



Incremental supervised learning: algorithms and applications in pattern recognition

Aida Chefrour^{1,2}

Received: 22 June 2018 / Revised: 16 January 2019 / Accepted: 17 January 2019 / Published online: 2 February 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

The most effective well-known methods in the context of static machine learning offer no alternative to evolution and dynamic adaptation to integrate new data or to restructure problems already partially learned. In this area, the incremental learning represents an interesting alternative and constitutes an open research field, becoming one of the major concerns of the machine learning and classification community. In this paper, we study incremental supervised learning techniques and their applications, especially in the field of pattern recognition. This article presents an overview of the main concepts and supervised algorithms of incremental learning, including a synthesis of research studies done in this field and focusing on neural networks, decision trees and support vector machines.

Keywords Machine learning · Incremental clustering · Pattern recognition · Neural network · Decision tree · SVM

1 Introduction

In the latest years, we observe a strong use of the classification methods in many application fields such as machine learning and data mining. These two areas are limited by the availability of training data. Indeed, if the data provided in the system is not representative of the problem to model, the answer of the system will not be reliable because the module of learning of this system will not be able to provide a model generalizing reality. The problem is that such a training set is not always available; it requires that the system is able to use and learn new data and information presented to it later to improve its performance with preservation of previously acquired knowledge, this is named incremental learning.

However, most methods found effective in the context of static learning do not propose any alternative of evolutionary and dynamic adaptation to integrate new knowledge, new data or to restructure problems already partially learned. In this case, a classical approach is to merge the old and new

data and to proceed a complete re-training of the system. The difficulties of such an approach are, of course, the need for having the old training data, in addition to prohibitory training times and the need for redefining the new parameters of the adopted model. Thus, an incremental learning system, able to learn new information (new data, new classes ...), without forgetting the previously acquired knowledge and without obligation to relearn the already learned data, becomes a very interesting alternative.

Moreover, the static methods cannot treat the large data sets that can be analyzed to find patterns, trends, requiring heavy computational power and memory. The deep learning is one technique that can be used for data analysis so that able to help us find abstract patterns in Big Data. The application of the Deep Learning to Big Data can find unknown and useful patterns that were impossible so far. With the help of Deep Learning, Artificial Intelligence is getting smart.

The incremental learning represents one of the major concerns of the community of machine learning and constitutes an open research area which is the subject of several types of work. Incremental methods provide solutions to process the learning data sequentially, using subsets of the training dataset. Incremental learning is a machine learning paradigm where the learning process takes place whenever new example(s) emerge and adjusts what has been learned according to the previous example(s) [11].

✉ Aida Chefrour
aida_chefrou@yaho.fr

¹ LISCO Laboratory, Computer Science Department, Badji Mokhtar-Annaba University, P.O. Box 12, 23000 Annaba, Algeria

² Computer Science Department, Mohamed Cherif Messaadia University, 41000 Souk Ahras, Algeria

In our research, we are interested in incremental supervised learning, which represents a class of machine learning techniques for which the output is known for all inputs and the system training algorithm uses the error to guide the training. Each example of the training set being is a couple (input, output). Supervised learning field includes classification methods (the output is a class) and regression methods (the output is a real number).

In our study, we provide an overview of incremental supervised algorithms and deep learning studies, commenting briefly on their application on pattern recognition since: (1) the domain of pattern recognition has become today one of the major areas in which more and more researchers are working and (2) there is little work that includes the application of incremental supervised learning in this area compared with incremental clustering.

The remainder of this article is organized as follows: in Sect. 2, we start with definitions and concepts related to the incremental learning. The principal incremental supervised learning algorithms are detailed in Sect. 3. We present the most recent works using these algorithms in the field of pattern recognition in Sect. 4 with a synthesis of all of this work in Sect. 5. Section 6 includes a conclusion and proposals for further research.

2 Basic concepts and definitions

Machine learning consists of automatically extracting of knowledge, rules, and models from available data (images, sounds,). It is a very large field representing a well-established discipline that had been investigated in the last decades by many researchers.

Several definitions have been proposed for incremental learning during the last few years. The most preferred and most synthesis is [1]: (1) in incremental learning, the training set is not fully available at the beginning of the learning process as in batch learning. Data can arrive at any time and the hypothesis concept has to be updated if necessary to capture the classes, (2) incremental learning requires an algorithm that is capable of learning from new data that may introduce new concept classes while retaining the previously acquired knowledge without requiring access to old datasets.

In incremental learning, the system must be able to answer the “stability/plasticity” dilemma:

- *Stability* The capacity to not forget the already learned data. A system of training is completely stable if it will learn only once the data and will retain the knowledge acquired without being able to learn from the new data;
- *Plasticity* The capacity to assimilate new data. A system of training is completely plastic if it will able to learn

from the new data but without forgetting those previously learned.

The objective of incremental learning is to propose a good “stability/plasticity” compromise by respecting the four following criteria [1] (as shown in Fig. 1):

- it should be able to learn additional information from new data (plasticity);
- it should not require access to the original data, used to train the existing classifier;
- it should preserve previously acquire knowledge it should not suffer from significant loss of originally learned knowledge (stability);
- it should be able to accommodate new classes that may be introduced with new data;

Thus, an incremental learning system can learn additional information from new data without having access to previously available data and without requiring any relearning of the system on the old and the new training data. The incremental learning is a type of training; we can say that a supervised or unsupervised learning algorithm can be incremental or not incremental.

3 Principal incremental supervised learning algorithms

Several incremental learning algorithms are proposed in the literature. Among these methods, some are not truly incremental because at least 1 of the 4 previously mentioned criteria is violated.

We can distinguish two main types of incremental learning algorithms (adaptive and evolving learning) according to their corresponding systems [1]:

- *Adaptive systems* For parameter incremental learning, the structure of the system is fixed and initialized at the beginning of the learning process and the system parameters are learned incrementally according to newly available data;

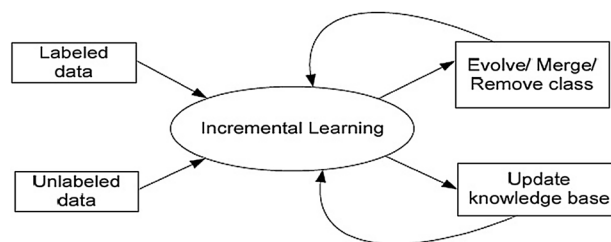


Fig. 1 Working on an incremental learning [15]

- *Evolving systems* Are more flexible than adaptive systems since they allow structure incremental learning which modifies their structure (by adding new classes if necessary).

