

## 무인점포 이상행동 인식을 위한 유전 알고리즘 기반 자세 추정 모델 최적화

### Optimization of Pose Estimation Model based on Genetic Algorithms for Anomaly Detection in Unmanned Stores

이상협<sup>1</sup>, 박장식<sup>2\*</sup>

Sang-Hyeop Lee<sup>1</sup>, Jang-Sik Park<sup>2\*</sup>

#### 〈Abstract〉

In this paper, we propose an optimization of a pose estimation deep learning model for recognition of abnormal behavior in unmanned stores using radio frequencies. The radio frequency use millimeter wave in the 30 GHz to 300 GHz band. Due to the short wavelength and strong straightness, it is a frequency with less grayness and less interference due to radio absorption on the object. A millimeter wave radar is used to solve the problem of personal information infringement that may occur in conventional CCTV image-based pose estimation. Deep learning-based pose estimation models generally use convolution neural networks. The convolution neural network is a combination of convolution layers and pooling layers of different types, and there are many cases of convolution filter size, number, and convolution operations, and more cases of combining components. Therefore, it is difficult to find the structure and components of the optimal posture estimation model for input data. Compared with conventional millimeter wave-based posture estimation studies, it is possible to explore the structure and components of the optimal posture estimation model for input data using genetic algorithms, and the performance of optimizing the proposed posture estimation model is excellent. Data are collected for actual unmanned stores, and point cloud data and three-dimensional keypoint information of Kinect Azure are collected using millimeter wave radar for collapse and property damage occurring in unmanned stores. As a result of the experiment, it was confirmed that the error was moored compared to the conventional posture estimation model.

*Keywords : Deep Learning, Pose Estimation, Human Activity Recognition, Genetic Algorithm, Neural Architecture Search*

1 경성대학교 전자공학과

2\* 교신저자, 경성대학교 전자공학과, 교수

E-mail: jsipark@ks.ac.kr

1 Dept. of Electronic Eng., Kyung Sung University, Student

2\* Corresponding Author, Dept. of Electronic Eng., Kyung Sung

University, Professor

## 1. 서론

최근 코로나 19의 영향과 사회적인 독립 가구가 늘어남에 따라 다양한 형태의 무인점포(Unmanned Stores)가 늘어나고 있다[1]. 최근 들어 무인점포 내 보안사고인 절도는 2019년 203건에서 2020년 367건으로 2021년 1~5월에만 686건으로 발생 빈도가 증가하였다. 현재 무인점포 내 이상행동 인지는 사건 발생 후 녹화된 단순 CCTV 영상을 사람이 육안으로 사건을 확인하고 검색하면서 정의를 내리는 경우가 많다[2].

무인점포 수의 확대에 따라 무인점포 내 이상 행동에 대한 AI 기반의 신뢰성이 보장되는 상황 인지 기술이 필요하다. 이러한 무인점포 내 이상 행동 인식을 위하여 영상 기반의 무인점포 내 이상 행동 인식에 관련된 연구들이 진행되었다[3-4].

CCTV를 기반으로 하는 무인점포 내 이상행동 인식에 대한 연구는 많이 진행이 되었다. 하지만 CCTV 영상은 개인정보 침해요소가 포함되어 있다[5]. 무인점포 내 발생 할 수 있는 개인정보 침해요소는 다음과 같다. IP 기반의 영상촬영 장치(CCTV 등)기기에 대한 해킹 등의 악의적인 접근이 가능하며, 최근 얼굴인식 기술을 활용한 출입 통제시스템과 고객관리 시스템이 증가하고 있으나 얼굴 인식 기술은 개인정보 침해의 우려가 있다[6]. 본인이 인지하지 못하는 사이에 개인의 얼굴 정보가 수집·저장 될 수 있으며, 개인 스스로가 서비스를 제공받기 위해 암묵적으로 동의하면 개인정보를 보호할 수 없다.

개인 정보 침해 문제를 해결하기 위해 영상 기반의 이상 행동 인식에서 자세 추정 기반의 이상 행동 인식 연구가 진행되었다[7]. 그 중 RF(Radio Frequency)를 사용하여 사람의 관절점(Keypoint)를 추정하고 추정된 관절점을 사용하여 이상 행동 인식을 수행하였다[8-10].

