Journal of KIIT. Vol. 16, No. 5, pp. 1-13, May 31, 2018. pISSN 1598-8619. eISSN 2093-7571 1
http://dx.doi.org/10.14801/jkiit.2018.16.5.1

CNN 기반의 소리 잡음에 강인한 돼지 호흡기 질병 탐지 및 식별 시스템

이종욱*, 최용주**, 박대희***, 정용화****

Sound Noise-Robust Porcine Wasting Diseases Detection and Classification System Using Convolutional Neural Network

Jonguk Lee*, Yongju Choi**, Daihee Park***, and Yongwha Chung****

이 논문은 2015년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2015R1D1A3A01018731)

요 약

상업적 시설에서 키우고 있는 돼지의 호흡기 질병을 초기에 정확하게 탐지하지 못한다면 해당 농가의 생산성에 문제를 발생시키는 심각한 요인이 된다. 본 논문에서는 돼지의 소리 데이터를 이용하여 현장에서 발생하는 소음에도 강인한 호흡기 질병 탐지 및 식별 시스템을 제안한다. 제안된 시스템은 먼저, 소리 신호에서 스펙트로그램 정보를 추출하고, 이를 CNN에 적용하여 돼지 호흡기 질병을 탐지 및 식별에 효과적인 특징 정보로 변환한다. 마지막으로, 변환된 정보는 MLP에 적용되어 해당 호흡기 질병을 탐지 및 식별과정을 수행한다. 본 연구의 실험 결과, 다양한 잡음 환경에서도 돼지 호흡기 질병 탐지 및 식별 성능이 안정적임을 확인하였다. 특히, 본 연구에서 제안한 방법과 MFCC와 SVM을 이용한 방법의 식별 성능을 비교한 결과 f-score 수치가 15.1% 향상됨을 확인하였다.

Abstract

Failure to detect pig wasting disease in a timely and accurate manner in the commercial pig farm can be a serious factor in achieving efficient livestock management. In this paper, we propose a noise-robust porcine wasting diseases detection and classification method in piglet farm monitoring system using sound data. First, we extract a spectrogram of sound signals and convert it into noise-robust features by a convolutional neural network (CNN), and lastly, use the multi-layer perceptron (MLP) as an early anomaly detector and classifier. On the basis of the experimental results, we confirmed that the proposed method could detect and classify the porcine wasting diseases with acceptable accuracy even under noise-environmental conditions. In particular, as a result of comparing the discrimination performance of the proposed method in this research and the MFCC-SVM method, it was confirmed that the f-score was improved by 15.1%.

Keywords
porcine wasting diseases classification, sound analysis, convolution neural network

* 고려대학교 컴퓨터융합소프트웨어학과 초빙교수
- ORCID: https://orcid.org/0000-0002-2077-4850
** 고려대학교 컴퓨터정보학과
- ORCID: https://orcid.org/0000-0003-4661-6196
*** 고려대학교 컴퓨터융합소프트웨어학과 교수(교신저자)
- ORCID: https://orcid.org/0000-0003-4726-4508
**** 고려대학교 컴퓨터융합소프트웨어학과 교수
- ORCID: https://orcid.org/0000-0001-6539-167X

· Received: Feb. 22, 2018, Revised: Mar. 26, 2018, Accepted: Mar. 29, 2018
· Corresponding Author: Daihee Park
Dept. of Computer Convergence Software, College of Science and Technology, Korea University, Sejong 30019, Korea
Tel.: +82-44-860-1344, Email: dhpark@korea.ac.kr

Ⅰ. 서 론

2017년 농림수산식품부 통계[1]에 의하면, 우리나라 축산업의 2016년도 총생산액은 약 18조 9천억 원으로 보고되고 있다. 그 중에서 양돈업계의 생산액은 약 6조 7천억 원으로 전체 농업에서 단일 업종으로는 벼 생산 다음이며, 축산 분야에서는 가장 많은 부분을 차지하고 있다. 또한, 우리나라 양돈농가의 사육규모는 2017년 현재 1,043만 여두이며, 소규모 사육농가는 급격히 감소하는 반면 중·대규모 사육농가는 급격히 증가하고 있다. 특히, 우리나라 농가의 60세 이상 농업종사자의 비중이 2000년 44.1%에서 급격히 증가하여, 2015년에는 62.2%로 고령화가 심화되어 영농인력의 부족현상이 발생하고 있다[2]. 즉, 양돈농가의 규모가 전업화·기업화·대형화 및 고령화에 따라 관리인 1인당 관리두수가 급격하게 증가되었기에, 개별 돼지의 건강을 보장하기 위한 세밀한 관찰 및 관리는 현실적으로 어려운 상황이다. 결국, 이로 인한 돼지의 질병 발생 가능성은 증가할 것으로 예상되며, 돼지 열병, 브루셀라, 호흡기 질병 등이 발생 시 대규모로 확산될 가능성이 매우 높은 것이 현실이다.

돼지 호흡기 질병을 유발하는 환경인자 중, 양돈장의 주요 세균성 및 바이러스성 호흡기 질병의 원인체는 MH(Mycoplasma Hyopneumonia), PCV2(Porcine Circo Virus 2) 및 돼지생식기호흡기증후군(PRRS, Porcine Reproductive and Respiratory Syndrome) 바이러스가 주된 병원체이며, 이유후전신성소모성증후군(PMWS, Postweaning Multisystemic Wasting Syndrome)을 유발한다. 이는 돼지의 체중감소, 만성 호흡기 질환, 설사, 황달 등의 심각한 질환으로 나타나며, 돼지의 폐사로까지 이어지게 된다[3]. 결국 이와 같은 질병들은 이유자돈을 키우는 축산농가에 막대한 경제적 손실을 초래함으로, 질병을 조기에 탐지할 수 있는 현실적인 방안이 요구된다. 그러나 현재 기술적, 경제적 문제 등으로 인하여 돼지 호흡기 질병을 위한 조기 방재 시스템은 전무한 실정이다.

최근 축산 농가의 적은 일손으로도 효과적으로 건강한 가축을 사육하고자 하는 노력의 일환으로, ICT(Information and Communication Technology) 기술과 농·축산업과의 융합 기술(Computers and Electronics in Agriculture)이라는 새로운 연구 분야가 선진 외국을 중심으로 활발하게 진행되고 있다[4]-[10]. 예를 들면, 돼지들의 사료 소모량, 습도와 온도 확인 등의 돈사 관리를 위한 기본적인 모니터링을 시작으로 가축의 행동을 보다 심층적으로 분석하기 위하여 영상 및 소리 센서를 활용하고 있다. 영상을 이용한 연구들을 살펴보면 RGB 카메라 또는 키넥트 카메라를 이용하여 이유자돈 사이의 이상행동 중 하나인 공격적인 행동을 탐지하기 위한 연구[4], 24시간 돈사의 이상상황을 모니터링하기 위해 선행되어야 하는 개별 돼지의 탐지 문제를 RGB 카메라에서 취득한 영상을 GMM(Gaussian Mixture Model)에 적용하여 해결하고자 하는 연구[5], 컨볼루션 신경망을 이용하여 영상정보에 나타난 개별 돼지를 탐지하는 방법[6], 빛에 의한 영향에서 벗어나 야간에도 탐지가 가능한 키넥트 카메라를 기반으로 개별 돼지를 탐지하려는 학술적 연구[7]들이 보고되고 있다. 반면, 영상정보를 활용한 경우와 달리 실제 동물의 생체 정보를 효과적으로 담고 있다고 알려진 소리 신호를 이용한 심층적인 연구들도 다수 발표되고 있다. 예로써, 산란계가 스트레스로 인한 계란 품질의 저하를 방지하기 위하여, 산란계의 발성음을 MFCC(Mel-Frequency Cepstrum Coefficeint) 소리 특징으로 변환 후 산란계의 스트레스를 탐지하는 연구[8], 닭의 발성음을 신호처리 기법과 기계학습으로 질병을 탐지하는 연구[9], 암소의 발성음을 기반으로 발정기를 탐지하는 방법[10]들이 수행되고 있다.

