Pattern Analysis / Recognition

Mihai Oltean

www.cs.ubbcluj.ro/~moltean

moltean@cs.ubbcluj.ro

What is pattern recognition?

"The assignment of a physical object or event to one of several prespecified categories" -- Duda & Hart

- A **pattern** is an object, process or event that can be given a name.
- A **pattern class** (or category) is a set of patterns sharing common attributes and usually originating from the same source.
- During **recognition** (or **classification**) given objects are assigned to prescribed classes.

- A **classifier** is a machine which performs classification.

Basic concepts

Pattern

$$\begin{bmatrix} x_1 \\ x_2 \\ \square \\ x_n \end{bmatrix} = \mathbf{x}$$

y
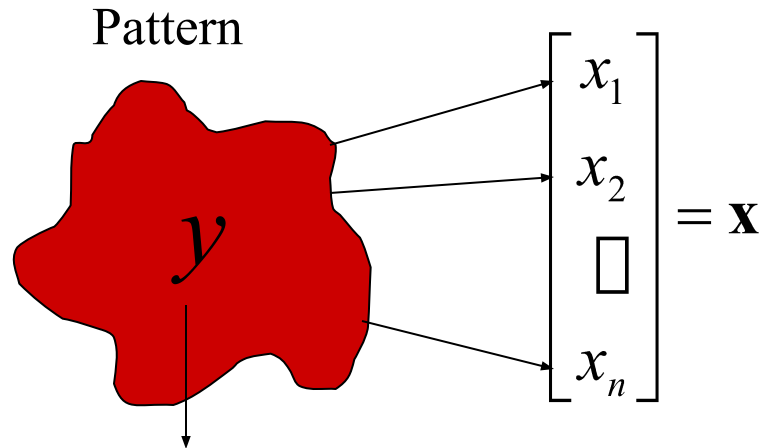
Feature vector $\quad \mathbf{x} \in X$

- A vector of observations (measurements).
- $\mathbf{x}$ is a point in feature space X.

Hidden state $\quad z \in Z$

- Cannot be directly measured.
- Patterns with equal hidden state belong to the same class.

Task

- To design a classifer (decision rule) $\quad q : X \rightarrow Z$

which decides about a hidden state based on an observation.

Examples of applications

• **Optical Character Recognition (OCR)**

• **Biometrics**

• **Diagnosis systems**

• **Military applications**

• Handwritten: sorting letters by postal code, input device for PDA's, writer identification

• Printed texts: reading machines for blind people, digitalization of text documents.

• Face recognition, verification, retrieval.

• Finger prints recognition.

• Speech recognition.

• Medical diagnosis: X-Ray, EKG analysis.

• Machine diagnostics, waster detection.

• Automated Target Recognition (ATR).

• Image segmentation and analysis (recognition from aerial or satelite photographs).

Handwritten / typewritten text recognition

- Translate (hand or type) written text into machine-editable text.

- Where?
  - Bank Cheques,
  - Postal Addresses,
  - Tax Forms,
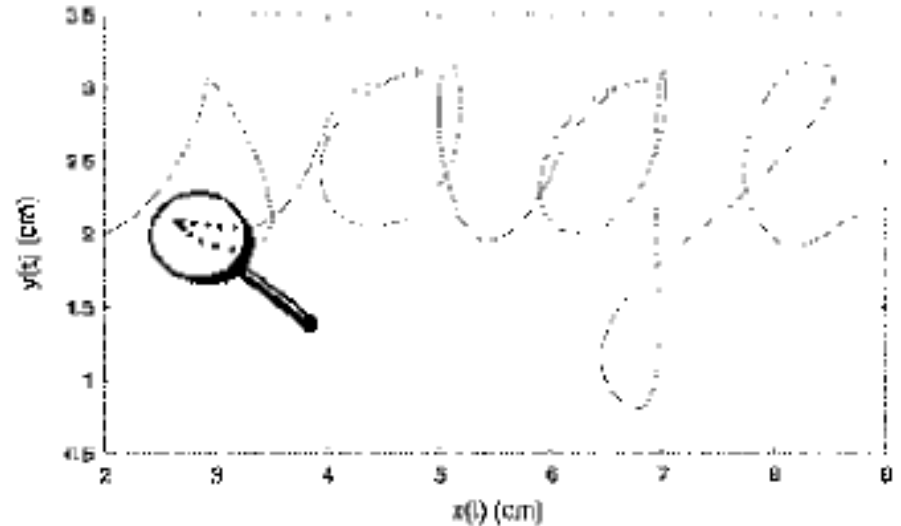 - Input Device for PDA's,
  - Book scanning.

