

Using Optical Character Recognition to Manage the Unstructured Disaster Data into Smart Disaster Management System

Authors : Dong Seop Lee, Byung Sik Kim

Abstract : In the 4th Industrial Revolution, various intelligent technologies have been developed in many fields. These artificial intelligence technologies are applied in various services, including disaster management. Disaster information management does not just support disaster work, but it is also the foundation of smart disaster management. Furthermore, it gets historical disaster information using artificial intelligence technology. Disaster information is one of important elements of entire disaster cycle. Disaster information management refers to the act of managing and processing electronic data about disaster cycle from its' occurrence to progress, response, and plan. However, information about status control, response, recovery from natural and social disaster events, etc. is mainly managed in the structured and unstructured form of reports. Those exist as handouts or hard-copies of reports. Such unstructured form of data is often lost or destroyed due to inefficient management. It is necessary to manage unstructured data for disaster information. In this paper, the Optical Character Recognition approach is used to convert handout, hard-copies, images or reports, which is printed or generated by scanners, etc. into electronic documents. Following that, the converted disaster data is organized into the disaster code system as disaster information. Those data are stored in the disaster database system. Gathering and creating disaster information based on Optical Character Recognition for unstructured data is important element as realm of the smart disaster management. In this paper, Korean characters were improved to over 90% character recognition rate by using upgraded OCR. In the case of character recognition, the recognition rate depends on the fonts, size, and special symbols of character. We improved it through the machine learning algorithm. These converted structured data is managed in a standardized disaster information form connected with the disaster code system. The disaster code system is covered that the structured information is stored and retrieve on entire disaster cycle such as historical disaster progress, damages, response, and recovery. The expected effect of this research will be able to apply it to smart disaster management and decision making by combining artificial intelligence technologies and historical big data.

Keywords : disaster information management, unstructured data, optical character recognition, machine learning

Conference Title : ICNHST 2019 : International Conference on Natural Hazard Science and Technology

Conference Location : Dubai, United Arab Emirates

Conference Dates : November 07-08, 2019