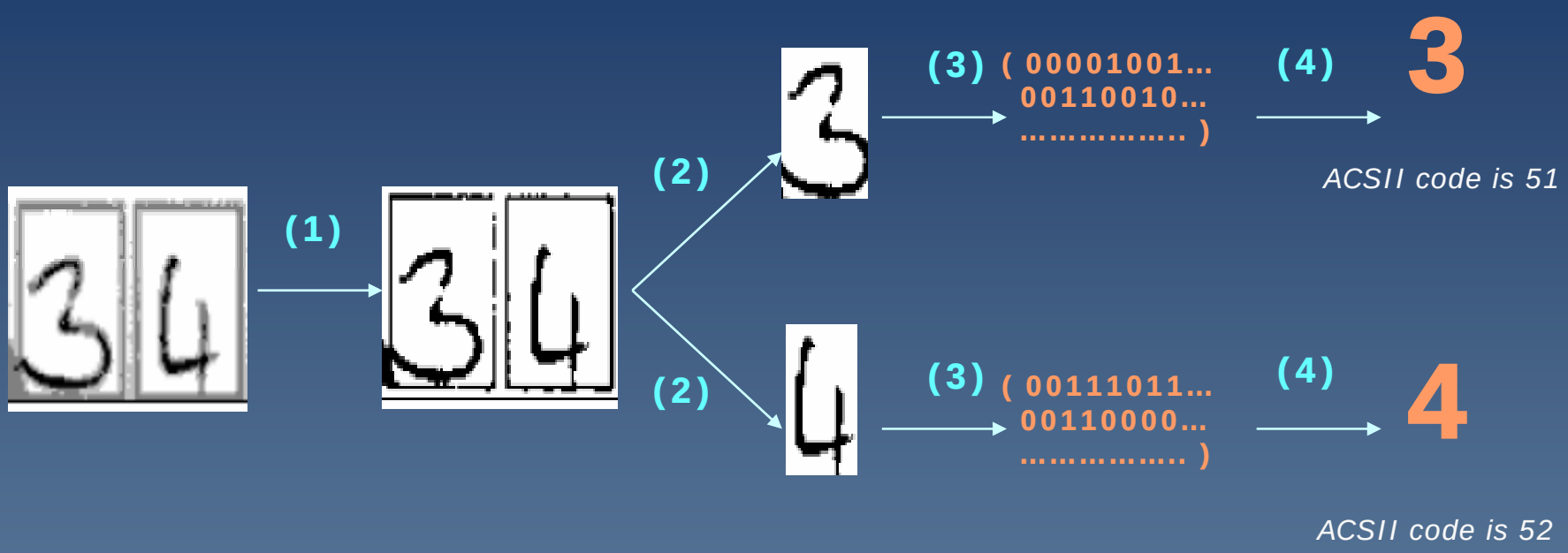


Figure 1: The main processes of the System

The processes of HCR application

1. ***Preprocessing:*** the hardcopy is scanned and the handwritten parts are found, separated and transformed into matrices with binary values (0s for the black pixels and 1s for the white ones).
2. ***Character segmentation:*** looking for the area of each character in the matrix.
3. ***Feature extraction:*** each character is divided vertically so that its primitive features are shown. Each primitive feature is represented by a vector.
4. ***Recognition procedure:*** follows which is based on a recognition model.

The results of each process



(1) Preprocessing
(2) Character segmentation
(3) Feature extraction
(4) Recognition procedure

Stage 1: Preprocessing

Why need this task?

- Scanning images have different quality...
- Recognition requires the best images for correct result.
- Simplify the recognition task

The processes of HCR application

- Before this process, document must be scanned into the computer through a scanner or some other image devices. The archived image can be stored in kinds of file (bitmap, jpeg ...) or in computer memory
- In the preprocessing stage, there are several algorithms used by the application to convert the raw image into proper image. The purpose of these transformation algorithms is to provide the best possible input data for the recognition algorithms.

The problems of scanning image

- Omitting the breaks of characters.
- Filtering the background information that is part of the original printed form and is not necessary or disturbs the recognition.
- Checking and correcting skewed pictures (in case of both rotated and slant pictures)
- Normalizing characters

The problems of scanning image

- Omitting the breaks of characters.
- Filtering the background information that is part of the original printed form and is not necessary or disturbs the recognition.
- Checking and correcting skewed pictures (in case of both rotated and slant pictures)
- Normalizing characters

The Commonly Used Algorithms

- **Binary Image Making Algorithm**

  To convert original image (color or gray scale) into binary image (0s for the black pixels and 1s for the white ones).

  Advantage:

  – Easily for other image processing algorithms

  – Reduce the complexity in computing.

