Image Classification

Aziz Qaroush

Birzeit University

Outline

- Introduction to Image Classification
- Challenges
- A simple Image Classification Pipeline
- Data Sets

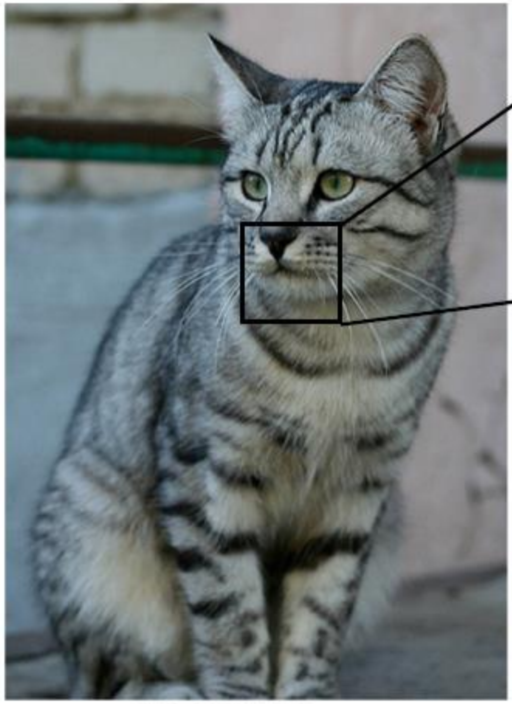
Image Classification: A core computer vision task

Input: image

Output: Assign image to one of a fixed set of categories

cat
bird
deer
dog
truck

Problem: Semantic Gap




```
[[105 112 108 111 104  99 106  99  96 103 112 119 104  97  93  87]
 [ 91  98 102 106 104  79  98 103  99 105 123 136 110 105  94  85]
 [ 76  85  90 105 128 105  87  96  95  99 115 112 106 103  99  85]
 [ 99  81  81  93 120 131 127 100  95  98 102  99  96  93 101  94]
 [106  91  61  64  69  91  88  85 101 107 109  98  75  84  96  95]
 [114 108  85  55  55  69  64  54  64  87 112 129  98  74  84  91]
 [133 137 147 103  65  81  80  65  52  54  74  84 102  93  85  82]
 [128 137 144 140 109  95  86  70  62  65  63  63  60  73  86 101]
 [125 133 148 137 119 121 117  94  65  79  80  65  54  64  72  98]
 [127 125 131 147 133 127 126 131 111  96  89  75  61  64  72  84]
 [115 114 109 123 150 148 131 118 113 109 100  92  74  65  72  78]
 [ 89  93  90  97 108 147 131 118 113 114 113 109 106  95  77  80]
 [ 63  77  86  81  77  79 102 123 117 115 117 125 125 130 115  87]
 [ 62  65  82  89  78  71  80 101 124 126 119 101 107 114 131 119]
 [ 63  65  75  88  89  71  62  81 120 138 135 105  81  98 110 118]
 [ 87  65  71  87 106  95  69  45  76 130 126 107  92  94 105 112]
 [118  97  82  86 117 123 116  66  41  51  95  93  89  95 102 107]
 [164 146 112  80  82 120 124 104  76  48  45  66  88 101 102 109]
 [157 170 157 120  93  86 114 132 112  97  69  55  70  82  99  94]
 [130 128 134 161 139 100 109 118 121 134 114  87  65  53  69  86]
 [128 112  96 117 150 144 120 115 104 107 102  93  87  81  72  79]
 [123 107  96  86  83 112 153 149 122 109 104  75  80 107 112  99]
 [122 121 102  80  82  86  94 117 145 148 153 102  58  78  92 107]
 [122 164 148 103  71  56  78  83  93 103 119 139 102  61  69  84]]
```

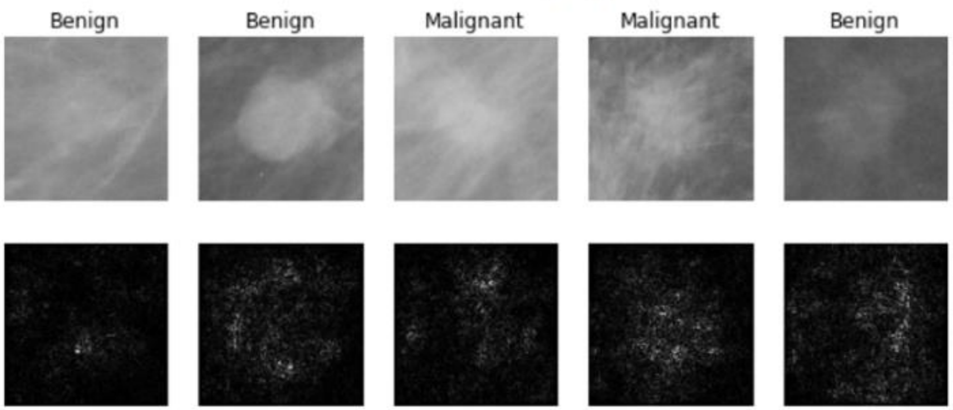
What the computer sees

An image is just a big grid of numbers between [0, 255]:

e.g. 800 x 600 x 3
(3 channels RGB)

Image Classification: Very Useful!

