TRANSNAV
http://www.transnav.eu
the International Journal on Marine Navigation and Safety of Sea Transportation
Volume 13
Number 4
December 2019
DOI: 10.12716/1001.13.04.22

Developing Generative Adversarial Nets to Extend Training Sets and Optimize Discrete Actions

R.L. Zhang & M. Furusho
Kobe University, Kobe, Japan

ABSTRACT: This study proposes the use of generative adversarial networks (GANs) to solve two crucial problems in the unmanned ship navigation: insufficient training data for neural networks and convergence of optimal actions under discrete conditions. To achieve smart collision avoidance of unmanned ships in various sea environments, first, this study proposes a collision avoidance decision model based on a deep reinforcement learning method. Then, it utilizes GANs to generate enough realistic image training sets to train the decision model. According to generative network learning, the conditional probability distribution of ship maneuvers is learnt (action units). Subsequently, the decision system can select a reasonable action to avoid the obstacles due to the discrete responses of the generated model to different actions and achieve the effect of intelligent collision avoidance. The experimental results showed that the generated target ship image set can be used as the training set of decision neural networks. Further, a theoretical reference to optimize the optimal convergence of discrete actions is provided.

1 INTRODUCTION

Unmanned ocean transportation is sure to revolutionize maritime unmanned navigation in the future. In October 2016, the Norwegian Maritime Administration and the Norwegian Coastal Authority established the world's first autonomous ship test zone in the Trondheim fiord as well as the Norwegian Forum for Autonomous Ships. It marked the promotion of unmanned ship research to the national level in Norway. In early December 2018, the "Suomenlinna II" polar passenger ferry successfully crossed the test area near the port of Helsinki, under the unmanned state, and passed the remote sea trial. The intelligent decision-making module is the "brain" of unmanned ships. It involves various technologies such as route optimization, risk warning, smart decision-making, and energy efficiency management. It can make most decisions based on the external navigation environment information, ship internal information, and shore-based support information. For example, it can take excellent navigation decisions and send control commands to the execution unit to make appropriate decisions (Finn et al., 2010).

Taking ship collision avoidance as an example, the intelligent decision-making module obtains the actual navigation situation around the ship according to the targets acquired by the radar, AIS, ship-borne infrared camera, visible light camera, and other sensors and its fusion information, and conducts a risk information analysis for the surrounding targets (Trucco, 2008). If there is a dangerous target, the collision avoidance decision is made through the intelligent collision avoidance technology combined with the current position, direction it is heading toward, and speed. The instructions formed by the decision, such as changing the course and changing the speed, are sent to the rudder control system. In the process of collision avoidance of ships, the information transmitted by multiple sensors and equipment is continuously integrated, making sure the collision avoidance scheme is adjusted in time (Wang, 2007). The core of the current research is how the decision-making module can satisfy the optimal navigational operations in all types of extreme offshore environments.

Therefore, in the risk assessment and early warning research of unmanned ship navigation, it is necessary to focus on the unmanned ship in complex navigation conditions (such as ports, straits, canals, and other intensive waters), ship collision avoidance and hydrometeorology, geographical environment, traffic situation, and other issues. This research is based on ship sensor data acquisition and training optimization of decision neural networks (Mazurowski, 2008). An intelligent risk warning model and method suitable for unmanned ships under complex navigation conditions is formed to approach real-time warning of ships (Scheffer, 2012).

In the intelligent decision-making research, an intelligent fusion correlation analysis is carried out on static and dynamic targets and navigation conditions around unmanned ships. Intelligent theories, such as deep learning, knowledge base, and situation calculation, are applied. Research on ship navigation intelligent decision theory based on ship navigation system information and shore-based support information, break through the key technologies of ship autonomous meteorological navigation technology. Technologies such as ship collision avoidance, reef avoidance, anti-shelf integration, and smart processing of navigation information support autonomous decision-making of ship navigation (Capraro, 2006).

