The Quiet Revolution in Machine Vision
A state-of-the-art survey paper, including historical review, perspectives, and future directions

Melvyn L Smith, Lyndon N Smith, Mark F Hansen

Centre for Machine Vision,
Bristol Robotics Laboratory, University of the West of England, Bristol, UK

Abstract

Over the past few years, what might not unreasonably be described as a true revolution has taken place in the field of machine vision, radically altering the way many things had previously been done and offering new and exciting opportunities for those able to quickly embrace and master the new techniques. Rapid developments in machine learning, largely enabled by faster GPU-equipped computing hardware, has facilitated an explosion of machine vision applications into hitherto extremely challenging or, in many cases, previously impossible to automate industrial tasks. Together with developments towards an internet of things and the availability of big data, these form key components of what many consider to be the fourth industrial revolution. This transformation has dramatically improved the efficacy of some existing machine vision activities, such as in manufacturing (e.g. inspection for quality control and quality assurance), security (e.g. facial biometrics) and in medicine (e.g. detecting cancers), while in other cases has opened up completely new areas of use, such as in agriculture and construction (as well as in the existing domains of manufacturing and medicine). Here we will explore the history and nature of this change, what underlies it, what enables it, and the impact it has had - the latter by reviewing several recent indicative applications described in the research literature. We will also consider the continuing role that traditional or classical machine vision might still play. Finally, the key future challenges and developing opportunities in machine vision will also be discussed.

Keywords: machine vision; machine learning; deep learning; state-of-the-art

1. Introduction – the purpose of this paper and what it brings that is new to the reader

The appropriate utilisation of visual data is the enabling component in many automated tasks, from industrial quality control to driverless cars. Here we consider the recent history of machine vision research, the huge changes that have, and are still taking place, and the dramatic impact this is having, both in terms of methodology and new applications. We introduce, define, and explain key terminology for those not completely familiar with the discipline, and comment on the most relevant and impactful developments by pinpointing ideas in the literature which have changed radically the approaches and techniques now being explored. Using insightful examples taken from representative applications, we will explain how these ideas are being applied now and consider the future directions.

2. What is machine vision?

The introduction of machine vision to industrial processes is often motivated by a desire to reduce costs by increasing efficiency (and so productivity), reduce errors (and so improve quality) or gather data. Equally importantly, it may also substitute for an absence of available skilled labour or release workers from dangerous, demanding or fatiguing industrial activities. In the past, the definition of the term has been somewhat unclear, however more recently ‘machine vision’ has, largely *de facto*, come to be understood as the practical realisation of image understanding, or more specifically computer vision techniques, to help solve practical industrial problems that involve a significant visual component. The more recent emphasis on the marriage between machine vision and machine learning, that has so revolutionised the discipline, has been made possible by transformative developments in the field of artificial intelligence (AI) and has served to move some machine vision capabilities closer to that of human vision – a long unmet ambition that has existed since the early days of computer vision as far back as the 1960s.

3. Machine learning and the artificial neural network (ANN)

To help appreciate why this transformation has come about, we first need to understand the role of machine learning and have an appreciation of some of the overlapping terminology. Deep learning is a particular kind of machine learning, which in turn is itself a subset of artificial intelligence or AI (Figure 1). The term general AI, also referred to as strong AI, captures the ambitious notion of a theoretical synthetic device able to learn, reason and behave as humans do. This does not (perhaps yet) exist. A form of AI that does exist is that of narrow AI. This includes software able to dramatically outperform humans, but only in very limited, usually single, tasks. It is generally the case that while machine learning software can automatically learn from historical data, and so improve with experience, and make informed decisions using a range of (often statistical) techniques, without needing to be explicitly manually programmed, the term deep learning is reserved for a special kind of machine learning that allows computers to solve much more challenging and complex problems.

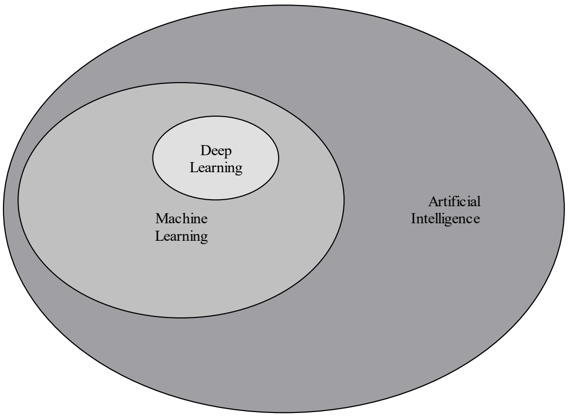

Figure 1 The relationship between the domains of AI, machine learning and deep learning

In practice, deep learning is realised using artificial neural networks (ANNs) to simulate human-like decision making. Inspired by, and loosely simulating the functioning of the human cortex, ANNs comprise a matrix of connected layers of nodes (Figure 2). The input layer is where data enter the network (for which the number of nodes is the same as the dimension of input features). This is followed by one or more hidden layers that transform the data flowing through the network towards the output layer (where the number of nodes is the same as the number of classes to be classified), that produces the network predictions. A deep network simply refers to an ANN with more than three hidden layers and in some cases can have millions of nodes (loosely analogous to neurons in the human brain). We should note here that in fact there are two different kinds of output prediction - either regression (a continuous quantity), or, as mentioned above and which is more often the case in many machine vision applications, a classification (a discrete class label or category).

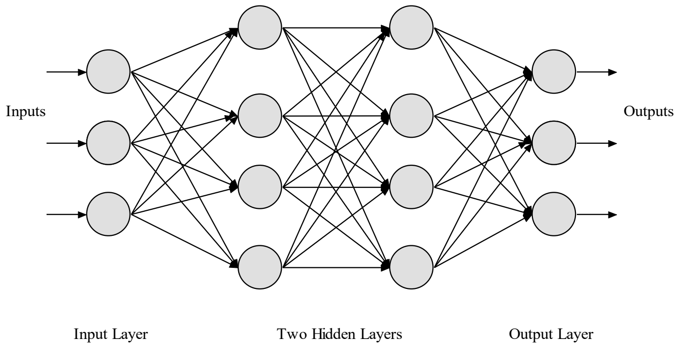

Figure 2 The connected node topology of an artificial neural network (ANN)

Connections between the layer nodes have weights and biases, known as the layer parameters (Figure 3), that control how the layers together transform data between the input and output of the network during a process known as feedforwarding. It is these parameters that represent the network model, and we shall see later how a feedback learning process is used to first establish these parameters during training. The development of ANNs has its origins as far back as the 1950s and for a while fell out of favour. However, a series of breakthroughs dating from the 1990s by Yann LeCun et. al. [1, 2], culminating in [3] and then later around 2010, most notably the work by Alex Krizhevsky [4] and others in image classification, led to a particular form of ANN known as a deep convolutional neural network (CNN) becoming established for many challenging computer vision tasks, particularly those involving perception for object recognition and localisation in natural images.

