

# 비대면 원격진단을 위한 디지털 검이경 청진기 헬스케어 플랫폼 개발

최수영<sup>1,2</sup> · 이 학<sup>2</sup> · 박찬용<sup>1</sup> · 주수빈<sup>1</sup> · 권오원<sup>1</sup> · 이동규<sup>1\*</sup>

<sup>1</sup>한국기계연구원 대구융합기술연구센터 의료기계연구실

<sup>2</sup>경북대학교 기계공학부

## Development of a Digital Otoscope-Stethoscope Healthcare Platform for Telemedicine

Su Young Choi<sup>1,2</sup>, Hak Yi<sup>2</sup>, Chanyong Park<sup>1</sup>, Subin Joo<sup>1</sup>, Ohwon Kwon<sup>1</sup> and Dongkyu Lee<sup>1\*</sup>

<sup>1</sup>Department of Medical Device, Korea Institute of Machinery & Materials (KIMM)

<sup>2</sup>School of Mechanical Engineering, Kyungpook National University

(Manuscript received 29 April 2024 ; revised 24 May 2024 ; accepted 4 June 2024)

**Abstract:** We developed a device that integrates digital otoscope and stethoscope for telemedicine. The integrated device was utilized for the collection of tympanic membrane images and cardiac auscultation data. Data accumulated on the platform server can support real-time diagnosis of heart and eardrum diseases using artificial intelligence. Public data from Kaggle were used for deep learning. After comparing with various deep learning models, the MobileNetV2 model showed superior performance in analyzing tympanic membrane data, and the VGG16 model excelled in analyzing cardiac data. The classification algorithm achieved an accuracy of 89.9% for eardrums data and 100% for heart sound data. These results demonstrate the possibility of diagnosing diseases without the limitations of time and space by using this platform.

**Key words:** Telemedicine, Digital healthcare platform, Integrated device, Artificial intelligence, EarDrum and heart sound

### 1. 서 론

최근 환경오염, 흡연, 바이러스 등의 문제로 만성 호흡기 질환, 알레르기 질환과 같은 유병율이 크게 증가하고 있으며, 지속적인 관리가 요구되고 있다[1-3]. 특히, 원격진료 서비스를 통해 만성 질환 환자의 건강을 효율적으로 관리할 수 있으며, 거동이 불편한 노령의 환자 및 의료 접근성이 낮은 의료 소외지역의 환자에 대한 지속적인 검진을 위해 원격진료의 필요성이 증가하고 있다[4]. 또한, 코로나19의 확산과 같은 전염병 상황에서 의료 소비자의 욕구 증가로 홈 헬스케어 서

비스 등으로 자신의 건강상태를 지속 관리하는 비대면 진료의 수요가 크게 증가하였고, 의료 서비스 산업의 패러다임이 빠르게 변화하고 있다[5].

지속적인 건강상태 확인을 위한 내과진료에 사용되는 진단도구는 체온계, 검이경 및 청진기를 활용하여 환자를 진찰하지만, 가정에서 쉽게 사용할 수 있는 체온계와 다르게 검이경과 청진기를 이용한 진찰은 의사의 전문적인 진단이 필요하기 때문에 환자들은 주기적으로 병원을 내원해야 한다. 최근 이러한 문제를 해결하기 위해 환자와 의료인이 직접적인 접촉 없이도 스마트폰과 연동하여 기초 진료 및 건강관리가 가능한 다양한 원격진료 기기가 개발되고 있다. 원격진료를 위한 대표적인 디지털 검이경 및 청진기로 셀스콥(Cellscope)사의 오토스콥(Otoscope), 스테소미(StethoMe)사의 스테소미(StethoMe), 타이토케어(TytoCare)사의 타이토홈(TytoHome) 복합장치가 있다. 오토스콥은 스마트폰과 연동되어 사용자가 직접 고막 사진을 촬영하여 고막의 상태를

\*Corresponding Author : Dongkyu Lee

Department of Medical Device, Korea Institute of Machinery & Materials (KIMM), Daegu 42994, Korea

Tel: +82-53-670-9110

E-mail: dongkyu@kimm.re.kr

본 연구는 산업통상자원부 바이오산업기술개발사업 글로벌진출형 디지털치료기 개발지원 세부사업(20018535)의 지원을 받아 수행되었습니다.

검사하고 기록할 수 있는 디지털 검이경으로, 조명 기술과 스마트폰 카메라를 통합하여 원격 진단 및 관리 장치이다[6-8]. 스테소미는 환자가 자신의 호흡음을 녹음하여 의사가 원격으로 청취할 수 있게 해주는 스마트 디지털 청진기로 심음을 이용하여 천식과 같은 기도 질환을 진단하는데 특화되어 있다[9-11]. 타이토홈은 하나의 기기에 디지털 검이경 및 청진기 등 다양한 진단도구로 구성된 복합 기능 의료 기기로 원격진료에 특화된 장치이다. 또한, 이 기기는 무선통신기반 디지털 청진기, 디지털 검이경 진단 도구로 환자가 수집한 검진 사진 및 심음을 의료진과 화상대화를 통해 원격진료를 수행하는 제품이다. 타이토홈은 환자가 집에서 편리하고 고품질의 기본 의료 서비스를 받을 수 있도록 하여 의료 접근성을 향상시키는 데 기여하고 있다[12].