In our study, we focus on the incremental supervised learning algorithms: neural networks, decision tree, SVM...

3.1 Decision trees

A decision tree is a classifier (the result of a training process of a learning algorithm on training data) which can divide a population of individuals into homogeneous groups according to discriminating attributes. It allows making predictions based on known data on the problem by reduction, level by level, the domain solutions. Each internal node of a decision tree is attributed discrimination, of the elements to classify, that divides these elements seamlessly between the son of a different node. The branches connecting a node to represent his son discriminant values of the attribute node. Finally, the leaves of a tree decision are his predictions for the data to classify.

It is a classifier which has the advantage of being readable and allows for analysis to determine the discriminative pairs (attribute, value) among a very large number of attributes and values. The main idea of decision trees is to recursively divide, and efficiently as possible, the examples of the training set by tests defined using attributes until you get sub-sets which contain only examples belonging to the same class. The process of building a decision tree is the learning, it can be done in different ways depending on various algorithms.

Among the algorithms, the best known and most commonly used is Quinlan's ID3 algorithm. The learning phase of decision trees is commonly divided into three stages: the tree growth (Separation into branches, selecting an attribute to partition the whole training set), stopping and pruning. These phases vary according to the decision tree building algorithm. The criterion of tree branches separation (choice of the most discriminating attribute) is the most important measure to categorizing the different algorithms to build decision trees [4]. The ID3 is an efficient learning algorithm that creates decision trees to represent classification rules. It uses concepts of entropy and information gain.

Decision trees are semi-related to incremental learning at all. They are built incrementally using a recursive procedure (ID3, C4.5...). A tree is built incrementally starting from a root node, but if you insert or remove only one training example, you need to rebuild the tree from scratch because the procedure (e.g. ID3) takes into account all the training examples at each iteration to calculate the entropy of each attribute, then its information gain.

The most known incremental decision tree algorithms are:

- ID4 is the first algorithm proposed by Schlimmer and Fisher [34]. It incorporates data incrementally. However, certain concepts were unlearnable, because ID4 discards subtrees when a new test is chosen for a node;
- ID5 developed by Utgoff [37], it did not discard sub-trees but also did not guarantee that it would produce the same tree as ID3;
- ID5R is an algorithm that performs learning by induction from examples in an incremental manner. This method generates the same decision tree found by ID3 for the same set of training [32];
- ITI is an efficient method for incrementally inducing decision trees. The same tree is produced for the training dataset. The method consists of storing the statistics in the leaves, which allows a restructuring of the tree during the arrival of new examples [32];
- VFDT (Very Fast Decision Tree) is based on the calculation of information gain for selecting the attributes. The VFDT learner reduces training time for large incremental datasets by sub-sampling the incoming data stream [32].

We have chosen to detail ID5R, one of the most used incremental decision trees. The reason behind choosing the ID5R is their embedded flexibility regarding a level of granularity and their robusted.

The structure of the tree induced by ID5R contains [37]:

- a leaf node (or answer node) that contains: a class name, and the set of instance descriptions at the node belonging to the class;
- a non-leaf node (or decision node) that contains: an attribute test, with a branch to another decision tree for each possible value of the attribute, the positive and negative counts for each possible value of the attribute, and the set of non-test attributes at the node, each with positive and negative counts for each possible value of the attribute.

The essential difference between ID5R and ID3 is that in ID5R, the examples are kept in the tree. This makes it possible to obtain a capacity of generalization without losing the specificity of the examples. The learning algorithm used by ID5R is the following:

- if the tree is empty, then define it as the unexpanded form, setting the class name to the class of the instance, and the set of instances to the singleton set containing the instance;
- otherwise, if the tree is in unexpanded form and the instance is from the same class, then add the instance to the set of instances kept in the node;
- otherwise;

- if the tree is in unexpanded form, then expand it to one level, choosing the test attribute for the root arbitrarily;
- for the test attribute and all no test attributes at the current node, update the count of positive or negative instances for the value of that attribute in the training instance;
- if the current node contains an attribute test that does not have the lowest E-score, then
 - restructure the tree an attribute with the lowest E-score will be at the root;
 - recursively repeat the test attribute in each subtree except the one that will be updated in step3 (recursively update);
 - recursively update the decision tree below the current decision node along the branch for the value of the test attribute that occurs in the instance description. Grow the branch if necessary.

3.2 Neural networks

An artificial neural network is a model inspired by biological neurons. A preliminary definition of a neural network is given [10] in his course package as follows: “neural network” is an interconnected assembly of simple processing elements, units or nodes, whose functionality is loosely based on the animal neuron. The processing ability of the network is stored in the inter-unit connection strengths, or weights, obtained by a process of adaptation to or learning from, a set of training patterns.

Each neuron makes a balanced sum of its entries and turns over a value based on the function of activation. This value can be used either like one of the entries of a new layer of neurons or as a result which is up to the use to interpret. The training phase of neural networks consists of updating the weights associated with the input neurons. Incremental versions of neural networks were developed, the most referred is:

Learn++ [27]: is an incremental learning algorithm using neural networks inspired by the AdaBoost (Adaptive Boosting). It is based on the principle of weak classifiers combination to make a decision. The system trains several classifiers on several subsets of the training set. The difficulties of this algorithm are the creation of training subsets and the combination of these classifiers.

3.3 Support vector machine (SVM)

Support vector machine or SVM is a supervised learning technique introduced by Vapnik in 1996, which causes much interest for its good performances in a large range of practical applications.

It allows making a binary decision, i.e. to carry out the classification of a form in a decision space containing

a maximum of two. The basic idea of this approach is to separate the training data, by seeking the separator which maximizes the margin (to choose the hyperplane which maximizes the minimal distance to the examples of training). In other words, the objective is to find the maximization criterion of the margin which gives it an excellent power of generalization. The principal motivation of the use of SVM is the research of the probabilistic limits which reduce to the maximum the classification error. The goal of such a learning technique is to model the probabilities of membership of an object to a class, but this approach consists in determining the borders between the discriminating classes and surfaces. The data of the whole of training closest to the hyperplane which separates the classes are called: vectors of support. There are two versions of the SVM: the linear SVM whose vectors of supports make it possible to discriminate the classes by defining a linear border of the decision and nonlinear SVM adapted to the classes whose separation cannot be linear without transformation of the space of representation (using kernel functions). Numerous SVM versions have been developed, we mention:

The first work was done by Syed et al. [35] who proposed to learn new data, do keep in addition to these, the support vectors identified during the earlier phases of learning without retaining the data themselves since support vector perfectly summarize the data already learned.

