A Deep Reinforcement Learning Blind AI in DareFightingICE

Thai Van Nguyen, Xincheng Dai, Ibrahim Khan
Graduate School of Information Science and Engineering
Ritsumeikan University
Kusatsu, Shiga, Japan
{gr0557fv, gr0502pv, gr0556vx}@ed.ritsumei.ac.jp

Ruck Thawonmas
College of Information Science and Engineering
Ritsumeikan University
Kusatsu, Shiga, Japan
ruck@is.ritsumei.ac.jp

Hai V. Pham
School of Information and Communication Technology
Hanoi University of Science And Technology
Hanoi, Vietnam
haipv@soict.hust.edu.vn

Abstract—This paper presents a deep reinforcement learning agent (AI) that uses sound as the input on the DareFightingICE platform at the DareFightingICE Competition in IEEE CoG 2022. In this work, an AI that only uses sound as the input is called blind AI. While state-of-the-art AIs rely mostly on visual or structured observations provided by their environments, learning to play games from only sound is still new and thus challenging. We propose different approaches to process audio data and use the Proximal Policy Optimization algorithm for our blind AI. We also propose to use our blind AI in evaluation of sound designs submitted to the competition and define two metrics for this task. The experimental results show the effectiveness of not only our blind AI but also the proposed two metrics.

Index Terms—Sound, Blind AI, Deep Reinforcement Learning, Proximal Policy Optimization, Fighting Game, FightingICE, DareFightingICE

I. INTRODUCTION

Sounds in video games have been an important factor for a long time [1]–[3]. There are different types of sound effects in a game, ranging from UI sounds, ambient sounds, background music, and so on. These different types of sound effects improve the immersion of the human players. Human players use this sound input from the game for various tasks, such as, finding the location of an item or an enemy, and recognizing objects by their specific sounds. Sounds in a video game can help human players in different ways, but the question that we address in this paper is: Can AIs learn from sound in a video game?

Recent research has shown that an AI can use sound as an input to detect the location and direction of an object [4]. An AI that uses sound as an input together with other information is shown to perform better than those who do not [5]. However, research in AIs that use sound as the input is still in its premature stage, and previous research either focused on learning to play simplified games from only audio cues or used visual and other inputs like game data along with audio data in their AI. Our research introduces an AI that only uses sound as the input to play a fighting game called "DareFightingICE".

DareFightingICE [6] is an enhanced version of FightingICE [7], a fighting game, which has been used as the platform for the Fighting Game AI Competition series since 2013. DareFightingICE is an official competition in the 2022 IEEE Conference on Games. DareFightingICE has a "Sound-Only" option in which AIs only receive audio data as the input.

The contributions of this work are as follows: first, the creation of the first fighting game AI, an official sample AI in the competition, that uses only sound as the input (blind AI); second, our usage of the AI in evaluating the effectiveness of a given sound design regarding its in-game-event representation ability; and third, our work opens a new door to research into blind AIs.

II. RELATED WORK

A. Artificial Intelligence techniques for sound

Audio signal processing is one of the most important areas of Artificial Intelligence. In recent years, because deep learning has become more and more ubiquitous, it has been applied to audio processing and therefore has gained successes in applications such as speech recognition [8], [9] and text-to-speech [10], [11]. One common way to apply audio processing in deep learning is to convert the audio data into images and then process them as with other images. This can be done by generating spectrograms, which are 2D images representing sequences of spectra with time and frequency along two axes and with color representing the strength of a frequency component. A spectrogram can be obtained by applying Short-Time Fourier Transform (STFT) to the audio signal, and the STFT of a signal can be calculated using Fast Fourier Transform (FFT). A Mel-frequency scale can also be applied to spectrograms to create Mel-spectrogram, which is better for human perception and is used in [8] and [9]. In our work, as audio encoders for our blind AI, we compare three types of transformations that use a convolutional neural network (CNN) containing two 1D convolutional layers, a combination of FFT and a 2-layer fully connected network (FCN), and a

combination of Mel-spectrogram and a CNN containing two 2D convolutional layers, respectively.