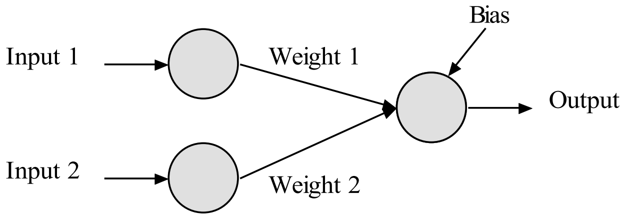

Figure 3 The layer node parameters (Output = Input 1 x Weight 1 + Input 2 x Weight 2 + Bias)

3.1 The convolutional neural network (CNN) – a special kind of ANN for images

Figure 4 shows the principal components of a convolutional neural network (CNN). The main difference between CNNs and regular ANNs, and the reason why they lend themselves so well to image analysis, is the addition of a convolutional layer - an initial filtering stage, implemented using conventional filter kernels that traverse the 2D image to detect local spatial features, such as edges, corners, and other patterns, at different positions – somewhat like processes in human vision. Outputs from the kernel are summed using a bias function to form the input to a non-linear activation function layer, often a rectified linear unit (ReLU) (a non-linear transformation that reduces negative input values to zero while positive values are passed as output). This non-linear characteristic of the ReLU is also similar to the firing behaviour of neurons in the human brain and so is suited to visual tasks (or more formally it offers a reduced sensitivity to the vanishing gradient problem – an issue that impedes CNN training). A data reduction stage (necessary as images may contain millions of pixels and so vast amounts of data – some potentially redundant to the task) is then performed by a pooling layer. This usually employs either average- or a max-pooling function to a defined neighbourhood of pixel values, reducing the data dimensionality to leave only the most significant pixels. Note here that the combined convolution and pooling layers may be repeated, to detect a hierarchy of image features, before input to a fully connected classification stage, giving a feature vector, and finally an output layer, for which a Softmax activation function allows multinomial labelling (where the probability of all outputs sums to 1.0). Hence, the network output gives a probability distribution over the output labels indicating which output case the network has most closely matched to the input. So, CNNs are very well adapted to machine vision tasks, as they are specifically designed to process large quantities of pixel data. The convolution and pooling layers act as feature extractors from the input image, while the fully connected layers act as a classifier. In this way, CNNs can usefully be divided into two parts: feature extraction and classification. Feature extraction is achieved by the convolution plus ReLU together with pooling layers, and classification is achieved by the fully connected layers.

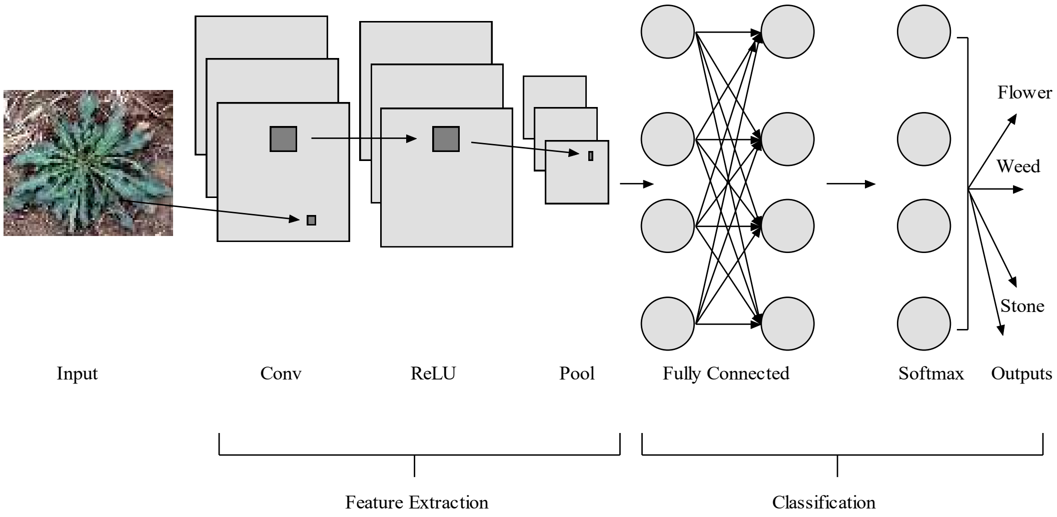

Figure 4 The architecture of a convolutional neural network (CNN). The CNN can be divided into two parts: feature extraction and linear classification

However, before all this can happen, the network must first be trained. Training enables the network to output predictions by identifying patterns in a set of labelled training data, fed through the network while the outputs are compared with the actual labels by an objective loss function. During training the network's parameters (the weight and bias of each neuron) are tuned over a series of iterations or epochs, until the patterns identified by the network result in good predictions for the training data. Thus, the parameters of the network (weights and biases) are found during a supervised training phase by minimizing a loss function, (calculated during a forward propagation), between prediction and ground truth labels at the output layer. To help this process, regularisation constraints are used to update the network weights and biases at each iteration (e.g. employing stochastic gradient descent – SGD) using backpropagation until convergence. The concept of backpropagation in training the network is very important and the key part of how an ANN is made to work. It is therefore worth spending a little time underlining what is going on. During backpropagation, within each layer, the node weights are adjusted in proportion to the amount of error they contributed to the loss function (the difference between predictions and labels) during forward propagation. The more a given node contributes to the error, the more it is adjusted. When the loss function reaches a convergence state, the current weights and biases are preserved as the well-trained model. It is in this manner that the weights and biases of the network are established via the learning process using real data to form the deep learnt model. This trained model may later be applied (a process known as inferencing) to unknown data (known as generalisation) in a practical application to identify various features of interest in images. CNNs have become the default choice for simultaneous detection of multi-class categorisation problems, typical of many industrial visual tasks.

4. Classical machine vision

Since the 1980s, increasing computational capacity at a reducing cost, combined with improvements in camera and lighting technologies, have made possible the practicable deployment of industrial grade cameras and microprocessors together with new capabilities in LED lighting (offering more control of the lighting intensity, colour or frequency and differing projected patterns), to undertake a multitude of industrial vision tasks. Examples include measuring the size and position of objects, assessing their surface quality in terms of colour or texture, identifying defects or contamination, sorting, checking the integrity of safety critical components and assemblies, or guiding industrial robots. In almost all cases, these tasks have required strict environmental structuring, whereby, for example, any external ambient lighting is eliminated and replaced, and the objects to be inspected are themselves unchanging and in known fixed locations. In many production processes, where manufactured objects are to be detected, measured, or inspected in some way, this requirement for strict environmental structuring is relatively easily accommodated, as the components being manufactured are already under control and most often in predictable fixed positions at repeatable known times during the manufacturing cycle. For example, in the case of objects on a production line conveyor or being handled by an industrial robot, there often exists good opportunity to introduce cameras and lighting at suitable physical locations that have a minimal impact on the production process. However, while this is fine in many manufacturing production lines, there are also a wide range of other industrial activities involving a visual sense where such environmental structuring is either problematic, undesirable, or even completely impossible.