Medical Imaging



Levy et al, 2016 Figure reproduced with permission

Galaxy Classification

Dieleman et al, 2014

Whale recognition

Kaggle Challenge

Image Classification: Building Block for other tasks!

- Object Detection

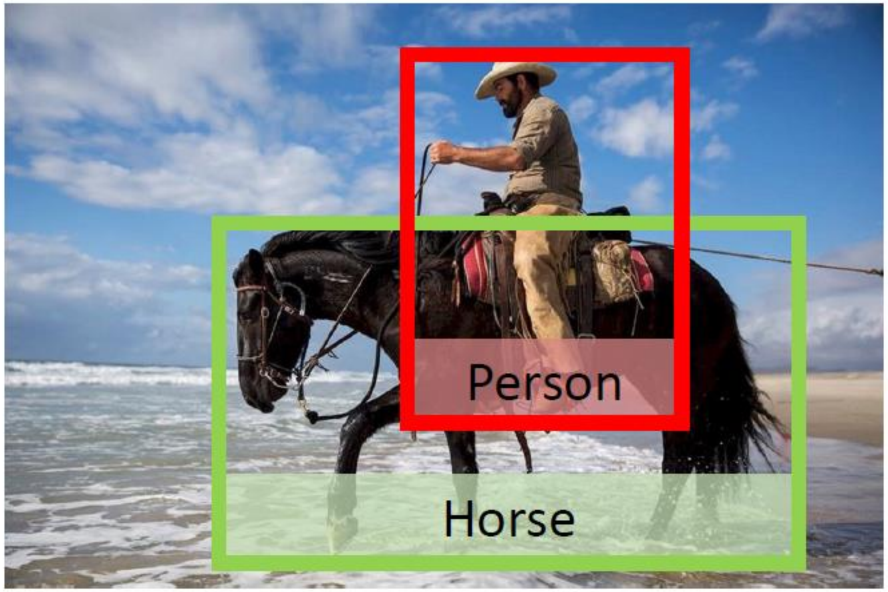


Image Classification: Building Block for other tasks!

- Image Captioning

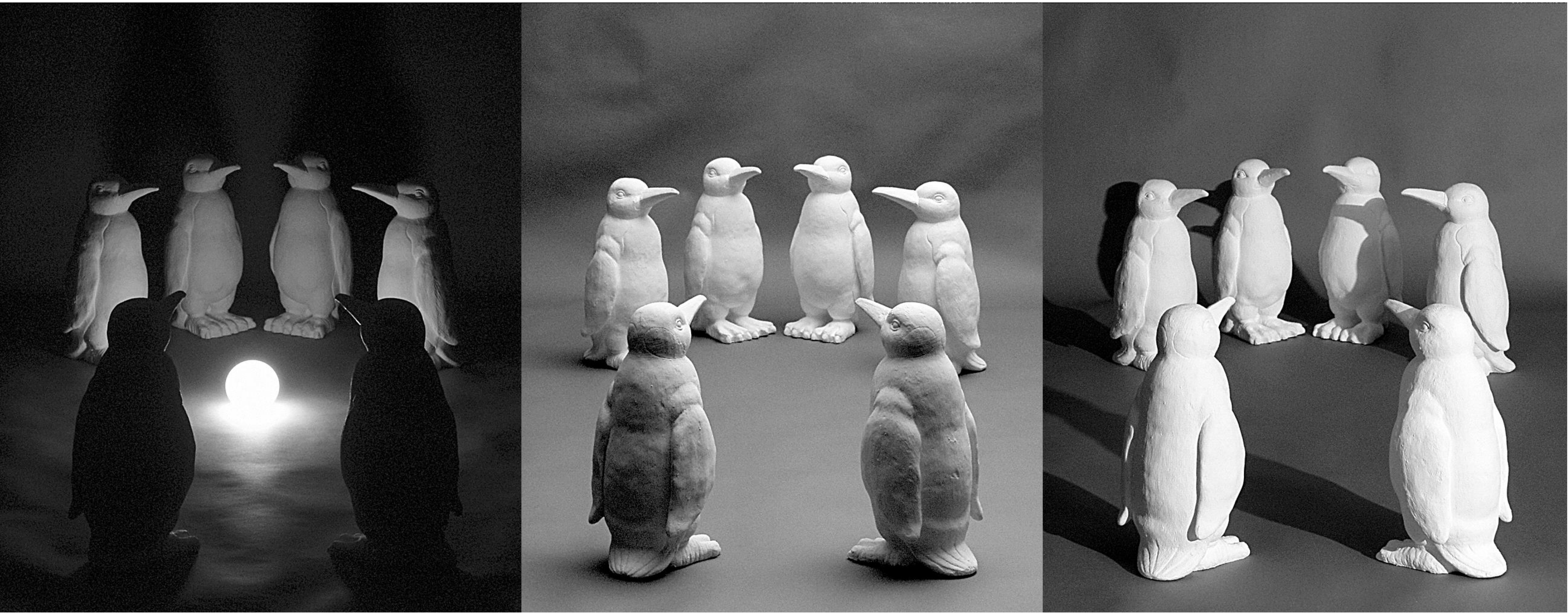
riding
cat
horse
man
when
...
<STOP>

What word to say next?

