

Article

AI-enabled Efficient and Safe Food Supply Chain

Ilianna Kollia ^{2,†} , Jack Stevenson ¹ , Stefanos Kollias ^{1,2} ¹ School of Computer Science, University of Lincoln, Lincoln, UK² School of Electrical and Computer Engineering, National Technical University of Athens, Athens, Greece

* Correspondence: ilianna2@mail.ntua.gr

Abstract: This paper provides a review of an emerging field in the food processing sector, referring to efficient and safe food supply chains, 'from farm to fork', as enabled by Artificial Intelligence (AI). Recent advances in machine and deep learning are used for effective food production, energy management and food labeling. Appropriate deep neural architectures are adopted and used for this purpose, including Fully Convolutional Networks, Long Short-Term Memories and Recurrent Neural Networks, Auto-Encoders and Attention mechanisms, Latent Variable extraction and clustering, as well as Domain Adaptation. Three experimental studies are presented, illustrating the ability of these AI methodologies to produce state-of-the-art performance in the whole food supply chain. In particular, these concern: (i) predicting plant growth and tomato yield in greenhouses, thus matching food production to market needs and reducing food waste or food unavailability; (ii) optimizing energy consumption across large networks of food retail refrigeration systems, through optimal selection of systems that can get shut-down and through prediction of the respective food de-freezing times, during peaks of power demand load; (iii) optical recognition and verification of food consumption expiry date in automatic inspection of retail packaged food, thus ensuring safety of food and people's health.

Keywords: deep learning; deep neural networks; recurrent LSTM models; attention layers; latent variable extraction; domain adaptation; yield and growth prediction in greenhouses; energy optimization in retail refrigerator systems; verification and recognition of expiry date in retail food packaging.

1. Introduction

Food and drink processing is one of the largest manufacturing sectors worldwide, including all processing steps 'from farm to fork' [1]. The economic values of the AI-enabled precision farming market is estimated to grow and reach EUR 11,8 billion by 2025 globally [2]. There are significant challenges within these processing steps, regarding waste, safety and energy use [3]. Artificial Intelligence (AI) and Machine Learning (ML) technologies offer a transformative solution and recent Deep Learning (DL) approaches have innovated AI-enabled efficient yield and food production, food conservation and supply, reducing food waste and improving food safety.

This paper presents recent progress in the food production and supply pipeline achieved through the use of AI and DL technologies, focusing on three main tasks of the 'from farm to fork' pipeline: a) accurate prediction of yield growth and production in greenhouses; b) optimization of power consumption in food retailing refrigeration systems; c) quality control in retail food packaging.

At first, accurately predicting yield and food production is of great significance for reducing food waste and achieving smooth supply of food to supermarkets. Crop growers nowadays prefer using greenhouses than field growing. By using greenhouses, growers can extend the growing season and protect crops against changes of weather and temperature. In addition, in greenhouses, environmental parameters, such as temperature, humidity, radiation, carbon dioxide, soil quality and fertilization, can be controlled, providing a safe environment for crop growing [4,5].

Developing models which can effectively predict growth and yield can help growers improve the environmental control for better production, match supply and market demand and lower costs [6,7]. Greenhouse farming operations, yield and production prediction still rely heavily on human expertise. Automated yield and production prediction systems can let growers effectively anticipate weekly fluctuations and avoid problems of both over-demand and overproduction arising if the yield cannot be accurately predicted [8,9].

As with many bio-systems, plant growth is a highly complex and dynamic environmentally linked system. Therefore, growth and yield modeling is a significant scientific challenge [10,11]. Modeling approaches vary in a number of aspects, including, scale of interest, level of description and integration of environmental stress. There is a large number of tools that can help farmers in making decisions. These can provide yield rate prediction, suggest climate control strategies, or synchronise crop production with market demands.

According to [12,13] two basic modeling approaches are possible, namely, "knowledge-driven" and "data-driven" modeling. The knowledge driven approach relies mainly on existing domain knowledge, including biophysical models for specific species and plants [14,15]. In contrast, a data-driven modeling approach is capable of formulating a model solely from gathered data without necessarily using domain knowledge [16]. In this paper we focus on recently developed deep learning models, which can be trained with environmental data (CO_2 , humidity, radiation, outside temperature, inside temperature), as well as data for the actual yield and for significant plant characteristics, so as to accurately predict the yield, or these characteristics.

In the following phase of the pipeline, food enters the retailing phase, during which is stored in refrigeration systems of supermarkets. Food refrigeration accounts for a large percentage of electricity demands in developed countries, e.g., it is over 14% of UK's electricity demands, and mass refrigeration is responsible for a large percentage of carbon emissions, e.g. around 12% in UK. When energy consumption exceeds certain limits, this can cause unmanageable spikes in energy usage, with unpredictable problems [17,18]. To reduce load, countries impose response times on large retail energy consumers, requiring them to react urgently or face high financial penalties. For large companies with thousands of refrigerators, meeting this requirement relies on staff reactivity, a highly manualized, resource intensive, non-synchronous procedure.

A procedure to cope with such cases is to turn off some of the refrigerators in such cases for some period of time [19,20]. Using AI and machine learning, it is possible to predict which refrigerators to select and for how long to turn them off, whilst maintaining food quality and safety. In this paper we focus on a deep learning approach that has been recently developed for coping with such situations.

In the third phase of the pipeline, after assuring a safe and efficient storage of food in retailing refrigeration systems, the food is packaged and delivered to the shelves of supermarkets, so that customers can buy and consume it. Serious problems are also met in this phase as well, due to large amounts of food that remain unsold beyond their expiration date. Up to 30% of food is wasted each year, and food poisoning (including over 64,000 annual incidents in the UK from *Campylobacter* alone) is costly for the national healthcare system and individuals affected [21].

Food manufacturing faces the risk of product recalls and emergency product withdrawals caused by human error on packaging lines; if the expiry date is incorrectly listed as too early, consumers may believe that the product has reached the end of its shelf life and not consume, i.e., waste it. Conversely, if the expiry date exceeds the actual date, consumers may use the product beyond its safe timeframe, risking illness or potential fatality.

Due to prevalence of inkjet printers in food industry, characterized by high degree of quality variability, traditional Optical-Character-Recognition based vision systems have not been widely implemented, as they struggle with varied or distorted text. In this

paper we focus on recently developed deep learning systems that can identify and verify the presence and legibility of expiry date on food packaging photos captured while the products pass along production lines [22,23].

Section 2 presents the methods developed and used to implement the above described frameworks. Section 3 describes the application of the developed deep learning methodologies to real life environments and the achieved performance in all three phases of the pipeline. Discussion of the results that have been obtained and description of the future directions are provided in Section 4.

2. Materials and Methods

2.1. Fully Convolutional Networks

Fully Convolutional Networks (FCN) [24] do not contain dense layers, like typical Convolutional Neural Networks (CNN), but contain 1×1 convolutions that perform the task of fully connected layers. They are used in the experimental Section for text detection in images. Their architecture includes three parts: a feature extractor stem part, a feature merging branch part and the output part.

The feature extractor stem part is a PVANet [25], including convolutional and pooling layers. In our implementation, four feature map are extracted from an input image through the convolutional layers, enabling multi-scale detection of text regions of different sizes. Moreover, each pooling layer down-samples a corresponding feature maps by a factor of 2.

The extracted four feature maps, $f_i, i = 1, \dots, 4$, are then fed in the feature-merging branch part. In the i -th merging stage, the feature map f_{i-1} obtained in the former stage is fed to an unpooling layer for size doubling and then concatenated with the feature map f_i into a new feature map. A convolutional 1×1 operator is then applied, followed by a convolutional 3×3 operator that produces the merging stage output. Finally, a convolutional 3×3 layer produces the merging output and feeds it to the output layer. Multiple convolutional 1×1 operations are then used to produce the network outputs.

2.2. Long Short-Term Memories

Long Short-Term Memories (LSTM) are a variation of the Recurrent Neural Network (RNN) architecture [26]. Networks composed of LSTM units have been able to solve the gradient vanishing problem met in long-term time series analysis.

To achieve this, the LSTM structure contains three modules: the forget gate, the input gate and the output gate. The forget and input gates control which part of the information should be removed, or reserved to the network; the output gate uses the processed information to generate the provided output. LSTM units also include a Cell State, which allows the information to be saved for a long time.

2.3. Convolutional-Recurrent Neural Networks

Convolutional-Recurrent Neural Networks (CRNN) are lightweighted networks, which are used in the experimental section for text recognition. They include three parts, i.e., a feature extraction part, a bidirectional LSTM-RNN part and a transcription layer part.

The feature extraction part consists of a VGG network [27]. An input image is divided into T different image patches; feature vectors x_1, x_2, \dots, x_T are extracted from the different patches through convolutional and pooling layers. These feature vectors are then fed to a deep bidirectional Recurrent Neural Network (RNN) composed of recurrent layers with LSTM units. The RNN part can captures contextual dependencies between text in consecutive image patches. It is also able to operate on arbitrary lengths of text sequences.

The final layer in the CRNN is used for transcription, converting the text predictions made by the bidirectional LSTM-RNN into a label sequence. This is achieved through

maximization of a conditional probability given the bidirectional LSTM-RNN predictions [28].

2.4. Encoder-Decoder Model

In LSTM-based encoder-decoder models, the encoder part compresses the information from the entire input sequence into a vector composed of the sequence of the LSTM hidden states. Consequently, the encoder summarizes the whole input sequence (x_1, \dots, x_t) into the cell (C_0, \dots, C_{t-1}) and memory (h_1, \dots, h_{t-1}) state vectors and passes them to the decoder [29]. The latter uses this representation as initial state to reconstruct the time series $(S_{t'})$. The architecture employs two LSTM networks called the encoder and decoder, as shown in Figure 1.