본 논문에서는 소규모의 농장에서도 저비용으로 구축이 가능한 소리 센서 환경에서의 돼지의 호흡기 질병 탐지 및 식별 시스템을 제안한다. 특히 본 시스템은 실제 돈사 환경에 설치되어 돼지 호흡기 질병의 탐지 및 식별이 현실적으로 실현 가능할 수 있도록, 실제 돈사 환경에서 발생하는 다양한 잡음에 강인한 시스템을 설계한다. 제안된 시스템은 학습 데이터의 증폭과 함께 최근 각광을 받고 있는 CNN의 커널 및 구조를 자체 설계하여 소리 신호에서 스펙트로그램 특징을 추출한 후 이를 CNN의 입력 정보로 사용한다. CNN 계층(Layer)의 학습 과정

을 거치면서 적응적으로 호흡기 질병을 식별하기 좋은 특징 벡터로 자동 변환한 후 호흡기 질병 식별기인 MLP(Multi Layer Perceptron)에 입력함으로써, 정상소리와 호흡기 질병 소리들 중 하나로 식별한다. 실제 돈사에서 취득한 소리신호를 이용하여 본 연구에서 제안하는 시스템의 성능을 실험적으로 검증하며, 기존 방법론과의 정성적·정량적 성능 비교를 통해 본 연구의 우수성을 검증한다.

본 논문의 구성은 다음과 같다. 2장에서는 돼지 호흡기 질병 탐지 시스템과 관련한 최근의 연구방법론들을 간략하게 소개하고, 3장에서는 본 연구에서 제안하는 잡음에 강인한 CNN 기반의 돼지 호흡기 질병 탐지 및 식별 시스템에 대해 기술한다. 4장에서는 제안한 알고리듬으로 실험한 결과를 객관적인 지표로 검증 및 성능을 평가하고, 마지막으로 5장에서 결론 및 향후 연구에 대해 언급한다.

II. 관련 연구

본 연구의 관심 대상인 동물의 발성음을 기반으로 호흡기 질병을 탐지하는 최근의 연구동향을 살펴보면 다음과 같다[11]-[14]. Vandermeulen 등[11]은 소가 내는 울음소리의 스펙트로그램을 분석하여 소의 호흡기 질병을 탐지할 수 있음을 확인하였으며, Guarino 등[12]은 소리 주파수 대역에서 필터링 기법과 진폭 변조 등의 기법을 적용하여 특징 벡터를 생성한 후 DTW(Dynamic Time Warping) 기법을 이용하여 돼지의 호흡기 질병을 식별하는 연구를 수행하였다. Chung 등[13]은 돼지 호흡기 질병으로 인한 기침 소리 여부를 MFCC 소리 특징정보와 단일 클래스 탐지기인 SVDD(Support Vector Data Description)를 이용하여 호흡기 질병을 탐지한 후, SRC(Sparse Representation Classifier)를 이용하여 해당 기침의 질병 종류를 식별하는 이중 구조를 제안하였다. Lee 등[14]은 돼지의 호흡기 질병을 효과적으로 탐지하기 위하여, 시간 영역과 주파수 영역의 다양한 소리 특징들 중 돼지 호흡기 질병 탐지에 유효한 특징들만을 선택 및 조합하는 방법에 관한 연구결과를 발표하였다. 이와 같이 소리 기반 돼지 호흡기 질병의 탐지가 현실적으로 가능함을 확인하는 학술적 노력들이 일정 부분 성공적인 결과를 도출하였으나, 실제 돈사 환경에서 발생하는 잡음의 영향력을 충분히 고려했다고는 볼 수 없다. 학술적 프로타입 시스템을 실제 돈사환경에서 성공적으로 운용하기 위해서는 소리 취득 시 발생하는 다양한 잡음에서도 시스템의 강인함이 보장되어야만 한다. 실제 돈사에서는 돼지들이 돈사에서 움직이는 소리, 돼지 성장을 위해 음악을 틀어놓는 등의 환경잡음과 소리 취득 시 소리 정보의 왜곡 등과 같은 인위적인 잡음 등이 존재한다.

한편, 영상 정보를 해석하기 위해 제안된 딥러닝 기법들을 소리 정보를 기반으로 특정 상황을 식별하는 연구에 적용하는 사례들이 최근 발표되고 있으며, 실제 식별 성능이 기존의 알고리즘들 보다 우수한 성능을 보이고 있다. 특히 일부 연구에서는 한정된 학습 데이터를 증폭함으로써 학습 데이터의 다양화와 함께, 풀고자 하는 문제에 최적화된 딥러닝 구조를 설계함으로써 소리 또는 진동 신호의 잡음이 혼재된 환경에서도 이상 상황을 강인하게 식별할 수 있음이 보고되고 있다[15][16]. 예를 들면, Salamon과 Bello[15]는 에어컨 소리, 경적 소리, 개 짖는 소리, 어린아이들이 놀면서 내는 소리 등 10가지의 소리를 소리 길이 변환, 피치 변환등의 소리 데이터 증폭 기법을 이용하여 데이터의 양을 인위적으로 늘린 후 해당 연구에서 제안한 CNN 구조에 적용함으로써, 잡음에 강인한 식별 성능을 확인하였다. Zhang 등[16]은 롤링 머신의 베어링에서 발생하는 고장 문제를 진동신호를 이용하여 탐지하는 방법을 제안하였다. 이 때 첫 번째 CNN 입력 계층의 커널(Kernel)을 넓게 하고, 진동 신호를 이동하여 학습 데이터의 양을 늘리는 방법을 적용한 후 CNN을 학습하였다.

그 결과 인위적인 백색 잡음(White Noise)의 강도가 다양한 상황에서도 베어링의 고장을 효과적으로 탐지함을 실험적으로 검증하였다. 결국 소리 잡음의 영향을 줄이기 위해 잡음 특성을 분석한 후 해당 분석 내용에 맞게 잡음을 줄일 수 있는 필터를 연구자가 설계하는 수고 없이, 학습을 위한 데이터양의 증폭과 함께 효과적으로 CNN의 커널 및 구조를 설계하게 된다면 잡음에도 강인한 시스템을 구축할 수 있음을 확인할 수 있는 중요한 연구 결과이다.

이와 같은 최근 학계의 연구 성과들을 참조하여, 본 논문에서는 딥러닝 기술의 하나인 컨볼루션 신경망을 이용하여 다양한 잡음이 발생하는 실제 돈사 환경에서도 강인한 호흡기 질병 탐지 및 식별 시스템을 제안하고자 한다.

