Pattern Recognition

CSE 802

Michigan State University

Spring 2017

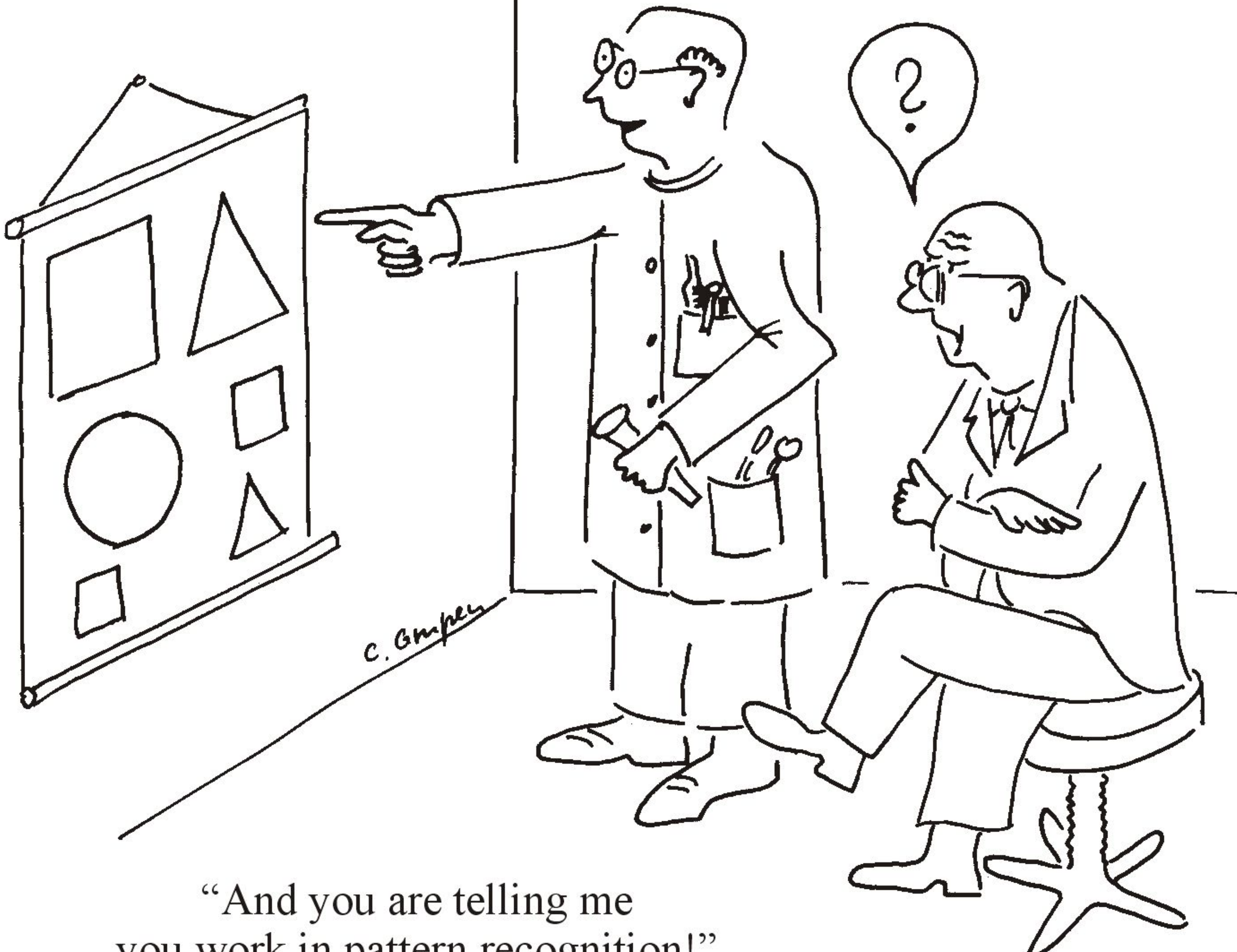
Lecture 1, January 9, 2017

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*



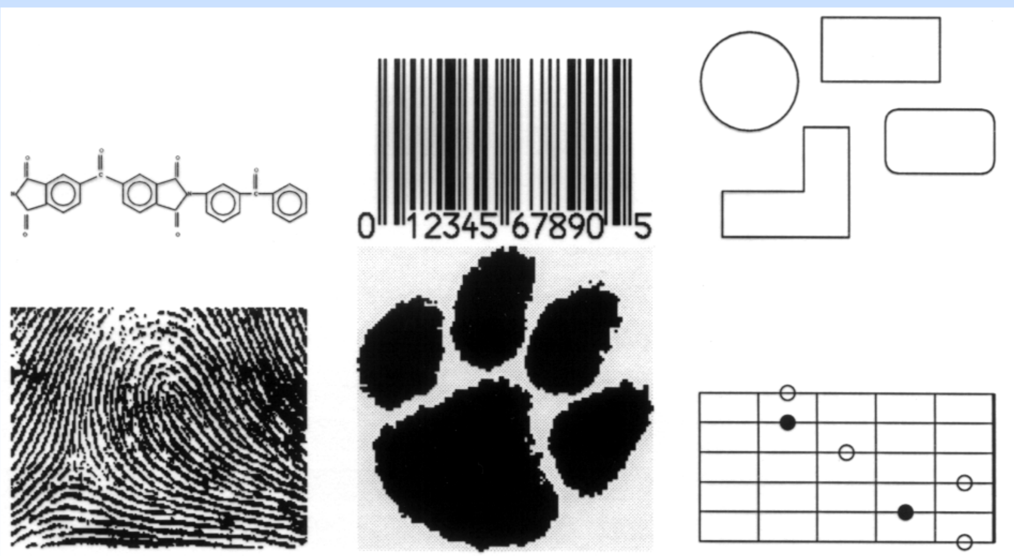
"And you are telling me
you work in pattern recognition!"

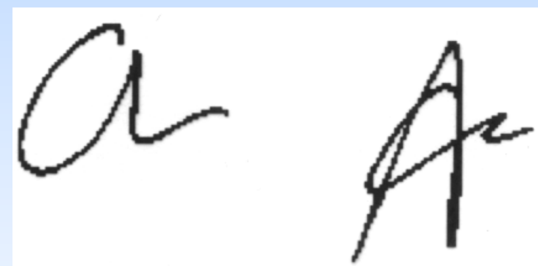
Pattern Recognition

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)

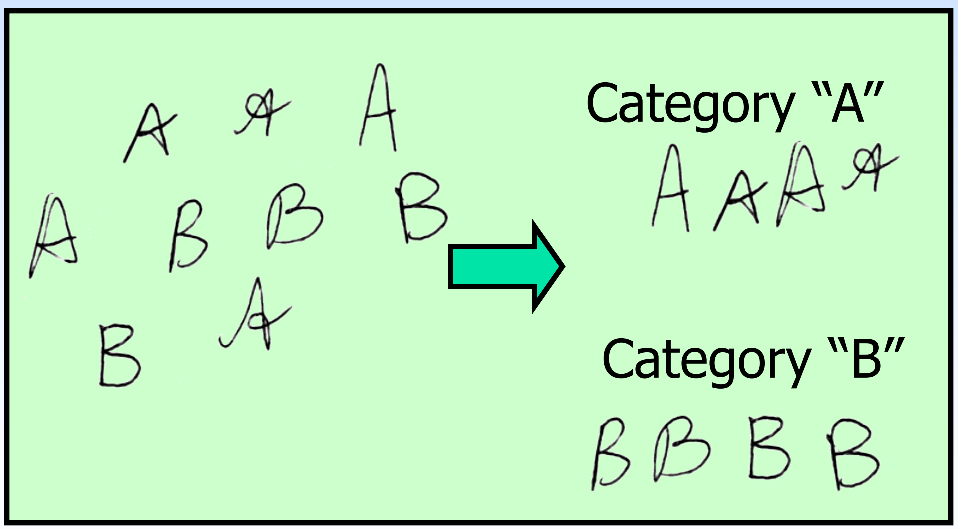




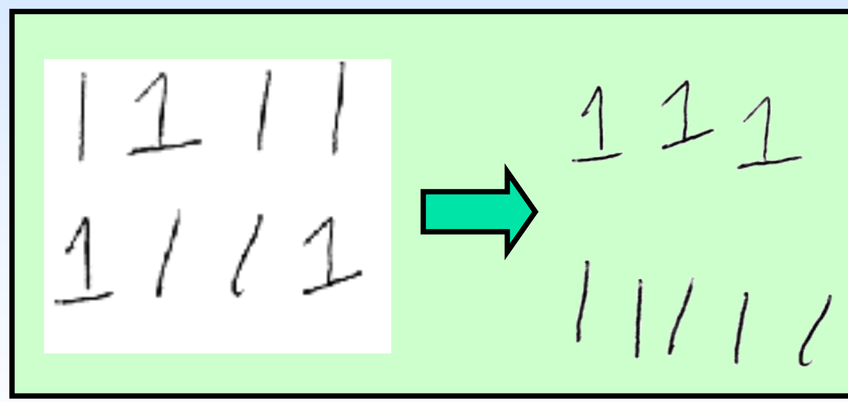
Recognition

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)



Classification



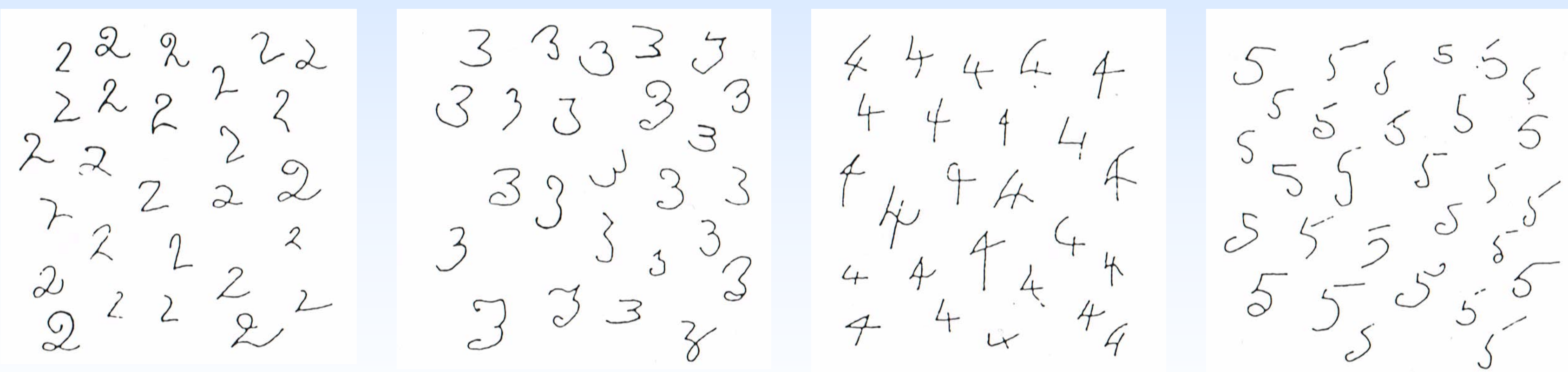
Clustering

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?

