

## **An Intelligent Edge Computing-based Scalable Architecture for Large-scale Smart Farm System**

Hee-Dong Park\*

Department of Smart IT, Joongbu University, #305 Dongheon-ro, Dugyang-gu,  
Goyang, Gyeonggi-do, Korea

*hdpark@joongbu.ac.kr*

**Abstract.** The recent smart farm system, which can be controlled remotely and automatically, is demanding for an intelligent strategy due to its limitations in data processing and storage capacity. Therefore, cloud server computing approach is generally used for such functions with complex algorithms, but the explosive increase of data traffic causes network congestion and overhead in central cloud. To overcome the limitations of the cloud-centric systems, the concept of edge computing has become important. In this paper, we propose an intelligent edge computing-based scalable architecture for smart farm, which is based on two-stage control system for intelligent management using conditional and correlation analysis. Since the cloud server collects preprocessed data from correlation stage, it processes smaller capacity data than the original of the edge computer, thereby reducing the congestion of the entire network and the processing burden on the server. It is possible to determine the optimal next state of farm environment according to the growth of the crop from the previous state in consideration of the correlation of input parameters using deep learning model. Cooperation or standalone management of large-scale farmland can be possible using network interface with lower conditional control stage and intelligent filtered data from upper cloud server. Our proposed edge computing-based architecture can be referenced as an architecture framework for intelligent smart farm standardization in the future, with scalable structure that can be connected and extended to other correlation control stages of edge computers.

**Keywords:** Edge computing, smart farming, intelligent farm control, two-stage structure, internet of things, correlation analysis, conditional control.

## **1. Introduction**

In order to perform farm activities such as watering or fertilizing, farmers need to visit their farms frequently. Due to the impacts from climate change, aging and declining populations, limitations of land for agricultural production make food security as a global issue. As there is labor shortage caused by the reduction of young people in the traditional agricultural industry, smart farms are attracting attention with the development of current technology related to the 4th Industrial Revolution. Smart farm refers to a farm which can properly manage and control the growing environment of crops and livestock remotely and automatically using Information Technology (IT). The farm environment can be managed with the measured parameters such as temperature, humidity, carbon dioxide (CO<sub>2</sub>), soil status, etc.

The types of actuators used in the greenhouse are nutrient supply, ventilation fans, coolers, heaters, and motors for side or ceiling, and they can be controlled according to actuator control algorithm, thus creating an optimal growth environment. There are several advantages of a smart farm using Internet of Things (IoT). Large amounts of data from various sensors, such as climate conditions, soil status, and plant growth processes, are useful for efficient farm management. Efficient processing with automated control reduces the risk in production management and make possible advanced standardized management of crop quality and growth speed. In general, the flow of a smart farm control could be divided into a device-layer where sensors and drivers are connected, a network-layer for data transmission, and a cloud-layer for data storage and analysis. As the connection between the device-layer and the cloud-layer becomes important, however, it required the advanced actuator control algorithm according to various environmental conditions between connected sensors, the analysis and prediction of collected data, and the operating structure for intelligent control with statistical correlation processing.

Until now, the important function of smart farm is the automatic control service based on real-time field data, which includes detecting internal and external environment and controlling various devices (Yeo et al., 2016; Mahendran et al., 2017). At the same time, data from sensors or control actions are stored in local server or controller in real time and used for statistical data or intelligent growing environment processing. The integrated control systems operated on the farm are getting higher performance and the functions are becoming more complex, and even when the network or cloud server is not connected, the control system must provide not only a whole farm management but also each farm's control functions. This ensures that even when a failure occurs, the minimum farm automation function needs to be carried out. It needs a smart farm control system that provides not only existing integrated management, but also intelligent farm monitoring, data collection and sharing, prediction and diagnosis functions. Therefore, it requires the

concept of edge computing, that is being developed with the lightweight big data processing as well as built-in artificial intelligence (AI) technology.

The collected data itself may not be important, but data analysis is the most important part of smart agriculture, and powerful analysis that can apply predictive algorithm and machine learning is required. Various studies (Lee et al., 2020; Park et al., 2017; Axel et al., 2020) are being conducted using big data and AI technology for such intelligent control functions, and the need for such research is also increasing, especially for storing and analyzing collected data at the device level and applying it to various correlated system status. For example, it is possible to make automatic control of tractors or agriculture machines and collect surrounding environment data from sensors. Such automated units can calculate the optimal work path or jobs to perform using computer vision and pre-obtained field data with the locations of various devices and can avoid collisions or obstacles with humans or animals. It also enables irrigation and sowing work, as well as automation of harvesting.

Edge computing (Inés and Juan, 2019; Ai et al., 2018) began to be an approach in the limitations of efficiency of cloud servers that demands distributed computing on IT environments where increased real-time or large-scale data processing is needed. The advantage of using edge computing applied to smart farms is precision agriculture. This approach can make efficiency of controls and processes in agriculture and thus reducing costs. Farms can be automated using IoT edge computing, which can handle data collection and various agricultural processes alone. Therefore, edge computing can make decisions within local area without a connection to the main server, which has reliability and cost savings, enabling more sustainable farming. Edge computing can also be prepared for potential environmental hazards and natural disasters but does not replace all the functions performed by cloud services. The concept of edge computing was developed to compensate for the shortcomings of cloud computing. Because of the complementary relationship between cloud and edge computing, they will continue to develop into a coexisting form, and since edge computing is not a mature technology yet, each solution will be developed in their own strategy. The cloud-edge-device architecture will make cloud computing infrastructure more efficient, and it is currently spreading to various industries.

Large-scale smart farms also need to use high-performance computing chips while processing with big data function. There can be several problems to consider in the existing cloud-based smart farm structure. It includes security issues while using network, network congestion due to data collection and transmission, high-performance processing issues due to enormous data analysis, and cloud computing costs as the number of connected precision smart farms increases. In case of CCTV (Closed-Circuit Television), most operating time is spent on standby, so most of the data are not meaningful. Edge computing technology can be used for motion

detection, filtering, and tracking to select the necessary data, but motion itself does not always mean that there occurs an abnormal status. Therefore, most of the video data transmitted to the cloud may be meaningless and incurs unnecessary network costs. If a camera detects a motion and sends it to a server or control station through the Internet, the actual network load will not be different from the CCTV for motion detection or tracking. Therefore, for the CCTV to be meaningful, it requires the technology that can analyze the video data more accurately and send the filtered data to the cloud.

