IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 8, 2019, accepted January 6, 2020, date of publication January 13, 2020, date of current version January 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2966237

Maneuver Decision-Making of Deep Learning for UCAV Thorough Azimuth Angles

HONGPENG ZHANG AND CHANGQIANG HUANG
Institute of Aeronautics Engineering, Air Force Engineering University, Xi'an 710038, China
Corresponding author: Hongpeng Zhang (1152951370@qq.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51579209.

ABSTRACT Maneuver decision-making directly determines the success or failure of air combat. To improve the dogfight ability of unmanned combat aerial vehicles and avoid the deficiencies of traditional methods, such as poor flexibility and a weak decision-making ability, a maneuver method using deep learning is proposed. A total of 72 different maneuvers are constructed, and 544320 states are designed. Flight simulations are conducted under these different states to obtain corresponding future azimuth angles. A deep neural network is trained with these offline data, and thus, the network possesses state prediction capability. A situation assessment function and a decision objective function based on azimuth angles are constructed. During air combat, the optimal maneuver is selected from the maneuver library according to the predicted state and the decision objective function. The results of air combat simulations indicate that the unmanned combat aerial vehicle (UCAV) can win the air combat game by the proposed method in a balanced situation and can meet missile launching conditions in an adverse situation. The operational time of this method has been reduced by 0.01 s compared with the comparison method.

INDEX TERMS Unmanned combat aerial vehicle, decision-making, deep learning, air combat simulation, situation assessment.

I. INTRODUCTION

Unmanned aircraft systems (UASs) have been successful in replacing manned aircraft in a variety of commercial and military aerial missions. UAVs have been used to travel over the sensed environment to collect data, and they exhibit faster data collection while achieving a high packet delivery rate and low energy usage [1]. However, because of the challenging and dynamic nature of air combat, these missions are solely accomplished by manned platforms.

Since the late 1970s, NASA has funded the development of a computer program for the simulation of a dogfight between two fighter planes. The goal of the program is to develop a solution technique for computing the near-optimal maneuvering decisions of unmanned combat aerial vehicles (UCAVs). Maneuvering decision-making can be summarized as follows: first, several elemental maneuvers are chosen based on human-pilot air combat experiences, which are indicated by three control variables (roll angle, normal overload and tangential overload); then, at each decision point, the fighter predicts the position of the elemental trial maneuvers after the same time period; finally, a value is placed on each of the maneuvers by answering questions about the state of each maneuver relative to the positions of the enemy and the UCAV itself. The maneuver that scores the highest value is chosen as the next to be performed. Austin *et al.* proposed seven elementary maneuvers and selected the most beneficial maneuver by means of a scoring matrix [2]–[4]. The authors stated that the maneuver selection only guaranteed optimality in the short term and only with respect to the chosen heuristic scoring function. Even so, the method produced some maneuvering decisions similar to those made by experienced human pilots. The elementary maneuvers were also used by Sun *et al.* [5] to build an air combat decision support system. Virtanen *et al.* [6], [7] proposed a moving horizon decision-making model to solve the air combat game. In this approach, the time horizon of the original game is truncated, and a feedback Nash equilibrium of the dynamic game lasting for only a limited planning horizon is determined and implemented at each decision stage. Although a limited planning horizon can mitigate the computational complexity, long planning horizons are essential to making good maneuver choices

The associate editor coordinating the review of this manuscript and approving it for publication was Yue Cao.

during air combat. Huang *et al.* [8] introduced fuzzy logic and Bayesian inference into a moving horizon decision-making model. By means of fuzzy logic, the authors built maneuver decision factor functions according to the state or situation, such as the relative distance vector between a UCAV and an opponent fighter, azimuth angles, height and velocity of fighters. The situation constitutes the maneuver decision objective function, the maneuver decision is to optimize the decision objective function, and the function weights can be changed adaptively using Bayesian inference. Ren *et al.* [9] proposed a decision-making model based on a structure-varied discrete dynamic Bayesian network (SVDDBN). This model is composed of three parts: threat evaluation, target value assessment and situation assessment. According to the evaluative results of the above three parts, the SVDDBN inference algorithm is applied for current mission decision-making.

Mcgrew *et al.* [10], [11] regarded the air combat game as a dynamic programming problem; further, assuming that computing the optimal policy using an exact DP is intractable because of the exponential growth of the state-space size with the number of state-space variables, they adopted approximate dynamic programming (ADP) to make decisions during the air combat game. However, ADP limited aircraft within the horizontal plane, and only three maneuvers were used. Wang *et al.* [12] proposed a robust maneuvering decision method. They improved the membership function of situation evaluation during the decision process on the basis of the MIN-MAX decision method to make the situation function have a certain insensitivity to changes in the air combat situation. The specific process of this decision method can be summarized as follows: based on the information of the UCAV and the enemy aircraft at the current decision time t, the control command of all the actions in the maneuver library is sent to the flight dynamical model for maneuver trial; all possible locations of the UCAV in the next stage are obtained, and the corresponding situation assessment function values are computed; finally, the maneuver with the highest situation assessment function value is chosen. A genetic fuzzy-based artificial intelligence algorithm is used for UCAV control in simulated air combat missions [13]. A combination of particle swarm optimization and game theory is utilized for the cooperative decision-making of multiple UCAVs [14]. Holsapple studied the autonomous decision-making method of the UCAV in uncertain environments for intelligence, surveillance, and reconnaissance (ISR) tasks [15].

Since Deepmind's team used recent advances in training deep neural networks to develop a novel artificial agent [16]–[18], termed a deep Q-network [19], which was able to surpass the performance of all previous algorithms and achieve a level comparable to that of a professional human games tester, deep reinforcement learning (DRL) has become a useful method to solve control problems and decision-making problems [20], [21]. Unlike the above decision-making method, DRL decision-making does not need to make trial maneuvers and predict the positions at each decision-making stage. It chooses a maneuver by means of a deep neural network that is able to output the maneuver with the highest situation assessment function value according to the state directly without a maneuver trial [22]. Zhang *et al.* [23] proposed a heuristic Q-Network method integrating expert experience and used expert experience as a heuristic signal to solve the super-horizon air combat maneuver decision-making problem. To find the shortest flight path of a UCAV to avoid enemy missiles, Lee and Kim [24] proposed a new reinforcement learning algorithm that enhances exploration by amplifying the imitation effect. However, these methods require plenty of time to sample, and convergence of the neural network that outputs the maneuver with the highest situation assessment function value is difficult.