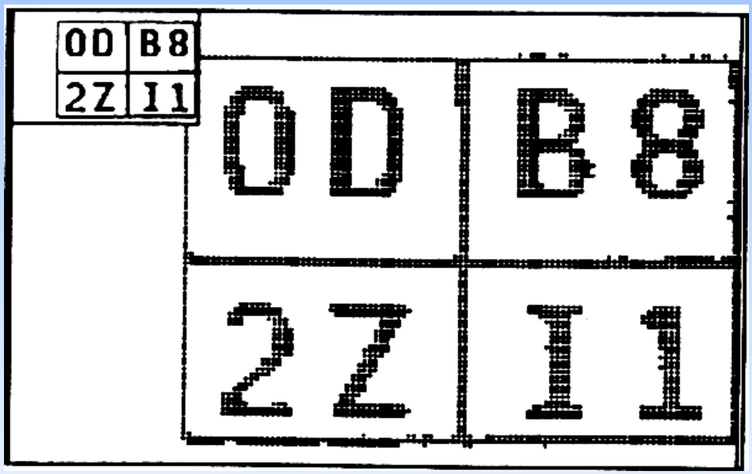
Intra-Class Variability

Handwritten numerals



Inter-class Similarity

Identical twins



Characters that look similar

Cat vs. Dog: 2-class Classification

(Supervised) Classification

Labeled training samples for classifier design

Clustering: Unsupervised Classification

Training samples are unlabeled

Cats and Dogs

Omkar M Parkhi[1,2] Andrea Vedaldi[1] Andrew Zisserman[1] C. V. Jawahar[2]

[1]Department of Engineering Science,
University of Oxford,
United Kingdom
{omkar,vedaldi,az}@robots.ox.ac.uk

[2]Center for Visual Information Technology,
International Institute of Information Technology,
Hyderabad, India
jawahar@iiit.ac.in

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)

Example Images of Cats & Dogs

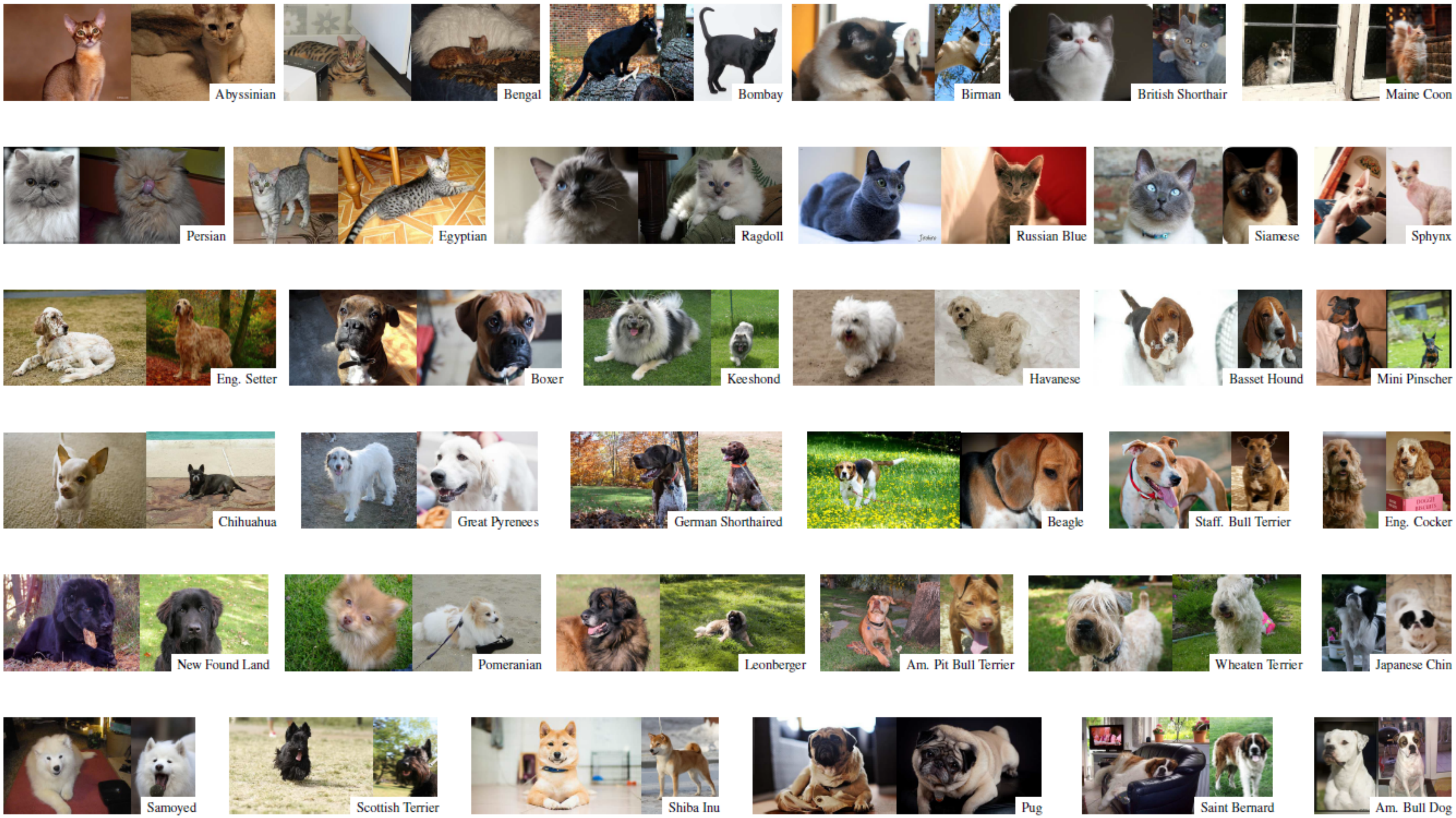


Figure 2. **Example images from the Oxford-IIIT Pet data.** Two images per breed are shown side by side to illustrate the data variability.

Segmentation: Foreground v. Background

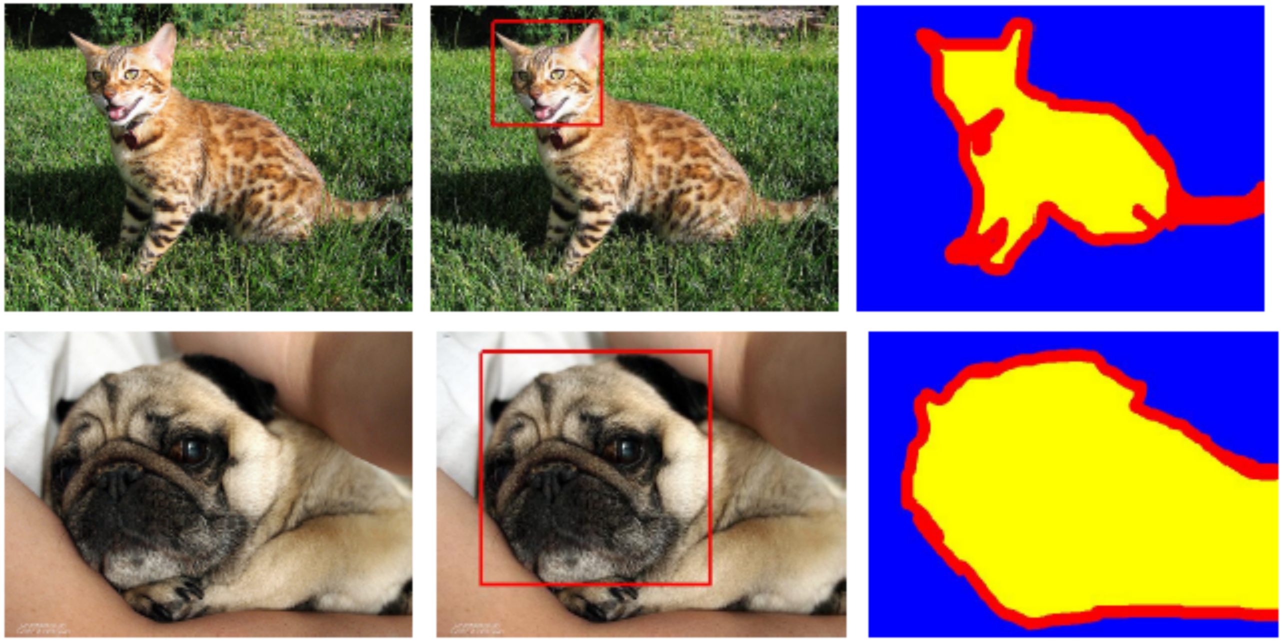


Figure 1. **Annotations in the Oxford-IIIT Pet data.** From left to right: pet image, head bounding box, and trimap segmentation (*blue*: background region; *red*: ambiguous region; *yellow*: foreground region).

Feature Extraction

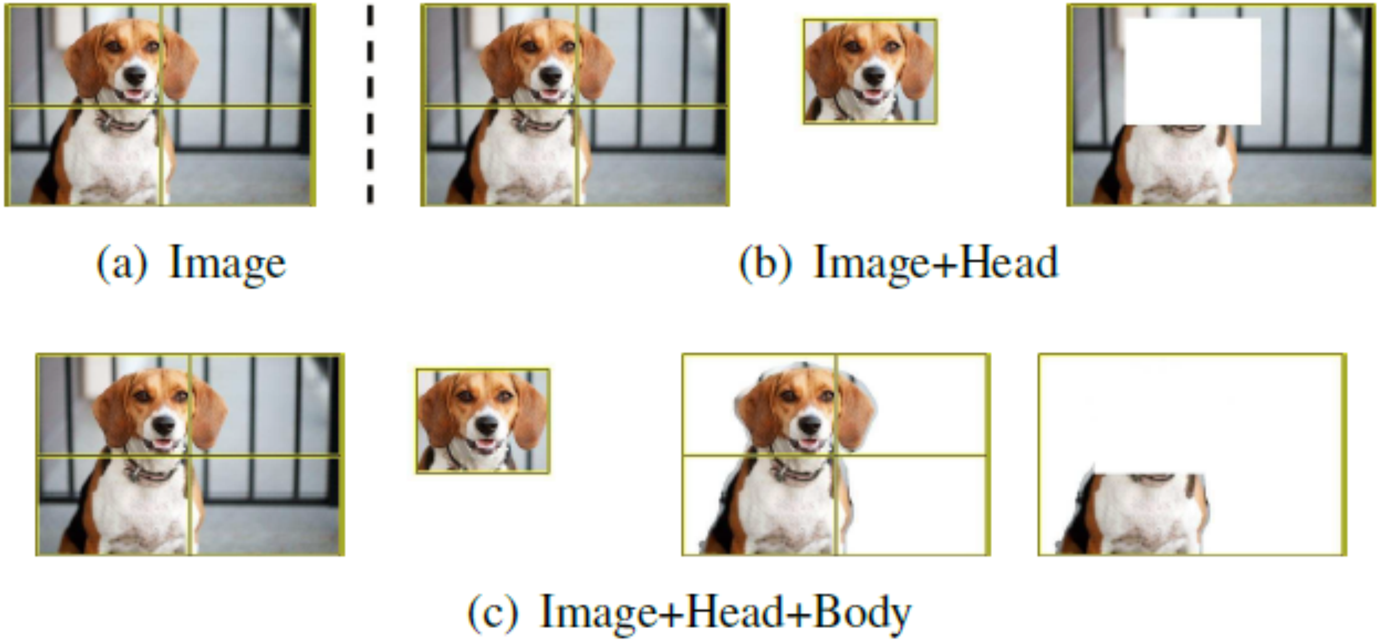


Figure 4. **Spatial histogram layouts.** The three different spatial layouts used for computing the image descriptors. The image descriptor in each case is formed by concatenating the histograms computed on the individual spatial components of the layout. The spatial bins are denoted by yellow-black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

Classification Performance

.	Shape	Appearance		Classification Accuracy (%)				
		layout type	using ground truth	family	breed (S. 4.2)		both (S. 4.3)	
				(S. 4.1)	cat	dog	hierarchical	flat
1	✓	–	–	94.21	NA	NA	NA	NA
2	–	Image	–	82.56	52.01	40.59	NA	39.64
3	–	Image+Head	–	85.06	60.37	52.10	NA	51.23
4	–	Image+Head+Body	–	87.78	64.27	54.31	NA	54.05
5	–	Image+Head+Body	✓	88.68	66.12	57.29	NA	56.60
6	✓	Image	–	94.88	50.27	42.94	42.29	43.30
7	✓	Image+Head	–	95.07	59.11	54.56	52.78	54.03
8	✓	Image+Head+Body	–	94.89	63.48	55.68	55.26	56.68
9	✓	Image+Head+Body	✓	95.37	66.07	59.18	57.77	59.21

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

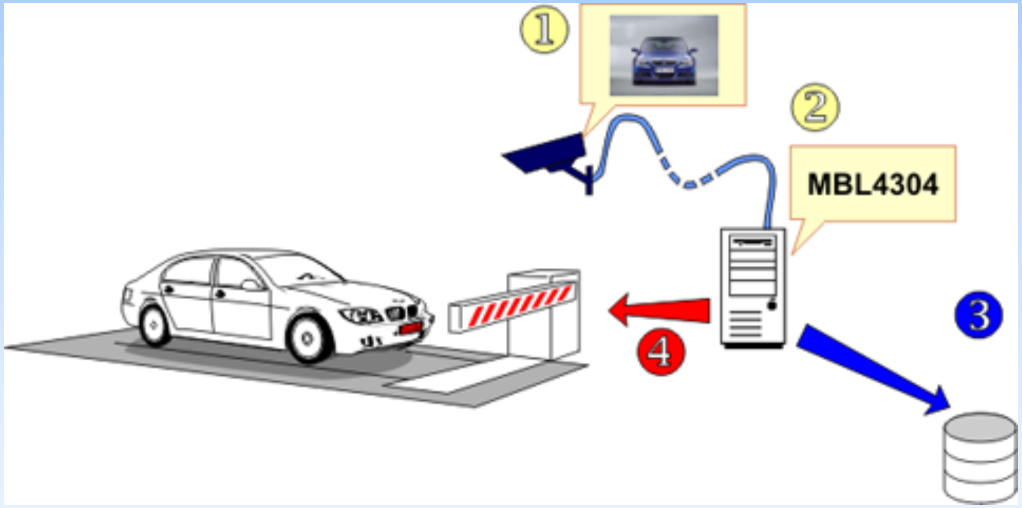
Problem	Input	Output
Speech recognition	Speech waveforms	Spoken words, speaker identity
Non-destructive testing	Ultrasound, eddy current, acoustic emission waveforms	Presence/absence of flaw, type of flaw
Medical waveform analysis	EKG, EEG waveforms	Types of cardiac conditions, classes of brain conditions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