III. 소리 정보를 이용한 돼지 호흡기 질병 식별 시스템

본 논문에서 제안하는 돼지 호흡기 질병의 탐지 및 식별 시스템의 구조는 그림 1과 같이 크게 3개의 모듈로 구성된다: 1) 소리 센서로 부터 소리를 취득하는 모듈, 2) 취득한 소리 데이터에서 스펙트로그램 정보를 추출하는 모듈, 3) 잡음 상황에도 강인하게 돼지의 호흡기 질병을 탐지 및 식별하는 CNN 모듈.

3.1 스펙트로그램 모듈

소리 정보를 특징 벡터화하는 방법 중 하나인 스펙트로그램은 시간에 따른 주파수 및 진폭 정보를 포함하여 다양한 유형의 소리 정보를 효과적으로 디지털 정보로 변환시킬 수 있다. 스펙트로그램은 2차원의 행렬로 표현되며, 이와 같이 표현된 디지털 정보는 CNN의 영상처리에서 사용되는 시각화 정보인 이미지와 유사하게 표현될 수 있다. 특히, 스펙트로그램은 서로 다른 소리 신호를 구분할 수 있는 중요한 정보들을 포함하고 있기에 다양한 소리 분석 연구 분야에서 범용적으로 사용되고 있다[17]. 본 연구에서도 소리 시그널을 이미지 정보로 표현이 가능한 스펙트로그램을 이용하여 CNN의 입력 정보로 사용한다.

3.2 호흡기 질병을 탐지 및 식별하는 CNN 모듈

권위 있는 영상 정보 분류 경연 중 하나인 ILSVRC(ImageNet Large Scale Visual Recognition Challenge) 2012에서 CNN을 기반으로 한 구조인 Alexnet[18]의 우승은 다양한 분야의 연구자들이 CNN에 관심을 기울이게 되는 계기가 되었다. 특히 연구자들의 흥미를 일으킨 부분은 기존의 사람에 의해 설계된 특징(Hand-craft Feature)이 아닌 CNN의 커널 설정 및 해당 구조에 의해 해결하고자 하는 문제에 적합한 고수준의 특징을 CNN이 자동으로 생성할 수 있다는 점이다. 또한 소리 시그널을 스펙트로그램 이미지로 변환하여 마치 영상 정보의 입력과 유사하게 처리가 가능하다는 점은 소리 기반의 연구자들에게도 매우 매력적인 부분이었으며, 실제 소리 연구 분야에 적용한 결과 안정적인 성능을 보임을 확인하였다.

CNN의 기본 구조는 그림 2와 같으며, 크게 3가지의 중요한 구성 요소가 존재한다[17]-[21]: 1) 커널을 이용한 컨벌루션 계층, 2) 풀링(Pooling) 계층, 3) MLP(Multi-Layer Perceptron)로 구현된 분류 계층이며, 해당 시스템의 설계에 따라 컨벌루션 계층과 풀링 계층을 다양하게 조합한 후 최종적으로 생성된 특징 벡터를 분류계층에 입력으로 사용한다.

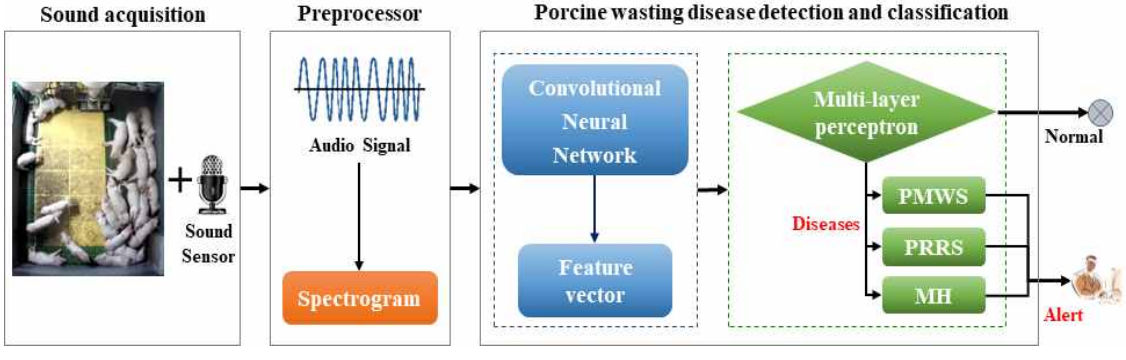


그림 1. 돼지 호흡기 질병 탐지 및 식별 시스템 구조
Fig. 1. System architecture for porcine wasting disease detection and classification

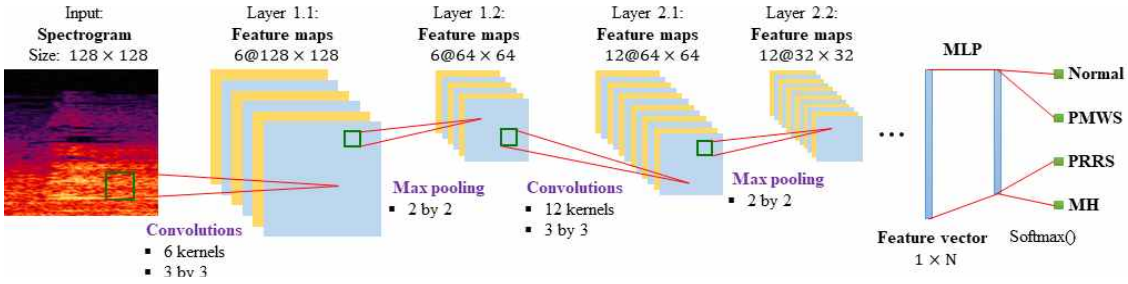


그림 2. CNN 기본 구조
Fig. 2. Basic structure of CNN

1) 커널을 이용한 컨벌루션 계층

전통적인 영상처리 또는 소리의 커널(또는 필터)은 전문가에 의해 해당 이미지 또는 소리의 특정한 특징이 두드러지게 나타나도록 커널의 값을 경험에 의해 설정 했다면, CNN의 커널은 초기에 랜덤으로 설정한 후 해당 컨벌루션 신경망의 역전파(Back-Propagation) 방법에 의해 주어진 클래스를 올바르게 식별할 수 있는 커널의 값이 학습 과정을 거쳐 구해지게 된다.

2) 풀링 계층

소리 정보가 컨벌루션 과정에 적용되어 생성된 특징 맵(Feature Map)은 입력된 소리 이미지와 커널의 개수 등에 따라 해당 정보의 차원이 입력 소리 정보에 비해 매우 크게 확장된다. 이와 같이 매우 큰 특징 차원의 정보를 보다 빠르게 연산하고, 이웃 영역에서 중요한 정보들을 보다 강조하는 방식으로 압축을 진행할 수 있도록 풀링 연산을 수행한다. 최대 풀링의 경우 주어진 크기가 2 by 2, stride가 2로 설정되었다면, 해당 특징 맵에서 2×2 영역에서 가장 큰 값을 대푯값으로 치환하게 된다. 이와 같은 과정을 거치면 특징 맵 면적의 크기는 1/4로 축소된다. 이 과정은 해당 소리의 식별을 위한 중요한 정보는 유지하되 데이터의 차원은 축소시키는 장점이 있다.