While the aforementioned approaches achieved some success, there are still several deficiencies: these methods utilized only seven typical flight maneuvers designed by NASA scholars, which cannot meet the requirement of air combat maneuvering; and a fighter predicts the position of the elemental trial maneuvers after the same time period through the dynamic model, but predicting the position after long time period takes much more time. Therefore, the goal of the paper is to improve upon these aspects in terms of real-time implementation, increased planning horizons, and increased optimality. These objectives are achieved via a deep neural network and a novel maneuver library.

To avoid the deficiencies of the above decision-making methods, we propose a novel tactical decision-making framework for the autonomous air combat of a UCAV using a deep neural network. The inputs of the neural network are the current state including the current roll angle, the target roll angle, the pitch angle and the velocity, and the outputs are the future pitch angle and the yaw angle. We choose only the pitch angle and the yaw angle as neural network outputs. Further, the other factors are contained abstractly in the changes of the pitch angle and the yaw angle. Then, the two angles constitute the maneuver decision objective function. To find the increased optimality of a maneuver, we construct a novel maneuver library consisting of 72 different maneuvers. These maneuvers are based on the seven element maneuvers designed by NASA scholars, and their normal overloads and tangential overloads are set to a maximum. The reason these overloads are set to a maximum is that maximum overloads ensure maneuverability. Finally, one on one air combat under different conditions is simulated with enemy fighters using the decision-making model in [12].

II. UCAV DYNAMICAL MODEL AND MANEUVER LIBRARY

A. DYNAMICAL UCAV MODEL

In the course of researching maneuver decision-making, the UCAV motion dynamics model adopts the normal overload, tangential overload and roll angle as control parameters. To simplify the complexity of the problem, the angle of attack and the sideslip angle are taken as zero, and the ground coordinate system is treated as the inertial system. Meanwhile, the effects of the rotation of the earth are overlooked. The UCAV dynamical model is shown as follows:

$$\begin{cases} \dot{x}_t = v_t \cos\gamma_t \cos\psi_t \\ \dot{y}_t = v_t \cos\gamma_t \sin\psi_t \\ \dot{z}_t = v_t \sin\gamma_t \\ \dot{v}_t = g(n_{tx} - \sin\gamma_t) \\ \dot{\gamma}_t = \frac{g}{v_t}(n_{tz}\cos\mu_t - \cos\gamma_t) \\ \dot{\psi}_t = \frac{g}{v_t\cos\gamma_t} n_{tz}\sin\mu_t \end{cases} \tag{1}$$

where x_t, y_t, and z_t indicate the positions of the UCAV in the inertial coordinate system; γ_t is the pitch angle; ψ_t is the yaw angle; v_t is the velocity; and g is the acceleration of gravity. Roll angle μ_t, tangential overload n_{tx}, and normal overload n_{tz} are control parameters. The simulation step is set to 0.01 s. At every simulation step, the derivatives of v_t, γ_t, and ψ_t are updated, and then new derivatives of x_t, y_t, and z_t are obtained. Thus, a new point consisting of three coordinates can be computed by their derivatives. After some time, numerous points are acquired, these points are connected into lines, and we can obtain the flight trajectory of the aircraft. The model parameters are shown in Fig. 1.

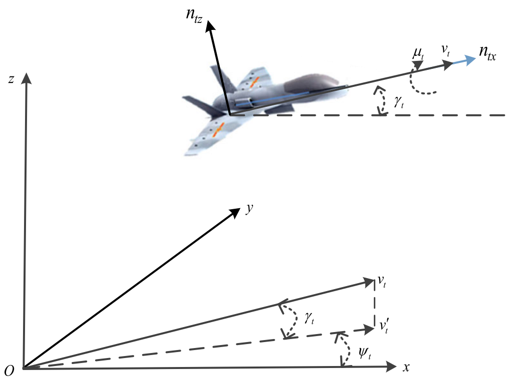

FIGURE 1. Definition diagram of the dynamical model parameters.

B. UCAV MANEUVER LIBRARY

NASA scholars have designed seven typical flight maneuvers: (1) continued stable flight, (2) maximum acceleration flight, (3) maximum deceleration flight, (4) maximum overload left-turn flight, (5) maximum overload right-turn flight, (6) maximum overload upward flight, and (7) maximum overload downward flight. However, although these maneuvers are usually applied in air combat simulations, they are not flexible enough. Thus, we design 72 different maneuvers according to the dynamical model. The specific method is to take a roll angle at intervals of 5° between −180° and 180° and maximize the normal overload at the same time. Therefore, 72 maneuvers are constructed. As shown in Fig. 2, there are 72 flight trajectories simulated for six seconds according to the dynamical model. Because of the existence of the angular speed of the roll angle, the aircraft needs to roll from zero; therefore, there is an opening in the graph composed of trajectories.

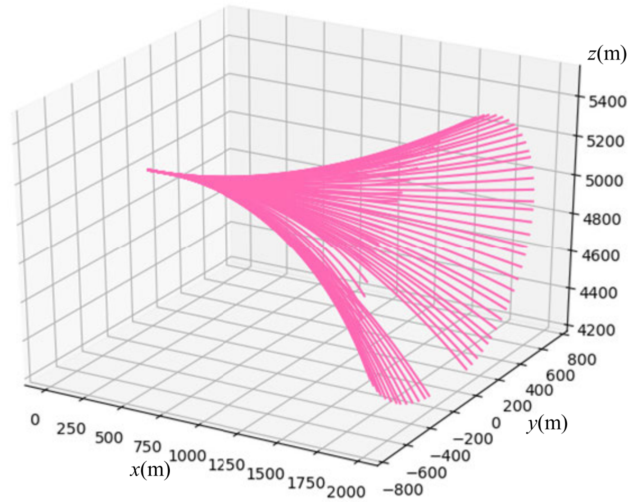

FIGURE 2. Seventy two different maneuvers.