Ruping [31] proposes an incremental SVM algorithm that aims to manage all support vectors incrementally. When a new training data is misclassified or inside margin, the separator is recalculated based on support vector and this new point.

Ralaivola and dAlché Buc [29] propose an incremental version of SVM's. They see SVM as a combination of experts, where each expert is a support vector. The kernel used is a Gaussian kernel or other kernel based on the notion of the neighborhood, the area of influence of a support vector is significant only in a neighborhood close to him. Upon the arrival of a new instance, it is not necessary to question the set of all support vectors, but only those whose influence in another example is the strongest.

In 2001, Cauwenberghs and Poggio [3] designed an exact on-line algorithm for incremental learning SVM, which updates the decision function parameters when adding or deleting one vector at a time.

A proximal SVM (PSVM) proposed by Fung and Mangasarian [8]. The basic idea is to use PSVM hyper-plane border examples around this hyperplane will be retained. Upon the arrival of a new example, the authors calculate the distance between the hyperplane and the new border, if it is close, it will be added (add new examples) and removes old examples, which makes incremental SVM.

In 2003, Diehl and Cauwenberghs [6] improve the previous work and presented a framework for exact SVM

incremental learning, adaptation and optimization, in order to simplify the model selection by perturbing the SVM solution when changing kernel parameters and doing regularization.

Loosli [19] shows that SVM has the ability to handle large databases online. This study is based on the enrichment of the learning initial bases by introducing variability models applied to these bases. Require that the SVM to learn new examples follow the same distributions as the examples already learned because the changes introduced by the new instance are local.

We synthesize the principal incremental supervised learning algorithms described above in Table 1, we choose a set of criteria used usually in any comparison between the classifiers:

4 Incremental learning and pattern recognition

In the last years, we have noticed that there are many applications in the field of pattern recognition using incremental learning. We now discuss some of the most common incremental supervised pattern recognition approaches.

Déniz et al. [5] propose an algorithm called: IRDB (Incremental Refinement of Decision Boundaries). The main characteristic of this algorithm is the use of a combination of classification results and its incremental nature, in contrast to a practical systems simple combination rules which are used as a combination rule: the mean, the maximum or the majority vote rules. IRDB use two classifiers: the Nearest Neighbor (NN) classifier and the SVM (Support Vector Machines) classifier, with radial basis function kernel.

Mańdziuk and Shastri [22] propose a new approach ICL (Incremental Class Learning) which is a supervised learning procedure for neural networks. This approach attempts to address the catastrophic interference problem and at the

same time offers a learning framework that promotes the sharing of previously learned knowledge structures.

Toh and Ozawa [36] propose an intelligent face recognition system which contains two features: one-pass incremental learning and automatic generation of training data. They use a neural network called Resource Allocating Network with Long-Term Memory (RANLTM), to realize efficient incremental learning without suffering from serious forgetting. In the face detection system, face localization is conducted based on the information about skin color and edges in the first stage. Then, facial features are searched for within the localized regions using a RAN, and the selected features are used for in the construction of face candidates. After the face detection, the face candidates are classified using RAN-LTM. The recognition performance evaluation shows that accuracy improves without increasing the false-positive rate even if the incremental learning proceeds. This fact suggests that incremental learning is a useful and efficient approach to face recognition tasks.

Ozawa et al. [26] propose a new approach to build adaptive face recognition systems in which a feature space and a classifier are incrementally trained. To learn a feature space incrementally, the authors adopt an extended version of Incremental Principal Component Analysis (IPCA) in which the augmentation of feature space dimensions is determined based on the accumulation ratio. When the feature space dynamically changes over the learning stages, the inputs of a neural classifier must also change in their values and the number of input variables. To adapt to the evolution of the feature space, an extended model of Resource Allocating Network (RAN) called RAN with Long-Term Memory (RAN-LTM) is adopted as a classifier, and they propose an efficient way to reconstruct RAN-LTM after updating the feature space.

Erdem et al. [7] propose a new incremental method (SVMLearn++) using an ensemble of SVMs trained with Learn++ learning rule. SVMLearn++ with two different

Table 1 Comparison between the incremental supervised learning methods

Classifier used	Decision tree	Neural networks	SVM
Type of learning	Incremental supervised learning	Incremental supervised learning (learn++, SApIN)	Incremental supervised learning
Type of input parameters	Numeric	Numeric	Numeric
Parameters to be chosen	Selection of attribute more discriminant	The rule combination of classifiers	The zone of the influence is the hyper-plane border
Type of classification	Strict	Strict	Strict
Satisfied the incremental learning?	Yes	Yes	Yes
Advantages	The simplicity of comprehension and interpretation-	Resistance to noise data	Strong mathematical foundations
Disadvantages	Some concepts are difficult	The absence of a systematic method	High calculation time

kernel functions has been tested on one real-world dataset and one benchmark datasets.

Zhao et al. [38] propose SVDU-IPCA algorithm like any incremental learning algorithm do not need to recompute the eigendecomposition from scratch. The aim of this work is to carry out an error analysis of the proposed IPCA algorithm, using a mathematical derivation. The proposed SVDU-IPCA algorithm provides the maximum error when evaluation the batch-mode PCA. It contains an easily extended to kernel space an IKPCA algorithm is also presented.

Prudent [28] works on two contributions: the first is PAO, an unsupervised neural network that uses a hybrid approach. This system was applied to the recognition of handwritten letters, it has a good performance but the PAO learning capabilities were not really exploited. The second contribution concerns the Learn++ algorithm. The prudent contribution was considered intermediate between the two ones: he proposed a new learning algorithm for the construction of a topological map to incrementally extract the data topology. This algorithm, called Incremental Growing Neural Gas (IGNG) introduces the probability notion of the existence of neurons and connections. This concept allows managing the relevance of a neuron in the arrival of a new example. He proposed another algorithm named: SApIn using data topology provided by the IGNG to sample training dataset and create several training subsets used to train several classifiers. SApIn is more efficient and better manage incremental learning than Learn++.

Ghassabeh and Moghaddam [9] apply the new IFR (Incremental Face Recognition) system based on new adaptive learning algorithms and networks. This system can be summarised in two stages: (1) image preprocessing, which include normalization, histogram equalization, mean centering and background omission and (2) adaptive LDA feature evaluation. The new IFR system was considered as a combination of a new adaptive 1/2 network in cascade with APCA network. All input images are cropped and prepared for the IFR system. Simulation results demonstrate the effectiveness of the proposed system for the adaptive estimation of feature space for online face recognition.