하지만 원격의료의 환자 와 의료인 사이의 지리적 접근성을 개선하는 중요한 역할을 수행할 수 있으나, 기존의 원격진료 방식은 여러 한계에 직면해 있다. 원격진료는 의료 전문가가 환자를 직접 관찰하고 검사할 수 없기 때문에 진단의 정확도가 낮고 질병에 대한 즉각적인 대응이 어렵다. 이는 만성질환과 같은 질병을 진단하고 관리함에 있어서 큰 단점이 되어 왔다. 이러한 한계점을 극복하기 위해서 최근 인공지능 기술을 비대면 진료에 도입하는 연구가 활발하게 이루어지고 있다[13-18]. 고막 관련 인공지능 논문에서는 고막 이미지로부터 옥타브 3D 컨볼루션(Octave 3D Convolution)과 지역 집중 모듈(Region-focus modules)의 조합을 통해 다양한 고막과 관련된 질환을 99.3%의 정확도로 진단하는 정량적인 결과를 보였다[13]. 또한, 심음과 관련된 인공지능 선행 연구에서는 수집된 심음의 주파수 분석을 통해 이상을 알고리즘과 다양한 인공지능 기법을 비교하여 심장질환의 이상유무를 비정상 97%, 정상 88%의 정확도로 판단하는 연구 결과가 보고되었다[14].

따라서, 본 연구에서는 비대면 원격진료 시 검진 효율성과 고사양 하드웨어의 가격 경쟁력을 고려하여 디지털 검이경-청진기 복합기능을 하나의 장치로 개발하였다. 개발된 복합장치는 검이경 모드에서 고막 이미지를 획득할 수 있고, 청진기 모드에서 심음을 녹음할 수 있다. 수집된 데이터는 각각의 모드에서 스마트폰 앱(App)으로 블루투스 통신을 통해 전송되어 관리된다. 또한, 스마트폰 앱에서 개별 데이터를 플랫폼 서버로 전송한 뒤 빅데이터 및 인공지능 분석을 수행하고 결과를 스마트폰 앱으로 다시 수신 받을 수 있는 디지털 검이경-청진기 헬스케어 플랫폼으로 개발했다.

플랫폼 내 수집된 고막이미지 및 심음데이터에 대해 이미지분석, 음향 분석 등의 빅데이터 분석 알고리즘과 인공지능 판별 분석 알고리즘을 적용하였다. 공개 데이터셋(Open Dataset)과 복합장치에서 수집된 고막이미지 및 심음 데이터를 이용하여 다양한 딥러닝 모델에서 학습시키고 가장 적합

한 모델을 선별하는 연구를 수행하였다. 이를 통해 정상과 비정상을 구분하는 모델로 개발하고, 선별한 모델을 플랫폼 서버와 스마트폰 앱에 적용하였다.

## II. 연구방법

### 1. 디지털 검이경-청진기 플랫폼 설계

그림 1은 디지털 검이경-청진기 복합 장치의 구성과 복합장치에서 획득된 이미지 데이터 및 음향 데이터를 스마트폰 앱과 연동하여 수집하고, 플랫폼 서버로 데이터를 전송하여 분석 결과를 스마트폰 앱으로 확인하는 플랫폼 기술의 전반적인 구조를 나타낸다. 복합장치는 Single Board Computer, Touch Screen, Audio ADC Converter로 구성되고 CMOS Camera module 및 Mems microphone이 장착되었다. 카메라로 실시간 이미지를 확인 후 이미지를 저장할 수 있고, 심음을 실시간으로 사용자가 듣고 원하는 음향을 40 초간 녹음할 수 있다. 복합장치에 저장된 이미지 및 심음 데이터는 블루투스 통신을 통해 스마트폰 앱으로 송신된다. 수집된 데이터는 스마트폰 앱에서 관리되고, Wi-Fi를 통해 플랫폼 서버로 전송하면 플랫폼 내 이미지 및 음향분석과 인공지능 분석을 플랫폼 서버에서 수행한다. 분석된 결과는 스마트폰 앱을 통해 확인할 수 있다.

디지털 검이경-청진기의 검이경 모드에서는 고막 이미지 데이터를 수집하고 청진기 모드에서는 심음을 수집하여 분석하는 전체 구동알고리즘을 구현하였다. 플랫폼 서버에는 검이경 모드에서 지속적인 사용자의 고막 데이터를 수집하고 분류하기 위해서 경계(Edge) 분석 및 색영역(RGB Histogram) 분석을 추가 수행하여 고막의 감염, 염증, 출혈 등과 같은 다양한 상태를 파악한다. 또한, 청진기 모드에서는 심음의 음향

110

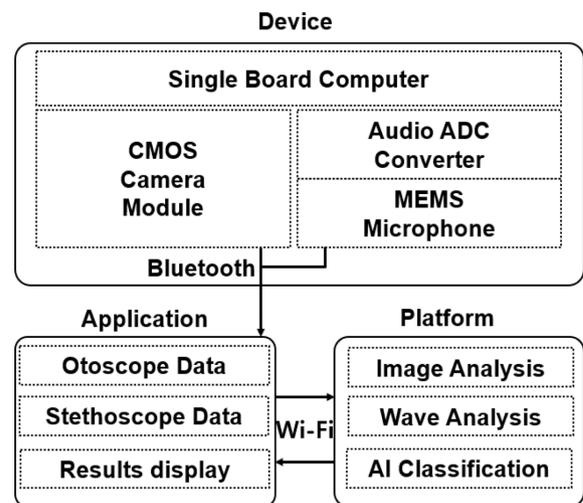


그림 1. 디지털 검이경-청진기 비대면 진료 플랫폼 설계  
Fig. 1. Design of a Digital Otoscope-Stethoscope Telemedicine Platform

파일을 분석하기 위해 20~150 Hz에서 BandPass Filtering을 통해 목표 주파수 이외의 노이즈를 감쇄한다. 심음 음향 파일에서 Dynamic analysis 프로그램을 통해 심박수를 추출한다[21]. 심음은 0~20 Hz 구간과 150~300 Hz 구간의 주파수 분석을 수행할 수 있다[22]. 인공지능 분석을 통해 현재 수집된 데이터의 고막 이미지 및 심음의 상태가 정상 및 비정상 여부를 판별하고 확률을 표시하도록 하였다. 플랫폼 서버에서 분석된 각각의 결과는 스마트폰 앱 결과 창에서 확인할 수 있다.