B. Game AI based on sound

Previous work existed where research was conducted on AIs that play games using sound as the input. In Gaina and Stephenson [4], sound input supports were introduced to the General Video Game AI framework. Hedge et al. [5] showed that AIs that used both sound and audio inputs outperformed those that did not in a first-person shooter game. These studies show the potential of research on an AI that learns to play games from sound. However, in these studies, AIs were trained with only one sound design and were not used to evaluate the effectiveness of sound designs in games. Therefore, our work, to the best of our knowledge, is the first time an AI is used to train with multiple sound designs and to evaluate their effectiveness.

C. Proximal Policy Optimization

The Proximal Policy Optimization (PPO) [13] algorithm has been a popular reinforcement learning approach in recent years. PPO provides a reliable Trust Region Policy Optimization based on previous policy gradient methods and outperforms traditional methods such as Q-learning. Its policy loss function is as follows:

$$L_t^{CLIP}\theta = \hat{E}_t[min(\rho_t(\theta)\hat{A}_t, clip(\rho_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t)] \quad (1)$$

$$\rho_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} \quad (2)$$

$$\hat{A}_t = \sigma_t + (\gamma\lambda)\sigma_{t+1} + ... + (\gamma\lambda)^{T-t+1}\sigma_{T-1} \quad (3)$$

$$\sigma_t = r_t + \gamma V_\theta(s_{t+1}) - V_\theta(s_t) \quad (4)$$

In the equations above, s_t and a_t are the state and action at timestep t, respectively, $\pi_\theta(a_t|s_t)$ and $\pi_{\theta_{old}}(a_t|s_t)$ are the probability of a_t given s_t of the current policy and the previous policy, respectively. $V_\theta(s_t)$ is the value function of state s_t, and ϵ, λ, γ are clipping, Generalized Advantage Estimate (GAE) and discount factor, respectively. In the Fighting Game AI competition, PPO had been used in a number of studies [14]–[16] and achieved a remarkable success, especially as can be seen from the 2021 champion and runner-up both using PPO as part of their AI[1]. In addition, PPO is outstanding in audio processing tasks, such as audio-based navigation in a multi-speaker environment [17] and semantic audio-visual navigation, where objects' sounds are consistent with their semantic meaning [18]. In this work, we use PPO to train our blind AI.

[1]https://www.ice.ci.ritsumei.ac.jp/ ftgaic/index-R.html

D. Fighting Game AI

Many state-of-the-art algorithms tested their performance in the Fighting Game AI Competition, including deep reinforcement learning (DRL). Kim et al. [14] created a fighting game AI agent using deep reinforcement learning with self-play and Monte Carlo Tree Search (MCTS), and they later proposed a reinforcement learning agent to tune the environment slightly by reusing latent space obtained from a different environment [16]. In 2020, Tang et al. [19] proposed a method that combined the Rolling Horizon Evolution Algorithm with an opponent model and won the competition that year. In 2021, Liang et al. [15] extended the said work by proposing PPO with an Elo-based selection mechanism in which strong historical AIs are more frequently chosen during training. However, all these AIs use frame data provided by the competition until 2021. From 2022, the DareFightingICE competition [6], using the titular game platform, has been launched where entry AIs are required to play the game from audio data only and have no access to frame data information. Therefore, our work is the first effort to train an AI to play a fighting game using only sound as the input.

III. METHODOLOGY

A. Preprocessing

Raw audio data provided by DareFightingICE are in the form of a vector $s \in [-1, 1]^n$ containing n normalized audio samples. The original raw audio data size is 800 for each of the two channels (left and right), but each channel is padded with zeros so that it has 1024 samples for the sake of FFT in Java [6]. In our work, however, we choose to use only the 800 original samples for each channel. Motivated by the work of Hedge et al. [5], we propose and compare three encoders to process the audio data, before feeding them to a deep neural network, in the following. Figure 1 shows the architectures of these encoders, where all the parameters are based on previous work [5], but empirically adapted for our work.

1) 1D-CNN: The audio data s are downsampled by taking every 8th sample and fed to two 1D convolution layers. Downsampling helps reduce the computational complexity. In the end, an audio-feature vector of 32x5 is obtained.