Caption: Man riding horse

Challenges: Viewpoint Variation

Challenges: Illumination



Challenges: Scale

Challenges: Deformation

Challenges: Occlusion

Challenges: Background Clutter

How many object categories are there?

Categorization vs Single instance recognition

Where is the crunchy nut?

Challenges: Intra-class variation



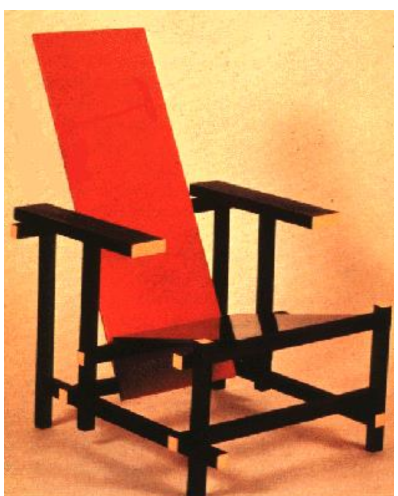
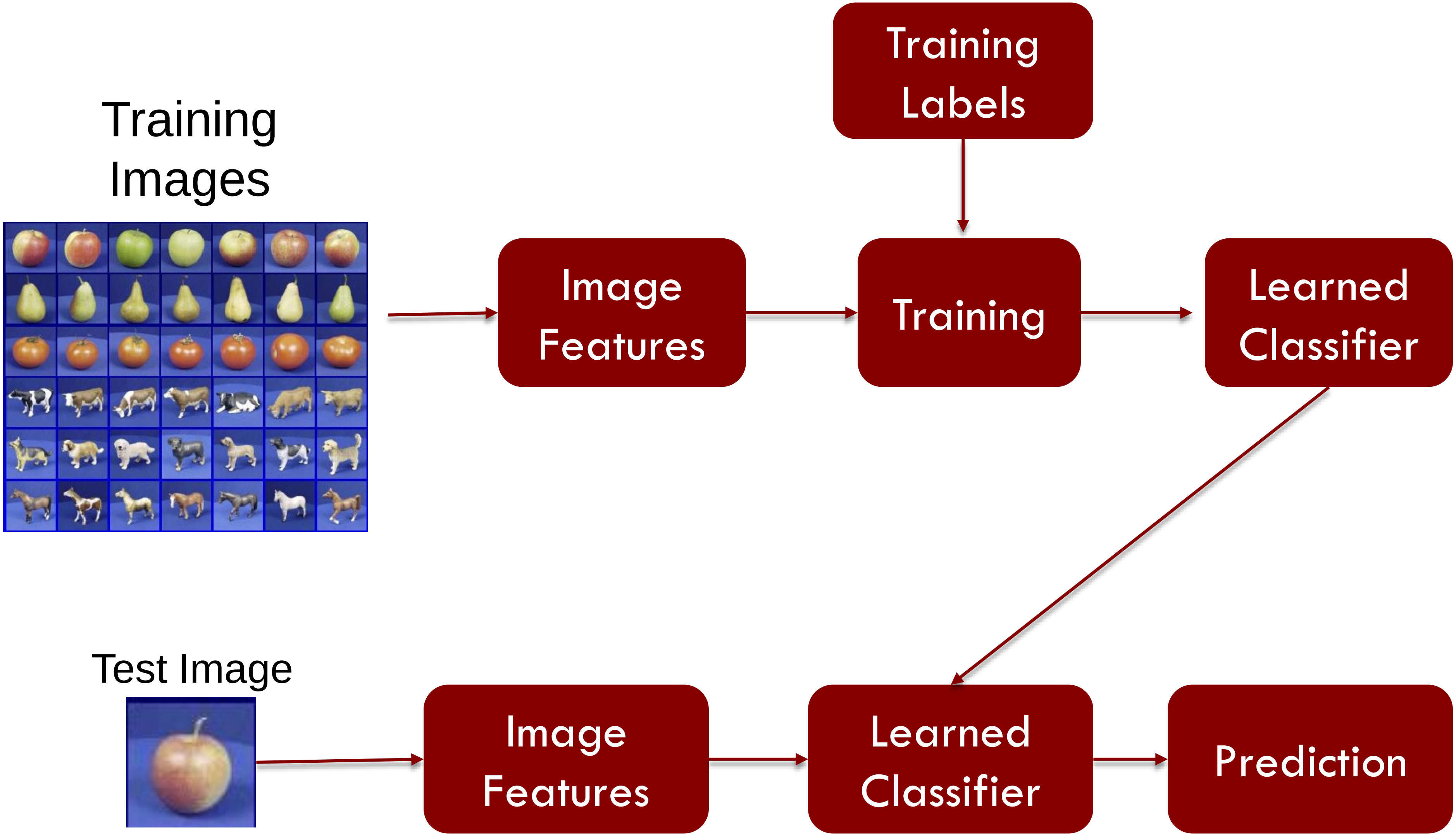