III. MANEUVER DECISION-MAKING MODEL USING DEEP LEARNING

A. REVIEW OF DEEP LEARNING

Deep learning is a new research direction in the field of machine learning. Deep learning originates from artificial neural networks. An artificial neural network will be in the local optimum and gradient disappearance or gradient explosion as the number of layers increases [25]. Deep learning can overcome the shortcomings of artificial neural networks and acquire distributed representation of data by means of combining low-level features to form more abstract high-level features. Unlike traditional shallow learning, deep learning emphasizes depth in neural networks and is able to transform the representation of a sample in the original space to the new space by layer-by-layer feature transformation, which consequently promotes the classification and regression accuracy. A deep neural network, which has a certain number of neural nodes and multilevel network structures, selects the appropriate input and output layers and is then trained to establish a functional relationship from input to output. The schematic structure of a deep neural network is shown in Fig. 3.

B. DECISION-MAKING MODEL USING DEEP LEARNING

1) AIR COMBAT SITUATION INFORMATION

The process of maneuver decision-making is the process of selecting the maneuver that is the most beneficial to one's own side on the basis of air combat situation information. The combat effectiveness of airborne weapons such as missiles and aircraft guns is affected by the air combat situation directly or indirectly, and different situations correspond to different decisions. Therefore, situation information is crucial to maneuver decision-making. There are three basic kinds of situation information:

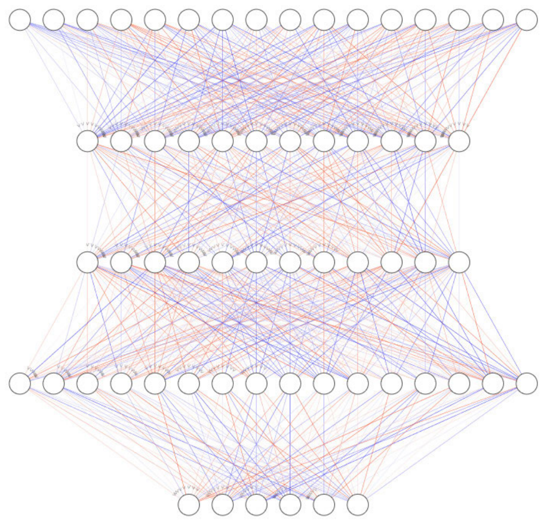

FIGURE 3. Schematic structure of a deep neural network.

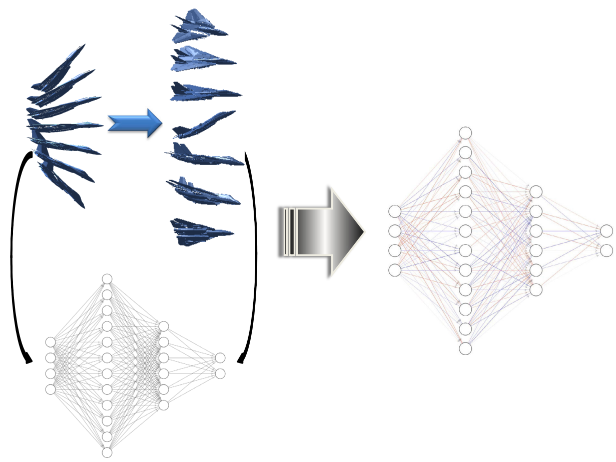

FIGURE 4. Deep neural network training.

1) Distance situation: the distance between the two sides of air combat in three dimensional space, the absolute altitude and the relative altitude of the two sides.
2) Velocity situation: the absolute and relative velocities of the two sides.
3) Angle situation: The angles between the velocity vector of one fighter and the line-of-sight vector in the horizontal and vertical directions, the angles between the velocity vector of another fighter and the line-of-sight vector in the horizontal and vertical directions, and the angular velocity of the line-of-sight vector.

2) DECISION-MAKING MODEL USING A DEEP NEURAL NETWORK

We propose a novel decision-making model by means of a deep neural network that includes two parts: deep neural network training and deep neural network decision-making. The network training procedure is shown in Fig. 4.

As shown in Fig. 4, the upper left indicate the process of sampling: regard the direction of the projection on the horizontal plane of the velocity vector of the UCAV as the direction of the x-axis; the z-axis is perpendicular to the horizontal plane; and the y-axis is defined according to the right hand rule. Thus, a coordinate system is established for sampling in the flight simulation. Every time sampling starts, the UCAV is put at the coordinate origin with different initial pitch angles γ_0, different initial roll angles μ_0, different initial velocities v_0, and different target roll angles μ_{tar}, which represent the 72 maneuvers. Then, a flight simulation lasting for a period of time T is conducted, and the pitch angle γ_{tar} and yaw angle ψ_{tar} are recorded at the end of each simulation. Therefore, samples consisting of inputs and outputs that are used for training the deep neural network are acquired. The inputs are made up of different groups of γ_0, μ_0, v_0, and μ_{tar}, and corresponding outputs are made up of different groups of γ_{tar} and ψ_{tar}.

By means of uniformly sampling four inputs and flight simulations under these different conditions, the corresponding outputs are acquired. As a result, a large number of flight samples are obtained. Then, the deep neural network is trained so that it has the ability to predict future states based on the current situation and maneuver. The sampling range and interval are shown in Table 1. The pitch angles, yaw angles, speeds, and control quantities are divided evenly. Then, flight is simulated under these different situations, and 544320 groups of data consisting of the current situation, the control quantity and the future state (pitch angle and yaw angle) are obtained. The deep neural network is trained with these data, and thus, the network is capable of predicting future situations according to the current situation. The maneuver with the best future situation can be selected from the maneuver library using the deep neural network. Meanwhile, the operational speed of the deep neural network is fast, which can meet the real-time requirement of decision-making.

Fig. 5 indicates the procedure of decision-making on the basis of the deep neural network. At each decision time-step, the UCAV figures out the outputs including the pitch angles and yaw angles of all maneuvers by means of the deep neural network. These angles in the sampling coordinate system are then converted into angles in the inertial coordinate system, and the maneuver with the maximum decision-making target function value can be found according to the converted angles. Therefore, the UCAV can use this maneuver to gradually bring the enemy fighter into the missile attack range.

