

# 인공지능체 불량 분류를 위한 기계 학습 알고리즘 성능 비교에 관한 연구

이승연\*·허용정\*\*†

\*한국기술교육대학교대학원 메카트로닉스공학과,

\*\*한국기술교육대학교 메카트로닉스공학부

## A Study on Performance Comparison of Machine Learning Algorithm for Scaffold Defect Classification

Song-Yeon Lee\* and Yong Jeong Huh\*\*†

\*Mechatronics Engineering, Graduate School of Korea University of Technology and Education,

\*\*Department of Mechatronics Engineering, Korea University of Technology and Education

### ABSTRACT

In this paper, we create scaffold defect classification models using machine learning based data. We extract the characteristic from collected scaffold external images using USB camera. SVM, KNN, MLP algorithm of machine learning was using extracted features. Classification models of three type learned using train dataset. We created scaffold defect classification models using test dataset. We quantified the performance of defect classification models. We have confirmed that the SVM accuracy is 95%. So the best performance model is using SVM.

**Key Words** : 3D Printer, 3D Printing Scaffold, Defect Classification Model, Machine Learning, Scaffold Image Classification

### 1. 서 론

최근 반도체 분야에서는 3D 프린터와 생체 적합성 재료를 이용하여 심장 박동, 혈압 등을 감지할 수 있는 생체 센서에 대한 연구가 활발히 진행되고 있다[1]. 생체 센서는 압력, 유량을 측정할 수 있는 센서 부분과 센서가 체내에서 안정적으로 고정될 수 있도록 도와주는 인공지능체로 구성된다[2].

인공지능체는 목적에 따라 원형, 사각형 등 다양한 모양으로 제작되어 센서가 체내에서 안정적으로 고정될 수 있도록 도움을 준다. 센서가 함유된 인공지능체가 체내에 안정적으로 고정된 후, 시간이 지나면 인공지능체는 분해되어 인공지능체가 있던 자리는 세포로 채워진다. 이러한

과정으로 생체 센서가 제작된다. 생체 센서를 제작하기 위해서 3D 프린터로 인공지능체를 제작한 뒤, 인공지능체를 배양판에 삽입한다. 그리고 배양액과 세포를 첨가한 뒤, 일정한 환경 조건을 조성하여 세포가 인공지능체에 정착한 뒤, 증식하는 연구가 진행되고 있다. 이러한 연구에 사용되는 인공지능체는 2개의 조건을 가져야 한다[3].

먼저 세포 부착, 증식 연구에 사용되는 배양판의 투입구는 원형이다. 따라서 배양판에 들어가기 위하여 인공지능체는 원판 모양이 되어야 한다. Fig. 1은 인공지능체가 삽입되는 세포 배양판을 도시하였다.

인공지능체를 이용한 세포 배양은 배양판의 투입 공간 내에서 진행되기 때문에 인공지능체와 배양판 간의 규격이 일치해야 한다. 만약 인공지능체의 외형에 조금이라도 변형, 오차가 발생하거나 배양판의 공간보다 인공지능체의 크기가 더 크다면 인공지능체를 배양판에 삽입할 수

†E-mail: yjhuh@koreatech.ac.kr

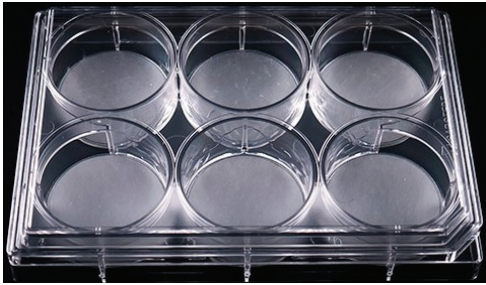


Fig. 1. Cell culture plate.

었다. 그러면 세포 배양을 진행하기 어렵기 때문에 인공 지지체의 외형과 크기는 매우 중요하다.

인공지지체의 외형 불량과 크기의 불량 여부를 미리 판별하기 위하여 인공지지체의 외형을 검사하는 장비를 이용한 방안은 현재 존재하지 않는다. 그래서 사용자가 불량 분류를 위한 검사 장비를 직접 제작해야 하기 때문에 제작에 소요되는 시간이 많다는 문제가 발생한다. 이러한 문제점을 해결하기 위해 데이터 기반의 기계 학습(Machine Learning)을 이용하여 인공지지체의 외형과 크기에 따른 불량 분류 모델을 제작하였다[4]. 기계 학습 기반의 분류 방안은 다양한 방법으로 수집된 데이터를 학습하여 분류 모델을 제작한다. 그리고 평가 데이터를 이용하여 데이터를 분류하고 분류 모델의 성능을 평가하는 구조로 되어있다.