To achieve intelligent collision avoidance function of unmanned ships in various environments, a collision avoidance decision module based on deep reinforcement learning is proposed to make autonomous decisions under various conditions (Mnih, 2015). In the Cyber confrontation game, the DeepMind team collects enough data for training; however, in the real navigation environment, it is difficult to obtain data in a rich and varied nautical environment. In particular, various types of encounter ships have different points of observation in different situations, and it is difficult to predict their future path of navigation (Sarukkai, 2000). In the process of calculating the global solution optimal solution, the decision model is difficult to differentiate due to the discrete action as a result, the global optimal solution cannot converge. Therefore, this study proposes a generative adversarial networks (GANs) model to solve the problem of neural network training data, and the combination of GAN and deep reinforcement learning to solve the convergence problem of optimal action under discrete action unit conditions.

2 RELATED WORK

2.1 *The principle of GAN*

GAN is a new method proposed by Goodfellow (2014) to train generated models. The method of GAN includes the generation and discrimination of two “adversarial” models. The generated model (G) is used to capture the data distribution, and the discriminant model (D) is used to estimate the probability that a sample is derived from real data rather than the generated data. Both the generator and discriminator are common convolutional networks as well as fully connected networks. The generator generates a sample from the stochastic vector, and the discriminator discriminates between the generated sample and the real training set sample.

This optimization process can be attributed to a two-player minimax game problem. Both purposes can be achieved through a backpropagation method. A well-trained generation network can transform any noise vector into a sample similar to the training set. This noise can be seen as the encoding of the sample in a low dimensional space. The generator generates meaningful data based on the stochastic vectors. In contrast, the discriminator learns how to determine real and generated data and then passes the learning experience to the generator, thereby, enabling the generator to generate more workable data based on the stochastic vectors. Such a trained generator can have many uses; one of them being environmental generation in automatic navigation.

The specific process to obtain various target ships is shown in Figure 1. First, a few stochastic vectors are fed as input in the generator network, and fake data are subsequently generated by the generator. The aforementioned fake data can correspond to a few ship state pictures or navigation data such as AIS data of a nearby encounter of the given ship or the path planning data after the ship route is updated. We input the fake data to the discriminator, and the discriminator determines whether the input data are real data or fake data generated by the generator. The similarity between the generated data and the real data gradually increases, then the discriminating ability required by the discriminator also increases accordingly. Furthermore, the generator and the discriminator share a mutually competitive and mutually adversarial equation. The generated data are considered to sufficiently mirror real data, and therefore, the fake data input by the generator appear sufficiently realistic. The approximate accuracy of the discriminator in this case is 50%. This corresponds to the target ship image data that are required in a critical sea environment.

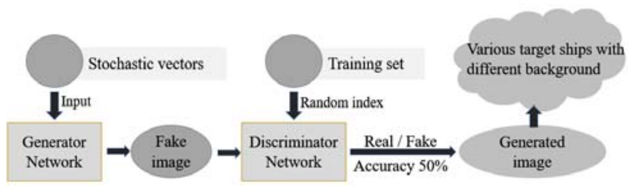

Figure 1. Applying GAN generate various target ships with different background

2.2 *GAN application examples*

In the previous study, the authors proposed using the GAN to build an executable method for maritime navigation route re-planning. The unmanned ship can independently generate new routes based on the environmental information around the ship before any possible danger occurs and can even interrupt the remote support. With the total generative time less than one second, the trained model helps the ship avoid obstacle or any latent disaster. In addition, GAN is easy to embed into the framework of reinforcement learning. For example, when using deep reinforcement learning to solve collision avoidance problems, GAN can be used to learn the conditional probability distribution of an action. The agent (own ship) can select a reasonable route based on the response of the generated model to different actions.

2.3 *Global optimality of discrete actions*

The mathematical equivalent of deep reinforcement learning can be considered as Markov decision processes in discrete time defined by five factors (S, 'a', P, 'r', γ) with a neural network instead of Q-value (Zhang, 2017). Here, S is the finite state space (state set) in which the unmanned ship is located; a is the behavior decision space of the unmanned ship; i.e., the set of all actions or reactions in a space in any state, for example, the left rudder, right rudder, acceleration, deceleration, heeling, and stopping, etc. P_a (S) = P(S'|S, a) , where P is a conditional probability indicating that the unmanned ship reaches the next state under state s and action 'a'. The probability of the state S', r_a (S| S') is a reward function, which represents the excitation obtained by the unmanned ship from the state S to the next state S' in the case of action a. $\gamma \in (0,1)$ is reward attenuation factor, the reward at the next instant time t is attenuated using this factor.