The Commonly Used Algorithms

- **Binary Image Making Algorithm**

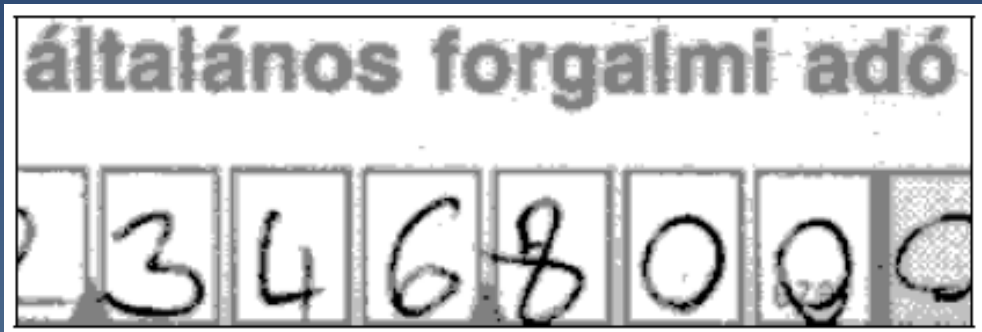


Figure 1 Grascale (8 bit) scanned image before color-filtering

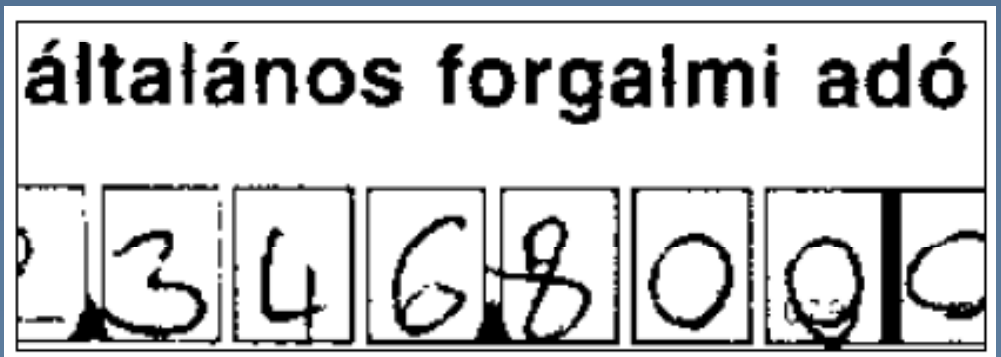


Figure 2 After filtering (binary image)

The Commonly Used Algorithms

- **Position-seeker algorithm**

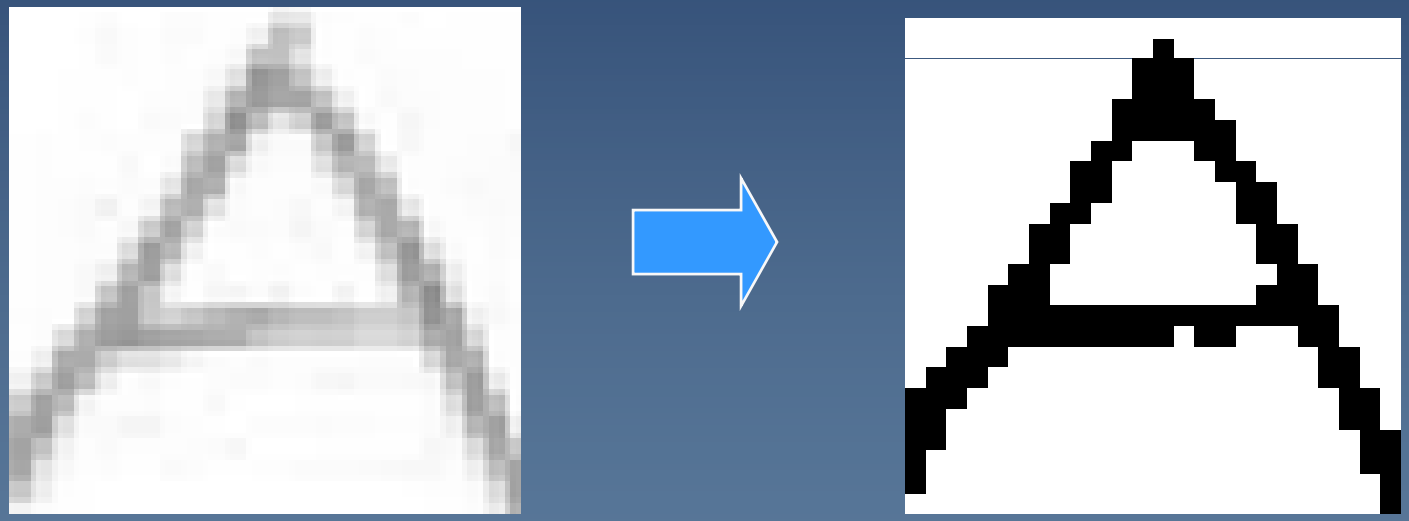
  Used for eliminating the sliding of pictures during the scanning. It searches for the top left corner of the picture.
- **Summary of the algorithm:**
  - The number of pixels/row & column are determined and stored in a matrix.
  - The largest number of pixels is sought in rows and columns.
  - The top left corner the first row & column is sought where the number of pixels reached the tenth of the picture's size or at least the half of the maximal pixel number.

