



ISSN: 3022-5388

JKAI website: <https://accesson.kr/jkaia>DOI: <http://doi.org/10.24225/jkaia.2024.2.1.25>

머신러닝을 활용한 코다이 학습장치의 인식을 변화

Changes in the Recognition Rate of Kodály Learning Devices using Machine Learning

YunJeong LEE¹, Min-Soo KANG², Dong Kun CHUNG³

Received: April 15, 2024. Revised: June 14, 2024. Accepted: June 14, 2024.

Abstract

Kodály hand signs are symbols that intuitively represent pitch and note names based on the shape and height of the hand. They are an excellent tool that can be easily expressed using the human body, making them highly engaging for children who are new to music. Traditional hand signs help beginners easily understand pitch and significantly aid in music learning and performance. However, Kodály hand signs have distinctive features, such as the ability to indicate key changes or chords using both hands and to clearly represent accidentals. These features enable the effective use of Kodály hand signs. In this paper, we aim to investigate the changes in recognition rates according to the complexity of scales by creating a device for learning Kodály hand signs, teaching simple Do-Re-Mi scales, and then gradually increasing the complexity of the scales and teaching complex scales and children's songs (such as "May Had A Little Lamb"). The learning device utilizes accelerometer and bending sensors. The accelerometer detects the tilt of the hand, while the bending sensor detects the degree of bending in the fingers. The utilized accelerometer is a 6-axis accelerometer that can also measure angular velocity, ensuring accurate data collection. The learning and performance evaluation of the Kodály learning device were conducted using Python.

Keywords : Kodály hand symbol, Scale complexity, Recognition ratio, Artificial Intelligence, etc

Major Classification: Machine Learning

1. Introduction

졸탄 코다이 (Zoltan Kodály, 1882~1926)는 음악 교육을 통해 모든 학생에게 음악을 읽고 쓸 수 있는 능력을 길러 줌으로써 음악성을 개발하고 그것을 바탕으로 학생 개개인이

음악을 통해 창조적인 활동을 영위하며, 더 나아가 바람직한 민족 음악 문화를 창건할 수 있도록 이끌어야 한다고 믿었던 음악 교육가이다(Choi,2009).

코다이 손기호는 사람의 신체 중에서 가장 손쉽게 표현할 수 있고 음악을 처음 접하는 어린이들에게 흥미를

1 First Author. Medical IT, Eulji University, Korea. Email: leeyunjeong010422@gmail.com

2 Second Author. Professor, Medical IT, Eulji University, Korea, Email: mskang@eulji.ac.kr

3 Co-Author. Professor, Medical IT, Eulji University, Korea, Email: tchung@eulji.ac.kr

© Copyright: The Author(s)

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

유발할 수 있는 좋은 재료로, 공간상의 음높이와 계이름을 확실하게 나타내어 정확한 음정 개념이 형성되도록 한다(Han, 2010).

한편, 모션 인식(Motion recognition or Gesture recognition)은 수학적 알고리즘을 통해서 사람의 동작을 해석하는 것을 목표로 하여 컴퓨터와 인간의 상호작용을 실현하는 기술이다. 이에는 손동작에 관한 인식 기술도 포함된다(Ahn,2019, Costante,2014, Kim,2013). 이들 대부분의 연구들은 손동작을 감지하거나 제어하는 것에 초점을 맞추고 있다.

손동작에 관한 연구는 접근 방식에 따라 크게 두 가지로 구분할 수 있을 것이다. 첫 번째는 손에 부착된 센서를 사용하여 손의 움직임 감지하는 방법이 있으며, 두 번째로는 영상 기반 방법으로 카메라 만을 활용하여 별도의 부가 장치 없이, 영상 분석으로 구현하는 것이다. 영상 분석을 통해 손가락의 모션을 감지하는 경우에, 실제 현장에서는 조명의 영향을 받을 뿐 더러, 촬영 장비와 피사체 간의 장애물 등 여러 영향을 받게 되므로 자칫 손동작 분석률이 떨어진다는 문제를 안게 된다. 본 논문에서는 음악 분야에서의 활용 가능성을 염두에 두고, 손과 손가락에 부착된 센서를 사용하는 방법을 통해, 손동작에 관한 연구를 진행시킴으로써, 편리하고 효율적으로 코다이 손기호를 학습하고, 실제 음악 수업에 손가락의 움직임을 인식하여, 학업에 활용할 수 있는 코다이 학습장치를 구현하고자 한다.