## 2. 디지털 검이경-청진기 플랫폼 제작

우선 제안한 플랫폼을 제작하기 위해 상용 부품을 선정하였다. 구성 부품으로는 Raspberry Pi Trading사의 Raspberry Pi 3 Model B+ 싱글 보드 컴퓨터를 사용하여 구동 소프트웨어를 구현하였다. 사용자 인터페이스 구현을 위해서는 Wolfson Microelectronics사의 5 인치 DSI LCD 모니터가 사용되었고, 영상 촬영을 위한 카메라로는 New Landing사의 CMOS Otoscope Camera를 이용하여 고막은 45 fps와 640×480의 해상도로 고막이미지를 촬영했다. Audio ADC Converter로는 Wolfson Microelectronics사의 WM8960G14J6PDR 제품을 장착하여 아날로그 음향 신호를 디지털 데이터로 변환하였다. TDK InvenSense사의 ICS-40310 MEMS 마이크로폰을 활용하여 심음을 정밀하게 감지하였으며, 3M사의 다이어프램(Diaphragm)으로는 Classic II S.E., Grey, 40006 부품을 이용하여 벨(Bell)구조를 설계하고 3D 프린터로 구조를 제작하여 조립했다. 그

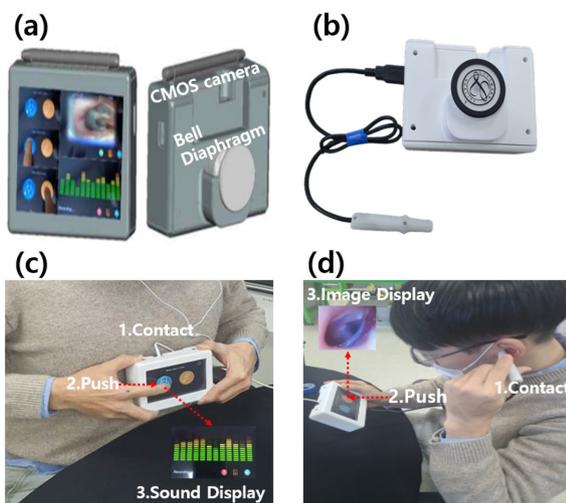


그림 2. 복합장치의 설계 및 제작 및 신호측정사진: (a) 복합장치3D 설계, (b) 실제복합장치 사진, (c) 심음측정사진, (d) 고막측정사진  
Fig. 2. Design, Fabrication, and Signal Measurement Images of the Device: (a) 3D Design of the Device, (b) Photo of the Actual Device, (c) Image of Heart Sound Measurement, (d) Image of Tympanic Measurement

림 2(a)는 디지털 검이경-청진기 복합 장치의 하드웨어 부품의 하우징 설계 이미지이다. 그림 2(b)는 CMOS camera 및 MEMS microphone 회로가 장착된 벨과 소리 신호 증폭을 위한 다이어프램을 조립한 디지털 검이경-청진기 복합장치 이미지이다. 그림 2(c)는 복합장치의 청진기 모드를 이용하여 사용자의 심박수를 측정하는 과정이다. 헤드셋을 통해 사용자의 심음을 청취하고 심박수 데이터를 수집하는 사진이다. 그림 2(d)는 복합장치의 검이경모드를 이용하여 고막을 관찰하고 고막 데이터를 수집하는 사진이다.

## 3. 학습 데이터 선별

복합장치의 검이경 모드를 통해 획득한 고막 이미지를 정상과 비정상으로 구별하기 위한 딥러닝 설계를 수행하였다. 딥러닝의 학습을 위한 데이터셋은 Kaggle의 Tympanic membrane / eardrum dataset / otitis media의 데이터를 활용하였다[23]. 이 데이터셋의 정상 및 고막질환 환자의 데이터의 이미지 해상도는 500×500 픽셀이다. 또한, 공개된 이미지 데이터에서 정상고막 534 개, 비정상 고막 421 개를 사용하여 학습데이터셋을 선정하였다. 비정상 고막의 데이터는 Earwax(귀지 과다 분비 질환) 140 개, Acute Otitis Media(급성 감염성 귀질환) 119 개, Chronic suppurative Otitis Media(만성 감염성 귀질환) 63 개, Otitis externa(외이도 염증 귀질환) 41 개, Tympanoskleros(고막 경화 질환) 28 개, Ear ventilation tube(만성 중이염) 16 개, Pseudo membranes(유사 고막 질환) 11 개, Foreign bodies in the ear(이물질 귀질환) 3 개로 구성되어 있다. 본 연구에서는 각각의 고막 데이터 분할 비율은 6:2:2로, 훈련(Train), 검증(Validation), 테스트(Test)셋으로 나누었다(표 1).

청진기 모드의 딥러닝 설계를 위한 학습 데이터는 심음 데이터를 기반으로 사용했다. 심음은 심장의 수축 및 이완 과정에서 판막의 개폐 시 발생하는 소리를 의미한다. 본 연구

표 1. 고막 학습 데이터셋

Table 1. Tympanic membrane training dataset

Class	Train	Validation	Test
Normal	321	107	106
Earwax Plug	84	28	28
Acute Otitis Media	71	24	24
Chronic Suppurative Otitis Media	37	13	13
Otitis Externa	25	8	8
Tympano skleros	16	6	6
Ear Ventilation	10	3	3
Pseudo Membranes	7	2	2
Foreign bodies in the ear	2	1	0