The Commonly Used Algorithms

- **Noise filtering algorithm**
  - The scanning the image may get noisy even if no noise is present in the original picture so noise filtering is definitely necessary.
  - Using this method only small, one-pixel errors can be corrected.
  - Cannot manage stronger noise filtering because we may loose some important pixels from the even originally not too good quality picture.

The Commonly Used Algorithms

- **Noise filtering algorithm**



The Commonly Used Algorithms

- **Straightening algorithm**

The Commonly Used Algorithms

- **Dilatation algorithm**

  To revise some breaking characters of scanning image.

  The summary of the dilatation: if a white pixels has at least one black neighbor-pixel than that pixel will be also black. Black pixels aren't needed to be examined.

Stage 2:
Character Segmentation

Character Segmentation

- This is the most difficult step of preprocessing.
- It is also the most important step.
- The goal of segmentation is to break the handwriting sample down into smaller entities representing individual characters or individual pieces of a character

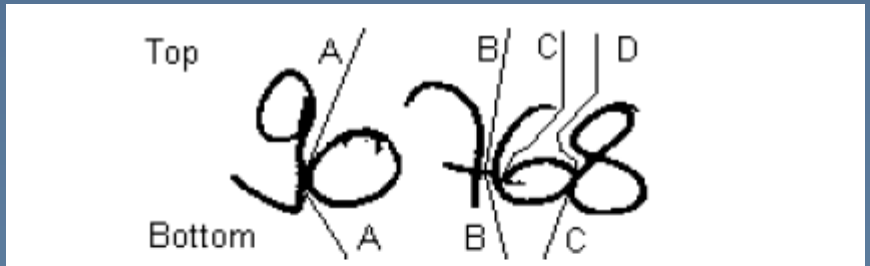


Fig. 15. The different possible segmentation paths for an example ZIP code block.