3) DECISION-MAKING TARGET FUNCTION

The current decision-making method usually makes decisions according to the situation assessment function, which is also used for assessing the situation. The situation assessment function includes the factors of angle, distance, and velocity. However, the goal of air combat within-the-horizon is to meet

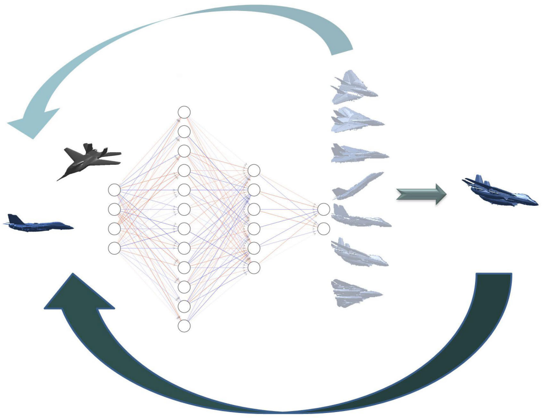

FIGURE 5. Deep neural network decision-making.

missile launch conditions as soon as possible. Only after meeting the angle conditions of the launching missile does it make sense to consider other factors. Thus, we separate the processes of assessment and decision-making by designing a novel decision objective function:

$$Tar = \alpha_1 \left|\gamma_{pre} - \gamma_{los}\right| + \alpha_2 \left|\psi_{pre} + \psi_t - \psi_{los}\right| \tag{2}$$

where γ_{pre} and ψ_{pre} are the pitch angle and the yaw angle of the UCAV in the sampling coordinate system, respectively; and $alpha_1$ and α_2 are the corresponding weights of the angles in decision-making. Different weights suggest different degrees of importance for the pitch angle and the yaw angle. Both the weights are set to 0.5. γ_{los} and ψ_{los} are the pitch angle and the yaw angle, respectively, of line-of-sight vector $\boldsymbol{R_{los}}$ in the inertial coordinate system. The goal of the deep neural network is to select the maneuver with the lowest decision objective function value from all maneuvers, and thus, the UCAV is able to rapidly meet the conditions of missile launching

IV. AIR COMBAT SITUATION ASSESSMENT AND VICTORY OR DEFEAT JUDGEMENT

A. SITUATION ASSESSMENT FUNCTION

To describe the air combat process objectively, a situation assessment function is designed. The existing decision-making method usually regards the maneuver with the highest situation assessment function value as the best maneuver. However, in our decision-making method, the situation assessment function is used only to evaluate the status in air combat and is not involved in maneuver decision-making.

1) ANGLE SITUATION

He most significant thing in air combat is to fire, followed by various tactical maneuvers and, finally, flight performance. Regardless of whether an aircraft gun or a missile is used to attack the enemy, it is necessary to get in the right place and meet the angle requirements. Thus, the angle situation is the most important thing in air combat. The angle situation function is

$$Q_\gamma = \begin{cases} \frac{3}{2}\cos\gamma_r + 1 - \frac{\sqrt{3}}{2}, & |\gamma_r| < \frac{\pi}{6} \\ \frac{6}{\pi}(\frac{\sqrt{3}}{2} - 1)(|\gamma_r| - \frac{\pi}{3}), & \frac{\pi}{6} \le |\gamma_r| \le \frac{\pi}{2} \end{cases} \tag{3}$$

$$Q_\psi = \begin{cases} \frac{3}{2}\cos\psi_r + 1 - \frac{\sqrt{3}}{2}, & |\psi_r| < \frac{\pi}{6} \\ \frac{5}{3} - \frac{4}{\pi}\left|\psi_r\right|, & \frac{\pi}{6} \le |\psi_r| < \frac{5\pi}{12} \\ 5 - \frac{12}{\pi}\left|\psi_r\right|, & \frac{5\pi}{12} \le |\psi_r| < \frac{\pi}{2} \\ -1, & \frac{\pi}{2} \le |\psi_r| \le \pi \end{cases} \tag{4}$$

where γ_r is the relative pitch angle of the velocity vector with respect to the line-of-sight vector, and ψ_r is the relative yaw angle of the velocity vector with respect to the line-of-sight vector. (x_a, y_a, z_a) are the coordinates of the UCAV in three dimensional space, and (x_b, y_b, z_b) are the coordinates of the enemy fighter in three dimensional space. When calculating the situation assessment function value of the UCAV, the line-of-sight vector is set as $\boldsymbol{R_{los}} = (x_b - x_a, y_b - y_a, z_b - z_a)$, and when calculating the situation assessment function value of the enemy fighter, the line-of-sight vector is set as $\boldsymbol{R_{los}} = (x_a - x_b, y_a - y_b, z_a - z_b)$.

2) DISTANCE SITUATION

The main factor that has an impact on the distance situation is the missile attack range. Thus, the distance situation function can be defined as

$$Q_R = \begin{cases} 1, & R < R_D \\ e^{\frac{-(R-R_D)^2}{2\sigma^2}}, & R_D \le R \end{cases} \tag{5}$$

where R_D is the missile attack range, σ is the standard deviation of the attack range, and R is the distance between both sides of air combat. When the enemy aircraft is within the range of the missile, the distance situation function value is always equal to 1. When the enemy aircraft is out of the range of the missile, the distance situation function value decreases as the distance increases. The real-time altitude of the aircraft is also important because if the altitude is too low, there may be a plane crash, and it is not conducive to the performance of the aircraft if the altitude is too high. Therefore, the altitude situation function can be defined as

$$Q_H = \begin{cases} -1, & H < H_L \ or \ H > H_U \\ 0, & H_L \le H \le H_U \end{cases} \tag{6}$$

where H_L is the lower limit of height, H_U is the upper limit of height, and H is the real-time height of the aircraft.

Taking into account the angle and distance, the air combat situation assessment function can be defined in the form of weighted summation as:

$$Q = \omega_1 Q_{\gamma_a} + \omega_2 Q_{\psi_a} + \omega_3 Q_{\gamma_b} + \omega_4 Q_{\psi_b} + \omega_5 Q_R + \omega_6 Q_H \tag{7}$$

where Q_{γ_a} and Q_{ψ_a} are the pitch angle situation value and the yaw angle situation value of the UCAV, respectively; Q_{γ_b} and Q_{ψ_b} are the pitch angle situation value and the yaw angle situation value of the enemy fighter, respectively; Q_{γ_a} is the distance situation value; Q_H is the altitude situation value; and ω_i is the weight of the corresponding situation value.