에서는 심음 진단을 위해 심장의 수축기에 해당하는 제1심음(S1)과 확장기에 해당하는 제2심음(S2)으로 나누어 청진 분석을 진행했다. 제1심음(S1)은 심방과 심실 사이에 있는 방실판막이 닫힐 때 발생하며, 제2심음(S2)은 좌우 심실과 대혈관 사이에 위치한 반월판이 닫히는 순간에 나는 소리로 정의된다. 이와 같은 심음의 특성을 파악하기 위하여, Kaggle에서 제공하는 Heartbeat Sounds(심음) 데이터셋을 활용했다[24]. 이 데이터셋은 Dataset-A와 Dataset-B로 구성되어 있다. Dataset-A는 스마트폰 앱을 통해 수집된 심장박동에 대한 음향 파일이며, Dataset-B는 디지털 청진기를 사용하여 병원의 임상 시험에서 수집한 음향 파일로 구성되어 있다. 또한 Dataset-A는 Normal(정상), Extrahls(심실 조기 박동 심잡음), Murmur(심장판막 결함 심잡음) 클래스를 포함하며, Dataset-B는 Normal(정상), Extrastole(심실 조기 박동 심잡음), Murmur(심장판막 결함 심잡음), MurmurNoise(심장판막 외 심잡음) 클래스로 이루어져 있다. 이러한 데이터를 BandPass Filtering(대역 통과 필터) 소프트웨어로 주파수 구간을 필터링했다. 심음의 주요 주파수 대역은 20~300 Hz 이지만 일상 소음 주파수 150~300 Hz 구간을 필터링 하고 FFT(Fast Fourier Transform)를 통해 분석한 결과 30~40 dB 이상 노이즈가 감쇄되는 것을 확인할 수 있

었다. BandPass Filtering 후, STFT(Short-Time Fourier Transform)를 적용하면 시간에 따른 주파수 대역폭의 변화를 분석할 수 있다. 필터링된 음성 데이터는 FFT로 확인한 주파수 감쇄율을 STFT를 이용하여 필터링 주파수 감쇄율을 시각적으로 나타내었다. 그림 3(a)는 정상 심음을 필터링했을 때 나타나는 STFT 그래프이며, 그림 3(b)는 비정상 심음을 필터링하고 STFT 로 변환했을 때 나타나는 그래프다.

표 2는 Kaggle에서 제공하는 5초 이상의 심박수 음향 파일을 0~5초 구간으로 변인 통제하여 필터링한 후, FFT와 STFT를 적용하여 생성된 학습 데이터셋을 나타낸다. 심음 데이터셋 역시 고막 데이터 데이터 분할 비율과 동일하게 6:2:2로, 훈련, 검증, 테스트셋으로 나누었다. 구체적으로, Dataset-A에서 Normal 클래스는 훈련용 17 개, 검증용 6 개, 테스트용 6 개, Extrahls 클래스는 훈련용 9 개, 검증용 3 개, 테스트용 3 개, Murmur 클래스는 훈련용 19 개, 검증용 7 개, 테스트용 6 개의 데이터를 할당했다. Dataset-B는 Normal 클래스에 훈련용 34 개, 검증용 11 개, 테스트용 11 개, Extrastole 클래스에 훈련용 14 개, 검증용 5 개, 테스트용 5 개, Murmur 클래스에 훈련용 23 개, 검증용 8 개, 테스트용 8 개, MurmurNoise 클래스에 훈련용 13 개, 검증용 5 개, 테스트용 4 개를 포함한다.

112

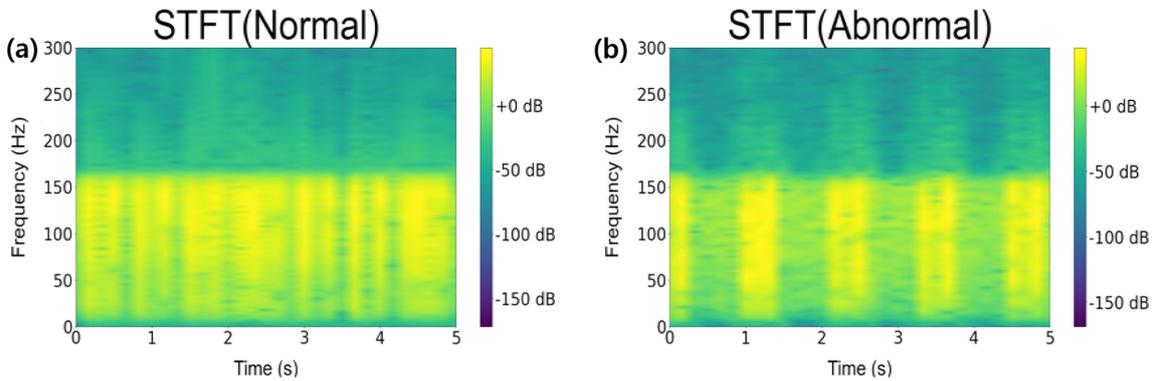


그림 3. 필터링 및 STFT 변환 후 정상 및 비정상 심음 비교 그래프: (a) Normal STFT 사진, (b) Abnormal STFT 사진  
 Fig. 3. Graph comparing normal and abnormal heart sounds after filtering and STFT transformation: (a) Normal STFT image, (b) Abnormal STFT image

표 2. 심음 질환 진단을 위한 데이터셋  
 Table 2. A dataset for diagnosing heart sound disorders.