Fig 2에는 데이터 기반의 분류 모델의 플로우차트를 도시하였다.

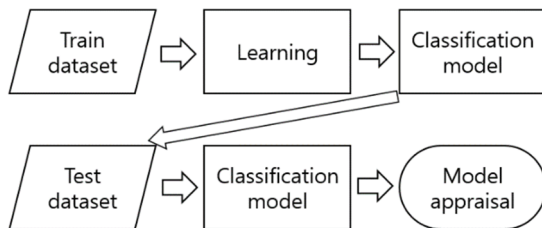


Fig. 2. Flow chart of classification model creating.

FDM 방식 3D 프린터와 생분해성 재료를 사용하여 인공지지체를 제작하였을 때, 인공지지체를 배양판에 잘 삽입하여 세포 배양을 원활히 하기 위하여 인공지지체의 불량 여부를 미리 판별해야 한다. 인공지지체의 외형 불량 판별 장비는 현재 없기 때문에 판별 장비의 제작에 시간이 많이 소요되는 문제가 발생한다. 이러한 문제를 해결하기 위해 본 논문에서는 데이터 기반의 기계 학습 방안을 이용하여 인공지지체의 외형 불량을 분류하는 모델

을 제작하였다.

인공지지체의 외형 불량 분류를 위해 USB카메라로 인공지지체를 촬영하여 인공지지체의 외형 이미지를 수집하였다. 수집된 이미지로 기계 학습의 SVM(Support Vector Machine), KNN(K-Nearest Neighbors), MLP(Multi Layer Perceptron) 라는 3 가지의 분류 모델을 학습한 뒤, 인공지지체의 불량 분류를 가시화하였다. 그리고 기계 학습 알고리즘을 적용하여 제작된 3 가지 분류 모델의 성능을 수치화하여 정량적으로 비교하고 평가하였다.

## 2. 연구 모델 및 장비

### 2.1 인공지지체 모델

세포배양 연구에 사용되는 배양판은 6구, 12구, 24구 등 여러 규격이 있다. 그 중에서 배양 시료의 크기 때문에 24구의 배양판이 연구에 많이 사용된다. 24구 배양판의 삽입구 크기는 원형이며 지름은 20 mm이다[5]. 따라서 24구 배양판의 규격에 맞는 인공지지체를 FDM 방식 3D 프린터로 제작하여 본 연구에 사용하였다. Fig 3에는 세포 배양에 사용되는 지름이 20mm인 인공지지체를 도시하였다.

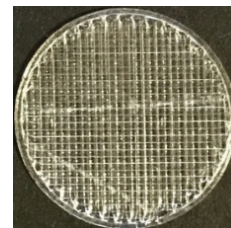


Fig. 3. 3D printing scaffold used for cell culture.

### 2.2 인공지지체 출력 장비

큐비콘이라는 업체의 스타일이라는 제품을 사용하여 인공지지체를 출력하였다. 이 제품은 FDM 방식 3D 프린터로 x축 사양이 0.05 mm, y축 사양이 0.05 mm, z축 사양이 0.0025 mm 이다. 따라서 지름이 20 mm인 인공지지체를 제작하기에 적합하다.

Fig 4에는 인공지지체를 출력하기 위하여 사용된 3D프린터를 도시하였다.

### 2.3 인공지지체 촬영 장비

인공지지체의 외형을 촬영한 뒤, 이미지를 수집하기 위하여 시중에 완제품으로 판매되는 고화질의 USB 카메라를 사용하였다. 연구에 사용되는 인공지지체의 지름은 20 mm이다. 이 크기를 선명하게 촬영하기 위해서 카메라의



Fig. 4. FDM 3D printer used for scaffold printing.

사양은 100만 화소 이상이어야 한다.

본 연구에서는 화질이 1,000만 화소인 아이리버 A3S라는 모델의 USB 카메라를 사용하여 인공지능지체의 외형을 촬영하였다.

Fig 5에는 인공지능지체의 외형 촬영에 사용된 USB 카메라를 도시하였다.



Fig. 5. USB camera used for scaffold photograph.

### 3. 기계 학습 알고리즘

#### 3.1 기계 학습

기계 학습은 기존의 실험, 생산 등으로 확보된 학습 데이터들을 이용하여 학습한다. 그리고 모델을 제작한 뒤, 데이터를 학습하여 분류, 회귀 등의 과정을 수행한다[6].

