Journal of Engineering and Applied Sciences 12 (10): 2560-2564, 2017
ISSN: 1816-949X

A Study on the Prediction Method of Emergency Room (ER) Pollution Level based on Deep Learning using Scattering Sensor

Mi-Lim Choi, Myung Jae Lim, Young-Man Kwon and Dong-Kun Chung
Department of Medical IT Marketing, Eulji University, Gyeonggi-do, Korea

Abstract: Regarding the seriousness of infection within the Emergency Room (ER), due to the various infectious diseases from virus such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) that occurred recently in the last few years, many patients that visited the ER for treatment were infected and resulted in death. To monitor the pollution level in the ER for preventing this occurrence of diseases in advance, the device is studied which is easy to install, counts the air particles by using the back scattering method and has IR sensor. The dangerous pollution level is alerted in advance to notify the hospital personnel, patients and guardians and the air particles within the ER are collected and performed with quick analysis to verify the pathogen and harmful virus before the contamination. To enable this, the decision tree algorithm of deep learning, a hot issue in the present and a part of machine learning is used to study the similar cases and to deliver the suspicious danger signals quickly. CHAID (Chi-squared Automatic Interaction Detection) of the decision tree has continuous target variables and stops the growth of the tree model in appropriate size to be considered of having advantage in saving the time to find the similar cases.

Key words: Emergency room, indoor air, virus, air particle, deep learning, IR sensor

INTRODUCTION

Virus refers to the infectious particles that are parasitic to the living cells like animals, plants and bacteria and various infectious diseases from virus such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) are suspicious of infecting the patients visiting ER for treatment which causes great social controversy. The indoor air contains various harmful bacteria and heavy metal particles in fine dust form to cause harmful influence to the human body and most of the viruses that had been discovered are in the nano-meter size of 1/100 million which is a size that cannot be seen through a general microscope. From cold, enteritis to SARS or MERS, there are many diseases that are infected by virus and the effects vary greatly. However, there are still more viral diseases that do not have vaccines for treatment. To measure the scattering value of the particles within the air in ER by using the IR sensor and process the data on the pollution level by using deep learning, a part of the machine learning of deep learning is to be used for monitoring. The machine learning is a part of Artificial Intelligence (AI): patters or characteristics from the given data are learned to be used for algorithm and in relevant fields to perform the work. The machine learning is applied in all fields based on algorithm from software technology to finance and economy and recently, it is being applied to the medical technology including medical data analysis. In addition, as the machine learning technique of deep learning is being focused as the core technology, the interests on the relevant technologies and applied fields are also increasing. The deep learning is a part of artificial neural networks that imitate the nervous system of organism and while conventional artificial neural network models consist of the connection of thin layers of neuron models, the deep learning is a model that increases the learning ability of the neural network by building up the layers of neuron models deeply. As an artificial neural network composed of many layers, the concept of deep learning was first proposed in 1970 but the study was stagnated due to its complexity in learning calculation and recently, various studies have been presented to improve its performance and show excellent results in the voice and phase recognition fields to quickly increase its demands. The purpose of this study is to use the back scattering sensor for the fine dust measurement sensor to enable efficient measurement even in narrow spaces especial ER. By mounting the dust sensor, he fine dust invisible to the eyes such as the house dust that is considered to be one

Corresponding Author: Myung-Jae Lim, Department of Medical IT Marketing, Eulji University, Gyeonggi-do, Korea

of the causes for asthma or allergy in the indoor air is to be detected to enable real-time monitoring of the pollution level.

Literature review

Deep learning: Deep learning which is a part of machine learning, became focused in the 2012 competition on the image recognition technology of ILSVRC (ImageNet Large Scale Visula Recognition Challenge). In the 2012 competition, the technique based on the deep learning of AlexNet dominated and won the competition to change the conventional approach methods on image recognition. In 2012, AlexNet reduced the error rate greatly and especially in 2015, in-depth neural network over 150 layers of ResNet reduced the error rate to 3.5% to exceed the recognition ability of normal humans (Gorky, 2017).

In machine learning, the neural network is a form that imitates the neuron of organism and calculates and outputs the data value into input layer, hidden layer and output layer as shown in the structure.