3) MLP로 구현된 분류 계층

컨벌루션과 풀링 계층을 반복하면서 생성된 2차원 특징 맵을 의 특징벡터로 변경한 후, 해당 입력 데이터에 맞는 클래스로 식별하기 위해서는 분류기가 필요하다. 일반적인 CNN 구조의 완결 연결 계층(Fully-Connected Layer)에서는 MLP를 분류기로 이용한다. 이 때 개별 클래스에 속할 확률을 소프트맥스(Softmax) 함수를 이용하여 계산한 후, 가장 큰 확률 값을 보인 클래스로 입력된 정보의 클래스를 판단한다.

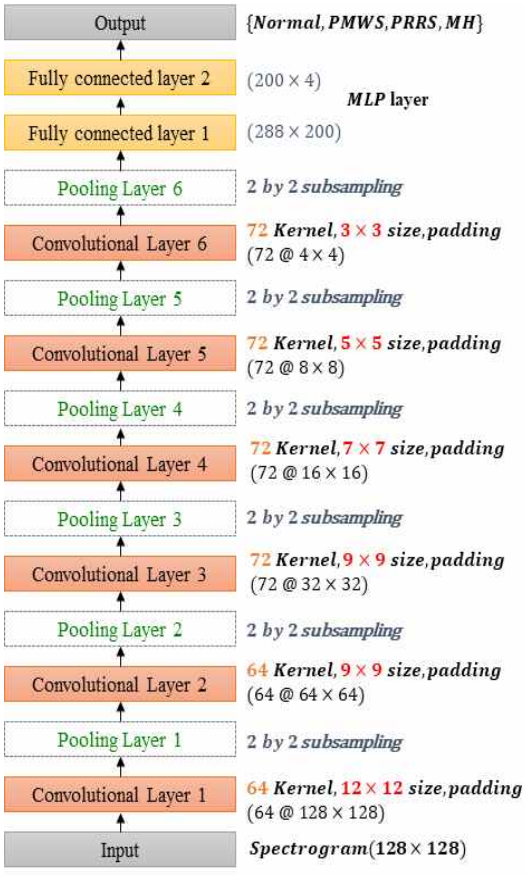


그림 3. 돼지 호흡기 질병 탐지 및 식별을 위한 CNN 구조
Fig. 3. Overall structure of CNN for porcine wasting disease detection and classification

본 논문에서 탐지 및 식별하고자 하는 돼지 소리는 정상인 경우의 소리와 호흡기 질병들 중 하나인 PMWS, PRRS, MH 이며, 본 연구에서 제안하는 CNN 구조는 그림 3과 같다. 소리 잡음에도 강인한 호흡기 질병의 탐지 및 식별 성능을 보일 수 있도록 기존 연구의 설계 구조와는 달리 본 연구에서 제안하는 CNN 구조는 첫 번째 계층에서부터 많은 수의 커널을 설정하였으며, 첫 번째 계층의 커널의 크기는 크게, 계층이 진행될수록 커널의 크기가 점차적으로 작아지도록 설계하였으며 컨볼루션과 풀링 구조가 조합된 6계층으로 CNN 구조를 설계하였다.

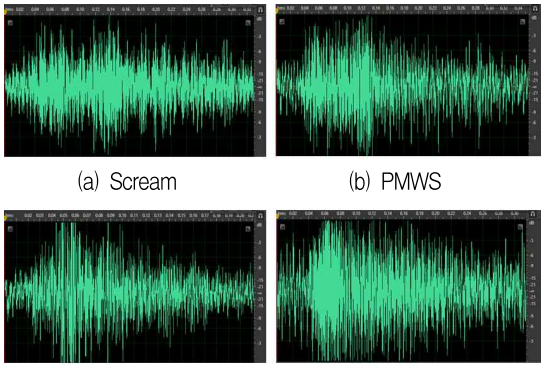
Ⅳ. 실험 및 결과 분석

4.1 실험 환경 및 소리 정보 데이터

실험을 위한 데이터는 충청남도에 위치한 4개의 돼지농장에서 평균적으로 25~30kg의 총 36마리의 돼지(Yorkshire×Landrace×Duroc)를 대상으로 호흡기 질병으로 의심되는 돼지의 혈액을 채취한 후, virus 분리와 serological analysis로 PMWS, PRRS, 그리고 MH에 감염된 개체 및 질병에 걸리지 않은 돼지를 확인하였다. 환축 및 정상축의 음성 수집은 각 개체로부터 1m의 거리에서 디지털 캠코더(JVC GR-DVL 520A, Japan)로 녹화하였다[13].

수집된 소리 정보를 편집한 결과 돼지 소리들은 0.13~2.66초, 샘플링은 44,100Hz, mono 타입이다. 이와 같이 수집된 소리 신호 원본에 현장에서 발생하는 실제 환경 잡음(돼지가 1~2마리가 움직일 때 발생하는 발자국 소리(Weak Footprint), 3마리 이상의 돼지들이 뛰면서 발생하는 보다 시끄러운 발자국 소리(Strong Footprint), 돼지 성장에 도움을 준다고 알려진 음악을 틀어둔 상태(Music))과 인위적인 백색 잡음을 SNR(Signal-to-Noise Ratio) {0, 3, 6, 9, 12, 15, 18, 21}dB로 합성한 데이터를 추가하였다. 이 때 weak footprint, music, strong footprint의 SNR 값은 각각 9.1, 8.7, 7.4dB이며, SNR 수치가 작을수록 잡음의 강도는 강하다. 돼지의 정상소리(Grunt, Healthy Cough, Scream) 중 하나인 scream과 호흡기 질환 소리의 종류별 신호 형태의 예는 그림 4와 같으며, 정상소리 중 하나인 scream과 호흡기 질병 소리인 PMWS, PRRS 그리고 MH 질환은 소리 시그널 자체로는 육안으로 쉽게 구분하기 어려움을 알 수 있다.

실험에 사용한 컴퓨터의 사양은 Windows 10, Intel® i7-7700K CPU@4.2GHz, 32GB 램, GeForce GTX 1080Ti 이다. 실험을 위하여 취득한 정상소리는 총 350 {grunt 110, healthy cough 100, scream 140}, 호흡기 질병 소리는 총 360 {PMWS 150, PRRS 140, MH 70}이다. 이 때 CNN구조는 학습 데이터의 양이 크면 클수록 잡음 환경에서도 식별 성능이 좋아진다는 연구 결과[15][16]에 따라, 본 논문에서는 Zhang 등[16]이 사용한 소리 데이터 확장 방법 중 하나인 데이터 이동(data shift) 기법을 이용하여 학습 데이터를 확장하였다. 소리 데이터 이동 기법은 그림 5와 같이 소리 시그널의 데이터 취득 시작 위치를 이동함으로써 여러 개의 데이터를 확보한다. 본 연구에서는 5번의 이동 과정을 거쳐 데이터를 5배 확장하였다.