B. VICTORY OR DEFEAT JUDGEMENT OF AIR COMBAT

1) DECISION-MAKING METHOD OF THE OPPONENT

The maneuver libraries of both sides are equal. The decision-making method of the enemy fighter is identical to the method in [12], which can be summarized as selecting the maneuver with the maximum situation function value in the next stage at each decision-making time-step.

2) VICTORY OR DEFEAT JUDGMENT

To win the air combat, it is indispensable to meet the missile launch condition first. Therefore, victory or defeat judgment of air combat can be constructed as

$$\begin{cases} R < R_{fire} \\ |\psi_{ra}| < \frac{\pi}{6} \\ |\gamma_{ra}| < \frac{\pi}{6} \\ Q_b < Q_a \end{cases} \tag{8}$$

where R represents the distance between both sides of air combat, R_{fire} is the optimum missile launching distance, γ_{ra} is the relative pitch angle of the velocity vector of the UCAV with respect to the line-of-sight vector, ψ_{ra} is the relative yaw angle of the velocity vector of the UCAV with respect to the line-of-sight vector, and Q_a and Q_b are the situation assessment functions of the UCAV and the enemy fighter, respectively. First, the missile launch condition must be satisfied, namely, the first three in (8). Then, to verify the effectiveness of this decision-making method, after meeting the missile launch condition, the situation function value of the UCAV must be greater than the situation function value of the enemy fighter. Then, the UCAV wins the combat, or vice versa. The block diagram of air combat judgment is shown in Fig. 6.

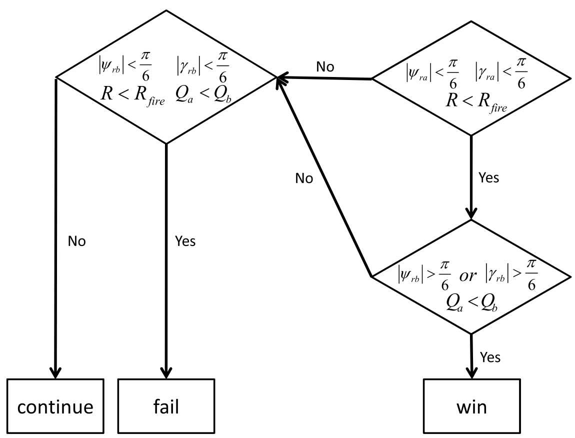

FIGURE 6. Judgment block diagram.

V. DEEP NEURAL NETWORK TRAINING

The sampling range and interval are shown in Table 1. Using dynamical model (1) to simulate the flying process of the UCAV, 544320 samples are acquired in total. Because of the large number of samples, a deep neural network is used to construct the mapping from the current flight situation and the control parameters to future flight situations. The inputs are the pitch angles, yaw angles, speeds, and control quantities, and the outputs are the pitch angle and yaw angle of the aircraft. A fully connected network is utilized due to its strong ability to fit a function. The input shape of the network is 4, and the output shape is 2. The network has five hidden layers. The hidden layers and the output layer are shown in Table. 2. To reduce overfitting, dropout layers are added in the network [25]. The neural network gradient descend algorithm is RMSprop, the loss function is the mean square error, and the learning rate is 0.001.

TABLE 1. Sampling range and interval.

Sampling	Range	Interval
Initial pitch angle(°)	(-50，50)	5
Initial roll angle(°)	(-180，180)	10
Initial velocity(m·s^{-1})	(250，340)	10
Target roll angle(°)	(-180，180)	10

TABLE 2. Structure of the deep neural network.

Serial number	Layer type	Activate function	Output shape
1	Dense	tanh	512
2	Dropout	-	512
3	Dense	ReLU	512
4	Dropout	-	512
5	Dense	ReLU	256
6	Dropout	-	256
7	Dense	ReLU	256
8	Dropout	-	256
9	Dense	ReLU	256
10	Dropout	-	256
11	Dense	tanh	2

Fig. 7 indicates the change of accuracy of the training set and the validation set. As shown in Fig. 7, because the training sample is evenly sampled and the network has a certain depth, the deep neural network has a good prediction accuracy.

VI. AIR COMBAT SIMULATIONS AND ANALYSIS

A. COMMON PARAMETERS

Common parameters: the initial flight speed is set to 250 m/s, the maximum flight speed is set to 400 m/s, and the minimum flight speed is set to 90 m/s. The optimum launching distance

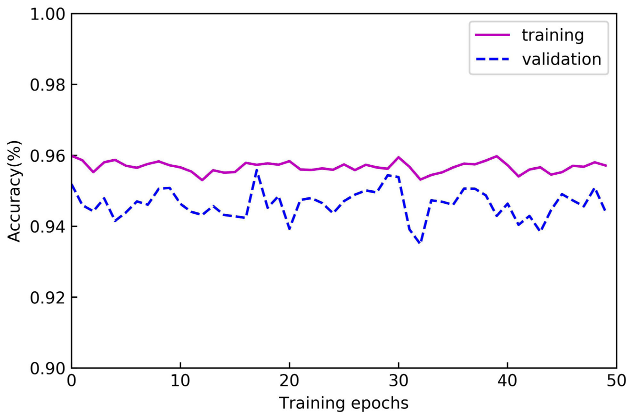

FIGURE 7. **Accuracy curves.**

of missiles R_{fire} is set to 4000 m, the attack range of missiles R_D is set to 5000 m, the standard deviation σ is set to 100 m, the lower limit of height H_L is set to 500 m, the upper limit of height H_U is set to 12000 m, and the initial altitude is set to 5000 m. The maximum flight time is set to 100 s, the decision-making period is set to 1 s, and the initial roll angle is set to 0.

