The Changing Fortunes of Pattern Recognition and Computer Vision

Rama Chellappa

PII: S0262-8856(16)30066-X
DOI: doi: 10.1016/j.imavis.2016.04.005
Reference: IMAVIS 3491

To appear in: *Image and Vision Computing*

Received date: 31 March 2016
Accepted date: 13 April 2016

Please cite this article as: Rama Chellappa, The Changing Fortunes of Pattern Recognition and Computer Vision, *Image and Vision Computing* (2016), doi: 10.1016/j.imavis.2016.04.005

The Changing Fortunes of Pattern Recognition and Computer Vision

Rama Chellappa

Department of Electrical and Computer Engineering and the Center for Automation Research, UMIACS

University of Maryland, College Park, MD 201742, USA

Abstract

As someone who been attending conferences on pattern recognition and computer vision since 1978, I have watched with interest the ups and downs of pattern recognition and computer vision areas and how they have been presented at various conferences such as PRIP, CVPR , ECCV, ICCV, ACCV, ICPR, IJCAI, AAAI, NIPS, ICASSP and ICIP. Given the recent successes of deep learning networks, it appears that the scale is tilting towards pattern recognition as is commonly understood. A good number of papers in recent vision conferences seem to push data through one or other deep learning networks and report improvements over state of the art. While one cannot argue against the remarkable (magical?) performance improvements obtained by deep learning network-based approaches, I worry that five years from now, most students in computer vision will only be aware of software that implements some deep learning network. After all, 2-D based detection and recognition problems for which the deep learning networks have shown their mettle are only a subset of the computer vision field. While enjoying the ride, I would like to caution that understanding of scene and image formation, invariants, interaction between light and matter, human vision, 3D recovery, and emerging concepts like common sense reasoning are too important to ignore for the long-term viability of the computer vision field. It will be a dream come true if we manage to integrate these computer vision concepts into deep learning networks so that more robust performance can be obtained with less data and cheaper computers.

A Brief History of Developments in Image Recognition Using Computer Vision and Neural Network-based Methods

Since the early sixties, when Robert's edge operator was introduced, computer vision researchers have been working on designing various object recognition systems. The goal has been to design an end-to-end automated system that will simply accept 2D, 3D or video inputs and spit out the class labels or identities of objects. Beginning with template matching approaches in the sixties and seventies, methods based on global and local shape descriptors were developed. In the seventies, methods based on representations such as Fourier descriptors, moments, Markov models, and statistical pattern recognizers were developed. Even in the early years, the need for making the global recognition approaches be invariant to various transformations such as scale, rotation, etc. were recognized. Unlike these global descriptors, local descriptors based on primitives such as line segments, arcs etc. were used in either structural or syntactic pattern recognition engines. For example, generative grammars of various types were designed to parse the given object contour into one of many classes. More information on these developments can be found in [1-2].