Huang et al. [13] employ an incremental learning algorithm to adjust a boosted strong classifier to online samples. The beginning of this algorithm is the modification of the likelihood of each category. The input of this algorithm is the observation (instance) space and the output is the prediction space. The proposed incremental learning algorithm takes Real AdaBoost with domain-partitioning weak hypotheses. It achieves the same level of classification accuracy in both offline testing set and online testing set without any offline sample and is more robust against ill conditions caused by the improper choice of online reinforcement ratio. It has two properties: the competence and stable offline estimation method based on Nave-Bayes-like factorization, and

global optimization of the central hybrid objective loss function by means of parallel updating of all weak hypotheses. The results experiments show that this incremental approach helps and performs and still maintains good generalization ability for the common environments (such as CMU frontal face testing set). It is not only extremely important for practical applications in diverse real-life environments but also theoretically meaningful in adaptive learning field.

Hulley and Marwala [14] propose a method of Incremental Learning Using Genetic Algorithm (ILUGA) using a binary SVM classifier that is trained to be strong classifiers using a genetic algorithm. Each of the classifiers is optimized using GA to find the optimal separating hyperplane. This is done by finding the best kernel and the best soft margin. The voting weights are then generated by GA using the strong classifiers.

Luo et al. [21] propose a version of incremental SVM to place recognition, which allows the memory to increase control as the system maintains achieving new data. This method consists to (a) forget the oldest support vectors to regard the newest data when updating the decision function, (b) recognize if a new database contains new information or not. Experiments results show that this method achieves recognition performances statistically equivalent to those of the batch algorithm while obtaining a memory reduction. Regardless of the algorithm complexity, updating the internal representation at every incremental stage is computationally expensive.

Ozawa et al. [25] present Chunk IPCA, an incremental learning scheme in which the extended IPCA algorithm is used for eigenspace learning and the combination of ECM (Evolving Connectionist Model) and K-Nearest Neighbor (K-NN) method is adopted as classifier learning. This learning scheme gives a new concept for pattern recognition systems: feature selection and classifier learning are simultaneously carried out online. The experimental results show that Chunk IPCA learns major eigenvectors without serious approximation error and that a designated accumulation ratio is maintained by increasing new Eigen-axes automatically.

Reddy et al. [30], an incremental action recognition approach was proposed based on the feature tree, which increases in size when additional training instances become accessible. It requires to store all the seen training instances in the form of a feature tree. The authors, propose to use in the first time, the SR-tree (sphere/rectangle) to generate the feature-tree using the features (spatiotemporal Dollar features) of the identified training examples. Instead of dividing the high-dimensional feature space by hyperplanes as is done in previous tree techniques, the feature points are organized into regions in the feature space, which are specified by the intersection of a bounding sphere and bounding rectangle. SR-tree can keep on smaller region volume to contribute disjoint regions and larger diameter to support

fast range and NN search in high-dimensional space. In the recognition phase, they extracted features from an unknown action video and classify each feature into an action category using SR-tree. The recognition of the action needs to adopt by a simple voting method by counting the labels of the features.

Zou et al. [39] propose an SVM's incremental learning algorithm based on the filtering fixed partition of the data set. The authors present “Two-class problems” algorithm and generalize it to the “Multiclass problems” algorithm. The experimental results show that the proposed incremental learning technique can greatly improve the efficiency of SVM learning and SVM incremental learning cannot only ensure the correct identification rate but also speed up the training process.

Almaksour [1] propose Evolve++, an incremental approach to learning classifiers based on first-order Takagi–Sugeno (TS) fuzzy inference systems. This approach includes, on the one hand, the adaptation of linear consequences of the fuzzy rules using the recursive least-squares method, and, on the other hand, incremental learning of the antecedent of these rules in order to modify the membership functions according to the evolution of data density in the input space. With this approach, the recognition system is able to learn new forms from very limited data, it can further adapt and improve for each new example available. The proposed method, Evolve++ solves the problems of instability incremental learning of such systems through a global learning paradigm where the premises and conclusions of rules are learned in SIF system synergy and not independently. In the second contribution, they propose the automatic generation of artificial data to accelerate the learning of new symbols. This generation technique is based on Sigma-lognormal theory, which proposes a new representation space of handwritten forms based on neuromuscular modeling of writing mechanism.

Mohammed et al. [23] evaluate the feasibility of SPAN learning (an incremental learning algorithm based on temporal coding for Spiking Neural Network SNN) on the MNIST database (classifying images of handwritten digit dataset). The reason in using SNN is that these last uses spikes to communicate, where data is encoded in the time of the spikes that make it be suitable for spatial–temporal data processing and the data in SPAN method is coded into the precise time of the spikes. The first stage in the learning process is to convert the images into spike patterns using the Virtual Retina VR -a software simulator that transforms image/video input into spike patterns. The generated patterns are used to train a single layer of SPANs for classification. The MNIST is a non-linear dataset which guarantees that the different spike pattern classes have more complicated inter/intraclass characteristics, i.e., patterns that belong to the same class are varied and there is overlap between different classes, giving

a clear indication of the feasibility of SPAN learning for practical applications. The results experiment encouraging considering the SPAN learning for practical temporal pattern recognition applications.

Kawewong et al. [16] propose an incremental framework for learning on indoor scene categories. This framework supports the incremental interactions with humans which provide feedback to the system for extended learning at any time. It is based on the proposed n-value self-organizing and incremental neural network (n-SOINN), which has been derived by modifying the original SOINN to be appropriate for use in scene recognition. The proposed n-SOINN-SVM performs fast incremental learning which makes it fit the framework. It offers a high accuracy on par with that of SOA method while being capable to learn additional feedback from human experts.

Molina et al. [24] propose an incremental learning ensemble algorithm using support vector machines (SVM) tackling this problem by employing multimodal MR images and a texture-based information strategy. The proposed system integrates anatomic, texture, and functional features. The dataset was preprocessed using B-Spline interpolation, bias field correction and intensity standardization. First- and second-order angular independent statistical approaches and rotation invariant local phase quantization (RI-LPQ) was used to quantify texture information. An incremental learning ensemble SVM was implemented to suit working conditions in medical applications and to improve effectiveness and robustness of the system.