해당 장치를 활용하여 먼저 간단한 도레미 음계를 학습시킨 후에 순차적으로 음계의 복잡도를 증가시켜 복잡한 음계 및 동요(비행기)를 학습시킴으로써 음계의 복잡도에 따른 인식을 변화를 탐구하고자 한다. 또한, 인식이 감소하는 경우, 이를 극복하고 인식을 높일 수 있는 방법을 찾아보고자 한다.

2. Related Research

2.1. Kodai Hand Symbol

손기호는 초보자들이 음의 높이를 쉽게 이해할 수 있도록 돕는 동시에, 음악 학습과 음악 연주에 많은 도움을 주는 도구이다. 이러한 손기호는 원래 19세기 중반 영국의 커웬 (John Spencer Curwen, 1816-1880)에 의해 개발되었다. 하지만 코다이 손기호는 두 손을 사용하여 조바꿈을 나타내거나 화음을 표시할 수 있으며, 임시 음들도 명확하게

표시할 수 있는 등 차별적인 특징을 가지고 있다. 이러한 특징들이 코다이 손기호의 효과적인 활용을 가능하게 한다.

코다이 손기호에 관련하여 제안된 다양한 국내 연구를 살펴보자면, 김정미는 음악 능력의 기초라 할 수 있는 청음력을 신장시키기 위해 코다이의 교수법을 활용한 청음 학습 프로그램을 구안하여 초등학교 6학년 아동들에게 적용함으로써 그 효과를 검증하였다(Kim,2004). 적용 결과, 코다이 교수법은 청음력 신장에 효과적임을 알 수 있었고, 특히 듣고 적는 능력에 조금 더 효과적임을 알 수 있었다. 강문진은 코다이 교수법이 유아의 음악성에 어떠한 효과를 미치는지에 대해 연구하였다(Kang,2009). 코다이 교수법으로 수업한 유아들은 리듬감뿐만 아니라 음정감에서도 향상되었음을 알 수 있었다. 음정감과 리듬감의 향상은 음악적인 기초 성장 단계로 이어져 성공적인 음악성 개발이 가능함을 의미한다. 마지막으로 한정수는 코다이 교수법을 활용한 가창지도가 장애아동의 음악 청취력에 미치는 효과에 대해 연구하였다(Han,2010). 코다이 교수법을 활용하여 가창 지도를 받은 장애아동은 일반적인 방법으로 가창 지도를 받은 장애아동보다 리듬감 청취력과 음정감 청취력에서 더 높은 향상도가 나타났다.

2.2. Gesture Recognition

기존에 손동작 인식 기술에 관한 다양한 연구들이 제안되었다. 몇가지 살펴보자면 Zhou Ren은 사람의 신체부위 및 움직임을 인식할 수 있는 키넥트(Kinect) 센서를 사용하여 연구하였다(Ren, 2013). 이 센서의 노이즈를 없애기 위해 Finger-Earth Mover's Distance(FEMD)라는 방식을 제안하였는데, 이 방식은 손바닥을 제외하고 손가락의 모양을 인식하여 손동작을 구분하여, 보다 더 세밀하고 작은 손동작의 차이를 인식할 수 있었다. Costante는 스마트워치를 기반으로 한 제스처 인터페이스 시스템을 제안했다(Costante,2014). Costante는 사용자 개인에 맞춘 제스처 인식에 초점을 맞췄고 transfer metric learning에 기반한 알고리즘을 디자인 했다. Yu는 비전 시스템 기반의 동작 인식 시스템을 제안하였고 동작을 분류하기 위하여 다층 퍼셉트론을 이용하였다(Yu,2010). Yu가 사용한 다층 퍼셉트론(Multi Layer Perceptron)은 97.4%의 정확도를 보였다. 안현진은 스마트워치의 가속도 센서에서 취득된 데이터를 인공지능망을 통해 처리하여 사용자의 손동작을 인식하고 분류하여 문을 열기 위한 손동작의 정확도 인식을 분석하였다(Ahn,2019). 다층 퍼셉트론을 사용하고 입력 값의