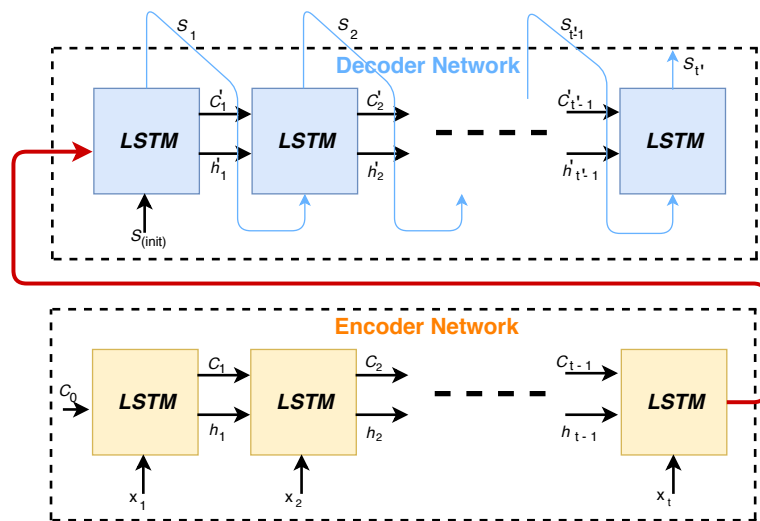


Figure 1. LSTM encoder decoder architecture

2.5. Attention Mechanisms

Attention mechanisms help to focus on feature segments of high significance [30]. An attention mechanism [31] models long-term dependencies by computing context vectors as weighted sums of all provided information. Such dependencies can be computed across the different internal LSTM layers, as well as over the LSTM output layers.

Output Predictions can be derived using the conditional probability distribution of the input signal and of the previous samples of the output. These are computed in terms of the current context, i.e., a vector holding information of which inputs are important at the current step. The context is derived from both the current state and the input sequence.

2.6. Performance Visualization

Two methodologies are examined in the paper for visualization of the performance of the developed prediction models, based either on class activation maps, or on latent variable extraction from trained deep models and adaptive clustering.

2.6.1. Class Activation Mapping

Deep neural networks are usually viewed as black boxes which do not provide means to visualize, or explain the decision making process. Class activation mapping (CAM) [32] and Gradient-weighted CAM [33] are efficient ways to visualize the significance of various parts of an input image for network training and decision making.

CAM-based methods provide such visualization by generating heat maps for each input image, focusing on the areas that influenced the network's prediction. In practice, they highlight the image pixels which are mainly used for classifying each image to a specific category.

2.6.2. Latent Variable Adaptive Clustering

Extraction of latent variables from trained deep neural networks and generation of concise representations through clustering them has been recently used as another means for visualising and explaining the decision making process [34–36].

Let us assume that for each input k , we extract M neuron outputs, as latent variables, from the trained deep neural network, forming a vector $\mathbf{v}(k)$. In total, we get:

$$\mathcal{V} = \{(\mathbf{v}(k), k = 1, \dots, N)\} \quad (1)$$

where N is the number of available training data.

We generate a concise representation of the \mathbf{v} vectors, that can be used as a backward model, to trace the most representative inputs for the performed prediction. This can be achieved using a clustering algorithm, e.g., k-means++ [?], which generates, say, L clusters $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_L\}$ by minimizing the following criterion:

$$\hat{\mathcal{Q}}_{k\text{-means}} = \arg \min_{\mathcal{Q}} \sum_{i=1}^L \sum_{\mathbf{v} \in V} \|\mathbf{v} - \mu_i\|^2 \quad (2)$$

where μ_i is the mean of \mathbf{v} values included in cluster i .

Then, we compute each cluster center $\mathbf{c}(i)$, generating the set \mathcal{C} , which constitutes the targeted concise representation:

$$\mathcal{C} = \{(\mathbf{c}(i), i = 1, \dots, L)\} \quad (3)$$

This representation can be used to provide analysis of variability between distributions and adapt information across different food datasets [37].

2.7. Domain Adaptation

Most of the deep learning models are developed through supervised learning over large annotated datasets. However, labeling large datasets consumes a lot of time and is cost ineffective. In addition, when deploying a trained model to a real-life processing pipeline, the assumption that both the source (training set) and the target one (application-specific environment) are drawn from the same distributions, may not be valid. In such cases, the deep learning model, trained on the source domain will not generalize well on the target domain. This is known as domain shift and training a model in the presence of domain shift is known as Domain Adaptation [38].

Many domain adaptation methods have been proposed in the last few years. Discrepancy, Adversarial and Reconstruction based approaches are the three primary domain adaptation approaches currently being applied to address the distribution shift [39]. Discrepancy-based approaches rely on aligning the distributions in order to minimize the divergence between them. The most commonly used discrepancy-based methods are Maximum Mean Discrepancy (MMD) [40], Correlation Alignment (CORAL) [41] and Kullback–Leibler divergence [42]. Adversarial-based approaches minimize the distance between the source and the target distributions using domain confusion, an adversarial method [43,44] inspired by Generative Adversarial Networks (GANs) [45]. Reconstruction-based approaches create a shared representation between the source and the target domains whilst preserving the individual characteristics of each domain [46,47].

In the experimental Section we focus on the use of domain adaptation for ensuring safety and waste reduction in the food supply chain. The domain adaptation model aims

at minimizing the feature discrepancy, for learning domain-invariant representations, the class boundary discrepancy, for minimizing the mismatch among different classifiers, and classification loss, for improving source data classification and leading to improved generalization on the target dataset [48].

The model jointly adapts features combining MMD with CORAL metrics, in order to align the underlying first and second order statistics. MMD defines the distance between the source and target distributions, with their mean embeddings in the Reproducing Kernel Hilbert Space. The CORAL Loss is also used to minimize the distance between the second order statistics (covariances) of the source and target features.

The total feature discrepancy loss is therefore given as:

$$Loss_{FD} = Loss_{MMD} + Loss_{CORAL} \quad (4)$$

In case we are considering domain adaptation from more than one source domains, the respective classifiers are likely to misclassify the target samples near the class boundary, since they are trained using different source domains. In such cases we also minimize the class discrepancy loss among classifiers, $Loss_{CD}$ [48], making their probabilistic outputs similar. Finally, since the network is trained with labeled source data, the classification, cross-entropy, loss, $Loss_{CL}$ is additionally minimized during training.

$$Loss_{TOTAL} = Loss_{FD} + Loss_{CD} + Loss_{CL} \quad (5)$$

3. Experimental Study

Figure 2 shows the described AI-enabled Food Supply Chain, including three pillars, which span food production, food storage and maintenance and food distribution and consumption:

- Food production in greenhouse environments, with a focus on predicting yield and optimizing crop growth and harvesting.
- Food storage and maintenance in retailing refrigerator systems, with a focus on reducing energy consumption and CO₂ production, whilst keeping food safe.
- Food distribution and consumption with a focus on quality control of retail packaging, through visual inspection of the food expiry date, while aiming to reduce food waste and avoid public health problems.

3.1. Food Production in Greenhouse Environments

3.1.1. Plant Growth Prediction

Machine learning techniques have been used to predict growth of Ficus plants using data collected from four cultivation tables in a greenhouse compartment of the Ornamental Plant Research Centre in Destelbergen, Belgium. Greenhouse microclimate was continuously monitored, while controlling the window openings, a thermal screen, an air heating system, assimilation light and a CO₂ adding system. In particular, ficus stem diameter was continuously monitored on three plants and data were collected corresponding to its hourly variation rate, i.e., as the difference between the current stem diameter and the stem diameter recorded on one hour earlier. The environmental data were also continuously recorded in hourly basis.

The prediction problem concerns one-step, two-step and three-step forecasting of the stem diameter [49,50]. In one-step forecasting, stem diameter measurements and environmental data collected, in a time window of the previous 15 hours, are used to predict the stem diameter value in the current hour. In two-step-ahead (6 hour) forecasting, the above data are used to predict the stem diameter six hours ahead. Three-step-ahead (12 hour) forecasting makes a stem diameter prediction 12 hours ahead.

Figure 3 presents an effective stem diameter prediction method, based on combination of the Encoder-Decoder (ED) model described in Section 2.4, the LSTM model described in Section 2.2. and the attention model (AM) described in Section 2.5 [51].