자세 추정 기반 이상 행동 인식의 성능은 자세 추정 모델의 성능에 영향을 받는다. 딥러닝(Deep Learning) 기반의 자세 추정 모델의 최적의 구조 및 성능은 입력 데이터에 따라 다를 수 있다.

본 논문에서는 유전 알고리즘을 사용하여 무인점포 내 이상 행동 인식을 위한 RF 기반 자세 추정 모델의 최적화를 제안한다. 사용하는 RF는 mmWave(Millimeter Wave)이며, 유전 알고리즘을 통해 최적의 딥러닝 모델의 구조와 하이퍼 파라미터(Hyperparameter)를 탐색한다. 최적의 딥러닝 모델을 사용하여 관절점을 추정하고 이를 LSTM(Long Short-Term Memory)와 GRU(Gated Recurrent Unit)을 사용하여 이상 행동인식을 한다.

## 2. 딥러닝 모델 최적화 및 자세 추정

### 2.1 유전 알고리즘

유전 알고리즘은 유전과 진화를 공학적으로 모형화한 확률 탐색 방법으로 근접 최적의 해를 찾기 쉽고, 해를 병렬적으로 탐색하여 지역 최적해 문제에 빠지지 않고, 전역 최적의 해를 찾아갈 수 있다. 하지만, 항상 전역 최적의 해를 찾을 수 있는 것은 아니며, 최적의 해를 찾더라도 그 과정에 대해서는 알 수 없다[11].

유전 알고리즘의 염색체(Chromosome)는 유전 정보를 담고 있는 집합이다. 유전자(Gene)는 구성하는 요소로 유전 정보를 나타내며, 자손(Offspring)은 이전 세대의 염색체로부터 생성된 새로운 염색체다. 적응도(Fitness)는 염색체가 가지고 있는 고유의 값이며, 엘리트주의(Elitism)는 이전 세대의 염색체 중 적응도가 좋은 염색체를 다음 세대에 그대로 보존하는 것이다.

유전 알고리즘은 Fig. 1과 같이 초기 세대에 대

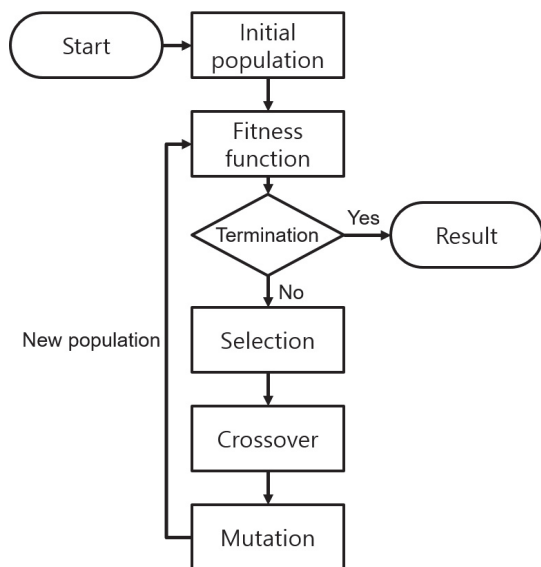


Fig. 1 Flowchart of genetic algorithm

한 염색체의 유전자 초기화(Initialize), 각 염색체의 고유의 적응도를 계산하는 평가 함수(Fitness Function), 자손 염색체 생성을 위한 부모 염색체를 선택하는 선택 연산(Selection), 선택 연산으로 선택된 부모 염색체로부터 새로운 자손 염색체를 생성하는 교차 연산(Crossover) 그리고 지역 최적화 문제를 해결하기 위해 부모의 형질과 다른 유전자를 가지게 하는 돌연변이(Mutation) 연산과 같은 단계를 통해 다음 세대를 생성하고 문제의 해를 도출할 때까지 반복한다.

유전 알고리즘의 선택연산을 통해 부모 세대의 염색체를 선택한다. 선택된 염색체들 간의 연산을 통해 자손을 생성한다. 선택연산에는 다양한 방법이 있지만, 기본적인 원칙은 적응도가 좋은 염색체의 선택확률을 높이고, 적응도가 낮은 염색체의 선택확률을 낮추는 것이다. 하지만, 적응도가 좋은 염색체만 선택한다면 실익은 수렴 문제(Premature Convergence)에 빠질 수 있기 때문에 주의가 필요하다.