In actual navigation practices, completing a collision avoidance process may require different operational coordination methods. These operations are incoherent and discrete. Further, the way different people respond to the same event may be different. Generally, both the reward 'r' and the attenuation function 'γ' are different. Thus, a method to converge an optimal global action group is required. This study will discuss the possibility of combining GAN and deep reinforcement learning to solve this problem.

3 GAN METHOD FOR GENERATING TRAINING DATA SETS

3.1 *Conventional acquisition of target ship data set method*

The acquisition of related target ship data is usually carried out at the position of the ship's bridge. Here, we can observe the state of the target ship's navigation, and then, photograph the target ship that will be encountered. The author carried out a seven-day summer research voyage on the university training ship "FUKAE MARU". A data set of a total of 4,000 images of valid target ships were obtained. However, this is not enough for the training of neural recognition networks. The target recognition and classification require a large number of data sets for both training as well as target identification, if the positional posture of the target ship is to be perceived. Therefore, a new approach to get training data is needed.

3.2 *Example of the GAN generate lifeboat image data*

This section mainly demonstrates the use of the demo provided by Big-GAN, as shown in Figure 2, which is the process of generating a lifeboat using GAN. In the images obtained at different times in the generation process, the sea surface appears in the image generated in Figure 2.a, and there are a group of fuzzy things in the middle that cannot identify the object. In Figure 2.b, the orange upper body and black rubber are common in the lifeboat. The hull, shown in Figure 2.c, is an almost completed generated image, and our eye and training model can roughly identify this as a lifeboat. We can input different stochastic vectors and combine the real data input by the discriminator to get a large number of high-quality image data sets. As shown in Figure 3, we got three different types of lifeboats. More importantly, the background of the lifeboat could also be changed to provide a large amount of training data for our unsupervised decision model. It was observed to be much richer than the target ship data collected from the real seas.

a. GAN generating… b. GAN generating… c. GAN almost generated

Figure 2. Target ship image data-generating process

GAN-generated image_1 GAN-generated image_2 GAN-generated image_3

Figure 3. Example of generated lifeboats with different backgrounds

The most important part of this research is to obtain a data set of the target ship with sufficiently high quality and quantity. As shown in Figure 4, a portion of the entire large-scale lifeboat target image data set is shown. These images were not taken by the camera and were generated entirely from the GAN model. Using different truncation and noise_seed, our model could generate various encounter situations at sea as shown in Figures 5 and 6, as well as the various forms possible for the target ship at the time of the encounter; including various types of accidents such as collision, stranding, fire and loss of goods. We obtained the training data set for lifeboats, ocean liner data sets as shown in Figures 7 and 8, as well as data sets for various other types of marine moving targets.

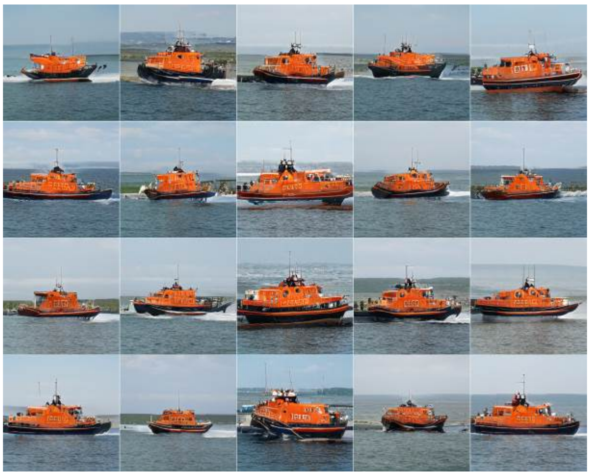

Figure 4. Large-scale data generated with different ship status and backgrounds. (Lifeboat)
Truncation: 0.14; Noise_seed: 0

