

Damaged Golf Ball Detection through CNN-based Impact Sound Classification

Seokho Chung*, Jinho Cho*, Woonho Lee*, Sangyul Lee**

*Student, Division of Mechanical and Electronic Engineering, Hansung University, Seoul, Korea

**Professor, Division of Mechanical and Electronic Engineering, Hansung University, Seoul, Korea

[Abstract]

In the lost golf ball market, damaged balls are often mixed into distribution, leading to quality issues. To address this, we developed a CNN-based acoustic classification system that detects damaged balls by analyzing impact sounds. A hardware device with triple-layered soundproofing enabled high-quality audio data collection. The signals were converted into Mel-Spectrograms and classified using EfficientNet-based CNN, with data augmentation enhancing generalization. Model optimization (46.9 MB→4.7 MB) enabled real-time classification in embedded environments. Field tests achieved 96.5 % accuracy, and an automated system integrating Jetson Nano and Arduino significantly improved sorting efficiency. This system enhances efficiency and product reliability in lost ball sorting, offering environmental and economic benefits.

▶ **Key words:** Deep Learning, Acoustic Classification, Mel-Spectrogram, CNN, Golf Ball Sorting

[요 약]

골프공 로스트볼 시장에서는 손상된 공이 혼입되어 유통되는 문제로 품질 불만이 지속되고 있다. 본 연구에서는 골프공 낙하 시 발생하는 충격음을 CNN 기반 음향 분류 기법으로 분석하여 정상 공과 손상 공을 자동 검출하는 시스템을 개발하였다. 삼중 방음 구조와 적정 낙하 속도 구현 하드웨어를 통해 고품질 음향 데이터를 수집하였으며, Mel-Spectrogram과 EfficientNet을 활용한 이진 분류를 수행하였다. 데이터 증강 기법 적용으로 모델 일반화 성능을 향상시켰고, 모델 경량화(46.9 MB→4.7 MB)를 통해 임베디드 환경에서도 실시간 분류가 가능함을 확인하였다. 실제 환경 실험에서 96.5 %의 정확도를 달성하였으며, Jetson Nano와 Arduino 기반 자동화 시스템으로 대량 분류 효율성을 크게 개선하였다. 본 시스템은 로스트볼 선별 작업의 효율성과 품질 보증을 통해 환경적·경제적 이점을 제공한다.

▶ **주제어:** 딥러닝, 음향 분류, 멜 스펙트로그램, 합성곱 신경망, 골프공 분류

-
- First Author: Seokho Chung, Corresponding Author: Sangyul Lee
 - *Seokho Chung (kdgb01@naver.com), Division of Mechanical and Electronic Engineering, Hansung University
 - *Jinho Cho (jinhoya130@gmail.com), Division of Mechanical and Electronic Engineering, Hansung University
 - *Woonho Lee (sfv678@gmail.com), Division of Mechanical and Electronic Engineering, Hansung University
 - **Sangyul Lee (sangyul.lee@hansung.ac.kr), Division of Mechanical and Electronic Engineering, Hansung University
 - Received: 2025. 10. 01, Revised: 2025. 12. 15, Accepted: 2025. 12. 15.

I. Introduction

매년 전 세계적으로 최대 50억 개에 달하는 골프공이 분실된다[1]. 하지만 이들 대부분은 자연적으로 분해되는데 굉장히 많은 시간이 소요되는 플라스틱과 합성고무로 제작되어, 수백 년간 환경에 계속 남아있게 된다. 일반적으로 분실된 골프공은 수거 후 소각을 통해 처리되지만, 이러한 과정에서 처리 비용이 발생하게 될 뿐만 아니라 대량 소각 시 유해 물질이 방출되어 환경적인 문제가 지속적으로 발생한다. 이러한 문제를 완화하기 위한 대안으로 미국의 언론사 CNN은 매년 전 세계적으로 수십억 개의 골프공이 분실되어 해양 및 자연환경을 오염시키고 있으며, 분해되는 데 100년에서 1,000년이 소요된다는 점을 보도하며 로스트볼 환경오염 문제의 심각성을 지적하였다[1].

그러나 현재 로스트볼 시장에서는 파손되거나 사용이 불가능한 공이 유입된 제품이 유통되는 사례가 종종 발생하고 있다. 이에 따른 소비자들의 불만이 꾸준히 제기되고 있으며, 이러한 문제는 주로 작업자의 육안을 통한 검사를 기반으로 한 기존 선별 방식의 한계에서 비롯된다. 육안 검사는 시간이 많이 소요되고 작업자의 피로도에 따라 정확도가 변동되며, 특히 내부 균열이나 구조적 손상과 같이 외관상 확인이 어려운 결함을 검출하기 어렵다는 근본적인 한계가 있다. 따라서 효율성과 정확도 측면에서 개선이 이루어진 자동화된 선별 시스템이 요구된다.

골프공의 품질 평가는 현재 주로 USGA와 R&A에서 정한 표준화된 물리적 테스트에 의존하고 있다[2]. 제조 품질 관리는 압축 테스트, 변형 측정, 표면 마감 분석, X-ray 검사 등 전통적인 방법을 사용하지만, 이러한 접근법은 시간 집약적이며 대량 선별 작업에 적용하기 어렵다는 한계가 있다.

최근 딥러닝 기술의 발전으로 음향 신호 분석을 통한 비파괴 품질 검사가 제조업 분야에서 주목받고 있다. 특히 CNN(Convolutional Neural Network) 기반 음향 분류 기술은 높은 정확도와 실시간 처리 능력으로 산업 현장에 적용되고 있다. 그러나 골프공 품질 검사에 음향 분석을 적용한 연구는 현재까지 보고된 바 없다.

본 연구에서는 골프공이 낙하할 때 발생하는 충격음을 CNN 기반 딥러닝 모델로 분석하여 정상 공과 손상 공을 자동으로 분류하는 시스템을 제안한다. 한국골프협회(KGA) 대회 규정을 참고하여 균열이나 형태 변형, 내부 노출과 같이 물리적으로 심각한 손상이 있는 공을 '손상 공'으로 분류하고, 그 외의 경우를 '정상 공'으로 분류하였다. 본 시스템은 삼중 방음 구조의 하드웨어 장치와 Jetson Nano 기

반 임베디드 시스템을 통합하여 실제 로스트볼 선별 작업에 적용 가능한 완전한 자동화 솔루션을 제시한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구를 검토한다. 3장에서는 하드웨어 설계와 소프트웨어 아키텍처를 포함한 전체 시스템 설계에 대해 기술한다. 4장에서는 데이터 수집 및 증강 기법과 CNN 기반 딥러닝 모델 설계 과정을 상세히 설명한다. 5장에서는 실험 환경과 성능 평가 결과를 제시하고, 6장에서는 결론 및 향후 연구 방향을 논의한다.