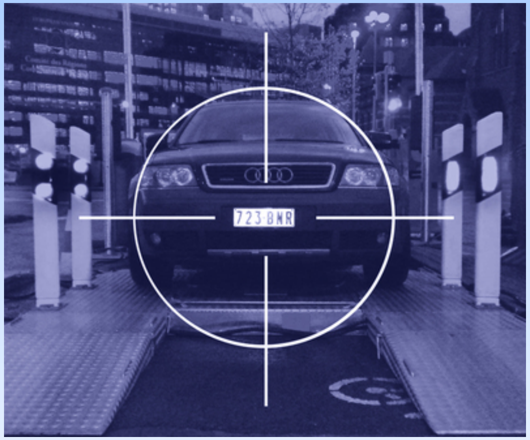
Pattern Recognition Applications

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

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

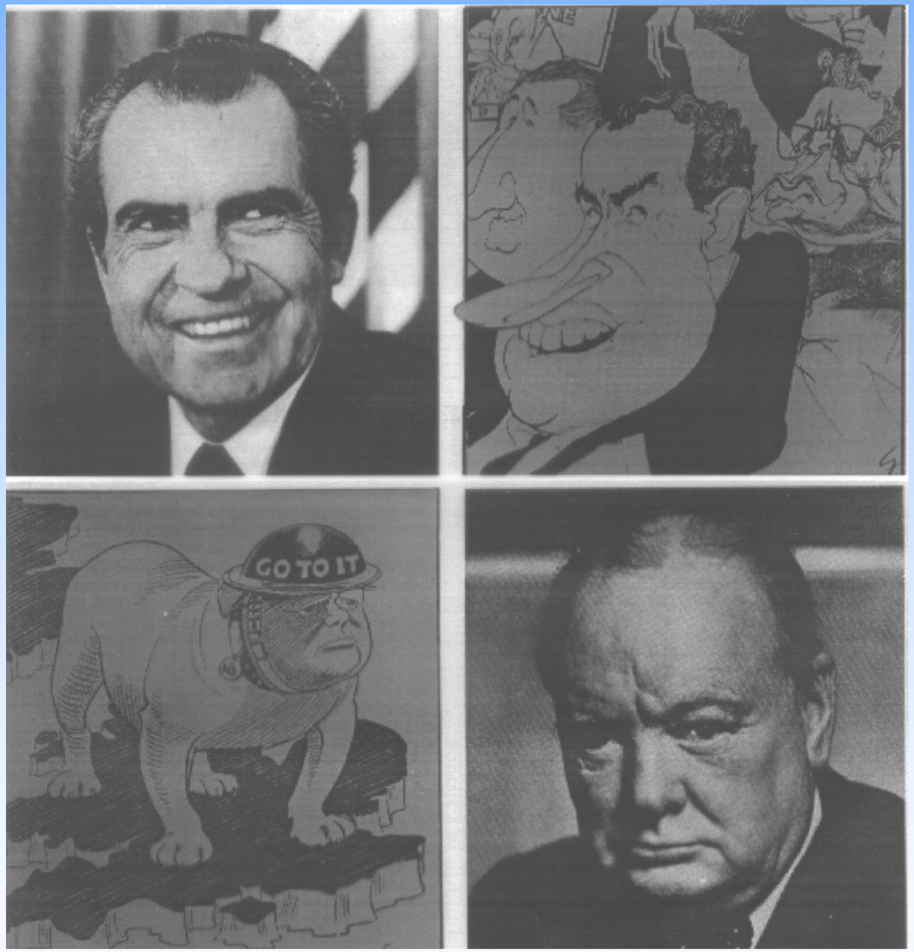
Challenges

- 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

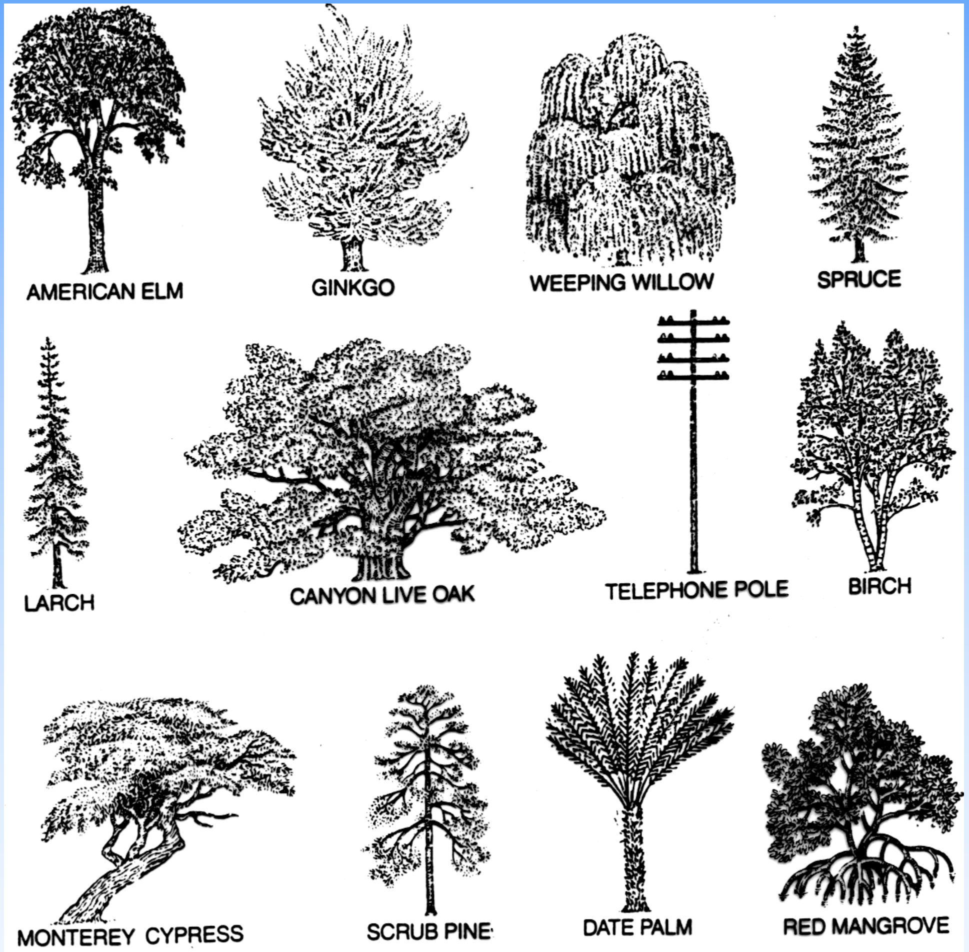
Representation



What facial features to use to account for the large intra-class variability?

John P. Frisby, *Seeing. Illusion, Brian and Mind*, Oxford University Press, 1980


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 Representation

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

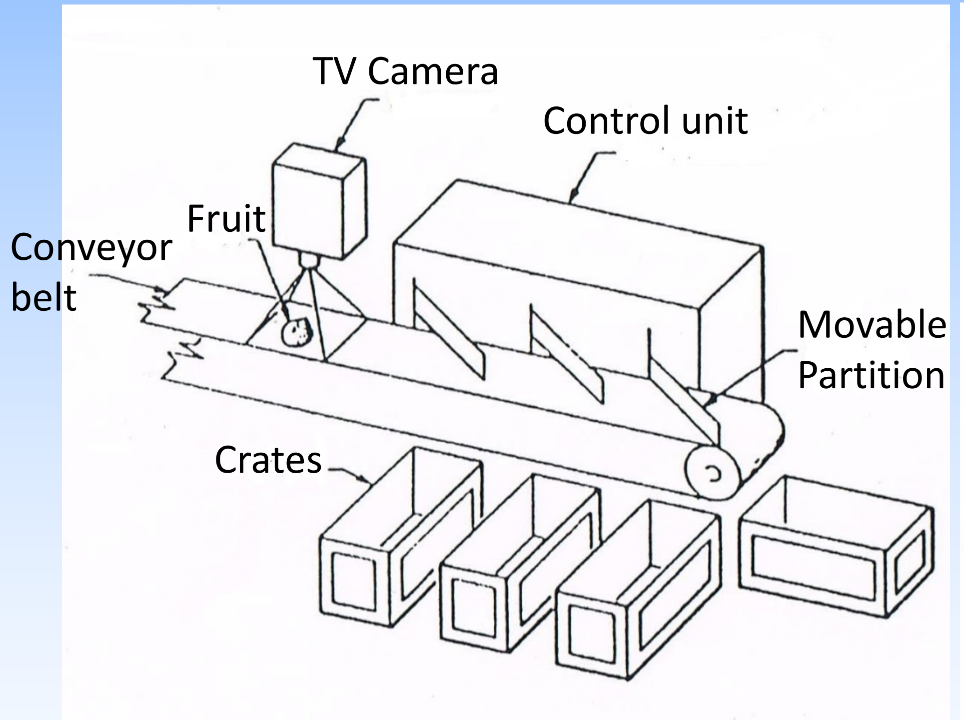
- Error rate; confusion matrix, RoC
- Speed (throughput)
- Cost
- Robustness
- Reject option
- Return on Investment (RoI)

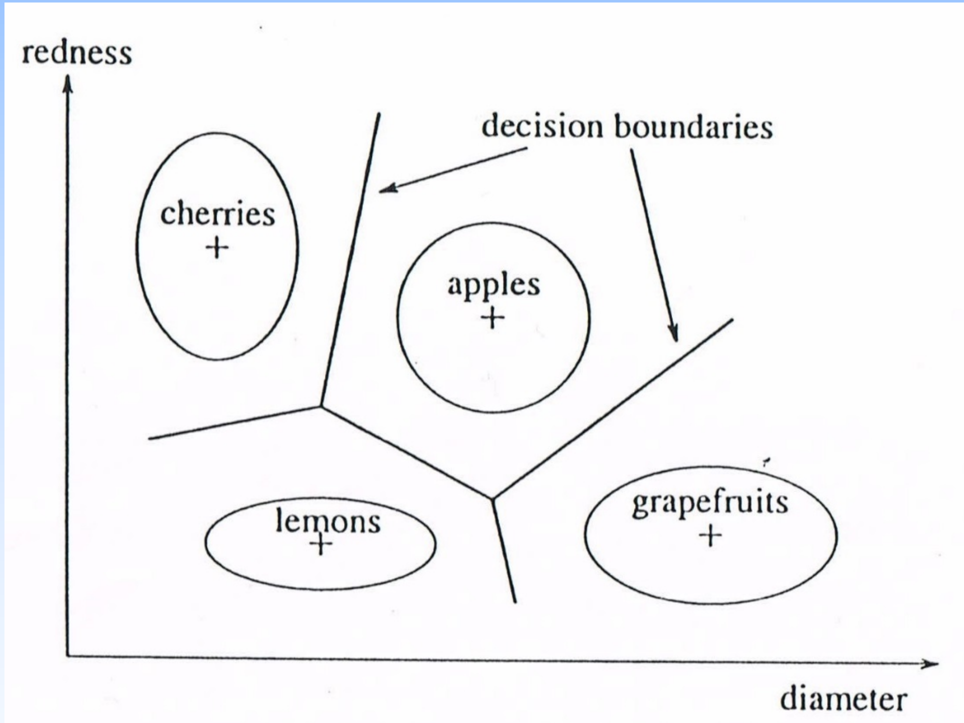
Reject Option

What if the system encounters a previously unseen class?