4.1 Why the need for structure in classical machine vision?

4.1.1 A rule driven vs a new data driven approach

For an automated system to be able to make sense of visual data, a model of the visual scene together with a set of task rules and parameters is required. Conventionally, this model is hand-crafted in the form of rules based on *a priori* information that is realised in code. In other words, in a manufacturing scenario a CAD model of the object being detected, measured, or inspected, is most likely readily available. In the case of defect detection, the type and range of defects will also be known in advance of inspection system hard and software design and installation. Also, given the nature of the manufacturing process, the position and pose of the objects to be inspected, when viewed by a camera, can also be established in advance, and are able to be readily fixed. Consider the example of the mass production of kitchen and bathroom ceramic tiles [5], where an inspection stage is needed to check conformity in terms of size and geometry and the presence of any unacceptable surface anomalies. The latter may take the form of 2D colour or 3D moulding defects [6]. In this case, a model describing the features of an ideal perfect tile can be obtained using existing data (or alternatively using suitable good quality examples, sometimes called a 'golden template' approach), and tolerances of acceptability applied to any variation in these features captured during the production process. The introduction of structured lighting, along with suitable camera and optics, can allow an ideal and repeatable view of the tiles in a fixed position as they pass along a moving conveyor. All this available a-priori knowledge and structuring reduces variation, offers predictability, and so greatly simplifies the machine vision activity. Any ambient lighting, that may be subject to change, can be excluded, and controlled artificial lighting substituted, allowing activities such as segmentation, i.e. the isolation of the tile object and its features, to, for example, be performed using simple fixed image intensity thresholds. Together, this structuring reduces the need to be able to accommodate change and uncertainty within the inspection process. This means that the machine vision software solution, that is the design of the algorithms realised in software, is relatively straightforward and so easily hand-crafted.

Next consider an entirely different class of industrial machine vision problem, for example that of identifying weeds within crops in a field as part of an agricultural farming application [7]. This is typical of many machine vision tasks that involve natural objects, such as plants (e.g. in farming, and in food and timber processing), minerals (e.g. polished marble and granite), animals and animal by-products (e.g. meat and leather products) or humans (e.g. in medical and security applications), as opposed to synthetic man-made objects – such as manufactured metal and plastic parts and assemblies. In our example case of automated weeding, the aim is to automatically provide a targeted precision dose of herbicide only to the weeds, and not the crop or surrounding area [8]. Such a visually informed system might not only reduce waste but could offer significant environmental benefits by minimising levels of toxic chemical runoff and so ground and river pollution. The machine vision system specification required here will need to be such that the device can both recognise the weeds, or even perhaps a range of differing weed species (dock, ragwort, etc), and determine their individual spatial locations. All will need to be completed in real-time in an outdoor environment, as a tractor-mounted vision-equipped computer-controlled spray boom traverses the crop. These uncontrollable variables, such as different invading species, the amount of weed and crop growth, uneven ground, changing lighting and occlusion from the crop itself, make the task inherently difficult as it includes a great deal of random variation. Until very recently such tasks were considered entirely beyond the capability of a hand-crafted rule-based model approach for any sort of practical application. Even so, such a visually complex task is representative of many that are now being successfully addressed by developments in machine learning techniques. More specifically, by deploying a form of specialised artificial neural network, known as a convolutional neural network (CNN), it is possible for the rules of such a complex machine vision model to be automatically learnt from the data by the system itself, in a so-called data-driven approach.

5. Deep learning is changing the way machine vision is done

5.1 The conventional machine vision process

In conventional or classical machine vision, there are several well-established logical steps necessary to realise almost any industrial machine vision application task. These can broadly be categorised into the five core activities of: 1. image acquisition; 2. pre-processing; 3. segmentation; 4. feature extraction; and 5. classification or interpretation (Figure 5). Image acquisition will normally include aspects of environmental structuring, involving scene constraints related to object positioning, camera and lens selection, and the design of a suitable structured lighting arrangement. This is arguably by far the most impactful stage. It needs to be very carefully considered as it can dramatically affect the complexity and so the cost, efficiency, and reliability of the subsequent stages. The output following classification / interpretation may take the form of an action, such as operating an actuator, guiding a robot, altering a process variable (in the case of an adaptive closed-loop feedback system) or generating a report (say in a quality or process control monitoring application). However, most significantly in terms of the overall performance of the machine vision solution, are the stages in between. This is where the creative thinking of the engineer developing a solution largely takes place. Realising an optimal solution, often via a time-consuming trial-and-error process to establish rules and tune parameters, depends strongly on the knowledge, experience, and judgment of the system developer. Note also that when combined, these stages produce a translation of data type, along with a filtering, to give a dramatic reduction of data quantity. What starts as an optical image, is converted into a numerical data 2D array, from which features and then scene descriptors are subsequently derived. So, while an acquired image may be comprised of millions of bytes of data, say from a high-resolution colour camera, the desired output from a notional machine vision system may take the form of only a single binary bit, 1 or 0, indicating that, say, an object is either of acceptable quality or not.

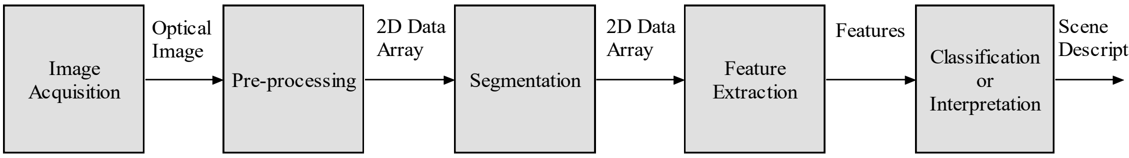

Figure 5 The five steps of classical machine vision

5.2 The deep learning machine vision process

5.2.1 Deep learning is changing the role of the machine vision solution developer

The development of a machine learning, or more specifically a deep learning solution to a machine vision task, is quite different from the processes involved for implementing the conventional machine vision pipeline. In some ways the task is greatly simplified by combining segmentation, feature extraction and classification stages, and it completely avoids the need to design and hand-code these steps. Instead, the system developer is required to possess a somewhat differing set of skills, including expertise in implementing deep learning network architectures (e.g. selection of an appropriate network architecture, methods of data augmentation, adjustment of network hyperparameters such as number of training epochs, learning rate, L1 and L2 (regularization parameters), etc) [9]. As we have already seen, the way the deep learning machine model is fashioned is via a process known as end-to-end (usually supervised) learning, in which the ANN model is created by presenting the network with labelled example image datasets (ground truth data). This labelling is used to identify the object classes present in the image for training the network. As described above, the labels represent the desired output for a given input and are used to create the loss function that the network tries to minimise by adjusting the weights during backpropagation. Because it is necessary to label the data, to explicitly tell the network what the data represent, this is known as a supervised form of learning. The network learns the significant image features needed to automatically map image inputs to target labels using a series of data transformations. Then later in application (the network inferencing stage), the output of the trained network takes the form of predictions, i.e. which of the possible outputs the network thinks most closely match the given input. However, while this all sounds relatively straightforward, there are a few downsides. Perhaps most significant of which is the amount of example, manually labelled in the case of supervised learning, data (perhaps tens of thousands of images) and time (often hours or days), needed to train the network. We will see next that there are various ways to tackle these issues (along with a few others), however the main takeaway here is that once trained, the inferencing (operational) stage is very much quicker (perhaps taking only a matter of milliseconds).

