Advanced Science and Technology Letters
Vol.145 (NGCIT 2017), pp.112-115
http://dx.doi.org/10.14257/astl.2017.145.22

A Study on Mushroom Growth Environment Analysis System based on Machine Learning for Efficient Operation of Mushroom Plantation

Kyoung-Jong Kim[1], Se-Hoon Jung[2] and Chun-Bo Sim[3*]

[1, 2] Dept. of Multimedia Engineering, Sunchon National University, Suncheon, Korea
[3*] School of Information Communication and Multimedia Engineering, Sunchon National University, Suncheon, Korea
E-mail : kkj0201@nate.com, iam1710@hanmail.net, cbsim@sunchon.ac.kr

Abstract. There are many technology researches being conducted based on integration with the artificial intelligence technologies in the 6th Industry, a recent topic of ICT convergence technology in the agriculture and life industry. Of the artificial intelligence technologies, machine learning, in particular, requires Big Data analysis techniques in the agriculture and life industry with various types of data and uses many different methodologies. This study set out to propose a Big Data-based integrated system to manage and analyze a mushroom growth environment for the efficient management of mushroom plantation.

Keywords: Mushroom, Analysis, Machine Learning, Plantation.

1 Introduction

The recent increase of Internet information network utilization and the continuous accumulation of data have raised a need for massive data processing techniques to deal with increasing volumes of data[1]. Big Data represents a technology to analyze useful information and outcomes out of massive data, and there are active research efforts on Big Data-based system and data management methods home and abroad. In South Korea, the fusion of ICT and agriculture and life industry focuses on how to solve all sorts of issues in the agriculture and life industry under the topic of 6th Industry. The data analysis technologies are oriented toward the artificial intelligence technologies and fragmented based machine learning. In agriculture, in particular, there are time series data influenced by the changing seasons and time. Those data are used in statistics-based regression analysis or machine learning-based analysis, and many researches are carries out to design an algorithm for them. Various types of data management/analysis open source tools are used to realize an analysis environment, and Apache Hadoop, one of their good examples, is utilized in systems across many fields. Hadoop saves and processes Big Data and builds a Hadoop ecosystem to provide scalable functions. A Hadoop ecosystem accommodates many different forms of data storage and various types of data and is capable of expanding a framework of

ISSN: 2287-1233 ASTL

data manipulation and analysis, thus offering a proper data management feature for the agriculture and life industry consisting of structured and unstructured data. The present study aimed to build a machine learning-based integrated system to analyze a mushroom growth environment, analyze the growth environments and disease and pest images of mushrooms, and develop a system that would provide integrated analysis information based on analysis results. For those purposes, the investigator built a Hadoop ecosystem capable of processing and analyzing the Big Data of environmental information and treating images and videos. Added to the ecosystem were Sqoop to link the Hadoop HDFS to a relation database for data processing and Spark Streaming to process data in real time. The proposed integrated system would establish a Spark MLlib library to analyze collected data based on machine learning and provide a service to promote the efficient operation of mushroom plantation.

2 Proposed Study

2.1 Mushroom Growth Environment Analysis

Analyzed in the study was a vinyl greenhouse with multiple layers of supplements built to cultivate mushroom. The vinyl greenhouse contained three- or four-story shelves where mushrooms were cultivated in the medium. A medium was made in a spawn vaccination and culture process or one that had been treated with the process was purchased. A medium would grow mushrooms for 40~60 days since it was introduced into the cultivation facility. During the period, a cycle of mushroom harvesting and medium water flowing would be repeated three to four times with a two-week interval. Fig. 1 shows the data affecting the growth factors of mushrooms including the mushroom species, medium state(importer), type of house(vinyl or panel), and time series data of temperature, humidity, illumination intensity, and Co2 in and outside.

Fig. 1. The Algorithm presents the altered K-means algorithm proposed in the study.

Since different mushroom species require different growth environment(medium and hardwood), the environmental data collected from plantation should be analyzed according to the mushroom species. Having a huge influence on a plantation, the great temperature changes between summer and winter also make an influential factor of growth environment. The outdoor environment is characterized by high temperature and humidity in summer and low temperature and humidity in winter, which causes differences to the environmental information of a cultivation facility. It is thus needed

to adjust the expected yield accordingly by taking the differences into consideration in the design of an analysis algorithm.

