Pattern Recognition

CSE 802
Michigan State University
Spring 2017
Lecture 1, January 9, 2017
Pattern Recognition

“The real power of human thinking is based on recognizing patterns. The better computers get at pattern recognition, the more humanlike they will become.”
Ray Kurzweil, NY Times, Nov 24, 2003

“The problem of searching for patterns in data is a fundamental one and has a long and successful history.” Bishop
"And you are telling me you work in pattern recognition!"
Pattern Recognition

The act of taking as input sensed data (measurements) and taking an action based on the “category” or “class” of the pattern.
What is a Pattern?

“A pattern is the **opposite of a chaos**; it is an entity vaguely defined, that could be given a name.”

(Watanabe)
Recognition

Identification of a pattern as a member of a category (class) we already know, or we are familiar with:

- **Classification** (known categories)
- **Clustering** (learning categories)
Pattern Class

- A collection of similar (not necessarily identical) objects
- A class is defined by class samples (exemplars, prototypes)
- Intra-class variability
- Inter-class similarity
- How to define similarity?
Intra-Class Variability

Handwritten numerals
Inter-class Similarity

Characters that look similar

Identical twins

Jain CSE 802, Spring 2017
Cat vs. Dog: 2-class Classification
(Supervised) Classification

Labeled training samples for classifier design

Jain CSE 802, Spring 2017
Clustering: Unsupervised Classification

Training samples are unlabeled

Jain CSE 802, Spring 2017
1. Use shape and appearance to classify a pet breed automatically from an image.

2. Shape is captured by a deformable part model detecting the pet face; appearance is captured by a bag-of-words model to describe the pet fur.

3. Automatically segmenting the animal in the image.

4. Two classification approaches: a hierarchical one, in which a pet is first assigned to the cat or dog family and then to a breed, and a flat one, in which the breed is obtained directly.

https://www.robots.ox.ac.uk/~vgg/publications/2012/parkhi12a/parkhi12a.pdf
Problem Definition & Data

Oxford-IIIT Pet dataset: 7,349 images of cats & dogs of 37 different breeds: 25 dogs & 12 cats. ~200 images/breed, split randomly into 50 for training, 50 for validation, and 100 for testing. Three tasks are defined:

• Pet family classification (Cat vs Dog, a 2-class problem)

• Breed classification given the family (a 12-class problem for cats and a 25-class problem for dogs)

• Breed and family classification (a 37-class problem)
Figure 2. Example images from the Oxford-IIIT Pet data. Two images per breed are shown side by side to illustrate the data variability.
Segmentation: Foreground v. Background

Figure 1. **Annotations in the Oxford-11IT Pet data.** From left to right: pet image, head bounding box, and trimap segmentation (*blue*: background region; *red*: ambiguous region; *yellow*: foreground region).
Feature Extraction

Figure 4. **Spatial histogram layouts.** The three different spatial layouts used for computing the image descriptors. The image descriptor in each case is formed by concatenating the histograms computed on the individual spatial components of the layout. The spatial bins are denoted by yellow-black lines.
### Table 4. Comparison between different models.

The table compares different models on the three tasks of discriminating the family, the breed given the family, and the breed and family of the pets in the Oxford-IIIT Pet dataset (Sect. 2). Different combinations of the shape features (deformable part model of the pet faces) and of the various appearance features are tested (Sect. 3.2, Fig. 4).
## Pattern Recognition Applications

<table>
<thead>
<tr>
<th>Problem</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech recognition</td>
<td>Speech waveforms</td>
<td>Spoken words, speaker identity</td>
</tr>
<tr>
<td>Non-destructive testing</td>
<td>Ultrasound, eddy current, acoustic emission waveforms</td>
<td>Presence/absence of flaw, type of flaw</td>
</tr>
<tr>
<td>Medical waveform analysis</td>
<td>EKG, EEG waveforms</td>
<td>Types of cardiac conditions, classes of brain conditions</td>
</tr>
<tr>
<td>Remote sensing</td>
<td>Multispectral images</td>
<td>Terrain forms, vegetation cover</td>
</tr>
<tr>
<td>Aerial reconnaissance</td>
<td>Visual, infrared, radar images</td>
<td>Tanks, airfields</td>
</tr>
<tr>
<td>Character recognition (page readers, zip code, license plate)</td>
<td>scanned image</td>
<td>Alphanumeric characters</td>
</tr>
</tbody>
</table>
## Pattern Recognition Applications