(a) Scream (b) PMWS

(c) PRRS (d) MH

그림 4. 돼지 소리 샘플
Fig. 4. Sample waveform of pig sounds

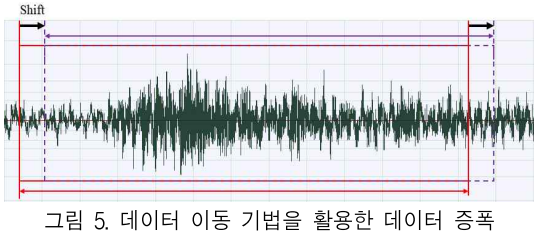


그림 5. 데이터 이동 기법을 활용한 데이터 증폭
Fig. 5. Data augmentation with a data shift technique

4.2 실험 및 결과 분석

잡음 환경에도 강인한 돼지 호흡기 질병 식별을 위한 실험에서는 최근 구글(Google)에서 개발한 오픈소스 기계학습 엔진인 텐서플로우(TensorFlow) 1.02[22]를 활용하여 실험을 진행하였으며, 전체 데이터의 80%를 무작위로 선택하여 학습하였고, 나머지 20%를 테스트 데이터로 사용하였다. 모든 데이터는 동일한 크기인 128×128 스펙트로그램 크기로 조정한 후 CNN의 입력으로 사용하였다. CNN 학습에 사용한 옵션들은 Xaiver.initializer, 학습률은 0.001, Max pooling, 각각의 컨벌루션 계층의 경우 padding 기법을 적용하였으며, drop out 비율은 50%, activation 함수는 ReLU(Rectified Linear Unit)로 설정하였다. 커널의 크기는 12×12로 시작한 후 점차 작은 크기로 설정하였으며, 최종 커널의 크기는 3×3이다.

실험에 사용한 평가지표는 precision, recall, f-score이며, precision은 특정 클래스로 예측된 결과들 중에서 실제 해당 클래스에 속하는 비율, recall은 특정 클래스를 예측과정에서 해당 클래스로 올바르게 판단한 정도이며, 마지막으로 f-score는 precision과 recall간의 트레이드오프(Trade Off)를 통합하여 조화평균을 구한 것이다. 해당 수식에서 tp는 true positive, fp는 false positive, fn는 false negative를 의미한다.

$$Precision = \frac{tp}{tp + fp} \times 100 \quad (1)$$

$$Recall = \frac{tp}{tp + fn} \times 100 \quad (2)$$

$$F-score = \frac{2 \times precision \times recall}{precision + recall} \quad (3)$$

실험의 목적은 잡음이 제어된 또는 추가되지 않은 깨끗한 데이터와 잡음이 합성된 잡음 데이터로 학습된 CNN이 적응적인 특징 벡터를 생성하고, 그 결과 다양한 잡음 환경에서도 강인한 호흡기 질병 식별 성능을 보장하는 지의 여부를 확인하기 위함이다. 본 연구에서 제안한 CNN 기반의 돼지 호흡기 질병 식별 성능 비교를 위하여, 대표적인 소리 특징인 MFCC와 이진 분류에 탁월한 성능을 보이는 SVM(Support Vector Machine) 분류기를 이용하여 돼지 호흡기 질병을 식별 실험을 수행하였다. SVM은 프로그램 R의 'e1071' 패키지를 사용하였으며, 사용한 커널 함수는 'RBF(radial)', 'Polynomial', 'Sigmoid' 이다. RBF 커널 함수를 이용한 경우가 실험 성능이 상대적으로 좋았으며, 이 때 gamma 값은 0.000059, SVM의 상대적 중요성을 조정하는 상수(trade-off constant) C 값은 6.5로 설정하였다.

본 실험에서는 백색 잡음(SNR0, SNR12, SNR21)으로 합성된 데이터들을 CNN 학습 과정에 추가하여 실험을 수행하였다. 테스트 수행 시에는 잡음이 제어된 데이터, 다양한 환경 잡음 및 백색 잡음을 합성한 데이터들을 모두 개별적으로 적용하여 식별 성능을 확인하였으며, 해당 결과를 그림 6과 표 1~3으로 정리하였다. 그림 6을 살펴보면, 백색 잡음의 강도가 가장 강하게 합성된 SNR0 뿐만 아니라 환경적인 잡음이 합성된 모든 테스트에서의 f-score 성능이 매우 높게 향상되었음을 확인할 수 있다.

표 1. 제안한 학습 구조에 'non-noise' 테스트용 데이터를 적용한 실험 결과 (단위 %)
Table 1. Classification results of porcine wasting diseases with 'non-noise and noise' train and 'non-noise' test data

	MFCC + SVM				Proposed method			
	PMWS	PRRS	MH	Normal	PMWS	PRRS	MH	Normal
Precision	87.9	100	90.9	98.6	100	100	100	100
Recall	96.7	96.4	71.4	100	100	100	100	100
F-score	96.4	98.2	80.0	99.3	100	100	100	100

표 2. 제안한 학습 구조에 'strong footprint' 테스트용 데이터를 적용한 실험 결과 (단위 %)
Table 2. Classification results of porcine wasting diseases with 'non-noise and noise' train and 'strong footprint' test data

	MFCC + SVM				Proposed method			
	PMWS	PRRS	MH	Normal	PMWS	PRRS	MH	Normal
Precision	61.3	81.5	65.0	98.1	94.9	98.3	88.6	97.2
Recall	90.0	78.6	92.9	74.3	100	84.3	100	98.0
F-score	75.0	80.0	76.5	84.6	97.4	90.8	98.0	94.9

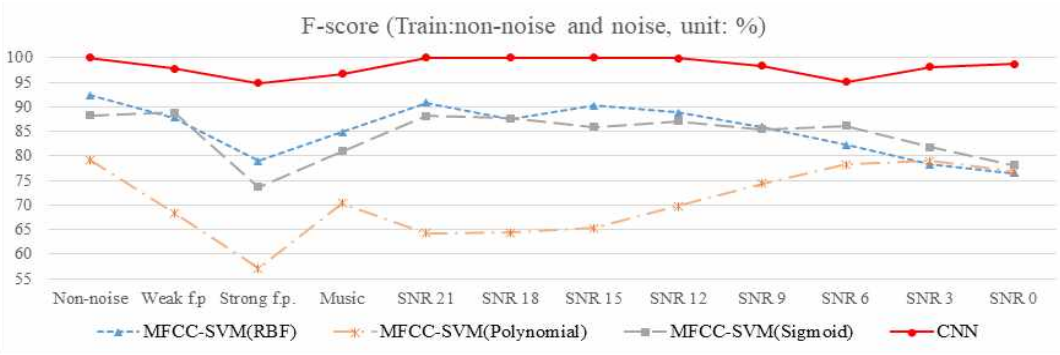


그림 6. 실험 결과 비교 그래프(학습: non-noise, SNR0, SNR12, and SNR12, 테스트: non-noise, 다양한 잡음 합성 데이터 각각 사용)

Fig. 6. Performance comparison graph. (Learning: non-noise, SNR0, SNR12, and SNR21 data, Testing: non-noise data and various noise data)