水

Fruit Sorter





Statistical Pattern Recognition, Anil K. Jain

Veggie Vision: A Produce Recognition System

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.

http://researcher.watson.ibm.com/researcher/files/us-smiyaza/jhc-waiat.pdf
http://researcher.watson.ibm.com/researcher/files/us-smiyaza/jhc-wacv.pdf

Pattern Recognition System

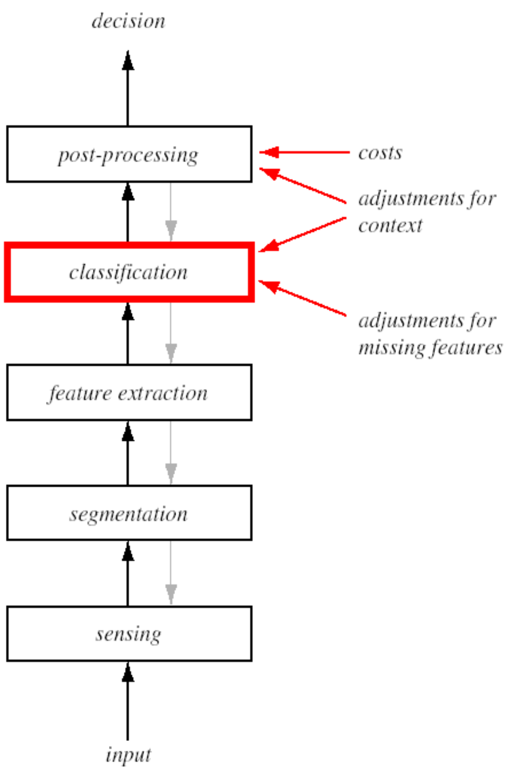


FIGURE 1.7. Many pattern recognition systems can be partitioned into components such as the ones shown here. A sensor converts images or sounds or other physical inputs into signal data. The segmentor isolates sensed objects from the background or from other objects. A feature extractor measures object 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

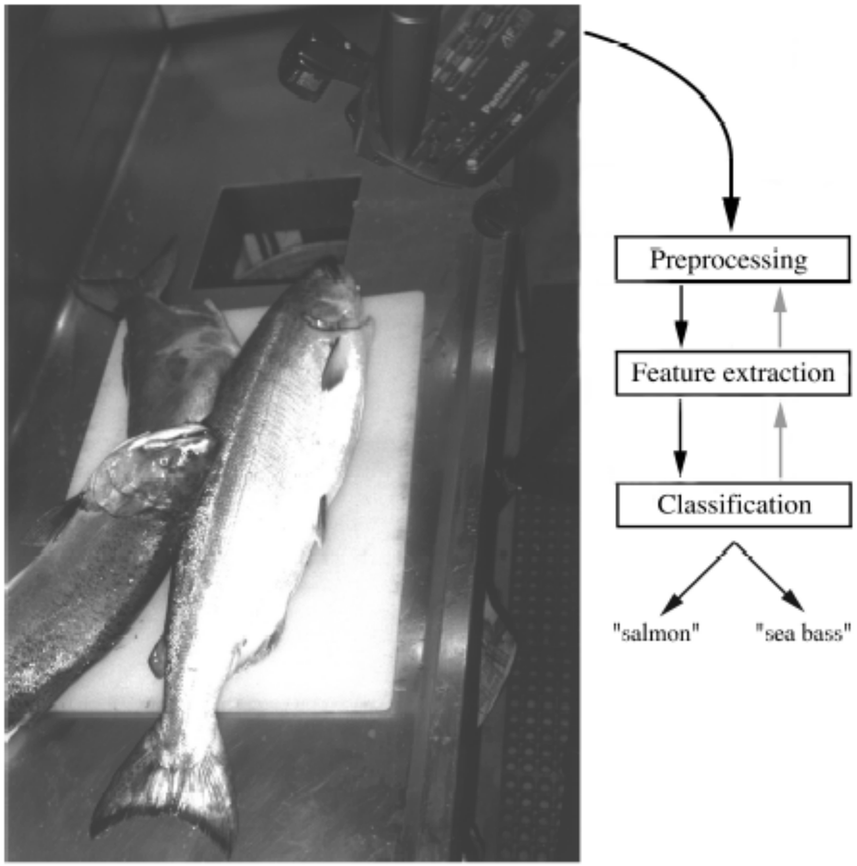
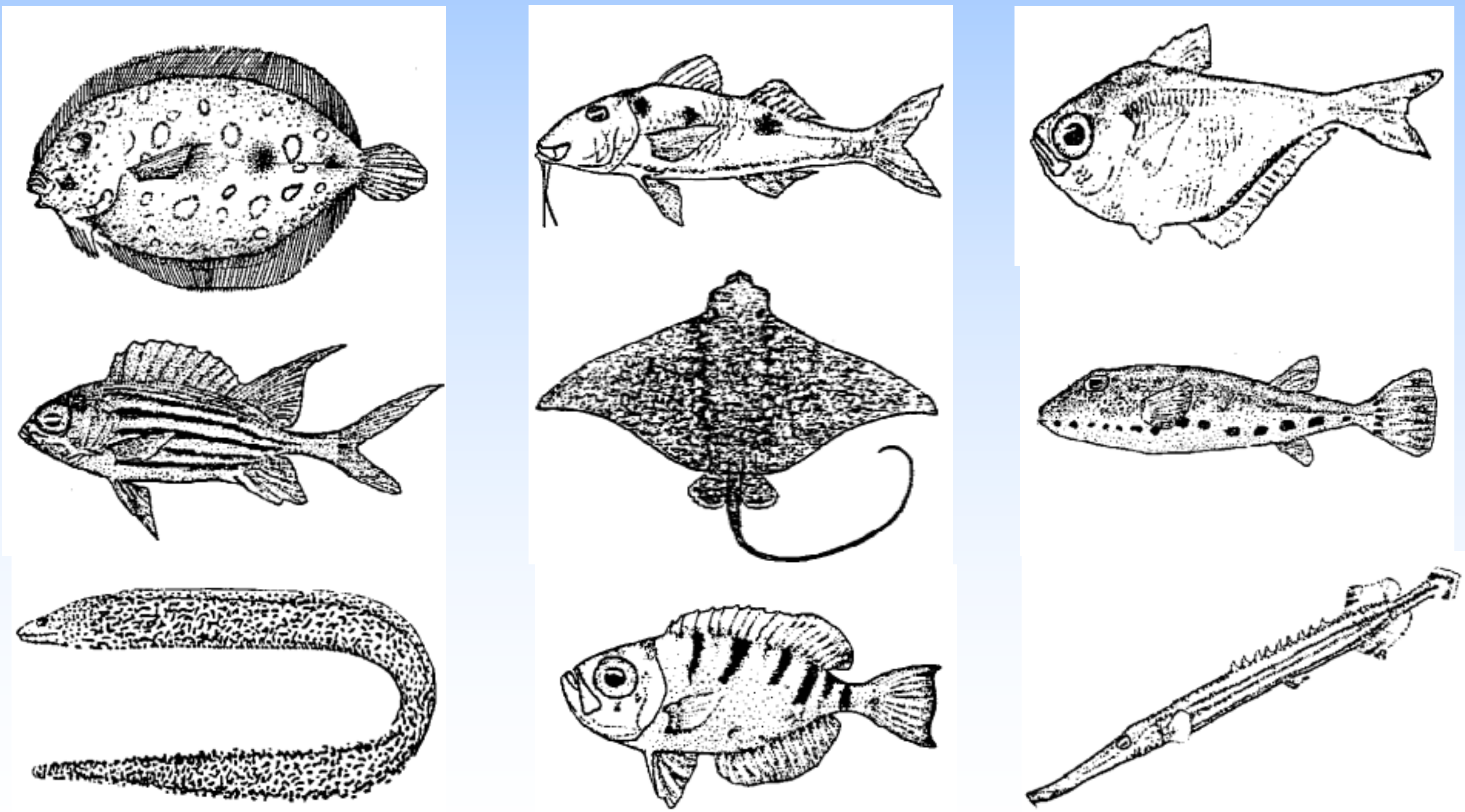


FIGURE 1.1. The objects to be classified are first sensed by a transducer (camera), whose signals are preprocessed. Next the features are extracted and finally the classification is emitted, here either "salmon" or "sea bass." Although the information flow is often chosen to be from the source to the 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

Treasures of the Great Barrier Reef NOVAactivity

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.



http://www-tc.pbs.org/wgbh/nova/education/activities/pdf/2215_reef.pdf

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**

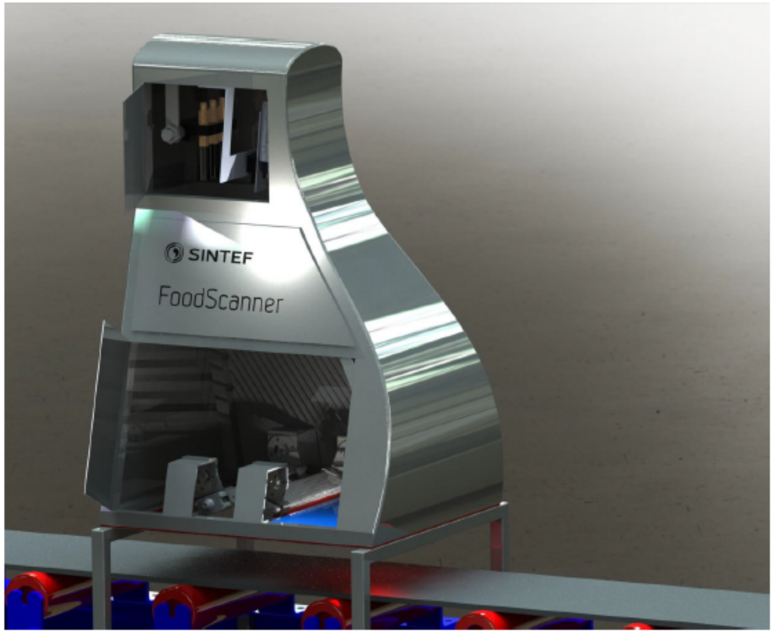
Step 7
If fish has stripes… then go to step 8

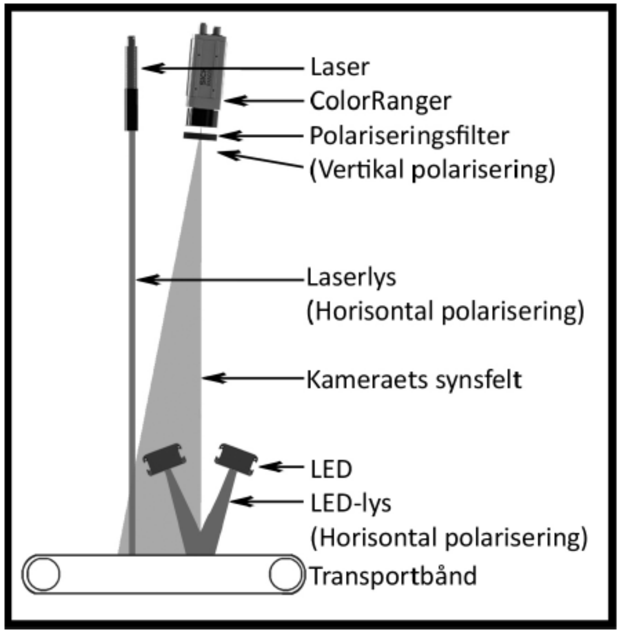
If fish does not have stripes, it is a **glassy sweeper**

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**

WP 5/6 Development of weight estimating and species sorting systems for wild fish





FoodScanner Mini – sorting (species, weight, quality)

Representation: Fish Length as a Feature

Training samples

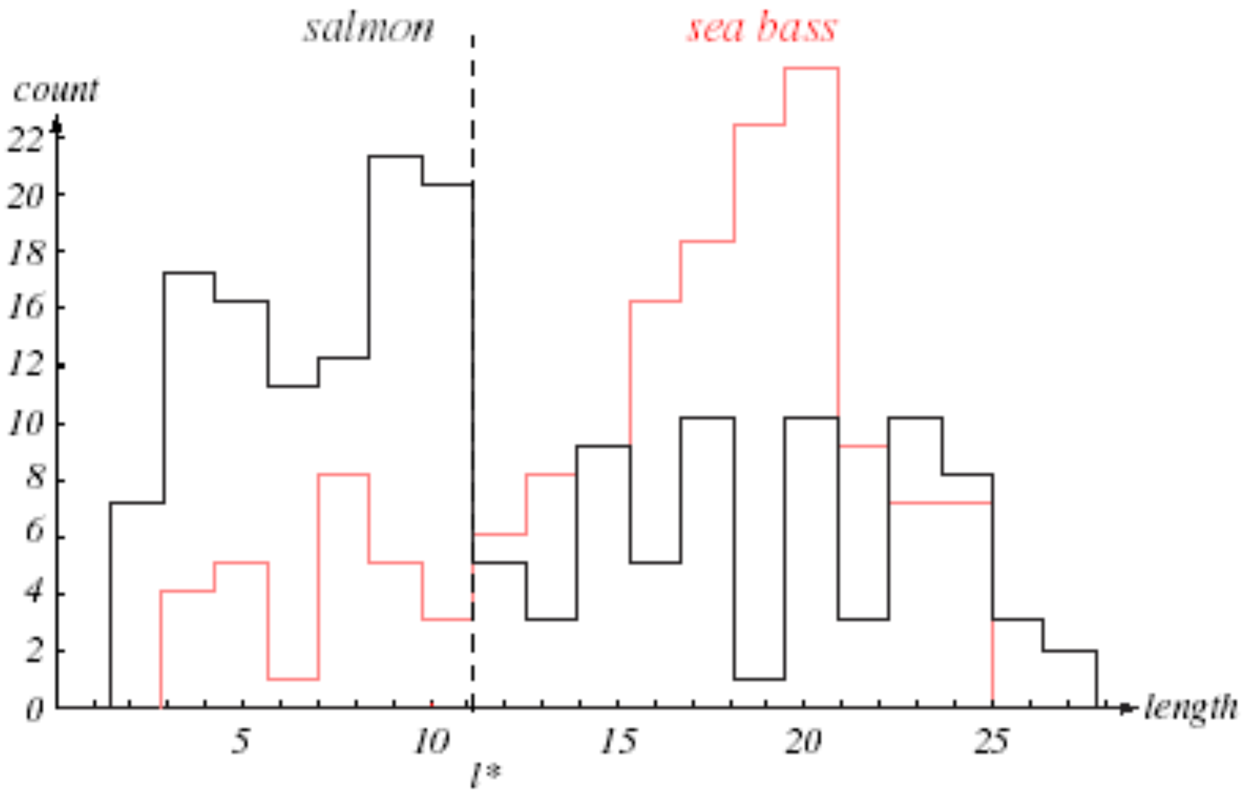


FIGURE 1.2. Histograms for the length feature for the two categories. No single threshold value of the length will serve to unambiguously discriminate between the two categories; using length alone, we will have some errors. The value marked l^* will lead to the smallest number of errors, on average. 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