Dataset	Class	Train Data	Validation Data	Test Data
Dataset-A	Normal	17	6	6
	Extrahls	9	3	3
	Murmur	19	7	6
Dataset-B	Normal	34	11	11
	Extrastole	14	5	5
	Murmur	23	8	8
	Murmur Noise	13	5	4

#### 4. 성능평가방법

Kaggle Dataset을 이용하여 취득한 생체데이터의 예측값과 결과값을 비교하기 위해서 혼동 행렬(Confusion Matrix)를 나타냈다. 혼동 행렬에서의 참 양성(True Positive, TP), 참 음성(True Negative, TN), 거짓 양성(False Positive, FP), 거짓 음성(False Negative, FN)을 기반으로 한 딥러닝 모델의 성능 평가 지표인 재현율(Recall), 정밀도(Precision), F1 점수(F1-Score), 정확도(Accuracy, Acc)는 모델의 실제 환경 적용 가능성을 예측하는 데 필수적인 역할을 수행한다. Acc는 모델이 양성 및 음성 샘플을 정확하게 분류한 비율을 나타내며, 구체적인 수식은 식 (1)로 표현된다. Recall은 모든 실제 양성 샘플 중 모델이 양성으로 예측한 비율을 나타내며, 식 (2)로 표현된다. Recall 값은 모델이 실제 양성 샘플을 누락 없이 잘 포착함을 의미한다. Precision은 모델이 양성으로 예측한 샘플 중 실제로 양성인 샘플의 비율을 나타내며, 식 (3)으로 표현된다. 높은 Precision 값은 모델의 양성 예측 샘플이 실제로 양성일 가능성이 높음을 의미한다. F1-Score는 Recall과 Precision의 조화 평균을 나타내며, 식 (4)로 표현된다. AUC 값은 ROC 곡선 아래의 면적을 의미하며, 성능이 우수한 모델일수록 이 값이 1에 근접한다.

$$Acc(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (1)$$

$$Recall(\%) = \frac{TP}{TP + FN} \times 100 \quad (2)$$

$$Precision(\%) = \frac{TP}{TP + FP} \times 100 \quad (3)$$

$$F1\ Score(\%) = \frac{2 \times Recall \times Precision}{(Recall + Precision)} \times 100 \quad (4)$$

#### 5. Classification 알고리즘 적용

학습 데이터 선정과 평가 방법을 선정한 이후에 고막과 심음 STFT Graph의 정상 및 비정상 상태를 분류하기 위한 딥러닝 기반의 분류 모델들을 적용했다. 이를 위해, PyTorch 라이브러리를 활용하여, 3 채널(RGB)의 이미지를 입력으로 받는 신경망 모델들(VGG16, MobileNetV2)을 적용했다. 신경망 모델들은 첫 번째 입력층을 수정하여 3 채널의 입력 이미지에서 16 개의 필터를 사용하는 컨볼루션 레이어로 제작하였고, 마지막 레이어에서 2 개의 출력 노드를 통해 이미지가 정상 또는 비정상 중 어느 것인지를 분류하도록 수정하였다. 이미지 전처리 과정에서는 청진기 모드는 STFT로 구현한 2 차원 그래프 이미지를 150×150 크기로 변환하였고, 검이경 모드는 이미지를 150×150 크기로 변환하여 생체신호 데이터를 판별하였다.

### III. 연구 결과

#### 1. 고막 질환 판별 모델 성능평가

검이경 모드에서 고막질환 판별을 위한 MobileNetV2, InceptionV3, ResNet50, VGG16 4 종의 딥러닝 모델을 선정하였다[25-28]. 4가지 딥러닝 모델 성능은 표 1의 학습 데이터셋을 통해 학습시키고 이후에 평가 데이터를 활용하여 그림 4와 같이 TN, TP, FP, FN를 혼동행렬로 시각화 하였다. ResNet50은 (TN=76), (FP=8), (FN=21), (TP=85)의 결과를 나타냈다. InceptionV3는 (TN=84), (FP=0), (FN=89), (TP=17)의 결과를 보여준다. MobileNetV2는 (TN=77), (FP=7), (FN=19), (TP=87)의 결과를 보여준다. 마지막으로 VGG16은 (TN=75), (FP=9), (FN=18), (TP=88)의 결과를 보여준다. 혼동행렬을 통해 평가 결과를 비교하면 ResNet50, MobileNetV2, VGG16의 성능이 InceptionV3와 비교했을 때 상대적으로 우수한 것으로 판단된다. 그림 5에서는 ROC 곡선과 AUC 지표를 활용하여 추가적인 성능 평가를 수행했다. 그림 5에서 ResNet50과 MobileNetV2가 AUC=0.93로 가장 높은 AUC 수치를 확인할 수 있었다.

표 3에서 각 모델을 Recall, Precision, F1-Score, Accuracy, 그리고 AUC의 평가지표를 활용해서 종합적인 성능을 분석하였다. InceptionV3 모델은 Precision = 100%를 보였으나, Recall과, F1-Score 및 Acc가 타 모델에 비해 상대적으로 낮아 모델로 적용하기 적합하지 않다. 최종적으로 MobileNetV2의 종합적인 성능수치가 가장 높은 것으로 확인되었다.

		ResNet50		VGG16	
True Label	Abnormal	76	8	75	9
	Normal	21	85	18	88
		Abnormal	Normal	Abnormal	Normal
		Predicted Label		Predicted Label	
		MobileNetV2		InceptionV3	
True Label	Abnormal	77	7	84	0
	Normal	19	87	89	17
		Abnormal	Normal	Abnormal	Normal
		Predicted Label		Predicted Label	

그림 4. 딥러닝 모델의 고막질환 진단 성능 비교를 위한 Confusion Matrix 분석

Fig. 4. Analysis of Confusion Matrix for Comparing the Performance of Deep Learning Models in Diagnosing Tympanic Membrane Disorders

