学　位　論　文　内　容　の　要　旨

博士の専攻分野の名称　　博士（工学）　　氏名　VACHMANUS SIRAWICH

学　位　論　文　題　名

Proposal of a Segmentation Algorithm using Multi-Sensor Fusion for Autonomous Driving in Snowy Environments

(積雪環境における自律走行を目的としたマルチセンサフュージョンによるセグメンテーションアルゴリズムの提案 )

Currently, research and development of autonomous driving technology is being actively conducted around the world. In addition to vehicle driving technology, autonomous driving requires complex technologies and knowledge, such as sensors to recognize the environment, artificial intelligence (AI) to process vast amounts of data and make appropriate decisions, and software to integrate these technologies. However, most of these research and development efforts have been conducted in non-snowy environments, and almost none have been conducted in snowy or snow-falling environments. On the other hand, one half of Japan's land area (24 prefectures and 532 municipalities) is designated as a heavy snowfall area, and it is essential to deal with snowy and snow-falling environments to maintain the transportation infrastructure. As evidenced by the fact that the top three cities in the world with populations of 100,000 or more in terms of annual snowfall are Japanese cities, research and development of autonomous driving technology for snowy and snow-falling environments is an important issue that Japan should proactively tackle.

In recent years, Japan's declining birthrate, aging population, and labor shortages have become serious social issues, and this has led to a sharp rise in labor costs. Snow removal work in areas with heavy snowfall, such as Sapporo, is essential for maintaining the transportation infrastructure, and is greatly affected by the rising cost of labor. If autonomous driving technology can support or partially automate snow removal work by adapting to snowy and snow-falling environments, labor costs can be greatly reduced.

The challenges of autonomous driving on snow-covered roads can be broadly classified into three categories as shown below.

(1) Infrastructure for automated driving: In cold regions, signs, road boundaries, and lanes are covered with snow due to snowfall, and self-position estimation using dynamic maps (highly accurate spatial information), a commonly used automated driving infrastructure, does not function.

(2) Autonomous driving in compliance with laws and regulations: Signals cannot be recognized due to snow on the traffic signals. Snow on traffic signs prevents recognition of speed limits and road signs.

(3) Vehicle control in cold weather: Anti-skid control (vehicle stability control) for curves and sudden steering. Anti-lock control of brakes. Traction control for starting and accelerating.

For (2), the Civil Engineering Research Institute for Cold Region and other organizations are conducting research on installing covers to prevent snow from sticking to traffic signals and changing the shape and materials of signs. The same issues are being studied for (3) as for conventional vehicles, and research and development of (3) is being conducted by automobile manufacturers. Based on the above, this research aims to solve the issues related to (1) infrastructure for autonomous driving.

The dynamic map method used in general autonomous driving realizes autonomous driving by using a pre-prepared high-precision 3D map and determining where the vehicle is located on the map (i.e., self-position estimation) in real time. However, the existing self-position estimation method may not work on snow-covered roads because of the large discrepancy between the map prepared in advance (non-snow-covered environment) and the actual environment (snow-covered environment). Furthermore, for example, a four-lane road may frequently be reduced to three or two lanes due to snow accumulation during the winter, making accurate self-location estimation by the dynamic map method meaningless. In other words, the surrounding environment on a snow-covered road is not always clear. In other words, on snow-covered roads, it is necessary to recognize in real time the area that can be traveled by sensing the surrounding environment. Currently, RGB cameras and 3D-LiDAR are mainly used as sensors for human detection in autonomous driving technology, and various methods have been proposed. However, in bad weather conditions such as snow, rain, and fog, these sensors are unable to accurately detect people, and noise remains a problem. Especially in poor visibility environments such as snowfall, it is not possible to observe all features with a single sensor because the road surface is often covered with snow. Therefore, this paper proposes a multimodal RGB-T semantic segmentation method using RGB images from an RGB camera and thermal information from a thermal imaging camera as a recognition method for snowy environments.

In this paper, I propose a single-input and multiple-input semantic segmentation system for snowfall environments. First, I propose the pyramid monitoring paths for single-input segmentation system to improve the intrinsic network performance so that the system can work in poor environments. Then I propose the second system, multiple-input segmentation. It utilizes thermal information as a second input and includes a pyramid monitoring path as a decoder. The proposed methods show robust human detection and high mIoU on a dataset of snowfall environments. The network with fusion modules, multiple inputs, and pyramid monitoring paths is a state-of-the-art network for snowfall environments.

The paper consists of six chapters, each of which is summarized below.

Chapter 1 is an introduction, which describes the background of this study, including an overview of snow removal on sidewalks in Sapporo City and the current autonomous driving technology.

Chapter 2 describes the research background and previous studies in detail. Since this research is based on the semantic segmentation, which is a method of deep learning, computer vision, deep learning, etc. are explained and related works are surveyed.

Chapter 3 describes the proposed algorithm for the single-input case, the datasets used to train and evaluate the network, and the results of training using these datasets.

Chapter 4 describes the case where the network in Chapter 3 is extended to multiple inputs. As in Chapter 3, it provides an overview of the proposed algorithm, describes the datasets for training and evaluating the network, and describes the training and results using these datasets.

In Chapter 5, I discuss the single-input and multiple-input semantic segmentation system for snowfall environments proposed in Chapters 3 and 4.

Chapter 6 is the conclusion, in which the results obtained in this study are described.