The agricultural environment still has several disadvantageous conditions, including the stability of Internet and farmers' lack of understanding on Information and Communications Technology (ICT) environment. In addition, there exists actual acceptance sense from smart farms in terms of social awareness and economic feasibility. As precision agriculture becomes advanced, the load on central cloud server increases and the amount of data transmitted increases rapidly, causing network congestion. Therefore, a lot of approaches have been proposed to overcome such problems, so attempts to work with big data analysis and artificial intelligence are increasing. Edge AI happens when AI techniques are embedded in Internet of Things endpoints, gateways, and other devices at the point of use. Edge computing brings the data and computes the closest to the point of interaction. Edge AI is a very real (and rapidly expanding) phenomenon, powering everything from smart devices to automotive sensors and security cameras. As IoT implementations have matured, there has been an increased interest in applying AI at the point of generation for real-time event processing.

Efforts to incorporate edge computing into smart farm systems are just beginning. Especially, research to increase efficiency and scalability by dividing the smart farm control into modular stages has not been studied yet. In this paper, an architecture of correlation control stage for smart farm system is proposed, which is based on two-stage control system for efficient and intelligent management using edge computing with lightweight big data and AI technology. We briefly introduce the background of our work, and describe relationship between smart farm parameters, and present an edge computing-based correlation control architecture based on the two-stage control level system, and finally draw our conclusion and future work.

## **2. Background**

Many studies on smart farm-related monitoring and control systems are ongoing. The basic remote-control system can manage the state of each device by extracting environmental status such as temperature, humidity, sunlight, carbon dioxide and wind speed in the greenhouse. However, studies on integrated intelligent control systems are increasing. An efficient system can control the greenhouse environment using fuzzy logic to determine on/off command of air conditioners and heaters (Guerbaoui et al., 2013), and a convergence service for field-oriented Internet of

Things-based smart farm system can conduct intelligent whole-cycle process of agricultural production using complex algorithms (Yeo et al., 2016). There also exist many studies such as implementation of a system that monitors environmental conditions using IoT technology and transfers it to servers and users (Mahendran et al., 2017), development of cloud-based farm environment data collection, analysis, and prediction system (Khattab et al., 2016), heterogeneous sensors and communication systems which operate as a single system with incorporated knowledge-based technology (Minh et al., 2017), and structure of three-layer smart farm with data collection, network and cloud back-end (Ibrahim et al., 2019).

Recently, for standardization of smart farm interfaces, a technology that provides recommendations to control actuators in greenhouses has been studied by collecting and analyzing sensor data using a scalable structure that can help system settings or heterogeneous interfaces (Lee et al., 2018; Yang et al., 2018). The two-stage architecture for greenhouse was proposed to provide an automatic system using conditional analysis control stage, and advanced architecture was also proposed by securing connectivity among greenhouses to apply in scalable smart farm system which covers large-scale farmland (Yang and Park, 2019). An advanced simple structure of intelligent control analysis for smart farm system is proposed, which is based on two-stage control system for efficient and intelligent management using correlation analysis with edge computing (Park, 2021).

Research was presented from the perspective of data management for the development of smart farms into agriculture 5.0 (Saiz-Rubio and Rovira-Más, 2020). Edge models can be used for the development of global food security and smart agriculture with systemic challenges affecting to present and future agriculture from field survey (O'Grady et al., 2019). A scalar data analysis framework was proposed which includes transmitting this to the server through edge pre-processing and the server provides harvest prediction results for crops (Park et al., 2017), and a technique to calculate the optimal layer of each fog node was executed by applying the optimization algorithm of the deep learning model using the node based on fog computing smart agriculture (Lee et al., 2020). A system architecture with network can support a wide area using edge computing and Low-Power Wide Area Network (LPWAN) for smart agriculture and farming activities using edge AI (Gia et al., 2019), and study on the overview of smart agriculture development and how artificial intelligence neural networks can be applied to greenhouse systems by integrating technologies such as IoT and machine learning (Axel et al., 2020) was proposed.

The edge approach is cost effective because its bandwidth is more efficient and transfers less data to the cloud. In addition, it improves communication latency, the responsiveness, and the scalability (Shi et al., 2016). An architecture was presented, which is based on the distributed edge or cloud paradigm and supports a balance between the cost of processing data and benefits at the edge versus at a central (Calo

et al., 2017). Edge computing technology will be gradually used in various social and industrial fields that require real-time and standalone control application, such as agricultural machinery control, factory management, mobile communication system, and disaster protection, and therefore, it will gradually move toward an open platform based on commercialization.

To resolve the challenges to the differential in Quality of Service (QoS) between urban and rural areas, vertical farming is one of the solutions, and has a lot of promise and predicts as the farm of the future. However, there are some stumbling blocks to consider into vertical farming (Rameshkumar et al., 2020; Royston and Pavithra, 2018). It provides a plan to handle future plant demands and allows crops to grow all year-round while requires less water without weather influence. More organic crops can be cultivated with less exposure to disease and chemical threats. However, it requires more cost to build and economic feasibility analysis have not yet been completed. Pollination and fertilization would be very difficult and costly, so would involve higher labor costs while it relies too much on technology. Moreover, one day of power loss would be devastating the farm. This vertical farm technologies are still relatively new. Companies are, therefore, yet to successfully produce crops or plants on a large scale and make it economically meet the growing food demand. The performance of farms will determine how important a role vertical farming will play in the future to face the challenge of food demand. It is worth noting, however, that technologies developed for vertical farms are also being adopted by other segments of the indoor farming sector, such as greenhouses, which can utilize natural sunlight, albeit requiring much more real estate and longer routes to market.