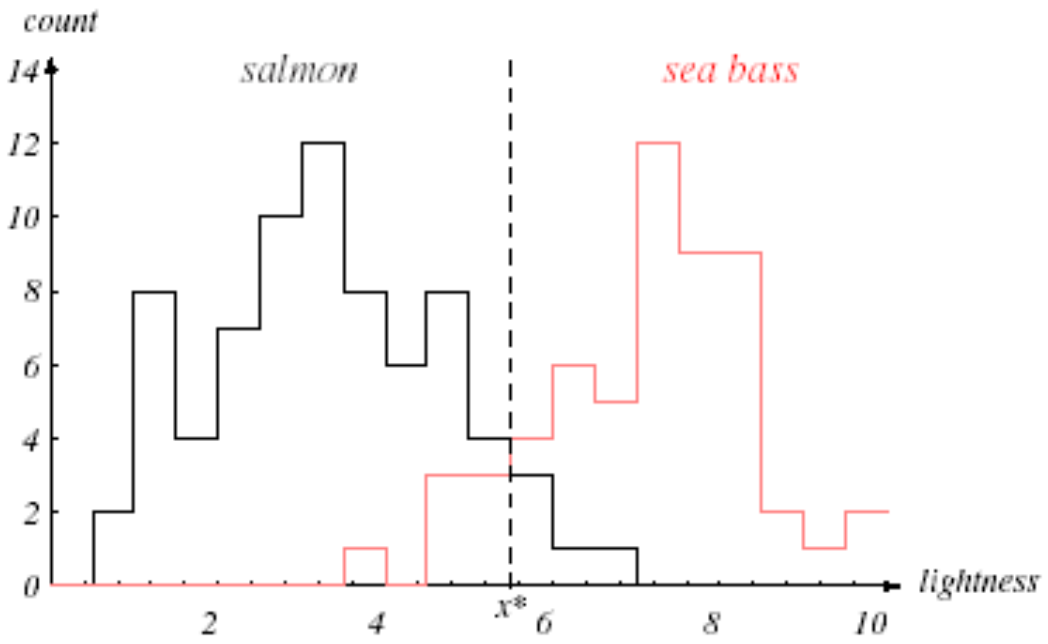


FIGURE 1.3. Histograms for the lightness feature for the two categories. No single threshold value x^* (decision boundary) will serve to unambiguously discriminate between the two categories; using lightness alone, we will have some errors. The value x^* marked will lead to the smallest number of errors,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

Overlap of these histograms is small compared to length feature

Two-dimensional Feature Space

Linear (simple) decision boundary; linear classifier

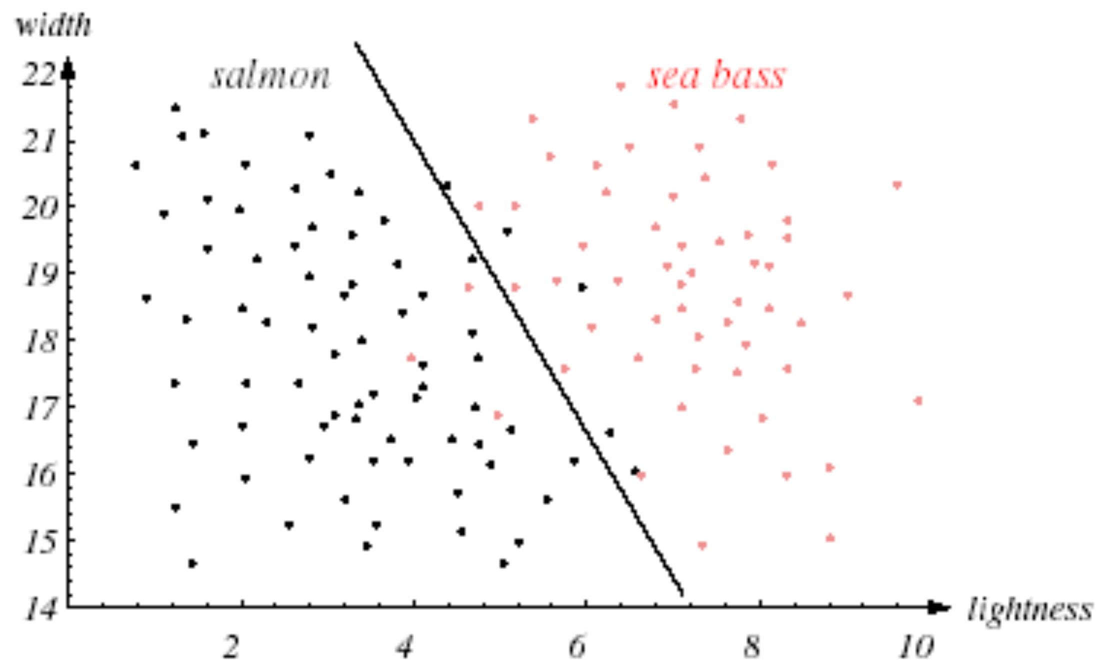


FIGURE 1.4. The two features of lightness and width for sea bass and salmon. The dark line could serve as a decision boundary of our classifier. Overall classification error on the data shown is lower than if we use only one feature as in Fig. 1.3, but there will still be some errors. From: Richard O.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

Joint distribution of two features leads to better separation

Complex Decision Boundary (Polynominal classifier)

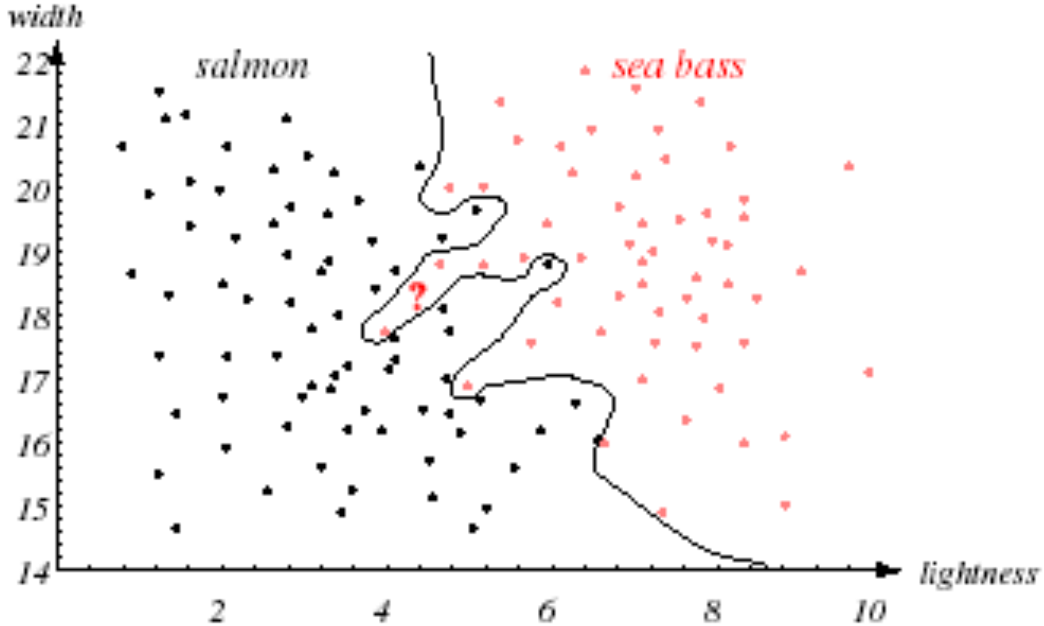


FIGURE 1.5. Overly complex models for the fish will lead to decision boundaries that are complicated. While such a decision may lead to perfect classification of our training samples, it would lead to poor performance on future patterns. The novel test point marked ? is evidently most likely a salmon, whereas the complex decision boundary shown leads it to be classified as a sea bass. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

What is the generalization ability of the classifier?

Good Generalization & Good Accuracy

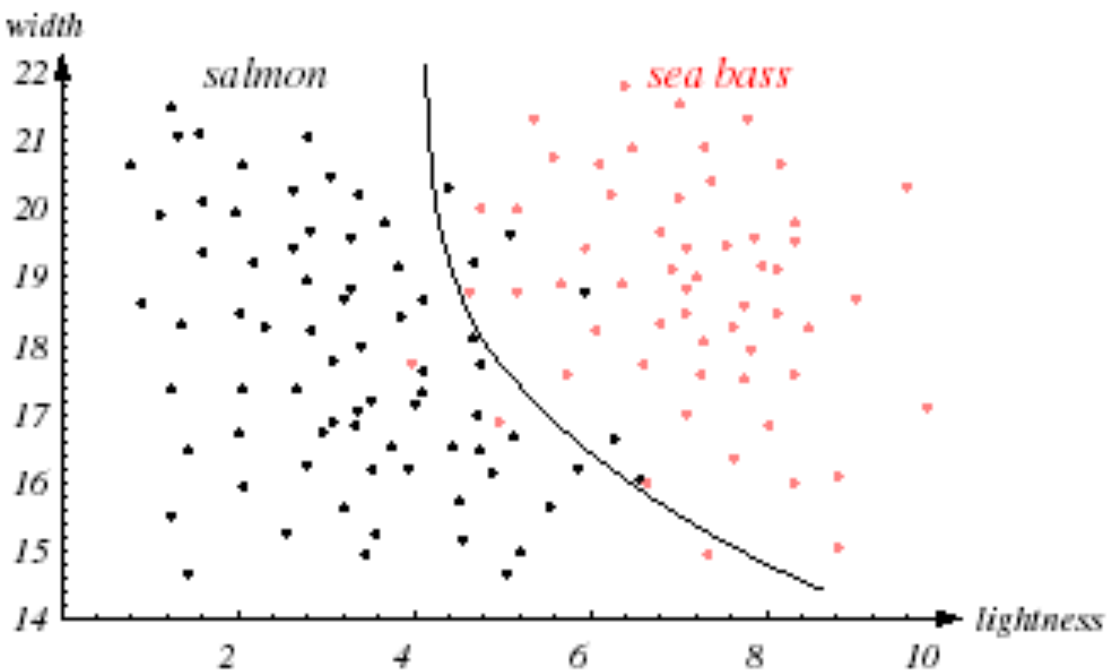


FIGURE 1.6. The decision boundary shown might represent the optimal tradeoff between performance on the training set and simplicity of classifier, thereby giving the highest accuracy on new patterns. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001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

Simple decision boundaries are preferred

Occam’s Razor (William of Ockham (c. 1287–1347)

"If you have two equally likely solutions to a problem, choose the simplest"