2) FFT: The input audio data s are transformed to frequency domain using FFT, and the FFT data are converted to the natural logarithm of the magnitudes $sFFT = logFFT(s) \in R^{n/2}$. The resulting data are then downsampled and fed to a two-layer FCN. In the end, a one-dimensional audio-feature vector of 256 is obtained.

3) Mel-spectrogram: The input data s are transformed to frequency domain spectrogram with short-term Fourier transform (STFT), which is a sequence of Fourier transform of a windowed signal moved with a given hop. The frequencies are then processed with Mel scale. For hyperparameter setting, we choose a hop size of 10 ms, a window size of 25ms and 80 mel-frequency components. The spectrogram data are then fed to a network of two 2D convolutional layers. In the end, an audio-feature vector of 32x40x1 is obtained.
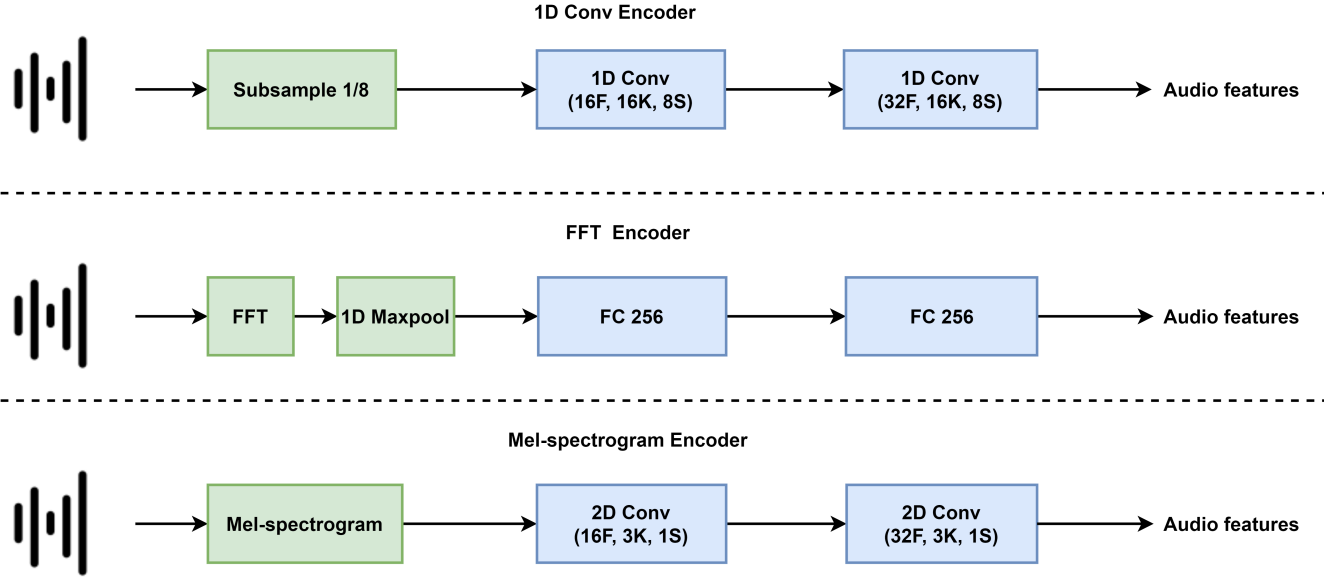
Fig. 1: Audio encoders: 1D-CNN (top), FFT (middle), and Mel-spectrogram (bottom).

TABLE I: Hyperparameters setting of PPO.

Hyperparameter	Value
Adam step size	$3 * 10^{-4}$
Number of epochs for optimizing surrogate	10
Mini batch size	64
Discount (γ)	0.99
GAE parameter (λ)	0.95
GRU hidden units	512

B. AI Design

DRL refers to a growing set of powerful algorithms which use deep neural networks to learn in environments that have high dimensional states and actions. In our work, as stated earlier at the end of II-C, we use PPO whose architecture and reward are described in the following.

1) Network Architecture: Our model consists of a chosen audio encoder given in the previous section, a gated recurrent unit [20], and a fully connected three-layered network to produce action probabilities. The fully connected layer network consists of three layers. The input of the network is the output of the audio encoder. There are two hidden layers, and each layer has 256 nodes. The output layer contains 40 nodes representing 40 actions[2] reused from [15], where the CROUCH action is omitted. We follow the PPO hyperparameters setting from [15], as shown in Table I. The architecture of our AI is depicted in Fig. 2.