<table>
<thead>
<tr>
<th>Problem</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification and counting of cells</td>
<td>Slides of blood samples, micro-sections of tissues</td>
<td>Type of cells</td>
</tr>
<tr>
<td>Industrial inspection (PC boards, IC masks, textiles)</td>
<td>Scanned image (visible, infrared)</td>
<td>Acceptable/unacceptable</td>
</tr>
<tr>
<td>Factory automation</td>
<td>3-D images (structured light, laser, stereo)</td>
<td>Identify objects, pose, assembly</td>
</tr>
<tr>
<td>Web search</td>
<td>Key words specified by a user</td>
<td>Text relevant to the user</td>
</tr>
<tr>
<td>Fingerprint identification</td>
<td>Input image from fingerprint sensors</td>
<td>Owner of the fingerprint, fingerprint classes</td>
</tr>
<tr>
<td>Signature recognition (off-line, on-line)</td>
<td>Signature</td>
<td>Financial transactions</td>
</tr>
</tbody>
</table>
License Plate Reading System

- Detect and read the license plates

- Modules: (i) acquisition, (ii) enhancement, (iii) segmentation, (iv) character recognition

- Accuracy, robustness & real-time
Processing Steps

- Plate localization: Isolate the plate in image
- Preprocessing: Plate orientation and sizing;
- Normalization: Adjust image brightness & contrast
- Segmentation: Find individual characters
- Character recognition: OCR
- Post-processing: Rules for character placement
Challenges

- Poor image resolution: plate too far; low-res. image
- Motion blur
- Low contrast: overexposure, reflection/shadows
- Viewpoint variation and occlusion
- Different fonts, background
Pattern Recognition System

- Challenges
  - Pattern representation
  - Pattern classification

- System design
  - System training or learning
  - System testing or evaluation
What facial features to use to account for the large intra-class variability?

Yes, Mr. Palmer, we here at Mega-Corp did try your "face recognition software," and frankly, it didn't work worth a darn...
ARE ALL THESE OBJECTS TREES? Even a young child can answer correctly; a conventional computer, however, has enormous difficulty in doing so. Although there is a fair amount of regularity among the trees shown (each has a trunk and branches, for example), there is also a major component of arboreal irregularity among them. A generalized definition of a tree based on the underlying regularity could lead to erroneous identifications (such as mistaking a telephone pole, which has a “trunk” and “branches,” for a tree). Hence any effective program designed to recognize trees would essentially have to be a list of all types of trees, which cannot be done in a few lines of computer code.
Representation: Desirable Properties

- Invariance
- Account for intra-class variations
- Ability to discriminate classes of interest; low inter-class similarity
- Robustness to noise, occlusion,..
- Provide simple decision making strategies
- Low measurement cost; real-time
Invariant to

- Translation
- Rotation
- Scale
- Skew
- Deformation
- Color

Not all invariant properties are needed for a given application
“Appleby believes he’s come up with a way to defeat facial recognition software.”
System Performance

- Error rate; confusion matrix, RoC
- Speed (throughput)
- Cost
- Robustness
- Reject option
- Return on Investment (RoI)
Reject Option

What if the system encounters a previously unseen class?
Fruit Sorter

Statistical Pattern Recognition, Anil K. Jain

Jain CSE 802, Spring 2017
A test of 145 items, every item on the shelf, was performed, using all produce items available in a supermarket. Ten images per item was collected for a total of 1,450 images. Leave-one-out method for evaluation was used. For color & texture features combined, 84% of the time, correct produce class was selected, 96% of the time, the correct class was present in the top four choices.
FIGURE 1.7. Many pattern recognition systems can be partitioned into components such as the ones shown here. A sensor converts images or sounds or other physical inputs into signal data. The segmentor isolates sensed objects from the background or from other objects. A feature extractor measures object properties that are useful for classification. The classifier uses these features to assign the sensed object to a category. Finally, a post processor can take account of other considerations, such as the effects of context and the costs of errors, to decide on the appropriate action. Although this description stresses a one-way or “bottom-up” flow of data, some systems employ feedback from higher levels back down to lower levels (gray arrows). From: Richard O. Duda, Peter E. Hart, and David G. Stork, Pattern Classification. Copyright © 2001 by John Wiley & Sons, Inc.
Fish Classification: Salmon v. Sea Bass

- Preprocessing (enhancement, segmentation)
- Separate touching or occluding fishes
- Extract fish contour