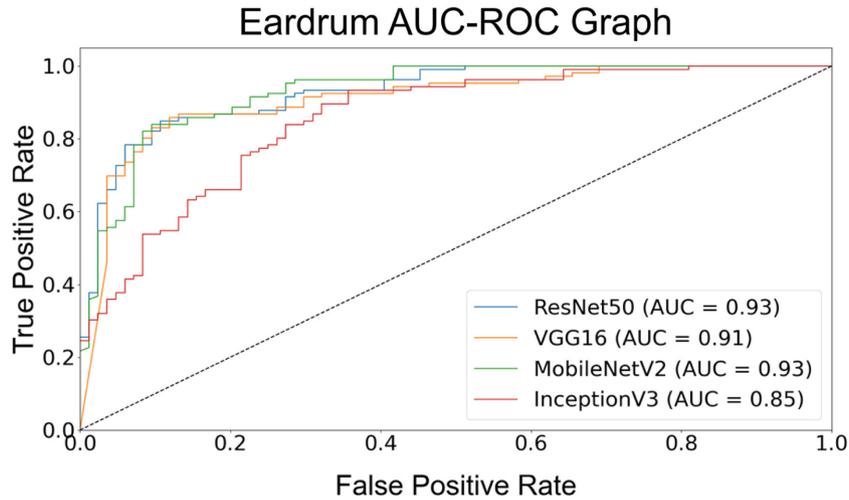


그림 5. 고막질환 진단 딥러닝 모델의 AUC-ROC 곡선  
 Fig. 5. The AUC-ROC curve of a deep learning model for diagnosing tympanic membrane disorders

표 3. 정상 및 비정상 고막 판별을 위한 네 가지 모델의 성능 비교

Table 3. A comparison of the performance of four models for distinguishing between normal and abnormal tympanic membranes

Model	Recall (%)	Precision (%)	F1 Score (%)	Accuracy (%)	AUC
Inception V3	16.04	100	27.61	53.16	0.85
ResNet50	80.19	91.39	85	84.74	0.93
VGG16	83.02	90.72	86.76	85.79	0.91
MobileNetV2	82.08	92.55	87.08	86.32	0.93

114

2. 심음 판별 모델 성능평가

청진기 모드에서 심음질환 판별을 위해 ResNet50, VGG16, SuffleNetV2, DenseNet121 모델을 선정하였다[27-30]. 이후에 모델들은 표 2의 심음 STFT Test 데이터셋을

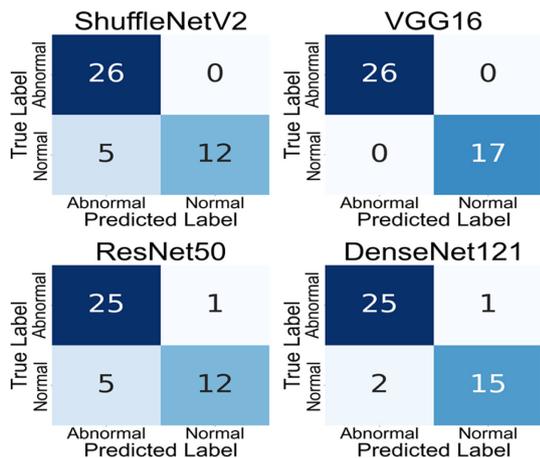


그림 6. 딥러닝 모델의 심음 질환 진단 성능 비교를 위한 Confusion Matrix 분석  
 Fig. 6. Analysis of the confusion matrix to compare the performance of deep learning models in diagnosing heart sound disorders

통해 그림 6과 같이 혼동행렬을 도출하였다. VGG16이 TP, TN 결과가 상대적으로 가장 우수하였으며 DenseNet121, ShuffleNetV2, ResNet50 순으로 성능이 높은 결과를 보였다. 그림 7의 ROC 곡선과 AUC 지표를 활용한 추가적인 성능 평가에서도 VGG16은 AUC 지표에서 1로 가장 높은 수치를 기록하며 최적의 성능을 보였다. DenseNet121과 ShuffleNetV2이 AUC = 0.98로 동일한 수치의 성능을 보였고, ResNet50이 AUC = 0.93으로 가장 낮은 성능수치를 보였다.

이후 종합적인 성능수치는 표 4에서 보듯이 VGG16은 5가지 모든 성능 수치에서 만점의 성능을 보였으며, DenseNet121, SuffleNetV2, ResNe50모델 순으로 높은 성능을 보였다.

3. 비대면 디지털 검이경-청진기 플랫폼 개발

학습된 최적의 모델들을 통해 디지털 검이경-청진기 플랫폼을 완성하였다. 그림 8(a)처럼 고막 이미지의 경우, 원본 이미지에 엣지 분석을 적용해 구조적 특징을 파악한 뒤, 색상 프로파일 분석으로 고막의 색상 변화와 패턴을 분석할 수 있다. 이를 바탕으로 인공지능 알고리즘을 통해 정상 혹은 비정상 여부를 판별한다. 그림 8(b)는 인공지능 알고리즘을 통

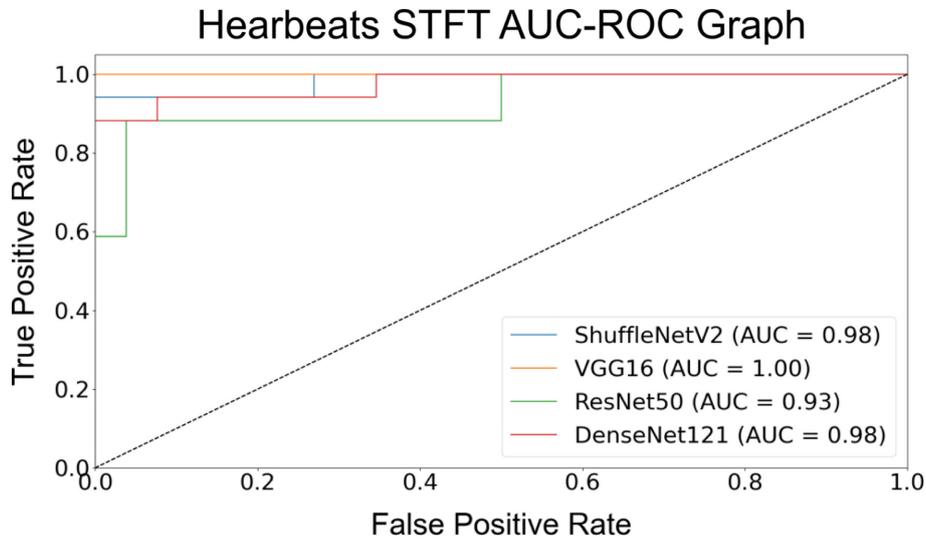


그림 7. 심음질환 진단 딥러닝 모델의 AUC-ROC 곡선  
 Fig. 7. The AUC-ROC curve for the deep learning model diagnosing heart sound disorders