2) Reward Definition: Following the recipe in previous work [21], we define the reward function as follows:

$$Reward = Reward_t^{offense} + Reward_t^{defense} \quad (5)$$

$$Reward_t^{offense} = HP_t^{opp} - HP_{t+1}^{opp} \quad (6)$$

$$Reward_t^{defense} = HP_{t+1}^{self} - HP_t^{self} \quad (7)$$

where t and $t+1$ represent the current frame step and the subsequent step, respectively; HP indicates the hit points of a character of interest and will be decreased if the character ($self$) receives damage from its opponent (opp), and in this work as well as the competition is initialized with a value of 400 at the beginning of a game round.

C. Competition Metrics

Here we propose two metrics for evaluating the effectiveness of a given sound design and/or a given audio encoder: win_{ratio}, and $avgHP_{diff}$. They are defined in a way that the more effective a sound design or an audio encoder is, the higher values these two metrics will be. We first train three blind AIs, each using a different audio encoder. The opponent AI in use is a weakened version of the MCTS sample AI in the preceding competition, MctsAi65 discussed in Khan et al. [6]. Each training lasts 900 game rounds. MctsAi65 is an AI whose execution time is reduced to 6.5 ms and uses frame data as the input. Because MctsAi65 was selected in [6] for human evaluation, as it is not too strong for visually impaired players to play, we choose it as the opponent AI.

We then evaluate the fighting performance of each trained blind AI by making it fight against the aforementioned opponent AI for 90 rounds. The ratio of the number of wins[3] over

---

[2]The actions in use consist of 2 throw-in-ground, 12 attack-in-ground, 3 skill-in-ground, 7 movement-in-ground, 2 guard-in-ground, 12 attack-in-air, and 2 skill-in-air actions.

[3]In the game, the round winner is either the one with a non-zero HP while its opponent's HP has reached zero or the one with the higher HP when the round-length limit of 60 s has reached.
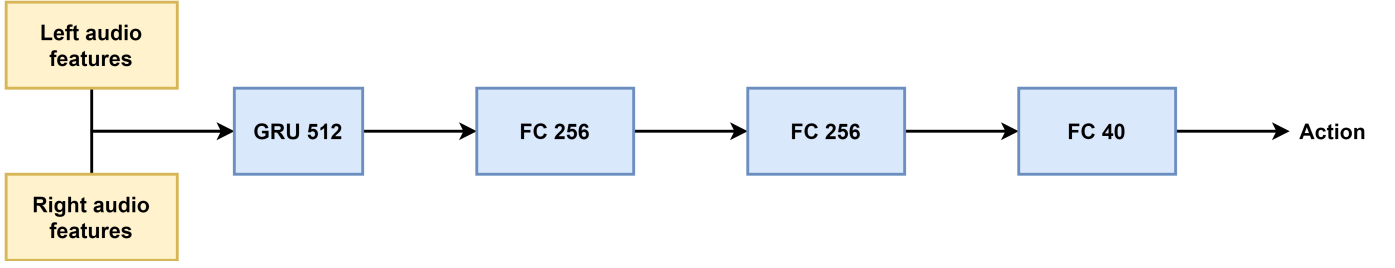
Fig. 2: Our blind AI architecture.

TABLE II: Performance of our blind AI with different sound designs and different audio encoders and when fighting against the opponent AI.

Sound design	Encoder	win_{ratio}	$avgHP_{diff}$
DareFightingICE	1D-CNN	0.33	-28.83
DareFightingICE	FFT	0.37	-40.5
DareFightingICE	Mel-spectrogram	**0.63**	**37.07**
FightingICE v4.5	1D-CNN	0.5	4.5
FightingICE v4.5	FFT	0.51	3.25
FightingICE v4.5	Mel-spectrogram	0.57	21.94

90 rounds, Eqn. (8), and the average HP difference at the end of a round between the trained AI and its opponent, Eqn. (9), are then calculated.

$$win_{ratio} = \frac{winning\ rounds}{total\ rounds} \qquad (8)$$

$$avgHP_{diff} = \frac{sum\ of\ HP_r^{self} - HP_r^{opp}\ for\ all\ r}{total\ rounds} \qquad (9)$$

IV. RESULTS AND DISCUSSIONS

We conduct experiments on two different sound designs, the sound design of DareFightingICE, in the 2022 DareFightingICE Competition, and the sound design of FightingICE, in the 2021 FightingICE. The competition metrics mentioned in the previous section are used to evaluate these sound designs. Each combination of encoder and sound design is assessed with three trials of training and performance evaluation.