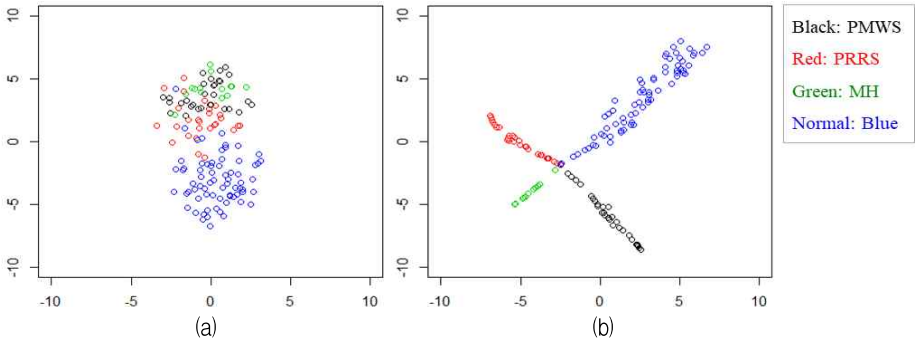


그림 7. t-SNE 시각화. (a) SNR0 합성된 데이터의 MFCC 특징을 시각화한 결과, (b) SNR0 합성된 데이터의 제안한 CNN 구조에 의해 생성된 특징을 시각화한 결과

Fig. 7. t-SNE visualization. (a) t-SNE visualization result of MFCC features with SNR0, (b) t-SNE visualization result of CNN feature with SNR0

표 3. 제안한 학습 구조에 'SNR0' 테스트용 데이터를 적용한 실험 결과 (단위 %)

Table 3. Classification results of porcine wasting diseases with 'non-noise and noise' train and 'SNR0' test data

	MFCC + SVM				Proposed method			
	PMWS	PRRS	MH	Normal	PMWS	PRRS	MH	Normal
Precision	70.3	73.1	72.7	95.6	99.3	98.6	95.9	100
Recall	86.7	67.9	57.1	92.9	98.7	98.6	100	99.4
F-score	77.6	70.4	64.0	94.2	99.0	98.6	97.9	99.7

결국, 추가된 일정 부분의 잡음 데이터로 학습된 CNN이 적응적인 특징 벡터를 자동 생성함으로써, 다양한 잡음 환경에서도 강인한 호흡기 질병 식별 성능을 보장할 수 있음을 실험적으로 확인하였다. 반면, MFCC를 소리 특징 벡터로 이용한 경우에는 백색 잡음이 강한 SNR0~SNR9의 실험 성능이 86% 미만의 f-score값을 보였으며, 학습 데이터에 추가되지 않은 환경적인 잡음인 발자국 소리와 음악 소리가 합성된 데이터를 테스트한 결과의 성능도 매우 낮음을 확인하였다.

본 연구에서 제안한 CNN이 학습을 통해서 적응적인 특징 벡터를 자동 생성함을 확인하기 위하여, 고차원의 벡터를 저차원으로 매핑하는 t-SNE[23] 시각화 기법을 활용하여 그 결과를 그림 7로 나타내었다. 비교를 위해 그림 7(a)는 소리 데이터에 백색 잡음인 SNR0를 합성한 후 MFCC 기법에 의해 추출된 특징 벡터를, 7(b)는 본 논문에서 제안한 학습이 완료된 CNN 구조에 SNR0가 합성된 소리 데이터를 적용한 후의 특징 벡터를 보여준다.

그림 7(a)에 의하면 MFCC는 소리 데이터의 잡음 상황에 대한 추가적인 고려 없이 기존에 정의된 MFCC 추출 방법에 의해 해당 소리의 MFCC 특징 정보가 추출된다. 그 결과 유사한 시그널을 보이는 호흡기 질병들(그림 4 참조)의 경우, MFCC 추출 특징들을 t-SNE 적용 후 유사도 정도를 2차원 그래프 상으로 표현한 그림 7(a)와 같이 PMWS(Black), PRRS(Red), MH (Green) 호흡기 질병들은 같은 영역에 군집으로 표현되어 손쉽게 구분이 어려운 것을 확인할 수 있다. 반면, 잡음 상황 및 각 클래스들이 효과적으로 식별되도록 학습이 수행 된 CNN에 적용되어 취득한 특징 벡터의 경우는 그림 7(b)에서 보는 것과 같이 극히 일부를 제외하고 정상 및 3가지 호흡기 질병 소리가 효과적으로 구분될 수 있도록 군집화된 것을 확인할 수 있다.

또한, 본 연구에서 제안한 CNN 구조에 의해 생성된 완전 연결 계층의 특징 벡터가 4개의 클래스를 구분하기 좋은 특징 벡터로 재구성되었음을 확인하기 위하여, CNN 구조의 첫 번째 계층(Layer 1)과 완전 연결 계층 1에서 취득한 특징 벡터를 행렬 시각화(Matrix Visualization) 기법을 활용하여 비교하였다.

행렬 시각화에 사용한 테스트 데이터는 잡음이 포함되지 않은 데이터, 백색 잡음들 중 가장 강한 SNR0가 합성된 데이터 그리고 환경 잡음들 중 잡음의 강도가 가장 센 돼지의 발자국 소리가 합성된 데이터이며, 이미 학습된 CNN 구조에 테스트 데이터를 적용하였다. 그리고 첫 번째 계층과 완전 연결 계층 1에서 각각 특징 벡터를 취득한 후 이를 행렬 시각화 기법에 적용한 결과를 그림 8과 그림 9로 나타내었다. 이 때 각 행은 하나의 테스트 데이터를 의미하며, 각 클래스별로 일정 수의 데이트 데이터를 사용하였다. 열은 CNN 학습을 거친 계층구조에서 취득한 특징 벡터의 개별 차원에 해당하는 정보이다.

CNN 첫 번째 계층에서 취득한 특징 벡터의 시각화 결과인 그림 8을 보면, 백색잡음 SNR0이 합성된 데이터 (b)의 경우 4개의 클래스(Normal, MH, PRRS, PMWS)를 손쉽게 구분 지을 수 있는 특징 벡터의 특정한 영역이 두드러지지 않는다. 또한 잡음이 포함되지 않은 (a)와 돼지의 매우 시끄러운 발자국 소리가 합성된 소리 인 (c)의 경우에도 클래스를 구분 지을 수 있는 특정한 패턴을 발견하기 어렵다. 즉, CNN의 첫 번째 계층에서는 정상과 호흡기 질병은 구분할 수 있지만 호흡기 질병간의 식별은 어려운 부분이 있다.

이와 다르게 CNN의 완전 연결 계층 1에서 취득한 특징 벡터의 행렬 시각화 결과를 확인해 보면, 잡음을 합성하지 않은 경우 (a), 백색 잡음 SNR0을 합성한 경우 (b) 그리고 매우 시끄러운 돼지의 발자국 소리가 합성된 (c) 결과 모두에서 4개의 클래스를 손쉽게 구분 지을 수 있는 두드러진 패턴들이 일정한 영역(검은색으로 덧그린)에서 나타남을 확인할 수 있다.