II. Related Works

1. Deep Learning for Audio Classification

딥러닝 기술의 발전으로 음향 분류 분야에서 CNN의 효과성이 입증되었다. Hershey 등[3]은 대규모 AudioSet 데이터셋에서 다양한 CNN 아키텍처를 Mel-Spectrogram과 결합하여 음향 분류의 기초를 확립하였다. Mel-Spectrogram은 인간의 청각 민감도를 근사하는 지각적 스케일로 변환된 주파수 표현으로, 음향 분류 작업에서 효과적인 특징으로 자리잡았다[4]. Piczak[5]는 Mel-Spectrogram을 2D 이미지로 처리하는 CNN 접근법이 환경음 분류에서 전통적인 기계 학습 기법을 능가함을 입증하였다.

제한된 학습 데이터의 한계를 극복하기 위한 데이터 증강 기법도 활발히 연구되고 있다. Salamon과 Bello[6]는 환경음 분류에서 피치 변환, 시간 확장 등의 데이터 증강과 심층 CNN을 결합하여 제한된 데이터로도 우수한 분류 성능을 달성하였다. Park 등[7]이 제안한 SpecAugment는 스펙트로그램 도메인에서 직접 시간-주파수 마스킹을 적용하여 추가 음향 데이터 생성 없이도 과적합을 효과적으로 방지하는 기법으로, 음성 인식과 음향 분류 분야에서 표준 기법으로 자리잡았다.

2. Acoustic-based Industrial Quality Inspection

음향 신호 기반 품질 검사는 제조업 분야에서 비파괴적 결함 탐지 방법으로 널리 활용되고 있다. Lee 등[8]은 환경 소음이 있는 실제 공장 환경에서 변분 오토인코더 기반 비지도 학습 방법으로 장비 고장을 진단하였으며, Sikdar 등[9]은 음향 방출(Acoustic Emission) 신호에 CNN을 적용하여 복합재료 패널의 손상 유형을 95 % 이상의 정확도로 분류하였다.

적층 제조 분야에서도 음향 기반 품질 검사가 활발히 연구되고 있다. Shevchik 등[10]은 적층 제조 공정에서 실

시간으로 수집한 음향 방출 신호와 CNN을 결합하여 78-91 %의 품질 분류 정확도를 달성하였다. Zhang 등 [11]은 콘크리트 구조물 검사에서 Impact-Echo 방법과 Extreme Learning Machine을 결합하여 100%의 결합 검출률과 0.87 %의 낮은 오검출률을 달성하였으며, 이는 충격음 분석이 재료 내부의 결함 검출에 매우 효과적임을 보여준다.

최근에는 더욱 진보된 딥러닝 아키텍처를 활용한 연구가 진행되고 있다. Ji 등[12]은 용접 품질 검사에 음향 신호와 Squeeze-and-Excitation CNN-LSTM 하이브리드 모델을 적용하여 91 %의 분류 정확도를 달성하였으며, Zhang 등[13]은 Wavelet 시간-주파수 변환과 MobileNetV3를 결합하여 적층 제조 공정의 결함을 98.31 %의 정확도로 실시간 분류하였다.

3. Spherical Object Quality Inspection

구형 물체의 품질 검사는 곡면 전체를 균일하게 검사해야 하는 기술적 난이도가 있다. 골프공 품질 검사와 직접적으로 관련된 연구로, Wu[14]는 AOI (Automated Optical Inspection) 시스템을 개발하여 딥لم 3형상을 고려한 특수 조명 설계를 통해 97.3 % 이상의 불량 검출률과 10 % 미만의 오탐률을 달성하였다. 그러나 이 방법은 표면 결함에만 국한되며 내부 손상 검출은 불가능하다는 한계가 있다.

유사한 구형 물체인 정밀 베어링 강구 검사에서 Wu와 Huang[15]은 DarkNet19 CNN과 수직 회전 메커니즘을 결합하여 전체 구면을 12장의 이미지로 전개하고 99.5 %의 표면 결함 분류 정확도를 달성하였다. 이는 구형 물체 검사에 CNN이 효과적임을 입증하지만, 영상 기반 방법의 특성상 내부 구조적 손상은 검출할 수 없다.

4. Lightweight CNN Architectures for Embedded Systems

임베디드 환경에서의 실시간 추론을 위해서는 경량화된 CNN 아키텍처가 필수적이다. Tan과 Le[16]가 개발한 EfficientNet은 깊이, 너비, 해상도를 복합적으로 스케일링하는 방법을 통해 기존 ConvNet보다 8.4배 작고 6.1배 빠르면서도 동등하거나 더 높은 정확도를 달성하였다. 특히 EfficientNet-B0은 5.3 M 파라미터로 가벼우면서도 우수한 성능을 보여 자원 제약이 있는 환경에 적합하다.

Berg 등[17]은 EfficientNet-B0을 제한된 학습 데이터 환경에서 미세 조정하여 복잡한 분류 작업에서도 95.95 %의 정확도를 달성하였으며, 이는 결함 샘플이 제한적인 산

업 현장에서 EfficientNet의 실용성을 뒷받침한다. 모델 경량화를 위한 양자화 기법도 중요한 기술이다. INT8 양자화는 FP32 대비 4배의 모델 크기 감소를 제공하며, 적절한 교정(calibration)을 통해 1 % 미만의 정확도 손실로 배포가 가능하다[18]. 이러한 최적화 기술은 NVIDIA Jetson Nano와 같은 엣지 디바이스에서 실시간 딥러닝 추론을 가능하게 한다.

5. Research Gap and Contributions of This Work

스포츠 장비의 자동화된 품질 검사에 관한 학술 연구는 극히 제한적이다. 광범위한 문헌 검토 결과, 골프공 검사 관련 연구는 광학 기반 표면 검사에 국한되어 있으며[14], 음향 신호를 활용한 골프공 결함 검출 연구는 전무한 것으로 확인되었다.

본 연구는 다음과 같은 측면에서 기존 연구와 차별화된다. 첫째, 음향 기반 골프공 결함 검출에 관한 최초의 학술 연구로서, 산업 분야의 음향 결함 검출 기술과 CNN 기반 오디오 분류 기술을 골프공 품질 검사에 통합한다. 둘째, 광학 검사로는 불가능한 내부 손상(균열, 내부 노출)을 충격음 분석을 통해 검출할 수 있는 새로운 접근법을 제시한다. 셋째, Mel-Spectrogram, EfficientNet, SpecAugment 등 검증된 최신 딥러닝 방법론을 골프공 음향 분석에 최초로 적용하여 높은 분류 정확도를 달성한다. 넷째, 삼중 방음 구조와 제어된 낙하 속도를 구현한 하드웨어 시스템과 Jetson Nano 기반 임베디드 추론 시스템을 통합하여 실제 로스트 볼 선별 작업에 즉시 적용 가능한 완전한 엔드-투-엔드 솔루션을 제시한다. 본 연구는 확립된 개별 기술들을 창의적으로 통합하여 연간 최대 50억 개가 분실되는 골프공의 효율적인 재활용과 품질 보증을 가능하게 하는 실용적 기여를 제공한다.

III. System Design

1. Hardware Design

본 연구에서는 골프공 자동 분류기를 제작하기에 앞서 전반적인 형상을 도식화하여(Fig. 1) 설계 방향을 수립하였다. 장치는 휴대성과 현장 적용성을 고려하여 승용차 트렁크에 적재 가능한 높이와 치수로 설계되었으며, 직육면체 프레임 내부에 주요 부품을 배치하였다(Fig. 2). 하드웨어는 크게 투입부, 가속부, 짐삼 및 분류부, 본체 구조로 구분된다.