Table II shows the average fighting performance among three trials of our blind AI in each combination of sound design and audio encoder. The DareFightingICE sound design outperforms the FightingICE sound design on both performance metrics for each encoder. This was expected because the sound design of DareFightingICE is an enhanced version of its predecessor and targets visually impaired players, although there is room for improvement due to its nature of being a sample and baseline sound design for the 2022 competition.

Now, we discuss AI behaviors[4]. In particular, we focus on differences in the AI behaviors when the sound designs are those of DareFightingICE and FightingICE, both using the

[4]Sample fight videos, where P1 is the blind AI and P2 is the opponent AI, of each of the six combinations of sound design and audio encoder are available on https://tinyurl.com/BlindAICoG2022.

Mel-spectrogram encoder. The blind AI cannot avoid a fireball skill of its opponent in the FightingICE sound design because there are no sound cues when the skill is fired. On the contrary, because the sound cue is played when the opponent releases a fireball skill in the DareFightingICE sound design, the blind AI seems to be able to recognize the sound cue and tries to avoid the skill as much as possible. Figures 4 and 5 show the resulting Mel-spectrogram of audio data when the opponent AI attacks with a fireball action in the DareFightingICE sound design and the FightingICE sound design, respectively. The game screen sequence when our blind AI tries to avoid a fireball attack is shown in Fig. 6.

The results above confirm that the proposed two metrics can be used to evaluate sound designs together with the evaluation done by the human judges in [6]. In the 2022 competition, the blind AI using the Mel-spectrogram encoder will be retrained from scratch to evaluate each entry sound design in the sound design track of the competition. In addition, the version trained in this work is made publicly available[5] as an official sample blind AI and will be used as a baseline AI in the AI track of the competition.

V. CONCLUSIONS

In this paper, we introduced a blind AI that only uses sound as the input on the DareFightingICE platform. We also evaluated the performance of the AI with different audio encoders when it fought against an opponent AI whose performance had been tuned for visually impaired players in previous work. Our blind AI was able to beat the opponent AI when the FFT encoder or the Mel-spectrogram encoder was used. It was also found that the Mel-spectrogram encoder was the best.

For evaluation of sound designs in the DareFightingICE Competition, we proposed two metrics that are the win ratio and the average HP difference when fighting against the aforementioned opponent AI. Our experiment results showed that the sound design of DareFightingICE was more effective than the sound design of FightingICE. This confirms that the proposed two metrics can be used in evaluation of entry sound designs in the competition.

In the future, we plan to improve our blind AI to make it better understand the game state from sound observations, and to use the AI in research to procedurally generate effective sound designs.

