**Line Grid Only，28 Lines Per Page,Single Line Spacing**

Deep learning is recently showing outstanding results for solving a wide variety of robotic tasks in the areas of perception, planning, localization, and control. Its excellent capabilities for learning representations from the complex data acquired in real environments make it extremely suitable for many kinds of autonomous robotic applications. In parallel, Unmanned Aerial Vehicles (UAVs) are currently being extensively applied for several types of civilian tasks in applications going from security, surveillance, and disaster rescue to parcel delivery or warehouse management.

In this paper, a thorough review has been performed on recent reported uses and applications of deep learning for UAVs, including the most relevant developments as well as their performances and limitations. In addition, a detailed explanation of the main deep learning techniques is provided. We conclude with a description of the main challenges for the application of deep learning for UAV-based solutions.

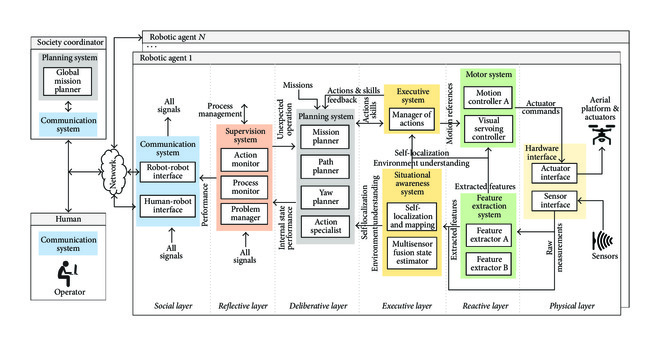
1. Introduction

Recent successes of deep learning techniques in solving many complex tasks by learning from raw sensor data have created a lot of excitement in the research community. However, deep learning is not a recent technology. It started being used back in 1971, when Ivakhnenko [1] trained an 8-layer neural network using the Group Method of Data Handling (GMDH) algorithm. The term deep learning began to be used during the 2000s, when Convolutional Neural Networks (CNNs), a computational original model from the 80s [2] but trained efficiently in the 90s [3], were able to provide decent results in visual object recognition tasks. At the time, datasets were small and computers were not powerful enough, so the performance was often similar to or worse than that of classical Computer Vision algorithms. The development of CUDA for Nvidia GPUs which enabled over 1000 GFLOPS per second and the publication of the ImageNet dataset, with 1.2 million images classified in 1000 categories [4], were important facts for the popularization of CNNs with several layers ( to  connections and  to  parameters). These deep models show great performance not only in Computer Vision tasks but also in other tasks such as speech recognition, signal processing, and natural language processing [5]. More details about recent advances in deep learning can be found in [6, 7].

An evidence of the suitability of deep learning for many kinds of autonomous robotic applications is the increasing trend in deep learning robot related scientific publications over the past decades, which is expected to continue growing [8].

Due to the versatility, automation capabilities, and low cost of Unmanned Aerial Vehicles (UAVs), civilian applications in diverse fields have experienced a drastic increase during the last years. Some examples include power line inspection [9], wildlife conservation [10], building inspection [11], and precision agriculture [12]. However, UAVs have limitations in the size, weight, and power consumption of the payload and limited range and endurance. These limitations cannot be overlooked and are particularly relevant when deep learning algorithms are required to run on board a UAV.

In this survey, we have grouped publications according to the taxonomy proposed in Aerostack [13], which is aerial robotics architecture consistent with the usual components related to perception, guidance, navigation, and control of unmanned rotorcraft systems. The purpose of referring to this architecture, depicted in Figure [1](https://www.hindawi.com/journals/js/2017/3296874/fig1/),+ is to achieve a better understanding about the nature of the components to the aerial robotic systems analyzed. Using this taxonomy also helps identify the components in which deep learning has not been applied yet. According to Aerostack, the components constituting an unmanned aerial robotic system can be classified into the following systems and interfaces:(i)Hardware interfaces: this category includes interfaces with both sensors and actuators(ii)Motor system: the components of a motor system are motion controllers, which typically receive commands of desired values for a variable (position, orientation, or speed). These desired values are translated into low-level commands that are sent to actuators(iii)Feature extraction system: feature extraction here refers to the extraction of useful features or representations from sensor data. The task of most deep learning algorithms is to learn data representations, so feature extraction systems are somewhat inherent to deep learning algorithms(iv)Situational awareness system: this system includes components that compile sensor information into state variables regarding the robot and its environment, pursuing environment understanding. An example component within the situational awareness system is SLAM algorithms(v)Executive system: this system receives high-level symbolic actions and generates detailed behaviour sequences(vi)Planning system: this type of system generates global solutions to complex tasks by means of planning (e.g., path planning and mission planning)(vii)Supervision system: components in the supervision system simulate self-awareness in the sense of ability to supervise other integrated systems. We can exemplify this type of component with an algorithm that checks whether the robot is actually making progress towards its goal and reacts in the presence of problems (unexpected obstacles, faults, etc.) with recovery actions(viii)Communication system: the components in the communication system are responsible for establishing an adequate communication with human operators and/or other robots



**Line Grid Only，36 Lines Per Page,Double Line Spacing**

Aerostack architecture, consisting of a layered structure, corresponding to the different abstraction levels in an unmanned aerial robotic system. The architecture has been applied here to systematically classify deep learning-based algorithms available in the state of the art which have been deployed for applications with Unmanned Aerial Vehicles.