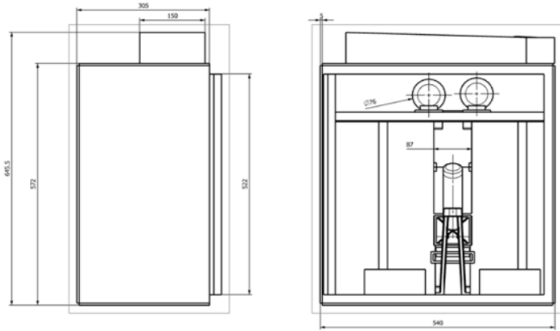


Fig. 1. Device design schematic

골프공 투입부는 다량의 골프공을 일괄 투입한 뒤 낙하 방향으로 이동시킬 수 있도록 경사 구조로 설계되었다. 약 60개 이상의 골프공을 동시에 적재할 수 있으며, 리니어 솔레노이드와 리미트 스위치를 이용해 공을 하나씩 낙하시킬 수 있도록 제어하였다(Fig. 3). 리니어 솔레노이드는 Arduino 보드에서 제어되는 신호를 기반으로 동작하며, Jetson Nano와 연동되어 공을 낙하부로 순차적으로 공급한다. 체결부는 3D 프린터로 제작된 맞춤형 부품을 사용하여 구조적 안정성을 확보하였다.

골프공 낙하 시 충격음 데이터의 정확도를 확보하기 위해 자유낙하 높이를 단계적으로 변화시키며 실험을 진행하였다. 실험 결과, 40 cm 낙하에서는 검증 정확도가 약 69.17%에 그쳤으나, 60 cm 낙하에서는 89.92%로 크게 향상되었다. 따라서 학습 모델의 성능을 높이기 위해 최소 60 cm에서의 자유낙하 속도가 필요함을 확인하였다. 낙하 속도는 식 (1)로 계산하였다.

$$v = \sqrt{v_0^2 + 2gh} \tag{1}$$

추가 속도를 구현하기 위해 DC775 모터와 실리콘 휠을 양측에 배치하여 낙하 공의 속도를 가속시키는 구조를 설계하였다(Fig. 4). 모터는 약 0.2 N·m 이상의 정격 토크를 제공할 수 있으며, 스러스트 베어링을 추가하여 마찰 부하

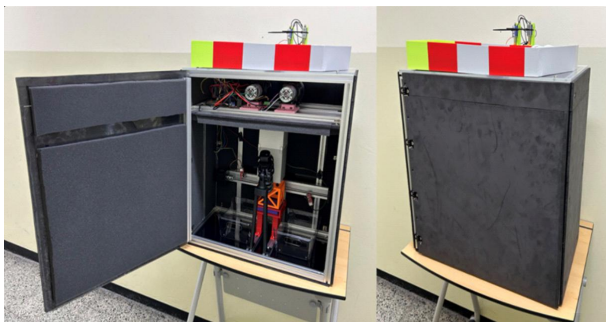


Fig. 2. Golf ball classifier open and closed configuration

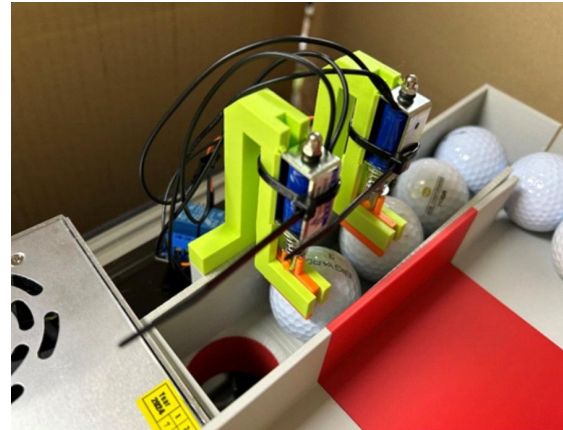


Fig. 3. Golf ball input unit

를 줄였다. 이를 통해 일정한 낙하 속도를 보장하여 충격음의 일관성을 확보하였다. 휠의 직경 3 inch에 모터 속도 2350 rpm을 적용하여 약 9.4 m/s의 초기속도를 갖게 하였다. 자유낙하는 3.4 m/s 이므로 약 3배의 속도 상승 효과가 발생했다.

집음부는 공이 일정 위치에 낙하하도록 유도하는 통로와 충격음을 효율적으로 수집할 수 있는 밀도 100% PLA 소재 판으로 구성하였다. 낙하음은 지향성 콘덴서 마이크를 통해 수집되며, 수집된 신호는 딥러닝 모델로 전달된다.

분류부는 리니어 레일 슬라이드와 스텝모터를 이용하여 설계하였다(Fig. 5). 딥러닝 모델의 판별 결과에 따라 슬라이드가 이동하며 골프공을 정상 공 및 손상 공 바구니로 자동 분류한다. 전원 공급은 12 V, 50 A 파워 서플라이(SMPS)를 사용하였으며, 14 AWG 전선을 통해 충분한 전류를 공급하였다. 모든 모터 및 센서는 Arduino Mega 보드에서 제어되며, Jetson Nano와 통신하여 실시간 분류가 이루어지도록 하였다.

고품질 음향 데이터 확보를 위해 본체를 밀폐형 삼중 방음 구조로 설계하였다(Fig. 6). 외부에는 두께 5 mm의 고무 차음재를 적용하여 외부 소음 유입을 최소화하였고, 중

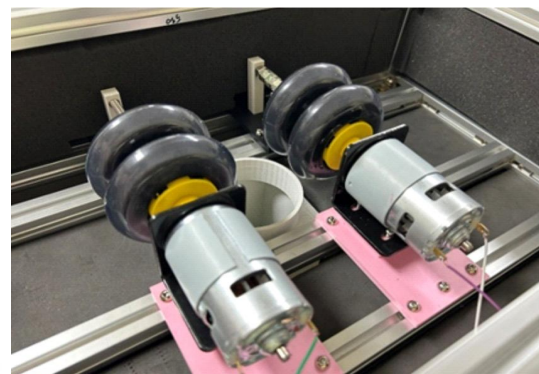


Fig. 4. Golf ball acceleration unit

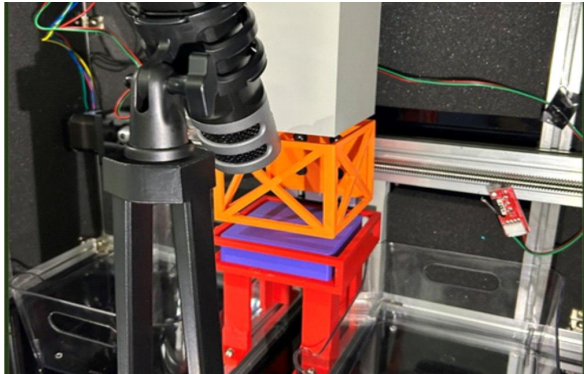


Fig. 5. Golf ball collection and classification unit

간층에는 아크릴 패널을 배치하여 완충재 역할을 수행하였다. 내부에는 20 mm 두께의 흡음 스펀지를 부착하여 반향음과 잔향을 흡수하였다. 이러한 삼중 방음 구조는 외부 소음과 내부 잡음을 효과적으로 차단하여 충격음 신호의 명확성을 확보하였으며, 결과적으로 딥러닝 모델 학습 및 분류 성능 향상에 기여하였다.