Lu et al. [20] present a new recognition system for detection of abnormal human actions from camera-based real-time scene monitoring, using an incremental SVM classifier. In this system, first, feature extraction and selection are implemented based on color and texture features (appearance of the person). They segmented each body of each person into three parts (head, top, and bottom), and these last of each silhouette, color, and texture features are extracted in the video sequences and then a set with 93 features is formed; then they applied a method to reduce the features space to obtain the best set. Second, they introduce the incremental SVM technique to recognize a person because the classes (persons) are not completely known at the beginning and the classical SVM cannot treat this case.

Bai et al. [2] describe how to build an incremental structured part model for object recognition. The proposed method analyzes the global structural information and multiple local attributes of objects for object model characterization. They use segment models to represent structure nodes, which cover the local information of an object. The segments are attained through segmentation and clustering method and are used to build the segment models in terms of multiple factor fusion and multi-class SVMs. The structured part model is then assembled by interacting different

segments through a deformable configuration. Additionally, they introduce an incremental learning strategy, which learns a part model by using only a small number of training samples. Explained images with high entropies are used to update the trained model. The benefit of this method is that it apprehends the inherent connections of the semantic parts of objects and characterizes the structural relationships between them.

Liu [18] proposes a novel incremental learning framework based on deep neural networks (DNN's) to solve the task of image classification. The basic idea is to use the trained network parameters with low-resolution images to improve the initial values of network parameters for images with high resolution. There are two solutions to implement this idea: one is to use the networks with scaled filters when the size of filters in deep networks is extended by upscaling parameters from the previously trained network with lower-resolution images. The other is to add convolutional filters to the network. The advantages of this method are: that a pre-trained model can be used as the initialization of other networks to shorten the training time, the time-consuming training of large-scale datasets, e.g. ImageNet dataset, is a disturbing problem. The incremental learning method can achieve superior efficiency for training without sacrificing any accuracy and performance can slightly improve the efficiency and performance through incremental learning methods.

Zribi and Boujelbene [40] use the incremental neural networks as a tool to detect human and technical errors of diagnostic of breast cancer. This incremental method help to improve the correct diagnostic in this domain, it allows classifying a mass in the breast (benign and malignant) using a selection of the most relevant risk factors and decision making of the breast cancer diagnosis.

Han et al. [12] propose a novel Incremental Boosting CNN (IB-CNN) to integrate boosting into the CNN via an incremental boosting layer that selects discriminative neurons from the lower layer and is incrementally updated on successive mini-batches. In addition, a novel loss function that accounts for errors from both the incremental boosted classifier and individual weak classifiers was proposed to fine-tune the IB-CNN. Experimental results on four benchmark AU databases have demonstrated that the IB-CNN yields significant improvement over the traditional CNN.

Hacene et al. [11] introduce a novel incremental algorithm based on pre-trained CNNs (Convolutional Neural Networks) and associative memories to classify images. The first ones use connection weights to process images, the second one uses the existence of connections to store them efficiently. This combination of methods allows learning and processing data using very few examples, memory usage, and computational power.

Lawal and Abdulkarim [17] present an adaptive SVM for the classification of data streams. They proposed an incremental learning-model selection (IL-MS) method for SVM, where they introduce the idea of incremental k-fold cross-validation to allow the incremental tuning of the hyper-parameters of the SVM in order to guarantee the effectiveness of the SVM while exploiting newly acquired streams of training data. The IL-MS was tested on the problem of online spam email filtering and human classification in video streams.

Sarwar et al. [33] develop a Deep convolutional neural network (DCNN). For learning a new set of classes, they form a new network by reusing previously learned convolutional layers (shared from the initial part of the base network) followed by new (added) trainable convolutional layers towards the later layers of the network. The shared convolutional layers work as fixed feature extractors (learning parameters are frozen) and minimize learning overhead for a new set of classes. The error resilience property of neural network ensures that even if we freeze some of the learning parameters, the network will be able to adapt to it and learn.

5 Discussion

Based on this short and selective survey of incremental supervised techniques for pattern recognition, we make the following observations:

- most incremental supervised techniques have been tested in the area of face recognition;
- performances of these techniques are great in terms of recognition accuracy, like the study of Zribi and Boujelbene [40], where obtained recognition accuracy achieves 99.95%;
- most studies enhance in SVM and Artificial Neural Network versions compared to the decision tree;
- we remark that the current research trend is the combination of incremental learning and deep learning;
- deep learning is a technology that continues to mature and has clearly been applied to pattern recognition to great effect;
- deep learning is more usable in image recognition;
- we have identified the domain application, a dataset of each approach with the methods of comparison, these are seen in Table 2;
- Table 2 summarizes the sorted works in chronological order.

We observed in Table 2, that the combined incremental algorithms provide the good results compared to other methods like:

Table 2 General comparison of various incremental supervised approaches for pattern recognition

References	Method	Application domain	Dataset	Compared the results with	Obtained results
Déniz et al. [5]	IRDB (Incremental Refinement of Decision Boundaries) algorithm is used. It uses two classifiers: the Nearest Neighbor (NN) classifier and the SVM (Support Vector Machines) classifier, with radial basis function kernel	Face recognition	2 classes were studied: 10 sequences, one for each individual, and each one with 167 frames. The images are provided by DESEO system. All the images used in the experiments are 39 × 43 pixels in size. For each sequence, 3 images were used for training the classifier, 50 as supervised information for the algorithm and the rest for the test	–	–
Mańdziuk and Shastri [22]	An ICL (Incremental Class Learning) approach for neural networks is proposed. This approach attempts to address the catastrophic interference problem and at the same time offers a learning framework that promotes the sharing of previously learned knowledge structures	Handwritten digit recognition	The USPS CEDAR database comprises of approximately 2400 handwritten zip codes collected from pieces of mail	The K-Nearest Neighbor classifier (K-NN)	The accuracy = 93.13%
Toh and Ozawa [36]	Two features are proposed in a proposed incremental learning algorithm: one-pass incremental learning and automatic generation of training data. The neural network called Resource Allocating Network with Long-Term Memory (RAN-LTM) is used to classify the face candidates	Face recognition	Video clips of seven people: four people (2 males and 2 females) are chosen as registered persons and the other three people are non-registered persons. The online dataset consists of 654 images detected from the video clips of the first week; on the other hand, the test data set consists of 499 images detected from the clips of the second week	The results of three RNNs	–
Ozawa et al. [26]	An extended version of Incremental Principal Component Analysis (IPCA) and Resource Allocating Network with Long-Term Memory (RAN-LTM) are combined to learn a feature space and a classifier incrementally	Face recognition	A self-compiled 224 video clips, which are taken from 22 people (19 males and 3 females), are given to the face detection part, and then the detected 7582 face images are used for the performance evaluation	–	The recognition accuracy = 0.97

