Intelligent Systems Reference Library 117

Diego Oliva
Erik Cuevas

Advances and Applications of Optimised Algorithms in Image Processing

Springer

Intelligent Systems Reference Library

Volume 117

About this Series

The aim of this series is to publish a Reference Library, including novel advances and developments in all aspects of Intelligent Systems in an easily accessible and well structured form. The series includes reference works, handbooks, compendia, textbooks, well-structured monographs, dictionaries, and encyclopedias. It contains well integrated knowledge and current information in the field of Intelligent Systems. The series covers the theory, applications, and design methods of Intelligent Systems. Virtually all disciplines such as engineering, computer science, avionics, business, e-commerce, environment, healthcare, physics and life science are included.

More information about this series at http://www.springer.com/series/8578

Diego Oliva · Erik Cuevas

Advances and Applications of Optimised Algorithms in Image Processing

Springer

Diego Oliva
Departamento de Electrónica, CUCEI
Universidad de Guadalajara
Guadalajara, Jalisco
Mexico

and

Tecnológico de Monterrey, Campus Guadalajara
Zapopan, Jalisco
Mexico

Erik Cuevas
Departamento de Electrónica, CUCEI
Universidad de Guadalajara
Guadalajara, Jalisco
Mexico

ISSN 1868-4394 ISSN 1868-4408 (electronic)
Intelligent Systems Reference Library
ISBN 978-3-319-48549-2 ISBN 978-3-319-48550-8 (eBook)
DOI 10.1007/978-3-319-48550-8

Library of Congress Control Number: 2016955427

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my family and Gosia Kijak, you are always my support

Foreword

This book brings together and explores possibilities for combining image processing and artificial intelligence, both focused on machine learning and optimization, two relevant areas and fields in computer science. Most books have been proposed about the major topics separately, but not in conjunction, giving it a special interest. The problems addressed and described in the different chapters were selected in order to demonstrate the capabilities of optimization and machine learning to solve different issues in image processing. These problems were selected considering the degree of relevance in the field providing important cues on particular applications domains. The topics include the study of different methods for image segmentation, and more specifically detection of geometrical shapes and object recognition, where their applications in medical image processing, based on the modification of optimization algorithms with machine learning techniques, provide a new point of view. In short, the book was intended with the purpose and motivation to show that optimization and machine learning main topics are attractive alternatives for image processing technique taking advantage over other existing strategies. Complex tasks can be addressed under these approaches providing new solutions or improving the existing ones thanks to the required foundation for solving problems in specific areas and applications.

Unlike other existing books in similar areas, the book proposed here introduces to the reader the new trends using optimization approaches about the use of optimization and machine learning techniques applied to image processing. Moreover, each chapter includes comparisons and updated references that support the results obtained by the proposed approaches, at the same time that provides the reader a practical guide to go to the reference sources.

The book was designed for graduate and postgraduate education, where students can find support for reinforcing or as the basis for their consolidation or deepening of knowledge, and for researchers. Also teachers can find support for the teaching process in the areas involving machine vision or as examples related to main techniques addressed. Additionally, professionals who want to learn and explore the advances on concepts and implementation of optimization and learning-based algorithms applied image processing find in this book an excellent guide for such purpose.

The content of this book has been organized considering an introduction to machine learning an optimization. After each chapter addresses and solves selected problems in image processing. In this regard, Chaps. 1 and 2 provides respectively introductions to machine learning and optimization, where the basic and main concepts related to image processing are addressed. Chapter 3, describes the electromagnetism-like optimization (EMO) algorithm, where the appropriate modifications are addressed to work properly in image processing. Moreover, its advantages and shortcomings are also explored. Chapter 4 addresses the digital image segmentation as an optimization problem. It explains how the image segmentation is treated as an optimization problem using different objective functions. Template matching using a physical inspired algorithm is addressed in Chap. 5, where indeed, template matching is considered as an optimization problem, based on a modification of EMO and considering the use of a memory to reduce the number of call functions. Chapter 6 addresses the detection of circular shapes problem in digital images, and again focused as an optimization problem. A practical medical application is proposed in Chap. 7, where blood cell segmentation by circle detection is the problem to be solved. This chapter introduces a new objective function to measure the match between the proposed solutions and the blood cells contained in the images. Finally, Chap. 8 proposes an improvement EMO applying the concept of opposition-based electromagnetism-like optimization. This chapter analyzes a modification of EMO used as a machine learning technique to improve its performance. An important advantage of this structure is that each chapter could be read separately. Although all chapters are interconnected, Chap. 3 serves as the basis for some of them.