B. AIR COMBAT SIMULATION

First initial condition: the initial position of the enemy is (0, 7000, 5000), the initial pitch angle is 0, and the initial yaw angle is $-25°$; the initial position of the UCAV is (0, 0, 5000), the initial pitch angle is 0, and the initial yaw angle is 0. As seen from the above parameters, the UCAV is at a disadvantage. Fig. 8 shows the results of air combat confrontation. The solid lines represent the trajectory and corresponding flight information of the UCAV, and the dashed lines represent the trajectory and corresponding flight information of the enemy fighter. (d) and (e) indicate the included angles and the distance between the two sides, respectively. In (d), the two black dashed lines represent 30° and $-30°$, the two solid lines represent projections of the included angles between the UCAV velocity vector and the line-of-sight vector in the horizontal and vertical planes (yaw angle and pitch angle), respectively. The two dashdot lines represent projections of the included angles between the enemy fighter velocity vector and the line-of-sight vector in the horizontal and vertical planes, respectively. In (e), the black dashed line represents 4000 m, and the solid line indicates the distance between the two sides. Once the distance is less than 4000 m and the two angles of a fighter are less than |30°|, the missile launching conditions are met. If only one side meets the launching conditions, it wins. If both sides meet the conditions, the side with the higher situation function value wins.

As shown in Fig. 8, the initial situation function value of the UCAV is less than that of the enemy. According to the curves of the roll angle change, it can be seen that the UCAV chose to change its roll angle to 100°, and at approximately 1.5 s, its roll angle was 100°. At 4 s, the UCAV changed its roll angle to 90° from 100°. Simultaneously, the enemy changed its roll angle to $-110°$ at 1 s and changed it to -70° at 3 s. At 4 s, it changed its roll angle to 90° as well. As shown in (d), the two angles of the enemy were less than |30°| at approximately 2 s, the two angles of the UCAV were less than |30°| at approximately 5.5 s, and because the distance between the two sides was more than 4000 m, the air combat game continued. At 7 s, the distance was less than 4000 m; therefore, both sides met the missile launching conditions. Since the situation function value of the UCAV was less than that of the enemy, the UCAV was defeated. However, this does not mean that the deep neural network decision making method is useless because the UCAV still met the missile launching conditions even if it was at a disadvantage at the beginning. Meanwhile, considering several minor adjustments during flying, it can be concluded that the designed maneuver library is effective compared with traditional maneuver library in which the difference between roll angles of different maneuvers is usually 45° or 90°.

Second initial condition: the initial position of the enemy fighter is (7000, 7000, 5000), the initial pitch angle is 0, and the initial yaw angle is -135°; the initial position of the UCAV is (0, 0, 5000), the initial pitch angle is 0, and the initial yaw angle is 0°. As seen from the above parameters, at the beginning of the simulation, there was an intersection angle of 45° between the flight direction of the UCAV and the line-of-sight, and the UCAV was at a disadvantage. Fig. 9 shows the results of the air combat game.

As shown in Fig. 9, because the UCAV was at a disadvantage at the beginning of the air combat simulation, the situation function value was less than that of the enemy fighter. The UCAV chose the maneuver with a 100° roll angle at the beginning and then changed the maneuver multiple times. As shown in (d), the two angles of the enemy were less than |30°| at the beginning. At approximately 2 s, the two angles of the UCAV were less than |30°|. At 9 s, the situation function value of the UCAV was more than that of the enemy; however, its advantage disappeared before long. Finally, both sides met the missile launching conditions.

Third initial condition: the initial position of the enemy fighter is (5000, 5000, 5000), the initial pitch angle is 0, and the initial yaw angle is 45°; the initial position of the UCAV is (0, 0, 5000), the initial pitch angle is 0, and the initial yaw angle is $-135°$. As seen from the above parameters, at the beginning of the simulation, neither side posed a threat to the other. The results of the air combat confrontation are shown in Fig. 10.

Because the azimuth angles of both sides were a maximum, much more time was required to meet the launching conditions, and there was a significant increase in the simulation time. According to the situation function value diagram, in the first four seconds, the situation function value gap between the two sides was small. According to the roll angle diagram, the UCAV flew to the right in the first 15 s. The enemy fighter flew to the right first and then flew to the left, causing itself to be chased by the UCAV