Artificial intelligence technology is being used to make agriculture more intelligent. Some AI paradigms are inherently distributed while using local computing power, and consist of distributed systems that communicate with each other using Internet, based on the behavior by some pre-programmed model. For scalability, the topology could be a tree structure with the edge nodes as the leaf and the centralized cloud server as the root, and this architecture can be sufficient to address a large range of practical smart farm IoT applications (Calo et al., 2017). Furthermore, consisting of only edge nodes and cloud server, it is adequate for most smart farm use-cases. Federated learning (Xu et al., 2013) is a method for training machine learning model on multiple clients without having access to the data itself. The models are trained locally on the devices and only the model updates are sent back to the cloud server, which then aggregates the updates and sends the updated basic and common model back to the client devices.

For traditional smart farm's reservation scheduling technique, an advanced reservation scheduling structure for base-level smart farm automation having environmental condition dependency was proposed (Yang et al., 2019), and a monitoring and control architecture using wireless network toward IoT applications

were also studied (Park and Kim, 2019). A method framework for emergency Internet of Things based on the resource slicing strategy of deep reinforcement learning (Mononteliza, 2020) can be used for multiple network slices application.

New technologies are developing rapidly, and studies on smart agriculture so far have made much progress in applying such technologies to data processing with artificial intelligence at the system level for more intelligent agriculture. However, research on the specific analysis, training, monitoring, and predictive control model for smart farm with edge computing environment may be insufficient yet. Therefore, for the structured and standardized smart agriculture system in the future, it needs to develop the integrated and locally customized smart farming support system for a wide range of regions, as well as smart sensor devices, data collection and analysis, and intelligent control of large-scale smart farms.

### **3. System Structure in Correlation Control Stage**

The environmental control system for smart agriculture requires a technology to improve efficiency and economics with sensors, control devices, network equipment and cloud servers. These control service should have a basic function to provide an automatic actuator control based on sensor data for an intelligent driving system. It has many service utilities such as ventilation, sun shading, cooling and heating, irrigation, and nutrient solution control, which can control actuators in order to provide optimal environment using sensor values or mode settings.

#### **3.1. Environment Parameter Item**

Most smart farm can be operated in a conditional control method having automated and remote-controlled devices based on the analysis of data from sensors for optimal growth environment. For intelligent and large-scale smart farm systems, however, it is necessary to evolve that the advanced management and control methods will be required using correlation analysis of many environmental parameters related to each other. Correlation analysis control stage has parameter items associated with environment of greenhouse for intelligent control and management, as shown in Table 1.

The conditional analysis control stage can be processed on complex controller or IoT gateway and is used to provide calculated results from measured values with pre-defined conditions in order to control each actuator with a combination of detailed control variables. The upper correlation analysis control stage has an intelligent structure that provides correlated result of advanced processing from conditional analysis control stage calculation and additional control parameters. This control stage includes function of big data processing that collects, stores, and manipulates a large amount of sensor data, and pre-processing step to cloud server. It also supports additional training, prediction, reasoning, and decision with the results from cloud to control the actuators with logical calculation and priority

reference.

For smart farm system in a specific area, it is efficient to process the automation process locally appropriate to the weather and environment in that area. In other words, it is necessary to extract optimal value by making a prediction model with the environment of a smart farm and applying appropriate algorithm. In edge computer or edge node in correlation analysis control stage, data analysis from various external and internal input parameters can be executed, and the predicted inference results are calculated and applied. It can be done by synthesizing the data inside and statistical data so far related to a specific crop or local environment. Some parameters for data analysis and prediction from sensors and conditions of a greenhouse can be described as a model with correlations as shown in Fig. 1.

Table 1: Parameter items of control stage in smart farm

Control level	Parameter items	Contents
Conditional analysis control stage	Light environment	Light intensity, Light quality (Spectrometry)
	Ventilation	Rate of air circulation, Number of air change
	Air conditioning	Min., Optimum temperature for crop, Max. temperature
	CO2	Optimum photosynthesis, Gasification of solid/liquid
	Watering, Nutrient	Soil water, Nutriculture (Nitrogen, Phosphoric acid, Kalium)
	Blight monitoring	Temperature, Humidity, Moisture
Correlation analysis control stage	Light environment	Light duration (Photoperiod) By region, Latitude/Longitude, By crop
	Air conditioning	Cooling technique, Circulation type Inside/Outside, Energy saving strategy
	Nutrient supplement	Ratio adjustment
	Pests	By seasonal/soil
	Management by crop	Optimal condition extraction
	Production	Planning, Regional, Economics
	Quality management	Appearance, Texture, Flavor, Harvest time



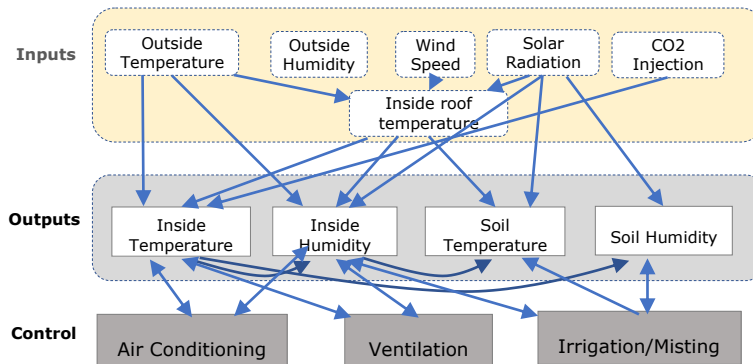


Fig. 1: Data analysis and prediction parameters in greenhouse environment

In our two-stage control processing architecture, the output of a conditional control stage is a calculated and processed result for one actuator driver. For example, to drive a ventilation fan, it is necessary to consider measured values such as temperature, humidity, carbon dioxide, light and other data from scheduling table. The correlation control stage processing, on the other hand, is to determine the results using conditional control stage with respect to the other intelligent data. The operation of the ventilation fan at this stage, for example, can be controlled by using more information such as training data before, the current state of side windows for the growing crops, the condition and soil quality with water pumps and heaters, regional position of greenhouse, seasonal sunlight, harvest time, etc. The study with only conditional control stage function uses the basic automatic function to control the smart farm system (Lee et al., 2018; Yang et al., 2018; Yang and Park, 2019). In here, we propose an automatic control structure at the correlation control level, which is useful when training and prediction with intelligent and statistical data are needed at the high-level using edge computing or lightweight big data and/or artificial intelligence.