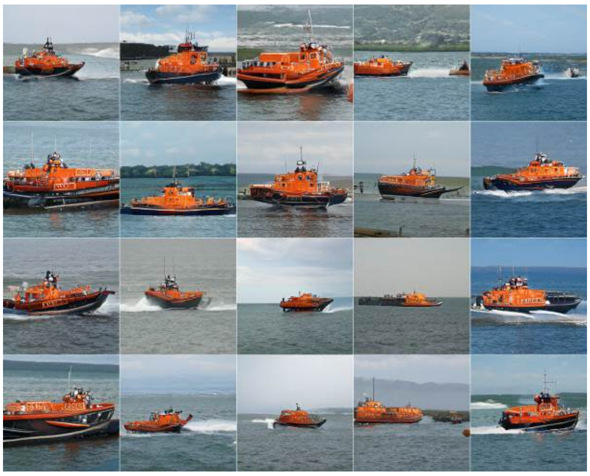

Figure 5. Large-scale data generated with different ship status and backgrounds. (Lifeboat)
Truncation: 0.28; Noise_seed: 0

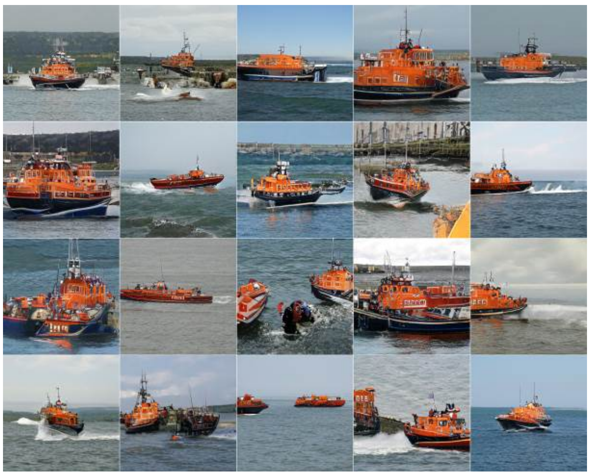

Figure 6. Large-scale data generated with different ship status and backgrounds. (Lifeboat)
Truncation: 0.56; Noise_seed: 0

Figure 7. Large-scale data generated with different ship status and backgrounds. (Ocean liner)
Truncation: 0.10; Noise_seed: 4

Figure8. Large-scale data generated with different ship status and backgrounds. (Ocean liner)
Truncation: 0.10; Noise_seed: 8

As shown in Figure 4–6, the larger the truncation, the greater the diversity of the generated samples. In fact, truncation controls the truncation distance of the hidden variable distribution (generally Gaussian), which is the sampling range. Therefore, it is not difficult to understand its role in diversity. As shown in Figures 7–8, the influence of the value of noise_seed on the generated result is the initial condition of each sample generation, and the final result will be different, which can be used to improve the generation diversity. When training image recognition models of convolutional neural networks and decision models, such as deep reinforcement learning, the quality of the input data considerably affects the effect of the training results. The target image data set generated by GAN has the same image size and image density, which can easily solve the problem of inconsistent input data during the training process. In addition, the GAN model solves many of the scene data that are difficult to obtain in a real navigation environment, making it possible to use large-scale data entry for deep reinforcement learning.

4 GAN WITH POLICY GRADIENT FOR OPTIMIZE DISCRETE ACTIONS.

Although the number of variants of GAN and their versatility is increasing, their adversarial-thinking has not changed. In other words, a discriminator that can identify the real data and generate the data is added in the generation process, so that the generator G and the discriminator D can compete with each other. The role of D is to try to distinguish the real data and the generated data to improve the generated data that can confuse D. When D can no longer separate the true and false data, it is considered that G has reached a stable state.

The numerous advantages are summarized as follows:

– It can generate better samples;
– No need to make inferences about hidden variables during training;
– The model only uses backpropagation without the need for a Markov chain;
– G's parameter update does not come directly from the data sample, however, uses backpropagation from D;
– In theory, as long as the differentiable function can be used to construct D and G, it can be combined with deep neural network to make a deep-generation model.

Part of the agent decision module to complete a collision avoidance evaluation may require different action coordination. These discrete actions are difficult to complete the mathematical differential operation; thus, it is necessary to find a way to converge a global optimal action group. The last of the above advantages is precisely its limitation. In discrete data, data are not continuous like image processing and can be differentiated. Therefore, GAN cannot be realized for discrete data.