The troubles

- Handwriting has various types which is hard to separate

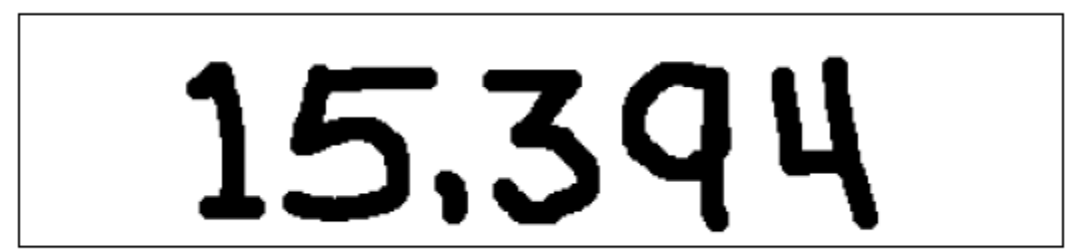


Figure1: Example of well written numerical amount

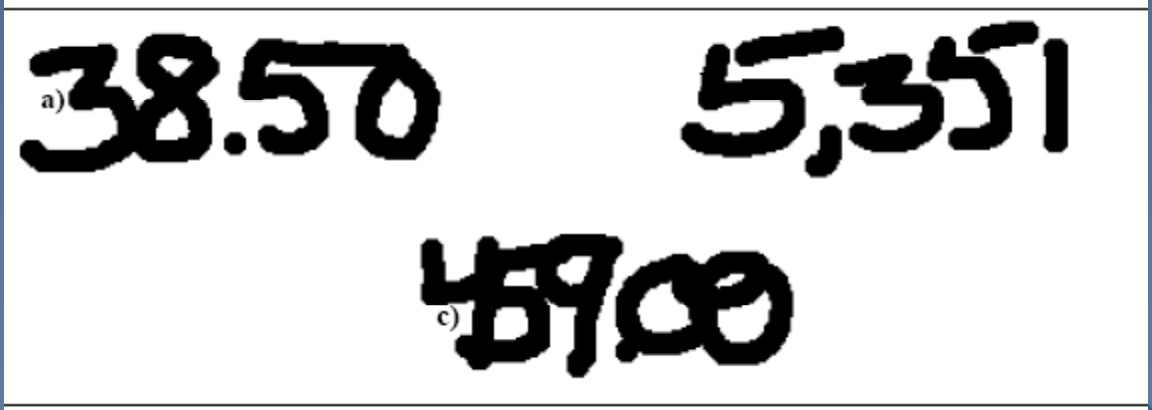


Figure 2: Examples of a) touching b) disjointed and c) overlapping numbers

The algorithms of segmentation

- **Drop Fall algorithms**

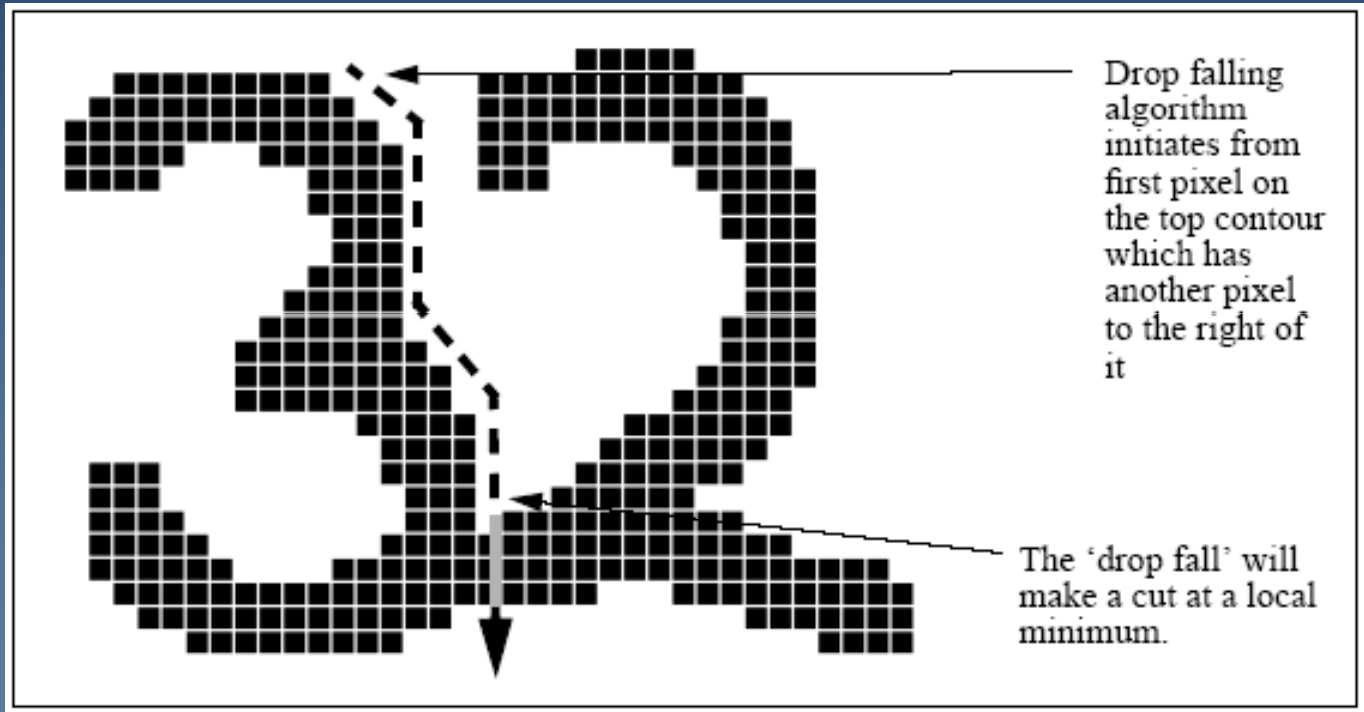


Figure 3: Example of the path of a 'drop-falling' algorithm

The algorithms of segmentation

- **Min-Max algorithms**

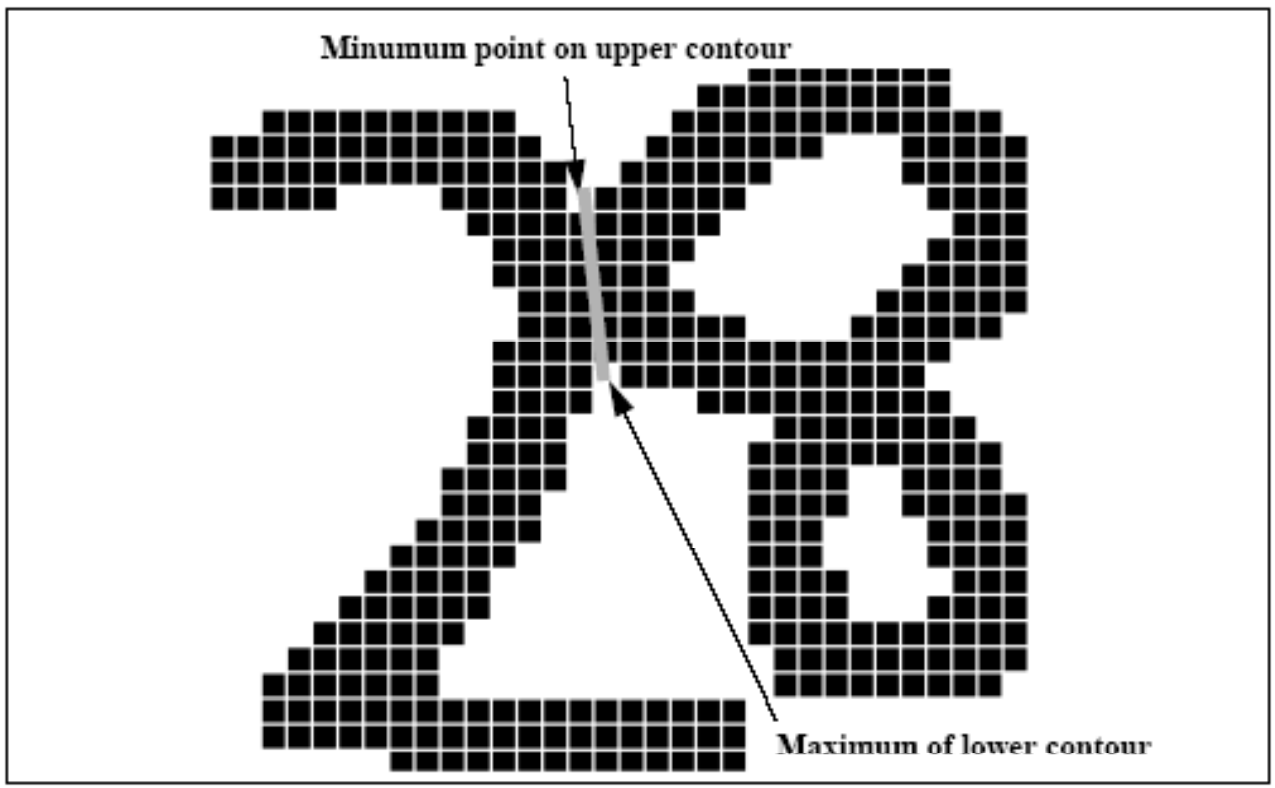