**Figure 1.1.** The objects to be classified are first sensed by a transducer (camera), whose signals are preprocessed. Next the features are extracted and finally the classification is emitted, here either “salmon” or “sea bass.” Although the information flow is often chosen to be from the source to the classifier, some systems employ information flow in which earlier levels of processing can be altered based on the tentative or preliminary response in later levels (gray arrows). Yet others combine two or more stages into a unified step, such as simultaneous segmentation and feature extraction. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.
Fish Sorting

Cut out each of the fish cards on this page, then follow your teacher’s instructions for sorting the fish into categories. After you have compared your classification system with your classmates, follow the steps in the fish key below to identify the names of the fish.


Jain CSE 82, Spring 2017
Fish Key (Rule-based System, Decision Tree)

Step 1
If fish shape is long and skinny...
then go to Step 2
If fish shape is not long and skinny...
then go to step 3

Step 2
If the fish has pointed fins, it is a **trumpet fish**
If the fish has smooth fins, it is a **spotted moray eel**

Step 3
If fish has both eyes on top of the head... then go to step 4
If fish has one eye on each side of the head... then go to step 5

Step 4
If the fish has long whip-like tail, it is a **spotted eagle ray**
If the fish has short, blunt tail, it is a **peacock flounder**

Step 5
If fish has spots... then go to step 6
If fish does not have spots... then go to step 7

Step 6
If fish has chin “whiskers,” it is a **spotted goat fish**
If fish does not have chin “whiskers,” it is a **band-tail puffer**

Step 7
If fish has stripes... then go to step 8
If fish does not have stripes, it is a **glassy sweeper**

Step 8
If fish has a v-shaped tail, it is a **squirrel fish**
If fish has a blunt tail, it is a **glass-eye snapper**

Jain CSE 802, Spring 2017
WP 5/6 Development of weight estimating and species sorting systems for wild fish

FoodScanner Mini – sorting (species, weight, quality)
FIGURE 1.2. Histograms for the length feature for the two categories. No single threshold value of the length will serve to unambiguously discriminate between the two categories; using length alone, we will have some errors. The value marked $l^*$ will lead to the smallest number of errors, on average. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.
Fish Lightness as a Feature

Overlap of these histograms is small compared to length feature

FIGURE 1.3. Histograms for the lightness feature for the two categories. No single threshold value $x^*$ (decision boundary) will serve to unambiguously discriminate between the two categories; using lightness alone, we will have some errors. The value $x^*$ marked will lead to the smallest number of errors, on average. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.
Two-dimensional Feature Space

Joint distribution of two features leads to better separation

FIGURE 1.4. The two features of lightness and width for sea bass and salmon. The dark line could serve as a decision boundary of our classifier. Overall classification error on the data shown is lower than if we use only one feature as in Fig. 1.3, but there will still be some errors. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

Linear (simple) decision boundary; linear classifier
Complex Decision Boundary
(Polynomial classifier)

What is the generalization ability of the classifier?

Jain CSE 802, Spring 2017
Good Generalization & Good Accuracy

Simple decision boundaries are preferred

**FIGURE 1.6.** The decision boundary shown might represent the optimal tradeoff between performance on the training set and simplicity of classifier, thereby giving the highest accuracy on new patterns. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.
Occam’s Razor (William of Ockham (c. 1287–1347))

"If you have two equally likely solutions to a problem, choose the simplest"
Feature Selection & Extraction

• Feature selection: which subset to use? Some features may be redundant

• Feature extraction: which combination of given features to use?

• Curse of dimensionality—Error rate may in fact increase with too many features in the case of small number of training samples
Pattern Recognition Models

- Template matching
  - Class-specific shape & appearance models
- Statistical (geometric)
  - Class-specific Prob. density function (pdf)
- Syntactic (structural)
  - Class-specific grammar
- Neural networks
Rigid Template Matching

Template
Prototype

Input scene
Deformable Template

Corpus callosum shape training set

Prototype and variation learning

Prototype registration to the low-level segmented image

Prototype warping
Statistical Pattern Recognition

Preprocessing ➔ Feature extraction ➔ Classification

Recognition

Training

Preprocessing ➔ Feature selection ➔ Learning

Patterns + Class labels

pattern ➔ Preprocessing
Each pattern is represented as a point in \textit{d-}
dimensional feature space

Choice of features and their desired invariance
properties are \textit{domain-specific}