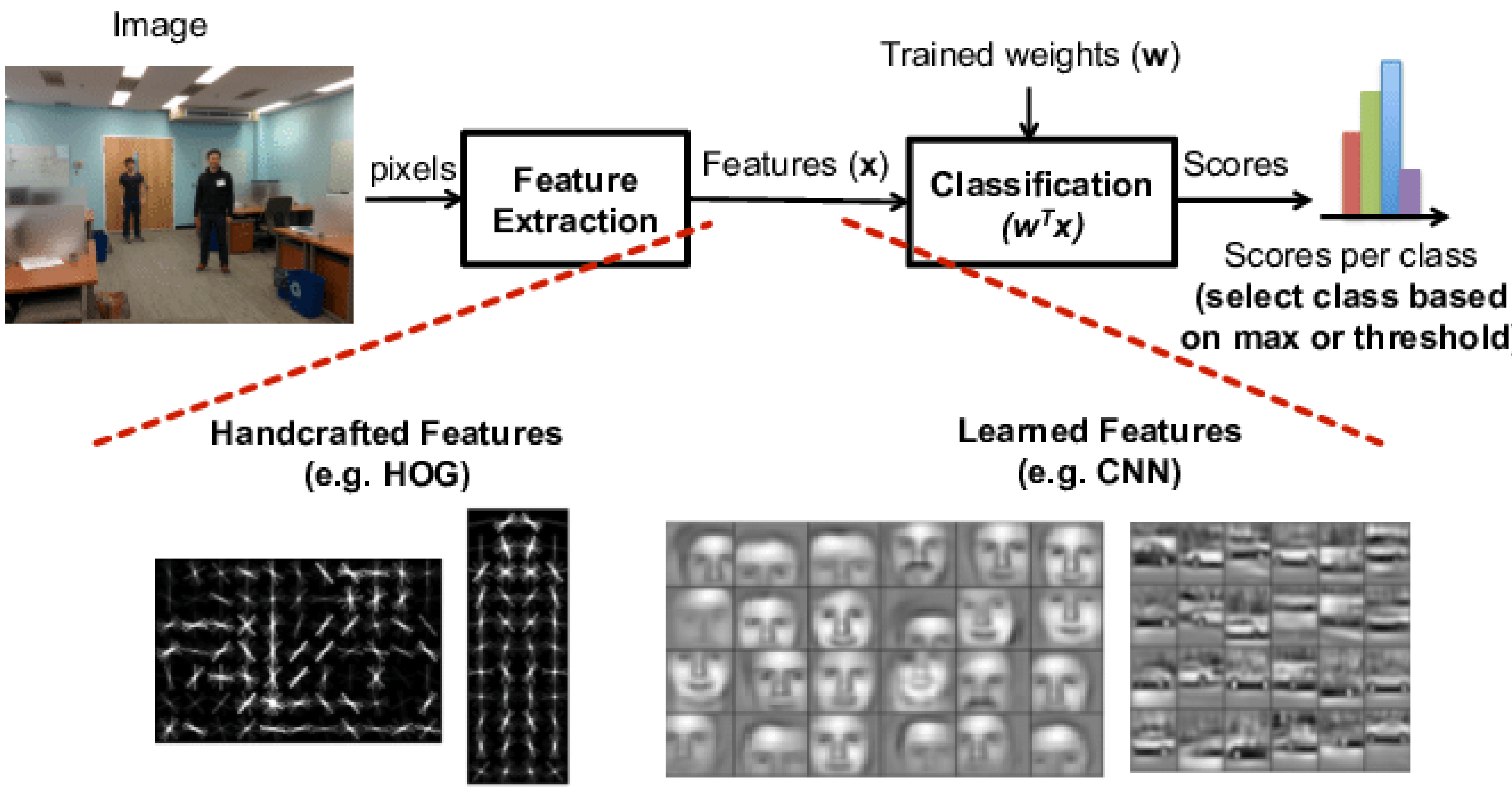
Image Classification Pipeline

- The pipeline typically consists of the following steps:
  1. **Data collection and preparation**: This involves collecting a large and diverse dataset of images, labeled with their corresponding categories. The images should be preprocessed to ensure that they are in a consistent format and size.
  2. **Feature extraction**: This involves extracting features from the images that can be used to train the machine learning model. Features can be extracted using a variety of methods, such as hand-crafted algorithms or deep learning models.
  3. **Model training**: This involves training a machine learning model on the extracted features. The model learns to identify the patterns in the features that are associated with different categories.
  4. **Model evaluation**: This involves evaluating the performance of the trained model on a held-out test set. The evaluation results are used to assess the model's accuracy and generalization ability.
  5. **Model deployment**: Once the model is trained and evaluated, it can be deployed to production to classify new images.

Image Classification Pipeline



Feature Extraction Techniques



Feature Extraction Techniques

Aspect	Handcrafted Features	Learned Features (Deep Learning)
Control Over Feature Selection	Full control based on domain knowledge.	Features are automatically learned from data.
Interpretability	More interpretable, as features are explicit.	Often considered as "black boxes" with less interpretability.
Performance on Large Datasets	May struggle with large datasets due to overfitting.	Excels on large datasets, capturing intricate patterns.
Data Representation Learning	Limited ability to learn data representations.	Learns hierarchical representations of data.
Problem Complexity	Suitable for well-understood problems.	Suitable for complex problems with unknown patterns.
Transfer Learning	Limited application due to handcrafted nature.	Can leverage pre-trained models for transfer learning.