Parameters that can be processed or controlled in the conditional and correlated control stage can be divided as shown in Fig. 2. In this figure, we can see that it is possible at the conditional control stage to control the farm environment by processing basic sensor inputs, but it is better to process by combining more intelligent technology such as bigdata and deep learning for more complex and efficient farm control.

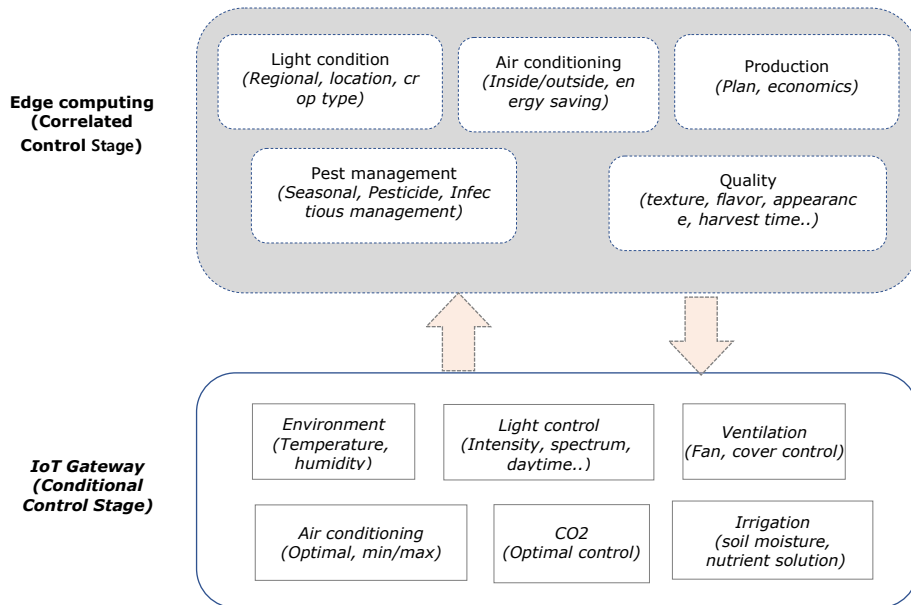


Fig. 2: Parameter items related to each control stage

In the proposed smart farm structure, the parameters related to each stage are processed in the most appropriate form according to the growth process to generate results. That is, it is possible to process various environmental parameters and make different results according to processes such as sowing, flowering, fruiting, harvesting and sales, and even in the period of preparation for the next cultivation.

The results from various input data should be appropriately applied according to the growth cycle of crops, which should be processed by considering the previous results as part of the input. Smart farms using these models can be applied to deep learning models. That is, by applying a Recurrent Neural Network (RNN), it is possible to determine the optimal next state according to the growth of the crop from the previous state in consideration of the correlation of various input parameters as shown in Fig. 3. In addition, due to the gradient loss problem, a technique to control the management of the value passed to the next layer using LSTM (Long Term Short Memory) should be utilized.

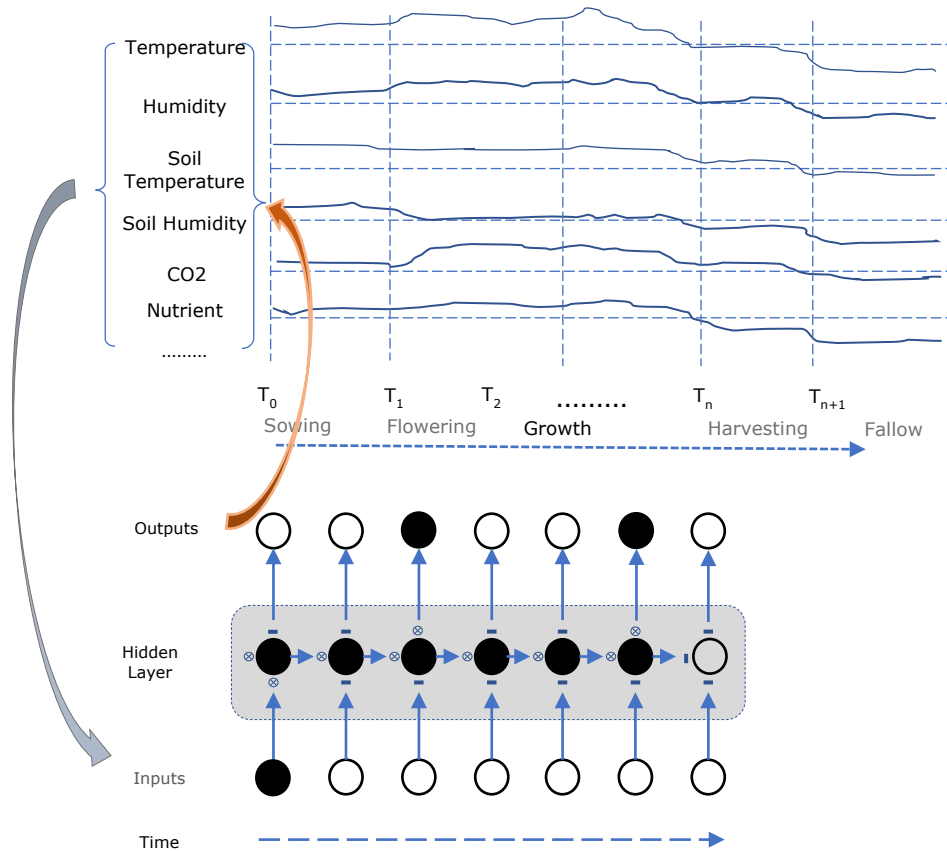


Fig. 3: Control of environment parameters over time and RNN model

As shown in Figure 3, the environmental control task should be processed according to the input parameters depending on the time of the plant growth process with some part of previous data. That is, the most suitable state should be calculated for the treatment result by the plant growth state such as sowing, flowering, and fruiting season. Looking at the entire cycle of the smart farm, it is appropriate to use RNN model to make intelligent control of smart farms. RNN and LSTM models have the feature of memory and time dimensions. They can be used to estimate the status of plant based on previously recorded data, such as assessing fruit yield or water needs. The Convolutional Neural Network(CNN), on the other hand, applications of deep learning in agriculture are more suitable for diseases detection, detection weeds, classification crop types, prediction yield, counting fruits, water stress estimation, and others. Both models can also be applied to the environmental factors, such as predicting climate change and phenomena.