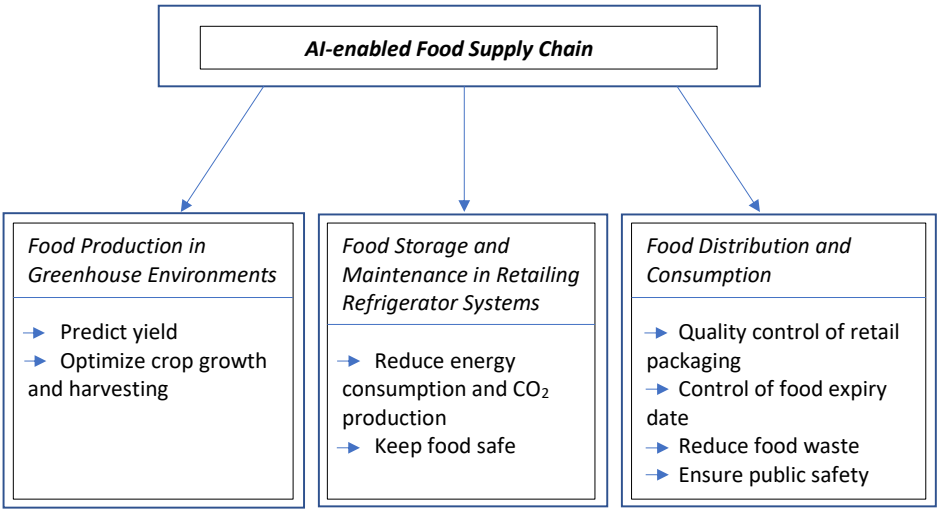


Figure 2. Food Supply Chain

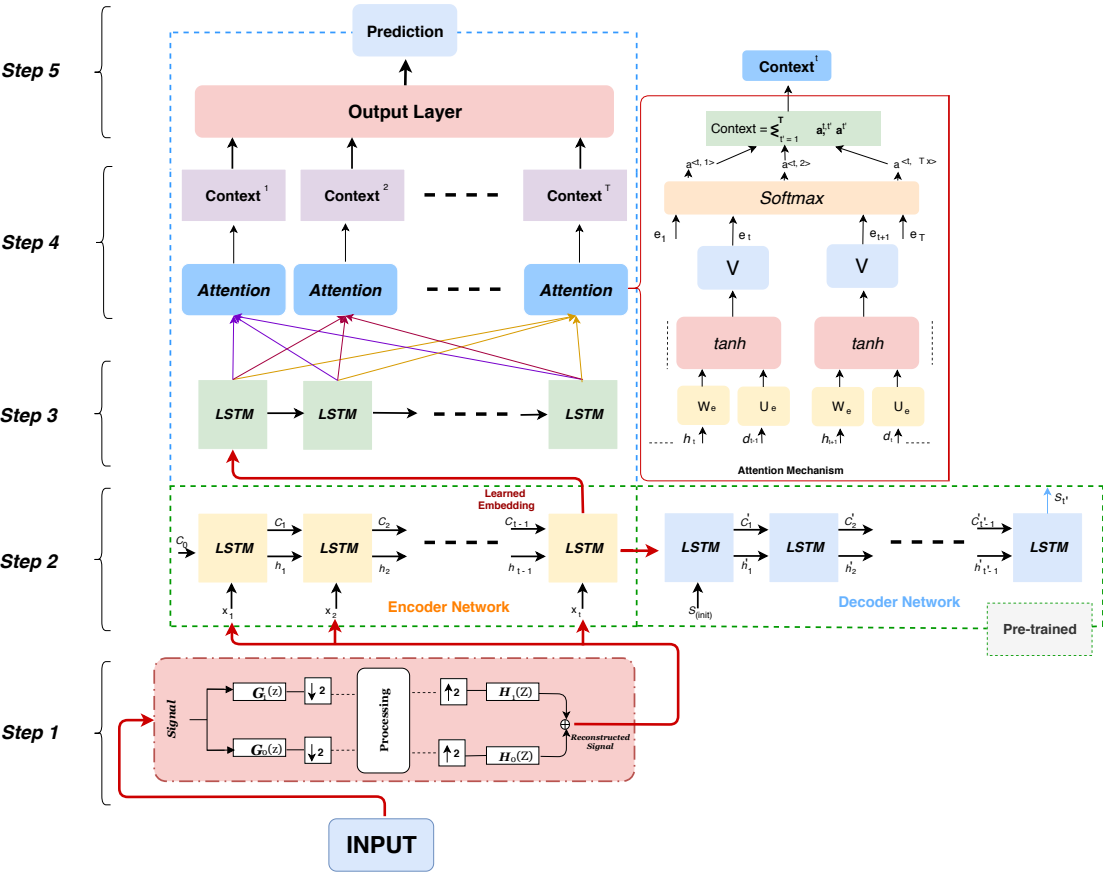


Figure 3. Deep architecture (WT-ED-LSTM-AM) for stem diameter prediction.

Table 1: Performance comparison of the WT-ED-LSTM-AM method to machine learning methods and ablation study

Method	RMSE		
	One Step Prediction	Two Step Prediction	Three Step Prediction
SVR	0.65	0.70	0.82
RFR	0.74	0.66	0.72
MLP	0.0034	0.0045	0.0048
LSTM	0.0031	0.0033	0.0054
WT-ED-LSTM	0.0028	0.0033	0.0042
ED-LSTM-AM	0.0034	0.0030	0.0046
WT-ED-LSTM-AM	0.0026	0.0028	0.0029

At first (Step 1), a denoising filter is applied to the input data, composed of the former stem diameter values and the respective environmental parameters. The filter is based on Wavelet Transform (WT) of the input data, removing the WT high frequency component, and generating a smoothed version of the input data. The ED model is then applied (Step 2). The encoder is pre-trained to extract useful and representative embeddings from the reconstructed time series data. A two-layer LSTM (128 and 32 neurons respectively) are used in the encoder shown in Figure 1. The decoder learns to generate the (reconstructed) input signal from the embeddings, thus optimizing this feature extraction procedure. Then, the LSTM network, with 128 neurons (Step 3), with attention mechanism (Step 4), is trained to model respective long-term dependencies, using the learned embedding as input features. A single layer neural network (Step 5) provides the final one-step, or multi-step ahead prediction provided by this WT-ED-LSTM-AM architecture.

In the accomplished experiments, the first 70% of data samples constituted the training set, the next 10% of data samples were the validation set and the rest 20% of data samples formed the test set. Min-max normalization was applied to the input data, scaling their values in the range [0, 1].

The experimental results illustrate the very good performance of the described method. Table 1 shows that the method outperforms all standard machine learning techniques, including a Support Vector Regressor (SVR), a Random Forest Regressor (RFR), a two-layer LSTM and a Multilayer Perceptron (MLP) with Stochastic Gradient Descent; a learning rate $1s = 0.001$ and a batch size of 32 were adopted. All models were trained for 100 epochs, using the same training, as well as validation and test data sets. The Root Mean Squared Error (RMSE) was used as performance evaluation criterion. To illustrate the contribution of the WT and AM steps, Table 1 also illustrates that the achieved performance gets worse, if either of these two steps is not included in the prediction approach.

Figure 4 shows the accuracy of Ficus growth one-step prediction by all methods for about 600 data samples. It can be seen that the described model successfully performs one-step ahead prediction, outperforming the other methods and providing accurate estimates of almost all peak values in the original data.

3.1.2. Yield Prediction

Tomato crop growing in greenhouse environments is a dynamic and complex system, with few models having been studied for it up to now [52,53]. These models are physics-dependent and represent biomass partitioning, crop growth, and yield as a function of several parameters. They are rather complex, with difficulty in estimating initial parameter values and need calibration in every environment.

The Tompousse model [14] was developed to predict tomato yield in terms of the weight of harvested fruits, based on linear relationship between flowering rate and fruit growth. Another tomato yield model [15] represented weekly yield fluctuations in terms

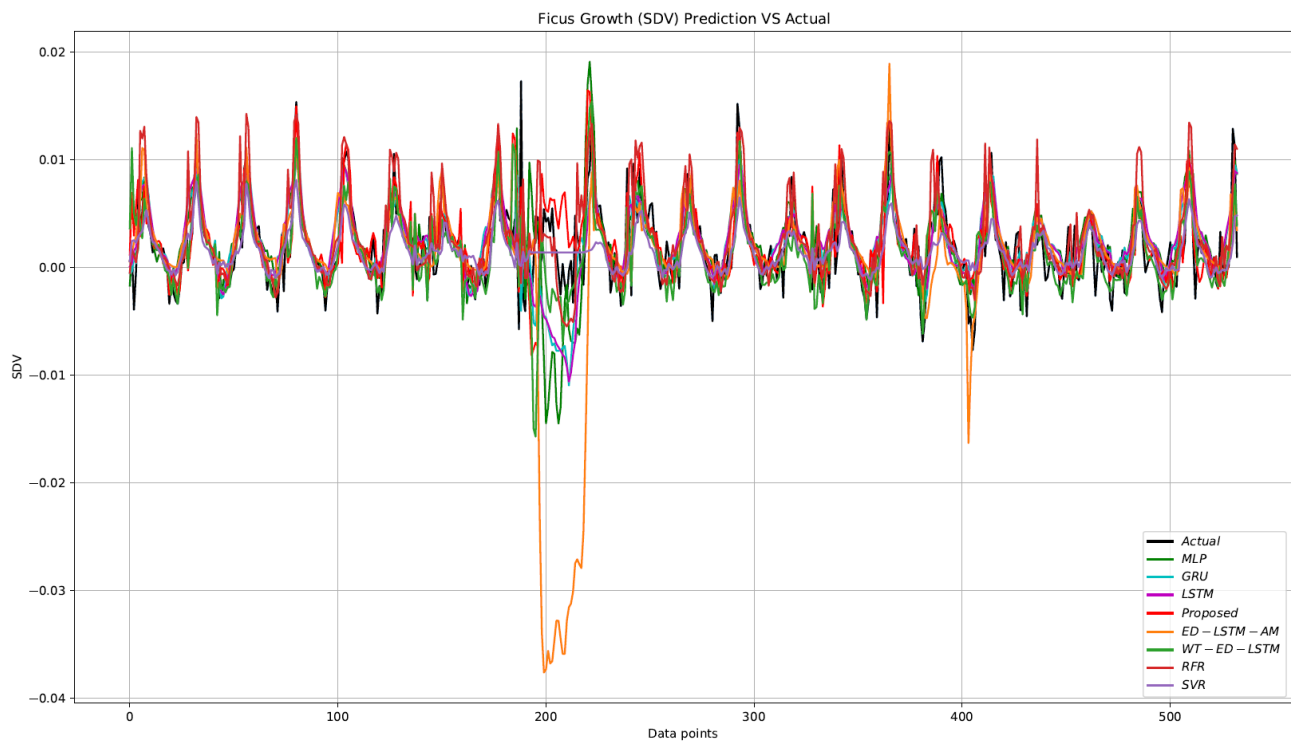


Figure 4. Performance comparison in one-step prediction (described and standard ML methods).