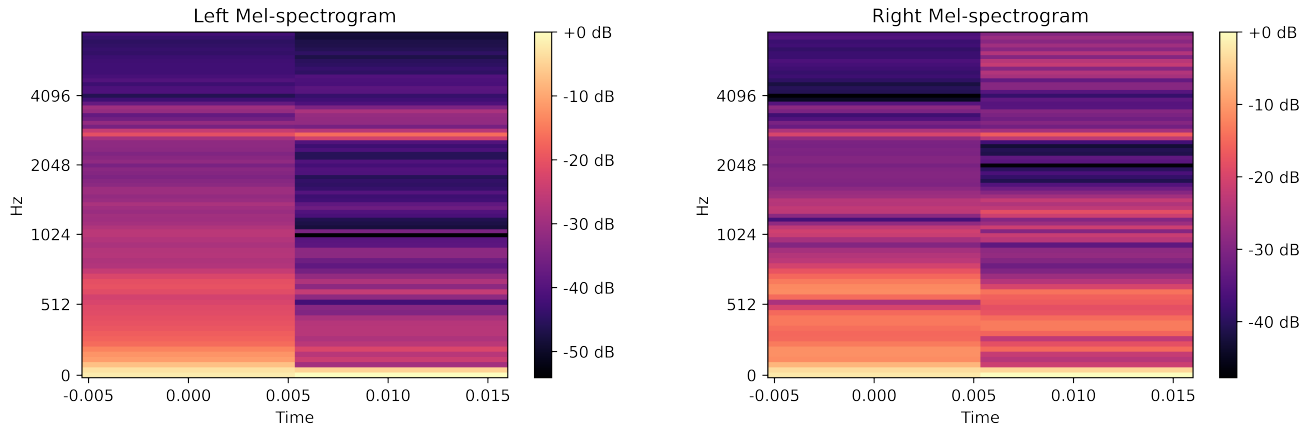[5]https://tinyurl.com/DareFightingICE/SampleAI/BlindAI

Fig. 4: Mel-spectrogram results of audio data when the opponent AI attacks with a fireball in the DareFightingICE sound design.
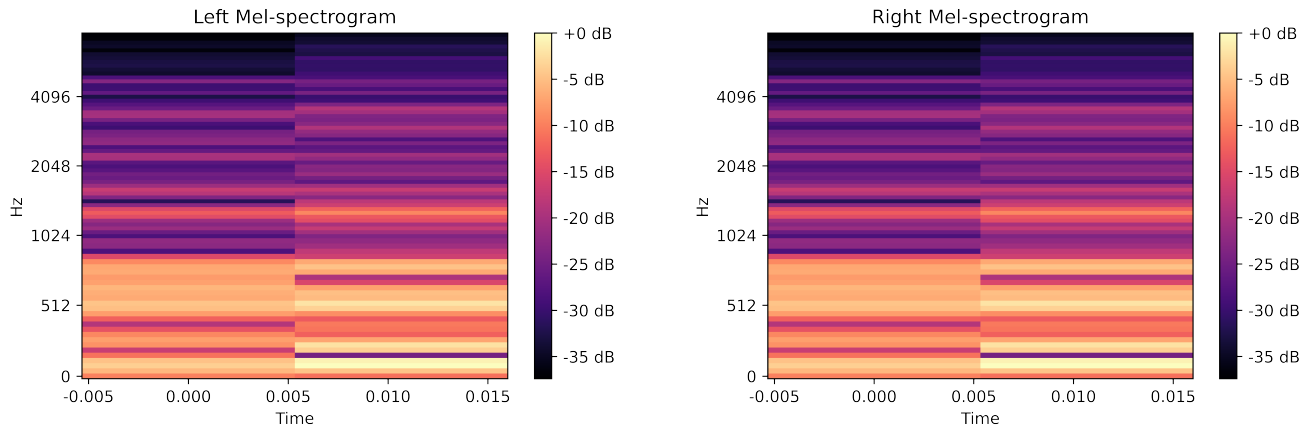

Fig. 5: Mel-spectrogram results of audio data when the opponent AI attacks with a fireball in the FightingICE sound design.

Fig. 6: Our blind AI tries to avoid a fireball attack of the opponent AI in the DareFightingICE sound design.

REFERENCES

[1] K. Tuuri, O. Koskela, J. Vahlo, and H. Tissari, "Identifying the Impact of Game Music both Within and Beyond Gameplay," in *Proceedings of International Conference on Entertainment Computing*, Springer, Cham, pp. 411-418, November 2021.

[2] J. Zhang, and X. Fu, "The influence of background music of video games on immersion," in *Journal of Psychology and Psychotherapy*, vol. 5, no. 4, p.1, 2015.

[3] F. Andersen, C. L. King, and A. A. Gunawan, "Audio Influence on Game Atmosphere during Various Game Events," in *Proceedings of Procedia Computer Science*, vol.179, pp.222-231, 2021

[4] R. D. Gaina and M. Stephenson, ""Did You Hear That?" Learning to Play Video Games from Audio Cues," in *Proceedings of the 2019 IEEE Conference on Games (CoG)*, pp. 1-4, 2019.