The edge node or computer stores data from sensors, and analyzes it and filters valid data to transmit it to the cloud. Many IoT devices continuously generate multimedia data, but in order to become useful information, the data must be

analyzed and processed. To perform such functions of correlation control level, it is effective to process using edge computing functions because existing integrated controller or IoT gateway have limited computing performance and storage. The development of platforms and ecosystem construction using high-performance computing cores such as NPU(Neural Processing Unit), TPU(Tensor Processing Unit) and SoC(System on Chip) to support edge computing is also accelerating, so it will be expanded to various services fields as well as smart farms systems. The rise of cheap computing and data storage resources with cloud infrastructure has given new opportunities to leverage machine learning at scale. However, this comes at the cost of latency and data transfer challenges due to bandwidth limitations. Training a machine learning model is a computationally expensive task well suited for cloud-based environment, whereas inference requires relatively low computing resources.

In the traditional setting, the inference is executed in a cloud computing platform. With edge AI, the model works in the edge device without requiring connection from the outside world at all times. The process of training a model on a consolidated dataset and then deploying it to production is still similar to cloud computing though. This approach can be problematic for multiple reasons. First, it requires building a dataset by transferring the data from the devices to a cloud database. This is problematic due to bandwidth limitations. Second, data from one device cannot be used to predict outcomes from other devices reliably. Finally, collecting and storing a centralized dataset is tricky from a privacy perspective. Legislative limitations such as GDPR(General Data Protection Regulation) are creating significant barriers to training machine learning models. Moreover, the centralized database is a lucrative target for attackers. Therefore, the popular statement that edge computing alone answers to privacy concerns is false.

### **3.2. Correlation Analysis Control Architecture**

Cultivation of crops using smart farm system should provide an optimal environment for crop growth. In addition, crop type selection, optimal environment according to geographical location, consideration of production plans and economics to resolve excesses and shortages with the demands of consumers may be considered. This should be considered as a social industry and intelligent concept beyond the basic automatic control system. We present a structure of correlation control stage, where storing collected data from devices, making automation strategies, and intelligent management of local farms can be performed in edge computer as shown in Fig. 4, while other high-level functions such as filtered data storage, integrated analysis and management of pre-processed data are handled in the cloud server.

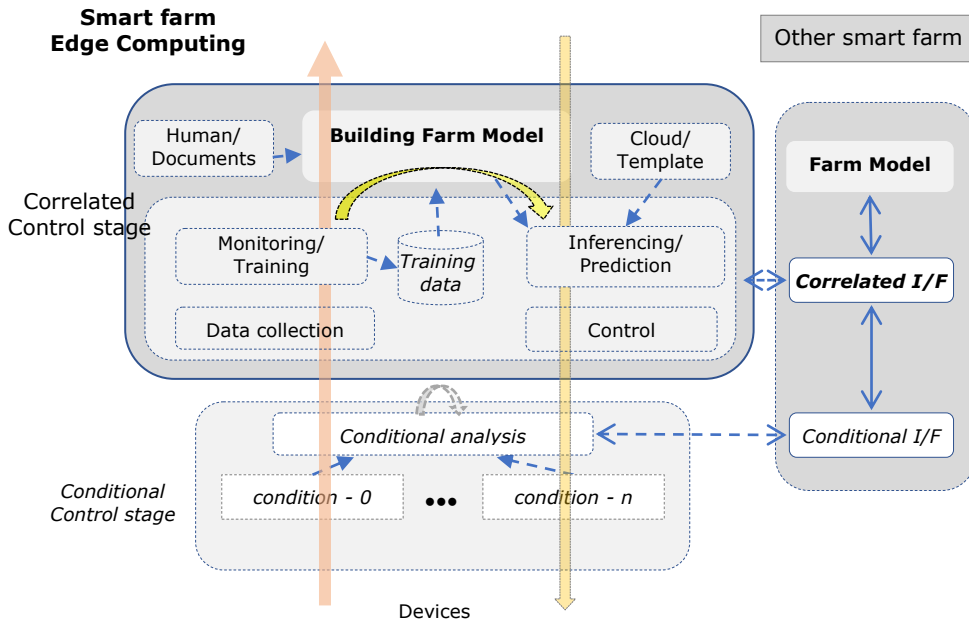


Fig. 4: Flow diagram of 2-stage control architecture for smart farm system.

From the proposed two-stage control architecture, the correlation control stage is connected to Internet by IoT gateway or complex control device within the farmland, which is associated with the edge node, so it can be extended to interconnect large-scale farmland operation. It processes the collected data from the lower conditional control stage and sends the extracted features to the upper stage, which is the cloud. It can support functions in cooperation with the cloud, such as adaptive light control by geographical and seasonal characteristic or crop type, and adaptive indoor environment control according to the external conditions with energy saving function. There exist additional high-level functions through the interface with the cloud, such as cooperating prevention for pests and diseases, production management by specific crop sowing area with harvest time and consumption prediction, and adjusting the harvest time for optimizing the texture, flavor, and appearance of crops or fruits by consumers' preferences.

In correlation processing module, modeling is performed by analyzing the input stored in the previous training process with the internal data collected actually and then run the algorithms, if necessary, with restricted conditions such as weight or bias according to the rules for crop, time, and farm condition. The result will be stored in internal database and can be used as input for intelligent control later. In a typical machine learning setting, we start by training a model for a specific task on a suitable dataset from sensors. Training the model basically means that it is programmed to find patterns in the training dataset in edge devices and then evaluated on a test dataset to validate its performance on other unseen datasets,

which should have similar properties to the ones that the model is trained on. The data selected and processed in the correlation control stage of the edge computer is transmitted to the cloud server, and the server executes functions such as additional processing, deep learning, and prediction to return the final control result. Since the cloud server collects preprocessed data from the edge computer, it processes smaller capacity data than the original of the edge computer, thereby reducing the congestion of the entire network and the processing burden on the server.