The concise comprehensive book on the topics addressed makes this work an important reference in image processing, which is an important area where a significant number of technologies are continuously emerging and sometimes untenable and scattered along the literature. Therefore, congratulations to authors for their diligence, oversight and dedication for assembling the topics addressed in the book. The computer vision community will be very grateful for this well-done work.

July 2016 Gonzalo Pajares

Universidad Complutense de Madrid

Preface

The use of cameras to obtain images or videos from the environment has been extended in the last years. Now these sensors are present in our lives, from cell phones to industrial, surveillance and medical applications. The tendency is to have automatic applications that can analyze the images obtained with the cameras. Such applications involve the use of image processing algorithms.

Image processing is a field in which the environment is analyzed using samples taken with a camera. The idea is to extract features that permit the identification of the objects contained in the image. To achieve this goal is necessary applying different operators that allow a correct analysis of a scene. Most of these operations are computationally expensive. On the other hand, optimization approaches are extensively used in different areas of engineering. They are used to explore complex search spaces and obtain the most appropriate solutions using an objective function. This book presents a study the uses of optimization algorithms in complex problems of image processing. The selected problems explore areas from the theory of image segmentation to the detection of complex objects in medical images. The concepts of machine learning and optimization are analyzed to provide an overview of the application of these tools in image processing.

The aim of this book is to present a study of the use of new tendencies to solve image processing problems. When we start working on those topics almost ten years ago, the related information was sparse. Now we realize that the researchers were divided and closed in their fields. On the other hand, the use of cameras was not popular then. This book presents in a practical way the task to adapt the traditional methods of a specific field to be solved using modern optimization algorithms. Moreover, in our study we notice that optimization algorithm could also be modified and hybridized with machine learning techniques. Such modifications are also included in some chapters. The reader could see that our goal is to show that exist a natural link between the image processing and optimization. To achieve this objective, the first three chapters introduce the concepts of machine learning, optimization and the optimization technique used to solve the problems. The structure of the rest of the sections is to first present an introduction to the problem to be solved and explain the basic ideas and concepts about the implementations.

The book was planned considering that, the readers could be students, researchers expert in the field and practitioners that are not completely involved with the topics.

This book has been structured so that each chapter can be read independently from the others. Chapter 1 describes the machine learning (ML). This chapter concentrates on elementary concepts of machine learning. Chapter 2 explains the theory related with global optimization (GO). Readers that are familiar with those topics may wish to skip these chapters.

In Chap. 3 the electromagnetism-like optimization (EMO) algorithm is introduced as a tool to solve complex optimization problems. The theory of physics behind the EMO operators is explained. Moreover, their pros and cons are widely analyzed, including some of the most significant modifications.

Chapter 4 presents three alternative methodologies for image segmentation considering different objective functions. The EMO algorithm is used to find the best thresholds that can segment the histogram of a digital image.

In Chap. 5 the problem template matching is introduced that consists in the detection of objects in an image using a template. Here the EMO algorithm optimizes an objective function. Moreover, improvements to reduce the number of evaluations and the convergence velocity are also explained.

Continuing with the object detection, Chap. 6 shows how EMO algorithm can be applied to detect circular shapes embedded in digital images. Meanwhile, in Chap. 7 a modified objective function is used to identify white blood cells in medical images using EMO.

Chapter 8 shows how a machine learning technique could improve the performance of an optimization algorithm without affecting its main features such as accuracy or convergence.

Writing this book was a very rewarding experience where many people were involved. We acknowledge Dr. Gonzalo Pajares for always being available to help us. We express our gratitude to Prof. Lakhmi Jain, who so warmly sustained this project. Acknowledgements also go to Dr. Thomas Ditzinger, who so kindly agreed to its appearance.

Finally, it is necessary to mention that this book is a small piece in the puzzles of image processing and optimization. We would like to encourage the reader to explore and expand the knowledge in order create their own implementations according their own necessities.

Zapopan, Mexico
Guadalajara, Mexico
July 2016

Diego Oliva
Erik Cuevas

Contents

Chapter 1
An Introduction to Machine Learning

1.1 Introduction

We already are in the era of big data. The overall amount of data is steadily growing. There are about one trillion of web pages; one hour of video is uploaded to YouTube every second, amounting to 10 years of content every day. Banks handle more than 1 M transactions per hour and has databases containing more than 2.5 petabytes (2.5×10^{15}) of information; and so on [1].