Table 2 (continued)

References	Method	Application domain	Dataset	Compared the results with	Obtained results
Erdem et al. [7]	A new incremental method using an ensemble of SVMs trained with Learn++ learning rule (SVMLearn++) is used. SVMLearn++ with two different kernel functions; polynomial and Gaussian	Optical character recognition	The optical character recognition (OCR) data has 10 classes with digits 0–9 and 64 attributes The Volatile Organic Compounds (VOC) dataset consists of 5 classes with 6 attributes	–	The generalization performance on OCR database improved from 78 to 80% The generalization performance on VOC database improved from 83 to 85%
Zhao et al. [38]	SVDU-IPCA algorithm is proposed to perform an error analysis of the proposed IPCA algorithm, using a mathematical derivation. It contains an easily extended to kernel space, an IKPCA algorithm	Face recognition	FERET database, ARface database	Eigenface and fisherface face recognition IPCA algorithm	The recognition accuracy = 93%
Prudent [28]	PAO, an unsupervised neural network that uses a hybrid approach. Incremental Growing Neural Gas (INGG) introduces the probability notion of the existence of neurons and connections SAPIn using data topology provided by the INGG	Recognition of handwritten letters	NIST and optdigits de l'UCI databases to validate the performance of IGNC, OCR database to test SAPIn	Learn++	–
Ghassabeh and Moghaddam [9]	A new adaptive learning algorithm for estimation of LDA feature during the incremental training process. It was considered as a combination of a new adaptive $\Sigma - I/2$ network in cascade with APCA network	Face recognition	YALE face database: contains grayscale images of 15 subjects in GIF format. They chose 5 individual subjects and considered 64 images per each subject (total 320 images) containing different illumination and different poses	–	The normalized errors = 0.121, 0.232 and 0.305
Huang et al. [13]	The proposed incremental learning algorithm takes Real AdaBoost with domain-partitioning weak hypotheses as for the basis, based on Discrete AdaBoost	Face recognition	20,000 frontal face patches of size 24×24 are collected from the FERET database. 1425 frontal face images from CMU PIE database	The combined learning approach	The detection rate = 0.97

Table 2 (continued)

References	Method	Application domain	Dataset	Compared the results with	Obtained results
Hulley and Marwala [14]	A new incremental learning approach is developed called Incremental Learning Using Genetic Algorithm (ILUGA) using a binary SVM classifier that is trained to be strong classifiers using a genetic algorithm	Optical character recognition	The OCR Dataset has 10 classes of digits 0–9 and 64 input attributes. Wine recognition database	Learn++, Learn++.MT, SVMLearn++, SVMLearn++.MT	An overall accuracy of 93% and 94%
Luo et al. [21]	A discriminative incremental learning approach to place recognition, using a version of incremental SVM, which allows controlling the memory growth as the system keeps acquiring new data	Place recognition in dynamic environments	The IDOL2 (Image Database for rObot Localization 2) database contains 24 image sequences acquired by a perspective camera, mounted on two mobile robot platforms PeopleBot Minnie and PowerBot Dumbo	Fixed-partition, error-driven, and memory-controlled techniques	The classification rate = 97%
Ozawa et al. [25]	Chunk IPCA, the extended IPCA algorithm is used for eigenspace learning and the combination of ECM (Evolving Connectionist Model) and K-Nearest Neighbor (K-NN) method is adopted as classifier learning	Pattern classification	Six standard datasets in the UCI Machine Learning Repository: Internet Advertisement, Isolet Spoken Letter Recog, MUSK, Spambase, Image Segmentation, and vowel	Fixed Eigenspace, original IPCA and Batch PCA	The classification accuracy = 88%
Reddy et al. [30]	A framework for incremental action recognition using feature-tree. It does not require intensive training (vocabulary construction and category modeling). It is implemented using disk-based SR-tree data structure	Action recognition	The KTH dataset and the IXMAS multi-view dataset for each video they extract 200 interest points and corresponding cuboids. Gradient-based feature descriptor	–	Accuracy = 90.3%
Zou et al. [39]	SVM's incremental learning algorithm based on the filtering fixed partition of the data set. "Two-class problems" algorithm is presented and then generalizes it to the "Multiclass problems" algorithm	Image recognition	Three types of data set Natural texture remote sensing images Infrared images of the plane The commonly used test data (XOR two-class testing data)	The standard SVM and the incremental SVM	The correct identification rate = 0.92

Table 2 (continued)

References	Method	Application domain	Dataset	Compared the results with	Obtained results
Almaksour [1]	Evolve++, incremental approach to learning classifiers based on fuzzy inference systems SIF Takagi–Sugeno of order 1. the automatic generation of artificial data to accelerate the learning of new symbols, based on the Sigma-lognormal theory	Recognition of handwritten gestures	SIGN'' dataset composed of 25 different gestures drawn by 11 different writers on Tablet PCs. Each writer has drawn 100 samples of each gesture. UCI databases: CoverType, PenDigits, Segmentation, Letters	A12P model two batch classification methods (KNN and MLP)	–
Mohammed et al. [23]	The proposed incremental learning algorithm SPAN is used for spiking neural networks which are based on encoding input information as precise time of spikes (Temporal coding)	Handwritten digit recognition	The MNIST handwritten digits where each image is 28×28 pixels. They use 200 sample images per digit (a total of 2000 images) for training and the network is tested on different 200 images per digit. The produced spike pattern consists of 784 spike trains	–	The average accuracy of 92% in the training set and 86.6% in the testing set
Kawewong et al. [16]	A new online incremental allows the machine to add new classes and update existing classes with new information at any time. A new incremental method fits the proposed framework well, based on the proposed n-value self-organizing and incremental neural networks (n-SOINN) which have been modified from the original SOINN	Indoor scene recognition	The MIT 67-category indoor scene datasets	RBF-SVM-k-means RBF-SVM-Random Original-SOINN-SVM	–
Molina et al. [24]	The proposed system using support vector machines (SVM) employing multimodal MR images and a texture-based information strategy and integrates anatomic, texture, and functional features, using B-Spline interpolation, bias field correction and intensity standardization	Multimodal classification of prostatic adenocarcinoma	MRImage database included MRI transversal slices with three modalities: T2-weighted, Dynamic-Contrast Enhanced plasma flow (DCE-PF) and DCE-mean transit time (DCE-MTT)	–	The average sensitivity (TPR) of 0.8446 ± 0.068 The average specificity (SPC) of 0.7806 ± 0.038