교차연산은 선택연산으로 선택된 부모 세대의

염색체 간의 교배를 통해 자손 염색체를 생성하는 연산이다. 즉, 부모세대의 염색체의 유전자를 서류 교환하여 새로운 염색체를 생성한다.

변이연산은 부모 염색체에 없는 유전자를 일정한 확률에 따라 자손의 유전자에 적용시킨다. 변이연산은 부모 염색체에 없는 유전자 형질을 적용시킨다는 점에서 교차연산으로 찾을 수 없는 해를 찾을 수 있다. 이처럼 변이연산은 해 집단을 다양화시켜 지역 최적해 문제에 빠지는 것을 방지하고, 전역 최적 해를 찾아낼 수 있다.

## 2.2 유전 알고리즘 기반 모델 최적화

딥러닝 모델은 생물의 신경세포에서 영감을 얻어 자유로운 연결 구조의 조합으로 다양한 구조의 모델을 구현 할 수 있다. 다양한 구조의 모델을 탐색하기 위해 NAS(Neural Architecture Search)이 제안되었다. 일반적으로 NAS에 시간 복잡도는  $O(nt)$ 로  $n$ 은 딥러닝 모델의 개수,  $t$ 는 신경망 모델의 평가 시간이다. NAS의 속도는 모델의 크기와 평가 시간에 비례하며, 주로 기울기를 사용하여 학습하는 딥러닝 모델은 1회 학습에 많은 시간이 소용된다. NAS를 위해 강화 학습, 경사 기반 접근법 그리고 확률 기반의 접근법 등과 같은 다양한 탐색 알고리즘을 사용한 연구가 진행되었다[12]. 신경망 구조 탐색은 신경망의 모든 구조에 대해 탐색하는 것이 아닌 특정 탐색 공간을 정의하고 이 공간에 대해 탐색을 수행한다. 탐색 공간을 정확하게 설정하여 정의 하는 것이 탐색 알고리즘의 성능에 영향을 끼친다.

## 2.3 자세 추정

자세 추정은 사람의 관절점의 위치를 추정하여 사람의 동작 자세를 인하는 기술이다. 주로 추정

하는 관절점은 머리, 목, 어깨, 팔꿈치, 손목, 엉덩이, 무릎, 발목이다. 추정하는 좌표계는 3차원 자세 추정(3D Pose Estimation)과 2차원 자세 추정(2D Pose Estimation)으로 나뉘지며 인원 추정 방식은 추정하는 사람의 수에 따라 단일 사람 자세 추정(Single-Person Pose Estimation)과 다중 사람 자세 추정(Multi-Human Pose Estimation)으로 구분한다.

자세 추정 방식에는 객체를 먼저 검출하고 자세를 추정하는 하향(Top-Down) 방식과 객체를 지역화 하여 관절을 추정하고 각 관절의 상관관계를 계산하여 연결하는 상향(Bottom-Up) 방식으로 나뉜다. 하향 방식은 상향 방식 비해 정확도가 높으나, 객체를 검출하는 과정이 요구되고 추정되는 다중 사람 자세 추정의 경우 각 객체 별 자세를 추정하기 때문에 추정 속도가 느리다. 반면 상향 방식은 하향식 방식에 비해 정확도가 낮으나 객체 검출 과정이 없어 추정 속도가 높아 실시간 처리에 적합하다.

### 3. 제안하는 유전 알고리즘 기반의 자세 추정 모델 최적화

본 논문에서는 무인점포 내 이상 행동 인식을 위해 유전 알고리즘 기반의 자세 추정 모델 최적화를 제안한다. 제안하는 이상 행동 인식을 위한 자세 추정 최적화 모델은 다음 Fig. 2와 같다. mmWave에서 x, y, z, v, I의 5차원 Point Cloud를 수집하여, x축 기준으로 정렬한다. 이후 유전 알고리즘을 사용하여 최적의 관저점 추정 모델이 구조와 하이퍼파라미터를 탐색한다.