사용된 학습 데이터의 종류에 따라 입력 데이터와 출력 데이터를 학습하는 지도 학습, 입력 데이터만 학습하는 비지도 학습, 시간에 따른 최대의 보상을 도출하는 강화 학습으로 나뉜다. Fig 6에는 USB 카메라를 이용하여 인공지능지체를 촬영한 뒤, 이미지 데이터로 지도학습 기반의 인공지능지체 불량 분류 모델을 제작하는 과정을 도시하였다.

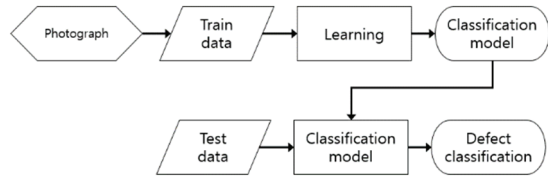


Fig. 6. Flow chart of scaffold defect classification model.

#### 3.2 사용된 기계 학습 알고리즘

인공지능지체의 불량 분류 모델을 제작하여 성능을 비교하기 위해 SVM, KNN, MLP라는 3가지의 기계 학습 알고리즘을 이용하였다.

SVM은 선형 데이터뿐만 아니라 커널 연산을 활용하여 비선형 데이터의 분류에도 사용할 수 있어 다양한 방면에서 활용되는 알고리즘 이다. KNN은 분류 기준이 없는 데이터를 임의로 분류 기준을 선정하여 데이터를 분류하는 알고리즘이다. MLP는 입력층(Input layer), 은닉층(Hidden layer), 출력층(Output layer)으로 구성된다. 측정된 인공지능지체 이미지를 전처리 하여 특징을 구분한다. 이렇게 처리된 특징은 입력 레이어로 설정된다. 그리고 입력 레이어들은 여러 개의 은닉층들을 거쳐 비용, 가중치, 바이어스를 학습한 뒤, 출력 레이어로 도출된다. 최종 도출된 출력 레이어로 불량, 정상 제품을 분류하는 구조이다.

Fig 7에는 입력 데이터가 퍼셉트론 층을 통과하여 출력 데이터로 도출되는 MLP의 구조를 도시하였다.

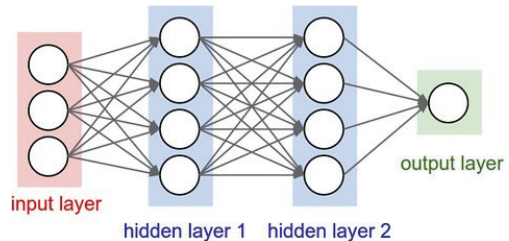


Fig. 7. Structure of MLP.

MLP는 데이터를 비선형적으로도 풀 수 있는 알고리즘이다. XOR 분류를 MLP를 이용한 분류의 대표적인 예로 들 수 있다[7]. 인공지능지체의 이미지로 외형과 내부 구조를 평가하여 불량, 정상을 분류할 때의 데이터 분포는 Fig 8과 같이 나타난다. Fig 8에는 선형으로 분류하지 못하는 데이터들을 MLP를 적용하여 XOR 게이트로 분류하는 과정을 도시하였다.

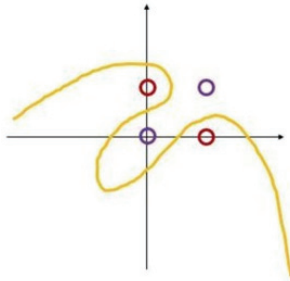


Fig. 8. XOR classification used MLP.

### 3.3 데이터 수집

FDM 방식 3D프린터로 지름 20 mm, 두께 2 mm의 원판 모양의 인공지지체를 제작한 뒤, 외형을 USB 카메라로 촬영하였다. 분류 모델을 학습하고 테스트 하기 위하여 352장의 사진을 이용하였다. 수집된 인공지지체의 제작 출력 환경을 추출하여 데이터화 한 뒤, 학습과 테스트에 사용하였다. 인공지지체의 출력 환경에는 노즐 온도, 베드 온도, 이동 속도, 채움 밀도, 토출량이라는 5개의 요인이 있다. 노즐 온도는 인공지지체를 제작할 때의 노즐 표면의 온도이고 베드 온도는 베드 표면의 온도이다. 이동 속도는 원료를 적층하기 위해 노즐이 이동하는 속도를 의미하며 채움 밀도는 적층하는 레이어(Layer)의 간격을 의미한다. 마지막으로 토출량은 기어가 원료를 밀어내는 압력을 의미한다. 이러한 여러 출력 환경에서 제작된 인공지지체를 촬영하여 외형 정보를 수집한다.