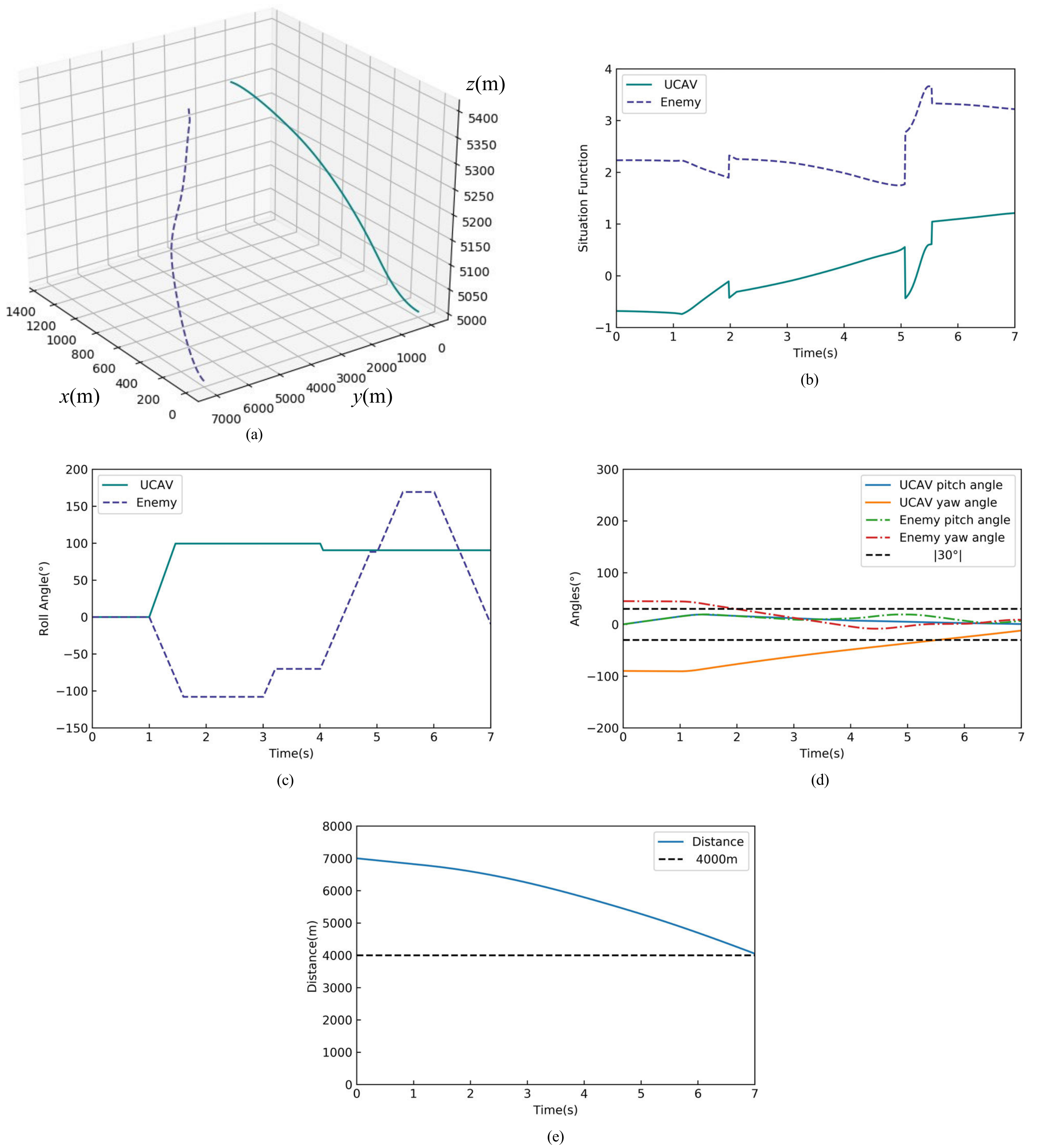

FIGURE 8. The results of an air combat confrontation under the first condition. (a) Air combat trajectories. (b) Simulation function value. (c) Change of the roll angle. (d) Projections of the included angles in the horizontal and vertical planes. (e) Distance between the two sides.

and putting the UCAV in a dominant position for nearly 40 s. Therefore, it can be concluded that it is beneficial to consider only the azimuth angles in decision-making. After that, the enemy aircraft kept changing its flight direction to get rid of the chase, and the UCAV followed it by means of turning. Because 180° is equivalent to $-180°$, the yaw angle of the two sides changed to $-180°$ from 180° and changed to 180° from $-180°$ immediately, as shown in (d). In (e), the distance between the two sides was less than 4000 m from approximately 44 s to 53 s; however, neither side met the angle conditions for missile launching. Finally, because the UCAV took the lead in meeting the missile launching conditions and the situation function value of the UCAV was greater than that of the enemy fighter,

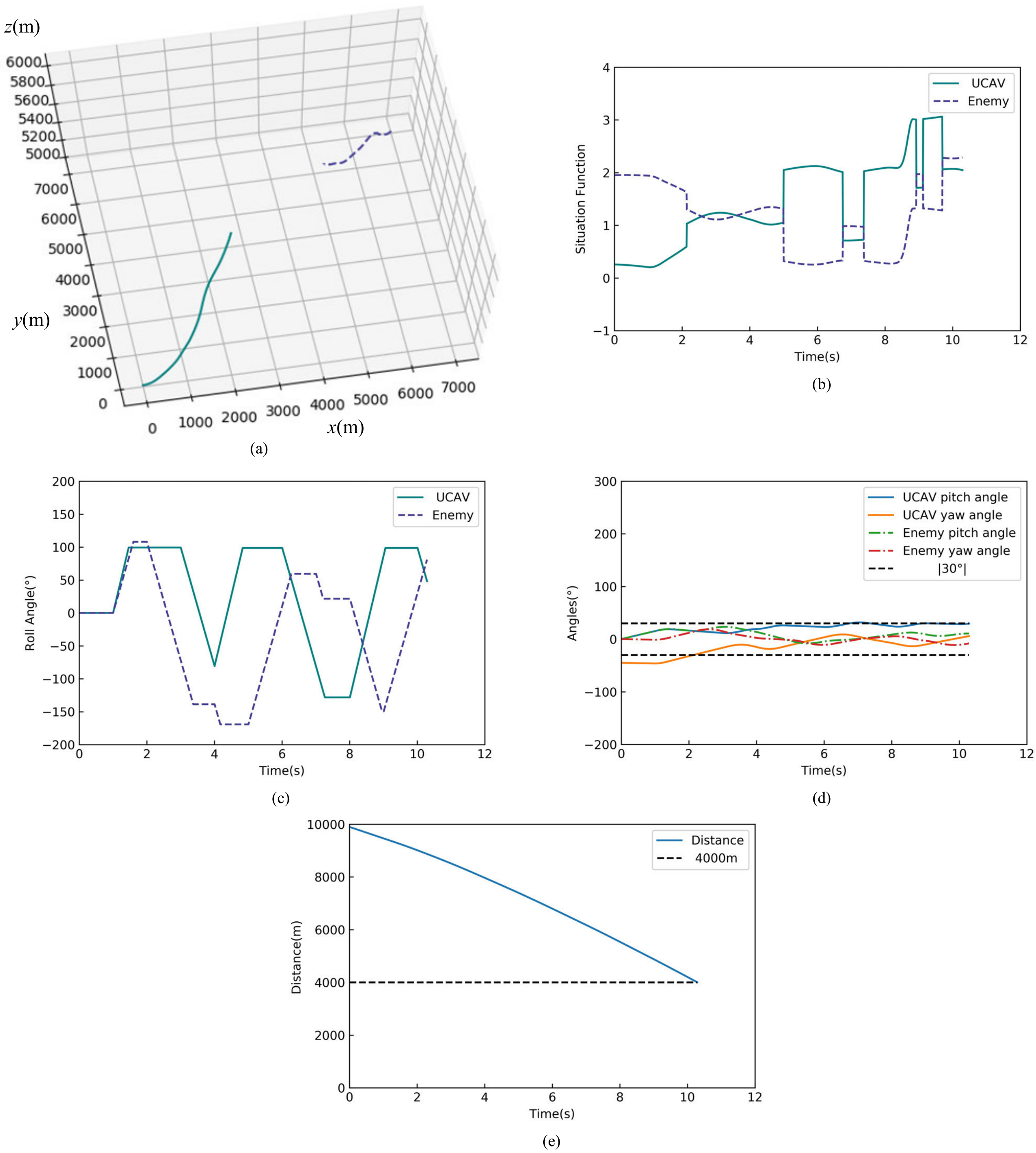

FIGURE 9. **The results of an air combat confrontation under the second condition. (a) Air combat trajectories. (b) Simulation function value. (c) Change of the roll angle. (d) Projections of the included angles in the horizontal and vertical planes. (e) Distance between the two sides.**

the UCAV won. Third initial condition is mutually beneficial, because neither side of the air combat posed a threat to each other, therefore, we can conclude that our decision-making method can win the air combat in mutually beneficial circumstances.

