Guest Editorial Applications of Artificial Neural Networks to Image Processing

ARTIFICIAL neural network (NN) architectures have been recognized for a number of years as a powerful technology for solving real-world image processing problems. The primary purpose of this special issue is to demonstrate some recent success in solving image processing problems and hopefully to motivate other image processing researchers to utilize this technology to solve their real-world problems. Finally, it is our hope that this special issue will increase the awareness of image processing researchers to the impact of the neural network-based algorithms.

From the response to the initial call for papers, ten manuscripts have been selected for inclusion in this special issue. Eight papers have been offered as full papers and two as correspondence items. These papers covered the following major topics:

1) neural network-based algorithms for character recognition;
2) automatic target recognition using artificial neural networks;
3) object identification, classification and segmentation;
4) image prediction and compression.

The first paper by Garris *et al.* provides an overview of the NN-based approaches to optical character recognition (OCR). In this paper the authors present results from the evaluation of several NN-based OCR systems. They also provide an end-to-end OCR recognition system based on an enhanced multilayer perceptron (MLP) classifier.

The next three papers deal with the topic of the automatic target recognition (ATR). The paper by Wang *et al.* proposes a new ATR classifier based on an NN architecture called the modular neural network (MNN) classifier. The MNN classifier consists of several independent neural networks trained on local features extracted from specific portions of the image. The final classification is achieved by combining the decision produced by each individual neural network by a method known as *stacked generalization*. This NN-based classifier is tested on a large set of real forward-looking infrared imagery.

Young *et al.* present a method for detecting and classifying a target from a multiresolution foveal image. In this algorithm target identification decisions are based on minimizing an energy function which is implemented by a novel multilayer Hopfield neural network. This energy function supports connections between nodes at the same level as well as interconnections between two sets of nodes at two different resolution levels. A simultaneous top-down-and-bottom-up search is implemented by using a multilayer Hopfield neural network to minimize this energy function.

The paper by Principe *et al.* focuses on target detection in synthetic aperture radar (SAR) imagery using linear and nonlinear adaptive neural networks. They propose to use a quadratic gamma detector (QGD) which is a nonparametrically trained classifier based on local image intensity. The linear processing element (output adder) of the QGD is further extended with a set of nonlinear processing elements to combine the input feature elements. Thus, they implemented a QGD as an MLP which was trained with a cost function based on a mixed norm weighting the false alarms and the missed detections differently.

The paper by Phillips presents a face identification algorithm that automatically processes an unknown image by locating and identifying the face. His algorithm is based on designing a net of matching pursuit filters optimized for face detection and identification. For identification, the filters find features that differentiate among faces, whereas, for detection, the filters encode the similarities among faces. This algorithm has been evaluated on three sets of images. The first set was images from the FERET data base (a well-known benchmarking data set for face recognition). The second set was infrared and visible images of the same people. This demonstration was done to compare performance on infrared and visible images individually, and on fusing the results from both modalities. The third set was mugshot data from a law enforcement application.

Wang *et al.* present a probabilistic neural network-based technique for unsupervised quantification and segmentation of brain tissues from magnetic resonance images (MRI's). The proposed technique uses suitable statistical models for both the pixel and context images and formulates the problem in terms of model-histogram fitting and global consistency labeling. The quantification is achieved by probabilistic self-organizing mixtures and the segmentation by a probabilistic constraint relaxation network. Experimental results are presented for sequence of MRI brain scans. It is also shown that it can be applied to clinical problems such as those encountered in tissue segmentation and quantitative diagnosis.

The paper by Zhu *et al.* proposes an integrated system called road understanding neural network (RUNN) for a autonomous mobile robot to move in an outdoor road environment. The RUNN consists of two major neural network modules, a single three-layer road classification network (RCN) to identify the

Manuscript received May 21, 1998.
Publisher Item Identifier S 1057-7149(98)05309-3.

road category (straight: road, intersection or T-junction), and a two-layer road orientation network (RON) for each road category. Several design issues, including the network model, the selection of input data, the number of the hidden units and the learning problems are studied. Experimental results are presented for real scene imagery.

Finally, the last three papers focus primarily on applications of neural networks to image prediction and compression. In the paper by Wang *et al.*, a modular neural network vector predictor is proposed that is based on a mixture of expert predictors. In this algorithm five expert vector predictors are designed where each predictor is optimized for a particular class of input vectors. An integrating unit is used to select or combine the outputs of the experts in order to form the final output of the modular predictor.

The correspondence item by Karayiannis *et al.* presents an image compression system based on wavelet decomposition and vector quantization (VQ). In this paper the VQ codebook is designed by various fuzzy learning vector quantization schemes. The performance of the codebooks designed by these algorithms is tested on several images and compared with the classical Linde–Buzo–Gray (LBG) algorithm.

In the correspondence by Tzovaras and Strintzis, a nonlinear principal component analysis (NLPCA) method is combined with VQ for coding of images. The NLPCA is realized using the backpropagation neural network, while VQ is performed using learning vector quantization.