길이를 동일하게 만드는 전처리 과정에서는 아키마 스플라인 알고리즘을 사용하였다. 문에 따라 차이가 있었지만 73.3%-98.5%의 정확도를 보였다.

이러한 연구들을 바탕으로 코다이 교수법은 장애 아동뿐 아니라 아동, 학생들에게 음악적인 표현을 자극하고 리듬과 음정 표현을 증진시켜주는 교수법이라 할 수 있으며, 이러한 코다이 교수법에 머신러닝을 활용한 장치를 제공함으로써, 아동 및 학생들에게 능동적으로 학업에 참여할 수 있는 기회를 제공한다면, 개개인에 대한 성공적인 음악성 개발로 이어질 것으로 사료된다.

본 논문에서는 이 코다이 손기호를 학습할 수 있는 학습장치를 구현하였고, 해당 장치를 학습시켜 순차적으로 음계의 복잡도를 증가시켜 음계의 복잡도에 따른 인식률 변화를 연구하고자 한다. 또한, 인식률이 감소하는 경우에는 이를 극복하고 인식률을 높일 수 있는 방법을 찾아보고자 한다.

3. Methods

3.1. Implementation of Kodai Learning Device

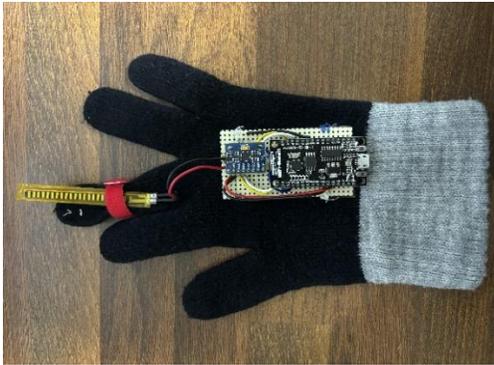


Figure 1: Kodály Learning Device

이 논문에서 활용된 코다이 학습장치는 가속도 센서와 구부림 센서를 이용하여 측정된 값을 분석함으로써 손기호를 식별한다. 가속도 센서는 손의 기울기를 감지하며, 구부림 센서는 손가락의 구부러짐 정도를 감지한다. 이용된 가속도 센서는 6축 가속도 센서로 각속도까지 측정할 수 있어 정확한 데이터를 확보할 수 있다. 전체 음계 중, 구현의 가능성을 살펴보는데 그 목적을 두고 있으므로, 음계는 도,레,미 3개음으로 한정하였다. 도, 레, 미 이 세 개 계이름의 손기호 센서값을 측정하고 그 값을 이용하여

학습을 진행하였으며, 해당 결과를 통해 코다이 학습장치의 인식률을 평가하였다.

특히, 손기호를 학습하고 손기호를 인식하는 장치의 구현에 있어 주요 요소는 휴대성을 동반한 소형화이다. 센서의 값을 읽고 이에 따른 학습을 진행시키며, 학습된 내용을 기억하여 실제 음악 수업에 활용토록 하기 위해, 32비트 MCU에 2.4GHz Wi-Fi, 그리고, 블루투스 5 가 장착된 소형 프로세서를 선택하였는데, 본 연구에서는 Espressif사의 ESP32 모듈을 사용하였다.