The edge node has such structure that expands the functions of data storage and intelligent computing by extending the functions that the IoT gateway controls the existing smart farm. Before sending all data to the cloud, the data should be analyzed and filtered to transmit the necessary data and related information to the cloud. All collected data is also stored on the edge computer and performs the automatic control function for the smart farm. Fig. 5 shows a correlation control structure within the two-stage architecture using edge computing, which includes various control parameters, step-by-step functions, data processing, and lightweight edge training and prediction functions for intelligent smart farms.

In the correlation prediction/control stage, it executes control of the smart farm while going down to the condition control stage based on the input of previously stored parameters and control values from the cloud server. Using the parameters analyzed in this step, the results containing environmental characteristics are predicted according to the control type through normalization, modeling, and prediction algorithms, and will be stored in the edge database. In addition, useful information and control values from the cloud server can drive many devices of the smart farm and informs the administrator or farmer of the status. Since edge nodes have low performance compared to cloud servers, lightweight technologies such as parameter removal, sharing or model compression, knowledge distillation and acceleration should be applied to edge computer. These functions can be operated in such way that using meta information of the model with input/output values, execution methods by registering a growth model that includes an environment control technique optimized for a specific crop. They also can be operated by processing it from input values in advance using data management services, and by passing the input value when the model is executed.

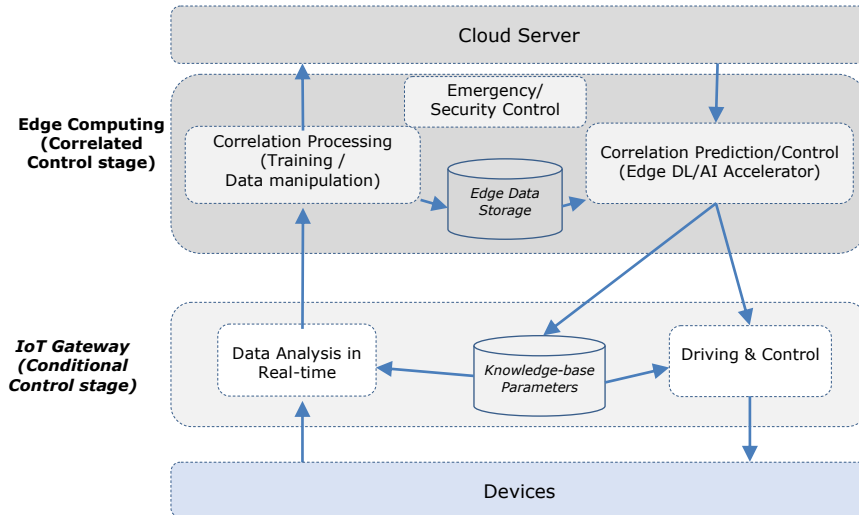


Fig. 5: Functional structure of conditional and correlation control stage in edge node and IoT gateway.

### 3.3. Intelligent Interface Model using Data Processing and Inferencing

In order to make scalable smart farm structure, it is necessary to include functions to connect and communicate with other smart farms distributed in various regions which include the structure to manage them in an integrated manner. In the case of different types of crops, the optimum environment and requirements for crops in each smart farm system may be different, so training data, learning, and inference models may also be different. Of course, the basic conditions for growing crops should be stored in the cloud server and, if necessary, transmitted to individual edge computers for use. In this case, it is necessary to support the connection and control management structure for large-scale farmland system, so it is appropriate to use the architecture with central cloud server and edge computing as shown in Fig. 6.

There can be untrained shared smart farm model on the cloud server, which will be allocated edge nodes for training. Edge nodes will train the model with collected and stored local data. After local learning in correlation control stage, changes of the model, i.e.,  $\Phi_x$  will be summarized as a small update, which will be transmitted to the cloud through the Internet. The server averages received changes from all edge nodes and updates the shared common model with the averaged result. Then, edge nodes will download the update for their local model, if needed, and repeat the processing step to improve the shared model for the next round of learning. Transfer learning and edge computing can be combined to learn a smarter model for smart farm systems.

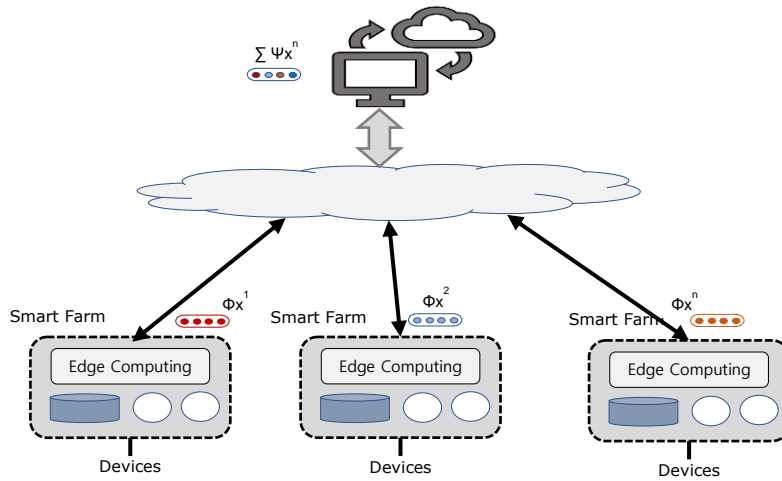


Fig. 6: Participants training data aggregation and shared model to edge nodes

This model allows for hyper-personalization - while preserving privacy. Central server or cloud chooses a statistical model for training. Initial model will be transmitted to several smart farm nodes or edge nodes. Each node trains its local model using asymmetric local data, and finally central server creates one shared model using parameters from nodes' models. Edge computing is not going to completely replace cloud computing, rather it is going to work in conjunction with it.

Fig. 7 shows a diagram structure for pre-processing data and inferencing in edge nodes. Data pre-processing can be operated based on the collected data from devices in real time, then extracts the necessary parts by using the stored local data, retrieved data from a remote location if necessary, and from human resources. This data can be transformed into data for smart farm, which includes vectorization, shuffling and labeling.