Figure 6: Example of a cut produced by a min-max based heuristic

The algorithms of segmentation

- **Hybrid Algorithms**

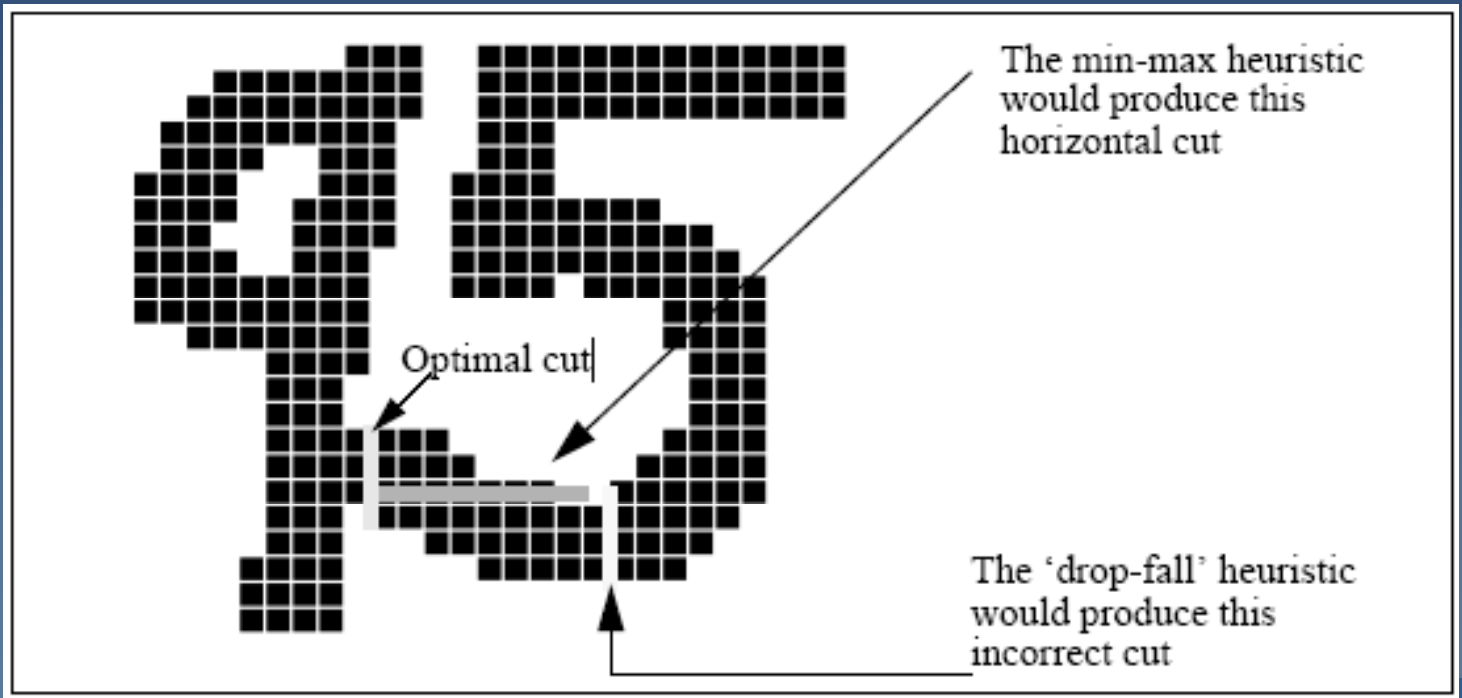


Figure 7: Both 'drop-fall' and min-max heuristics fail on the above example

The algorithms of segmentation

- Handwriting has various types which is hard to separate

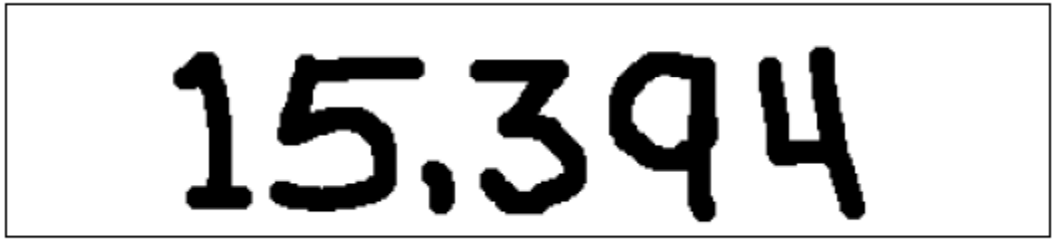


Figure1: Example of well written numerical amount

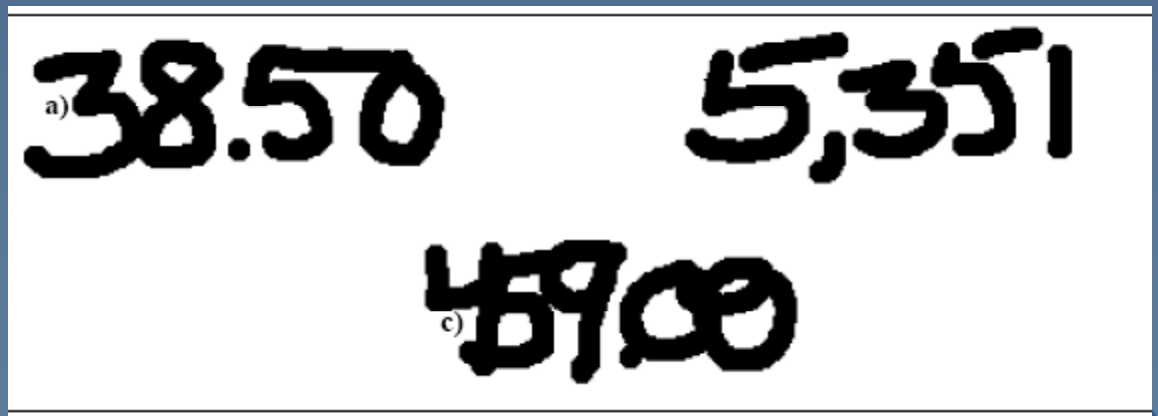


Figure 2: Examples of a) touching b) disjointed and c) overlapping numbers

Stage 3:
