Chapter VII

Bayesian Networks for Image Understanding

Andreas Savakis, Rochester Institute of Technology, USA
Jiebo Luo, Kodak Research Laboratories, USA
Michael Kane, Yale University, USA

Abstract

Image understanding deals with extracting and interpreting scene content for use in various applications. In this chapter, we illustrate that Bayesian networks are particularly well-suited for image understanding problems, and present case studies in indoor-outdoor scene classification and parts-based object detection. First, improved scene classification is accomplished using both low-level features, such as color and texture, and semantic features, such as the presence of sky and grass. Integration of low-level and semantic features is achieved using a Bayesian network framework. The network structure can be determined by expert opinion or by automated structure learning methods. Second, object detection at multiple views relies on a parts-based approach, where specialized detectors locate object parts and a Bayesian network acts as the arbitrator in order to determine the object presence. In general, Bayesian networks are found to be powerful integrators of different features and help improve the performance of image understanding systems.
Introduction

Bayesian networks, also known as belief networks or Bayes nets, have emerged as an effective tool for knowledge representation and inference (Neapolitan, 2003; Pearl, 1988). A Bayesian network is a directed, acyclic graph that can be used to represent the dependency between random variables, represented by nodes. Links between nodes represent conditional probabilities and link directions represent causality between the parent and children nodes. A distinct advantage of Bayesian networks is the ability to incorporate domain-specific knowledge in the network structure, so that the overall joint probability distribution is expressed as a set of conditionally independent relationships that are easier to characterize. According to Bayes’ rule, the posterior probability can be expressed in terms of the joint probability, which can be further expressed by conditional probability and prior probability:

$$P(S|E) = \frac{P(S,E)}{P(E)} = \frac{P(E|S)P(S)}{P(E)},$$

where $S$ denotes semantic task and $E$ denotes evidence. Probabilistic reasoning uses the joint probability distribution of a given domain to answer a question about this domain. However, as the number of variables grows, the joint probability can become intractable. With Bayesian networks, the computation of the joint probability distribution over the entire system, given partial evidences about the state of the system, is greatly simplified by using Bayes’ rule to exploit the conditional independence relationships among variables.

A Bayes network can be viewed as a knowledge representation and an inference engine that can be useful for many problems. Its advantages include explicit uncertainty characterization, representation of domain-specific knowledge in a human reasoning framework, efficient computation, quick training, easy construction, adaptability, good generalization with limited training data, and easy retraining when pruning or adding new features or new training data. These advantages make them particularly suitable for real-world applications where information can be incomplete or inaccurate. In this chapter, we discuss the application of Bayes nets to two major types of image understanding problems, namely, scene classification and object detection.

Image understanding is the highest processing level in computer vision (Sonka, Hlavac, & Boyle, 1999), where semantic information is extracted from the image, in contrast to image processing, which converts one image representation to another (e.g., by converting an intensity image to an edge map). Early successes in image understanding were limited to applications dealing with constrained environments, for example, military target recognition (Dudgeon & Lacoss, 1993), document processing (Schurmann, Bartneck, Bayer, Franke, Mandler, & Oberlander, 1992), and medical imaging (Robinson & Colchester, 1994). While image understanding in unconstrained environments remains a challenging problem, progress is being made in object detection and scene classification. Object detection deals with identifying known objects within the image, while scene classification characterizes an image into one of the known categories, for example, indoor or outdoor, city or landscape.
A Dynamic Neural Field Model of Word Learning
www.igi-global.com/chapter/dynamic-neural-field-model-word/74887?camid=4v1a