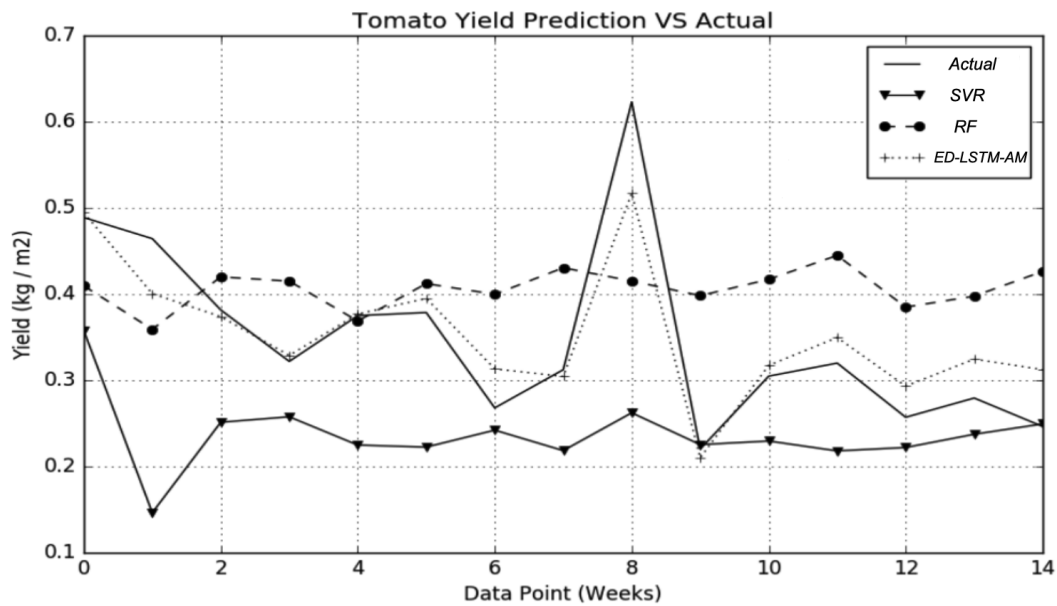


Figure 5. Performance comparison in tomato yield prediction (by ED-LSTM-AM and standard ML methods).

of fruit size and harvest rate. Such models [54] can help farmers in making yield rate prediction, suggesting climate control strategies and synchronising crop production with market demands.

Machine learning models have been recently developed for tomato yield prediction [55]. Such models have been trained with data collected from Greenhouse farms in United Kingdom, over a period of two years, including both environmental parameters, i.e., CO₂, humidity, radiation, outside temperature, inside temperature, as well as, yield actual measurements.

The environmental data were collected on an hourly basis, while the yield on a weekly basis. To deal with these data characteristics, data augmentation was performed, through interpolation, on the yield measurements, thus resulting in daily data measurements. Averaging of the hourly environmental data was also performed, so as to achieve similar daily representations.

Experiments with these data also showed the ability of the above-described ED-LSTM-AM model (WT did not provide any significant improvement in this case) to provide accurate prediction of tomato yield. Figure 5 shows that this model outperforms standard ML methods, including SVR and RFR, in tomato yield prediction. The MSE in actual yield prediction has been 0.015 for SVR, 0.040 for RFR and only 0.02 for ED-LSTM-AM.

3.2. Food Retailing Refrigeration Systems

Currently the majority of deep learning frameworks lack, or are still in their infancy, with regard to distribution. Novel solutions such as TensorFlow-distribute, or IBMs distributed deep learning are emerging to face this issue. The ability to share data between distributed nodes implementing, for example, the LSTM networks described in Section 2.2, is a critical bottleneck, despite the prevalence of databases providing very mature and advanced functionalities, such as MongoDB's aggregate pipeline. There are few parallelisable deep learning frameworks, which can broadly be categorised as follows [56]:

- Model parallelism (TensorFlow, Microsoft CNTK), where a single deep model is trained using a group of hardware instances and a single data set.
- Data parallelism, where each hardware instance is trained across different data.
- Hybrid parallelism, where a group of hardware trains a single model, but multiple groups can be trained simultaneously with independent data sets.
- Automatic selection, where different parts of the training/test process are tiled, with different forms of parallelism between tiles.

In the food supply chain, machine learning methods can be effectively used in a hybrid parallel framework for controlling massive amounts of food retailing refrigeration systems. This was shown in a recent implementation [57], in which multiple LSTM models were coupled with a MongoDB database [58] and were trained, using 110,000 real-life datasets - from about 1000 refrigerators.

The generated Nemesyst system [59] has been capable of predicting which refrigerators to select and how long to turn them off, whilst maintaining food quality and safety, in a Demand Side Response setting that modifies power demand load proportionally to available energy, in the National Grid of United Kingdom [60,61]. The target of the research is to show how to optimise refrigeration systems with machine learning at scale, whilst ascertaining that food temperature does not pass certain thresholds.

The Nemesyst system simultaneously wrangles, learns, and infers multiple models, with multiple and distinct data sets, across a network of refrigerator systems. In such systems the thermal inertia/mass of food acts as a store of cold energy. However, the thermal inertia in a retail refrigeration case changes, if food is actively shopped by consumers and then refilled, if ambient temperature changes, if networks of stores possess multiple refrigeration systems, or if different types of food need specific control mechanisms [62,63].

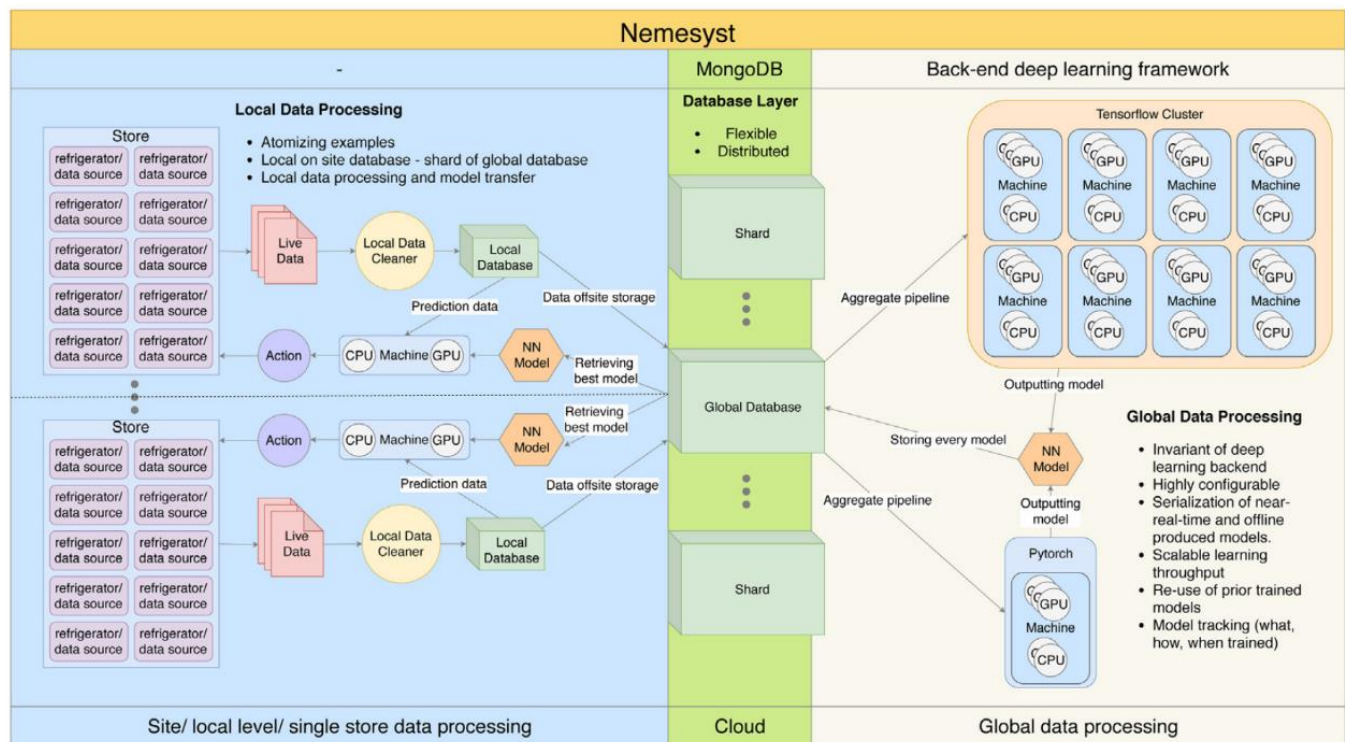


Figure 6. The Nemesyst system for AI-enabled power supply control of retailing refrigerators

Deep learning models, such as LSTMs, can model thousands of assets simultaneously, whilst being retrained to select which refrigerators to shed electrical load/turnoff across a massive pool. This requires an algorithm that can predict the thermal inertia in individual cases and therefore how long they can be shutdown (typically up to a maximum of up to 30 min), ensuring food safety and contributing to decarbonisation.

As shown in Figure 6, a backend deep learning framework aggregates data from a distributed global database layer and trains with them multiple deep models, such as LSTMs. It then stores the trained models in the database layer. The latter interacts with each local store and its refrigerator systems, which share local database instances and retrieve the best deep network from the global database layer.

In the experimental study the target was to predict the time (in seconds) until the refrigerator temperature rises from the point it is switched off until it breaches a food safety threshold. This duration varies between cases, for example the threshold is heuristically 8°C for high temperature refrigerators, and -183°C in low temperature refrigerators (freezers). Since the threshold for each fridge is not known, a two layer LSTM network is trained to predict the final point of defrost from the data as this is the most consistent point we can rely on.

As a consequence, a subset of fridges can be selected from the whole population, which are capable of remaining cool through the event requiring energy reduction. Knowing how long the refrigerators can remain off, before reaching this threshold, identifies the best candidates which reduce the power consumption sufficiently enough, while minimising the total number of affected refrigerators and potentially affected stock. The 110,000 defrost examples were split such that 10,000 were used for final testing, 10,000 were randomly selected each time for validation and the rest were used for training the system.

Figure 7 illustrates that an excellent prediction of the defrost time is achieved in four different cases. The difference between the last "observed" and final "defrost" value is the time (in seconds) that defrost can occur for before reaching the unsafe threshold of

8°C. Images (a) and (b) show cases where prediction takes place at the point of defrost; images (c) and (d) show cases where prediction occurs 2 minutes in advance of defrost.

3.3. *Quality Control in Retail Food Packaging*

The requirement for food high quality and safety of food and the public is a very important issue during the whole food supply chain [64]. The food product information printed on the food package is very important for food safety. Incorrectly labelled product information on food packages, such as the expiry date, can cause food safety incidents, like food poisoning. Moreover, such faults will incur high reputation and financial cost to food manufacturers, also causing large product recalls. As a consequence, verification of the correctness of the expiry date printed on food packages is of great significance.