Intensive learning method (Deep Q-Network) one of the deep learning methods, refers to the learning process which is human's trial-and-error when learning something and the basic frame of the Deep Q-Network is that learning is enabled to obtain better compensation of the agent. The main data structure of pandas is dataframe which is similar to the form of a table or spreadsheet. The dataframe is created through the records. Fundamentally, IPyton was designed to obtain optimum productivity for both conversational computing and software development and unlike other programming languages, the 'run-search' method is promoted rather than the 'edit-compile-run' method. Also, it is well integrated with the file system, the cell of the Operating System (OS). Due to these characteristics, IPython can quickly process the data search, test, error decoding and repeat, etc., that occupy large part of the data analysis programming. Moreover, the commands used previous are stores separately in the on-disk database to be used for the following purposes (Kinney, 2016):

- Rerun previous commands through search, auto complete and minimum input
- Maintain command history between sessions
- Input, output history file recording

MATERIALS AND METHODS

Fine dust: Materials floating in the atmosphere are called aerosol and the general size is in the 0.001~100 μm range. The material on the particle that is >1.0 or 2.5 μm in equivalent diameter of the aerosol in the atmosphere is called coarse mode particle and it is mostly the particle material that is created naturally. The particles smaller than the coarse mode particle are mostly from human activities and defined as fine mode particle and various harmful bacteria and heavy metal particles exist inside the indoor air in the fine dust form.

Due to the characteristics of the ER, the gaps between the beds are insufficient which causes the transfer of droplet nuclei and increases the risk of aerial infection. When physical separation is not possible, the air current must be controlled to prevent the dispersion of droplet nuclei (Cheong *et al*., 2016).

Backward-scattering measurement: To count the fine dust, the IR sensor and the back scattering among the scattering methods will be used to count the particles. The reason that the back scattering is used is because it is a method that can measure even when the size of the device is small to enable installation in the indoor space. For the particle extraction of the IR sensor, the Mie Theory is used.

In this case, in the ideal atmosphere with only scattering distribution from the molecules, the meteorological optic range is 227 km. This indicates that the molecular scattering can mostly be ignored. In visible light, it is sensitive to green light (Wavelength λ = 550 nm) and this wavelength is used as the standard wavelength in measuring the visibility (Tjugum *et al*., 2005; Hun, 2012).

Koschmieder considered the general sensitivity of the human eye to study the effect of atmospheric optical characteristics on the visibility and the meteorological Range, Rm can be indicated through the useful Eq. 1 (Cartney, 1976):

$$R_m = \frac{3.912}{\beta_{sc}} \quad (1)$$

Where:

β_{sc} = The scattering coefficient

β_{sc} = First calculated based on the Mie theory

The average density of the aerosol is assumed with average density aerosol of general metropolitan cities and the impact of the scattering coefficient is the 550 nm attenuation coefficient of the visible ray. The absorption is considered to use the Mie theory.

RESULTS AND DISCUSSION

Pollution level prediction system

Realization of the pollution level prediction system: In Fig. 1, through comparing the back scatter technique and frequently used methods, the forward scattering method