Table 2 (continued)

References	Method	Application domain	Dataset	Compared the results with	Obtained results
Lu et al. [20]	Feature extraction and selection are implemented based on color and texture features (appearance of the person). The incremental SVM classifier has been introduced to recognize a person	On-line human recognition from video surveillance	CASIA-Gait database contains a set of 93 features	The classical SVM	A correct classification rate of 98.46%
Bai et al. [2]	A generic incremental structured part-based model, which not only enables an extensive hierarchical object characterization but also allows the model to be updated with respect to new training samples, is developed	Object recognition	The Caltech-256 dataset contains 30,607 images in 256 categories, with each class containing at least 80 images. The Pascal VOC 2007 dataset consists of 9,963 images from 20 classes	PASCAL_07 Best, LLC and Su	The miss rate = 0.1 false positives per image Average precision (AP) = 40.9% for the first category, 13.1% for the second and 32.6 for the third
Liu [18]	A novel incremental learning framework based on deep neural networks (DNN's) to solve the task of image classification. It uses the trained network parameters with low-resolution images to improve the initial values of network parameters for images with high resolution	Image classification	The ImageNet dataset and the Places205 dataset	The traditional method	The accuracy rate = 88%
Zribi and Boujelbene [40]	The proposed incremental Artificial Neural Network method help to improve the correct diagnostic in this domain, it allows classifying a mass in the breast (benign and malignant) using a selection of the most relevant risk factors and decision making of the breast cancer diagnosis	Human and technical errors of diagnostic of breast cancer	Wisconsin Breast Cancer Database (WBCD) contains a sample of 410 patients compound for learning and 273 for the test sample, the network consists of an input terminal and 9 input layer neurons, 4 hidden neurons associated with a sigmoidal activation function in the output layer (malignant and benign)	The Artificial metaphlasticity Multilayer Perceptron (AMMLP) algorithm	The classification accuracy = 99.95%

Table 2 (continued)

References	Method	Application domain	Dataset	Compared the results with	Obtained results
Han et al. [12]	A novel IB-CNN was proposed to integrate boosting classification into a CNN for the application of AU recognition. An incremental boosting algorithm was developed to accumulate information from multiple batches over time. A novel loss function that accounts for errors from both the incremental strong classifier and individual weak classifiers is proposed to fine-tune the IB-CNN	Facial action unit recognition	Four benchmark AU-coded databases The CK database which contains 486 image sequences from 97 subjects The FER2015 SEMAINE database which contains 6 AUs and 31 subjects with 93,000 images The FER2015 BP4D database has 11 AUs and 41 subjects with 146,847 images The DISFA database has 12 labeled AUs and 27 subjects with 130,814 images	The traditional CNN	Average classification rate For CK database = 0.951 For SEMAINE database = 0.416 For BP4D database = 0.578 For DISFA database = 0.858
Hacene et al. [11]	The proposed algorithm combines a pre-trained CNNs (Convolutional Neural Networks) and associative memories to classify images	Image recognition	ImageNet database (containing 10 classes each distinct from the 1000 ones that were used to train the CNN) and Cifar10 database	The 1-NN (Nearest Neighbor) and 5-NN	The accuracy rate = 86.07%
Lawal and Abdulkarim [17]	An incremental learning-model selection (IL-MS) method for SVM introduced using the incremental k-fold cross-validation to allow the incremental tuning of the hyper-parameters of the SVM in order to guarantee the effectiveness of the SVM while exploiting newly acquired streams of training data	On-line spam email filtering and human action classification from videos	The spam email dataset (9324 mails) consists of two classes: legitimate emails and spam The KTH action video dataset: 600 video clips for all combinations of 25 individuals, 6 actions, and 4 scenarios	The NO-MS SVM and C-MS SVM	The average classification accuracy from 79.1 to 91.6%
Sarwar et al. [33]	A Deep convolutional neural network (DCNN) developed to accommodate a new set of classes by network sharing, without forgetting the old classes	Image classification	The ImageNet, CIFAR100 and CIFAR10 databases for object recognition and TiCH for Character Recognition	–	The classification accuracy = 84.13%

- the incremental algorithm of Ozawa et al. [26] that combines the Incremental Principal Component Analysis (IPCA) and Resource Allocating Network with Long-Term Memory (RAN-LTM) and gives a recognition accuracy = 97%;
- the SVDU-IPCA algorithm of Zhao et al. [38] uses the IPCA algorithm, and an IKPCA algorithm with a recognition accuracy = 93%;
- an Incremental Learning algorithm [14] using Genetic Algorithm (ILUGA) and a binary SVM classifier is trained to be strong classifiers using a genetic algorithm with the overall accuracy of 93% and 94%.

We have found that incremental neural networks also give good recognition rates compared to other supervised learning algorithms, we identify:

- an ICL (Incremental Class Learning) approach for neural networks [22], the accuracy = 93.13%;
- the incremental learning algorithm SPAN [23] is used for spiking neural networks, the average accuracy = 92%;
- the proposed incremental Artificial Neural Network method of Zribi and Boujelbene [40] with the classification accuracy = 99.95%.

6 Conclusion and perspectives

In the context of incremental learning, we have presented, in this article, an introductory study of the main incremental supervised algorithms that could be found in the literature. We have identified the basic concepts of incremental learning and we have presented an overview of supervised methods and algorithms of incremental learning. We noticed the lack of contributions developed in this area and their orientation towards certain types of classifiers and we studied and synthesis different recent work in this context.

This study is the first step of our research for which we can consider several future extensions such as exploring the possibilities of hybridization between different incremental learning approaches and their application in evolving pattern recognition systems.