In the following, it is shown that, instead of relying on operators to check the date code, an automated solution, taking photos of each date code, can verify the programmed date code for that product run, allowing food processing businesses to introduce unmanned operations and achieve very high inspection with full traceability, without compromising product safety. The production line will stop, if date code is incorrect, ensuring that the respective products are not released into the supply chain, protecting consumers, business margins and their brand, while reducing labour costs and food waste.

The Food Packaging Image dataset used next consists of more than 30,000 images classified in two categories (existing valid date and non-existing or non-valid date) from six different locations in Midlands in United Kingdom. The training process was carried out on a 70 % sample with another 10% used for the validation process. Finally, the remaining 20% of the images formed the test set. Representative images are shown in Figure 8, showing complete date (a), partial dates (b-c), unreadable date (d) and no existing date (e).

3.3.1. The FCN-CRNN Approach for Expiry Date Recognition

Optical character recognition (OCR) systems [65,66] can be applied to automatically recognize expiry date characters based on food package images captured by RGB cameras. However, existing OCR systems can not perform effectively in real-world expiry date recognition scenarios, with high variability, different fonts/angles, complicated designs with rich colours/textures, blurred characters poor lighting conditions in food manufacturing/retailer sites. Deep neural networks have been recently used as a means to tackle such problems [67,68].

In the following we present an approach composed of the FCN network, described in Section 2.1 and the CRNN network described in Section 2.3. Both networks are lightweighted and achieve good performance for text detection and recognition 'in-the-wild', are fine-tuned and combined together to detect and recognize the expiry date. Fine-tuning is performed through transfer learning [69], by adapting a model pre-trained for recognition of text, to recognition of expiry date.

The FCN-RCNN architecture is shown in Figure 9. It includes the FCN part, which is responsible for the detection of the Region-Of-Interest (ROI) of the expiry date. This acts as a filter to identify the image patch including the ROI from a whole food package photograph, so that the recognition task is performed on that specific small image patch, instead of the whole image. The second part includes the CRNN network, which performs date character recognition on the image patch obtained from the first network. Multiple levels of features are extracted from equally divided regions in the expiry date ROI, so as to recognize the characters within each region, while contextual relationships between characters in consecutive regions are modelled by the recurrent layers of CRNN.

The accuracy of expiry date detection by the fine-tuned FCN model reaches 98,2 %. It is higher than the accuracy obtained with two other similarly fine-tuned popular

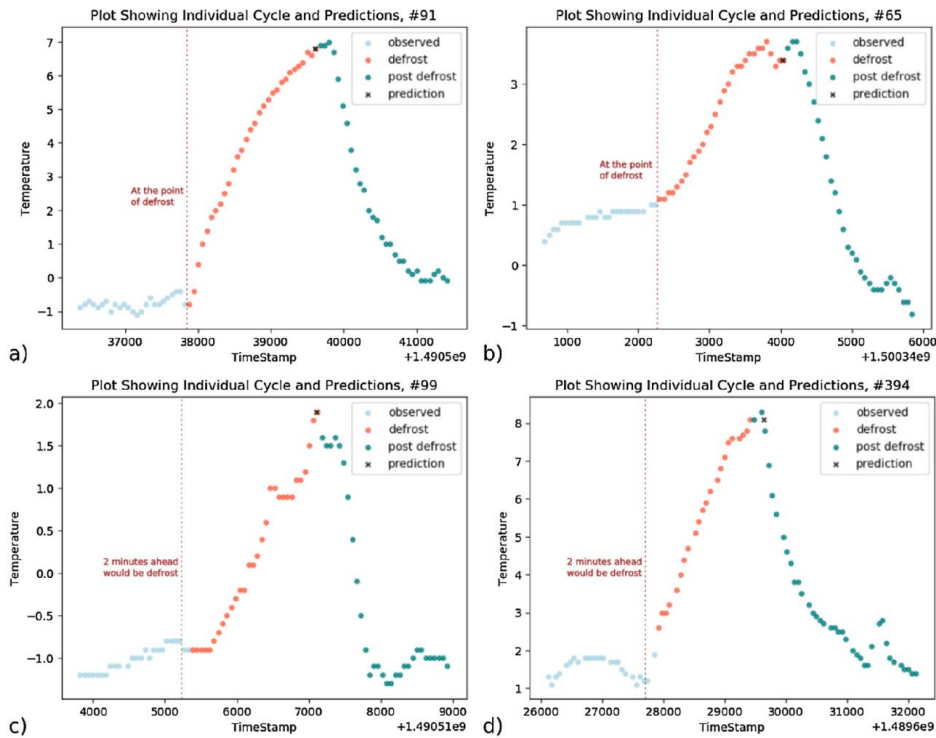


Figure 7. Prediction of the defrost time in four refrigerators: “Observed” (light blue) denotes training data; “defrost” (orange) is the ground truth; “prediction” (x) is the prediction of the final ground truth (orange) value.



Figure 8. Representative examples of food packaging images: (a) Complete date (day and month visible); (b) Partial date (no day visible); (c) Partial date (no month visible); (d) Unreadable; (e) No date.

deep neural networks, i.e., CTPN [70] and Seglink [71], while providing much fewer false alarms and miss detections, as shown in Table 2.

Table 2: Performance of the FCN detection sub-system

Method	Missing Detection (%)	False Alarm (%)	Accuracy (%)
FCN	1.67	0.28	98.20
CTPN [70]	2.79	16.57	92.20
Seglink [71]	5.71	12.53	93.73

Then the CRNN part of the method was fine-tuned using annotated (rectangular) image patches extracted by the FCN part. As shown in Table 3, the achieved recognition accuracy is 95,44 %, much higher than that provided by the Tesseract OCR tool [72] and higher than the performance of the TPS-ResNet-BiLSTM-Att network [73], which is much more complex (containing 6 times more parameters than the CRNN model).

Table 3: Performance of the CRNN recognition sub-system

Method	Accuracy (%)
CRNN	95.44
TPS-ResNet-BiLSTM-Att [73]	94.57
Tesseract OCR [72]	31.12

3.3.2. Latent Variable based Expiry Date Verification

A high variability exists in the image characteristics captured in different environments. As a consequence, an FCN model trained with images from a dataset collected in one location does not perform well when applied to images collected in another location. A method to cope with this problem is by extracting latent variables from each trained FCN; clustering them as described in Section 2.6.2 to produce respective cluster centroids; using these centroids as an equivalent model for evaluating images from other locations through nearest neighbor classification [37].

It has been shown that best results are achieved in this way, if latent variables are extracted by a dense layer preceding the output network layer [35,36,74]. For this reason, some dense layers are included on top of the FCN network, as shown in Figure 10. In particular, an averaging pooling and two dense layers, each containing 2048 ReLU units are added and the 2048 neuron outputs of the last dense layer are clustered using the k-means algorithm.

Let us consider the case of two datasets obtained from different locations. 7 clusters are generated per dataset citeref40, after applying the above procedure to each dataset. The respective 14 centroids are then merged, following an algorithm based on [35,75], which iteratively removes the lowest performing cluster, resulting in adaptation of some of the remaining centroids, until the expiry date verification stops improving.

Figure 11 shows the effect of the algorithm on the 14 original cluster centroids. Two of them (black stars) were excluded; four centroids (red squares) in class 1 and two (green circles) in class 2 were adapted and the rest CNN centroids remained unaffected. It should be mentioned that this produced an improvement in the verification accuracy of 13.8 % in total accuracy over the datasets of the two locations.

3.3.3. Domain Adaptation for Multi-source Expiry Date Recognition

The problem tackled in the former subsection, of high variability across datasets collected in different environments, as well as the general unavailability of labeled data across the whole food supply chain, is further examined next, using the domain adaptation approach described in Section 2.7.

Figure 12 shows the developed domain adaptation procedure. It comprises a feature extractor and a classification part. The feature extractor part learns useful representations,

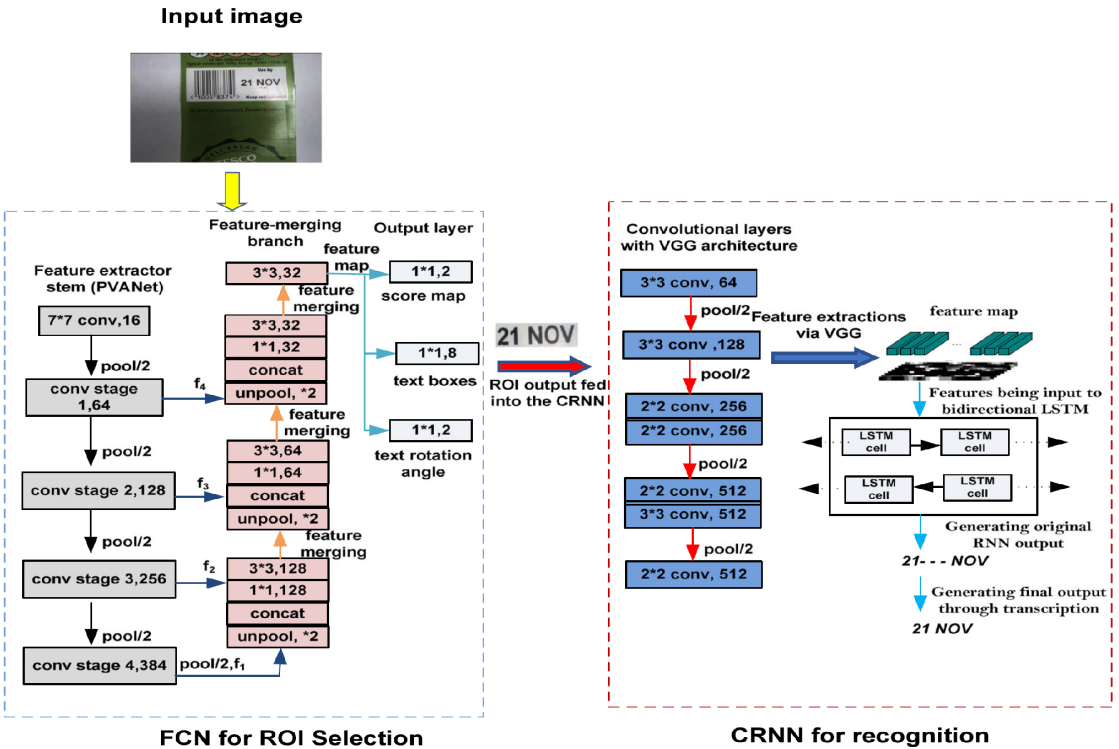


Figure 9. Architecture of the FCN-CRNN system

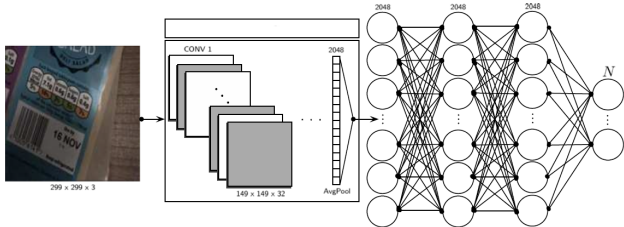


Figure 10. FCN part followed by dense layers - clustering the last layer neurons outputs