[5] S. Hegde, A. Kanervisto and A. Petrenko, "Agents that Listen: High-Throughput Reinforcement Learning with Multiple Sensory Systems," in *Proceedings of the 2021 IEEE Conference on Games (CoG)*, pp. 1-5, 2021.

[6] I. Khan, T. V. Nguyen, X. Dai, R. Thawonmas, "DareFightingICE

Competition: A Fighting Game Sound Design and AI Competition," *arXiv preprint arXiv:*2203.01556, 2022. (accepted for oral presentation at 2022 IEEE Conference on Games, August 2022).

[7] F. Lu, K. Yamamoto, L. H. Nomura, S. Mizuno, Y. Lee, and R. Thawonmas, "Fighting Game Artificial Intelligence Competition Platform," in *Proceedings of the 2nd IEEE Global Conference on Consumer Electronics (GCCE)*, pp. 320-323, October 2013.

[8] Q. Xu, A. Baevski, T. Likhomanenko, P. Tomasello, A. Conneau, R. Collobert, ... and M. Auli, "Self-Training and Pre-Training are Complementary for Speech Recognition," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3030-3034, 2021, doi: 10.1109/ICASSP39728.2021.9414641.

[9] G. Saon, Z. Tüske, D. Bolanos and B. Kingsbury, "Advancing RNN Transducer Technology for Speech Recognition," ICASSP 2021 - 2021 IEEE International Conference on Acoustics, in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5654-5658, 2021, doi: 10.1109/ICASSP39728.2021.9414716.

[10] T. Kaneko, K. Tanaka, H. Kameoka and S. Seki, "ISTFTNET: Fast and Lightweight Mel-Spectrogram Vocoder Incorporating Inverse Short-Time Fourier Transform," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6207-6211, 2022, doi: 10.1109/ICASSP43922.2022.9746713.

[11] Y. Ren, J. Liu, and Z. Zhao, "Portaspeech: Portable and high-quality generative text-tospeech," in *Proceedings of Conference and Workshop on Neural Information Processing Systems (NeurIPS)*, 2021.

[12] M. Kempka, M. Wydmuch, G. Runc, J. Toczek, and W. Jaskowski, "Vizdoom: A doom-based ai research platform for visual reinforcement learning," in *2016 IEEE conference on computational intelligence and games (CIG)*, pp. 1–8, 2016.

[13] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:*1707.06347, 2017.

[14] D. W. Kim, S. Park and S. I. Yang, "Mastering Fighting Game Using Deep Reinforcement Learning With Self-play," in *Proceedings of IEEE Conference on Games (CoG)*, pp. 576-583, 2020, doi: 10.1109/CoG47356.2020.9231639.

[15] R. Liang, Y. Zhu, Z. Tang, M. Yang and X. Zhu, "Proximal Policy Optimization with Elo-based Opponent Selection and Combination with Enhanced Rolling Horizon Evolution Algorithm," in *Proceedings of the 2021 IEEE Conference on Games (CoG)*, pp. 1-4, 2021.

[16] D. W. Kim, S. Y. Park, and S. I. Yang, "Reusing Agent's Representations for Adaptation to Tuned-environment in Fighting Game," in *Proceedings of International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 1120-1124, October 2021.

[17] P. Giannakopoulos, A. Pikrakis and Y. Cotronis, "A Deep Reinforcement Learning Approach To Audio-Based Navigation In A Multi-Speaker Environment," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3475-3479, 2021.

[18] C. Chen, Z. Al-Halah, and K. Grauman, "Semantic audio-visual navigation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 15516-15525, 2021.

[19] Z. Tang, Y. Zhu, D. Zhao and S. M. Lucas, "Enhanced Rolling Horizon Evolution Algorithm with Opponent Model Learning," in *IEEE Transactions on Games*, doi: 10.1109/TG.2020.3022698, 2020.

[20] Y. Takano, W. Ouyangy, S. Ito, T. Harada and R. Thawonmas, "Applying Hybrid Reward Architecture to a Fighting Game AI," in *Proceedings of IEEE Conference on Computational Intelligence and Games (CIG)*, pp. 433-436, August 2018.

[21] K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1724–1734, 2014.