Pattern Recognition of Human Arm Movement Using Deep Reinforcement Learning

W. Seok[1], Y. Kim[2], C. Park[3]

[1]Intelligent information system and embedded software engineering, Kwangwoon University, Seoul, Korea

[2, 3]Department of Computer Engineering, Kwangwoon University, Seoul, Korea

[1]wjseok@kw.ac.kr,[2]lyoungjookiml@gmail.com,[3]parkcheolsoo@kw.ac.kr

Abstract—Hand gesture recognition is one of the major research areas in the field of Human computer interaction (HCI). This paper proposes a deep reinforcement learning algorithm to recognize the human arm movement patterns using an IoT sensor device. Recent studies have explored supervised learning based methods, such as CNN and RNN to implement the HCI device. On the other hand, the deep reinforcement learning approach has also been investigated. Algorithms using this approach, learn the patterns from sensors using only the reward feedback with no class labels. This allows users to control the IoT device and produce the desired arm movement patterns without creating any labels. In this paper, the performance of convolutional neural network (CNN) with the DQN model was compared with that of long short-term memory (LSTM) models with DQN. Results show that the CNN based DQN model was more stable compared to the LSTM based model, and yielded a high classification accuracy of 98.33% to predict the arm movement patterns.

Keywords—Human-Computer Interaction (HCI), patterm recognition, deep reinforcemnet learning, deep Q-Netwrok, Myo armband

I. INTRODUCTION

The importance of Human-Computer Interaction (HCI) is increasing due to the advent of the 4th industrial revolution and improvement of the IoT technology. Especially, voice and image recognition studies are actively conducted, compared to the field of pattern recognition. In this study, we focus on the recognition of human arm gesture patterns using Electromyography (EMG). Thalmic Lab's Myo arm band, which embed 9-axis IMU sensors and eight surface electromyography sensors for EMG recording, are used to measure human arm movements. There are recent studies using the Myo arm band to classify hand gestures with EMG signals [1-2]. It is easy to classify the hand gestures when the human arm is still. However, in order to recognize the hand movement patterns, the IMU sensor data is necessary. The Myo arm band gives us 3-axis accelerometer data, 3-axis gyroscope data, and Quaternion data, and we used the Accelerometer and Quaternion data as an input of the DQN model. The following are the main advantages of the proposed pattern recognition method using deep Q-Network:

- End-to-end, with raw 9-axis IMU sensor data

- Learns from rewards not from labels

II. METHODS

A. Myo arm band data

It is possible to extract raw accelerometer data (Quaternion, roll, pitch and yaw) from the Myo Arm band.

B. Experiment for getting human arm movement pattern

We collected the human arm movement patterns using the Myo arm band. The experiment consists of three gestures, where subjects are instructed to draw a circle, rectangle and triangle in the air. Each of these three classes has 30 training data and 20 test data.
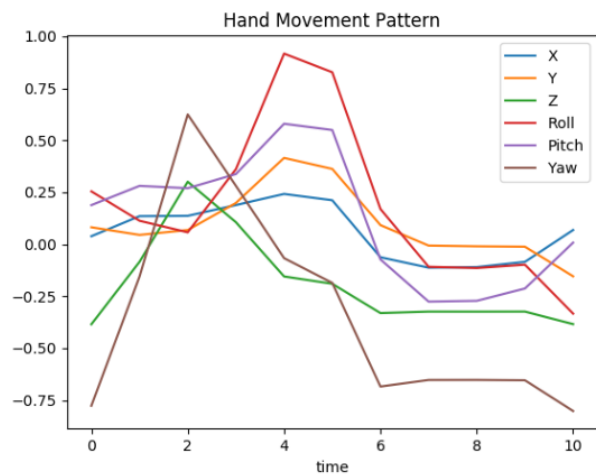

Fig. 1. Hand movement Pattern for drawing Triangle pattern

III. APPLICATION TO DEEP REINFORCEMENT LEARNING

Reinforcement learning has become very popular since the advent of AlphaGo, which is an artificial intelligence Go program developed by Google Deep Mind.

A. Reinforcement learning

Reinforcement learning differs from supervised learning and unsupervised learning. The answer is not given, but it is not just learning about the given data. Reinforcement learning is taught through rewards. In supervised learning, the model learns by calculating errors through direct answers, but in reinforcement learning, the agent learns from compensation that is the result of agent's own action. For more detailed concepts and formulas of reinforcement learning, refer to Sutton and Barto's book, "Reinforcement Learning: An Introduction" [3].

B. Markov Decision Process (MDP)

To solve sequential behavior decision problems, the problem must be mathematically defined as follows:

---

[3]Corresponding author

- State: A state is a set of states observable by an agent. In this paper we set the one trial arm movement pattern as a single state. Three states are defined: drawing triangles, drawing squares, and drawing circles.