Postal letter

Bank check

- US and Brazilian banks
  - 68 billion checks / year in US

Offline   vs.   Online

In the off-line case, only the completed writing is available as an image.

In the on-line case, the two-dimensional coordinates of successive points of the writing as a function of time are stored in order, i.e., The order of strokes made by the writer is readily available.

Difficulties for typewritten

- New fonts
- Strange / Noisy backgrounds
- Locate the characters on the page
- Compute the size of characters
- Fuzzy profile
- …

Image Quality is important !

Difficulties for typewritten

Difficulties for offline handwritten

- Isolated characters:
  - Ambiguity, different styles.
- Strings:
  - Unknown length, broken, overlapping, and touching digits.

Bank Cheques

#500.000,50#

#620.000,52#

#691.502,35

##1,60## (#1,60#)

332.263,98 (332.263,93)

7165.501,84# (#165.501,84#)

Fuzzy digits
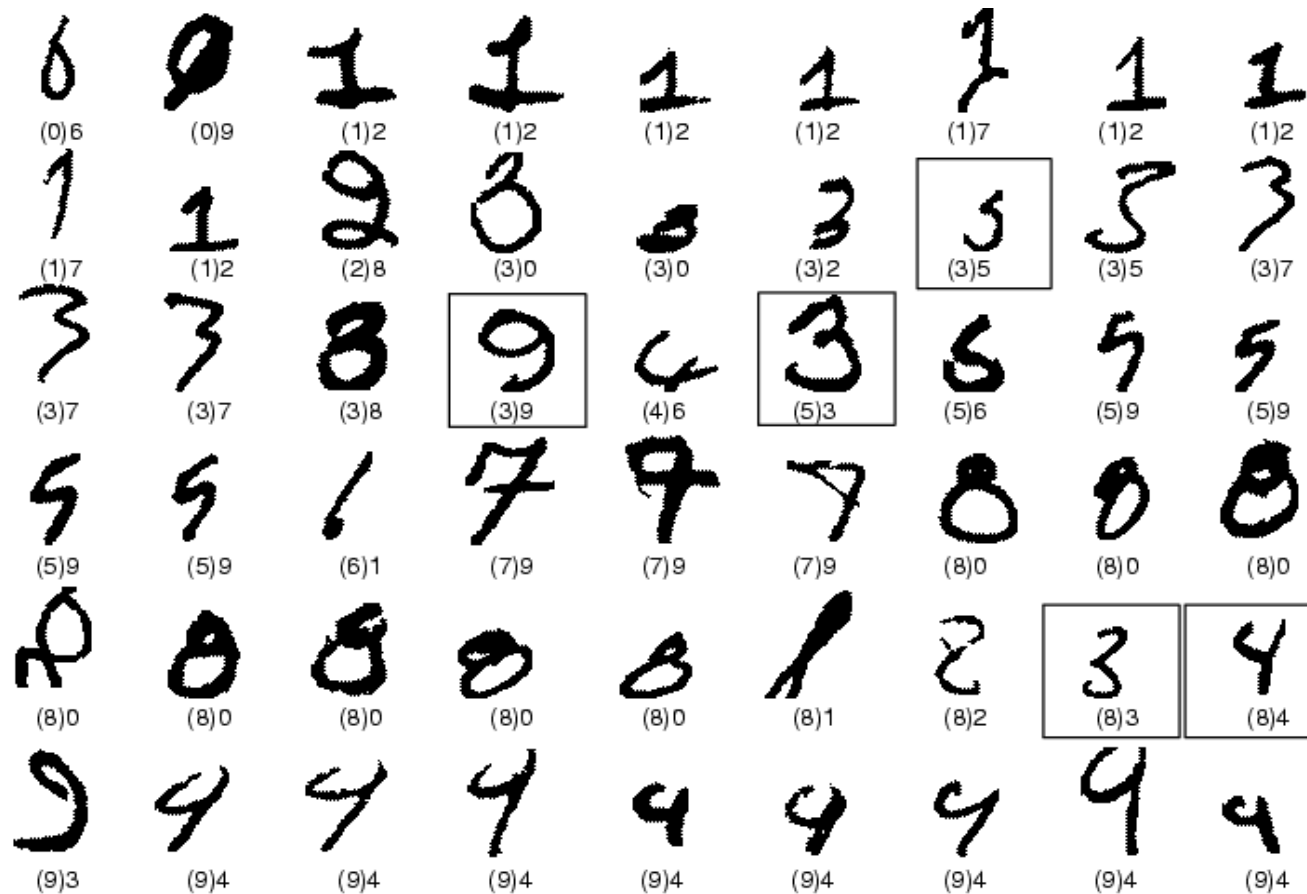
Ambiguity Between Digits And Pieces Of Digits.