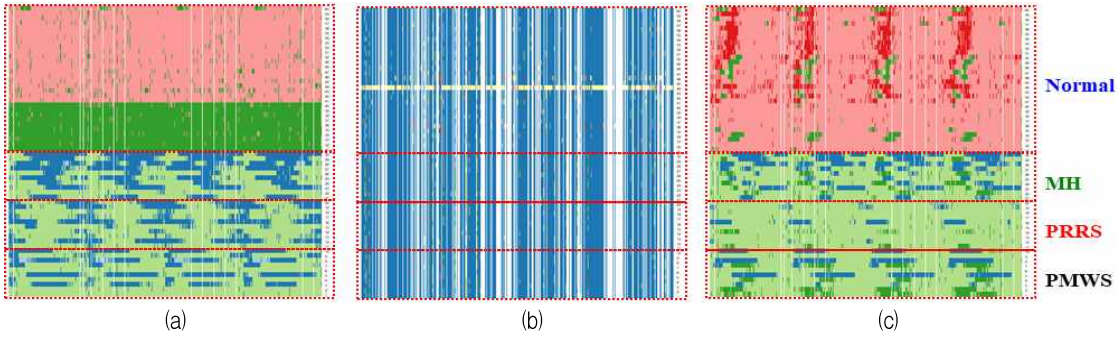


그림 8. CNN 첫 번째 계층의 특징 벡터에 대한 행렬 시각화. (a) non-noise 테스트 데이터의 행렬 시각화 결과, (b) SNR0 테스트 데이터의 행렬 시각화 결과, (c) strong footprint 테스트 데이터의 행렬 시각화 결과

Fig. 8. Matrix visualization of CNN layer 1's feature vector. (a) matrix visualization result of feature vector with non-noise test data, (b) matrix visualization result of feature vector with SNR0 test data, (c) matrix visualization result of feature vector with strong footprint test data

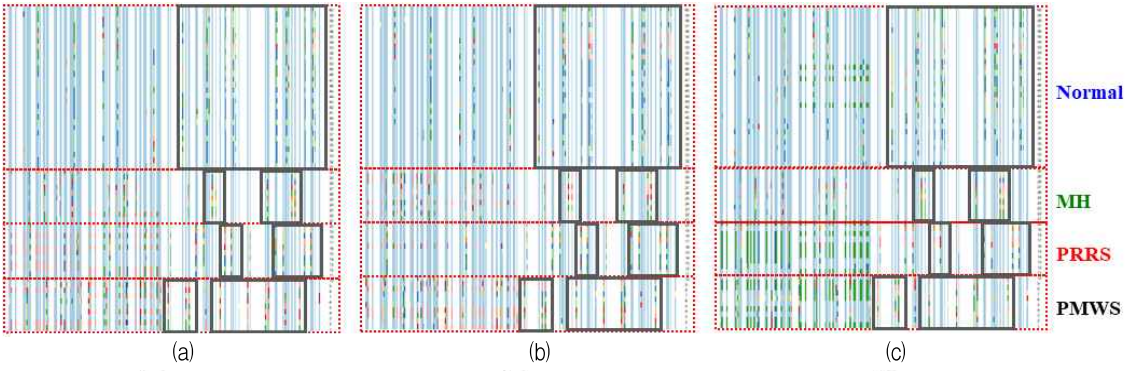


그림 9. CNN fully-connected 1 계층의 특징 벡터에 대한 행렬 시각화. (a) non-noise 테스트 데이터의 행렬 시각화 결과, (b) SNR0 테스트 데이터의 행렬 시각화 결과, (c) strong footprint 테스트 데이터의 행렬 시각화 결과

Fig. 9. Matrix visualization of CNN fully-connected layer 1's feature vector. (a) matrix visualization result of feature vector with non-noise test data, (b) matrix visualization result of feature vector with SNR0 test data, (c) matrix visualization result of feature vector with strong footprint test data

표 4. 기존 연구와의 성능 비교 표

Table 4. Performance comparison with existing methods

	Guarino et al.[12]	Chung et al.[13]	Lee et al.[14]	Proposed method
Feature extraction	Filtering method and AM demodulation	MFCC	Time and frequency sound features	Spectrogram + CNN
Consideration of noise	N/A	N/A	N/A	Yes
Disease detector	DTW	SVDD	SVM	MLP
Disease classifier	N/A	SRC	N/A	MLP
Number of normal data	433	200	449	350×5
Number of disease data	Cough: 159	PMWS: 150 PRRS: 120 MH: 30	PMWS: 146 PRRS: 126 MH: 31	PMWS: 150×5 PRRS: 140×5 MH: 70×5
Disease detection rate	85.5%	94.0%	97.7%	100%
Mean precision	82.2%	90.8%	N/A	99.2%
Mean recall	86.1%	92.0%	N/A	99.3%

이것은 본 연구에서 제안한 CNN 계층을 단계별로 거치면서 다양한 잡음에 강인하며, 각 클래스를 보다 손쉽게 구분 지을 수 있도록 일반화된 특징벡터를 생성하는 것으로 판단할 수 있다. 또한 1번째 계층의 262,144에 달하는 큰 특징 차원에 비해 매우 작은 수치인 완전 연결 계층 1의 288 차원의 특징 벡터만으로도 4개의 클래스를 효과적으로 구분 지을 수 있다는 큰 장점이 있다.

마지막으로, 표 4는 기존에 소리 정보를 이용하여 돼지의 호흡기 질병을 식별한 기존 연구들과 본 연구에서 제안한 기법을 정량적/정성적으로 비교 분석한 내용을 정리하였다.

V. 결론 및 향후 연구

동물의 발성음을 기반으로 호흡기 질병을 탐지하는 최근의 연구동향을 살펴보면, 소리 신호해석에 의한 돼지 호흡기 질병의 탐지가 현실적으로 가능함을 확인하는 학술적 노력들이 일정 부분 성공적인 결과를 도출하였으나, 실제 돈사 환경에서 발생하는 잡음의 영향력을 충분히 고려했다고는 볼 수 없다. 학술적 프로타입 시스템을 실제 돈사환경에서 성공적으로 운용하기 위해서는 소리 취득 시 발생하는 다양한 잡음에서도 강인함이 보장되어야만 한다. 본 논문에서는 소규모의 농장에서도 저비용으로 구축이 가능한 소리 센서 환경에서의 돼지의 호흡

기 질병 탐지 및 식별 시스템을 제안하였다. 특히 본 시스템은 실제 돈사 환경에 설치되어 돼지 호흡기 질병의 탐지 및 식별이 현실적으로 실현 가능할 수 있도록, 실제 돈사 환경에서 발생하는 다양한 잡음에 강인한 시스템을 설계하였다. 제안된 시스템은 학습 데이터의 증폭과 함께 최근 각광을 받고 있는 CNN의 커널 및 구조를 자체 설계하여 소리 신호에서 스펙트로그램 특징을 추출한 후 이를 CNN의 입력 정보로 사용함으로써, 학습 과정을 거치면서 적응적으로 호흡기 질병을 식별하기 좋은 특징 벡터로 자동 변환한 후 정상소리와 호흡기 질병 소리들 중 하나로 식별하였다.

향후 연구 과제로는 본 연구에서 제안한 프로토타입의 시스템을 실세계에서 구현·운용하기 위한 실용화 차원의 후속 연구들이 요구되며, 소리 신호뿐 만이 아니라 영상 신호를 동시에 사용하는 멀티모달 기반의 융합 연구를 후속 연구로 진행할 계획이다.