- Action: An action is a set of behaviors that an agent can have within a state. In this paper, an action predicts how one has acted within a state.

- Reward: A reward is the only information that an agent can learn and this information given to an agent by the environment. If the agent matches the intended pattern of the person, it will be compensated.

C. Q- learning

Q-learning was first introduced by Watkins *et al* [4]. To update the Q function, it must have a sample (s, a, r, s') The following is an updated formula of the Q function through Q-learning.

$$Q(S,A) \leftarrow Q(S,A) + \alpha \left(R + \gamma \max_{a'} Q(S',a') - Q(S,A)\right)^2 \tag{1}$$

D. Deep Q-Network

MNIH, Volodymyr, *et al* trained the agent to play Atari games better than the human player [6]. Equation 2 is a loss function for deep Q-Network.

$$\text{MSE} = \left(R_{t+1} + \gamma \max_{a'} Q(S_{t+1}, a', \Theta^-) - Q(S_t, A_t, \Theta)\right)^2 \tag{2}$$

We assumed the left term is the correct answer and the right term is the agent's prediction. For a stable learning, we fixed the θ^- until the end of the episode and updated the main network for each time step. Fig. 2 is the LSTM based DQN agent.
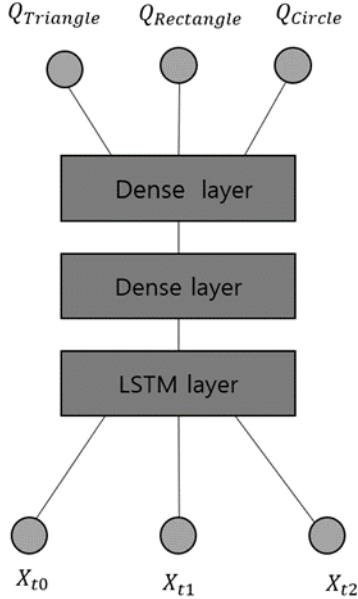

Fig. 2. LSTM based DQN agent structure

The LSTM based DQN agent structure was first introduced by Adam Cobb *et al* [5]. LSTM is one kind of recurrent neural network (RNN) and it solves the RNN's vanishing gradient problem by including a gating mechanism within each layer. The state, which is a single arm movement pattern, will take into LSTM layer. The last layer has no activation function, but instead has a linear function, since the Q value is not a probability but a constant value. Then we used the CNN based DQN agent.
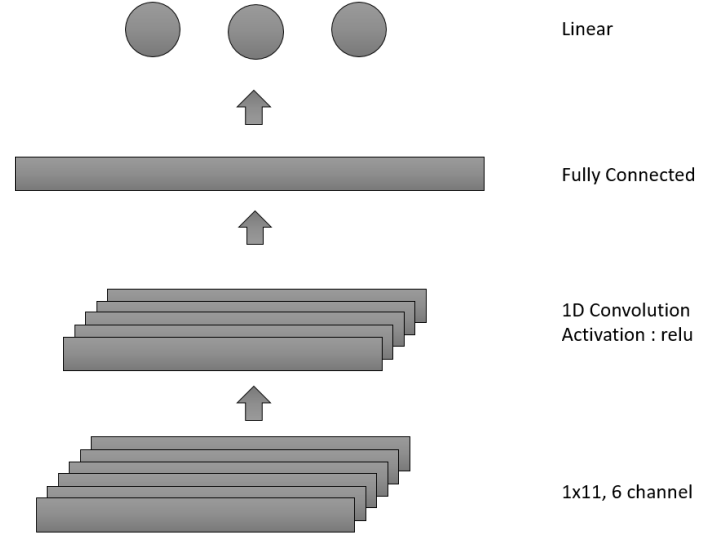

Fig. 3. 1-D CNN based DQN agent structure

MNIH, Volodymyr, *et al* introduced the experience replay [7]. Experience replay is a way to break the correlation between each sample. We used experience replay with a size of 2000.
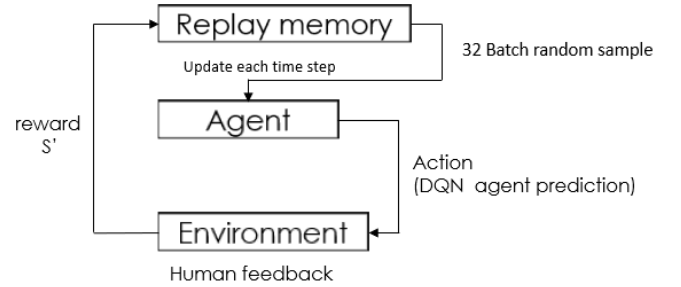

Fig. 4. Agent's learning process

Fig. 4 shows the agent's learning process. The agent predicts the human arm moving patterns and the result of the prediction is an action. Then the environment (human) gives the feedback (reward) back to the agent. If the prediction is correct, the human would give the + reward and if the prediction was incorrect, the environment gives a penalty. The episode is made up of 10 lives and if the agent predicts an incorrect action, then the life are reduced.