3.2. How to Collect and Learn Data

Algorithm 1: Part of the data learning course source code

```

Input : Dataset, Number of Folds  $n\_folds$ , Learning Rate  $l\_rate$ ,
        Number of Epochs  $n\_epoch$ , Number of Hidden Units
         $n\_hidden$ 
Output: List of Scores
 $n\_folds \leftarrow 5$ ;
 $l\_rate \leftarrow 0.3$ ;
 $n\_epoch \leftarrow 500$ ;
 $n\_hidden \leftarrow 5$ ;
 $scores \leftarrow$ 
    evaluate.algorithm(dataset, back.propagation,  $n\_folds$ ,  $l\_rate$ ,  $n\_epoch$ ,  $n\_hidden$ );

print('Scores: %s' % scores);
print('Mean Accuracy: %3f%
```

Figure 2: Data Collection Source Code (Pseudo code)

데이터 수집은 ms단위로 현재 시간을 측정하고, 이전 시간과의 차이를 계산하여 각 계산 간격을 조정하였다. 상수 alpha는 각속도 값을 보정하는데 사용되었으며, 이를 기반으로 각속도 값을 보정하여 temp_x 및 temp_y를 계산하였다. 이러한 값들은 센서 데이터의 정확성을 높이기 위해 사용되었다. 마지막으로, 이전 각도 및 보정된 각속도 값을 기반으로 현재 각도를 계산하여 출력하였다.

이러한 과정을 통해 시간에 따른 각속도 값을 각도로 변환하여 정확한 기울기를 계산할 수 있었으며, 센서의 데이터를 읽고, 각속도를 계산하여 시간에 따른 각도를 추정하는데 사용되고, 각 도, 레, 미에 대한 손기호 데이터를 수집할 수 있었다.

데이터 학습은 역전파 알고리즘을 사용하여 데이터를 학습하는 함수(그림 3)를 사용하였다. 이 함수는 주어진 훈련 데이터를 사용하여 신경망을 초기화하고 학습하는 과정을 수행한다. 학습률('l_rate'), 에포크 수('n_epoch'), 및 은닉층의

크기('n_hidden') 등의 매개변수를 통해 학습 과정을 조정할 수 있다. 이 함수를 통해 훈련된 결과를 바탕으로 테스트 데이터에 대한 예측값을 생성하여 반환하였다. 이 예측값은 신경망이 학습한 모델을 기반으로 구축되며, 이 모델을 사용하여 데이터를 분류하거나 예측한다.

데이터를 학습하는 과정은 다음과 같이 수행하였다. 먼저 CSV 파일을 불러온 후, 데이터를 적절한 형태로 준비한 후, 문자열 형태의 열을 실수로 변환하고, 클래스 열을 정수로 변환하는 과정을 거쳤다. 이후 입력 변수를 정규화하여 데이터의 스케일을 조정하였다.

Algorithm 1: Part of the data learning course source code

```

Input : Current Time curtime, Old Time oldtime, Last Pitch Angle last_pi_x, Last Roll Angle last_ro.y, IMU Sensor Data sd, Correction Factor alpha
Output: Updated Pitch Angle angle_x, Updated Roll Angle angle_y
curtime ← millis();
stime ← curtime - oldtime;
temp_x ←  $-\frac{1}{\alpha} \cdot (-pi_x + last\_pi\_x) + sd.GyX$ ;
temp_y ←  $-\frac{1}{\alpha} \cdot (-ro.y + last\_ro.y) + sd.GyY$ ;
angle_x ← last_pi_x + temp_x · stime · 0.001;
angle_y ← last_ro.y + temp_y · stime · 0.001;
last_pi_x ← angle_x;
last_ro.y ← angle_y;
oldtime ← curtime;
return angle_x, angle_y;

```

Figure 3: Data Training Process Source Code (Pseudo code)