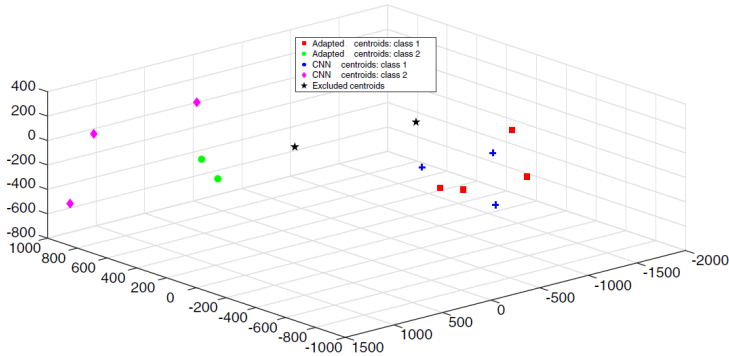


Figure 11. 3-D Visualization of the adapted, constant and excluded cluster centroids; class 1(2) includes good(bad) quality images

Table 4: Comparison of Performance of Domain Adaptation (DA) Methods

Method	Accuracy (%)
Single Source DA	84.14
Two Source Combined DA	85.05
Three Source Combined DA	86.13
Multi(2)-Source DA	90.53
Multi(3)-Source DA	92.50

with its sub-networks learning features specific to each source-target domain pairs. The classification part of the model learns domain-specific attributes for each target image. A Class Activation Map is included, as described in Section 2.6.1 to provide visualization of which parts of the input images affected most the provided predictions.

Experiments were made using a labeled single source dataset and an unlabeled single target dataset for all six locations. The goal of this experiment has been to establish a baseline for images that would be classified as readable and acceptable according to human standards. Further experiments were conducted using the multi-source domain adaptation approach, by using either two labeled source datasets and a single unlabeled target domain, or three labeled source datasets with a single unlabeled target domain.

For comparison purposes, additional experiments were carried out by combining the two, or the three source datasets into a single source dataset and adapting them towards the single target domain dataset. In the experiments, either the FCN - or the ResNet-50 [76] pretrained on ImageNet [77] - was the backbone network, by adding, a dense layer, and fine-tuning all convolutional and pooling layers.

Table 4 presents the average classification accuracy for all examined domain adaptation methods. As shown in the Table, the results of source combined methods are comparatively better than single source methods. This can be justified by considering this case as data enrichment, indicating that combining multiple source domains into single source domain is helpful. The multi-source domain adaptation approach significantly outperforms both other methods with an average classification accuracy improvement by more than 6 %.

4. Discussion

The presented experimental studies illustrated that the use of AI and ML methodologies can provide the food supply chain with efficiency and safety, reducing food waste and environment pollution; this is the main target of the data platforms required for food systems [78,79]. Most of the presented work has taken advantage of the recent progress in deep learning and deep neural networks.

In the first case study, DL systems were trained to predict growth and yield in time series data collected in greenhouses, taking into account environmental data and historical growth, or yield data. It has been shown that RNN/LSTM models combined with autoencoding, or attention mechanisms, were able to provide state-of-the-art performance compared with existing methods for multi-step prediction of stem diameter based growth of plants, as well as of tomato yield weekly prediction. Similar approaches can be used for prediction of yield of other categories, for example fruits, and in particular strawberry yield and their harvesting using robots [80,81].

In the second case study, it has been shown that control of the energy consumed by big food suppliers, e.g., large supermarket companies, is feasible, using deep learning based prediction. A large scale study, for automatic control, i.e., temporary shutdown of retailing refrigerator systems in peak energy demand hours, has been implemented, based on prediction of the specific refrigerator defrost period in each considered location. RNN/LSTM models were mainly used to perform the prediction task; GAN models were also successfully used to perform data augmentation and prediction. Similar methods

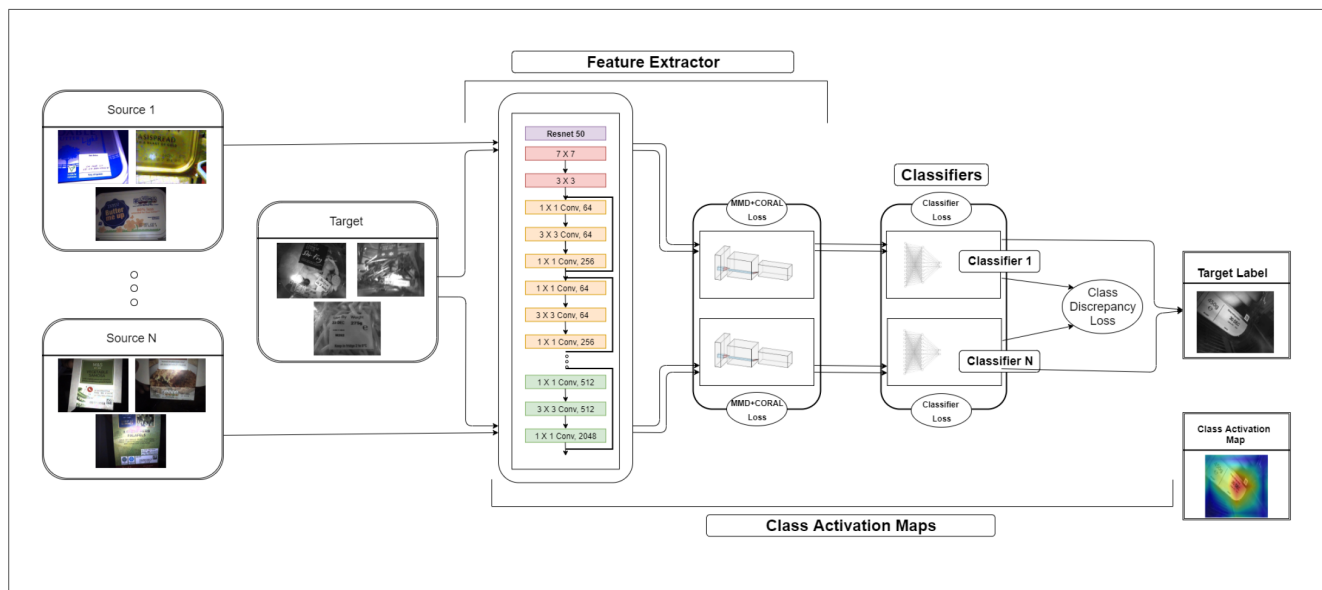


Figure 12. The Multi-source Domain Adaptation architecture

have been used to predict anomalies, i.e., instances when the power demand load gets higher than the available power in a country [82].

Current targets related to food supply chain, worldwide, include [3]: cut of greenhouse gas emissions by at least 55 % by 2030; development and use of energy-efficient approaches; optimization of connectivity to energy; analysis of production data; efficient management of resources, such as energy, water, soil, biodiversity. The presented approaches illustrate how AI and DL can be used to assist towards these goals.

The third case study illustrated that automatic recognition of expiry date in retail food packaging can also be successfully implemented through deep learning methodologies. The impact of this task is also crucial for ensuring food safety and consumers' health, whilst reducing food waste and related financial loss.

It was shown that FCN and CRNN models can be used to effectively recognize the expiry date on real life food packaging. Verification of the expiry date, with visualization of the decision making procedure can be achieved through latent variable extraction of the trained deep neural architectures and a related clustering procedure. Moreover, domain adaptation of the deep neural architectures can be performed so as to achieve expiry date verification and recognition across different retail environments.

The presented deep learning approaches can also be applied on wider aspects of food package control, such as the verification of the allergen labeling barcode, or other nutritional information. These can create a further impact towards people's safety, by automatically informing customers which allergens are included in the food products they are thinking of buying in retail supermarkets.

All above experiments have used the state-of-the-art in deep learning methodologies. In future research, the developments will focus on extending the described frameworks, with a target to combine deep learning with AI knowledge representation and reasoning, as well as multi-objective optimization technologies [83,84].

In this context, the methodologies that were presented in this paper for achieving an efficient and safe food supply chain will be extended, so as to create trustworthy AI-enabled food chain supply, that offer transparency and explainability to their users, i.e., food producers and farmers, retail supermarket owners, customers and the general public.

In particular, the deep neural architectures, especially the approach deriving the backward model of extracted cluster characteristics, described in Section 3.3.2, can

be interweaved with ontological representation [85] of this backward model, so as to provide explainable cues of the decision making procedure by using query answering techniques [86,87].

Moreover, the derived models will be adapted and contextualized across different environments, using general-purpose methodologies, as developed for other classification and prediction fields [88,89]. In addition, simultaneous maximization of yield and minimization of the consumed energy in greenhouses can be tackled through dynamic multi-objective optimization [90,91].

Author Contributions: The individual contributions of the authors are as follows: Conceptualization, Stefanos Kollias; methodology, Ilianna Kolli; software, Jack Stevenson; validation, Stefanos Kollias, Jack Stevenson; formal analysis, Ilianna Kolli; writing—original draft preparation, Ilianna Kolli and Jack Stevenson; writing—review and editing, Stefanos Kollias. All authors have read and agreed to the published version of the manuscript.