Original Image
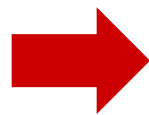
Segmentation Rules

Optimal Segmentation

A segmentation hypotheses

How to solve it?

- Data preprocessing
  - Very important
- Solve it

Data pre-processing

- Extract ink from paper (Thresholding)
- Noise removal
- Line segmentation
- Word/Character segmentation (splitting)
- Normalization (slant, thickness, size)

Thresholding
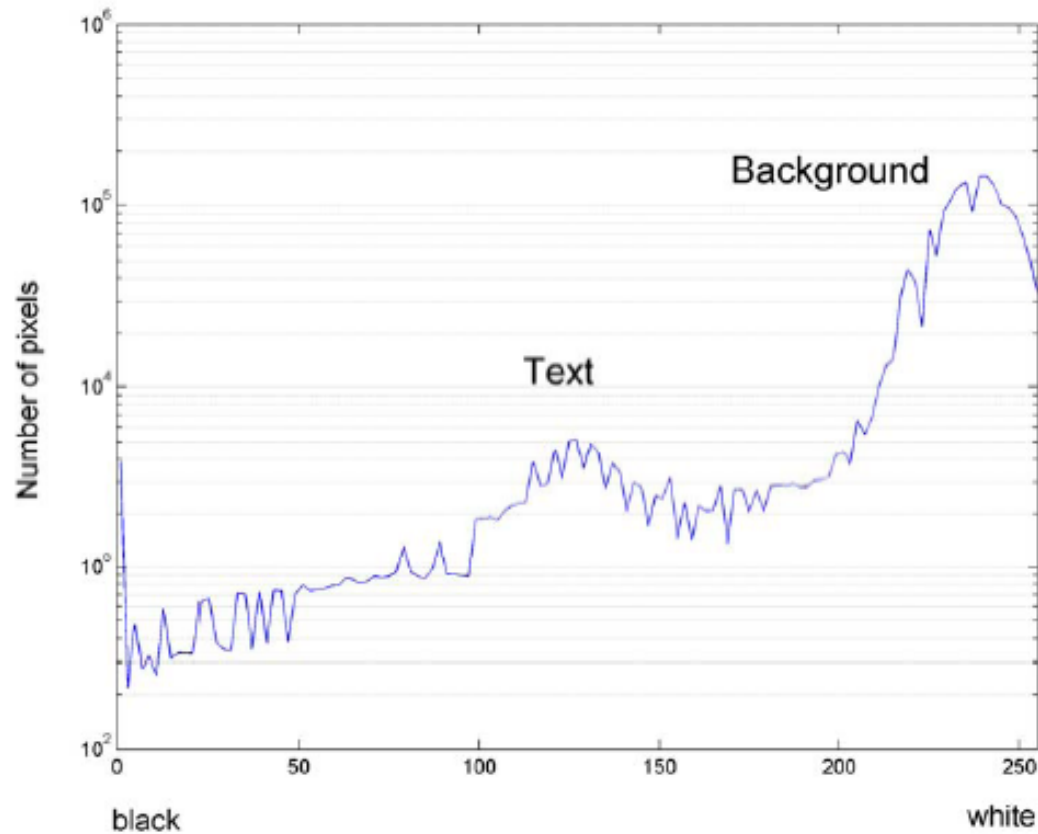
Figure 3: Histogram of Image of Check