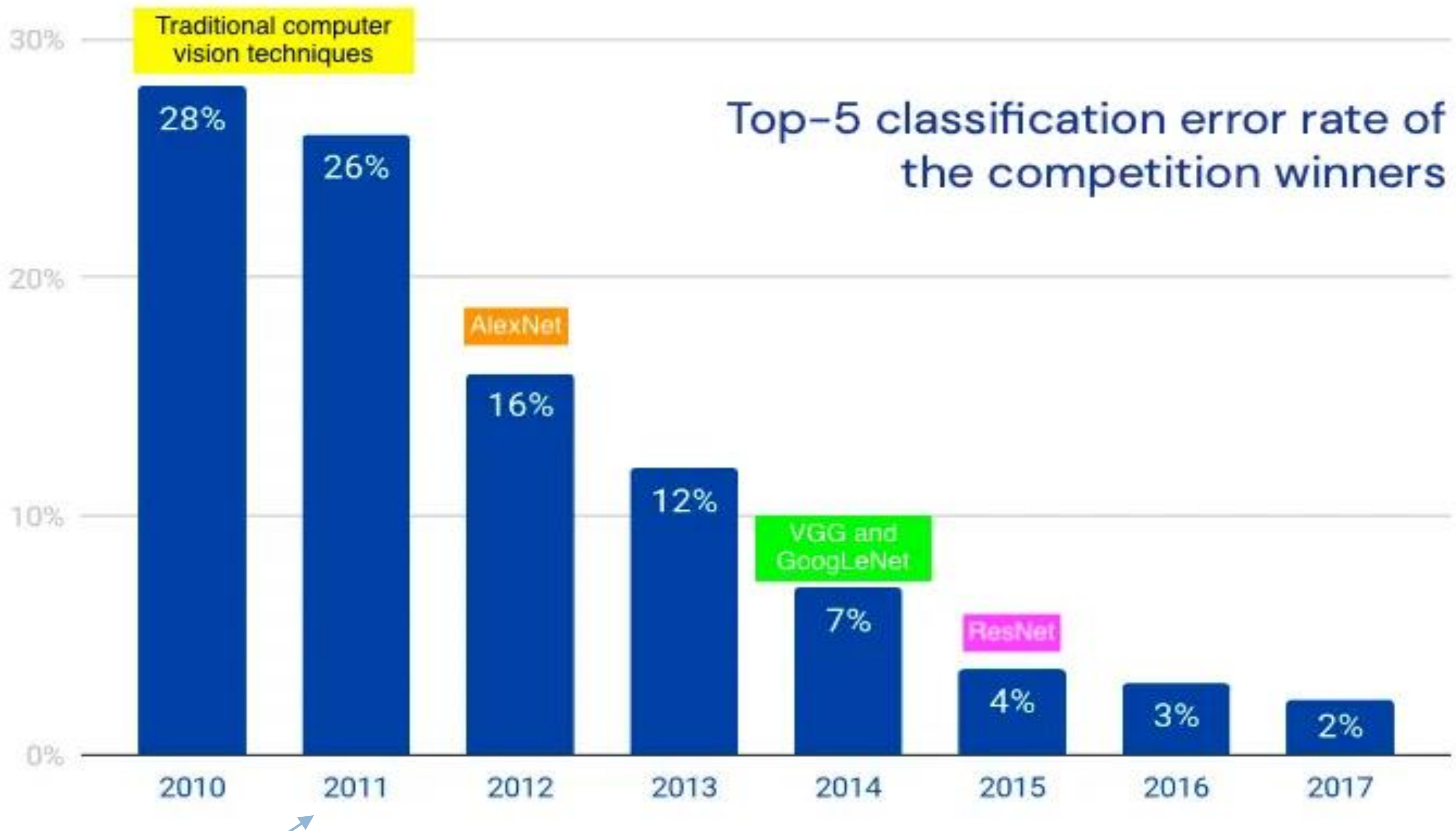
Many classifiers to choose from

- K-nearest Neighbor
- Decision Trees
- Naive Bayes
- SVM
- Random Forests
- XGBoost
- Neural Networks
- Convolutional neural network.
- .....

What is the best one?

- CNNs have been shown to achieve state-of-the-art results on a variety of image classification tasks.