References

1. European Commission Communication, A 'farm to fork' strategy for a fair, h.; environmentally-friendly food system (COM(2020) 81 final). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52020DC0381>.
2. European Commission Sectoral Watch, Technological trends of the agri-food industry, S.. <https://ati.ec.europa.eu/reports/sectoral-watch/technological-trends-agri-food-industry>.
3. European Commission Coordinated Plan on Artificial Intelligence 2021 Review: <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review>.
4. Vandegehuchte, M.W.; Guyot, A.; Hubau, M.; De Groote, S.R.; De Baerdemaeker, N.J.; Hayes, M.; Welti, N.; Lovelock, C.E.; Lockington, D.A.; Steppe, K. Long-term versus daily stem diameter variation in co-occurring mangrove species: Environmental versus ecophysiological drivers. *Agricultural and Forest Meteorology* **2014**, *192*, 51–58.
5. Cohen, S.; Gijzen, H. The implementation of software engineering concepts in the greenhouse crop model hortisim1. *Acta Horticulturae* **1998**, pp. 431–440.
6. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture* **2018**, *147*, 70–90.
7. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*.
8. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and electronics in agriculture* **2018**, *151*, 61–69.
9. Daniel, J.; Andrés, P.U.; Héctor, S.; Miguel, B.; Marco, T.; others. A survey of artificial neural network-based modeling in agroecology. In *Soft Computing applications in industry*; Springer, 2008; pp. 247–269.
10. Kanai, S.; Adu-Gymfi, J.; Lei, K.; Ito, J.; Ohkura, K.; Moghaieb, R.; El-Shemy, H.; Mohapatra, R.; Mohapatra, P.; Saneoka, H.; Fujita, K. N-deficiency damps out circadian rhythmic changes of stem diameter dynamics in tomato plant. *Plant Science* **2008**, *174*, 183–191.
11. Moon, J.G.; Berglund, L.J.; Domire, Z.; An, K.N.; O'Driscoll, S.W. Stem diameter and micromotion of press fit radial head prosthesis: A biomechanical study. *Journal of Shoulder and Elbow Surgery* **2009**, *18*, 785–790.
12. Todorovski, L.; Džeroski, S. Integrating knowledge-driven and data-driven approaches to modeling. *Ecological Modelling* **2006**, *194*, 3–13. Special Issue on the Fourth European Conference on Ecological Modelling.
13. Nataša Atanasova, Ljupčo Todorovski, e.a. Application of automated model discovery from data and expert knowledge to a real-world domain: Lake Glumsø. *Ecological Modeling* **2008**, *212*, 92–98.
14. P. Abreu, J.F. Meneses, C.G. Tompousse, a model of yield prediction for tomato crops: calibration study for unheated plastic greenhouses. XXV International Horticultural Congress, Part 9: Computers and Automation, Electronic Information in Horticulture **1998**, pp. 141–150.
15. P. Abreu, J.F. Meneses, C.G. Predicting the weekly fluctuations in glasshouse tomato yields. IV International Symposium on Models for Plant Growth and Control in Greenhouses: Modeling for the 21st Century - Agronomic and Greenhouse Crop Models **1998**, pp. 19–23.
16. Fan, X.R.; Kang, M.Z.; Heuvelink, E.; de Reffye, P.; Hu, B.G. A knowledge-and-data-driven modeling approach for simulating plant growth: A case study on tomato growth. *Ecological Modelling* **2015**, *312*, 363–373.
17. Zhang, Q.; Yang, L.T.; Yan, Z.; Chen, Z.; Li, P. An Efficient Deep Learning Model to Predict Cloud Workload for Industry Informatics. *IEEE Trans. Ind. Informatics* **2018**, *14*, 3170–3178.
18. Granell, R.; Axon, C.; Wallom, D.; Layberry, R. Power-use profile analysis of non-domestic consumers for electricity tariff switching. *Energy Efficiency* **2016**, *9*.
19. Pallonetto, F.; De Rosa, M.; Milano, F.; Finn, D.P. Demand response algorithms for smart-grid ready residential buildings using machine learning models. *Applied Energy* **2019**, *239*, 1265–1282.
20. Panagiotidis, P.; Efraimis, A.; Xydis, G.A. An R-based forecasting approach for efficient demand response strategies in autonomous micro-grids. *Energy & Environment* **2019**, *30*, 63–80.

21. Pearson, S.; May, D.; Leontidis, G.; Swainson, M.; Brewer, S.; Bidaut, L.; Frey, J.G.; Parr, G.; Maull, R.; Zisman, A. Are Distributed Ledger Technologies the panacea for food traceability? *Global Food Security* **2019**, *20*, 145–149.
22. Thota, M.; Kollias, S.; Swainson, M.; Leontidis, G. Multi-source domain adaptation for quality control in retail food packaging. *Computers in Industry* **2020**, 123.
23. Gong, L.; Thota, M.; Yu, M.; Duan, W.; Swainson, M.; Ye, X.; Kollias, S. A novel unified deep neural networks methodology for use by date recognition in retail food package image. *Signal, Image and Video Processing* **2021**, 15.
24. Zhou, X.; Yao, C.; Wen, H.; Wang, Y.; Zhou, S.; He, W.; Liang, J. EAST: An Efficient and Accurate Scene Text Detector. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2642–2651.
25. Kim, K.; Cheon, Y.; Hong, S.; Roh, B.; Park, M. PVANET: Deep but Lightweight Neural Networks for Real-time Object Detection. *CoRR* **2016**, *abs/1608.08021*, [1608.08021].
26. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Computation* **1997**, *9*, 1735–1780.
27. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings; Bengio, Y.; LeCun, Y., Eds., 2015.
28. Shi, B.; Bai, X.; Yao, C. An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2298–2304.
29. Srivastava, N.; Mansimov, E.; Salakhudinov, R. Unsupervised Learning of Video Representations using LSTMs. Proceedings of the 32nd International Conference on Machine Learning; Bach, F.; Blei, D., Eds.; PMLR: Lille, France, 2015; Vol. 37, *Proceedings of Machine Learning Research*, pp. 843–852.
30. Geng, Z.; Chen, G.; Han, Y.; Lu, G.; Li, F. Semantic relation extraction using sequential and tree-structured LSTM with attention. *Information Sciences* **2020**, *509*, 183–192.
31. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473* **2014**.
32. Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning Deep Features for Discriminative Localization. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
33. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618–626. doi:10.1109/ICCV.2017.74.
34. Kollias, D.; Tagaris, A.; Stafylopatis, A.; Kollias, S.; Tagaris, G. Deep neural architectures for prediction in healthcare. *Complex & Intelligent Systems* **2018**, *4*, 119–131.
35. Wingate, J.; Kollia, I.; Bidaut, L.; Kollias, S. Unified deep learning approach for prediction of Parkinson's disease. *IET Image Processing* **2020**, *14*, 1980–1989.
36. Kollias, D.; Vlaxos, Y.; Seferis, M.; Kollia, I.; Sukissian, L.; Wingate, J.; Kollias, S.D. Transparent Adaptation in Deep Medical Image Diagnosis. Trustworthy AI - Integrating Learning, Optimization and Reasoning - First International Workshop, TAILOR 2020, Virtual Event, September 4-5, 2020, Revised Selected Papers; Heintz, F.; Milano, M.; O'Sullivan, B., Eds. Springer, 2020, Vol. 12641, *Lecture Notes in Computer Science*, pp. 251–267.
37. Ribeiro, F.D.S.; Calivá, F.; Swainson, M.; Gudmundsson, K.; Leontidis, G.; Kollias, S.D. Deep Bayesian Self-Training. *Neural Comput. Appl.* **2020**, *32*, 4275–4291.
38. Wang, M.; Deng, W. Deep visual domain adaptation: A survey. *Neurocomputing* **2018**, *312*, 135–153.
39. Wang, M.; Deng, W. Deep visual domain adaptation: A survey. *Neurocomputing* **2018**, *312*, 135–153.
40. Long, M.; Cao, Y.; Wang, J.; Jordan, M.I. Learning Transferable Features with Deep Adaptation Networks. *ArXiv* **2015**, *abs/1502.02791*.
41. Sun, B.; Saenko, K. Deep coral: Correlation alignment for deep domain adaptation. European conference on computer vision. Springer, 2016, pp. 443–450.
42. Zhuang, F.; Cheng, X.; Luo, P.; Pan, S.J.; He, Q. Supervised Representation Learning: Transfer Learning with Deep Autoencoders. Proceedings of the 24th International Conference on Artificial Intelligence. AAAI Press, 2015, IJCAI'15, p. 4119–4125.
43. Ganin, Y.; Lempitsky, V. Unsupervised Domain Adaptation by Backpropagation. Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37. JMLR.org, 2015, ICML'15, p. 1180–1189.
44. Shen, J.; Qu, Y.; Zhang, W.; Yu, Y. Wasserstein Distance Guided Representation Learning for Domain Adaptation. Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018; McIlraith, S.A.; Weinberger, K.Q., Eds. AAAI Press, 2018, pp. 4058–4065.
45. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2; MIT Press: Cambridge, MA, USA, 2014; NIPS'14, p. 2672–2680.
46. Ghifary, M.; Kleijn, W.B.; Zhang, M.; Balduzzi, D.; Li, W. Deep reconstruction-classification networks for unsupervised domain adaptation. Computer Vision – ECCV 2016; Leibe, B.; Matas, J.; Sebe, N.; Welling, M., Eds.; Springer: Berlin, 2016; Vol. 9908, pp. 597 – 613. 14th European Conference on Computer Vision (ECCV 2016), doi:10.1007/978-3-319-46493-0_36.