Feature Extraction

Feature extraction

Contour Making

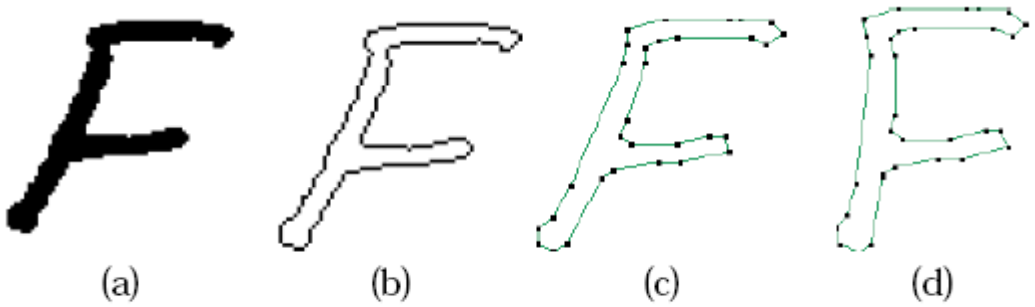


Figure 3.9: Contour representation and preprocessing; (a) input image, (b) pixel based contour representation, (c) piece-wise linearized contour and (d) piece-wise linear contour representation after normalization

Feature extraction

***E*ach character is divided vertically so that its primitive features are shown. Each primitive feature is represented by a vector.**

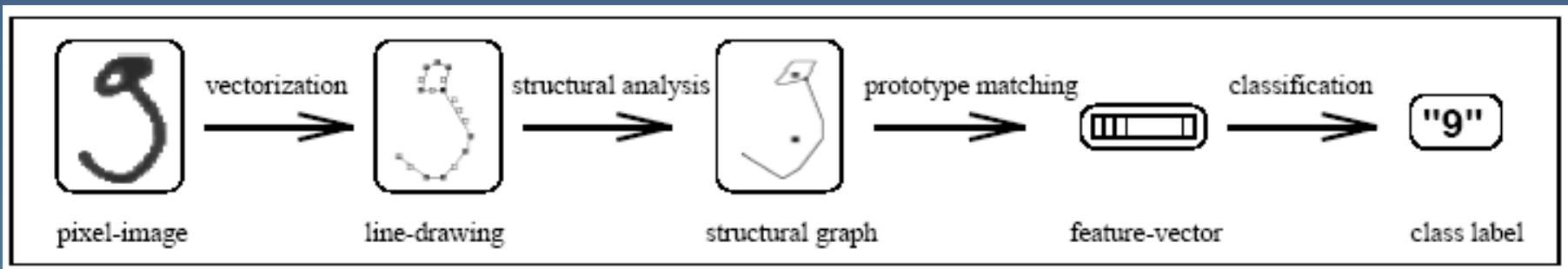


Fig. 10. Stages of the Structural Digit Recognition.