표 4. 정상 및 비정상 심음 판별을 위한 네 가지 모델의 성능 비교  
 Table 4. Comparing the performance of four models to distinguish between normal and abnormal heart sounds

Model	Recall(%)	Precision(%)	F1 Score(%)	Accuracy(%)	AUC
ShuffleNetV2	70.59	100	82.76	88.37	1
ResNet50	70.59	92.31	80.04	86.05	0.93
VGG16	100	100	100	100	1
DenseNet121	88.24	93.75	90.87	93.02	0.98

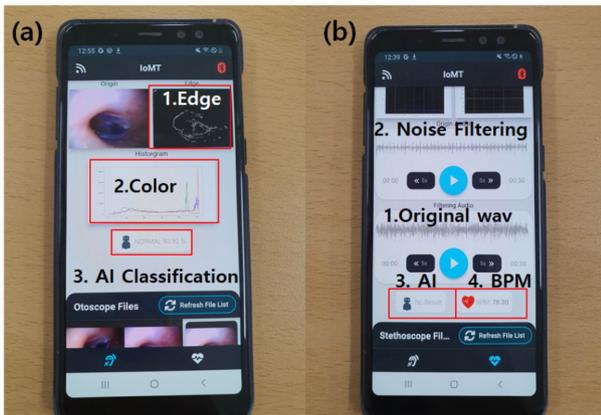


그림 8. 복합장치를 이용한 생체신호 딥러닝 판별 결과의 앱 전송 확인: (a) 고막분석사진, (b) 심음분석사진  
 Fig. 8. Verification of app transmission of biometric signal deep learning discrimination results using a device: (a) tympanic membrane analysis photo, (b) heart sound analysis photo

한 심음의 정상 혹은 비정상에 대한 판별 결과와 dynamic analysis를 통해 계산된 분당 심박수( beats per minutes, bpm)를 나타낸다. 분석 결과는 앱으로 다시 전송되어 사용자가 손쉽게 자신의 생체 데이터를 확인할 수 있다. 고막 및 심

음 질환이 없는 연구자를 대상으로 디지털 검이경-청진기 복합 장치를 활용하여 생체 신호를 수집하고, 이를 통해 연구자들의 정상 상태 여부를 판별하였다. 3 명의 연구자들의 고막 이미지 및 심음을 총 3 회 측정하였으며, 총 9 개의 고막 이미지 및 심음 데이터에 대해 정상 여부를 인공지능으로 판별하는 검증 연구를 수행하였다. 고막 판별 정확도는  $89.9 \pm 3.64\%$ 로 나타났으며, 심음 판별 정확도는 100%로 나타났다.

#### IV. 고찰 및 결론

본 연구에서는 하나의 장치에 디지털 검이경-청진기 복합 기능을 구현하는 통합 플랫폼을 제작하고, 검이경 및 청진기 각각의 모드에서 사용자의 고막 이미지 및 심음을 수집 및 환자의 건강 이상 여부를 판별 보조할 수 있는 비대면 원격의료 플랫폼을 개발했다. 환자의 건강 이상여부를 판단하기 위해, 고막 이미지와 심음 데이터 각각에 대해 4가지 딥러닝 모델을 선정하였다. 고막 모델에서는 MobileNetV2가 Recall = 82.08%, Precision = 92.55%, F1 Score = 87.08%, Acc = 86.32%, AUC = 0.93으로 가장 높은 성능을 보였고, 심음 모델에서는 VGG 16이 모든 성능 지표에서 만점으로 가장 높

은 성능을 보였다. 두 모델을 바탕으로 Classification Model을 개발하고 이를 플랫폼 서버에 연동하여 질병을 판별해주는 디지털 헬스케어 플랫폼을 완성했다. 정상 연구자의 고막 이미지와 심음 데이터를 판별하는 검증을 수행하였을 때 정상고막 및 정상 심음을 89.90±3.64% 및 100%의 정확도로 판별했다. 비대면 디지털 검이경-청진기 원격의료 플랫폼을 활용하여 의료 소외 계층이나 의료 접근성이 낮은 지역에서도 환자의 데이터를 지속적으로 수집 관리할 수 있고, 인공지능 기술을 이용하여 상태를 실시간으로 진단 보조할 수 있을 것으로 기대된다.