The remainder of this paper is as follows: firstly, Section [2](https://www.hindawi.com/journals/js/2017/3296874/#sec2) covers a description of the currently relevant and prominent deep learning algorithms. For the sake of completeness, deep learning algorithms have been included regardless of their direct use in UAV applications. Section [3](https://www.hindawi.com/journals/js/2017/3296874/#sec3) presents the state of the art in deep learning for feature extraction in UAV applications. Section [4](https://www.hindawi.com/journals/js/2017/3296874/#sec4) surveys UAV applications of deep learning for the development of components of planning and situation awareness systems. Reported applications of deep learning for motion control in UAVs are presented in Section [5](https://www.hindawi.com/journals/js/2017/3296874/#sec5). Finally, a discussion of the main challenges for the application of deep learning for UAVs is covered in Section [6](https://www.hindawi.com/journals/js/2017/3296874/#sec6).

2. Deep Learning in the Context of Machine Learning

Machine Learning is a capability enabling Artificial Intelligence (AI) systems to learn from data. A good definition for what learning involves is the following: “a computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E” [15]. The nature of this experience E is typically considered for classifying Machine Learning algorithms into the following three categories: supervised, unsupervised, and reinforcement learning:(i)In supervised learning, algorithms are presented with a dataset containing a collection of features. Additionally, labels or target values are provided for each sample. This mapping of features to labels of target values is where the knowledge is encoded. Once it has learned, the algorithm is expected to find the mapping from the features of unseen samples to their correct labels or target values.(ii)The purpose in unsupervised learning is to extract meaningful representations and explain key features of the data. No labels or target values are necessary in this case in order to learn from the data.(iii)In reinforcement learning algorithms, an AI agent interacts with a real or simulated environment. This interaction provides feedback between the learning system and the interaction experience which is useful to improve performance in the task being learned.

Deep learning algorithms are a subset of Machine Learning algorithms that typically involve learning representations at different hierarchy levels to enable building complex concepts out of simpler ones. The following paragraphs cover the most relevant deep learning technologies currently available in supervised, unsupervised, and reinforcement learning.

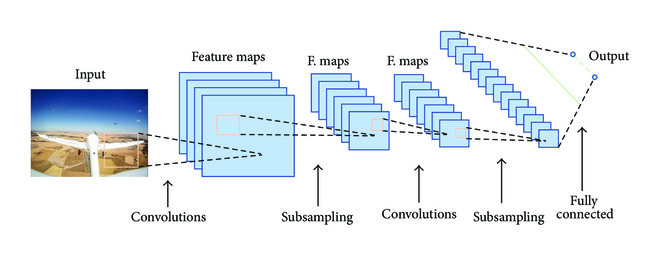
2.1. Supervised Learning

Supervised learning algorithms learn how to associate an input with some output, given a training set of examples of inputs and outputs [16]. The following paragraphs cover the most relevant algorithms nowadays in supervised learning: Feedforward Neural Networks, a popular variation of these called Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and a variation of RNNs called Long Short-Term Memory (LSTM) models.

Feedforward Neural Networks, also known as Multilayer Perceptrons (MLPs), are the most common supervised learning models. Their purpose is to work as function approximators: given a sample vector  with  features, a trained algorithm is expected to produce an output value or classification category  that is consistent with the mapping of inputs and outputs provided in the training set. The approximated function is usually built by stacking together several hidden layers that are activated in chain to obtain the desired output. The number of hidden layers is usually referred to as the depth of the model, which explains the origin of the term deep learning: learning using models with several layers. These layers are made up of neurons or units whose activation given an input vector  is given by the following equation:where  is a vector of  weights and  is an activation function that is usually chosen to be nonlinear. The activation of unit  in layer  given its  inputs (outputs of the previous layer ) is given by the following equation:

During the process of learning, the weights in each unit are updated using backpropagation in order to optimize a cost function, which generally indicates the similarity between the desired outputs and the actual ones.

Convolutional Neural Networks (CNNs), depicted in Figure [2](https://www.hindawi.com/journals/js/2017/3296874/fig2/), are a specific type of models conceived to accept 2-dimensional input data, such as images or time series data. These models take their name from the mathematical linear operation of convolution which is always present in at least one of the layers of the network. The most typical convolution operation used in deep learning is 2D convolution of a 2-dimensional image  with a 2-dimensional kernel , given by the following equation:

[](https://www.hindawi.com/journals/js/2017/3296874/fig2/" \t "https://www.hindawi.com/journals/js/2017/3296874/_blank)