2. Software Architecture

본 연구에서는 음향 신호 기반으로 골프공의 손상 여부를 분류하기 위한 자동화된 검사 시스템을 설계하였다. 전체 시스템은 물리적 구동 장치(Arduino Mega 2560), 신호 처리 및 제어용 엷지 디바이스(NVIDIA Jetson Nano), 그리고 이를 연동하는 통신 프로토콜로 구성된다(Fig. 7).

2.1 System Operation Sequence

골프공 분류 기기의 분류 프로세스는 다음과 같다(Fig. 8). 첫째, Arduino Mega 보드는 투입부의 리니어 솔레노이드를 제어하여 골프공을 낙하시킨다. 그 후 즉각적으로 Jetson Nano로 낙하 발생 신호를 전송한다. 발생 신호는 이후 데이터 수집 및 예측 절차의 시작을 알리는 동기 신호로 사용된다.

둘째, Jetson Nano에 연결된 오디오 모듈은 수신된 트리거 신호에 반응하여 2초간의 낙하음을 음성 파일 형식으로 녹음하고, 이를 이미지 기반 학습 모델인 CNN 모델의 입력에 적합하도록 음향을 시각화한 스펙트로그램 이미지 형태로 변환한다.

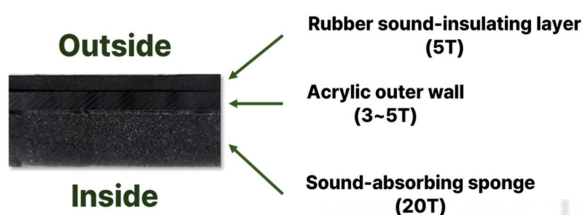


Fig. 6. Triple soundproofing structure

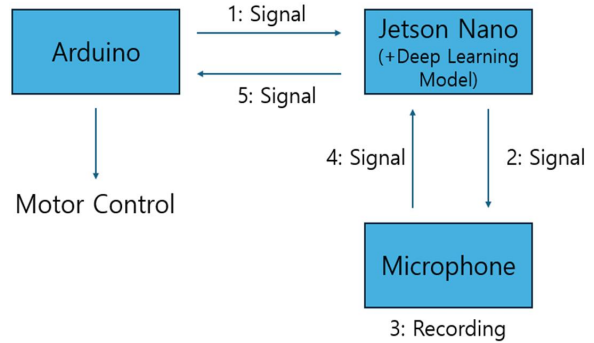


Fig. 7. Software Architecture

셋째, 변환된 데이터는 Jetson Nano에 사전 탑재된 딥러닝 기반 분류 모델의 입력으로 전달되며, 모델은 이를 분석해 해당 골프공의 상태를 정상(Normal: 1) 또는 손상(Broken: 0)으로 이진 분류하여 출력한다. 마지막으로, 예측 결과는 다시 Arduino Mega 보드로 전송되며, 아두이노 측에서는 이 정보를 바탕으로 골프공을 정상/손상부로 이동하도록 제어한다.

2.2 Hardware-Software Integration

Arduino Mega 2560은 하드웨어 제어를 담당하며, 리니어 솔레노이드, 스텝모터, 리미트 스위치 등의 물리적 구성요소를 제어한다. 동시에 Jetson Nano와의 시리얼 통신을 통해 골프공 낙하 시점을 알리고 분류 결과를 수신한다.

NVIDIA Jetson Nano는 저전력으로 동작하면서도 GPU 가속 추론을 지원하여, 본 시스템의 휴대성과 현장 적용성 요구에 부합하는 엷지 컴퓨팅 플랫폼이다. Jetson Nano는 오디오 입력 모듈을 통해 충격음을 수집하고, 사전 훈련된 CNN 모델을 사용하여 실시간 분류를 수행한다. 또한 TensorFlow Lite를 활용한 모델 최적화를 통해 제한된 연산 자원에서도 효율적인 추론이 가능하도록 구현하였다.

실시간 처리를 위해 멀티스레딩 구조를 채택하였다. 메인 스레드는 Arduino와의 통신을 담당하고, 오디오 처리 스레드는 충격음 녹음과 전처리를 수행하며, 추론 스레드는 CNN 모델을 사용한 분류 작업을 처리한다. 이러한 병렬 처리 구조를 통해 전체 분류 과정이 3초 이내에 완료되도록 최적화하였다.

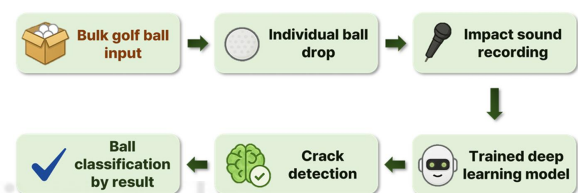


Fig. 8. Overall process flow

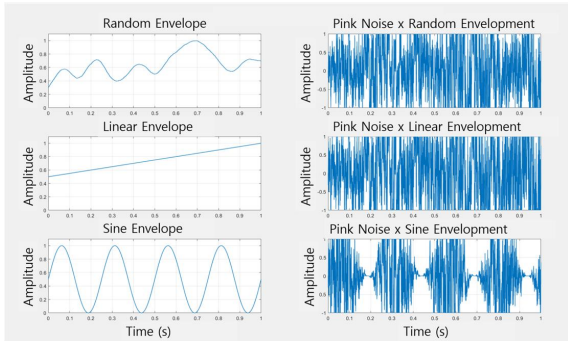


Fig. 9. Dynamic noise injection examples: noise signals modulated by random, linear, and sinusoidal envelopes

IV. Deep Learning Model Development

1. Data Collection and Augmentation

1.1 Acoustic Data Collection

본 연구에서는 스크린 골프장에서 사용 후 폐기 예정인 골프공을 수집하였다. 다양한 브랜드와 피스 구조의 골프공 100개를 한국골프협회(KGA) 경기 규정을 참고하여 정상 공 50개와 손상 공 50개로 분류하였다. 손상 공은 균열, 변형, 내부 노출 등 사용 불가능한 손상이 있는 공으로 정의하였다.

각 골프공을 2장의 하드웨어로 60 cm 높이에서 6회씩 낙하시켜 총 600개의 원본 충격음을 수집하였다(정상 300개, 손상 300개). 각 샘플은 44.1 kHz 샘플링 레이트로 2 초 동안 녹음되었으며, WAV 형식으로 저장되었다.

1.2 Data augmentation techniques

원본 600개 샘플만으로는 딥러닝 모델 학습에 충분하지 않으므로, 다양한 데이터 증강 기법을 적용하였다. 각 원본 샘플에 대해 피치 변형, 동적 노이즈 삽입, 밴드패스 필터링 등을 무작위로 조합하여 증강 샘플을 생성하여, 최종적으로 정상 공 6,000개, 손상 공 6,000개, 총 12,000개의 학습 데이터셋을 구축하였다.