## References

- [1] Jung KH, Torrone D, Lovinsky-Desir S, Perzanowski M, Bautista J, Jezioro JR, et al. Short-term exposure to PM(2.5) and vanadium and changes in asthma gene DNA methylation and lung function decrements among urban children. *Respir Res.* 2017;18(1):63.
- [2] Labarca G, Drake L, Horta G, Jantz MA, Mehta HJ, Fernandez-Bussy S, et al. Association between inflammatory bowel disease and chronic obstructive pulmonary disease: a systematic review and meta-analysis. *BMC Pulm Med.* 2019;19(1):186.
- [3] Schraufnagel DE, Balmes JR, Cowl CT, De Matteis S, Jung SH, Mortimer K, et al. Air Pollution and Noncommunicable Diseases: A Review by the Forum of International Respiratory Societies' Environmental Committee, Part 1: The Damaging Effects of Air Pollution. *Chest.* 2019;155(2):409-16.
- [4] Ma Y, Zhao C, Zhao Y, Lu J, Jiang H, Cao Y, et al. Telemedicine application in patients with chronic disease: a systematic review and meta-analysis. *BMC Med Inform Decis Mak.* 2022;22(1):105.
- [5] Wang H, Yuan X, Wang J, Sun C, Wang G. Telemedicine maybe an effective solution for management of chronic disease during the COVID-19 epidemic. *Prim Health Care Res Dev.* 2021;22:e48.
- [6] D'Onofrio KL, Zeng FG. Tele-Audiology: Current State and Future Directions. *Front Digit Health.* 2021;3:788103.
- [7] Dendere R, Myburg N, Douglas TS. A review of cellphone microscopy for disease detection. *J Microsc.* 2015;260(3):248-59.
- [8] Mousseau S, Lapointe A, Gravel J. Diagnosing acute otitis media using a smartphone otoscope; a randomized controlled trial. *Am J Emerg Med.* 2018;36(10):1796-801.
- [9] Hafke-Dys H, Kuznar-Kaminska B, Grzywalski T, Maciaszek A, Szarzynski K, Kocinski J. Artificial Intelligence Approach to the Monitoring of Respiratory Sounds in Asthmatic Patients. *Front Physiol.* 2021;12:745635.
- [10] Kim Y, Hyon Y, Lee S, Woo SD, Ha T, Chung C. The coming era of a new auscultation system for analyzing respiratory sounds. *BMC Pulm Med.* 2022;22(1):119.
- [11] Kim Y, Hyon Y, Woo SD, Lee S, Lee SI, Ha T, et al. Evolution of the Stethoscope: Advances with the Adoption of Machine Learning and Development of Wearable Devices. *Tuberc Respir Dis (Seoul).* 2023;86(4):251-63.
- [12] McDaniel NL, Novicoff W, Gunnell B, Cattell Gordon D. Comparison of a Novel Handheld Telehealth Device with Stand-Alone Examination Tools in a Clinic Setting. *Telemed J E Health.* 2019;25(12):1225-30.
- [13] Cha D, Pae C, Seong SB, Choi JY, Park HJ. Automated diagnosis of ear disease using ensemble deep learning with a big otoendoscopy image database. *EBioMedicine.* 2019;45:606-14.
- [14] Chen J, Sun K, Sun Y, Li X. Signal Quality Assessment of PPG Signals using STFT Time-Frequency Spectra and Deep Learning Approaches. *Annu Int Conf IEEE Eng Med Biol Soc.* 2021;2021:1153-6.
- [15] Goto R, Horimoto T, Koyama S, Suzuki T, Tsutsumi J, Matsuyama T, et al. Detection of Heartbeat Components Through Doppler Radar Systems Using Semantic Segmentation and Non-Harmonic Analysis. *IEEE Access.* 2024;12:32349-60.
- [16] Khan MA, Kwon S, Choo J, Hong SM, Kang SH, Park IH, et al. Automatic detection of tympanic membrane and middle ear infection from oto-endoscopic images via convolutional neural networks. *Neural Netw.* 2020;126:384-94.
- [17] Park YS, Jeon JH, Kong TH, Chung TY, Seo YJ. Deep Learning Techniques for Ear Diseases Based on Segmentation of the Normal Tympanic Membrane. *Clin Exp Otorhinolaryngol.* 2023;16(1):28-36.
- [18] Yadav K, Tiwari S, Jain A, Dafhalla AKY. Deep learning based cardiovascular disease diagnosis system from heartbeat sound. *International Journal of Speech Technology.* 2021.
- [19] Rao D, Singh R, Kamath SK, Pendekanti SK, Pai D, Kolekar SV, et al. OTONet: Deep Neural Network for Precise Otoscopy Image Classification. *IEEE Access.* 2024;12:7734-46.
- [20] Chowdhury MEH, Khandakar A, Alzoubi K, Mansoor S, A MT, Reaz MBI, et al. Real-Time Smart-Digital Stethoscope System for Heart Diseases Monitoring. *Sensors (Basel).* 2019;19(12).
- [21] Quiceno AF, Delgado E, Vallverd M, Matijasevic AM, Castellanos-Domnguez G. Effective phonocardiogram segmentation using nonlinear dynamic analysis and high-frequency decomposition. In: 2008 Computers in Cardiology. *IEEE;* 2008.
- [22] Lee K, Ji Y, Jeon Y, Park YC. Development and Implementation of Noise-Canceling Technology for Digital Stethoscope. *Journal of Biomedical Engineering Research.* 2013;34(4):204-11.
- [23] Tympanic membrane / eardrum dataset / otitis media. Published online June 6, 2022.
- [24] King E. Heartbeat Sounds. Published online November 27, 2016.
- [25] Akay M, Du Y, Sershen CL, Wu M, Chen TY, Assassi S, et al. Deep Learning Classification of Systemic Sclerosis Skin Using the MobileNetV2 Model. *IEEE Open J Eng Med Biol.* 2021;2:104-10.
- [26] Dong N, Zhao L, Wu CH, Chang JF. Inception v3 based cervical cell classification combined with artificially extracted features. *Applied Soft Computing.* 2020;93.
- [27] Guefrechi S, Jabra MB, Ammar A, Koubaa A, Hamam H. Deep learning based detection of COVID-19 from chest X-ray images. *Multimed Tools Appl.* 2021;80(21-23):31803-20.
- [28] Younis A, Qiang L, Nyatega CO, Adamu MJ, Kawuwa HB. Brain Tumor Analysis Using Deep Learning and VGG-

- 16 Ensembling Learning Approaches. *Applied Sciences*. 2022;12(14).
- [29] Ma N, Zhang X, Zheng H-T, Sun J. ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. *Computer Vision – ECCV 2018. Lecture Notes in Computer Science* 2018. p. 122-38.
- [30] Ji Q, Huang J, He W, Sun Y. Optimized Deep Convolutional Neural Networks for Identification of Macular Diseases from Optical Coherence Tomography Images. *Algorithms*. 2019;12(3).