USB 카메라를 이용하여 여러 형태의 인공지지체를 촬영하였다. 촬영하여 수집된 이미지는 분류 모델 제작을 위한 학습, 테스트 데이터로 사용된다. Fig. 9에는 USB 카메라를 이용하여 촬영된 여러 종류의 인공지지체 외형을 도시하였다.

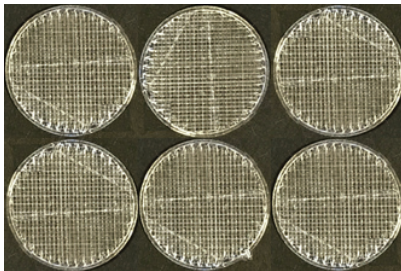


Fig. 9. Photographed scaffold using USB camera.

### 3.4 분류 모델 제작

Window 10 환경에서 Python으로 Jupyter notebook을 이용하여 인공지지체 불량 분류 모델들을 제작하였다.

인공지지체 불량 분류 모델들을 제작함에 있어 SVM, KNN, MLP 3가지의 분류 알고리즘을 이용하였다.

SVM은 선형, 비선형 방식이 있지만 본 연구에서는 선형 방식을 이용하였다. 그리고 SVM에서 C-value를 이용하여 분류 성능을 조절할 수 있다. 본 연구에서의 C-value는 1에서 10까지 설정할 수 있으며 숫자가 커질수록 분류 폭이 커지기 때문에 본 연구에서는 C-value를 중간 값인 5로 설정하였다.

KNN은 가장 간단한 기계 학습 모델 중 하나이다. 분류한 데이터들과 유사도를 측정하고 가장 비슷한 카테고리로 입력 데이터를 분류하는 방법을 의미한다. KNN은 주변 영역을 정하는 반지름의 크기에 따라 분류 성능이 결정된다. 본 연구에서는 특징 공간상에서의 반지름을 경험적으로 1로 설정하였다. MLP는 분류 문제를 해결할 수 있는 간단한 신경망으로 입력층, 은닉층, 출력층을 이용하여 데이터들의 특성을 찾아내어 가중치를 스스로 학습하는 알고리즘이다. 신경망을 구성하는 것이 다른 모델들에 비하여 간단하다. 따라서 학습에 소요되는 시간이 다른 모델에 비하여 짧은 것이 장점이다.

본 연구에서 MLP를 이용한 분류 모델은 1개의 입력층, 5개의 은닉층, 1개의 출력층으로 구성되어 있다. 사용된 활성화 함수는 분류 속도를 더욱 단축하기 위하여 ReLU 함수를 사용하였다.

## 4. 평가

### 4.1 분류 모델 성능 평가

인공지지체 불량 분류 모델 성능을 평가하기 위한 지표로 F1스코어를 사용하였다. F1스코어는 분류 모델의 정확도를 나타내는 대표적인 평가지표 중의 하나이다. 정밀도(Precision)와 재현율(Recall)의 조화 평균 값으로 계산되며 F1 스코어가 높을수록 분류 성능이 높다.

본 연구에서 사용된 데이터에 대한 인공지지체 불량 분류 모델의 일반화 성능 평가를 위하여 많이 사용되는 교차 검증 방법인 10겹 교차 검증(10 fold cross validation) 방법을 사용하였다. 10겹 교차 검증 방법은 전체 데이터를 무작위로 섞은 후 10개의 세트무작위로 나눈다. 나뉘어진 10 세트 중 9 세트는 학습 데이터로 사용하고, 1 세트는 테스트 데이터로 사용한다. 이러한 과정을 중복 없이 10회 진행한다. 각 회마다 테스트 결과를 F1 스코어로 출력한다. 출력된 결과를 분석하여 인공지지체 분류 모델의 성능을 비교한다. Table 1은 인공지지체 불량 분류 모델의 학습 성능을 평가하기 위하여 사용된 SVM, KNN, MLP 각 모델 별 F1 score를 나타내었다. Average는 평균 분류율을 나타내며 Total은 전체 분류율을 나타낸다.