2.2 Mushroom Disease and Insect Pest Image Analysis Method

Mushrooms have irregular shapes and colors, which means there are limitations with analyzing mushrooms in feature points and patterns based on the old video processing method. Fig. 2 shows a convolution neural network(CNN) to design and apply an algorithm for the analysis of mushroom diseases and pests. Convolution calculations were obtained by getting a feature map according to the mushroom characteristics and applying a filter to the map according to the weight. Whenever an image to be analyzed was moved to the right shift with a mask of designated size, a convolution calculation would be done to reduce the image. The reduced feature maps were categorized according to the characteristics so that the system would be able to identify diseases and pests according to the mushroom species. The images were saved in the server based on the classification and analysis results to be retrieved along with the analysis results of cultivation facility environment.

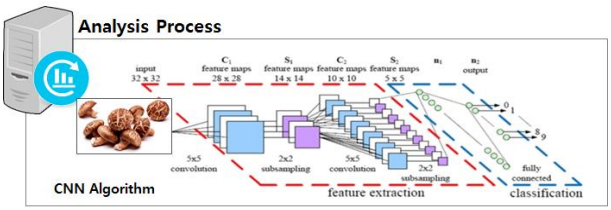

Fig. 2. Object Analysis Algorithm of Mushroom using convolution neural network.

2.3 Big Data Integration Analysis System

Since different mushroom plantation have different environment according to the plantation, the data of growth environments obtained from the facilities should be analyzed according to the concerned environment. The investigator built a Big Data analysis system to process the growth environment data collected from many different plantation. A framework was added according to the functional elements based on the Hadoop ecosystem structure. MySQL DBMS was used to build a Hadoop lined Spark framework to save and process data, a Sqoop application to collect data, and a relation database to process structured data. A multiple regression equation was used based on Spark MLlib(machine learning library) to analyze environmental data. The images of mushroom diseases and pests were provided as analysis data. A CNN algorithm was implemented with an image analyzer for the service of analyzing the images of mushroom diseases and pests. A JSP based on web service was used to receive the service from the environment analysis system of a mushroom cultivation facility. The data would be provided to mushroom plantation via the service UI. The system was developed to help the mushroom growers select the date and cultivation facility of

analysis to analyze the environmental information of their plantation and receive the information about the prevention and extermination of diseases and pests according to the images and conditions of mushroom diseases and pests based on the environmental analysis results. Fig. 3 shows the diagram of a structure to design an integrated Big Data analysis environment.

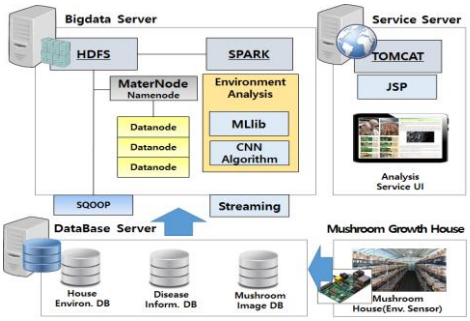

Fig. 3. Structure of Mushroom Integration Analysis System.

3 Conclusion

In this paper proposed a management service system based on machine learning, an artificial intelligence technology, to analyze the environments of mushroom plantation in real time and replace people judgments with artificial intelligence for efficient operation. It was difficult to analyze an environment accurately with the data collected for a short period of time after the system was implemented, which raises a need to collect data for a long period of time. Future study plans to use the *Zepplein* web service of Apache Project along with a rich data collection and add a visualization function for analysis results to provide an easily accessible analysis environment.

Acknowledgments. The research was supported by 'Area Software Convergence Commercialization Program', through the Ministry of Science, ICT and Future Planning(S0417161012).

References

1. K. S. Shim, Y. H. Kim, J. H. Lee, J. H. Kim, Y. J. Park,: Recent Research Trends of Mapreduce Algorithm for Big Data Analysis, Journal of the Korean Institute of Information Scientists and Engineers, vol.32, no.1, pp.27-32(2014)