6. A few notable key challenges with deep learning

Apart from the amount of time needed for training, there are one or two other challenging issues regularly encountered when using deep learning to solve industrial problems that deserve some special attention.

6.1 Limited training data: Often there may only be limited training data available for the desired task - not enough to train a network from scratch. In such cases it may be possible to deploy transfer learning with fine tuning as a feature extractor. Transfer learning is a machine learning technique where a model (network) previously trained on one task is re-purposed on a second (usually) related task. The advantage is that huge quantities of training data are not needed. This can also help reduce hardware costs for the computationally expensive training stage [10]. Here a pre-trained CNN (where the weights and biases are already established, often by training on a large natural image dataset, for example the ImageNet dataset) is subject to further supervised training for part, or all, of the network layers using the new available data for the task under solution. Often, the last two layers in the original network are replaced by two fully connected layers. The last layer is the output layer and is matched with the number of classes in the dataset. Examples of differing pre-trained classifiers include VGG-16 and VGG-19 networks [11], the ResNet50 network [12], Mobilenet [13], and Xception [14]. Available data will usually be divided into training, validation, and test sets for learning from examples, establishing the soundness of learning results, and evaluating the generalisation ability of a developed algorithm on unseen data, respectively. Also, in the case of limited data availability, cross validation methods (e.g. one-leave out, fivefold, or tenfold validations) can be used.

6.2 The need to manually label training data: Most applications of CNNs involve supervised training with large quantities of data – thousands of images. This can pose two problems. Firstly, as mentioned above, is the limited availability of such large datasets for training, and secondly there is the need to label (i.e. say what the image contains) the data (usually a manual process). Both issues can to some extent be addressed using data augmentation. This is where a limited labelled dataset is slightly altered using a range of automated image transformations, such as rotations, reflections, height and width shifts, zooming, horizontal-flipping, shear intensity changes, cropping, and adding noise, to artificially create variation in the images and so a much larger training dataset. It is worth noting that while such techniques can help improve training with a limited dataset, it is often the case that use of transfer learning and data augmentation would not be expected to result in as good a performance as training from scratch using a large dataset.

6.3 Overfitting the training data: Another well-known problem in machine learning is that of overfitting. This is where the model fits the training data well but does not generalise to unseen data (the useful part). Overfitting is one of the most frustrating problems in machine learning projects and most practitioners experience it in their work. There are thankfully a range of tools that can help, including just simplifying the network by reducing the number of hidden layers, regularisation (adding a cost to the loss function for large weights) and the use of dropout layers.

6.4 The mysterious black box: It is worth noting that it can be difficult to know what the ANN is actually doing - for example which parts of the image are being used? Consider an object recognition task. Could the network be using something unexpected, such as a timestamp in the corner of the image and no part of the object itself? This can be of most concern in medical applications, where even if performance is good, this lack of accountability could have important legal consequences. Thankfully, a prediction process debugging tool exists in the form of the Grad-CAM method. This gives a heatmap of class activation that distinguishes the similarity of each image location with respect to a particular class. More specifically, the heatmap of class activation is a feature map of the last convolutional layer for a given input image, with the feature map channels weighted by a class gradient calculated with regards to the feature map [15]. Figure 6 shows an example heat map indicating the areas of a pig's face used in a biometric recognition task [16]. This shows in some detail which parts of the face the network was using for recognition and helps us, to some extent, understand what the network was doing and the predictions it made.

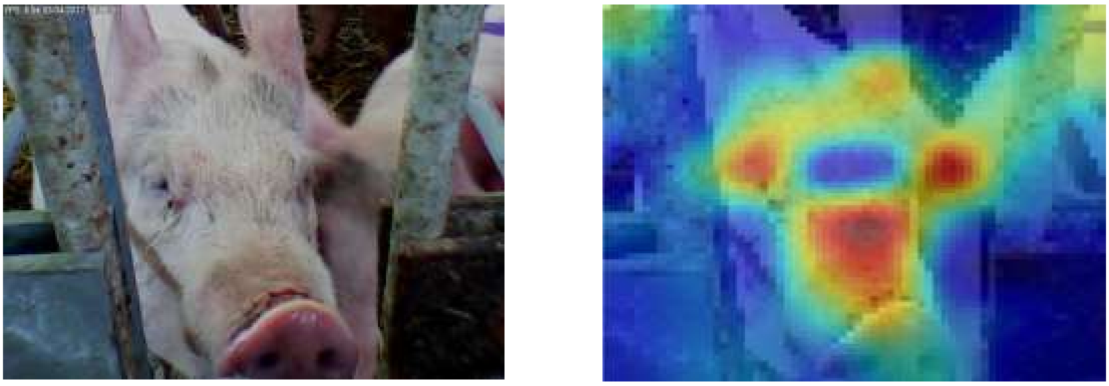

Figure 6 Example heat map (right) indicating the areas of a pig's face (left) used in a biometric face recognition task where yellows and reds progressively show more significant areas

7. Examples of machine vision tasks enabled by deep learning - a new class of machine vision tasks

We next delve a little deeper into the details of how deep learning is currently being applied to real machine vision industrial tasks via a range of seminal state-of-the-art real-world prototypical example applications. We explore applications across differing sectors, consider their context, including the role of big data [17] and the internet of things [18], the motivations, and the potential advantages. We also review the common architectures in use today; and look to the future by considering the most significant remaining problematic issues still to be addressed.

7.1 Agriculture – pre-farm gate

As previously discussed, deep learning has improved performance in many existing industrial tasks as well as allowing new areas of application. Perhaps none more so than in agriculture, where, in so called ‘pre-farm gate’ applications, the imposition of any form of environmental structuring is most often extremely difficult, and so in practice tends to be minimal. These applications may include, crop assessment (detecting abiotic and biotic stress), pruning (including weeding) and picking or harvesting (identifying ripeness and location). Uncertainty and variation can exist both in terms of the natural objects being imaged and their outdoor context.