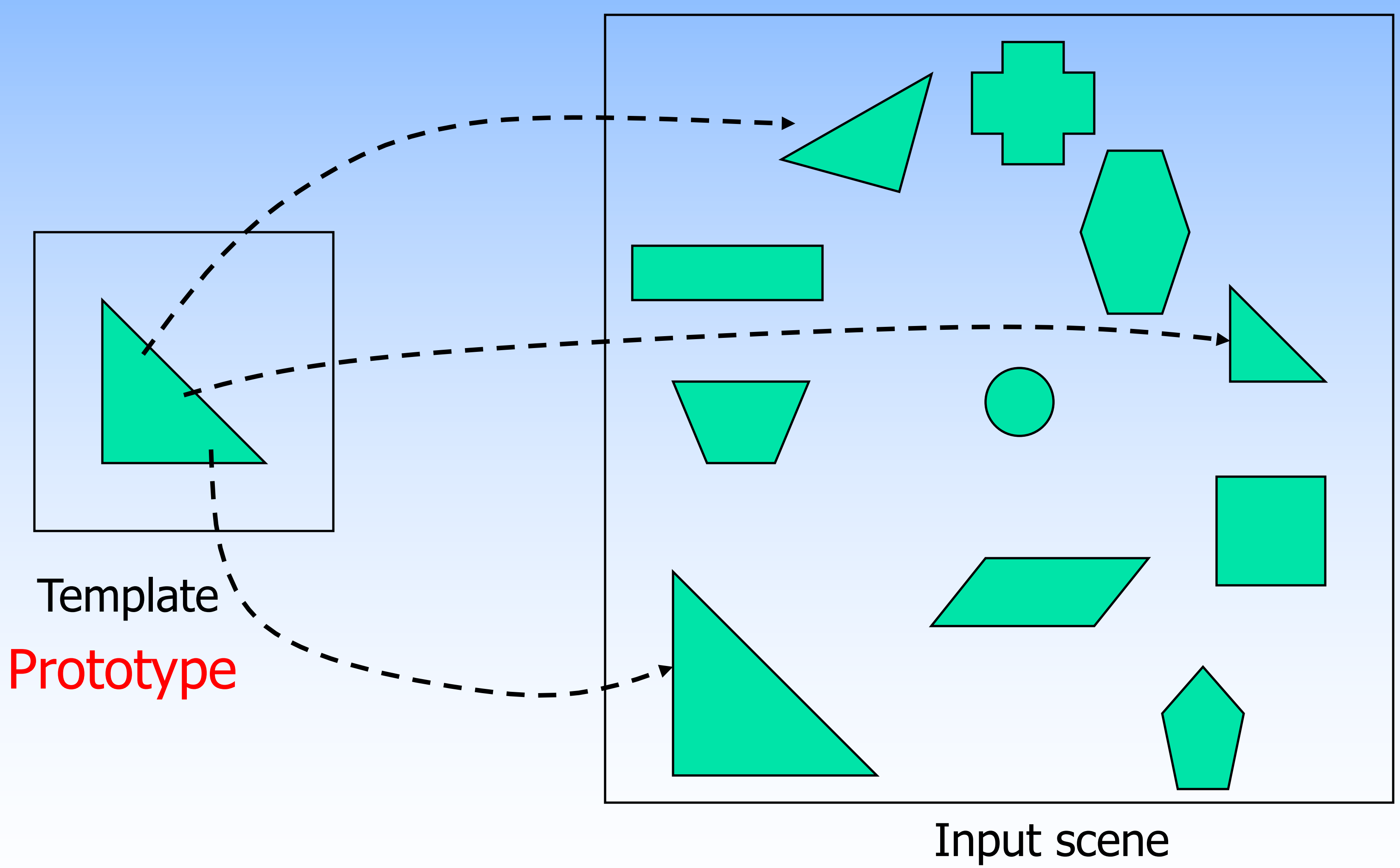
Feature Selection & Extraction

- 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

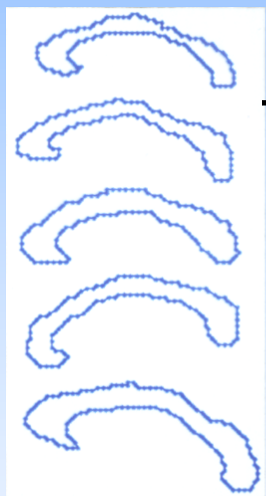
- Template matching
  - Class-specific shape & appearance models
- Statistical (geometric)
  - Class-specific Prob. density function (pdf)
- Syntactic (structural)
  - Class-specific grammar
- Neural networks

Rigid Template Matching

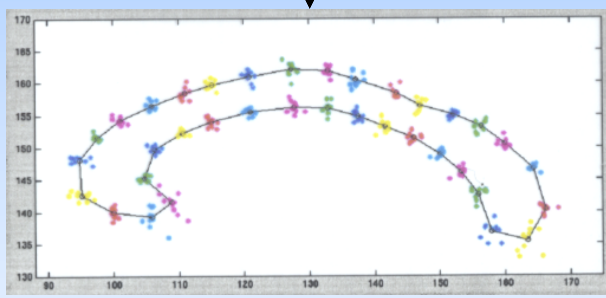


Deformable Template

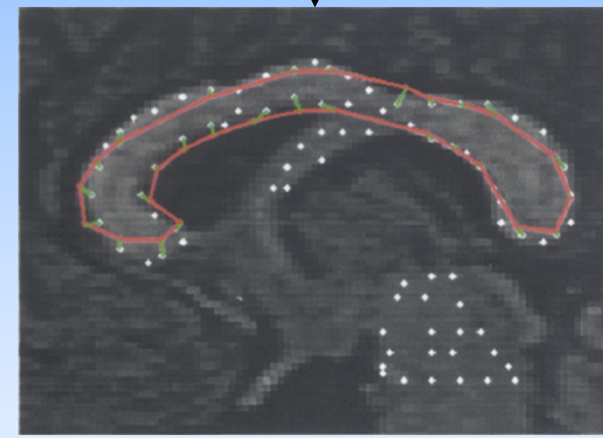
Corpus callosum shape training set



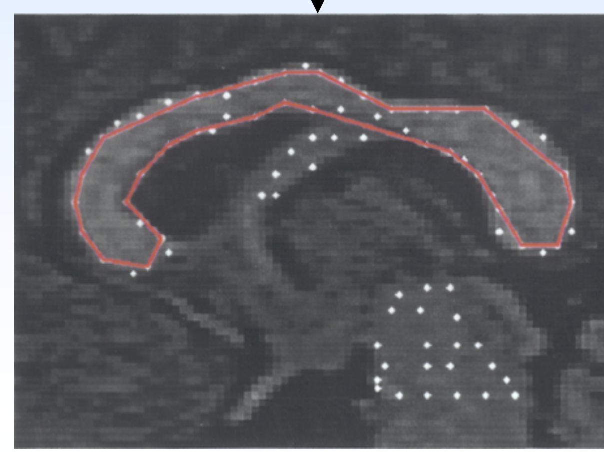
Prototype and variation learning

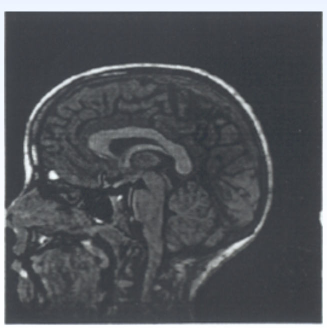


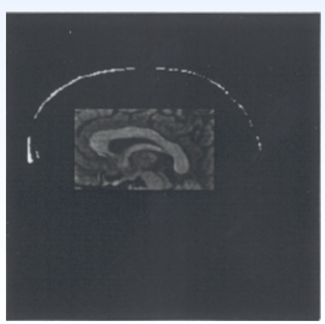
Prototype registration to the low-level segmented image



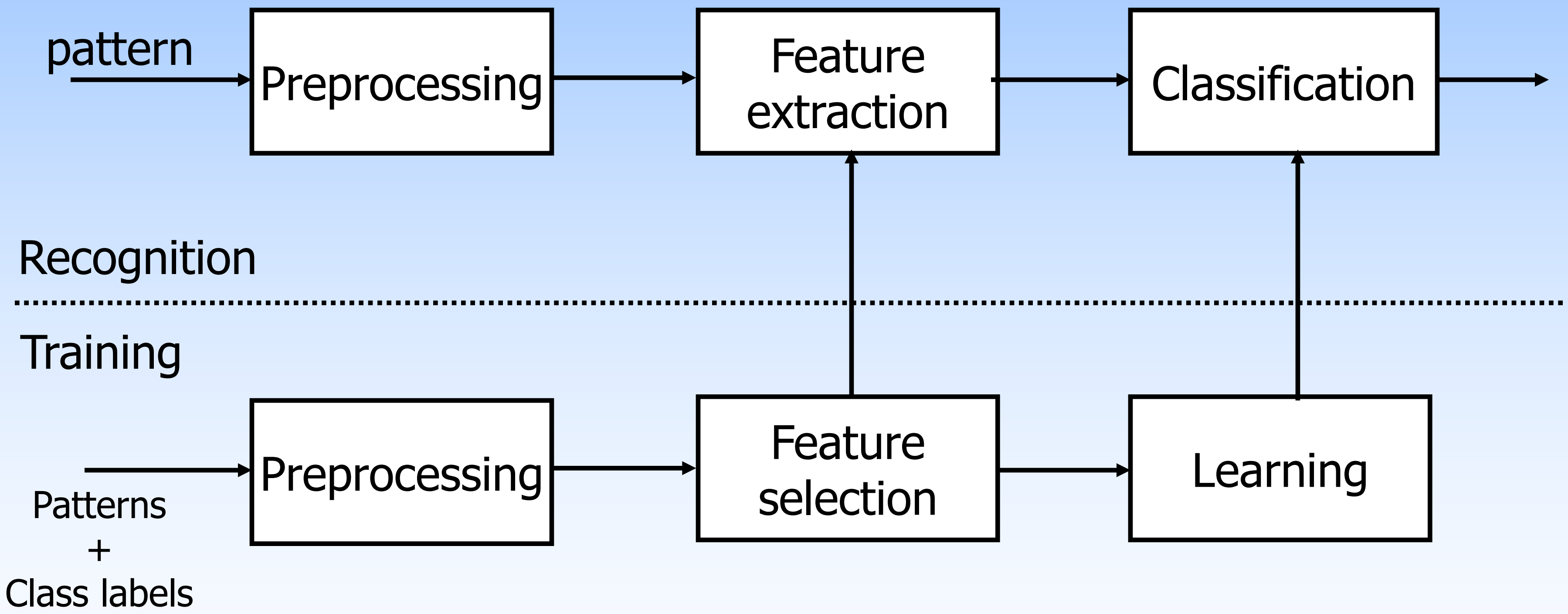
Prototype warping





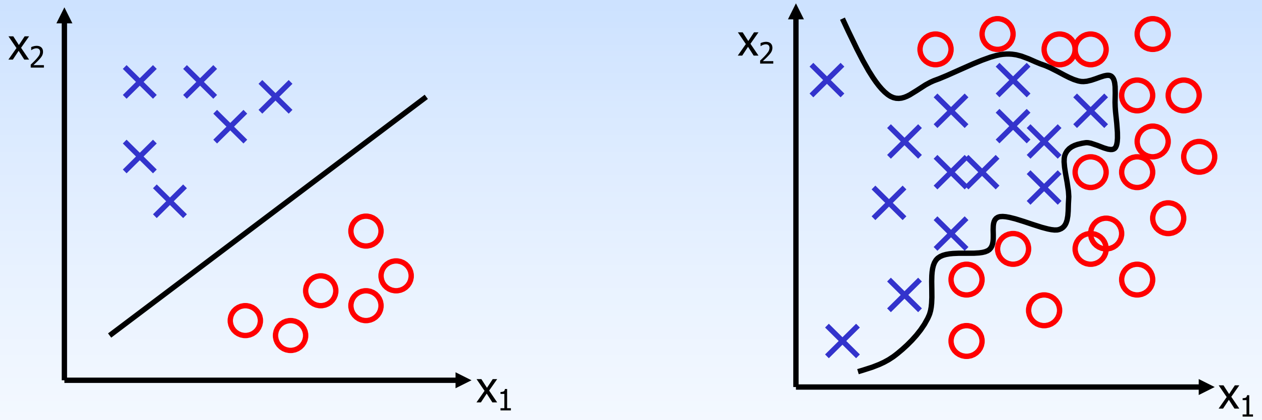


Statistical Pattern Recognition



Representation

- Each pattern is represented as a point in *d*-dimensional feature space
- Choice of features and their desired invariance properties are domain-specific



- Good representation implies (i) small intra-class variation, (ii) large interclass separation and (iii) simple decision boundary