References

[1] Ministry of Agriculture, Food and Rural Affairs. http://www.mifaff.go.kr [Accessed Oct. 24, 2017]

[2] B. Kim, Y. Lee, Y. Kim, T. Kim, J. Park, and S. Lee et al., "Top 10 Agriculture Issues in 2017", Korea Rural Economic Institute, Focus on Agricultural Affairs, Vol. 142, pp. 1-27, Jan. 2017.

[3] K. S. Chu, M. S. Kang, Y. S. Jo, and J. W. Lee, "Detection of Porcine Circovirus 2, Porcine Reproductive and Respiratory Syndrome Virus and Mycoplasma Hyopneumoniae from Swine Lungs with Lesions by PCR", Korean Journal of Veterinary Service, Vol. 31, No. 1, pp. 71-77, Mar. 2008.

[4] L. Lee, L. Jin, D. Park, and Y. Chung, "Automatic Recognition of Aggressive Behavior in Pigs using a Kinect Depth Sensor", Sensors, Vol. 16, No. 5, 631, Nov. 2016.

[5] M. Ju, H. Baek, J. Sa, H. Kim, Y. Chung, and D. Park, "Real-Time Pig Segmentation for Individual Pig Monitoring in a Weaning Pig Room", Journal of Korea Multimedia Society, Vol. 19, No. 2, pp. 215-223, Feb. 2016.

[6] J. Choi, L. Lee, Y. Chung, and D. Park, "Individual Pig Detection using Fast Region-based Convolution Neural Network", Journal of Korea Multimedia Society, Vol. 20, No. 2, pp. 216-224, Feb. 2017.

[7] J. Choi, L. Lee, Y. Chung, and D. Park, "Individual Pig Detection using Kinect Depth Information", KIPS Transactions on Computer and Communication Systems, Vol. 5, No. 10, pp. 319-326, Aug. 2016.

[8] J. Lee, B. Noh, S. Jang, D. Park, Y. Chung, and H.H. Chang, "Stress Detection and Classification of Laying Hens by Sound Analysis", Asian-Australasian Journal of Animal Sciences, Vol. 28, No. 4, 592, Apr. 2015.

[9] M. Rizwan, B. T. Carroll, D. V. Anderson, W. Daley, S. Harbert, and D. F. Britton et al., "Identifying Rale Sounds in Chickens using Audio Signals for Early Disease Detection in Poultry", Proceeding of Global Conference on Signal and Information Processing, pp. 55-59, Dec. 2016.

[10] Y. Chung, J. Lee, S. Oh, D. Park, H. H. Chang, and S. Kim, "Automatic Detection of Cow's Oestrus in Audio Surveillance System", Asian-Australasian Journal of Animal Sciences, Vol. 26, No. 7, pp. 1030-037, Jul. 2013.

[11] J. Vandermeulen, C. Bahr, D. Johnston, B. Earley, E. Tullo, and I. Fontana et al., "Early Recognition of Bovine Respiratory Disease in Calves using Automated Continuous Monitoring of Cough Sounds", Computers and Electronics in Agriculture, Vol. 129, pp. 15-26, Nov. 2016.

[12] M. Guarino, P. Jans, A. Costa, J. M. Aerts, and D. Berckmans, "Field Test of Algorithm for Automatic Cough Detection in Pig Houses", Computers and Electronics in Agriculture, Vol. 62, No. 1, pp. 22-28, Feb. 2008.

[13] Y. Chung, S. Oh, J. Lee, D. Park, H. H. Chang, and S. Kim, "Automatic Detection and Recognition

of Pig Wasting Diseases using Sound Data in Audio Surveillance", Sensors, Vol. 13, No. 10, pp. 12929-12942, Sep. 2013.

[14] J. Lee, L. Jin, D. Park, Y. Chung, and H. H. Chang, "Acoustic Features for Pig Wasting Disease Detection", International Journal of Information Processing and Management, Vol. 6, No. 1, pp. 37-46, Apr. 2015.

[15] J. Salamon and J. P. Bello, "Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification", Signal Processing Letters, Vol. 24, No. 3, pp. 279-283, Jan, 2017.

[16] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, "A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals", Sensors, Vol. 17, No. 2, 425, Feb. 2017.

[17] T. L. Nwe, H. D. Tran, W. Z. T. Ng, and B. Ma, "An Integrated Solution for Snoring Sound Classification Using Bhattacharyya Distance based GMM Supervectors with SVM, Feature Selection with Random Forest and Spectrogram with CNN", Processing of Interspeech, pp. 3467-3471, Aug. 2017.

[18] A. Krizhevsky, I., Sutskever, and G. E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems, pp. 1097-1105, Dec. 2012.

[19] I. S. Kim and S. Y. Hwang, "A Effective Online Training Algorithm by Partitioning Bounding Box for Visual Object Tracking using Convolutional Neural Network", The Journal of Korean Institute of Communications and Information Sciences, Vol. 42, No. 6, pp. 1117-1128, Jul. 2017.

[20] K. H. Kong and D. S. Kang, "A Study of Face Detection Algorithm Using CNN with Mixed-LGP and Hippocampus Structure", Journal of KIIT, Vol. 16, No. 1, pp. 11-17, Jan. 2018.

[21] K. K. Yeo and D. S. Kang, "CNN-based Alzheimer's Disease Image Learning and Accuracy Measurement Using ^{18}F-Florbetaben PET Image Editted Database", Journal of KIIT, Vol. 15, No. 10, pp. 1-8, Oct. 2017.

[22] TensorFlow ver.1.02. http://www.tensorflow.org [Accessed: Apr. 09, 2017]

[23] L.V.D. Maaten and G. Hinton, "Visualizing Data using t-SNE", Journal of Machine Learning Research, Vol. 9, pp. 2579-2605, Nov. 2008.

저자소개

이 종 욱 (Jonguk Lee)

2002년 : 고려대학교 전산학과 (학사)
2005년 : 고려대학교 전산학과 (석사)
2014년 : 고려대학교 전산학과 (박사)
2014년 ~ 현재 : 고려대학교 컴퓨터 융합소프트웨어학과 초빙교수
관심분야 : 딥러닝, 기계학습, 멀티미디어 마이닝, 융합 IT

최 용 주 (Yongju Choi)

2007년 : 고려대학교 컴퓨터정보학과(학사)
2017년 ~ 현재 : 고려대학교 컴퓨터정보학과 석사과정
관심분야 : 기계학습, 딥러닝, 데이터마이닝, 인공지능, 융합IT

박 대 희 (Daihee Park)

1982년 : 고려대학교 수학과(학사)
1984년 : 고려대학교 수학과(석사)
1989년 : 플로리다 주립대학 전산학과(석사)
1992년 : 플로리다 주립대학 전산학과(박사)
1993년 ~ 현재 : 고려대학교 컴퓨터융합소프트웨어학과교수
관심분야 : 데이터마이닝, 인공지능, 축산 IT

정 용 화 (Yongwha Chung)

1984년 : 한양대학교
전자통신공학과(학사)
1986년 : 한양대학교
전자통신공학과(석사)
1997년 : 남가주대학 전기공학과
(박사)
1986년 ~ 2003년 : 한국전자통신
연구원 생체인식기술연구팀(팀장)
2003년 ~ 현재 : 고려대학교 컴터융합소프트웨어학과(교수)
관심분야 : 영상처리, 병렬처리, 축산 IT