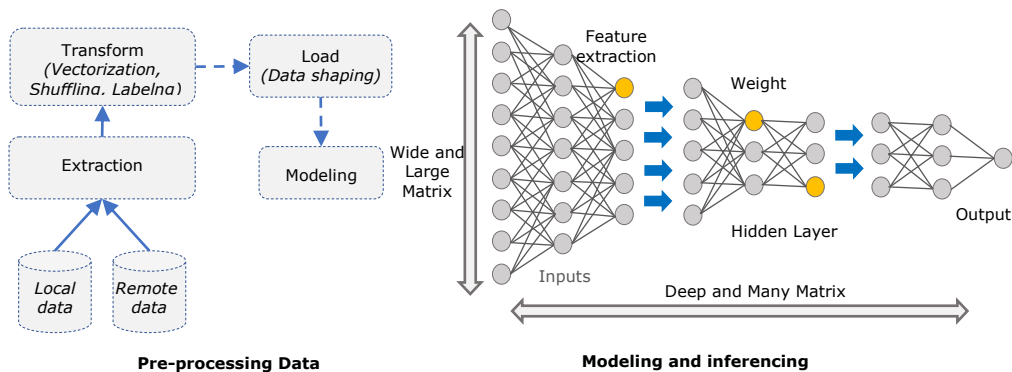


Fig. 7: Pre-processing data and inferencing structure in edge nodes



For smart farm operation, the stored data is imported, and modeling is performed based on it. The optimized model for each smart farm system can be made by integrating additional model suitable for the plants and processes growing on the farm with the base model transmitted from the cloud and make inference job based on it. In the modeling process, a set of AI models could be built using machine learning algorithms with training data, by a set of documents, or by encoding from human expertise. Models can be demonstrated in different ways, such as decision trees, inference rules, or neural networks. Once modeling process is completed, the models can be used to make inferences from the input data from the upper cloud server and control the operation of the smart farm system.

Prototyping edge deployment can be constructed associated with one or more AI models. This can be obtained with the basic elements for performing data collection, feature extraction, aggregation, and reporting as shown in Fig. 8. The processing task can also be increased or augmented with user defined filters. There can also be system defined functions that are required for executing the model on the edge nodes, monitoring the specific algorithms, and setting up triggers to make report machine learning model scores or classification.

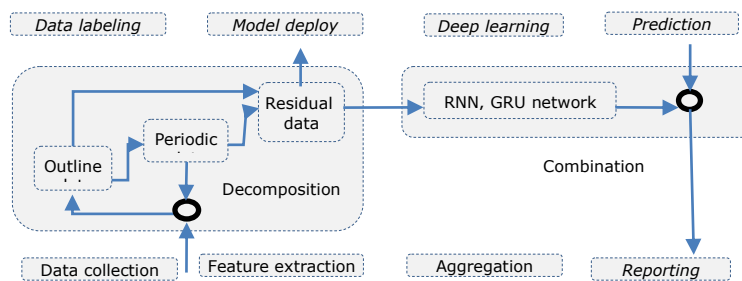


Fig. 8: The structure of decomposition, modeling, and combination flow work

The collected data could be decomposed into components, for example, then, each component is treated separately to obtain different GRU(Gated Recurrent Units) sub-predictors in the network training stage. In the prediction stage, different GRUs are used to predict the different components, respectively. Lastly, all the predictions will be combined to get the final predicted results in the output node. The first-level decomposition is some period of outline data, period per day, and residual were obtained. Because the residual obtained by the first-level decomposition still had periodicity, it then makes the next decomposition to decompose the residual into further components.

It can be used for inference in a specific context, for example as a microservice. Inference refers to the process of making predictions using trained machine learning algorithms. Once the model works as wanted, predictions produced by the model can be utilized in improving farm management processes. The model output is then either communicated to another correlation software component in edge node, or in

some cases, visualized on the application front-end for the end user.

Edge AI speeds up decision-making, makes data processing more secure, improves user experience with hyper-personalization, and lowers costs by speeding up processes and making devices more energy efficient. When most of the data processing is done locally, on the edge, centralized service or data transfer will not become a bottleneck. Edge AI use cases typically involve large amounts of data. It is necessary to prepare the information immediately before training the neural network models. The present real-world data obtained from various smart farm sensors are distributed non-uniformly, and henceforth, the data cannot be utilized between training and testing of the neural network models. In this way, the input characteristics are normalized, and classified parameters are converted to numeric data through labeling encoder for efficient processing. The regulated dataset is then divided into training and independent test sets at a given fixed rate. The results obtained are discovered in terms of multiclass classification, as well as individual class basis. The process of training and testing will be repeated for some number of iterations until the convergence optimization is met, and for the number of neurons and hidden layers for each model.

The smart farm technology has been developed mainly on the automation of the crop growth environment until now, but it is becoming more intelligent such as prevention of pests, fertilizer supply, sowing, and harvesting, and even prediction and management in advance are being developed. We are now at the point of innovative smart farm system, which includes, through the application of big data and edge AI, optimal farm and environment management that refers to the past items, control of shipment and distribution by prediction of crop type and production volume, and customized agricultural product according to consumer preferences. Therefore, this rapid development of smart farm technology (Saiz-Rubio and Rovira-Más, 2020) has made precision farming possible due to the use of big data and AI, and the need for scalable architecture and standardization framework of smart farm has increased.

In particular, since the outdoor smart farm has more sophisticated parameters to consider, unlike indoor greenhouse control, this requires more function at a higher level different from existing greenhouse control strategy. In other words, it needs to control using statistics, big data and AI technologies which encompass management of the sunlight environment according to geographic location or season, irregular changes in weather, response to various and sudden natural disasters, control of irrigation and nutrient solution according to the soil status, adaptive soil management by short-term precipitation, and pest control. However, indoor smart farms also gradually increase in scale and require intelligent analysis and control rather than basic automated control, so the advanced architecture that can be processed with the help of intelligence technologies will be needed. The step-by-step correlation functions in edge computing-based architecture described in this

paper can be used for intelligent smart farm to control devices that have not yet been standardized. And efficiency can be increased by interworking with edge computer to reduce the burden of network traffic with the cloud server.