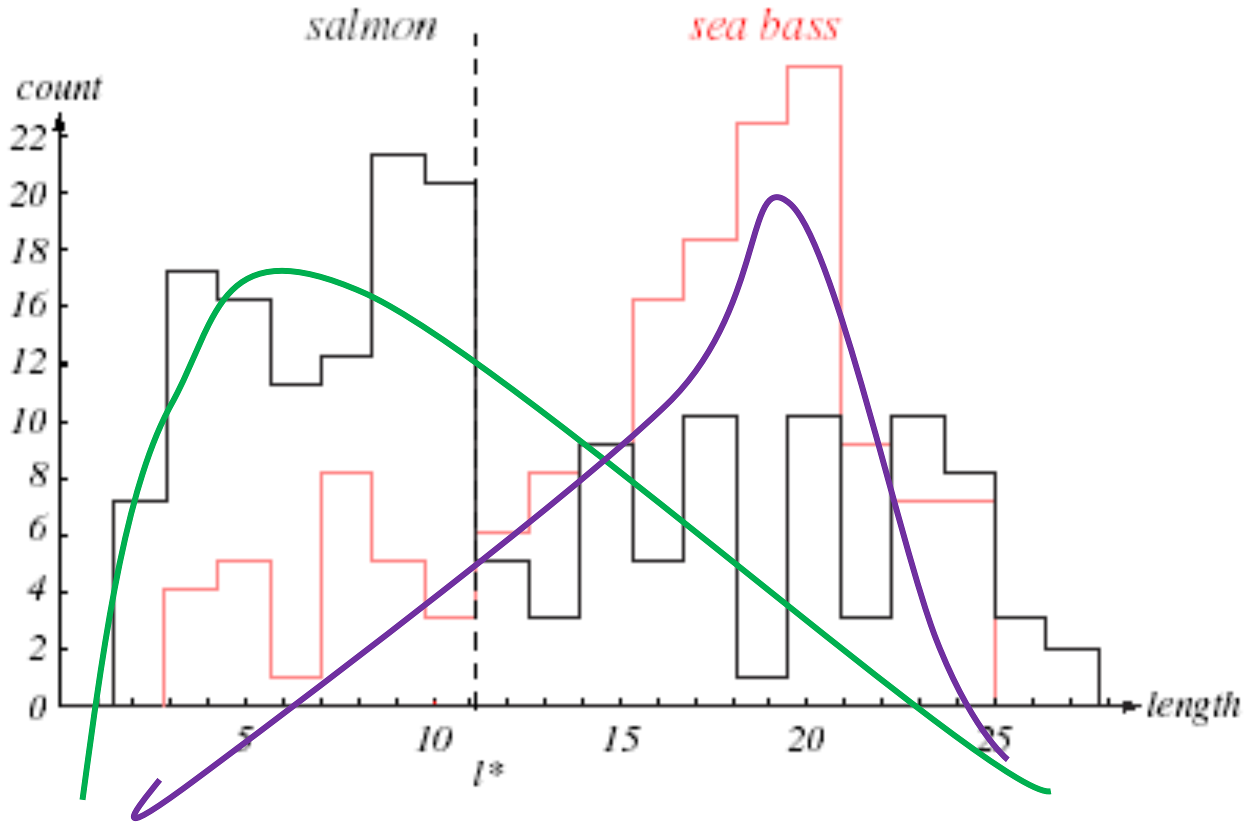


FIGURE 1.2. Histograms for the length feature for the two categories. No single threshold value of the length will serve to unambiguously discriminate between the two categories; using length alone, we will have some errors. The value marked l^* will lead to the smallest number of errors, on average.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

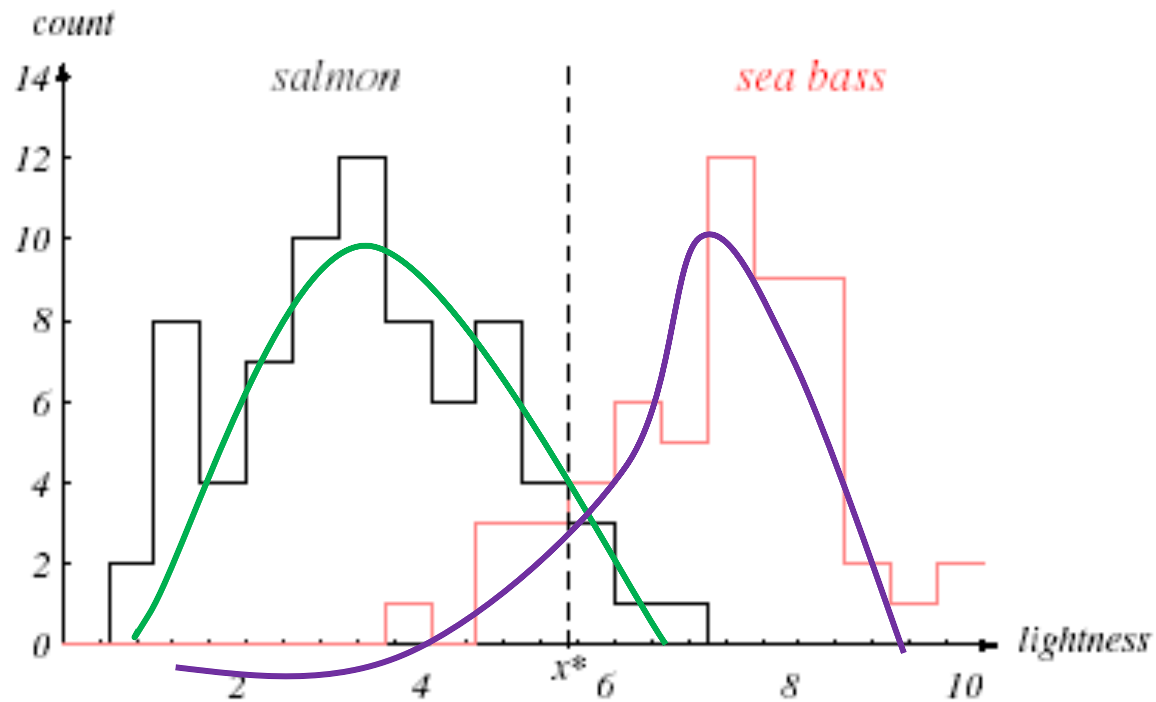


FIGURE 1.3. Histograms for the lightness feature for the two categories. No single threshold value x^* (decision boundary) will serve to unambiguously discriminate between the two categories; using lightness alone, we will have some errors. The value x^* marked will lead to the smallest number of errors,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

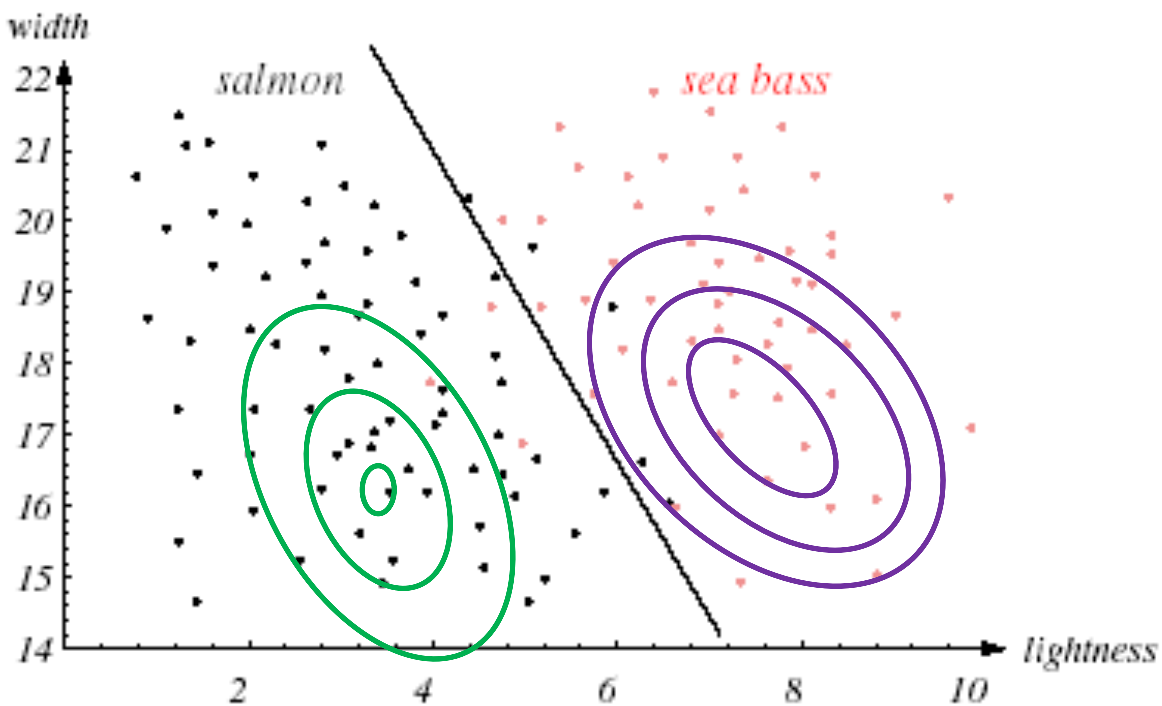
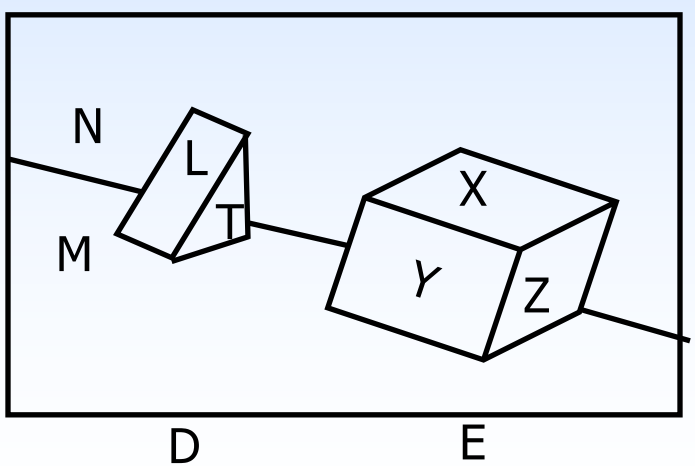
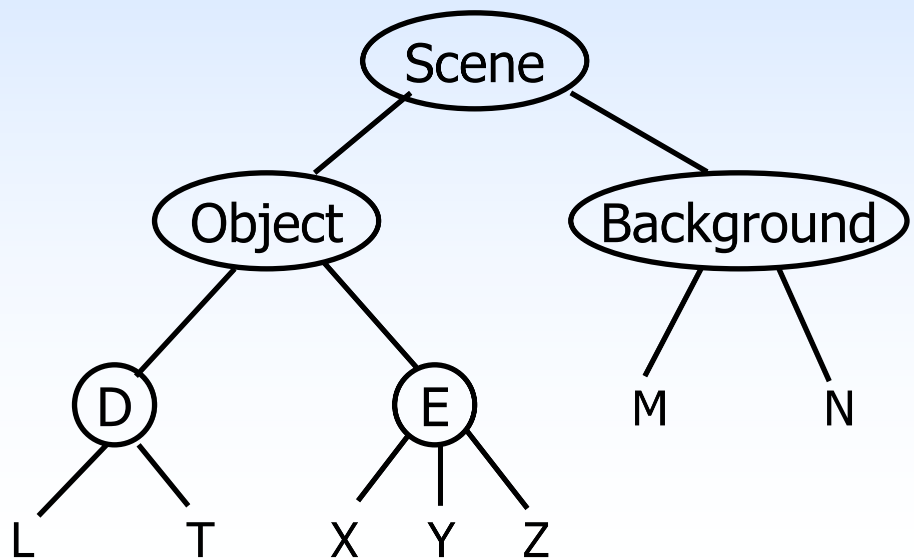


FIGURE 1.4. The two features of lightness and width for sea bass and salmon. The dark line could serve as a decision boundary of our classifier. Overall classification error on the data shown is lower than if we use only one feature as in Fig. 1.3, but there will still be some errors. From: Richard O.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