RAMA CHELLAPPA, *Guest Editor*
University of Maryland
College Park, MD

KUNIHIKO FUKUSHIMA, *Guest Editor*
Osaka University
Osaka, Japan

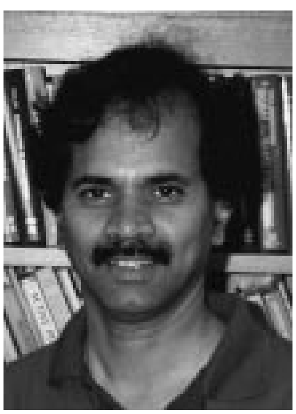
AGGELOS K. KATSAGGELOS, *Guest Editor*
Northwestern University
Evanston, IL

SUN-YUAN KUNG, *Guest Editor*
Princeton University
Princeton, NJ

YANN LECUN, *Guest Editor*
AT&T Laboratories-Research
Holmdel, NJ

NASSER M. NASRABADI, *Guest Editor*
U.S. Army Research Laboratory
Adelphi, MD

TOMAS A. POGGIO, *Guest Editor*
Massachusetts Institute of Technology
Cambridge, MA



Rama Chellappa (S'78–M'81–SM'83–F'92) received the B.E. degree (hons.) from the University of Madras, India, in 1975, the M.E. (distinction) degree from the Indian Institute of Science, Bangalore, in 1977, and the M.S.E.E. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1978 and 1981, respectively.

Since 1991, he has been a Professor of Electrical Engineering and an Affiliate Professor of Computer Science at the University of Maryland, College Park. He is also affiliated with the Center for Automation Research (Associate Director) and the Institute for Advanced Computer Studies. Prior to joining the University of Maryland, he was an Associate Professor and Director of the Signal and Image Processing Institute, University of Southern California (USC), Los Angeles. His current research interests are image compression, automatic target recognition from stationary and moving platforms, surveillance and monitoring, automatic design of vision algorithms, synthetic aperture radar image understanding, and commercial applications of image processing and understanding. Over the last 16 years he has published numerous book chapters and peer-reviewed journal papers. Several of his journal papers have been reproduced in *Collected Works* (IEEE Press, IEEE Computer Society Press, and MIT Press). He has edited a collection of papers on digital image processing (IEEE Computer Society Press, 1992), co-authored a research monograph on artificial neural networks for computer vision, with Y. T. Zhou, (Springer-Verlag, 1994), and co-edited *Markov Random Fields*, with A. K. Jain (Academic, 1996).

Dr. Chellappa has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON IMAGE PROCESSING, and IEEE TRANSACTIONS ON NEURAL NETWORKS, and as a co-Editor-in-Chief of *Graphical Models and Image Processing*. He currently serves as an Associate Editor of IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. He has served as a General and Technical program Chair for several IEEE International and National Conferences and Workshops. He has received several awards, including the 1985 NSF Presidential Young Investigator Award, a 1985 IBM Faculty Development Award, the 1991 Excellence in Teaching Award from the School of Engineering, USC, and the 1992 Best Industry Related Paper Award from the International Association of Pattern Recognition (with Q. Zheng). He was recently elected a Distinguished Faculty Research Fellow (1996–1998) at the University of Maryland. He is a Fellow of the International Association for Pattern Recognition.

Kunihiko Fukushima received the B.Eng. degree in electronics in 1958, and the Ph.D. degree in electrical engineering in 1966 from Kyoto University, Japan.

He is a Professor of the Department of Systems and Human Science, Graduate School of Engineering Science, Osaka University, Japan, a position he has held since 1989. Prior to his Professorship, he was a Senior Research Scientist at the NHK Science and Technical Research Laboratories. He is a pioneer in the field of neural networks and has been engaged in modeling neural networks of the brain since 1965. His special interests lie in the mechanisms for visual information processing, selective attention, learning, and memory. He invented the "Neocognitron" for deformation invariant pattern recognition, and the "Selective Attention Model," which can recognize and segment individual objects in a complex visual scene.

Dr. Fukushima is the author of many books on neural networks, including *Neural Networks and Information Processing*, *Neural Networks and Self-Organization*, and *Physiology and Bionics of the Visual System*. He was the founding President of the Japanese Neural Network Society and is a founding member on the Board of Governors of the International Neural Network Society.

Aggelos K. Katsaggelos (S'80–M'85–SM'92–F'98) received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1979, and the M.S. and Ph.D. degrees, both in electrical engineering, from the Georgia Institute of Technology, Atlanta, in 1981 and 1985, respectively.

In 1985, he joined the Department of Electrical Engineering and Computer Science at Northwestern University, Evanston, IL, where he is currently Professor, holding the Ameritech Chair of Information Technology. During the 1986–1987 academic, year he was an Assistant Professor at Polytechnic University, Department of Electrical Engineering and Computer Science, Brooklyn, NY. His current research interests include image recovery, processing of moving images (motion estimation, enhancement, very low bit rate compression), computational vision, and multimedia signal processing.