학습 과정에 K-Fold 교차검증(K=5)을 실시하였고, 이때의 학습률은 0.3으로 설정하였다. 또한, 500번의 에포크와 5개의 은닉층을 설정하여 신경망 모델을 초기화하였고, 역전파 알고리즘을 사용하여 학습을 진행하였다. 이를 통해 훈련된 모델을 기반으로 테스트 데이터에 대한 예측값을 생성하고, 이를 통해 인식률을 평가하였다. 최종적으로는 각 폴드에서 계산된 점수를 평균 내어 전체적인 평가 점수를 도출하였다. 이 과정을 통해 데이터를 학습하고, 해당 모델을 활용하여 정확한 분류와 예측을 수행할 수 있음을 확인하였다.

4. Results

그림 5의 D-R-M은 가장 기본적인 '도,레,미', 세 개의 음을 그림 4의 변동구간 S를 포함하여 한 번씩 연속적으로 학습시켰을 경우이며, 전체 인식률은 96.9%로 산출되었다.

변동구간 S를 제외하고 도, 레, 미를 각각 개별적으로 학습시켰을 때 도는 99.6%, 레는 99.1%, 미는 98.7%로 음을 연속적으로 학습시켰을 때 보다 더 높은 인식률이 나타났다.

이전 결과보다 계이름 별로 학습시킨 경우 더 높은 인식률이 나타나는 이유를 확인하고자 여러 가설을 고려하였다. 그 결과, 연속적으로 학습시킨 경우에는 음계 사이에서 결측값이 발생하기 때문에, 인식률이 감소할 수 있다는 가설이 나왔다. 이 가설이 사실로 확인된다면, 음계의 복잡도가 증가하거나 결측값의 종류가 다양해짐에 따라 인식률이 저하될 것이다.

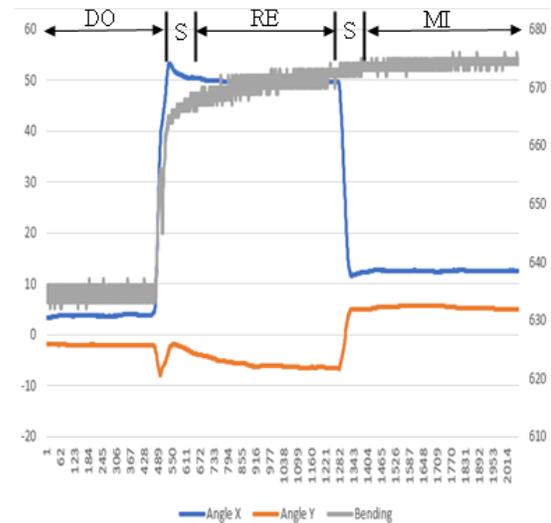


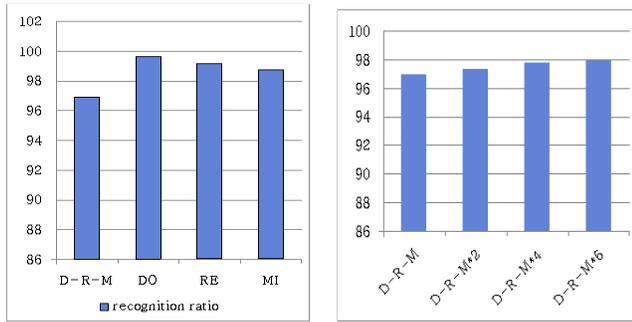
Figure 4: Data from Do-Re-Mi Hand Gestures

도, 레, 미를 연속적으로 여러 번 학습시켰을 경우에는 초기 한 번 학습시켰을 때 보다 모델이 결측값들을 더 정확하게 학습할 수 있기 때문에 인식률은 증가할 것이라고 예상하였다. 결측값의 종류가 다양해지는 것이 아니라 유사한 결측값의 데이터 수가 많아지는 것이기 때문이다. 이 가설을 확인하기 위해 도, 레, 미를 연속적으로 2번, 4번, 6번으로 나누어서 학습시켰다.