In general, we define machine learning as a set of methods that can automatically detect patterns in data, and then use the uncovered patterns to predict future data, or to perform other kinds of decision making under uncertainty. Learning means that novel knowledge is generated from observations and that this knowledge is used to achieve defined objectives. Data itself is already knowledge. But for certain applications and for human understanding, large data sets cannot directly be applied in their raw form. Learning from data means that new condensed knowledge is extracted from the large amount of information [2].

Some typical machine learning problems include, for example in bioinformatics, the analysis of large genome data sets to detect illnesses and for the development of drugs. In economics, the study of large data sets of market data can improve the behavior of decision makers. Prediction and inference can help to improve planning strategies for efficient market behavior. The analysis of share markets and stock time series can be used to learn models that allow the prediction of future developments. There are thousands of further examples that require the development of efficient data mining and machine learning techniques. Machine learning tasks vary in various kinds of ways, e.g., the type of learning task, the number of patterns, and their size [2].

D. Oliva and E. Cuevas, *Advances and Applications of Optimised Algorithms in Image Processing*, Intelligent Systems Reference Library 117,
DOI 10.1007/978-3-319-48550-8_1

1.2 Typed of Machine Learning Strategies

The Machine learning methods are usually divided into three main types: supervised, unsupervised and reinforcement learning [3]. In the predictive or supervised learning approach, the goal is to learn a mapping from inputs $\mathbf{x}$ to outputs y, given a labeled set of input-output pairs $\mathbf{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}, \mathbf{x}_i = (x_i^1, \ldots, x_i^d)$. Here $\mathbf{D}$ is called the training data set, and N represents the number of training examples.

In the simplest formulation, each training vector $\mathbf{x}$ is a d-dimensional vector, where each dimension represents a feature or attribute of $\mathbf{x}$. Similarly, y_i symbolizes the category assigned to $\mathbf{x}_i$. Such categories integrate a set defined as $y_i \in \{1, \ldots, C\}$. When y_i is categorical, the problem is known as classification and when y_i is real-valued, the problem is known as regression. Figure 1.1 shows a schematic representation of the supervised learning.

The second main method of machine learning is the unsupervised learning. In unsupervised learning, it is only necessary to provide the data $\mathbf{D} = \{\mathbf{x}_i\}_{i=1}^{N}$. Therefore, the objective of an unsupervised algorithm is to automatically find patterns from the data, which are not initially apparent. This process is sometimes called knowledge discovery. Under such conditions, this process is a much less well-defined problem, since we are not told what kinds of patterns to look for, and there is no obvious error metric to use (unlike supervised learning, where we can compare our prediction of y_i for a given $\mathbf{x}_i$ to the observed value). Figure 1.2 illustrate the process of unsupervised learning. In the figure, data are automatically classified according to their distances in two categories, such as clustering algorithms.

Reinforcement Learning is the third method of machine learning. It is less popular compared with supervised and unsupervised methods. Under, Reinforcement learning, an agent learns to behave in an unknown scenario through the signals of reward and punishment provided by a critic. Different to supervised learning, the reward and punishment signals give less information, in most of the cases only failure or success. The final objective of the agent is to maximize the total reward obtained in a complete learning episode. Figure 1.3 illustrate the process of reinforcement learning.

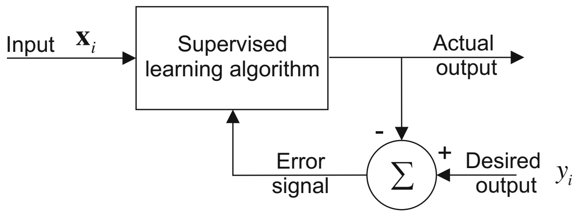

Fig. 1.1 Schematic representation of the supervised learning

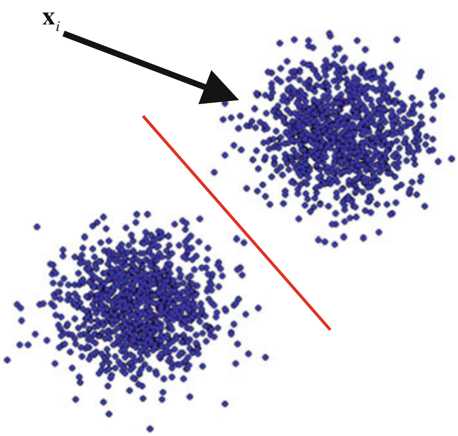

Fig. 1.2 Process of unsupervised learning. Data are automatically classified according to their distances in two categories, such as clustering algorithms