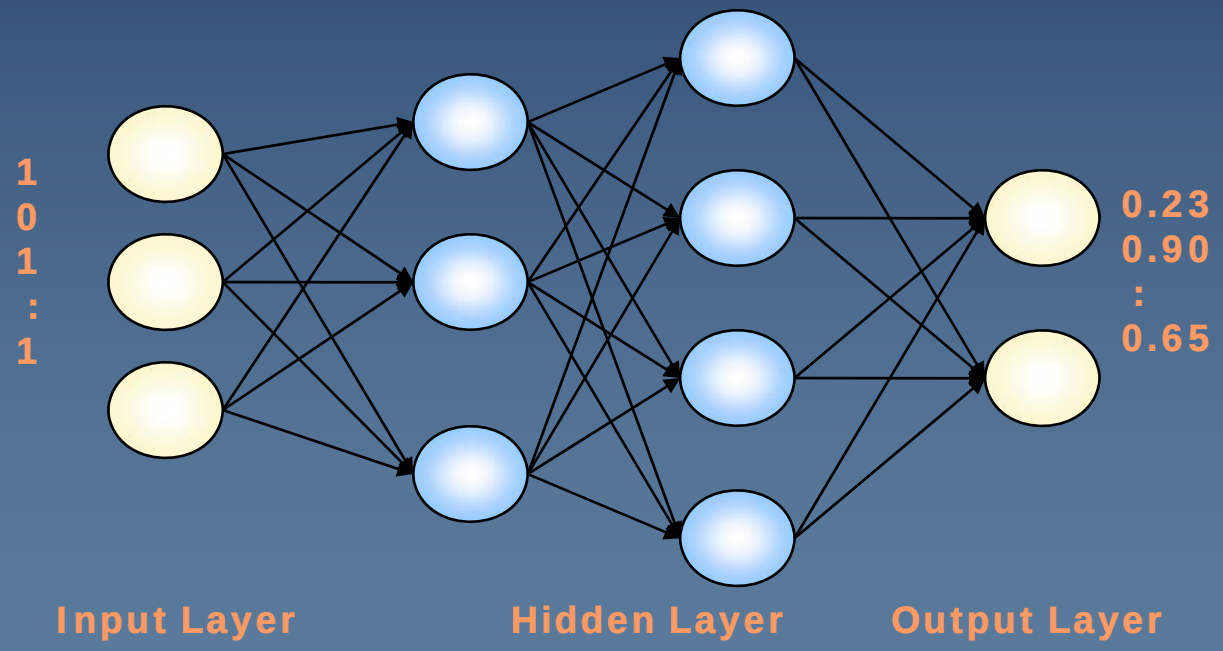
Stage 4:
Character Recognition

Some methods for recognition

- Neural Network (FNN, ANN, Combined NN...)
- Hidden Markov model
- Wavelet packet transform and Neuro-Fuzzy
- Genetic-Algorithm
- Adaptive Statistical Similarity (using Statistical tools such as Bayesian and Gaussian)
- Hybrid methods