Dr. Katsaggelos is a member of the Steering Committees of the IEEE TRANSACTIONS ON MEDICAL IMAGING and the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE Technical Committees on Visual Signal Processing and Communications, Image and Multidimensional Signal Processing, and Multimedia Signal Processing, the Technical Chamber of Commerce of Greece, and Sigma Xi. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1990–1992), an area editor for the journal *Graphical Models and Image Processing* (1992–1995), and is currently the Editor-in-Chief of the *IEEE Signal Processing Magazine*. He is the editor of *Digital Image Restoration* (Springer-Verlag, 1991), and co-author of *Rate-Distortion Based Video Compression* (Kluwer, 1997). He has served as the General Chairman of the 1994 Visual Communications and Image Processing Conference (Chicago, IL), and he will serve as the Technical Program Co-chair of the 1998 IEEE International Conference on Image Processing (Chicago, IL). He is an Ameritech Fellow, a member of the Associate Staff, Department of Medicine, Evanston Hospital, and a member of SPIE.

Sun-Yuan Kung (S'74–M'78–SM'84–F'88) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA.

In 1974, he was an Associate Engineer at Amdahl Corporation, Sunnyvale, CA. From 1977 to 1987, he was a Professor of Electrical Engineering—Systems of the University of Southern California, Los Angeles. In 1984, he was a Visiting Professor of the Stanford University and Delft University of Technology, The Netherlands. Since 1987, he has been a Professor of Electrical Engineering at Princeton University, Princeton, NJ. His research interests include VLSI array processors, neural networks and multimedia signal processing. He has authored more than 300 technical publications and three books, *VLSI Array Processors* (Prentice-Hall, 1988), *Digital Neural Networks* (Prentice–Hall, 1993), and *Principal Component Neural Networks* (Wiley, 1996).

Dr. Kung has served as an Editor-In-Chief of the *Journal of VLSI Signal Processing* (since 1990) and associate editor on several IEEE Transactions. He was appointed as an IEEE-SP Distinguished Lecturer during 1994–1996. He received the 1996 IEEE Signal Processing Society's Best Paper Award, and the 1992 IEEE Signal Processing Society's Technical Achievement Award.

Yann LeCun (S'87–M'90) was born near Paris, France, in 1960. He received the Diplôme d'Ingénieur from the Ecole Supérieure d'Ingénieur en Electrotechnique et Electronique, Paris, in 1983, and the Ph.D. in computer science from the Université Pierre et Marie Curie, Paris, in 1987, during which he proposed an early version of the backpropagation learning algorithm for neural networks.

He then joined the Department of Computer Science at the University of Toronto, Toronto, Ont., Canada, as a Research Associate. In 1988, he joined the Adaptive Systems Research Department, AT&T Bell Laboratories, Holmdel, NJ, where he worked on neural networks, machine learning, and handwriting recognition. Following AT&T's split up in 1996, he became head of the Image Processing Services Research Department, AT&T Labs.-Research. In addition to the above topics, his current interests include image-based digital libraries, video-based user interfaces, image compression, and content-based indexing of multimedia material. He has published over 70 technical papers and book chapters on neural networks, machine learning, pattern recognition, handwriting recognition, document understanding, image processing, VLSI design, and information theory.

Dr. LeCun serves on the board of the *Machine Learning Journal*, and has served as Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS. He is general chair of the "Machines that Learn" Workshop, held every year since 1986 in Snowbird, UT. He has served as program co-chair of IJCNN 89, INNC 90, NIPS 90, 94, and 95.

Nasser M. Nasrabadi (S'80–M'84–SM'92) received the B.Sc. (Eng.) and Ph.D. degrees in electrical engineering from Imperial College of Science and Technology, University of London, London, U.K., in 1980 and 1984, respectively.

From October to December 1984, he worked for IBM (U.K.) as a Senior Programmer. During 1985 to 1986, he worked with Philips Research Laboratory, NY, as a Member of Technical Staff. From 1986 to 1991 he was an Assistant Professor in the Department of Electrical Engineering at Worcester Polytechnic Institute, Worcester, MA. From 1991 to 1996 he was an Associate Professor with the Department of Electrical and Computer Engineering, State University of New York at Buffalo. Since September 1996 he has been a Senior Research Scientist with the U.S. Army Research Laboratory, Adelphi, MD, working on image compression and automatic target recognition. His current research interests are in image and video compression, packet video, automatic target recognition, and neural networks applications to image processing. He has over 50 journal publications, appearing mostly in IEEE Transactions.

Dr. Nasrabadi has served as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and is currently an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He is a Fellow of SPIE.

Tomas A. Poggio (A'86) received the doctorate degree in theoretical physics from the University of Genoa, Italy, in 1970.

He is the Uncas and Helen Whitaker Professor in the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge. He is doing research in computational learning and vision at the MIT Center for Biological and Computational Learning, of which he is Co-Director. He is also a member of the Artificial Intelligence Laboratory. He is the author of several papers in areas ranging from psychophysics and biophysics to information processing in man and machine, artificial intelligence, machine vision, and learning.

Dr. Poggio has received several awards and is on the editorial board of a number of interdisciplinary journals. He is a Fellow of the American Association for Artificial Intelligence and of the American Academy of Arts and Sciences, and an Honorary Associate of the Neuroscience Research Program, Rockefeller University, New York.