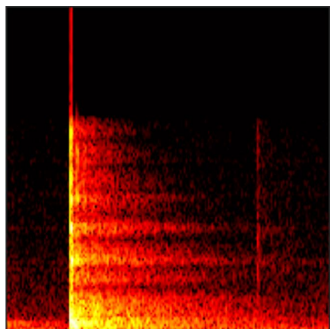


Fig. 10. Decibel threshold-based spectrogram

○ 피치 변형 (Pitch Shifting)

입력 오디오의 주파수 성분을 $\pm 2\sim 6$ 세미톤 범위에서 상하로 이동시키는 방식으로 구현되었다. 세미톤 변화 값은 사전에 정의한 확률 분포를 기반으로 무작위로 선택된다. 예를 들어 ± 2 세미톤은 상대적으로 높은 확률로, ± 6 세미톤은 낮은 확률로 선택되어 과도한 왜곡은 방지하면서도 변이 다양성을 확보하였다.

○ 동적 노이즈 삽입

다양한 스펙트럼 특성을 지닌 노이즈(핑크, 브라운, 블루, 임펄스 노이즈 등)를 주어진 음원에 가중하여 삽입하였다(Fig. 9). 각 노이즈는 시간 축 상에서 랜덤, 선형, 사인 형태의 다양한 스케일링 엔벨로프를 적용함으로써 시간적으로 가변적인 강도의 잡음을 생성하였다. 노이즈의 세기는 무작위로 조정되며, 이는 학습 과정에서의 잡음 내성 확보에 기여한다.

○ 밴드패스 필터링 (Bandpass Filtering)

Butterworth 필터를 적용하여 충격음 중 특정 주파수 범위에 해당하는 정보를 강조하거나 제거하였다. 필터 범위는 실험적으로 효과가 입증된 85 Hz~3400 Hz 내에서 무작위로 설정되었다.

○ 데시벨 임계값 기반 스펙트로그램 생성

낙하음 신호의 명확한 특징을 추출하고 배경 잡음의 영향을 최소화하기 위한 전처리 방법으로 적용하였다(Fig. 10). 스펙트로그램 상에서 음압이 특정 임계값보다 낮은 부분을 0으로 마스킹 처리하여 배경 노이즈 및 비정형적 신호의 영향을 제거하고 낙하음 중심의 고에너지 주파수 성분만을 부각시켰다.

Fig. 11은 원본 데이터와 서술된 방법을 적용해 증강된 데이터의 Mel-Spectrogram을 비교한 것으로, 증강 기법 적용 후 시간-주파수 특성이 다양하게 변화됨을 확인할 수 있다.

1.3 Effect Analysis of Data Augmentation Techniques

본 연구에서 적용한 데이터 증강 기법의 효과를 검증하기 위해 각 기법별 실험을 수행하였다.

피치 변형 실험에서 ± 10 세미톤과 같이 과도한 변형은 실제 충격음 특성과 과도하게 괴리되어 성능 저하를 유발하였다. 반면, ± 6 세미톤 범위에서 생성된 데이터는 실제 낙하음 특성과의 유사성을 유지하면서도 충분한 다양성을 확보할 수 있었다.

노이즈 삽입 실험에서 초기에 고정 세기의 가우시안 노이즈를 사용하였으나, 모델이 일정한 노이즈 패턴에 과도하게 적응하여 과적합 현상이 발생하였다. 동적 노이즈 삽

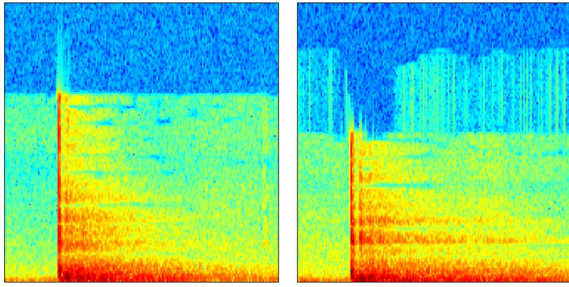


Fig. 11. Original (L) and Augmented Data (R)

입 기법을 도입하여 사인, 선형, 랜덤 형태의 엔벨로프를 적용한 결과, 실제 소음 환경에서의 예측 성능이 약 2.4 % 향상되었다.

밴드패스 필터링 실험을 통해 85-3400 Hz 구간이 가장 이상적인 대역으로 확인되었으며, 저주파 기계음이나 고주파 잡음이 효과적으로 제거되었다. 해당 필터를 적용한 경우 검증 정확도가 평균 1.6 % 향상되었으며, 특히 손상 공판별 민감도가 개선되었다.

데시벨 임계값 기반 스펙트로그램 생성은 배경 잡음을 효과적으로 제거하고 충격음 중심의 고에너지 성분을 강조함으로써 평균 정확도가 약 2 % 향상되었으며, 합성곱 신경망의 수렴 속도도 개선되는 효과가 있었다.

2. CNN Model Design using Mel-Spectrogram

2.1 Mel-Spectrogram feature extraction

골프공의 상태(정상/손상)에 따라 분류할 때 오디오 신호의 Mel-Spectrogram을 입력받아 예측을 진행하기 위하여 CNN을 활용하였다. CNN은 이미지 데이터를 처리하는 데 특화된 딥러닝 알고리즘으로 이미지 데이터의 특징을 추출하는 합성곱 연산을 기본으로 하며, 여러 필터를 통해 입력 데이터에서 패턴이나 특징을 학습한다.

Mel-Spectrogram은 시간 축과 주파수 축의 2차원 구조를 가지는데, CNN은 이러한 시간-주파수 특성에서 패턴을 추출하는데 매우 적합하다. 수집된 음향 데이터는 Librosa 라이브러리를 사용하여 Mel-Spectrogram으로 변환되었으며, 변환 파라미터는 다음과 같이 설정하였다: 샘플링 레이트 44,100 Hz, FFT 윈도우 크기 2048, 홉 길이가 512, Mel 필터뱅크 128개. 생성된 Mel-Spectrogram은 224x224 픽셀 크기로 정규화되어 CNN 모델의 입력으로 사용되었다(Fig. 11).

2.2 Model Architecture Selection

본 연구에서는 임베디드 환경에서의 실시간 추론을 위해 최신 CNN 아키텍처 3종을 동일 조건에서 비교 평가하였다. 비교 대상 모델로는 복합 스케일링 기법을 적용한

Table 1. Performance comparison across all models.

Models	Resnet50	EfficientNet	MobileNetV2
Test Acc.	99.98 %	99.96 %	99.88 %
Test Std.	±0.027 %	±0.042 %	±0.126 %

EfficientNet-B0[16], 깊은 잔차 연결 구조의 ResNet50[19], 그리고 경량화를 위해 설계된 MobileNetV2[20]를 선정하였다. 모든 모델은 ImageNet 데이터셋으로 사전 학습된 가중치를 사용하였으며, 본 연구의 골프공 분류 작업에 맞게 전이학습을 수행하였다.

공정한 비교를 위해 12,000개의 전체 데이터셋을 계층적 샘플링을 통해 학습+검증용 10,000개와 테스트용 2,000개로 분할하였다. 학습+검증 데이터에 대해서는 5-fold 교차검증을 적용하여 각 fold당 학습 8,000개, 검증 2,000개로 구성하였으며, 테스트 세트는 모든 모델에 동일하게 적용하였다. 데이터 분할 시 random seed를 42로 고정하여 재현성을 보장하였다.