종래의 RF 기반 자세 추정은 CNN 또는 RNN을 사용하여 자세를 추정했다. 하지만 정형화 된 모델의 구조와 하이퍼파라미터를 사용하였기 때문

에 입력 데이터에 대한 최적의 딥러닝 모델은 아니다. 최적의 딥러닝 모델을 선택하기 위해 다음과 같은 방법을 제안한다.

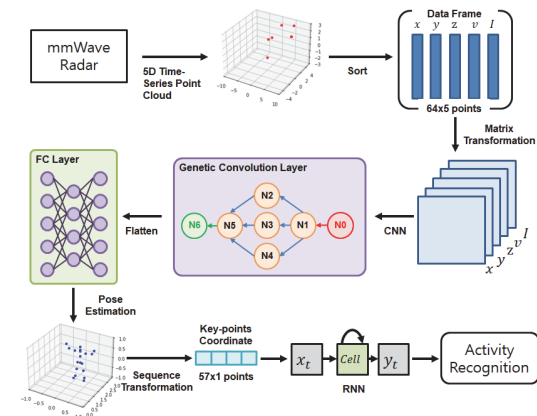


Fig. 2 Pose estimation for anomaly detection in the proposed unmanned store

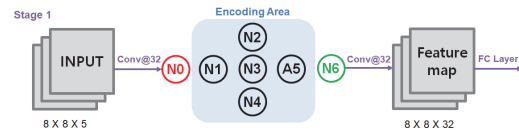


Fig. 3 A search space for deep learning model structure optimization

Table 1. Encoding rules for hyperparameters

Activation Function	
Hyperparameter	Bit Representation
ReLU	00
Leaky ReLU	01
4-Clipped ReLU	10
8-Clipped ReLU	11
Optimization Algorithm	
Hyperparameter	Bit Representation
SGD	00
SGD-M	01
RMSProp	10
Adam	11

딥러닝 모델 최적화를 위해 탐색 공간을 길이가 15인 유전자로 표현한다. 딥러닝 모델의 구조는 Fig. 3과 같이 1-Stage 5-Node로 구성되어 있으며, 하이퍼파라미터는 활성 함수와 최적화 알고리즘은 Table 1과 같이 표현하고 컨볼루션 필터의 종류는 Fig. 4와 같이 구성된다.

유전 알고리즘의 평가 함수는 탐색체의 유전자 정보를 바탕으로 딥러닝 모델을 생성한다. 생성된 딥러닝 모델의 지역화 오차를 계산하여, 낮은 오차를 가지는 탐색체에 대해 높은 적응도를 부여하였다.

초기화는 SRR(Simple Random Reverse) 방식을 사용한다. SRR은 탐색체의 길이의 절반을 계산하여, 계산된 값보다 작은 순번(Index)를 가지는 유전자에 임의의 값을 부여한다. 이후 대칭되는 유전자 위치에 보수를 취한다.

선택 연산은 룰렛 휠 선택(Roulette Wheel Selection) 방식을 사용하며, 교차 연산 균등 교차(Uniform Crossover) 방식을 사용한다.

변이 연산은 고정된 5세대마다 교차 연산을 통해 새로 생성된 탐색체에 대해 변이 연산을 사용한다. 변이 연산은 균등 교차와 비슷한 방식으로 다음과 같이 사용된다. 0에서 1 사이의 난수를 가지는 탐색체 길이의 배열을 생성한다. 임계 확률  $P$ 를 임의의 값으로 생성하여, 각 유전자 위치의

난수가  $P$ 이상인 경우 역수를 취하여 변이된 탐색체를 생성한다.

## 4. 실험 및 검토

### 4.1 데이터셋

mmWave 기반의 자세 추정에 사용되는 데이터셋은 Fig. 5와 같이 mmWave Radar에서 검출된 포인트 클라우드(Point Cloud)의 좌표 정보 및 수신 세기와 GT를 위한 관절점의 좌표 정보로 구성된다. 먼저 mmWave Radar에서 전파를 방출한다. 방출된 전파는 객체에 반사되어 수신기를 통해 수신되게 된다. 이후 Radar Signal Processing Chain를 통해 Point Cloud의 3차원 좌표를 계산한다. 동일한 프레임에 촬영된 영상 기반의 자세 추정 결과를 정답으로 하여 학습용 데이터셋을 구성한다.