ImageNet Competition Winners



Low-level feature extraction ≈ 10k patches per image

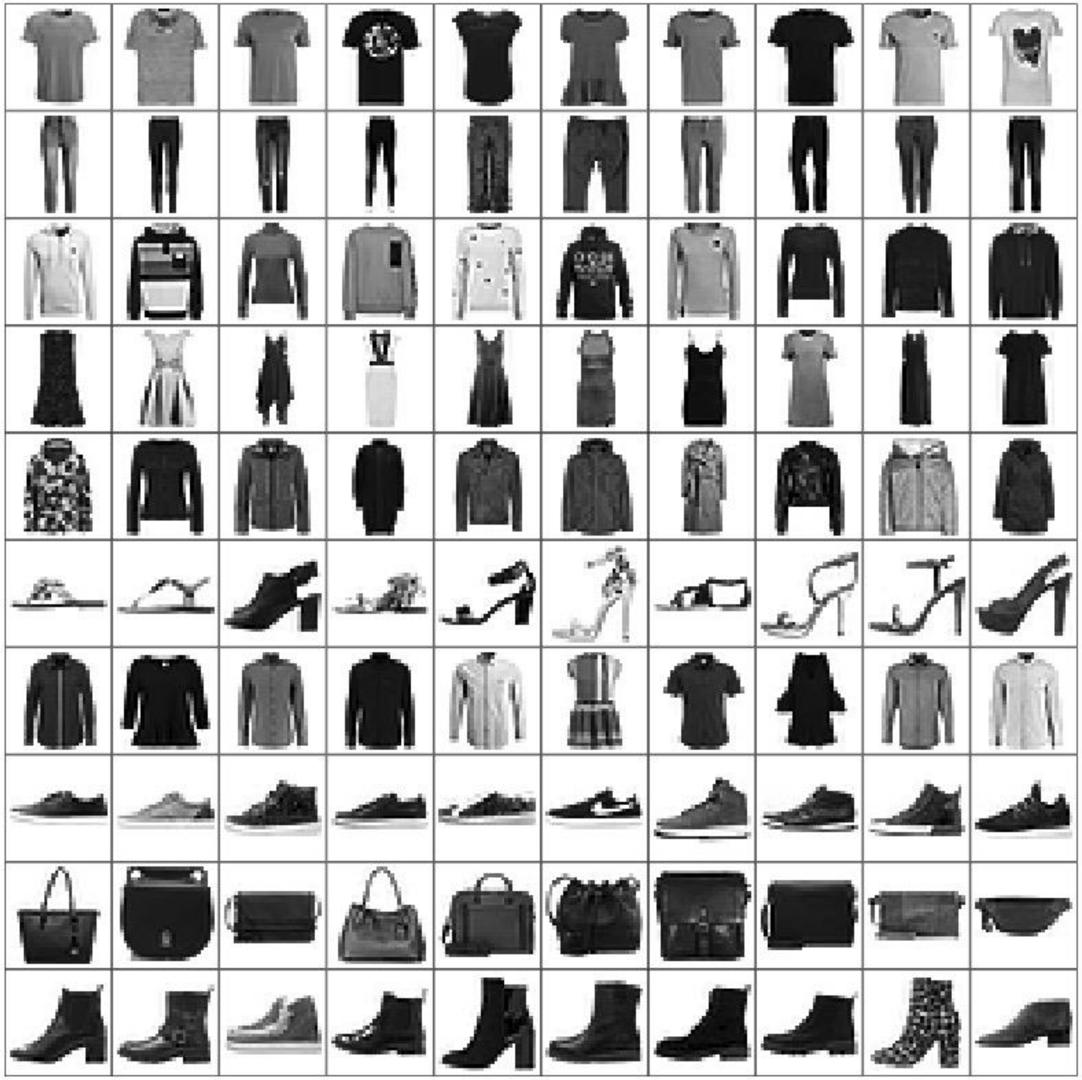
- SIFT: 128-dim
- color: 96-dim

reduced to 64-dim with PCA

Image Classification Datasets: MNIST

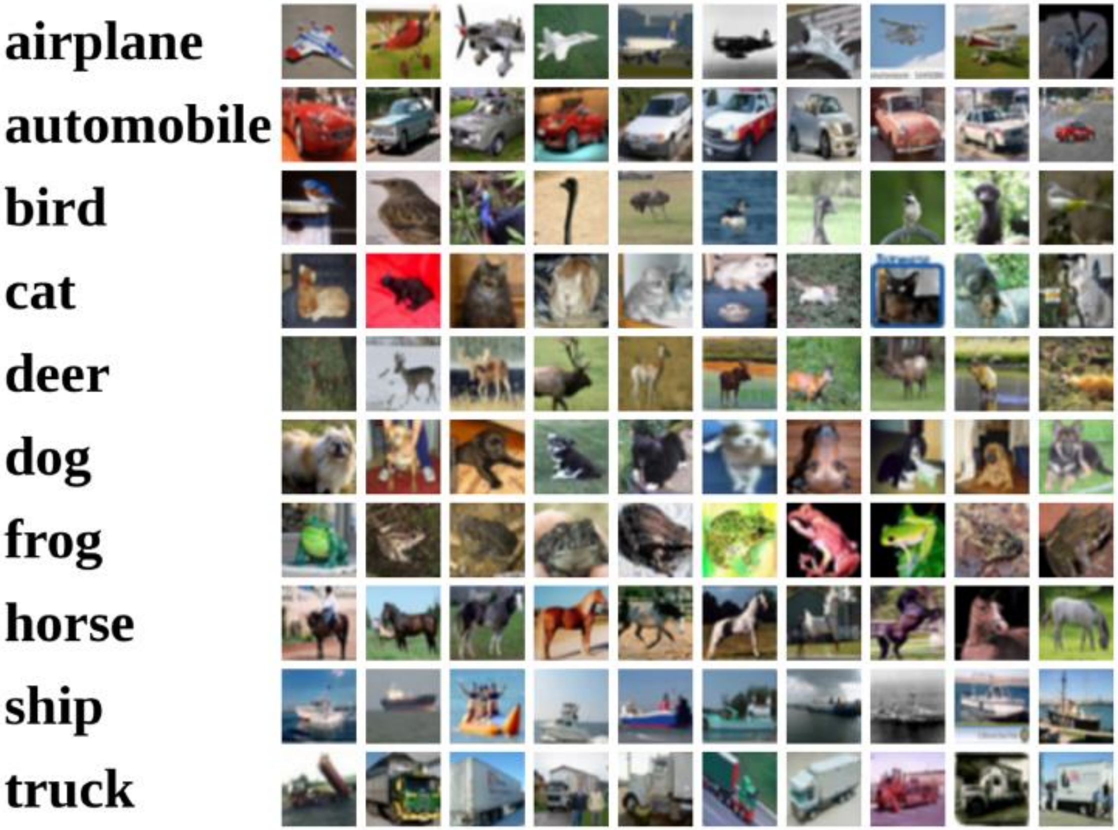
- **10 classes**: Digits 0 to 9
- **28x28** grayscale images
- **50k** training images
- **10k** test images