그 결과 2번 반복했을 때 97.3%, 4번 반복했을 때 97.7%, 6번 반복했을 때 97.9%로 초기 인식률보다 증가하였다. 반복 횟수를 더욱 증가시킨 결과, 8번 반복했을 때 98.1%, 10번 반복했을 때 98.2%의 인식률이 산출되었으며, 반복 횟수가 계속 증가함에 따라 인식률의 증가 폭은 점차 감소하는 경향을 보였다. 이러한 결과는 반복 횟수가 증가함에 따라 인식률이 일정 수준 이상에서는 더 이상 크게 향상되지 않는다는 것을 보여준다.

위 그림 6의 결과가 보여주는 인식을 향상이 결측치의 학습 인식을 증가로 전체 인식이 증가가 된 것이 맞는지 정확하게 확인하기 위해 결측값의 학습 인식을 고려하여 비교하였다.

그림 5의 결과가 보여주는 인식을 향상이 결측치의 학습 인식을 증가로 전체 인식이 증가가 된 것이 맞는지 정확하게 확인하기 위해 결측값의 학습 인식을 고려하여 비교하였다.



(a) Each (b) Repeated
Figure 5: Recognition Rate Comparison of Do-Re-Mi

도, 레, 미를 한 번 학습시켰을 경우에 결측값 인식을 97%, 연속적으로 여러 번 학습시켰을 경우에는 98.2%로 여러 번 학습시켰을 경우에 결측값을 더 정확하게 인식했음을 보여준다.

도, 레, 미를 한 번만 학습시켰을 때 보다 여러 차례 반복하여 학습시켰을 경우에 인식이 증가하는 이유를 전체 학습 과정 중에서 음계 사이의 결측값을 학습할 때 다수의 유사한 결측값이 존재하여 결측값을 더욱 정확하게 학습할 수 있기 때문이라고 해석하였다.

이러한 증가된 인식은 데이터 학습 과정에서의 결측값들을 반복적으로 학습시킴으로써 모델이 음계 사이 결측값의 연속적인 패턴을 더욱 정확하게 이해할 수 있다는 결론을 얻을 수 있다.

앞에서는 유사한 결측값이 다수 존재할 때 인식이 증가한다는 것을 분석하였다. 다음으로는 음계의 복잡도가 증가함에 따라 인식이 저하되는지 즉, 결측값의 종류가 다양해진 경우, 인식이 저하되는지에 대한 결과를 분석하였다. 더 다양한 도, 레, 미의 연주를 고려하여 가능한 6가지 경우(도미레, 레도미, 레미도 등)를 학습시켰다.

이에 따른 결과는 초기 학습 인식률보다 감소한 93.8%의 인식이 측정됐다. 이에 이어 동요'비행기'를

연주하여 해당 데이터를 학습시킨 결과, 현재까지 측정된 인식을 중에서 가장 낮은 수치인 90.3%가 산출되었다. 이러한 결과는 복잡한 음계가 모델의 학습 인식률에 미치는 영향을 명확하게 보여준다.

따라서 인식이 복잡도에 따라 감소하는 이유는 결측값의 다양성이 증가하기 때문으로 해석하였다. 6가지의 경우의 수를 학습시켰을 때 보다 동요'비행기'를 학습시켰을 때 나오는 결측값의 종류가 더 다양하고 복잡하다. 다시 말해, 결측값이 복잡하면 인식이 감소하고, 결측값이 유사한 값 다수로 구성되어 있을 때는 인식이 증가한다고 볼 수 있다.

하지만 감소된 인식을 높이기 위한 방법도 고려해야 한다. 은닉층 구성을 조정하거나 다양한 학습 매개변수를 조정하여 인식을 향상시킬 수 있을 것이라고 예상하였다. 따라서 본 논문에서는 은닉층 구성을 변경하여 인식률의 변화를 확인하였다.