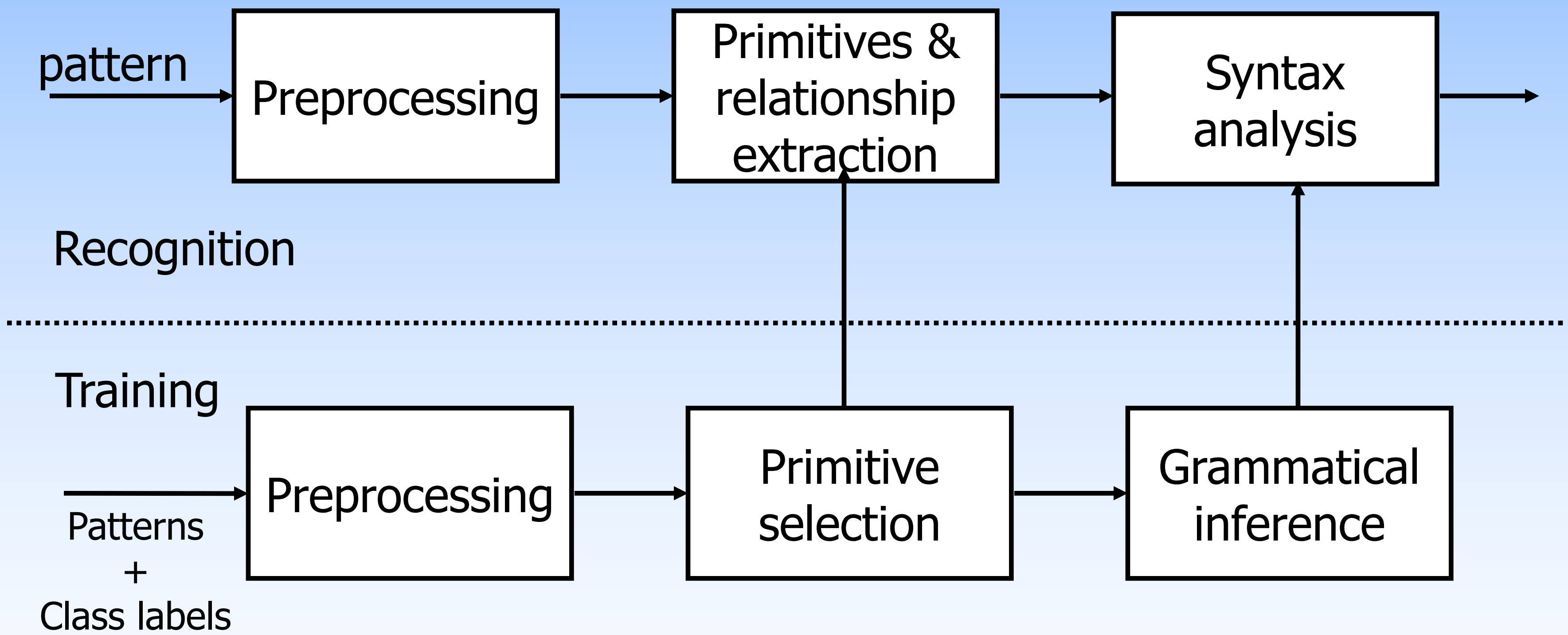
Structural Patten Recognition

- 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,…)



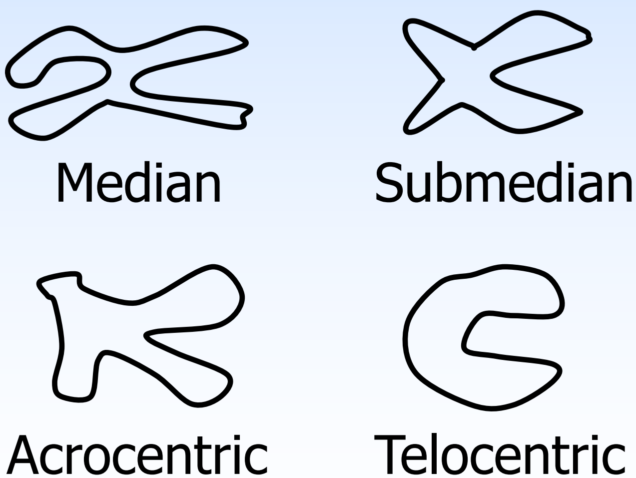


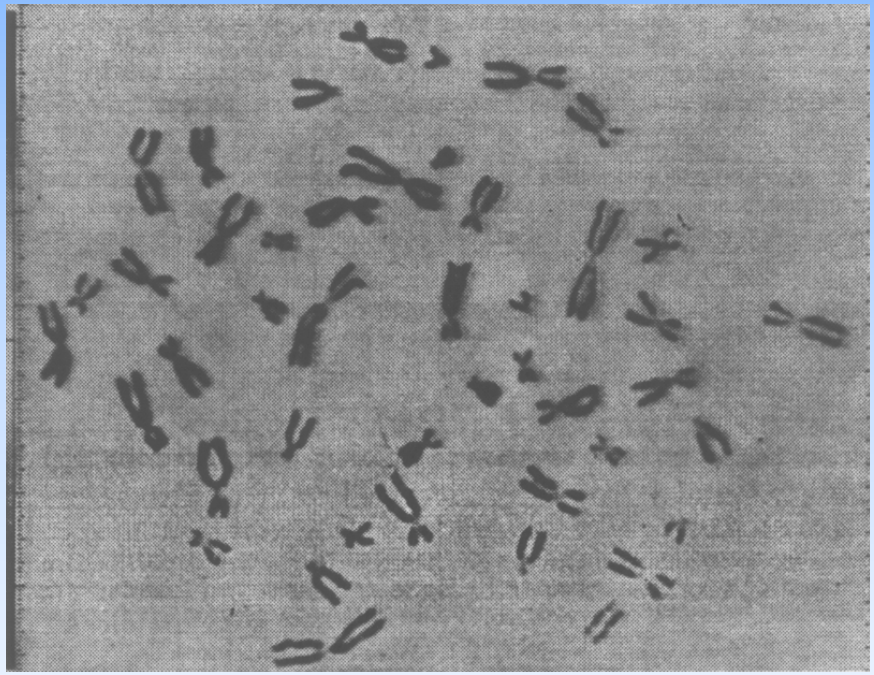
Syntactic Pattern Recognition



Chromosome Grammars

- Terminals:
  $V_T=\{\cap, |, \cup, \wr, \smile\}$
- Non-terminals:
  $V_N=\{A,B,C,D,E,F\}$
- Pattern Classes:





Chromosome Grammars

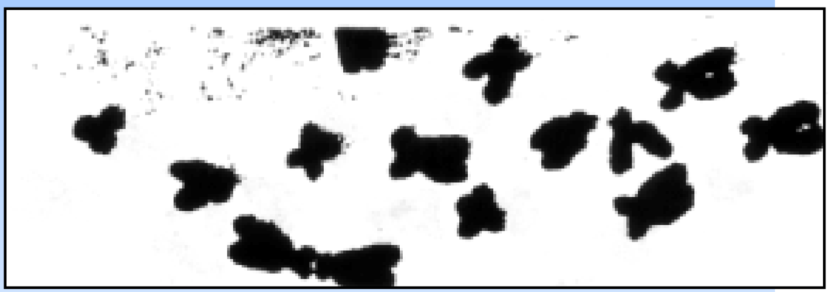
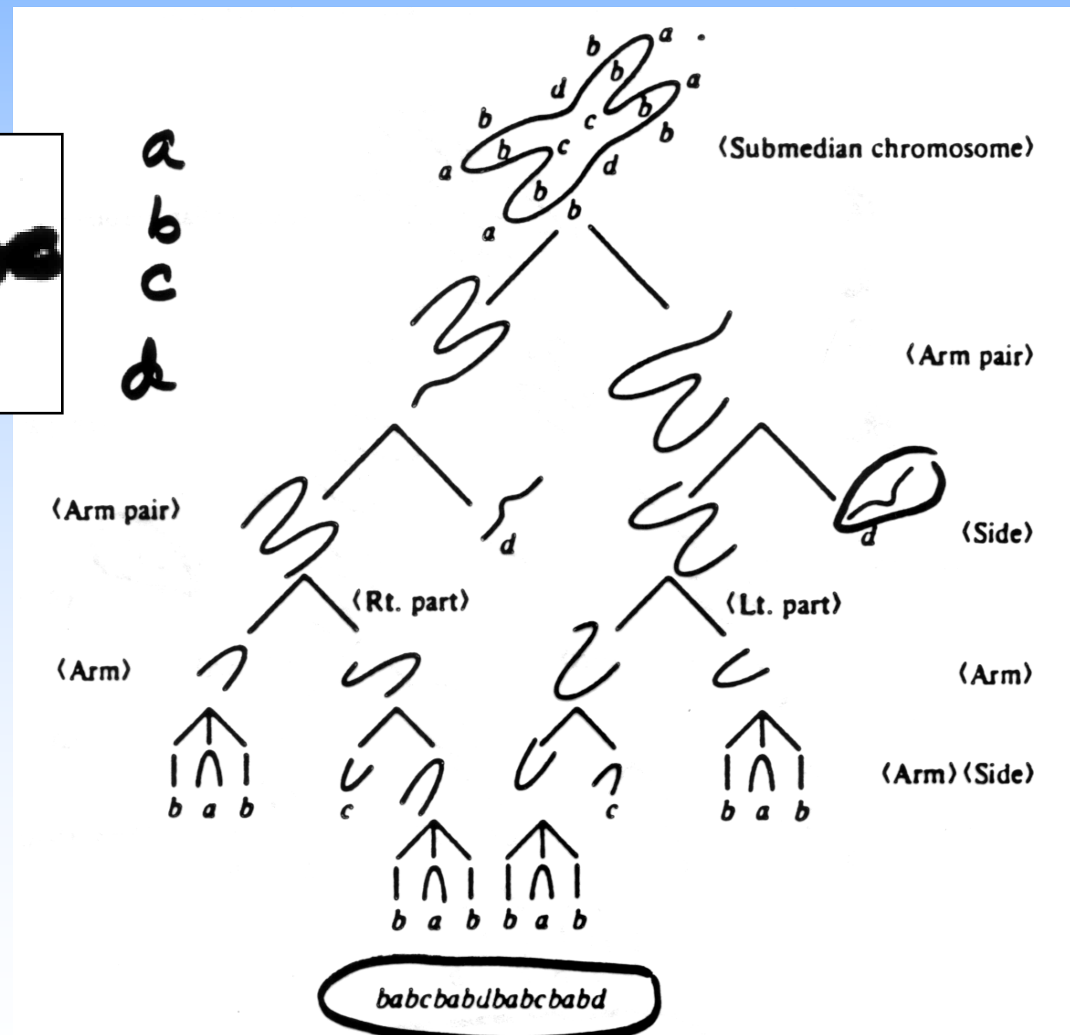


Image of human chromosomes



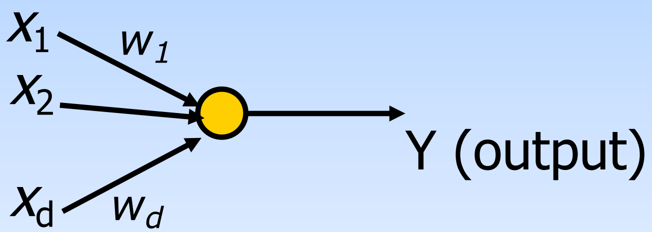
Hierarchical-structure description of a submedium chromosome

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 $\approx 10^{14}$

Neuron

- Nodes are nonlinear, typically analog

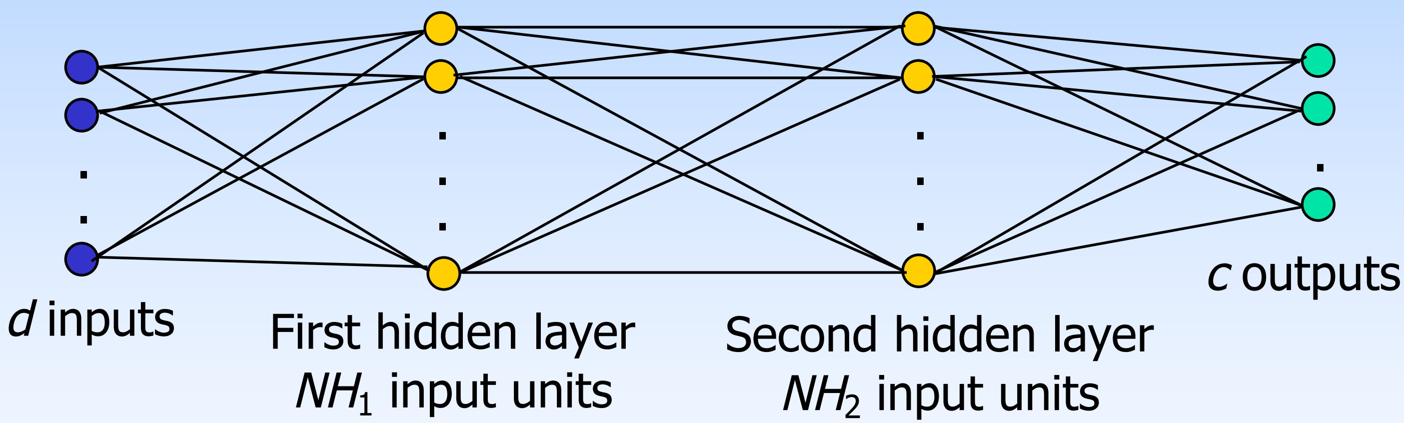


$$Y = f\left(\sum_{i=1}^{d} w_i x_i - \theta\right)$$

where θ is internal threshold or offset

Neural Networks

- Feed-forward networks with one or more layers (hidden) between input & output nodes
- How many nodes & hidden layers?



- Network training

Context: Post-processing

How m ch info mation are y u mi sing

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

Summary

- 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