환경 차이에 따른 돼지 발성음 분류 모델의 일반화 성능 평가

Evaluation of Generalization Performance of Pig Vocalization Classification Models under Different Environmental Conditions

판반뎃, 권경석, 김병현, 장동화, 김재은, 김종복*

Vandet Pann, Kyeong-Seok Kwon, Byeonghyeon Kim, Dong-Hwa Jang, Jaeeun Kim, Jong-Bok Kim*

농촌진흥청 국립축산과학원 스마트축산환경과

National Institute of Animal Science, Rural Development Administration, Wanju, Korea

***교신저자:** 김종복 (jbkimj@korea.kr)

While recent progress in classifying pig sounds has produced several strong and reliable methods, building a model that works well across different farms is still a major challenge. Factors like background noise, variations in pig vocalizations, limited labeled data, the need for real−time processing, and system integration issues all make it difficult to create a truly general solution. This study aims to evaluate the performance of the pig sound classification model under different environmental conditions on real pig farms using a convolutional neural network (CNN) model. For the experiment, two datasets were collected from two actual pig farms: Jeongeup and Iksan. Each dataset consists of 4000 WAV files (1,000 non−vocal, 1,000 coughing, 1,000 screaming, and 1,000 other) with a duration of three seconds. A CNN model with three convolutional layers, max−pooling, and two fully connected layers is trained using Mel−spectrogram features, batch normalization, and dropout to enhance training speed and reduce overfitting. This study used the k−fold cross−validation (k=5) technique to assess the model performance. By conducting rigorous experiments, the model achieved accuracy on the Jeongeup and Iksan datasets, with rates of 98.22 % and 96.93 %, respectively. A generalization experiment was performed to evaluate the effectiveness of the model by using the Jeongeup dataset as a training set and the Iksan dataset as a testing set. The average performance of the model in terms of accuracy, precision, recall, and F1−score achieved 88.86 %, 90.08 %, 88.87 %, and 89.47 %, respectively. The experimental results show that the model performs well classifying pig sounds on the same farm dataset. However, its performance slightly declines when tested on a dataset from a different pig farm. Based on the observations in this study, a well−generalizable model requires a robust network architecture along with numerous and diverse pig sound data.

***키워드(Keywords):** precision livestock farming, pig sounds classification, audio feature extraction, audio data augmentation, convolutional neural networks (CNNs)

***사사(Acknowledgment):** This work was carried out with the support of "Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01681003)" Rural Development Administration, Republic of Korea. This study was supported by 2025 the RDA Fellowship Program of National Institute of Animal Science, Rural Development Administration, Republic of Korea. This study was supported by 2025 collaborative research program between university and Rural Development Administration, Republic of Korea.