Image Classification Datasets: Fashion MNIST



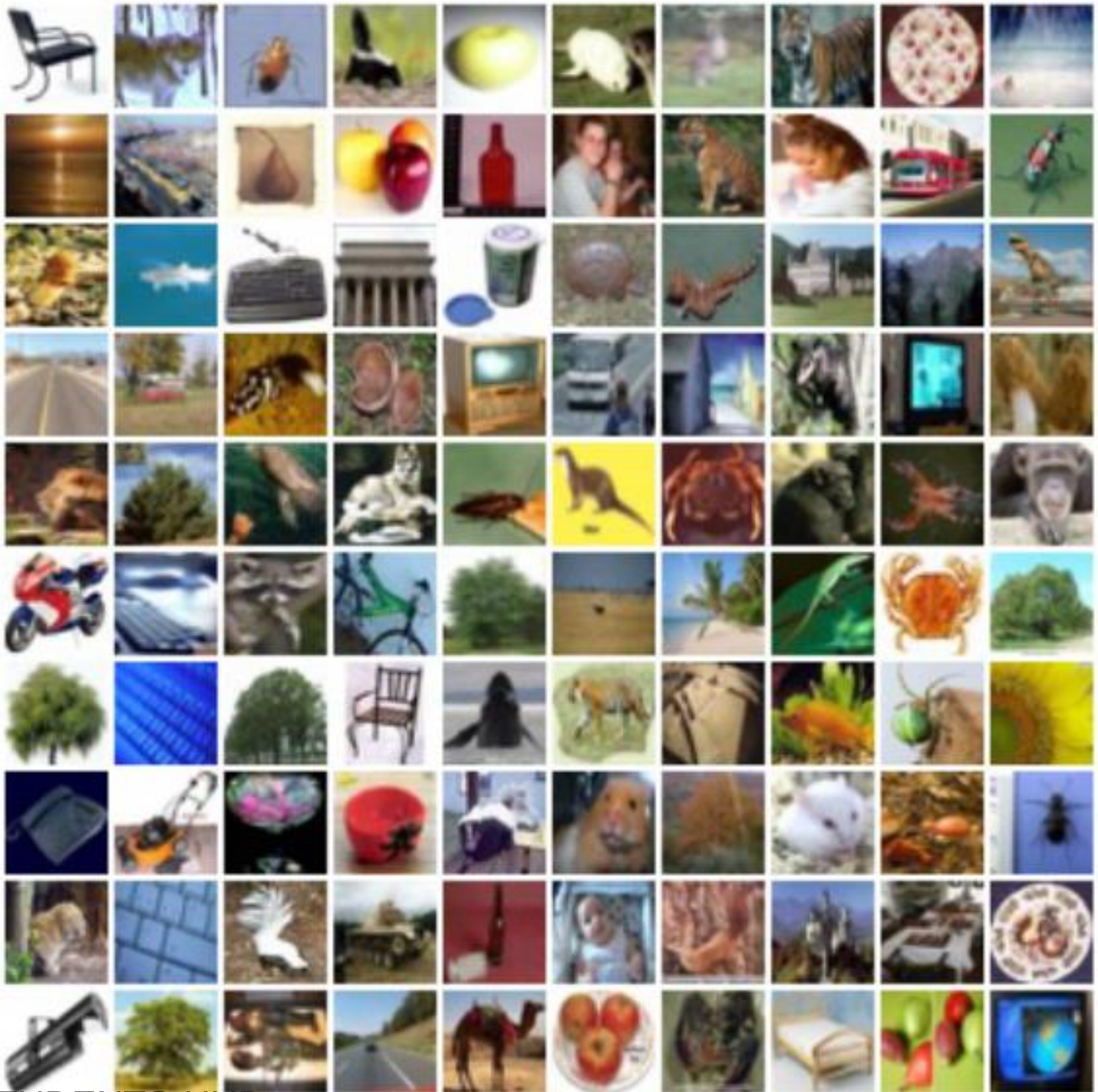
- The dataset contains 60,000 training images and 10,000 testing images
- Consists of 28x28 grayscale images of fashion items,
- 10 categories.
- Each category represents a different type of clothing or accessory.