Williams et al. [19] present the design and evaluation of a kiwifruit harvesting robot intended to operate autonomously in pergola style orchards. The technology is typical of many agri-tech solutions currently under development and the application highly topical as the work is in part motivated within the context of a reduced availability of farm labour. A deep neural network with stereo matching was used to detect and locate kiwifruit under real-world conditions (with some added illumination). The machine supervised learning model utilised here is typical of many similar applications and so worth exploring in a little detail. Semantic segmentation (where each image pixel is given a class label) was performed using a fully-convolutional network (FCN) adapted from the VGG-16 (available via Github as FCN-8S). The network was trained using an Nvidia GTX-1070 8GB graphics card on 63 (48 training and 15 validation) 200x200 pixel hand-labelled images, collected across multiple orchards under a range of typical conditions. Commercial orchard field trials gave an average pick cycle time of 5.5s/fruit, of which 3.0s was for the detection step. Some 89.6% of the pickable kiwifruit could be detected. This performance was reported as significantly better than other comparable automated systems. Wang et al. [20] provide a general review of ground-based machine vision for weed detection that included CNN deep learning-based approaches for detection and segmentation in the field. Many of the key challenges indicative of agricultural applications are identified in this work, such as occlusion and overlap of leaves, varying lighting conditions and the effects of different growth stages. Palacios et al. [21] describe a system for quantification of grapevine flowers using a field based mobile platform for early crop yield forecasting. Here artificial illumination and night-time operation offered some limited structuring. A deep fully-convolutional neural network SegNet architecture (two deep neural networks) with a VGG-19 network as the encoder were employed for semantic segmentation. As with other examples, images were also pixel-wise manually labelled and image augmentation (rotations and flipping) used to increase variability in the limited training data. As discussed above, we see that in practice the use of such data augmentation techniques are very commonly used in an effort to cope with, or avoid the need for, large quantities of hand labelled training images. Impressive F1 (F scores are a measure of accuracy calculated from the precision and recall) score values of 0.93 and 0.73 were obtained for multiscale flower segmentation. A strong correlation was reported between the estimated number of flowers and the final produce yield. Kakani [22] take a step back and explain how agri-tech start-ups are choosing task-specific AI and vision solutions in the context of improving yields and the goal of sustainable food supplies by 2050. Switching from largely static crops to moving farm animals, Nasirahmadi et al. [23] demonstrate how deep learning can be used to detect standing and lying behaviour of pigs to better than an average precision of 0.93 and 0.95 respectively, for monitoring animal health and welfare under varied real-world farm conditions. Use of Faster R-CNN (faster regions with convolutional neural network features), R-FCN (region-based fully convolutional network) and SSD (single shot multi-box detector) methods were all explored. In related work, Hansen et al. [16] show how deep learning achieved better than 96% accuracy in biometrically recognising pig faces on the farm and point to advantages of facial biometrics when compared to the use of conventional RFID ear-tags. They use both a transfer learning approach, centred on the pre-trained VGG-Face model (based on the VGG-Very-Deep-16 CNN and trained on the Labelled (human) Faces in the Wild dataset), and a CNN model, trained from scratch with their own artificially augmented pig face data. Class Activated Mapping using Grad-CAM was deployed to show the regions that the network used to discriminate between pig faces (Figure 6). In an example of the benefits of transfer learning, it is very interesting that the VGG-Face pre-trained model performed as well as it did (91% accuracy), given that it was trained only on human and not pig faces. In addition to the outdoor applications mentioned above, other rapidly developing areas include the use of visual data from drones [24]. Here multispectral, thermal and visible cameras can be used for land surveys (soil characteristics), inspection of crops for yield prediction, disease detection, and when to irrigate (smart irrigation), or otherwise treat and harvest.

7.2 Food processing or agriculture - post-farmgate

A greater level of environmental structuring is possible in post-farm gate processing, in the form of material handling (e.g. conveyors) and indoor controlled lighting. However, it is the product or, in the case of quality control, the contaminates themselves that are the main source of the variation. This makes conventional hardcoded feature extraction a significant challenge. Applications here may include, quality control, sorting, grading, and in packing.

Rong et al. [25] present an approach for detecting foreign objects in walnuts, pointing to the issue of irregular shapes and complex features making the manual design of suitable feature detectors extremely problematic. Their aim was the detection of different sized natural foreign objects (flesh leaf debris, dried leaf debris and gravel dust) and man-made foreign objects (paper and plastic scraps, packing material and metal parts). Two different convolutional neural network structures were used for segmenting the nuts and detecting foreign bodies. The proposed method was able to correctly classify 95% of the foreign objects in 277 validation images. Here, as with the production processing applications described below, speed of operation is an important consideration. The combined segmentation and detection processing time was less than 50ms. Similar issues were described by Aslam et al. [26], who offer a survey of inspection methods for leather defect detection with various deep learning architectures compared. Nasiri et al. [27] describe an egg sorting system using a CNN to perform a classification task with an emphasis on highly structured illumination. They classify unwashed egg images into three classes: intact, bloody, or broken. A pre-trained CNN, VGG-16 (with weights pre-trained on the ImageNet dataset), was modified by replacing the last fully connected layers with a classifier block. This block included a global average pooling layer (to minimise over-fitting through decreasing the number of parameters), two dense layers with 512 neurons and the ReLU function, batch normalization (to maintain all inputs of the layer to the same range), dropout (as a regularization technique to prevent overfitting), and a final dense layer in the form of the Softmax classifier. This final layer computed the normalized probability value of the three classes. Once again, common augmentation techniques were used to increase the size of the manually labelled training dataset. Performance was shown to outperform traditional machine vision-based models, achieving an average overall accuracy of 94.84% (by 5-fold cross-validation).

7.3 Traditional manufacturing applications

While factory-based manufacturing of man-made objects offers the best opportunity for greater levels of environmental structuring, there are still many examples of how the introduction of deep learning can offer significant advantages.

Würschinger et al. [28] consider the application of deep learning in a range of more traditional machine vision applications involving manufacturing production lines. In these applications deep learning offers flexibility in terms of adaptation to differing feature extraction tasks using transfer learning, allowing for a low-cost and fast set-up. Transferred features can be effective for object recognition, subcategory recognition, and so lend themselves to this kind of domain adaptation. They review a range of representative examples where a classifier is used to detect defects, including in castings and printed circuit boards and for cutting tool wear. Interestingly, they point to industrial developers of machine vision solutions being initially hesitant to invest in deep learning projects, aware of the challenges and risks but also the great opportunities. They also identify the burgeoning added potential for deep learning systems to increase the transparency of manufacturing processes and so enhance process understanding, contributing to root cause analysis, and helping in the evaluation of quality control results, all aimed at improving efficiency. Ren et al. [29] present an interesting generic deep learning-based technique for automated surface inspection – a huge application area in manufactured products. Their classifier uses features transferred from a pre-trained deep learning network for image classification and defect segmentation. Results on publicly available datasets, including texture and colour industrial defects in hot-rolled strip, x-ray images of metal pipe, in welds and in timber products, outperformed that of hand-crafted features. AI systems are sometimes applied once the traditional defect inspection sequence completes, where their main purpose is to either classify the defects or reduce false positives – a concept we shall further explore later. Zhihong et al. [30] propose a deep learning method for robotic rubbish sorting, in which bottles are identified traveling on a conveyor within a complex cluttered background. A Fast R-CNN composed of two sub-nets: a region proposal generation (RPN) and VGG-16 model were used for object recognition and grasp pose estimation. In related work, Bahaghighat et al. [31] report 99% accuracy using an optimised fine-tuning method based on the VGG-19 CNN as an end-to-end deep learning method for inspection of bottle caps, classified in three groups (normal cap, unfixed cap, and no cap). Against the background of deep learning's need for large datasets, Li et al. [17] set machine vision for manufacturing in the context of big data, where computing efficiency has become a bottleneck to implementing real-time inspection systems for smart industries. They propose a multibranch deep model for a manufacturing inspection application for which a common deep model structure is modified and adapted to a fog environment - a decentralised computing structure located between the cloud and devices that produce data. Other examples of deep learning in manufacturing applications, that point to a wide base of application, include for quality control in the printing industry [32] and for crack detection in welds [33] where high accuracies of 98% and 96% respectively were reported.