학습 조건은 모든 모델에서 동일하게 설정하였다. 옵티마이저는 Adam을 사용하였으며 초기 학습률은 0.0003으로 설정하였다. 배치 크기는 32, 최대 에폭은 50으로 설정하였고, 과적합 방지를 위해 0.5의 드롭아웃을 적용하였다. 학습 중 검증 손실이 5 에폭 동안 개선되지 않으면 학습률을 0.5배로 감소시키는 방법을 사용하였다. 학습은 4개의 NVIDIA RTX 4090 GPU를 활용하여 병렬로 수행하였다.

Table 1은 5-fold 교차검증을 통해 얻은 각 모델의 테스트 정확도, 표준편차, 모델 크기를 나타낸다. 실험 결과, ResNet50이 99.98 %의 가장 높은 테스트 정확도를 달성하였으나, 270 MB의 큰 모델 크기와 높은 연산량으로 인해 Jetson Nano 기반 임베디드 환경에는 부적합하여 정확도의 기준으로만 사용하였고 배포 후보에서 제외하였다.

실제 배포 가능한 경량 모델인 EfficientNet-B0과 MobileNetV2를 비교한 결과, EfficientNet-B0이 99.96%의 정확도와 ±0.042 %의 표준편차를 기록하여 MobileNetV2 보다 우수한 성능을 보였다. 따라서 임베디드 환경에 적합하면서도 높은 정확도와 안정성을 제공하는 EfficientNet-B0을 최종 모델로 선택하였다.

2.3 EfficientNetB0 architecture

앞서 언급한 바와 같이, EfficientNet-B0은 경량성과 높은 정확도를 동시에 제공하여 임베디드 환경에 적합하다[16]. EfficientNet-B0 아키텍처는 MBConv 블록을 기본 구성 요소로 사용하며, 각 블록은 확장 합성곱, 깊이별 합성곱, 축소 합성곱으로 구성된다(Fig. 12). 모델은 총

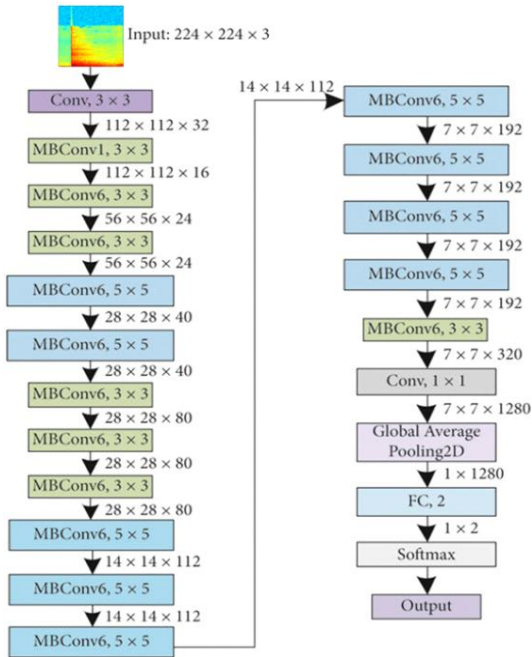


Fig. 12. EfficientNet-B0 network architecture

237개의 레이어로 구성되어 있으며, 약 5.3M개의 학습 가능한 파라미터를 가진다.

EfficientNet-B0은 기본적으로 3채널(RGB) 이미지 입력을 필요로 하기 때문에 단일 채널 Mel-Spectrogram 데이터를 3채널 형태로 변환하였다. 변환된 입력은 EfficientNet의 특성 추출 모듈을 통해 주요 시각적 특징을 추출한 뒤, 이를 간결하게 요약하기 위해 평균 풀링 연산을 적용하였고, 과적합 방지를 위해 0.5의 드롭아웃을 적용하였다. 최종 출력층은 2개의 클래스를 구분하는 Dense 레이어로 구성하였고, Softmax 함수를 통해 분류 확률을 출력하도록 설계하였다.

V. Experiment and Performance Evaluation

1. Experimental Setup

선정된 EfficientNet-B0 모델을 실제 골프공 분류 장치에 배포하기 위해 NVIDIA Jetson Nano 기반 임베디드 시스템을 구성하였다. Jetson Nano는 ARM Cortex-A57 쿼드코어 프로세서와 128-core Maxwell GPU를 탑재하여 실시간 딥러닝 추론에 적합하다.

제한된 연산 자원에서 실시간 분류가 가능하도록 TensorFlow Lite를 사용한 INT8(8-bit Integer) 양자화를 수행하였다. 양자화를 통해 모델 크기를 46.9 MB에서 4.7 MB로 약 10분의 1 수준으로 감소시켰다. 추론 속도는

원본 모델 대비 약 3.5배 향상되어 Jetson Nano에서 평균 100 ms 이내에 단일 샘플에 대한 예측이 완료된다.

2. Classification Performance Analysis

양자화에 따른 성능 변화를 검증하기 위해 Section 4에서 사용한 학습 데이터와 독립적인 테스트 데이터셋을 별도로 구축하였다. 동일한 하드웨어 장치를 사용하여 정상 공 630개, 손상 공 618개, 총 1,248개의 충격음 샘플을 새로 수집하였다. 이 데이터셋을 사용하여 양자화 전 모델(비최적화 모델)과 INT8 양자화 모델의 분류 성능을 비교하였다.

Fig. 13은 두 모델의 혼동행렬 비교를 보여준다. 비최적화 모델의 경우 전체 1,248개의 테스트 샘플 중 1,238개를 정확히 분류하여 99.20 %의 정확도를 달성하였다. 정상 공 630개 중 626개를 올바르게 분류하였고, 손상 공

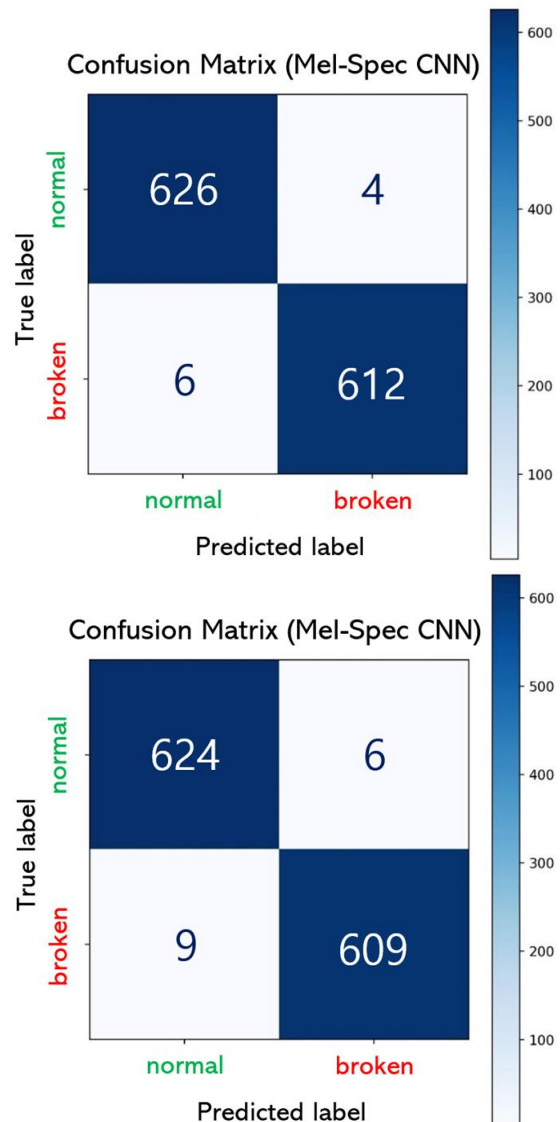


Fig. 13. Confusion matrix comparison between non-optimized (up) and INT8 quantized (down) models

Table 2. Classification experiment results

Normal to Normal	Normal to Broken	Broken to Broken	Broken to Normal
97.97 %	2.03 %	95.06 %	4.94 %

618개 중 612개를 정확히 판별하였다.