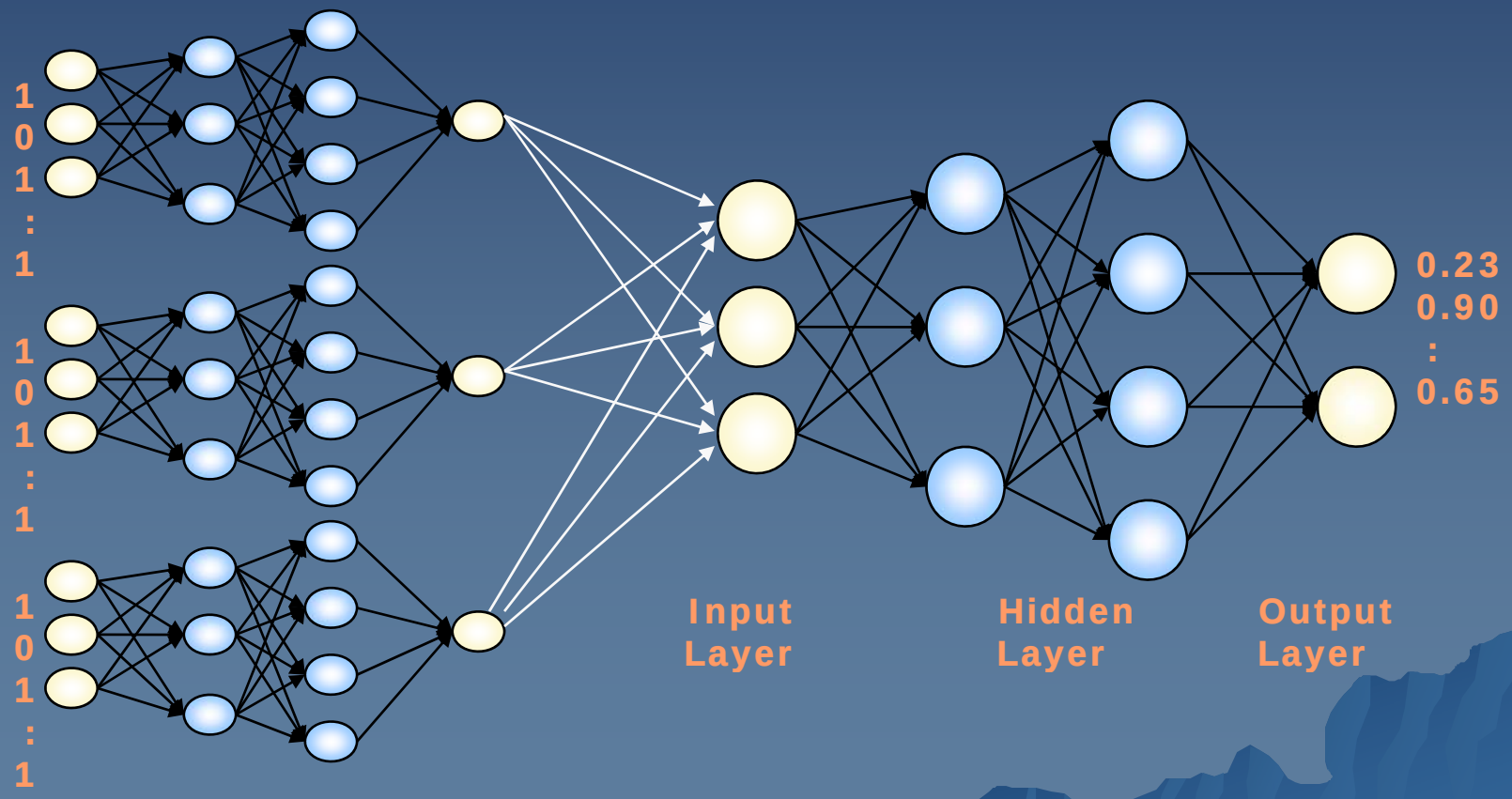
Our proposed Method

- Apply Feedforward Neural Network for recognition (**Multilayer Neural Network**)



Our proposed Method

- Or using **Combined Neural Networks**



Structure of FNN

- Input Layer : 576 Nodes (each node represent for a pixel of binary image)
- Hidden Layer : variety (about 3 → 5 layers, each layer and 256 → 1024 nodes per layers)
- Output Layer: 256 Nodes (each node represents for a character)

Applications

Address readers

- The address reader in a postal mail sorter locates the destination address block on a mail piece and reads the ZIP Code in this address block.
- If additional fields in the address block are read with high confidence the system may generate a 9 digit ZIP Code for the piece. The resulting ZIP Code is used to generate a bar code which is sprayed on the envelope.
- The Multiline Optical Character Reader (MLOCR) used by the United States Postal Service
- (USPS) The character classifier recognizes up to 400 fonts and the system can process up to 45,000 mail pieces per hour.

Form readers

- A form reading system needs to discriminate between pre-printed form instructions and filled in data. The system is first trained with a blank form. The system registers those areas on the form where the data should be printed. During the form recognition phase, the system uses the spatial information obtained from training to scan the regions that should be filled with data.
- Some readers read hand-printed data as well as various machine written text. They can read data on a form without being confused with the form instructions. Some systems can process forms at a rate of 5,800 forms per hour.

Check readers

- A check reader captures check images and recognize courtesy amounts and account information on the checks. Some readers also recognize the legal amount on checks and use the information in both fields to cross-check the recognition results. An operator can correct misclassified characters by cross-validating the recognition results with the check image that appears on a system console.

Airline ticket readers

- In order to claim revenue from a airline passenger ticket, an airline needs to have three records matched: reservation record, the travel agent record and the passenger ticket.
- However, it is impossible to match all three records for every ticket sold. Current methods which use manual random sampling of tickets is far from accurate in claiming the maximal amount of revenue.
- Several airlines are using a passenger revenue accounting system to account accurately for passenger revenues. The system reads the ticket number on a passenger ticket and matches it with the one in the airline reservation database. It scans tickets up to 260,000 tickets per day and achieves a sorting rate of 17 tickets per second.

Bill processing systems

- In general a bill processing system is used to read payment slips, utility bills and inventory documents. The system focuses on certain regions on a document where the expected information are located, e.g. account number and payment value.

Passport readers

- An automated passport reader is used to speed up the returning American passengers through custom inspections. The Reader reads a traveler's name, date of birth, and passport number on the passport and checks these against the database records that contain information on fugitive felons and smugglers.

Discussion

Discussion

- How can this problem be applied in our life? (Which fields? Their benefits?)
- What are the difficulties in accessing this problem?
- How many methods are applied for the process of recognition? What they are? And what is the best one?
- What are the achievements of this problem at present and its trend of research in the future?