7.4 Medical applications

A great deal of pioneering work in applying AI to image analysis has taken place in the medical field. Given that many of the concepts currently being explored have transferable application to other industries, this sector is worthy of some consideration. The large body of published work to date indicates that deep learning is having a huge impact within the medical sector, for medical imaging, medical data analysis, medical diagnostics and in wider healthcare. Previously, medical applications of machine vision have largely been limited to image processing, in which an image is enhanced in some way and then manually interpreted by a trained clinician. However, this requires a great deal of skill, is costly, and can be subjective and subject to error. Researchers have therefore started to explore how the image interpretation process might benefit from AI, potentially offering improved diagnostic performance.

Ayan et al. [34] use two well-known CNNs, Xception and VGG-16 with transfer learning and fine-tuning in the training stage for diagnosing of pneumonia from chest x-rays. It is notable that the Xception network was found to be more successful for detecting pneumonia cases, while the VGG-16 network was better at detecting normal cases, indicating that different networks have their own special capabilities on the same dataset and so the need for expertise in selection (see section 5.2.1 above). Litjens et al. [35] provide a general survey of deep learning in medical image analysis for organ detection, classification, segmentation, registration, and other tasks. Importantly, they, and others [36], point to future interest in unsupervised learning, such as generative adversarial networks (GANs). GANs use two competing CNNs - one generating artificial data samples and the other discriminating artificial from real samples, and can be trained to learn representative features in an unsupervised manner. This allows existing unlabelled data to be exploited more easily by leveraging the accessibility to big data, and so potentially has wide application well beyond that of medicine. The cost, time and effort involved in obtaining large sets of labelled data is one of the main limitations in deep learning – as we discussed above. Lundervold et al. [37] provide an overview of deep learning applied to analyse medical images, with a particular focus on one of the main areas – the MRI processing chain, from acquisition to image retrieval, and from segmentation to disease prediction. The authors conclude that even though there remain many challenges, the introduction of deep learning in clinical settings has quickly produced valuable results that are reflected in a large body of high-impact publications. One such example is that of Akkus et al. [38], who provide a review of deep learning for MRI brain segmentation, concluding how such techniques are outperforming previous state of the art classical machine learning algorithms. Liu et al. [39] also review deep learning in medicine, this time for the analysis of medical ultrasound images. They consider applications in biometrics, diagnosis, image guided intervention and therapy. Supervised deep models are already becoming widely used for the classification, segmentation, and detection of anatomical structures in medical ultrasound images, where CNNs and RNNs are the two most popular architectures. Low image quality, another issue which occurs across many applications, for example from poor target contrast, moving, non-rigid organs and the limited availability of medical labelled training data across multiple imaging modalities (such as MRI, X-ray, and ultrasound) are all major challenges typical of many medical applications. Application to 3D images is also seen as an important future area for development in this context. Inhomogeneity, varied intensity ranges, poor registration and contrast, together with noise, are all situations where the application of conventional techniques for pre-processing can help and we consider below other application examples of where conventional machine vision and deep learning techniques can benefit from a close integration.

8. Combining conventional machine vision with deep learning – the continuing important role of conventional techniques

Next, we consider recent examples across a range of application sectors where there has been an emphasis on the benefits of combining conventional machine vision techniques to improve the performance of deep learning. We have already seen one major example of this in the widespread use of image augmentation. Here we consider others, including for image pre-processing.

A study by Xie et al. [40] had the aim of combining a convolutional neural network with conventional computer vision techniques to automatically recognise and locate bones in

Atlantic salmon and to explore the impact of different image quality (obtained using differing levels of image compression) on performance, in what is considered a safety critical application. A Faster-RCNN object detection algorithm with three different convolutional neural network models (Alexnet, VGG-16 and VGG-19) were studied. An image compression ratio of 25% was identified as maximising costs and benefits with respect to detection accuracy, equipment prices and detection speed. Jiang et al. [41] describe the use of hyperspectral data for the detection of pesticide residue in post-harvest apples. They also fuse deep learning in the form of an Alex-Net-CNN with conventional techniques by using an Otsu segmentation and a Hough space transformation to first establish region of interest masks, centred on the apples. Detection accuracy was better than 95% and compared favourably with traditional k-nearest neighbour and support vector machine (SVM) classification algorithms. Related work for online fruit sorting using deep learning to detect internal mechanical damage of blueberries using hyperspectral data is presented by Wang et al. [42]. Another example implementation that combines conventional machine vision techniques with deep learning models for rapid defect identification in industrial process line inspection was presented by Wang et al. [43]. Their factory setting allowed the use of a fixed object pose and structured lighting to help reveal contamination defects in bottles. After applying an image processing stage in the form of a Gaussian filter and (Canny) edge detector, to generate an edge graph, a Hough transform was utilised to reduce the deep learning defect detection task complexity by establishing regions of interest to which a 'lightweight CNN' was applied. This illustrates how the availability of big data in the form of large numbers of example defects and defect free samples in many manufacturing applications lends itself to a deep learning solution. Furthermore, as with many manufacturing quality control applications of machine vision, the emphasis here was on efficiency in terms of speed of operation in a trade off with acceptable detection accuracy. Overall accuracy was reported to be as high as 99.6%, with a cycle time of just 47.6ms, allowing inspection of 21 products per second. These results compared favourably both with traditional feature based shallow machine learning, and other deep learning solutions. Li et al. [44] present a hybrid of conventional techniques and deep learning for detecting a wide range of defect types (scratches, floaters, light stains, and dark stains) on mobile phone screens during production. Conventional techniques were deployed in a pre-examination stage and regions of interest containing defect targets established using shape-based template matching. A two-stage approach, in which several deep learning models (VGG, ResNet, GoogLeNet, ResNeXt, SeNet, NasNet) were explored; all of which reached an impressive 99% detection accuracy, with the simpler VGG model achieving a cycle time of 4.56s and 2.47s on a CPU and GPU, respectively. Tang et al. [45] present an electrical component recognition method based on deep learning and conventional machine vision techniques. Conventional pre-processing operators, such as grayscale conversion, mean filtering, pose correction and other techniques were used (from the OpenCV library). Component coding of different types and materials were recognised by the CNN, with recognition results that compared favourably with traditional techniques across a wide range of components. Jang et al. [46] also propose a surface defect inspection solution but with an emphasis on a small training dataset in the context of a lack of defect sample images - typical of some manufacturing applications. Their proposed method for wafer/PCB defect inspection, a major application area from machine vision, exploited conventional defect inspection techniques to estimate a defect probability image. That, together with the original grey level image, formed the input to a CNN. Their integrating of conventional inspection techniques with a CNN model was shown to be highly effective for conventional defect inspection problems.