Image after thresholding

Splitting

- **Split Characters (if in pair – see infinity)**
  - **Multiple Character Detection**
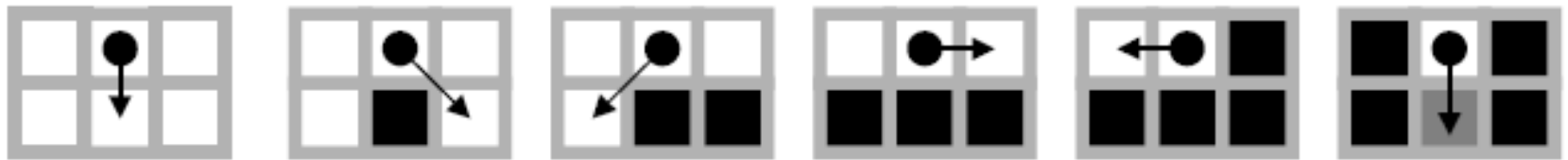 - **Vertically Straight Cut**

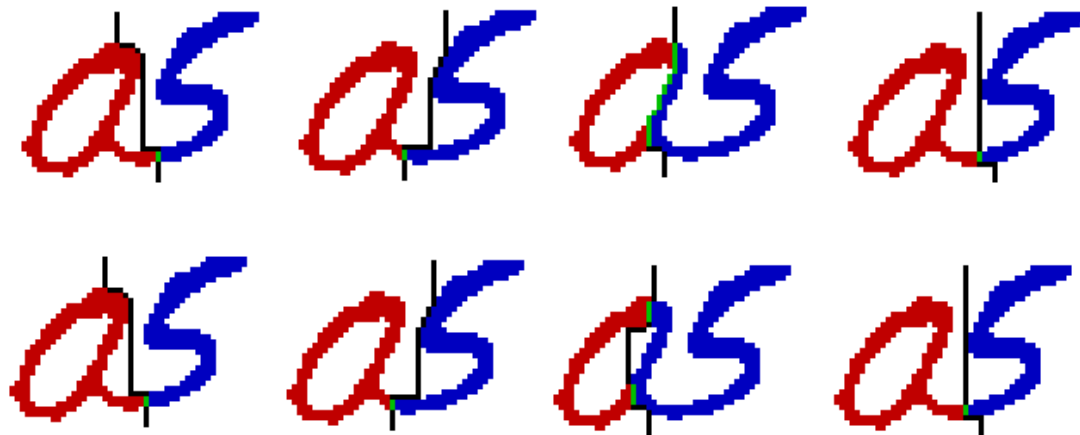(A) Original Image   (B) Vertical Cut

Drop fall algorithm

Starting point is important !!!

This gives you the cutting point. An incorrect cutting point leads to poor results

Splitting - Drop-Fall Method

8 possible paths to divide a connected component depending on three aspects: a starting point, movement rules, and orientation
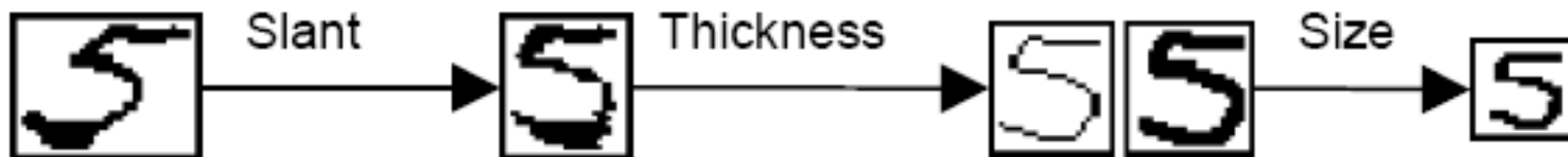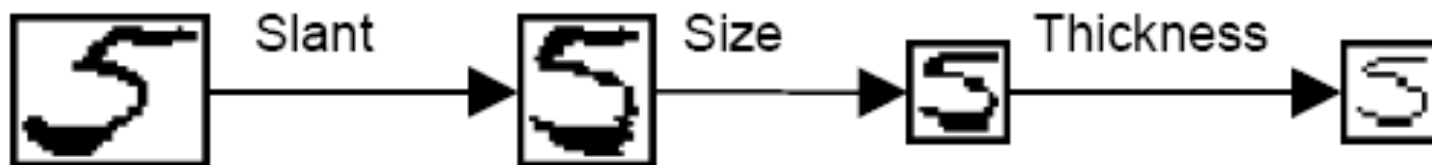
Difficulties with Drop-fall method