References

1. Almaksour A (2011) Incremental learning of evolving fuzzy inference systems: application to handwritten gesture recognition. Ph.D. thesis, INSA de Rennes
2. Bai X, Ren P, Zhang H, Zhou J (2015) An incremental structured part model for object recognition. *Neurocomputing* 154:189–199
3. Cauwenberghs G, Poggio T (2001) Incremental and decremental support vector machine learning. In: *Advances in neural information processing systems*, pp 409–415
4. Chefrour A, Souici-Meslati L (2013) Un panorama de méthodes d'apprentissage incrémental. In: *International conference on extraction and knowledge management*, Toulouse, France, pp 57–70
5. Déniz O, Castrillón M, Lorenzo J, Hernández M (2002) An incremental learning algorithm for face recognition. *International workshop on biometric authentication*. Springer, New York, pp 1–9
6. Diehl CP, Cauwenberghs G (2003) SVM incremental learning adaptation and optimization. In: *Proceedings of neural networks conference*, vol 4, pp 2685–2690
7. Erdem Z, Polikar R, Gurgen F, Yumusak N (2005) Ensemble of SVMs for incremental learning. In: *Proceedings of international workshop on multiple classifier systems*. Springer, New York, pp 246–256
8. Fung G, Mangasarian OL (2002) Incremental support vector machine classification. In: *Proceedings of the international conference on data mining (SIAM)*. Society for Industrial and Applied Mathematics, pp 247–260
9. Ghassebeh YA, Moghaddam HA (2007) A face recognition system using neural networks with incremental learning ability. In: *Proceedings of the international symposium on computational intelligence in robotics and automation*, pp 291–296
10. Gurney KR, Baker D, Rayner P, Denning S (2008) Interannual variations in continental-scale net carbon exchange and sensitivity to observing networks estimated from atmospheric CO₂ inversions for the period 1980 to 2005. *Global Biogeochem Cycles* 22(3):1–17
11. Hacene G, Gripon V, Farrugia N, Arzel M, Jezequel M (2017) Incremental learning with pretrained convolutional neural networks and binary associative memories
12. Han S, Meng Z, Khan AS, Tong Y (2016) Incremental boosting convolutional neural network for facial action unit recognition. In: *Advances in neural information processing systems*, pp 109–117
13. Huang C, Ai H, Yamashita T, Lao S, Kawade M (2007) Incremental learning of boosted face detector. In: *Proceedings of the 11th on computer vision IEEE*, pp 1–8
14. Hulley G, Marwala T (2007) Evolving classifiers: Methods for incremental learning. arXiv preprint [arXiv:0709.3965](https://arxiv.org/abs/0709.3965)
15. Joshi P, Kulkarni P (2012) Incremental learning: areas and methods-a survey. *Int J Data Min Knowl Manag Process* 2(5):43
16. Kawewong A, Pimup R, Hasegawa O (2013) Incremental learning framework for indoor scene recognition. In: *AAAI*, pp 1–7
17. Lawal IA, Abdulkarim SA (2017) Adaptive svm for data stream classification. *S Afr Comput J* 29(1):27–42
18. Liu Y (2015) Incremental learning in deep neural networks. Master of Science Thesis, Tampere University of Technology
19. Loosli G (2010) Méthodes à noyaux pour la détection de contexte. Academic editions E
20. Lu Y, Boukharouba K, Boonært J, Fleury A, Lecoeuche S (2014) Application of an incremental svm algorithm for on-line human recognition from video surveillance using texture and color features. *Neurocomputing* 126:132–140
21. Luo J, Pronobis A, Caputo B, Jensfelt P (2007) Incremental learning for place recognition in dynamic environments. In: *Proceedings of IEEE/RSJ international conference on intelligent robots and systems*, pp 721–728
22. Mańdziuk J, Shastri L (2002) Incremental class learning approach and its application to handwritten digit recognition. *Inf Sci* 141(3–4):193–217.
23. Mohammed A, Lu G, Kasabov N (2012) Evaluating span incremental learning for handwritten digit recognition. *International conference on neural information processing*. Springer, New York, pp 670–677

24. Molina JFG, Zheng L, Sertdemir M, Dinter DJ, Schönberg S, Rädle M (2014) Incremental learning with svm for multimodal classification of prostatic adenocarcinoma, PLoS One 9(4):e93600
25. Ozawa S, Pang S, Kasabov N (2008) Incremental learning of chunk data for online pattern classification systems. *IEEE Trans Neural Netw* 19(6):1061–1074
26. Ozawa S, Toh SL, Abe S, Pang S, Kasabov N (2005) Incremental learning of feature space and classifier for face recognition. *Neural Netw* 18(5–6):575–584
27. Polikar R, Upda L, Upda SS, Honavar V (2001) Learn++: an incremental learning algorithm for supervised neural networks. *IEEE Trans Syst Man Cybern Part C* 31(4):497–508
28. Prudent Y (2006) *Système d'apprentissage incrémental et hybride*. PhD thesis, INSA University, Rouen
29. Ralaivola L, dAlché Buc F (2001) Incremental support vector machine learning: A local approach. In: *Proceedings of the international conference on artificial neural networks*. Springer, New York, pp 322–330
30. Reddy KK, Liu J, Shah M (2009) Incremental action recognition using feature-tree. In: *Proceedings of the 12th international conference on computer vision IEEE*, pp 1010–1017
31. Ruping S (2001) Incremental learning with support vector machines. In: *Proceedings of the international conference on Data Mining IEEE*, pp 641–642
32. Salperwyck C, Lemaire V, de Bois DUDP (2010) Classification incrémentale supervisée: un panel introductif. In: *AAFD*, pp 121–148
33. Sarwar SS, Ankit A, Roy K (2017) Incremental learning in deep convolutional neural networks using partial network sharing. [arXiv:1712.02719](https://arxiv.org/abs/1712.02719)
34. Schlimmer JC, Fisher D (1986) A case study of incremental concept induction. *AAAI* 86, pp 496–501
35. Syed NA, Liu H, Sung KK (1999) Handling concept drifts in incremental learning with support vector machines. In: *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp 317–321
36. Toh SL, Ozawa S (2003) A face recognition system using neural networks with incremental learning ability. In: *Proceedings of the 8th Australian and New Zealand Conference on intelligent information systems*, Citeseer, pp 389–394
37. Utgoff PE (1989) Incremental induction of decision trees. *Mach Learn* 4(2):161–186
38. Zhao H, Yuen PC, Kwok JT (2006) A novel incremental principal component analysis and its application for face recognition. *IEEE Trans Syst Man Cybern Part B* 36(4):873–886
39. Zou L, Zhang T, Cao Z (2009) An incremental learning algorithm based on Support Vector Machine for pattern recognition. In: *Proceedings of the International Society for Optics and Photonics*, vol 7496
40. Zribi M, Boujelbene Y (2016) The neural networks with an incremental learning algorithm approach for mass classification in breast cancer. *Biomed Data Min* 5(118):2

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.