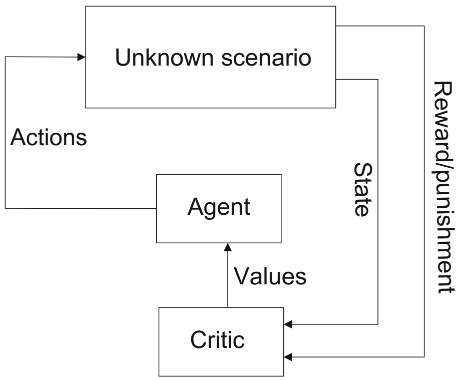

Fig. 1.3 Process of reinforcement learning

1.3 Classification

Classification considers the problem of determining categorical labels for unlabeled patterns based on observations. Let $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_N, y_N)$ be observations of d-dimensional continuous patterns, i.e., $\mathbf{x}_i \in \mathbb{R}^d$ with discrete labels $y_1, \ldots, y_N$. The objective in classification is to obtain a functional model f that allows a reasonable prediction of unknown class labels y' for a new pattern $\mathbf{x}'$. Patterns without labels should be assigned to labels of patterns that are enough similar, e.g., that are close to the target pattern in data space, that come from the same distribution, or that lie on the same side of a separating decision function. But learning from observed patterns can be difficult. Training sets can be noisy, important features may be unknown, similarities between patterns may not be easy to define, and observations may not be sufficiently described by simple distributions. Further, learning functional models can be tedious task, as classes may not be linearly separable or may be difficult to separate with simple rules or mathematical equations.

1.3.1 Nearest Neighbors

The Nearest neighbor (NN) method is the most popular method used in machine learning for classification. Its best characteristic is its simplicity. It is based on the idea that the closest patterns to a target pattern $\mathbf{x}'$, for which we seek the label, deliver useful information of its description. Based on this idea, NN assigns the class label of the majority of the k-nearest patterns in data space. Figure 1.4 show the classification process under the NN method, considering a 4-nearest approach. Analyzing Fig. 1.4, it is clear that the novel pattern $\mathbf{x}'$ will be classified as element of the class A, since most of the nearest element are of the A category.

1.4 Parametric and Non-parametric Models

The objective of a machine learning algorithm is to obtain a functional model f that allows a reasonable prediction or description of a data set. There are many ways to define such models, but the most important distinction is this: does the model have a fixed number of parameters, or does the number of parameters grow with the amount of training data? The former is called a parametric model, and the latter is called a nonparametric model. Parametric models have the advantage of often being faster to use, but the disadvantage of making stronger assumptions about the nature of the data distributions. Nonparametric models are more flexible, but often computationally intractable for large datasets. We will give examples of both kinds of models in the sections below. We focus on supervised learning for simplicity, although much of our discussion also applies to unsupervised learning. Figure 1.5 represents graphically the architectures from both approaches.

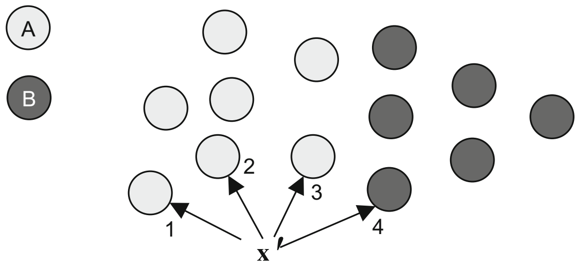

Fig. 1.4 Classification process under the NN method, considering a 4-nearest approach

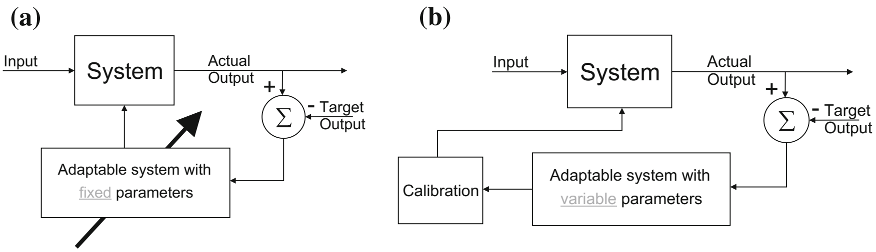

Fig. 1.5 Graphical representation of the learning process in Parametric and non-parametric models