- A random sampling of 400 checks showed that about 50 % of all connected digits consisted of a pair of zeros and 80-85% of successive zeros are connected.
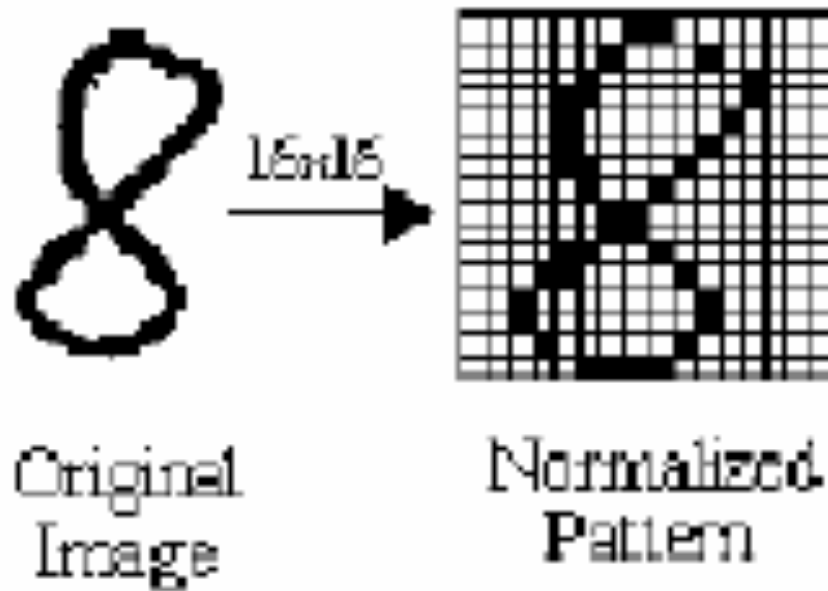
- 00 can be read as 07 or 02 or 60.

Normalization

Representation of input

- Bit maps
- Grid maps
- Black – white ratio
- Histograms
- Semantic

Bit Maps

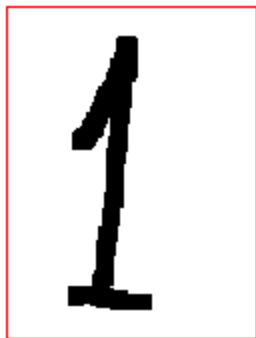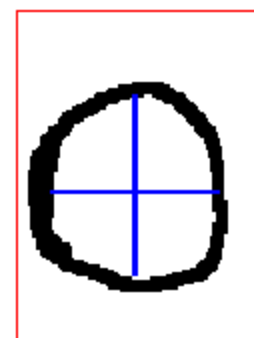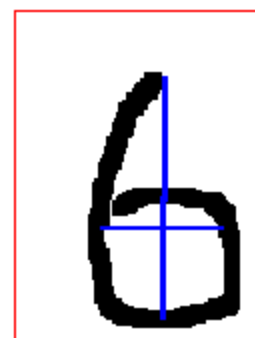Original Image → 16x16 → Normalized Pattern

Grid maps

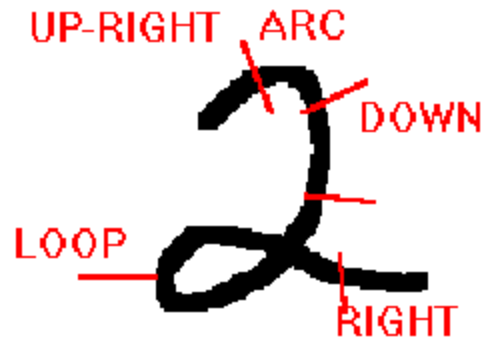# Black/White ratio

- (a) ratio of black pixels to white pixels

(b) ratio of height to width

Histograms

Semantic Approach - characteristics based on contours

Example characteristics of the contour:

"up", "down", "diagonal up", "arc", "loop", etc

Training

- Neural Networks
- GP techniques
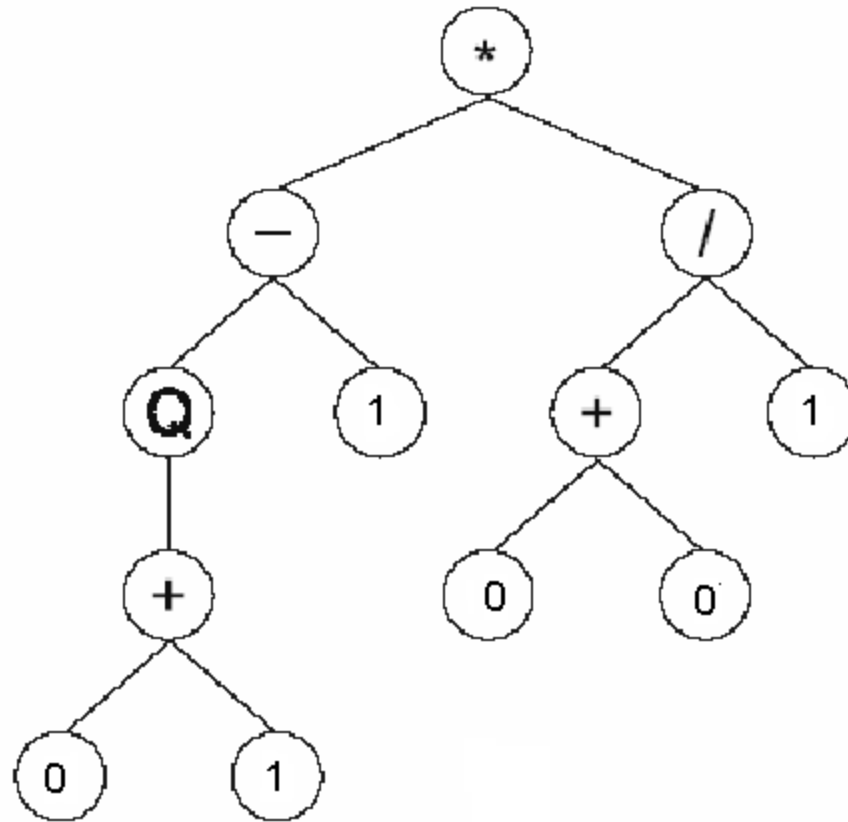
GP/NNs for handwritten reco

- Inputs are bit maps, grid maps, etc

- The output is a numerical value:
    - [0,1] we have A
    - [1,2] we have B
    - [2,3] we have C
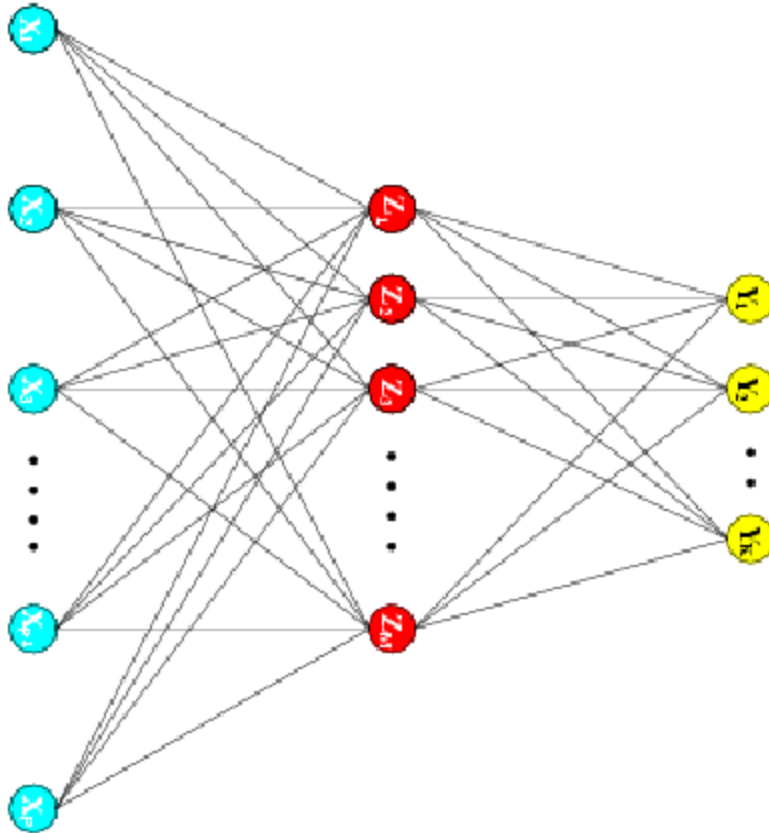   - …

Multiple outputs reco…

☐ Take the output with the greatest value

GP techniques

Neural Networks

3-layers fully connected neural networks are usually used:

256-80-10 (for digits recognition.

Improving recognition

- Using a lexicon of words
- Information about grammars

Existing Applications