Image Classification Datasets: CIFAR10



- **10** classes
- **50k** training images (5k per class)
- **10k** testing images (1k per class)
- **32x32 RGB** images

Image Classification Datasets: CIFAR100



- **100** classes
- **50k** training images (500 per class)
- **10k** testing images (100 per class)
- **32x32 RGB** images
- **20 superclasses** with 5 classes each:
- Aquatic mammals: beaver, dolphin, otter, seal, whale
- Trees: Maple, oak, palm, pine, willow

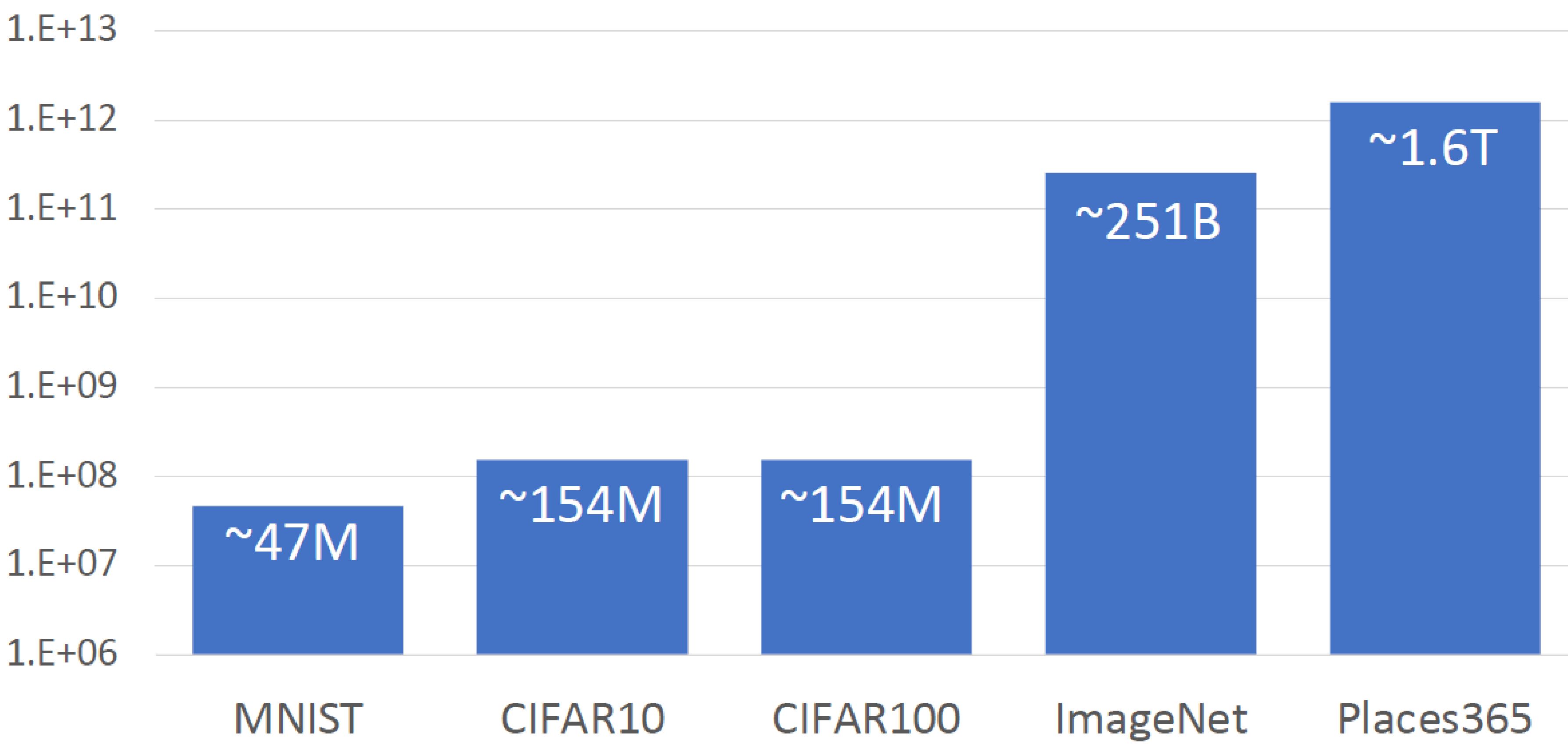
Image Classification Datasets: ImageNet

- **1000** classes
- **~1.3M** training images (~1.3K per class)
- **50K** validation images (50 per class)
- **100K** test images (100 per class)
- Performance metric: **Top 5 accuracy**
- Algorithm predicts 5 labels for each image; one of them needs to be right
- There is also a 22k category version of ImageNet, but less commonly used

Image Classification Datasets: MIT Places

- **365 classes** of different scene types
- ~**8M** training images
- **18.25K** val images (50 per class)
- **328.5K** test images (900 per class)
- Images have variable size, often
- resize to **256x256** for training

Classification Datasets: Number of Training Pixels



Acknowledgement

- The material in these slides are based on:
  - Digital Image Processing: Rafael C. Gonzalez, and Richard
  - Forsythe and Ponce: Computer Vision: A Modern Approach
  - Rick Szeliski's book: Computer Vision: Algorithms and Applications
  - cs131@ Stanford University
  - cs131n@ Stanford University
  - CS198-126@ University of California, Berkely
  - CAP5415@ University of Central Florida
  - CSW182 @ University of California, Berkely
  - Deep Learning Lecture Series @UCL
  - EECS 498.008 @ University of Michigan
  - CSE576 @ Washington University
  - 11-785@ Carnegie Mellon University
  - CSCI1430@ Brown University
  - Computer Vision@ Bonn University
  - ICS 505@ KFUPM
  - Digital Image Processing@ University of Jordan