9. Future key challenges - promises and opportunities

The literature demonstrates that machine vision systems are being used across a plethora of industries for a wide range of tasks. Of these, quality assurance and inspection form the largest segment and this alone is expected to grow at a rate of 9.5% over the next four years [47]. New markets are also developing, and given the new capabilities offer by deep learning, there is a particular growing demand in outdoor applications. These evolving applications mean that the global machine vision system market is expected to rapidly grow from $8.6 Billion in 2020 to $17.7 Billion by 2027 [47], accelerating the demand for improved performance and new capabilities.

We have seen how the application of deep ANNs in industrial machine vision tasks can use biologically informed modelling of the vision system to achieve substantial leaps in performance. Deep learning can also help automate much of the model creation steps during development, and in application can be more robust and flexible, more adaptable to change and offer greater generality in application. However, it is also apparent that classical machine vision still has an important role to play. While some tasks are not suited to deep learning [48], for those that are, we have seen how classical techniques can often be combined with deep learning to substantially improve performance.

However, it is also clear from the literature that there are challenges remaining, and it seems appropriate to attempt to prioritise these as opportunities for future research. Firstly, deep learning necessitates a large amount of, usually labelled, training data. This may not always be accessible and there can be a scarcity of publicly available data for training CNNs. This means, as we have seen, that in practice few researchers train deep CNNs from scratch. Instead, many use networks that have been pre-trained on large-scale image data, where the trained weights are applied as a feature extractor to a smaller dataset in the solution domain. We have seen how this can be achieved by removing the last few layers of the pre-trained CNN. This allows a system developed for one application domain to be relatively easily transferred to another (e.g. human face to pig face biometric recognition [16]). However, this prompts the question as to whether future research should focus on devising new data augmentation methods for expanding often limited available data, or on how better to acquire real big data for each task, or will transfer learning be sufficient? Secondly, datasets have to be labelled, and in some applications (e.g. medical) by domain experts. We have also seen how this problem can be partly solved by leveraging effective data augmentation (and to a lesser extent automated annotation) techniques. Perhaps a more promising area of research, aimed at addressing both limited training data and the need for manual labelling, could be to explore methods for unsupervised learning, such as that already offered by generative adversarial networks (GANs). As we have seen, GANs offer an unsupervised form of learning, in which a generator network works in partnership with a discriminator network to, for example, mimic a human expert. Unsupervised learning avoids the need for manual data labelling by automatically discovering patterns in the data such that the network model can generate new outputs that plausibly could have been drawn from the original real dataset. In this regard, it is rather like generating augmented data. Impressive examples of the latter exist, where GANs have been used to create very realistic, but completely artificial, fake human faces [49]. Unsupervised learning and the ability to create realistic training data with wide variation could have a huge impact across many industrial applications where training data are limited.

It is widely recognised that deep learning is now the state-of-the-art in machine learning for machine vision and is being increasingly widely deployed across industrial applications. We have seen how conventional machine vision techniques can support deep learning applications and have highlighted some of the key challenges that will need to be addressed in the design and development of future CNN based solutions.

References

1. Y. LeCun, O. Matan, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel and H. S. Baird: Handwritten Zip Code Recognition with Multilayer Networks, in IAPR (Eds), Proc. of the International Conference on Pattern Recognition, II:35-40, IEEE, Atlantic City, invited paper, 1990

2. Y. LeCun, Y. Bengio, Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), 1995

3. Y. LeCun, L. Bottou, Y. Bengio and P. Haffner: Gradient-Based Learning Applied to Document Recognition, Proceedings of the IEEE, 86(11), 1998

4. A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems 25(2), DOI: 10.1145/3065386, 2012

5. A. Sioma, Automated Control of Surface Defects on Ceramic Tiles Using 3D Image Analysis, Materials, 13(5), 2020

6. M. L. Smith, L. N. Smith, Dynamic Photometric Stereo - A New Technique for Moving Surface Analysis, Image and Vision Computing, 23(9), 2005

7. B. L. Steward, J. Gai, L. Tang, The use of agricultural robots in weed management and control, Agriculture and Engineering Biosystems Publications, Iowa State University, DOI:10.19103/AS.2019.0056.13, 2019

8. L. N. Smith, A. Byrne, M. F. Hansen, W. Zhang, M. L. Smith, Weed classification in grasslands using convolutional neural networks, Applications of Machine Learning, SPIE Optics Photonics, DOI.org/10.1117/12.2530092, 2019

9. N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, J. Liang, Convolutional neural networks for medical image analysis: Full training or fine tuning?, IEEE Transactions on Medical Imaging, 35(5), 2016.

10. D. Erhan, Y. Bengio, A. Courville, P. A. Manzagol, P. Vincent, S. Bengio, Why does unsupervised pre-training help deep learning? Journal of Machine Learning Research, 11, Online: http://dl.acm.org/citation.cfm?id=1756006.1756025, 2010

11. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv 1409.1556, 09, 2014

12. K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016

13. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Efficient convolutional neural networks for mobile vision applications, CoRR, abs/1704.04861, Online: http://arxiv.org/abs/1704.04861, 2017

14. F. Chollet, Xception: Deep learning with depthwise separable convolutions, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), DOI.org/10.1109/CVPR.2017.195, 2017

15. A. Nasiria, A.Taheri-Garavand, Y. Zhang, Image-based deep learning automated sorting of date fruit, Postharvest Biology and Technology, 153, 2019

16. M. F. Hansen, M. L. Smith, L. N. Smith, M. G. Salter, E. M. Baxter, M. Farish, B. Grieve, Towards on-farm pig face recognition using convolutional neural networks, Computers in Industry, 98, 2018

17. L. Li, K. Ota, M. Dong, Deep Learning for Smart Industry: Efficient Manufacture Inspection System with Fog Computing, IEEE Transactions on Industrial Informatics, 14(10), 2018

18. H. Yang, S. Kumara, S. T.S. Bukkapatnam, F. Tsung, The internet of things for smart manufacturing: A review, IISE Transactions, Vol. 51, Iss. 11, 2019

19. H. A. M. Williams, M. H. Jones, M. Nejati, M. J. Seabright, J. Bell, N. D. Penhall, J. J. Barnett, M. D. Duke, A. J. Scarfe, H. S. Ahn, J. Y. Lim, B. Macdonald, Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms, Biosystems Engineering, 181, 2019

20. A. Wang, W. Zhang, X. Wei, A review on weed detection using ground-based machine vision and image, Computers and Electronics in Agriculture, 158 2019

21. F. Palaciosa, G. Buenoc, J. Salidoc, M. P. Diago, I. Hernández, J. Tardaguilaa, Automated grapevine flower detection and quantification method based on computer vision and deep learning from on-the-go imaging using a mobile sensing platform under field conditions, Computers and Electronics in Agriculture, 178, 2020

22. V. Kakani, V. H. Nguyen, B. P. Kumar H. Kim, V. R. Pasupuleti, A critical review on computer vision and artificial intelligence in food industry, Journal of Agriculture and Food Research, 2, 2020

23. A. Nasirahmadi, B. Sturm, S. Edwards, K-H. Jeppsson, A-C. Olsson, S. Müller, O. Hensel, Deep Learning and Machine Vision Approaches for Posture Detection of Individual Pigs, Sensors, 19(17), 2019.