As shown in Figure 9, when using the deep reinforcement learning model to solve the optimal decision problem, GAN and deep reinforcement learning are combined to select a reasonable global optimal action combination. In Figure 9.a, an adversarial idea is portrayed, where real data from the sea environment plus generated data of G are required to train D. However, from the content described in the related work section, the discrete output of G is obtained, which makes it difficult for D to return a gradient to update G, and therefore, a few changes need to be made. As shown in Figure 9.b, the value returned by the policy network is G. The existing dot is called current state. The generated next dot operation is called action, because D needs to be a complete sequence score. Thus, the Monte Carlo tree search (MCTS) is used to complete the various possibilities of each action. D rewards these complete sequences, passes information back to G, and updates G by enhanced learning. This is done to use the reinforcement learning method to train a generation network that can generate the global action set.

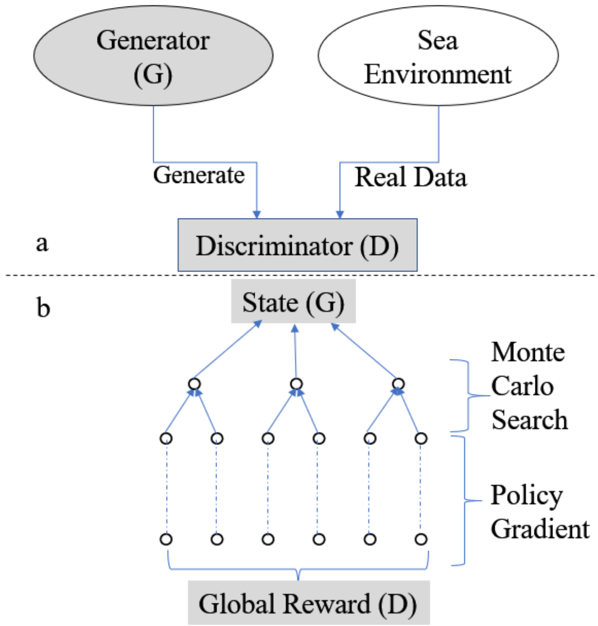

Figure 9. Policy gradient convergence of discrete data

5 CONCLUSION

This study uses the GAN method to implement a large number of generations of decision model training data sets. In the generated data set, according to the settings of the two parameters of truncation and noise_seed, different target ship image data can be obtained. Apart from the different positions of the target ship and state data, the encounter situation of different backgrounds and scenes and image data of various target ships under dangerous conditions are obtained. The target ship image data set generated by the generative adversarial model is useful to train the ship target recognition neural network under different environmental backgrounds, however, for the ship's motion situation prediction, collision avoidance decision, etc., it provided discontinuous data. At the time of processing, GAN did not satisfy this demand. Therefore, this study combines GAN with the idea of policy gradient in deep reinforcement learning, and a method for solving the convergence problem of discrete global action set is creatively proposed.

REFERENCES

Capraro, G. T., Farina, A., Griffiths, H. & Wicks, M. C. (2006). Knowledge-based radar signal and data processing: a tutorial review. IEEE Signal Processing Magazine, 23(1), 18-29.

Finn, A. & Scheding, S. (2010). Developments and challenges for autonomous unmanned vehicles. Intelligent Systems Reference Library, 3, 128-154.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S. ... & Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680).

Mazurowski, M. A., Habas, P. A., Zurada, J. M., Lo, J. Y., Baker, J. A. & Tourassi, G. D. (2008). Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. Neural networks, 21(2-3), 427-436.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G. ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529.

Sarukkai, R. R. (2000). Link prediction and path analysis using Markov chains1. Computer Networks, 33(1-6), 377-386.

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., ... & Pascual, M. (2012). Anticipating critical transitions. science, 338(6105), 344-348.

Trucco, P., Cagno, E., Ruggeri, F. & Grande, O. (2008). A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation. Reliability Engineering & System Safety, 93(6), 845-856.

Wang, X., Yadav, V. & Balakrishnan, S. N. (2007). Cooperative UAV formation flying with obstacle/collision avoidance.

Zhang, R. L. & Furusho, M. (2017). Conversion timing of seafarer’s decision-making for unmanned ship navigation. TransNav: International Journal on Marine Navigation and Safety of Sea Transportation, 11.