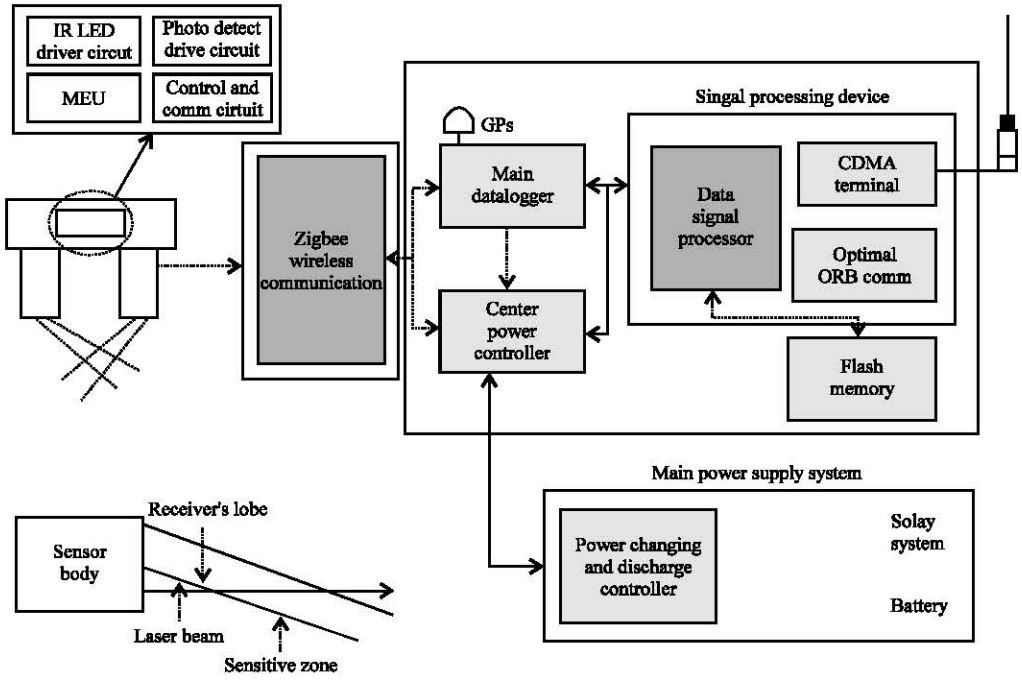

Fig. 1: Operation principle of the dust sensor and system configuration diagram

creates another document. If there is light in the sensitive zone, the back scattering receives signal in the sensor (Kim, 2013).

Signal is created from the sensor output and the sensitive zone is located approximately 30 cm in front of the sensor. In addition, the volume is to be 1 cm^3 or less. The Back scattered light includes the information on atmospheric amount of rainfall and fog particles (Gorky, 2017).

The back scattering measurement method is calculating the coefficient scattered in the opposite direction to the incident light and progression direction and the scattering coefficient is calculated through the Lidar Equation. In this study, the reflected particle is outputted as AD and the amount of fine particles are measured to estimate the abnormality (Grus, 2016).

Dust monitor: The scattered lights in multiple particles are synthesized to be detected as the brightness (illumination). In addition, the photoelectric current of the light receiving element and the concentration of the particle are in proportion but the proportional constant may differ depending on the particle type. Even in the overall clean state, high photoelectric current may flow when big particles such as dust flow in.

To analyze the atmospheric particles, the current space detector using the back scattering technique is applied. The detector outputs analog and digital RS232 providing the information. The output from the detector is given as the ratio of amount outputted for the last 5 seconds to the 1/1000 of 1 mm (= μm). In addition, the test box is to be designed to test cigarette smoke, one of the fine dusts.

Pollution level prediction algorithm

Decision tree: The decision tree algorithm uses the tree structure to show various decision paths and outcomes for the prediction with local data (Grus, 2016).

In addition, to create the decision tree, the question that includes the most information on the target for prediction is selected which is called entropy.

As the classification model that analyzes the data collected in the past to show in the combination of properties on the patterns (characteristics for each category) existing between them, the new data is classified or the relevant category value is used for the purpose of prediction and to extract the knowledge of the tree structure generalized from the data. According to the target variable type, the classification type is classified as classification tree and the continuous type is classified into regression tree.

The analysis process of the decision tree includes repetitive division (data for training is used to divide the dimensional space of the independent variable, repetitively) and branching (data for evaluation is used to perform branching).

Table 1: Test data sample

Dates	Hour	Min.	Sec	Transmitter AD value	Receiver AD value
2017-01-09	10	12	33	52	23
2017-01-09	10	12	38	52	23
2017-01-09	10	12	43	52	22
2017-01-09	10	12	48	52	23
2017-01-09	10	12	53	52	23
2017-01-09	10	12	58	52	23
2017-01-09	10	13	03	52	23
2017-01-09	10	13	08	52	23

Decision tree algorithm

CART (Classification and Regressing Trees): Most widely used algorithm. The target is possible for both category type and continuous type and has binary separation. For the impurity measurement, the category-type uses Gini's coefficient and the continuous-type uses dispersion. The optimum separation can be found not only in the individual input variables but also among the linear coupling of the input variables:

- C4.5 and C5.0
- Multiple split possible, Nominal target variable for the impurity measurement, the entropy index is used
- CHAID (chi-squared automatic interaction detection)
- The growth of tree model is stopped in appropriate size. Continuous target variable. For the impurity measurement, the chi-squared statistics are used

In the test, the CHAID (Chi-squared Automatic Interaction Detection) model will be used and the measured data will be measured in 5 sec interval through time: year-month-day/hour-minute-second/AD output value. In the left and right side of the IR sensor, it is shown that the transmitter values of the IR sensor maintain 52 and the values measured from the receiver are reflected by the fine dust in the space to show multiple values.

Sample data output type is time: year (4 digits) month (2 digits) day (2 digits)/h (2 digits): minute (2 digits): second (2 digits), ad: (3digits×2), 5 sec. The measured data are shown in Table 1 and in order to compare the contamination in the emergency room, it will be installed in the a-entrance, the b-bed and the c-vent.