IV. RESULTS

We compared CNN and LSTM models to construct the optimal DQN agent, which resulted in the CNN model performing better than the LSTM model. Fig. 5 shows that the

LSTM based DQN agent is unstable, since the scores fluctuates under 300. Fig. 6 shows that the 1-D CNN based DQN agent receives scores higher than 400. However, the LSTM DQN agent gains scores which range from 100 to 300. After learning with deep reinforcement learning, whenever an agent received a high score, we used weights in order to predict a new human arm movement pattern. Using 90 training data, where the human draws a circle, triangle and rectangle, the agent predicted the human arm movement patterns. As a result, it showed a 98.33% classification accuracy.
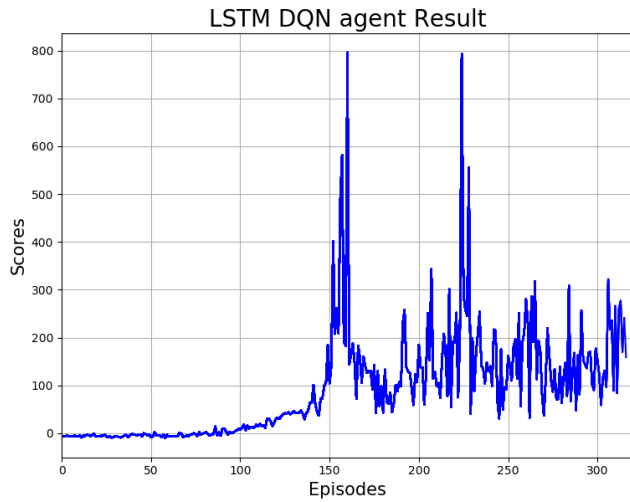

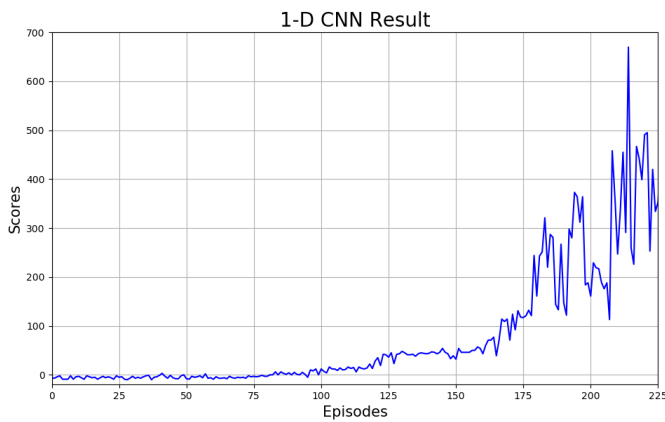Fig. 5.   Results using the LSTM based DQN agent

Fig. 6. Results using the CNN based DQN agent

V.   CONCLUSION

The result shows that deep Q-Network can successfully learn the human arm movement pattern. Additionally, the 1-D CNN structure yielded a higher performance compared to the LSTM structure. For future work, the deep reinforcement's unstable leaning can be investigated.

Authors and Affiliations

REFERENCE

[1] Ali BOYALI, "Spectral Collaborative Representation based Classication for Hand Gestures recognition on Electromyography Signals", 2015 IEEE 4th Global Conference on Consumer Electronics (GCCE), Volume 24, Pages 11-18, February 2016

[2] Mahfud Ridho Pambudi, Riyanto Sigit and Tri Harsono, "The Bionic Hand Movement Using Myo Sensor and Neural Networks", Knowledge Creation and Intelligent Computing (KCIC),2016

[3] Richard S Sutton and Andrew G Barto. "Reinforcement learning: An introduction", Vol. 1. No. 1. Cambridge: MIT press, 1998.

[4] WATKINS, Christopher JCH; DAYAN, Peter. "Q-learning". Machine learning, 8.3-4: 279-292 1992

[5] Adam Cobb and Andrew Markham, "Adaptive sampling of lion accelerometer data," Oxford, 2016

[6] MNIH, Volodymyr, et al. "Playing atari with deep reinforcement learning", arXiv preprint arXiv:1312.5602, 2013.

[7] MNIH, Volodymyr, et al. "Human-level control through deep reinforcement learning. Nature", 518.7540: 529-53, 2015

[8]