#### **4. Conclusion**

In this paper, for efficient control and intelligent management of smart farms, we propose a two-stage scalable smart farm architecture including correlation control stage, which is one of the two-stage control structure with higher level of basic automation step with edge computing. This architecture can reduce the burden on the cloud server and enables local control in case of cloud or communication failure. It supports advanced functions such as operation of large-scale smart farms through interfacing with network and cloud servers, intelligence increasing through intermediate data processing and training before delivery to the server using the lower condition control results. Cooperated management of large-scale farmland can be possible using processed averaged data from upper cloud server. With its own lightweight big data and edge AI model, it can be referenced as an architecture framework for intelligent smart farm standardization in the future, with scalable structure that can be connected and extended to other correlation control stages of edge computers. This architecture can provide more control functions through, the association of detailed variables and light-weight big data processing, training and inferencing, so our future work will include implementation of our structure with edge computing by applying it to the advanced smart farm systems.

#### **5. Acknowledgements**

This paper was supported by Joongbu University Research Program in 2020.

#### **References**

Ai Y., Peng M., & Zhang K. (2018). Edge cloud computing technologies for Internet of Things: A primer. *Digital Communications and Networks*, 4(2), 77-86.

Axel E., Genaro M. S., Manuel T., Edgar R., & Abraham G. (2020). Applications of Artificial Neural Networks in Greenhouse Technology and Overview for Smart Agriculture Development. *Applied Sciences*, 10(11), 3835.

Calo S.B., Touna M., Verma D. C., & Cullen A. (2017). Edge Computing Architecture for applying AI to IoT. *IEEE International Conference on Big Data (BIGDATA)*, USA, 3012-3016.

Gia T. N., L. Q., Queralta J. P., Zou Z., Tenhunen H. & Westerlund T. (2019). Edge AI in Smart Farming IoT: CNNs at the Edge and Fog Computing with LoRa. *2019 IEEE AFRICON*, Accra, Ghana, 25(27), 1-6.

Guerbaoui, M., Ed-dahhak A., EIAfou Y., Lachhab A., Belkoura L., & Bouchikhi B. (2013). Implementation of direct fuzzy controller in greenhouse based on labview. *International journal of electrical and electronics engineering studies*, 1(1), 1-13.

Ibrahim, H., Mostafa N., Halawa H., Elsalamouny M., Daoud R., Amer H., Adel Y., Shaarawi A., Khattab A., & Elsayed H. (2019). A Layered IoT Architecture for Greenhouse Monitoring and Remote Control. *SN - Applied Sciences*, 1(3), 223.

Inés S., & Juan M. C. (2019). An Edge Computing Tutorial. *Oriental Journal of Computer Science and Technology*, 12(2), 34-38.

Khattab A., Abdelgawad A., & Yelmarthi K. (2016). Design and implementation of a cloud-based IoT scheme for precision agriculture. *Proceedings of 28th International Conference on Microelectronics (ICM)*, Giza, Egypt, 201-204.

Lee K., Silva B. N. & Han K. (2020). Deep Learning Entrusted to Fog Nodes (DLEFN) Based Smart Agriculture. *Applied Sciences*, 10(4), 1544.

Lee S., Cho B., & Park H. (2018). Design of Scalable Sensor and Actuator Interface Module for smartfarm. *International Journal of Smart Home*, 12(4), 1-6.

Mahendran M., Sivakannu G., Balaji Sriraman(2017). Implementation of Smart Farm Monitoring using IoT. *International Journal of Current Engineering and Scientific Research (IJCESR)*, 4(6), 21-27.

Minh Q. T., Phan T. N., & Takahashi A. (2017). A Cost-effective Smart Farming System with Knowledge Base. *Proceedings of the Eighth International Symposium on Information and Communication Technology (SoICT 2017)*, NhaTrang, Vietnam, 309-316.

Mononteliza J. (2020). Research on EIoT Reservation Algorithm Based on Deep Learning. *Asia-pacific Journal of Convergent Research Interchange*, 6(9), 191-205.

O'Grady M.J., Langton D., & O'Hare G.M.P. (2019). Edge computing: A tractable model for smart agriculture? *Artificial Intelligence in Agriculture*, 3, 42-51.

Park H. D. (2021). A Scalable Structure of Correlation Control with Edge Computing for Large Scale Smart Farm. *International Journal of Hybrid Information Technology*, 14(2), 2021.

- Park H., Kim S. (2019). A Study on Monitoring and Control Architecture for Smart Lighting System in IoT Environment. *Asia-pacific Journal of Convergent Research Interchange, SoCoRI*, 5(3), 91-100.
- Park J., Choi J., Lee Y., & Min O. (2017). A Layered Features Analysis in Smart Farm Environments. *Proceedings of the International Conference on Big Data and Internet of Thing*, London, United Kingdom, 20(22), 169-173.
- Rameshkumar D., Jagathjothi N., Easwari S., Rajesh R., Muthuselvi R., Kumar P. N., Minithra R., Suresh R., & Krishnakumare B. (2020). Vertical Farming - Agriculture of the Future. *Indian Farmer*, 7(11), 1013-1017.
- Royston R. M., & Pavithra M. P. (2018). Vertical farming: A Concept. *International Journal of Engineering and Techniques*, 4(3), 500-506.
- Saiz-Rubio, V., & Rovira-Más, F. (2020). From Smart Farming towards Agriculture 5.0: A Review on Crop Data Management. *Agronomy*, 10(2), 207.
- Shi,W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, 3, 637-646.
- Xu, D., Li T., Li Y., Su X., Tarkoma S., Jiang T., Crowcroft J., & Pan H. (2020). Edge Intelligence: Architectures, Challenges, and Applications. *arXiv: Networking and Internet Architecture*.
- Yang A., & Park H. (2019). An Advanced Automatic Control Architecture with Conditional Results in Scalable Smartfarm System. *The Asian International Journal of Life Sciences, Supplement*, 20(2), 275-285.
- Yang A., Kim J., Cho B., & Park H. (2018). An architecture and Design of Data Converter for IoT-Based smartfarm. *International Journal of Smart Home*, 12(4), 7-12.
- Yang A., Kim S., Park H. (2019). A Study of Advanced Reservation Scheduling Structure for Smartfarm Actuator Control. *Asia-pacific Journal of Convergent Research Interchange, SoCoRI*, 5(4), 19-27.
- Yeo U., Lee I., Kwon K., Ha T., Park S., Kim R., and Lee S. (2016). Analysis of Research Trend and Core Technologies Based on ICT to Materialize Smart-farm. *Protected Horticulture and Plant Factory*, 25(1), 30-41.