INT8 양자화 모델의 경우 테스트 정확도는 99.20%에서 98.80 %로 0.4 %p 소폭 하락하였다. 전체 1,248개 샘플 중 1,233개를 정확히 분류하였으며, 정상 공 630개 중 624개, 손상 공 618개 중 609개를 정확히 판별하였다. 이러한 정확도 하락은 실제 응용에서 허용 가능한 수준이며, 모델 크기 90% 감소와 추론 속도 3.5배 향상이라는 이점을 고려할 때 양자화의 효용성이 충분히 입증되었다.

두 모델 모두 손상 공을 정상 공으로 오분류한 경우(비최적화: 6건, 양자화: 9건)가 그 반대(비최적화: 4건, 양자화: 6건)보다 다소 많았다. 이는 실제 현장에서 손상 공을 놓칠 위험성을 의미하며, 향후 손상 공 검출 민감도를 높이기 위한 임계값 조정이나 데이터 불균형 보정이 필요하다.

3. Real-world Environment Verification

실험 환경에서 수집한 데이터는 노이즈가 적어 모델 성능이 과대 평가될 가능성이 존재하였다. 이를 검증하기 위해 실제 환경과 유사한 조건에서 추가 실험을 수행하였다. 다양한 생활 소음이 존재하는 강의실 복도를 실험 장소로 선정하고, 다양한 브랜드와 구조(2피스, 3피스, 4피스)의 골프공을 포함하였다(Fig. 14).

정상 공과 손상 공을 각각 344회씩, 총 688회 투입하여 분류 실험을 진행하였다. 실험 결과, 정상 공의 분류 정확도는 97.97 %, 손상 공의 분류 정확도는 95.06 %, 전체 평균은 96.5 %로 나타났다(Table 2). 이는 본 연구의 목표치인 90 % 이상의 정확도를 상회하는 성과이다.

오분류 사례를 심층 분석한 결과, 다음과 같은 패턴을 발견하였다. 첫째, 미세한 내부 균열을 가진 골프공의 경우 외관상으로는 정상으로 보이나 음향 특성이 손상 공과 유사하여 오분류되는 경향이 있었다. 둘째, 심하게 마모된

정상 공의 경우 표면 거칠기로 인해 충격음의 주파수 특성이 변화하여 일부 오분류가 발생하였다. 셋째, 다양한 피스 구조에 따라 내부 구조가 상이하여 음향 신호의 변동성이 증가하였다.

VI. Conclusion

본 연구에서는 로스트볼 시장에서 발생하는 품질 불량 문제를 해결하고자, 골프공 낙하 시 발생하는 충격음을 활용한 딥러닝 기반 자동 분류 시스템을 설계 및 구현하였다. 하드웨어와 소프트웨어를 통합한 본 시스템은 하드웨어 설계, 삼중 방음 구조, 낙하 가속 장치, 고감도 마이크로 구성된 음향 수집부와, CNN 기반 분류 모델로 구성되었다.

하드웨어 측면에서는 60 cm 낙하 높이에서의 충격음이 최적의 분류 성능을 제공함을 실험적으로 확인하였으며, DC775 모터와 실리콘 휠을 활용한 가속 장치를 통해 일정한 낙하 속도를 보장하였다. 삼중 방음 구조를 적용하여 외부 소음과 내부 잡음을 효과적으로 차단하고, 고품질 음향 데이터를 확보할 수 있었다.

소프트웨어 측면에서는 Mel-Spectrogram 기반 특징 추출과 EfficientNetB0 CNN 아키텍처를 활용하여 높은 분류 정확도를 달성하였다. 데이터 증강 기법으로 피치 변형, 동적 노이즈 삽입, 밴드패스 필터링, 데시벨 임계값 기반 스펙트로그램 생성을 적용하여 모델의 일반화 성능을 크게 향상시켰다. 특히 동적 노이즈 삽입 기법은 실제 소음 환경에서의 성능을 2.4 % 향상시키는 효과를 보였다.

실험 결과, 약 60 cm 높이에서 낙하시킨 골프공의 충격음을 분석한 CNN 모델은 모델 선정 단계에서 5-fold 교차검증을 통해 99.96 %의 정확도를 달성하였으며, 양자화 검증을 위해 별도 수집한 1,248개 샘플에서 비최적화 모델 99.20 %의 정확도를 기록하였다. 모델 경량화를 통해 46.9 MB에서 4.7 MB로 크기를 감소시켰으며, 정확도는 98.80 %로, Jetson Nano 환경에서의 실시간 추론이 가능함을 확인하였다.

실제 환경 검증 실험에서는 생활 소음이 존재하는 강의실 복도에서 688회의 분류 테스트를 수행하여 96.5 %의 평균 정확도를 달성하였다. 정상 공 분류 정확도는 97.97 %, 손상 공 분류 정확도는 95.06 %로, 본 연구의 목표치인 90 % 이상을 상회하는 성과를 거두었다. 이는 시스템의 상용화 가능성을 입증하는 결과이다.



Fig. 14. Structure of golf balls by piece type

본 시스템은 환경 측면에서 파손 골프공의 재활용률 향상과 폐기물 감축에 효과적으로 기여할 수 있으며, 산업 측면에서는 물류 대량 처리와 품질 선별 자동화를 통해 제품 신뢰도 향상에 실질적인 도움을 줄 수 있다. Arduino Mega와 Jetson Nano를 연동한 자동화 시스템은 시간당 1,200개의 골프공을 처리할 수 있어 대량 선별 작업의 효율성을 크게 개선하였다.

그러나 본 연구에는 몇 가지 한계점이 존재한다. 첫째, 손상 공을 정상 공으로 오분류하는 경우가 상대적으로 많아, 실제 현장에서 손상 공을 놓칠 위험성이 있다. 둘째, 골프공의 피스 구조에 따른 음향 신호 변동성이 크게 나타났다. 셋째, 미세한 내부 균열을 가진 골프공의 경우 분류 정확도가 다소 낮았다.