1.5 Overfitting

The objective of learning is to obtain better predictions as outputs, being they class labels or continuous regression values. The process to know how successfully the algorithm has learnt is to compare the actual predictions with known target labels, which in fact is how the training is done in supervised learning. If we want to generalize the performance of the learning algorithm to examples that were not seen during the training process, we obviously can't test by using the same data set used in the learning stage. Therefore, it is necessary a different data, a test set, to prove the generalization ability of the learning method. This test set is used by the learning algorithm and compared with the predicted outputs produced during the learning process. In this test, the parameters obtained in the learning process are not modified.

In fact, during the learning process, there is at least as much danger in over-training as there is in under-training. The number of degrees of variability in most machine learning algorithms is huge—for a neural network there are lots of weights, and each of them can vary. This is undoubtedly more variation than there is in the function we are learning, so we need to be careful: if we train for too long, then we will overfit the data, which means that we have learnt about the noise and inaccuracies in the data as well as the actual function. Therefore, the model that we learn will be much too complicated, and won't be able to generalize.

Figure 1.6 illustrates this problem by plotting the predictions of some algorithm (as the curve) at two different points in the learning process. On the Fig. 1.6a the curve fits the overall trend of the data well (it has generalized to the underlying general function), but the training error would still not be that close to zero since it passes near, but not through, the training data. As the network continues to learn, it will eventually produce a much more complex model that has a lower training error (close to zero), meaning that it has memorized the training examples, including any noise component of them, so that is has overfitted the training data (see Fig. 1.6b).

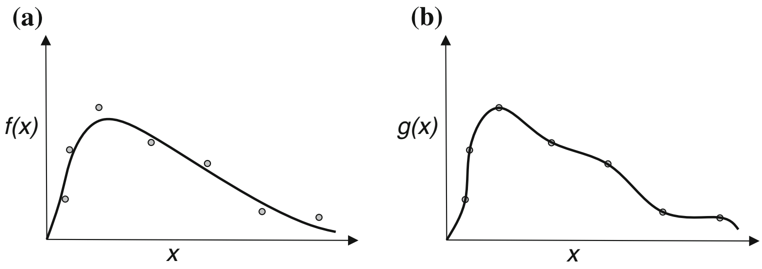

Fig. 1.6 Examples of **a** generalization and **b** overfitting

We want to stop the learning process before the algorithm overfits, which means that we need to know how well it is generalizing at each iteration. We can't use the training data for this, because we wouldn't detect overfitting, but we can't use the testing data either, because we're saving that for the final tests. So we need a third set of data to use for this purpose, which is called the validation set because we're using it to validate the learning so far. This is known as cross-validation in statistics. It is part of model selection: choosing the right parameters for the model so that it generalizes as well as possible.

1.6 The Curse of Dimensionality

The NN classifier is simple and can work quite well, when it is given a representative distance metric and an enough training data. In fact, it can be shown that the NN classifier can come within a factor of 2 of the best possible performance if $N \to \infty$.

However, the main problem with NN classifiers is that they do not work well with high dimensional data $\mathbf{x}$. The poor performance in high dimensional settings is due to the curse of dimensionality.

To explain the curse, we give a simple example. Consider applying a NN classifier to data where the inputs are uniformly distributed in the d-dimensional unit cube. Suppose we estimate the density of class labels around a test point $\mathbf{x}'$ by "growing" a hyper-cube around $\mathbf{x}'$ until it contains a desired fraction F of the data points. The expected edge length of this cube will be $e_d(F) = F^{1/d}$. If d = 10 and we want to compute our estimate on 10 % of the data, we have $e_{10}(0.1) = 0.8$, so we need to extend the cube 80 % along each dimension around $\mathbf{x}'$. Even if we only use 1 % of the data, we find $e_{10}(0.01) = 0.63$, see Fig. 1.7. Since the entire range of the data is only 1 along each dimension, we see that the method is no longer very local, despite the name "nearest neighbor". The trouble with looking at neighbors that are so far away is that they may not be good predictors about the behavior of the input-output function at a given point.