Good representation implies (i) small intra-class
variation, (ii) large interclass separation and (iii)
simple decision boundary
FIGURE 1.2. Histograms for the length feature for the two categories. No single threshold value of the length will serve to unambiguously discriminate between the two categories; using length alone, we will have some errors. The value marked $l^*$ will lead to the smallest number of errors, on average. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.
**FIGURE 1.3.** Histograms for the lightness feature for the two categories. No single threshold value $x^*$ (decision boundary) will serve to unambiguously discriminate between the two categories; using lightness alone, we will have some errors. The value $x^*$ marked will lead to the smallest number of errors, on average. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.
FIGURE 1.4. The two features of lightness and width for sea bass and salmon. The dark line could serve as a decision boundary of our classifier. Overall classification error on the data shown is lower than if we use only one feature as in Fig. 1.3, but there will still be some errors. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.
Instead of describing an object in terms of a feature vector, describe it by its structure.

Complex objects are represented in terms of simple primitives (shapes) and their relationship; parts-based representation (represent face as eyes, mouth, nose,...)
Syntactic Pattern Recognition

Recognition

- Preprocessing
- Primitives & relationship extraction
- Syntax analysis

Training

- Preprocessing
- Primitive selection
- Grammatical inference

- Pattern
- Patterns + Class labels

Jain CSE 802, Spring 2017
Chromosome Grammars

- **Terminals:**
  \[ V_T = \{\cap, |, \cup, \}, , \bowtie \} \]

- **Non-terminals:**
  \[ V_N = \{A, B, C, D, E, F\} \]

- **Pattern Classes:**
  - Median
  - Submedian
  - Acrocentric
  - Telocentric
Chromosome Grammars

Image of human chromosomes

Hierarchical-structure description of a submedium chromosome

Jain CSE 802, Spring 2017
Neural Networks

- Massive parallelism essential for complex recognition tasks (speech & image recognition)
  - Humans take only ~200ms. for most cognitive tasks; this suggests parallel computation in human brain

- Biological networks achieve excellent recognition performance via dense interconnection of simple computational elements (neurons)
  - Number of neurons $\approx 10^{10} - 10^{12}$
  - Number of interconnections/neuron $\approx 10^3 - 10^4$
  - Total number of interconnections $\approx 10^{14}$
Nodes are **nonlinear**, typically analog

\[ Y = f \left( \sum_{i=1}^{d} w_i x_i - \theta \right) \]

where \( \theta \) is internal threshold or offset
Feed-forward networks with one or more layers (hidden) between input & output nodes

How many nodes & hidden layers?

Network training
How much information are you missing

Qvest
Constraining the Recognition Problem

- Graffiti alphabet

GRAFFITI’S MODIFIED alphabet is largely based on single pen strokes, starting at the dots. As soon as the pen is lifted from the screen, the letter is immediately translated into normal text. The letter “X” is the exception
Super Classifier

Pool the evidence from component recognizers; also known as classifier combination, mixture of experts, evidence accumulation
Pattern recognition

- Automatic decision making
- Assist human decision makers

General-purpose PR is difficult; systems available for constrained domains: mugshot face recognition

No single recognition approach is optimal for all PR problems; toolbox of classifiers

Use of object models, constraints and context helps

Careful sensor design and feature extraction lead to simple classifiers
Same Problems, Many Approaches

Recognition, Classification, Clustering, Regression

- Fisher Linear Discriminant (1936)
- Perceptron, Rumelhart (1958)
- Adaptive multilayer networks, Widrow (1960s)
- Backpropagation learning algorithm, Werbos (1974)
- Artificial Intelligence (AI), McCarthy (1956)
- Pattern recognition
- Artificial neural networks
- Data mining
- Machine learning
- Knowledge discovery, expert systems
- Deep networks, convolution neural networks
Key Concepts

- Pattern class
- Representation, feature set
- Feature selection
- Feature extraction
- Linear transformation (PCA, LDA)
- Feature invariance
- Preprocessing
- Segmentation
- Training set
- Validation set
- Test set
- Error rate
- Reject rate
- Curse of dimensionality
Key Concepts

- Supervised learning
- Decision boundary
- Classifier
- unsupervised learning
- Clustering
- Density Estimation
- Cost of misclassification/Risk
- Feature space partitioning
- Generalization (overfitting)
- Contextual information
- Combination of classifiers
- Prior knowledge