무인점포 내 이상 행동 인식을 위해 쓰러짐과 기물파손에 대해 mmWave Radar와 Kinect Azure를 사용하여 포인트 클라우드 정보와 관절점 좌표 정보를 수집하였다. 센서를 지면과 높이 2m, 각도 75°로 설치하였다. 데이터 수집에는 5명의 실험자가 2가지 동작에 대해 각 동작 당 약 30초간 촬영을 하였다. 키포인트 정보는 Kinect에서 수집된 32개와 BlazePose에서 수집된 33개 그리고 OpenPose에서 수집된 25개로 구성되어 있

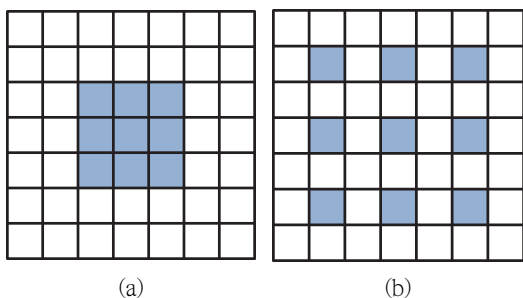


Fig. 4 Types of convolution filters  
(a) Normal convolution (b) Dilated convolution

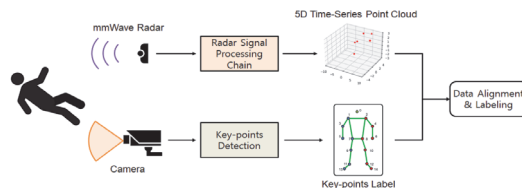


Fig. 5 mmWave based-on pose estimation dataset

으며, 데이터셋 수는 Kinect는 176,142프레임, BlazePose는 189,686 프레임, OpenPose는 96,293 프레임으로 구성된다.

## 4.2 실험 결과 및 고찰

실험 환경은 OS Ubuntu 18.04에서 Intel Xeon Silver 4208 CPU와 NVIDIA Quadro 5000P GPU 2대, RAM 32GB를 사용하였으며, 사용 언어는 Python 3.8과 Keras를 사용하여 자세 추정 모델의 최적화 실험 환경을 구축하였다. 딥러닝 모델 학습에는 학습률 0.001, 배치 크기 24, 그리고 에포크(epoch) 200 으로 사용하였다.

Table 2는 무인점포 데이터셋에 대한 자세 추정 실험의 결과이다. 종래의 mmWave 기반의 자세 추정 모델은 mmPose-NLP의 지역화 오차가 21.25로 가장 성능이 우수하였지만, 제안하는 최적 모델이 지역화 오차가 17.27로 가장 성능이 우수하였다. 종래의 mmWave 기반의 자세 추정 모델을 사용하는 것 보다 제안하는 유전 알고리즘 기반으로 최적의 모델을 탐색하는 것이 타당함을 보였다. 또한 영상 기반의 자세 추정과 성능을 비교하였을 때, 지역화 오차가 6.31로 가장 성능이 우수한 BlazePose 와는 성능 차이가 다소 나지만,

Table 2. Pose estimation experiment results using the unmanned store dataset

Pose Estimation Model	Localization Error(Pixel)			
	X	Y	Z	Avg.
Proposed Model	<b>16.92</b>	<b>17.81</b>	<b>17.09</b>	<b>17.27</b>
MARS	20.67	25.31	24.78	23.59
mmPose	20.63	23.46	23.62	22.57
mmPose-NLP	<b>20.78</b>	<b>21.33</b>	<b>21.63</b>	<b>21.25</b>
Kinect Azure	12.29	16.04	5.28	11.20
BlazePose	<b>6.42</b>	<b>8.92</b>	<b>3.58</b>	<b>6.31</b>
OpenPose	7.23	9.62	4.81	7.22

지역화 오차 11.20의 Kinect Azure와는 근소한 차이를 나타냈다. 종래의 영상 기반의 자세 추정에 비해서도 성능이 유사함을 확인했다.

시뮬레이션 결과 종래의 mmWave Radar 기반의 자세 추정은 z축이 다른 축에 비해 오차가 크게 발생하는 것을 확인하였다. 반면 제안하는 자세 추정 모델 최적화는 모든 축들이 비슷한 오차를 가지며, 종래의 mmWave Radar 기반의 자세 추정 보다 성능이 개선된 것을 확인 할 수 있다.

