Chapter 45
Machine Learning Applications in Computer Vision

Mehrtash Harandi
NICTA, Australia & The University of Queensland, Australia

Javid Taheri
The University of Sydney, Australia

Brian C. Lovell
NICTA, Australia & The University of Queensland, Australia

ABSTRACT

Recognizing objects based on their appearance (visual recognition) is one of the most significant abilities of many living creatures. In this study, recent advances in the area of automated object recognition are reviewed; the authors specifically look into several learning frameworks to discuss how they can be utilized in solving object recognition paradigms. This includes reinforcement learning, a biologically-inspired machine learning technique to solve sequential decision problems and transductive learning, and a framework where the learner observes query data and potentially exploits its structure for classification. The authors also discuss local and global appearance models for object recognition, as well as how similarities between objects can be learnt and evaluated.

INTRODUCTION

The very first question comes into mind when thinking about machine learning and computer vision is: "How can a machine recognize and interpret images/scenes?" Without a doubt, among the various capabilities that humans beings have, recognizing objects based on their appearance (visual recognition) is one of our most significant abilities. Human beings can recognize familiar objects with little difficulty; whereas, artificial vision systems are far from matching the accuracy, speed and generality of human vision (Ponce, Hebert, Schmid, & Zisserman, 2007). In fact, creating a machine with similar capabilities to human beings is one of the main goals of computer vision research. Upon achieving this ultimate goal, a wide variety of powerful applications would emerge. Face recognition – the task of identifying a person from his/her images – is probably the best

known application of computer vision and pattern recognition. In fact, any task that makes use of a scene to identify objects is a potential application; content-based image retrieval, object tracking, robot navigation, automated surveillance, etc are among many that come to mind.

In this chapter, several machine learning approaches devised for recognizing objects will be introduced. Influential works along recent topics in the area of object recognition and machine learning will also be discussed.

OBJECT DESCRIPTION AND REPRESENTATION IN COMPUTER VISION

Objects can be described by different cues. For example, objects can be described by geometrical primitives like boxes, spheres, cones, and cylinders. Describing and representing objects based on their appearance is a widely studied approach in the literature (Bartlett, Movellan, & Sejnowski, 2002; Belhumeur, Hespanha, & Kriegman, 1997; Dorko & Schmid, 2005; Lowe, 2004; Mikolajczyk, Leibe, & Schiele, 2005; Serre, Wolf, Bileschi, Riesenhuber, & Poggio, 2007; Shechtman & Irani, 2007; Turk & Pentland, 1991; H. Zhang, Gao, Chen, & Zhao, 2006). The general idea of appearance-based object recognition is to extract useful and robust information from only the appearance of the object-of-interest that is usually captured by different two-dimensional views. Appearance-based methods can be sub-divided into two main classes: local and global approaches.

Global Appearance Models for Object Recognition

Global appearance-based methods for object recognition try to project original input images onto a suitable lower dimensional representation. By choosing different optimization criteria for the projected data different methods can be derived.

Principal Component Analysis (PCA)

PCA was introduced to Computer Vision by Kirby and Sirovich (Sirovich & Kirby, 1987) and is the most well-known projection in this area. It was popularized as the influential Eigenface method of Turk and Pentland in face recognition in 1991 (Turk & Pentland, 1991). In the Eigenface method, images are considered to be high dimensional vectors and a given image x_i of size $h{\times}w$ is arranged as a vector $\in \mathbb{R}^d$, where $d{=}hw$. If a set of images $\mathbf{X}{=}\{\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N\}$, $\mathbf{x}_i \in \mathbb{R}^d$ are available for training then the Eigenface method uses PCA projection to derive a lower dimension representation of images in the form of $\mathbf{y}_i = \mathbf{U}^T_{d\times m}\mathbf{x}_i$, $m{\ll}d$. Not only has PCA and its extensions been widely used as object descriptors (Gottumukkal & Asari, 2004; K. I. Kim, Jung, & Kim, 2002; C. Liu, 2004; C. Liu & Wechsler, 2000; Moghaddam, 2002; Vidal, Ma, & Sastry, 2005; J. Yang, Frangi, Yang, Zhang, & Jin, 2005; J. Yang, Zhang, Frangi, & Yang, 2004) but it has been also considered as a pre-processing step for other methods like Fisherface (Belhumeur, et al., 1997) or ICA-based approaches (Bartlett, et al., 2002). A numerically robust way to compute PCA projection is based on Singular Value Decomposition (SVD). This algorithm is illustrated in Table 1.

Linear Discriminant Analysis

As PCA technique finds orthogonal dimensions of maximal variance (or energy), if substantial changes in illumination, pose or expression are presented in the training set, the resulting descriptor then does not necessary encode similarities between faces (Jain & Li, 2005). To alleviate this problem, if class labels are provided in the training set, this additional information can be used for learning. Linear Discriminant Analysis (LDA) is a linear projection that simultaneously (1) maximizes distances among different classes, and (2) reduces distances among samples inside

29 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/machine-learning-applications-computer-vision/77579

Related Content

Image Enhancement and Restoration Methods for Underwater Images
C. J. Prabhakarand P. U. Praveen Kumar (2014). *Research Developments in Computer Vision and Image Processing: Methodologies and Applications (pp. 93-110).*
www.irma-international.org/chapter/image-enhancement-and-restoration-methods-for-underwater-images/79722

Analysis of Text Identification Techniques Using Scene Text and Optical Character Recognition
Monica Gupta, Alka Choudharyand Jyotsna Parmar (2021). *International Journal of Computer Vision and Image Processing (pp. 39-62).*
www.irma-international.org/article/analysis-of-text-identification-techniques-using-scene-text-and-optical-character-recognition/288385

Evolutionary Image Thresholding for Image Segmentation
 Phanindra Kumar N.S.R.and Prasad Reddy P.V.G.D. (2019). *International Journal of Computer Vision and Image Processing (pp. 17-34).*
www.irma-international.org/article/evolutionary-image-thresholding-for-image-segmentation/224377

From Face to Facial Expression
Zakia Hammal (2013). *Image Processing: Concepts, Methodologies, Tools, and Applications (pp. 1508-1531).*
www.irma-international.org/chapter/face-facial-expression/77608

Advances in Artificial Intelligence for Image Processing: Techniques, Applications, and Optimization
Sampath Boopathi, Binay Kumar Pandeyand Digvijay Pandey (2023). *Handbook of Research on Thrust Technologies' Effect on Image Processing (pp. 73-95).*
www.irma-international.org/chapter/advances-in-artificial-intelligence-for-image-processing/328027