24. D. Liuzz, G. Silano, F. Picariello, L. Iannelli, L. Glielmo, L. De Vito, P. Daponte, A review on the use of drones for precision agriculture, IOP Conference Series Earth and Environmental Science, 275, DOI: 10.1088/1755-1315/275/1/012022, October 2018

25. D. Rong, L. Xie, Y. Ying, Computer vision detection of foreign objects in walnuts using deep learning, Computers and Electronics in Agriculture, 162, 2019

26. M. Aslam, T. M. Khan, S. S. Naqvi, G. Holmes, R. Naffa, On the Application of Automated Machine Vision for Leather Defect Inspection and Grading: A Survey, IEEE Access, DIO:10.1109/ACCESS.2019.2957427, 7, 2019

27. A. Nasiri, M. Omid, A. Taheri-Garavand, An automatic sorting system for unwashed eggs using deep learning, Journal of Food Engineering 283 (2020)

28. H. Würschinger, M. Mühlbauer, M. Winter, M. Engelbrecht, Project Administration, N. Hanenkamp, Implementation and potentials of a machine vision system in a series production using deep learning and low-cost hardware, Procedia CIRP, 90, 2020

29. R. Ren, T. Hung, and K. C. Tan, A Generic Deep-Learning-Based Approach for Automated Surface Inspection, IEEE Transactions on Cybernetics, 48(3), 2018

30. C. Zhihong, Z. Hebin, W. Yanbo, L. Binyan, L. Yu, A Vision-based Robotic Grasping System Using Deep Learning for Garbage Sorting, Proceedings of the 36th Chinese Control Conference, China, July 2017

31. M. Bahaghighat, F. Abedini, M. S'hoyan, A. Molnar, Vision Inspection of Bottle Caps in Drink Factories Using Convolutional Neural Networks, IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, DOI:10.1109/ICCP48234.2019.8959737, 2019

32. J. Villalba-Diez, D. Schmidt, R. Gevers, J. Ordieres-Meré, M. Buchwitz, W. Wellbrock, Deep Learning for Industrial Computer Vision Quality Control in the Printing Industry 4.0, Sensors, 19, 3987, DOI:10.3390/s19183987, 2019

33. R. Moreno, E. Gorostegui-Colinas, P. López de Uralde, A. Muniategui, Towards Automatic Crack Detection by Deep Learning and Active Thermography, International Work-Conference on Artificial Neural Networks, IWANN, Advances in Computational Intelligence, DOI: 10.1007/978-3-030-20518-8_13, 2019

34. E. Ayan and H. M. Ünver, Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning, 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Turkey, DOI: 10.1109/EBBT.2019.8741582,2019

35. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A.W.M. van der Laak, B. van Ginneken, C. I. Sánchez, A survey on deep learning in medical image analysis, Medical Image Analysis, 42, 2017

36. M. Ruppel, R. Persad, A. Bahl, S. Dogramadzi, C. Melhuish, and L. N. Smith, NANCY: Combining Adversarial Networks with Cycle-Consistency for Robust Multi-Modal Image Registration, International Journal of Computer and Information Engineering, 14.8, 2020

37. A. S. Lundervold, A. Lundervold, An overview of deep learning in medical imaging focusing on MRI, Z MedPhys, 29, DOI.org/10.1016/j.zemedi.2018.11.002, 2019

38. Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, B. J. Erickson, Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions. J Digit Imaging, DOI: 10.1007/s10278-017-9983-4, 2017.

39. S. Liu, Y. Wanga, X. Yang, B. Lei, L. Liu, S. Xiang Li, D. Ni, T. Wang, Deep Learning in Medical Ultrasound Analysis: A Review, Engineering, 5, 2019

40. T. Xie, X. Li, X. Zhang, J. Hu, Y. Fang, Detection of Atlantic salmon bone residues using machine vision technology, Food Control, DOI.org/10.1016/j.foodcont.2020.10778730, 2020

41. B. Jiang, J. He, S. Yang, H. Fu, T. Li, H. Song, D.He, Fusion of machine vision technology and AlexNet-CNNs deep learning network for the detection of postharvest apple pesticide residues, Artificial Intelligence in Agriculture, 1, 2019

42. Z. Wang, M. Hu, G. Zhai, Application of Deep Learning Architectures for Accurate and Rapid Detection of Internal Mechanical Damage of Blueberry Using Hyperspectral Transmittance Data, Sensors, 4, DOI: 10.3390/s18041126, 2018

43. J. Wang, P. Fu, R. X. Gao, Machine vision intelligence for product defect inspection based on deep learning and Hough transform, Journal of Manufacturing Systems, 51, DOI:10.1016/j.jmsy.2019.03.002, 2019

44. C. Li, X. Zhang, Y. Huang, C. Tang, S. Fatikow, A Novel Algorithm for Defect Extraction and Classification of Mobile Phone Screen Based on Machine Vision, Computers & Industrial Engineering, 146, DOI:10.1016/j.cie.2020.106530, 2020

45. H. Tang, J. Chen, X. Zhen, Component recognition method based on deep learning and machine vision, ICIGP '19: Proceedings of the 2nd International Conference on Image and Graphics Processing, DOI.org/10.1145/3313950.3313962, 2019

46. C. Jang, S. Yun, H. Hwang, H. Shin, S. Kim, Y. Park, A Defect Inspection Method for Machine Vision Using Defect Probability Image with Deep Convolutional Neural Network, Computer Vision – ACCV 2018, DOI:10.1007/978-3-030-20887-5_9, 2019

47. Machine Vision Systems - Global Market Trajectory & Analytics, Global Industry Analysts, Inc, ID: 338506, July 2020

48. J. Bier, Is deep learning the solution to all computer vision problems?, Vision System Design, 1 March 2019

49. Horev, Rani. "Style-based GANs–Generating and tuning realistic artificial faces." LyrnAI: Deep Learning Explained. Available at: www.lyrn.ai/2018/12/26/a-style-based-generator-architecture-for-generative-adversarial-networks (accessed 29 January 2021), 2018

End of Paper

8888888888888888888888888888888888888888888888888888888888888888888