## 5. 결론

제안하는 유전 알고리즘 기반의 자세 추정 모델 최적화는 입력 데이터에 대하여 1-stage 5-Node를 가지는 딥러닝 모델을 생성하여 성능을 비교한 결과 축 기준 정렬의 지역화 오차는 20.61이고 수신 세기 기준 정렬의 지역화 오차가 15.77로 오차가 4.84 화소만큼 개선되었다. 종래의 mmWave 기반의 자세 추정 모델과 비교하면, mmPose-NLP의 지역화 오차가 21.25로 종래의 연구에서 가장 성능이 우수하였지만, 제안된 최적화 모델과의 지역화 오차 차이가 5.48 화소만큼 난다. 또한 영상 기반의 자세 추정과 성능을 비교하였을 때, 지역화 오차가 6.31로 가장 성능이 우수한 BlazePose 와는 성능 차이가 다소 나지만, 지역화 오차 11.20의 Kinect Azure와는 근소한 차이를 나타냈다. 실험 결과 종래의 영상 기반의 자세 추정에 비해서도 성능이 근소한 차이를 나타내는 것을 확인했다.

## 감사의 글

본 논문은 부산광역시 및 (재)부산인재평생교육진흥원의 BB21플러스 사업 지원과 2022년도 인



공지능 학습용 데이터 구축 지원사업의 2-97. 실내(편의점, 매장) 사람행동 영상 데이터 과제의 연구 결과임.

## 참고문헌

- [1] 이진천, “무인점포 시대의 도래,” 설비저널, vol. 47, no. 12, pp. 144-145, 2018.
- [2] 김진희, 장철영, “빅데이터를 이용한 무인점포 범죄 연구.” Crisisonomy, vol. 18, no. 9, pp. 95-110, 2022.
- [3] 이지애, 문성철, “스마트 관제를 위한 딥러닝 기반 이상행동 기술 동향 분석.” 한국방송미디어공학회 학술대회, vol. 29, no. 8, pp. 14-16, 2019.
- [4] S.-J. Horng and P.-S. Huang, “Building Unmanned Store Identification Systems Using YOLOv4 And Siamese Network.” Applied Sciences, vol. 12, no. 8, pp. 3826-3843, 2022.
- [5] D. Greene, “Facial recognition technology ‘inherently biased’ says Google cloud boss.” Google, 2018.
- [6] 개인정보침해 조사팀, “얼굴인식 기술의 최근 개인정보보호 논란 관련 동향 분석.” KISA 동향보고서, pp. 1-10, 2019.
- [7] B. Raj N., A. Subramanian, K. Ravichandran, and N. Venkateswaran, “Exploring Techniques To Improve Activity Recognition Using Human Pose Skeletons.” 2020 IEEE Winter Applications of Computer Vision Workshops (WACVW), pp. 165-172, 2020.
- [8] S. An and U. Y. Ogras, “MARS: mmWave-Based Assistive Rehabilitation System For Smart Healthcare.” ACM Transactions on Embedded Computing Systems, vol. 20, no. 5, pp. 1-22, 2021.
- [9] A. Sengupta, F. Jin, R. Zhang, and S. Cao, “mm-Pose: Real-Time Human Skeletal Posture Estimation Using mmWave Radars And CNNs.” IEEE Sensors Journal, vol. 20, no. 17, pp. 10032-10044, 2020.
- [10] A. Sengupta and S. Cao, “mmPose-NLP: A Natural Language Processing Approach To Precise Skeletal Pose Estimation Using mmWave Radars.” IEEE Transactions on Neural Networks and Learning Systems, pp. 1-12, 2022.
- [11] S. Katoch, S. S. Chauhan, and V. Kumar, “A Review On Genetic Algorithm: Past, Present, And Future.” Multimedia Tools and Applications, vol. 80, no. 5, pp. 8091-8126, 2020.
- [12] T. Elsken, J. H. Metzen, and F. Hutter, “Neural Architecture Search.” Automated Machine Learning, vol. 20, no. 1, pp. 63-77, 2019.

(접수: 2022.11.10. 수정: 2022.12.27. 게재확장: 2022.12.30.)