Rain Echo Classification by Deep Learning Using Observation Big Data of Phased Array Weather Radar

*Shinsuke Satoh[1], Hiroshi HANADO[1], Seiji Kawamura[1], Tomohiko Konno[1], Michiaki Iwazume[1], Shigenori Otsuka[2], Takemasa Miyoshi[2]

1. National Institute of Information and Communications Technology, 2. RIKEN Center for Computational Science

The phased array weather radar (PAWR) was developed with the goal of early detection and prediction of sudden weather disasters caused by torrential rain and tornados and gusts. The PAWRs installed in Suita, Kobe, Okinawa, are operated regularly and generate three-dimensional data every 30 seconds. The observation big data exceeding 2 PB has been accumulated. The real-time observation data is used by the smartphone app "3D Rain Watch" and "RIKEN 3D Nowcast" and the research results of data analysis and data assimilation have been published, but the accurate prediction of torrential rain could not be realized. Also, many of archived past data cannot be used effectively. In recent years, it has become an era when a computer can realize judgment using a deep learning which has developed rapidly. Although deep learning needs to be learned using a large amount of data, new results which cannot obtain by human data analysis, are expected. In this study, we aim to automatic extraction of the precursor phenomena leading to torrential rain prediction by using deep learning technology called convolutional neural network (CNN). As the first attempt, we try to classify rain echoes from two-dimensional image data of rainfall distribution observed by PAWR. In terms of rain echo classification, the difference in rainfall type, convective or stratiform, is also important, but in addition to that, we attempt to classify rainfall distribution patterns such as isolated, linear, and mass.

The original data of PAWR observation is archived as binary data of polar coordinates, but for this deep learning we use a quick look (QL) image showing the horizontal distribution of the radar echo intensity of the altitude of 2 km which is published in real time on the Web Page. The number of accumulated QL images of Kobe PAWR reaches 4 million in total, but more than half are QL images without rain in the radar observation range of 120 km in diameter and cannot be used for rain echo classification. In order to perform deep learning, it is necessary to label the QL images, but it is not easy for a man to manually label 10000 images. Fortunately, in order to create a rain summary graph on the Web, numerical information of the average rainfall amount, maximum rainfall amount, and rainfall area corresponding to each QL image is saved as a text file. After rough labeling using the information, eventually it is judged by human eyes. Finally, we classify into eight categories of strong / weak isolated convective, linear convective, mass convective, stratified, and no rainfall. Using more than 900 samples in each category from 50 days observation data in June, July, August 2016, we repeatedly learned them about 1000 times using a simple 7-layer CNN, and as a result it was able to perform rain echo classification with an accuracy of 72%. We believe it is possible to increase accuracy by removing ambiguous labeling samples or using a more complicated CNN network. In the future, we want to advance research that leads to heavy rain prediction by using three-dimensional data and time-series data.

Keywords: Phased Array Weather Radar, Observation Big Data, Deep Learning, Rain Echo Classification

©2019. Japan Geoscience Union. All Right Reserved.