향후 연구에서는 다음과 같은 개선 방향을 제시한다. 첫째, 다양한 소음 환경에서도 견고하게 작동할 수 있도록 노이즈 억제 알고리즘을 고도화하고, 적응형 임계값 조정 기법을 도입할 예정이다. 둘째, 정상/손상의 이진 분류를 넘어 손상 유형별 다중 분류로 확장하여 시스템의 정밀성을 향상시킬 계획이다. 셋째, 피스 구조별로 최적화된 분류 모델을 개발하거나, 멀티태스킹 러닝 기법을 적용하여 다양한 골프공 타입에 대한 일반화 성능을 강화할 예정이다.

본 연구는 음향 기반 비파괴 검사 기법과 딥러닝 기술을 결합하여 로스트볼 선별 작업의 효율성과 품질 보증을 실현함으로써, 환경적·경제적 이점을 제공하는 실용적인 시스템을 제시하였다. 이는 제조업 품질 관리 분야에서 음향 분석 기반 자동화 시스템의 적용 가능성을 확장하는 데 기여할 것으로 기대된다.

ACKNOWLEDGEMENT

This research was financially supported by Hansung University.

REFERENCES

- [1] J. Bantock, "Billions of golf balls could be lost every year. Where do they end up?," CNN, <https://edition.cnn.com/sport/lost-golf-balls-environment-spt-spc>
- [2] USGA and R&A, "Equipment Rules", United States Golf Association, <https://www.usga.org/equipment-standards/equipment-rules-2019/equipment-rules/equipment-rules.html>, 2023.
- [3] S. Hershey, S. Chaudhuri, D.P.W. Ellis, J.F. Gemmeke, A. Jansen, R.C. Moore, M. Plakal, D. Platt, R.A. Saurous, B. Seybold, M. Slaney, R.J. Weiss, and K. Wilson, "CNN Architectures for Large-Scale Audio Classification," in Proc. IEEE ICASSP, pp. 131-135, New Orleans, LA, USA, March 2017. DOI: 10.1109/ICASSP.2017.7952132
- [4] J. Abeßer, "A Review of Deep Learning Based Methods for Acoustic Scene Classification," Applied Sciences, Vol. 10, No. 6, p. 2020, March 2020. DOI: 10.3390/app10062020
- [5] K.J. Piczak, "Environmental Sound Classification with Convolutional Neural Networks," in Proc. IEEE 25th Int. Workshop Mach. Learn. Signal Process. (MLSP), pp. 1-6, Boston, MA, USA, September 2015. DOI: 10.1109/MLSP.2015.7324337
- [6] J. Salamon and J.P. Bello, "Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification," IEEE Signal Process. Lett., Vol. 24, No. 3, pp. 279-283, March 2017. DOI: 10.1109/LSP.2017.2657381
- [7] D.S. Park, W. Chan, Y. Zhang, C.C. Chiu, B. Zoph, E.D. Cubuk, and Q.V. Le, "SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition," in Proc. Interspeech 2019, pp. 2613-2617, Graz, Austria, September 2019. DOI: 10.21437/Interspeech.2019-2680
- [8] J.G. Lee, K.S. Kim, and J.H. Lee, "Sound-Based Unsupervised Fault Diagnosis of Industrial Equipment Considering Environmental Noise," Sensors, Vol. 24, No. 22, pp. 7319, November 2024. DOI: 10.3390/s24227319
- [9] S. Sikdar, D. Liu, and A. Kundu, "Acoustic Emission Data Based Deep Learning Approach for Classification and Detection of Damage-Sources in a Composite Panel," Composites Part B: Engineering, Vol. 228, pp. 109450, January 2022. DOI: 10.1016/j.compositesb.2021.109450
- [10] S.A. Shevchik, G. Masinelli, C. Kenel, C. Leinenbach, and K. Wasmer, "Deep Learning for In Situ and Real-Time Quality Monitoring in Additive Manufacturing Using Acoustic Emission," IEEE Trans. Ind. Inform., Vol. 15, No. 9, pp. 5194-5203, September 2019. DOI: 10.1109/TII.2019.2910524
- [11] J.K. Zhang, W. Yan, and D.M. Cui, "Concrete Condition Assessment Using Impact-Echo Method and Extreme Learning Machines," Sensors, Vol. 16, No. 4, p. 447, March 2016. DOI: 10.3390/s16040447
- [12] T. Ji and N.M. Nor, "Deep Learning-Empowered Digital Twin Using Acoustic Signal for Welding Quality Inspection," Sensors, Vol. 23, No. 5, pp. 2643, March 2023. DOI: 10.3390/s23052643
- [13] H. Zhang, Q. Wu, W. Tang and J. Yang, "Acoustic Signal-Based Defect Identification for Directed Energy Deposition-Arc Using Wavelet Time-Frequency Diagrams," Sensors, Vol. 24, No. 13, p. 4397, July 2024. DOI: 10.3390/s24134397
- [14] H.H. Wu, J.W. Su, and C.L. Chen, "Automatic Optical Inspection System Design for Golf Ball," in Proc. SPIE Applications of

Digital Image Processing XXXIX, Vol. 9971, 2016. DOI: 10.1117/12.2237286

- [15] Y.Z. Wu and Y.C. Huang, "Steel Ball Defect Detection System Using Automatic Vertical Rotating Mechanism and Convolutional Neural Network," *Computers, Materials & Continua*, Vol. 83, No. 1, pp. 97-114, 2025. DOI: 10.32604/cmc.2025.063441
- [16] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proc. ICML*, pp. 6105-6114, Long Beach, CA, USA, June 2019. <https://proceedings.mlr.press/v97/tan19a.html>
- [17] A.P. Berg, Q. Zhang, and M.Y. Wang, "15,500 seconds: Lean UAV classification using EfficientNet and lightweight fine-tuning," *arXiv:2506.11049*, 2025.
- [18] R. Krishnamoorthi, "Quantizing deep convolutional networks for efficient inference: A whitepaper," *arXiv:1806.08342*, 2018.
- [19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 770-778, Las Vegas, NV, USA, June 2016. DOI: 10.1109/CVPR.2016.90
- [20] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 4510-4520, Salt Lake City, UT, USA, June 2018. DOI: 10.1109/CVPR.2018.00474

Authors



Seokho Chung is currently pursuing the B.S. degree with the Division of Mechanical and Electronic Engineering, Hansung University, Seoul, Korea. Mr. Chung's research interests focus on mechanical system design and learning-based methods.



Jinho Cho is currently pursuing the B.S. degree with the Division of Mechanical and Electronic Engineering, Hansung University, Seoul, Korea. Mr. Cho's research interests focus on the mobility system design and control systems.



Woonho Lee is currently pursuing the B.S. degree with the Division of Mechanical and Electronic Engineering, Hansung University, Seoul, Korea. Mr. Lee's research interests focus on the mechanical design and control.



Sangyul Lee received the B.S., M.S. and Ph.D. degrees in Mechanical Engineering from Seoul National University, Korea, in 2006, 2008 and 2013, respectively. Dr. Lee joined the faculty of the Department of

Robotics and Automation Engineering at Hoseo University, Asan, Korea, in 2017. He is currently a Associate Professor in the Division of Mechanical and Electronic Engineering, Hansung University. His research interests include the integration of mechanical systems and AI, computed tomography, and design and development of Unmanned Ground Vehicles.