The transmitter and the receiver were designed to be in the form of same direction for the measurement. In the test box produced for the transmitter and the receiver test of the IR sensor, the measured values are converted and expressed in A/D and cigarette smoke was used to demonstrate the fine dust randomly. The cigarette smoke demonstrated as the fine dust is flowed into the air, increasing the fine dust particles to show its significance.

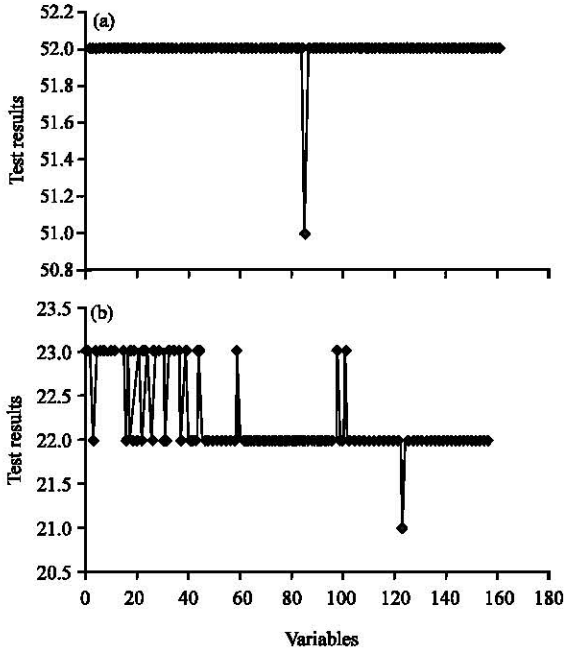

Fig. 2: Results of test in the test box

The value investigated in the transmitter of the left graph in Fig. 2 causes back scattering and becomes much smaller than the absorption value of the particle measured in the receiver of the right graph which signifies that there is a great amount of the fine dust in the air and very little particles are measured. The measured values can be used to assume the occurrence of problems due to the increase in the fine dust.

CONCLUSION

Regarding the seriousness of infection within the Emergency Room (ER) due to the various infectious diseases from virus such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) that occurred recently in the last few years, many patients that visited the ER for treatment were infected and resulted in death.

As the infectious viral disease infecting the ER through the atmosphere became suspicious, the installation of the monitoring that can measure the air pollution level within the ER to prevent the spread is being required. Installation is simple for measuring the air and with simple IR sensor, the back scattering is used to count the atmospheric particles to notify the danger of the pollution level in advance. Also, when alarming the pollution level suspicious of having problems, the ER air is collected to identify the existence of harmful virus and the type of viruses so the spread of virus can be

prevented in advance. In addition, deep learning is used to identify the types of virus that may be in the fine dust particles inside the collected air. And the measures of individuals and government-related agencies on the past occurrences of virus are compared and analyzed to have the advantage of notifying in advance on the actions that must be taken to supply and demand the vaccine for preventing the spread and to prevent the virus infection. For the test, simple test box and the measurement system were used and cigarette smoke was used to demonstrate the fine dust but in further studies, the system will be installed in the actual ER. For the deep learning method, the decision method of CHAID will be used but in the future, Pyton library will be used for the analysis of measurement data.

REFERENCES

Cartney, E.J.M., 1976. Optics of the Atmosphere: Scattering by Molecules and Particles. Wiley, Hoboken, New Jersey, ISBN:9780471015260, Pages: 408.

Cheong, C.H., Y.M. Seo, S.H. Lee and J.J. Kang, 2016. Improvement of Airborne Infection Prevention Methods in Emergency Room by Design Case Study. Kinetic Concepts Corporation, San Antonio, Texas, Pages: 433.

Gorky, S., 2017. Deep Learning from Scratch. O'Reilly Media, Sebastopol, California, Pages: 268.

Grus, J., 2016. Data Science from Scratch. O'Reilly Media, Sebastopol, California, Pages: 203.

Hun, K., 2012. Implementation o f a fog concentration measurement system using infrared communications. Masters Thesis, Mokpo National Maritime University, Mokpo, South Korea.

Kim, S.W., 2013. Visibility of the transparency of the atmosphere. Stedelijk Museum Bureau Amsterdam, Amsterdam, Netherlands.

Kinney, W.M., 2016. Python for Data Analysis. O'REILLY, Sebastopol, California, Pages: 40.

Tjugum, S.A., J.S. Vaagen, T. Jakobsen and B. Hamre, 2005. Use of optical scatter sensors for measurement of visibility. J. Environ. Monit., 7: 608-611.