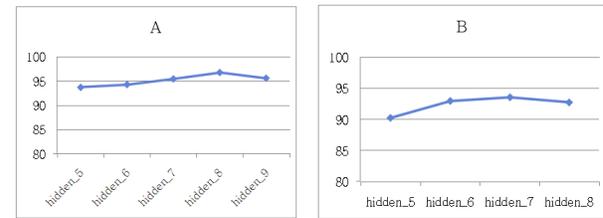


Figure 6: Recognition Rate Changes with Different Hidden Layers

그림 6에 나와 있는 두 개의 그림을 살펴보면, 은닉층에 변화에 따라 인식률도 변화하는 것을 확인할 수 있다. A에서는 은닉층이 8일 때, B에서는 은닉층이 7일 때 인식이 가장 높게 나타났다. 하지만 은닉층의 수를 계속해서 증가시키는 것이 인식을 계속해서 향상시키는 것은 아니다.

5. Conclusion

본 논문에서는 가속도 센서와 구부림 센서를 활용하여 코다이 손기호를 학습하고 인식할 수 있는 장치를 개발하였다. 개발한 코다이 손기호 학습장치를 이용하여 복잡도에 따른 인식률 변화를 분석하고, 감소된 인식률을 높이는 방법을 제안하였다. 복잡한 음계와 동요를 학습시키는 과정에서 인식이 감소하였고, 이는 결측치의 다양성과 복잡성이 인식률에 미치는 영향을 나타냈다. 또한, 특정 은닉층의 구성 변화로 인식을 향상 확인하였다.

음악 교육 및 음악 분야에서의 손동작 인식 기술에 대한 이해를 높일 것으로 기대되며, 정확하고 안정적인 손동작 인식 시스템을 구축하는데 큰 도움이 될 것이다.

향후 연구에서는 은닉층 구성 변경뿐 아니라 다른 방법을 통해 인식률을 향상시킬 수 있는 방법을 연구하고, 이를 통해 실제 응용 환경에서의 성능 향상과 함께 정확하고 신뢰할 수 있는 손동작 인식 시스템을 구축할 수 있도록 보완해 나갈 예정이다.

References

- Ahn, H. J. (2019). *Door opening hand gesture recognition system using artificial neural networks* (Master's thesis). University of Seoul, Graduate School.
- Choi, S. G. (2009). *A study on effective teaching methods of Sijo singing using the Kodaly method* (Master's thesis). Yonsei University, Graduate School of Education.
- Costante, G., Porzi, L., Lanz, O., Valigi, P., & Ricci, E. (2014). Personalizing a smartwatch-based gesture interface with transfer learning. *In Proceedings of Signal Processing Conference (EUSIPCO)* (pp. 2530–2534).
- Han, J. S. (2010). *The effect of singing instruction using the Kodály method on the music listening ability of children with disabilities* (Master's thesis). Korea National University of Education, Graduate School of Education.
- Kang, M. J. (2009). *The effect of the Kodály method on young children's musicality* (Master's thesis). Chongshin University.
- Kim, J. H. (2013). Implementation of a real-time number recognition system based on hand gesture using gradient and position information. *Journal of KIISE: Software and Data Engineering*, 2(3), 199-204.
- Kim, J. M. (2004). *The impact of the Kodály method on improving aural skills* (Master's thesis). Gyeongin National University of Education, Graduate School of Education.
- Ren, Z., Yuan, J., Meng, J., & Zhang, Z. (2013). Robust part-based hand gesture recognition using Kinect sensor. *IEEE Transactions on Multimedia*, 15(5), 1110–1120.
- Song, G. J. (2006). *A study on Zoltan Kodaly's music education method* (Master's thesis). Mokpo National University, Graduate School of Education.
- Yu, C., Wang, X., Huang, H., Shen, J., & Wu, K. (2010, October). Vision-based hand gesture recognition using combinational features. *In Proceedings of 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing* (pp. 543–546).