47. Kollias, D.; Zafeiriou, S. Training Deep Neural Networks with Different Datasets In-the-wild: The Emotion Recognition Paradigm. 2018 International Joint Conference on Neural Networks, IJCNN 2018, Rio de Janeiro, Brazil, July 8-13, 2018. IEEE, 2018, pp. 1–8.
48. Thota, M.; Kollias, S.; Swainson, M.; Leontidis, G. Multi-source domain adaptation for quality control in retail food packaging. *Computers in Industry* **2020**, 123.
49. Ben Taieb, S.; Bontempi, G.; Atiya, A.; Sorjamaa, A. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert Systems with Applications* **2011**, 39.
50. Duchesne, L.; Houle, D. Modelling day-to-day stem diameter variation and annual growth of balsam fir (*Abies balsamea* (L.) Mill.) from daily climate. *Forest Ecology and Management* **2011**, 262, 863–872.
51. Alhnaity, B.; Kollias, S.; Leontidis, G.; Jiang, S.; Schamp, B.; Pearson, S. An autoencoder wavelet based deep neural network with attention mechanism for multi-step prediction of plant growth. *Information Sciences* **2021**, 560, 35–50.
52. Jones, J.; Kenig, A.; Vallejos, C. Reduced State-Variable Tomato Growth Model. *Transactions of the ASABE* **1999**, 42, 255–265.
53. Heuvelink, E. Tomato growth and yield : quantitative analysis and synthesis. 1996.
54. K. Qaddoum, E. Hines, D.I. Yield Prediction for Tomato Greenhouse Using EFuNN. *ISRN Artificial Intelligence*, 2013.
55. Alhnaity, B.; Pearson, S.; Leontidis, G.; Kollias, S. Using Deep Learning to Predict Plant Growth and Yield in Greenhouse Environments. *Acta Horticulturae*. International Society for Horticultural Science, 2020, Vol. 1296, *Acta Horticulturae*, pp. 425–431. Funding Information: This work is part of EU Interreg SMARTGREEN project (2017-2021). We would like to thank all the growers (UK & EU), for providing the data. Their valuable feedback, suggestions and comments are highly appreciated to increase the overall quality of this work.
56. Panda, D.; Awan, A.A.; Subramoni, H. High performance distributed deep learning: a beginner's guide. 2019, pp. 452–454.
57. Onoufriou, G.; Bickerton, R.; Pearson, S.; Leontidis, G. Nemesyst: A hybrid parallelism deep learning-based framework applied for internet of things enabled food retailing refrigeration systems. *Comput. Ind.* **2019**, 113.
58. Dede, E.; Govindaraju, M.; Gunter, D.; Canon, R.S.; Ramakrishnan, L. Performance Evaluation of a MongoDB and Hadoop Platform for Scientific Data Analysis. Proceedings of the 4th ACM Workshop on Scientific Cloud Computing; Association for Computing Machinery: New York, NY, USA, 2013; Science Cloud '13, p. 13–20.
59. Nemesyst. <https://github.com/DreamingRaven/Nemesyst>.
60. Bradley, P.; Coke, A.; Leach, M. Financial incentive approaches for reducing peak electricity demand, experience from pilot trials with a UK energy provider. *Energy Policy* **2016**, 98, 108–120.
61. Grünewald, P.; Torriti, J. Demand response from the non-domestic sector: Early UK experiences and future opportunities. *Energy Policy* **2013**, 61, 423–429. doi:10.1016/j.enpol.2013.06.0.
62. Albayati, I.; Postnikov, A.; Bingham, C.; Bickerton, R.; Zolotas, A.; Pearson, S. Aggregated power profile of a large network of refrigeration compressors following FFR DSR events. International Conference on Energy Engineering, 2018.
63. Saleh, I.M.; Postnikov, A.; Arsene, C.; Zolotas, A.C.; Bingham, C.; Bickerton, R.; Pearson, S. Impact of Demand Side Response on a Commercial Retail Refrigeration System. *Energies* **2018**, 11.
64. FAO, W. <http://www.fao.org/docrep/012/a1552e/a1552e00.html> (2009).
65. Mori, S.; Suen, C.; Yamamoto, K. Historical review of OCR research and development. *Proceedings of the IEEE* **1992**, 80, 1029–1058.
66. Ribeiro, F.D.S.; Gong, L.; Calivá, F.; Swainson, M.; Gudmundsson, K.; Yu, M.; Leontidis, G.; Ye, X.; Kollias, S.D. An End-to-End Deep Neural Architecture for Optical Character Verification and Recognition in Retail Food Packaging. 2018 IEEE International Conference on Image Processing, ICIP 2018, Athens, Greece, October 7-10, 2018. IEEE, 2018, pp. 2376–2380.
67. Suh, S.; Lee, H.; Lee, Y.O.; Lukowicz, P.; Hwang, J. Robust Shipping Label Recognition and Validation for Logistics by Using Deep Neural Networks. 2019 IEEE International Conference on Image Processing, ICIP 2019, Taipei, Taiwan, September 22-25, 2019. IEEE, 2019, pp. 4509–4513.
68. Katyal, N.; Kumar, M.; Deshmukh, P.; Ruban, N. Automated Detection and Rectification of Defects in Fluid-Based Packaging using Machine Vision. IEEE International Symposium on Circuits and Systems, ISCAS 2019, Sapporo, Japan, May 26-29, 2019. IEEE, 2019, pp. 1–5.
69. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* **2010**, 22, 1345–1359.
70. Tian, Z.; Huang, W.; He, T.; He, P.; Qiao, Y. Detecting Text in Natural Image with Connectionist Text Proposal Network. ECCV, 2016.
71. Shi, B.; Bai, X.; Belongie, S. Detecting Oriented Text in Natural Images by Linking Segments. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3482–3490.
72. Tesseract-ocr. <https://github.com/tesseract-ocr/tesseract> (2018).
73. Baek, J.; Kim, G.; Lee, J.; Park, S.; Han, D.; Yun, S.; Oh, S.J.; Lee, H. What Is Wrong With Scene Text Recognition Model Comparisons? Dataset and Model Analysis. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 4714–4722.
74. Kollias, D.; Bouas, N.; Vlastos, Y.; Brillakis, V.; Seferis, M.; Kollia, I.; Sukissian, L.; Wingate, J.; Kollias, S. Deep Transparent Prediction through Latent Representation Analysis, 2020, [[arXiv:cs.LG/2009.07044](https://arxiv.org/abs/2009.07044)].
75. Kollias, S.D.; Anastassiou, D. A unified neural network approach to digital image halftoning. *IEEE Trans. Signal Process.* **1991**, 39, 980–984.
76. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity Mappings in Deep Residual Networks. Computer Vision – ECCV 2016; Leibe, B.; Matas, J.; Sebe, N.; Welling, M., Eds.; Springer International Publishing: Cham, 2016; pp. 630–645.

77. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; Berg, A.C.; Fei-Fei, L. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)* **2015**, *115*, 211–252.
78. Martindale, W.; Duong, L.; Hollands, T.; Swainson, M. Testing the data platforms required for the 21st century food system using an industry ecosystem approach. *The Science of the total environment* **2020**, *724*, 137871.
79. Martindale W, Wright I, K.L.e.a. Framing food security and food loss statistics for incisive supply chain improvement and knowledge transfer between Kenyan, Indian and United Kingdom food manufacturers. *Emerald Open Res* **2020**, *2*.
80. Kirk, R.; Cielniak, G.; Mangan, M. L*a*b*Fruits: A Rapid and Robust Outdoor Fruit Detection System Combining Bio-Inspired Features with One-Stage Deep Learning Networks. *Sensors* **2020**, *20*, 275.
81. Xiong, Y.; Ge, Y.; Grimstad, L.; From, P.J. An autonomous strawberry-harvesting robot: Design, development, integration, and field evaluation. *J. Field Robotics* **2020**, *37*, 202–224.
82. Kollias, I.; Kollias, S.D. A Deep Learning Approach for Load Demand Forecasting of Power Systems. IEEE Symposium Series on Computational Intelligence, SSCI 2018, Bangalore, India, November 18-21, 2018. IEEE, 2018, pp. 912–919.
83. Trustworthy AI - Integrating Learning, Optimisation and Reasoning: <https://tailor-network.eu>.
84. Network, C.R. <https://claire-ai.org/network/>.
85. Kollias, I.; Simou, N.; Stamou, G.; Stafylopatis, A. Interweaving Knowledge Representation and Adaptive Neural Networks.
86. Glimm, B.; Kazakov, Y.; Kollias, I.; Stamou, G.B. Using the TBox to Optimise SPARQL Queries. Informal Proceedings of the 26th International Workshop on Description Logics, Ulm, Germany, July 23 - 26, 2013; Eiter, T.; Glimm, B.; Kazakov, Y.; Krötzsch, M., Eds. CEUR-WS.org, 2013, Vol. 1014, *CEUR Workshop Proceedings*, pp. 181–196.
87. Kollias, I.; Glimm, B.; Horrocks, I. Answering Queries over OWL Ontologies with SPARQL. Proceedings of the 8th International Workshop on OWL: Experiences and Directions (OWLED 2011), San Francisco, California, USA, June 5-6, 2011; Dumontier, M.; Courtot, M., Eds. CEUR-WS.org, 2011, Vol. 796, *CEUR Workshop Proceedings*.
88. Kollias, D.; Yu, M.; Tagaris, A.; Leontidis, G.; Stafylopatis, A.; Kollias, S.D. Adaptation and contextualization of deep neural network models. 2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017, Honolulu, HI, USA, November 27 - Dec. 1, 2017. IEEE, 2017, pp. 1–8.
89. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 2018. cite arxiv:1810.04805Comment: 13 pages.
90. Jiang, S.; Kaiser, M.; Yang, S.; Kollias, S.D.; Krasnogor, N. A Scalable Test Suite for Continuous Dynamic Multiobjective Optimization. *IEEE Trans. Cybern.* **2020**, *50*, 2814–2826.
91. Jiang, S.; Yang, S. A Strength Pareto Evolutionary Algorithm Based on Reference Direction for Multiobjective and Many-Objective Optimization. *IEEE Transactions on Evolutionary Computation* **2017**, *21*, 329–346.