Abstract
The report includes some research activities carried out by the Computer Vision and Pattern Recognition group of the Department of Science and Technology, University Parthenope of Naples (cvprlab.uniparthenope.it) in 2015-2016. The activities cover different aspects related to machine learning and computer vision, and are carried out in the context of a variety of applied projects, where results and know-how from those activities are exploited.

Keywords: Computer Vision, Machine Learning, Pattern Recognition.

1. Machine Learning

Research activity of the group in this area concerns the design of computational models that take advantage of soft computing techniques and their adoption for different application problems, summarized in the following.

Semantic graphs and machine learning techniques have been adopted for web pages categorization [2]. A semantic graph provides a compact and structured representation of the concepts present in a document in order to take into account the semantic information. It allows determining a map of the semantic areas contained in the document and their relationships w.r.t. a particular concept or term. The document categorization is accomplished by a machine learning supervised (SVM) or unsupervised (SOM) technique.

A methodology to solar photovoltaic power forecasting based on a neural network (NN), trained in a Bayesian framework, has been proposed [9]. Here, an ahead prediction Multi-Layer Perceptron NN is used, whose parameters are estimated by a probabilistic Bayesian learning technique. The Bayesian framework allows obtaining the confidence intervals and to estimate the error bars of the NN predictions.
Social networks have gained great attention in the last decade, causing the generation of massive data characterized by computational issues, such as size, noise, and reliability, that make social network data complex to analyze manually. A recent software architecture [11], named lambda-architecture, modified with the introduction of machine learning components, allows to perform sentiment analysis on big data streams, as those provided by the Twitter social network. Moreover, a method to represent and synthesize the information content of Twitter conversations in form of semantic maps has been proposed [8], from which the main topics and the main orientations of Tweeters may easily be read. After a preliminary grouping of tweets in conversations, relevant keywords and Named Entities are extracted, disambiguated and clustered. Annotations have been made using extensive knowledge bases and state-of-the-art techniques from natural language processing and machine learning.

Recent researches concern small motifs that represent the local structure of networks and play a relevant role in determining their dynamics and evolution. Overrepresented motifs are the low-level building blocks of networks, useful to explain many of their properties. Triadic motifs have been found first partitioning a network by strength of connections and then analyzing the partitions separately. One of the main applications is studying motifs in the World Trade Web (WTW), the directed graph connecting world Countries with trade relationships, with the aim of finding its topological characterization and of isolating the key factors underlying its evolution. In [18], the WTW has been split based on the graph weights, to highlight structural differences between the big players in terms of trade volumes and the rest of the world.

A new scheme for the reconstruction of lost data in multimedia streaming, based on a compressive sampling technique, has been introduced [6]. The audio streaming data are encapsulated in different packets, at the sender, by using an interleaving technique. The compressive sampling technique is used to recover audio information in case of lost packets, at the receiver, and it has been corroborated by a fast dictionary learning approach, with a sparsifying basis provided by a greedy adaptive dictionary learning algorithm [7].

The identification of causes of genetic diseases is carried out by approaches with increasing complexity, producing large amounts of data that needs the support of statistical and computational methods to be correctly processed. An overview of statistical and computational methods has been provided [4], paying attention to methods for sequence analysis and complex diseases. Moreover, multivariate logistic regression-based researches were conducted for identifying single nucleotides in lipid-related genes associated with Familial Combined Hyperlipidemia [19], a polygenic and multifactorial disease, characterized by a variable phenotype showing increased levels of triglycerides and/or cholesterol.
Dimensionality reduction methods are preprocessing techniques used for coping with high dimensionality. They project the original data set of dimensionality $N$, without information loss, onto a lower $M$-dimensional submanifold. A review of the state-of-the-art methods of intrinsic dimension estimation, whose aim is to discover in advance the unknown value of $M$, has been provided [5], underlining the recent advances and the open problems. In the case of clustering, dimensionality reduction through unsupervised feature selection has been achieved based on the hybrid notion of rough fuzzy sets [11], that comes from the combination of the two models of uncertainty provided by rough set theory and fuzzy logic, and helps to exploit, at the same time, properties like coarseness and vagueness.

The recognition of arm movements performed by jugglers during their exercises is approached based on information on the arm orientation provided by Euler Angles, measured with a cheap sensor [3]. The recognition is obtained through a linear SVM after a feature extraction phase in which the dynamics of the system is reconstructed and three Correlation Dimensions, each associated to a given Euler Angle, are estimated.

Deep Convolutional Neural Networks (CNN) brought revolution to various challenging tasks, mainly in computer vision. However, their model designing still requires attention to reduce number of learnable parameters, with no meaningful reduction in performance. An investigation concerning to what extend CNN may take advantage of pyramid structure typical of biological neurons is provided in [22]. A generalized statement over convolutional layers from input till fully connected layer is introduced, that helps further understanding and designing a successful deep network. It reduces ambiguity, number of parameters, and their disk size, without degrading overall accuracy.

2. Computer Vision

Research activity in the field of computer vision is devoted to the analysis, design and implementation of algorithms for the detection, tracking and recognition of objects and of their activities in motion sequences.

Recent activities are focused on methods for Scene Background Modeling and Initialization (SBMI) from image sequences, that are the prerequisite for many applications, ranging from video surveillance to computational photography [16, 17]. Two datasets have been made publicly available, in order to advance the development of algorithms and methods for scene background modeling through objective evaluation on common data, providing rigorous and comprehensive academic benchmarking for testing and ranking existing and new algorithms for the problem. The SBI dataset (sbmi2015.na.icar.cnr.it/SBIdataset.html) has been adopted by participants to the SBMI workshop, organized in conjunction with ICIAP2015, and by au-
A feature independent method for **temporal segmentation** via unsupervised learning, called GRUNTS, has been proposed [1]. It employs graphs, through skeletonization and polygonal approximation, to represent objects in each frame, and graph matching to efficiently compute a Frame Kernel Matrix able to encode the similarities between frames. Temporal segmentation is adopted for human action recognition, using the Aligned Cluster Analysis as unsupervised learning strategy.

Previously proposed frameworks for the detection and tracking in real-time of unknown objects have been adopted for the 2015 **Visual Object Tracking** Challenge [15]. BDF (Best Displacement Flow) is a short-term tracking algorithm in which a set of local tracker responses are robustly combined to track the object. MatFlow (Matrioska BDF) enhances the performance of the first version of Matrioska, with response given by the BDF tracker.

Even though **iris recognition** is one of the most promising fields in biometrics, there are only few research works addressing it by machine learning techniques. A survey focusing on iris recognition has been proposed [10], taking specifically into account the kind of features used to code the iris pattern that significantly influence the methods complexity and performance.

A **person re-identification** non-learning based approach has been proposed [13], that uses symmetry principles, as well as structural relations among salient features. The re-identification problem is formulated as a graph matching problem, where each person is represented by a graph aimed not only at rejecting erroneous matches, but also at selecting additional useful ones. The problem has also been approached through a strict 3D pyramidal NN model based on convolutional neural networks, using the concept of pyramid images [14].

A **human action recognition** method is reported in [21], where pose representation is based on the contour points of the human silhouette and actions are learned by a strict 3d pyramidal NN (3DPyraNet) model which is based on CNNs and the image pyramids concept. 3DPyraNet extracts features from both spatial and temporal dimensions by keeping biological structure; thereby, it is capable to capture the motion information encoded in multiple adjacent frames. It proves to preserve the spatial topology of the input image and presents a simple connection scheme with lower computational and memory costs compared to other NNs.

### References