### 4.2 평가 결과

기계 학습 기반의 분류 알고리즘에 대한 각각의 방법으로 인공지능지체 불량 분류 모델을 제작하여 학습 데이터와 테스트 데이터의 분류 성능을 평가하였다. 평균 분류율에서의 신뢰도가 높기 때문에 평가 결과는 Table 1에서 도출된 결과 중에서 평균 분류율에 해당하는 결과에 100을 곱한 후, 소수점 첫번째 자리에서 반올림을 하였다. Table 2는 학습 데이터와 테스트 데이터를 이용하여 SVM, KNN, MLP 각 분류 모델들의 성능을 나타내었다.

**Table 1.** F1 score of SVM, KNN, MLP models

Model \ Score type	Average	Total
Support vector machine	0.982	0.985
K-nearest neighbors	0.871	0.869
Multi-layer perceptron	0.951	0.963

**Table 2.** Precision of classification model performance

Model \ Data type	Training data accuracy	Test data Accuracy
Support vector machine	98%	97%
K-nearest neighbors	87%	85%
Multi-layer perceptron	95%	93%

SVM, KNN, MLP를 이용하여 인공지능지체 불량 분류 모델들을 제작하였고, 각 모델들의 성능을 테스트 데이터를 이용하여 분석한 결과 SVM을 이용한 분류 모델이 97%로 가장 높았으며, 그 다음으로 MLP를 이용한 분류 모델이 93%이었다. 마지막으로 KNN을 이용한 분류 모델이 85%로 가장 낮음을 확인하였다.

### 5. 결 론

본 연구에서 USB카메라를 이용하여 인공지능지체의 외형 데이터를 수집하고 특징을 추출하였다. 그리고 기계 학습의 분류 방안을 이용하여 인공지능지체 불량 분류 모델들을 제작함으로써 다음과 같은 결론을 얻을 수 있었다.

1. USB 카메라를 이용하여 인공지능지체의 외형을 촬영하여 인공지능지체의 외형 이미지를 수집할 수 있었다.

2. 수집된 이미지로 기계 학습 기반 SVM, KNN, MLP라는 3가지의 분류 알고리즘을 이용하여 인공지능지체 불량 분류 모델을 제작할 수 있었다.
3. SVM, KNN, MLP로 제작된 3가지 모델의 성능을 수치화한 결과 SVM을 이용한 분류 모델이 97%로 가장 높음을 확인하였다. 따라서 SVM, KNN, MLP 모델 중, 성능이 가장 높은 모델은 SVM을 이용한 분류 모델이라는 것을 확인할 수 있었다.

### 참고문헌

1. Seung-Hyeok Choi, Min-Woo Sa and Jong-Young Kim, "New Fabricatio Method of Bio-Ceramic Scaffolds Based on Mold using a FDM 3D Printer", J. of The Korean Society of Precision Engineering, Vol.18, pp. 957-963, 2018.
2. Song-Yeon Lee and Yong-Jeong Huh, "A Study on Prediction Model Performance of Scaffold Pore Size Using Machine Learning Regression Method", J. of The Korean Society of Semiconductor & Display Technology, Vol.19, pp. 36-41, 2020.
3. Song-Yeon Lee and Yong-Jeong Huh, "A Study on Prediction Model of Scaffold Appearance Defect Using Machine Learning", J. of The Korean Society of Semiconductor & Display Technology, Vol.18, pp. 46-50, 2020.
4. Young-Ho Lee and Seong-Yun Hong, "A Machine Learning Approach to the Prediction of Individual Travel Mode Choices", J of the Koreaa Data and Information Science Society, Vol. 30, pp. 1011-1024, 2019.
5. Song-Yeon Lee and Yong-Jeong Huh, "A Study on Prediction Model of Scaffold Pore Size Using Machine Learning", J. of The Korean Society of Semiconductor & Display Technology, Vol.18, pp. 46-50, 2019.
6. Yong-Beom Park, Dong-Bin Choi and In-Soo Cho, "Taxation Analysis Using Machine Learning", J. of The Korean Society of Semiconductor & Display Technology, Vol.18, pp. 73-77, 2019.
7. Yu-Sun Ahn, Hue-Jin Kim, Sang-Kyu Lee and Byung sean Kim, "Prediction of Heating Energy Consumption Using Machine Learning and Parameters in Combined Heat and Power Generation", J of Air-Conditioning and Refrigeration Engineering, Vol. 31, pp. 352-360, 2019.

접수일: 2020년 9월 1일, 심사일: 2020년 9월 8일,  
 게재확정일: 2020년 9월 22일