Fourth initial condition: the initial position of the enemy fighter is (7000, 7000, 5000), the initial pitch angle is 0, and the initial yaw angle is $-135°$; the initial position of UCAV is (0, 0, 5000), the initial pitch angle is 0, and the initial yaw angle is $-135°$. As seen from the above parameters, at the beginning of the simulation, the UCAV was chased by the enemy fighter. The result of the air combat confrontation is shown in Fig. 11.

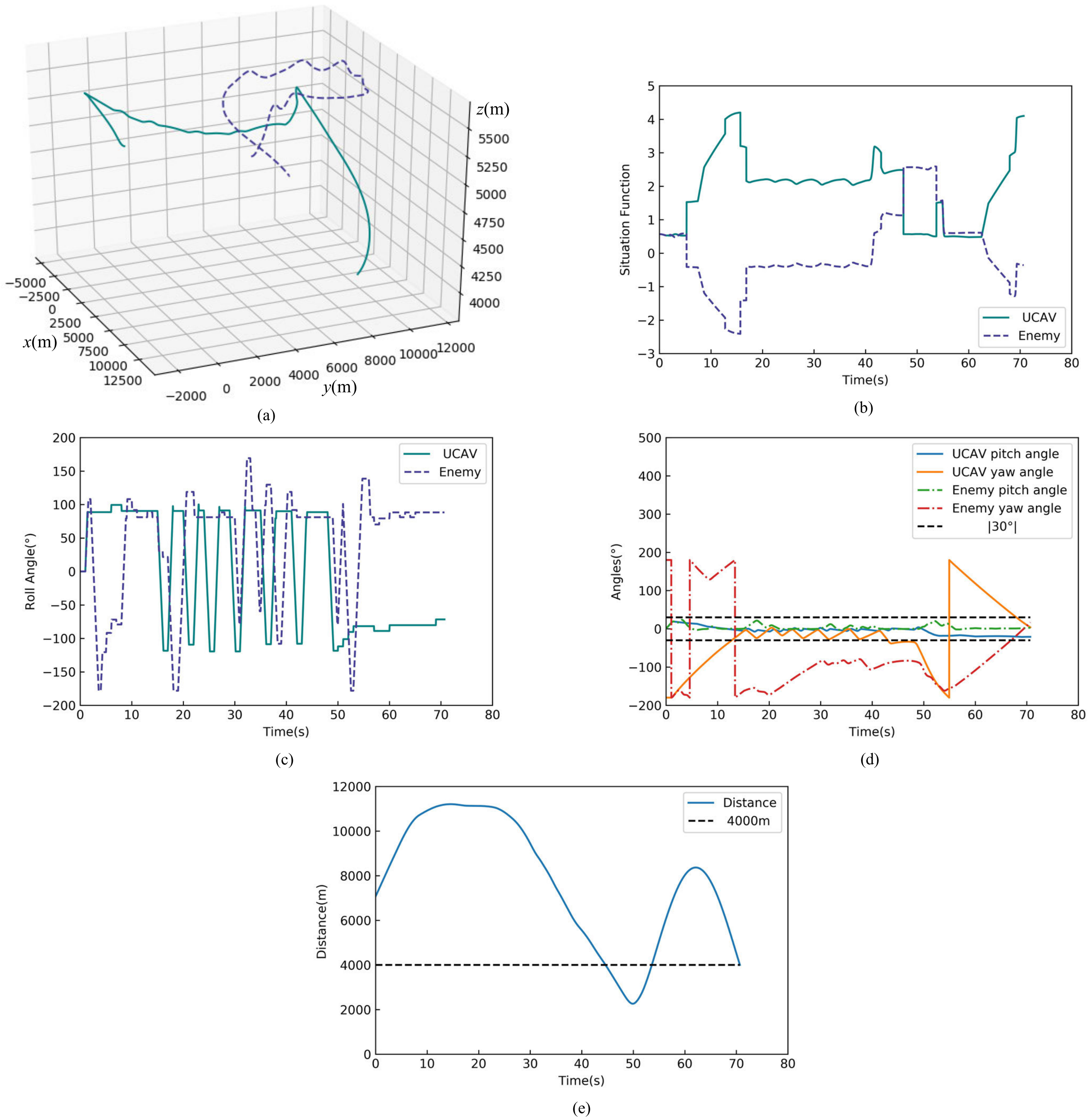

FIGURE 10. **The results of an air combat confrontation under the third condition. (a) Air combat trajectories. (b) Simulation function value. (c) Change of the roll angle. (d) Projections of the included angles in the horizontal and vertical planes. (e) Distance between the two sides.**

Because the UCAV was chased by the enemy fighter at the beginning of the air combat simulation, its situation function value was lower than that of the enemy fighter. Then, the UCAV chose to fly to the left to avoid being chased. The enemy fighter chose different flying directions, as shown in Fig. 11, because it needed to maximize its situation function value, which contains not only its own azimuth angles but also the azimuth angles of the UCAV. However, even if the enemy fighter was at a strong advantage at the beginning and its situation function value was always higher than that of the UCAV, the UCAV still met the conditions of missile launching after approximately 14 s, as shown in (d). Because both sides met the conditions of missile launching at the same time and the enemy fighter's situation function value was higher than that of the UCAV, the enemy fighter won. Finally, the average time taken to select the optimal maneuver is 0.06 s in the deep neural network decision-making and 0.07 s in the enemy fighter's decision-making.

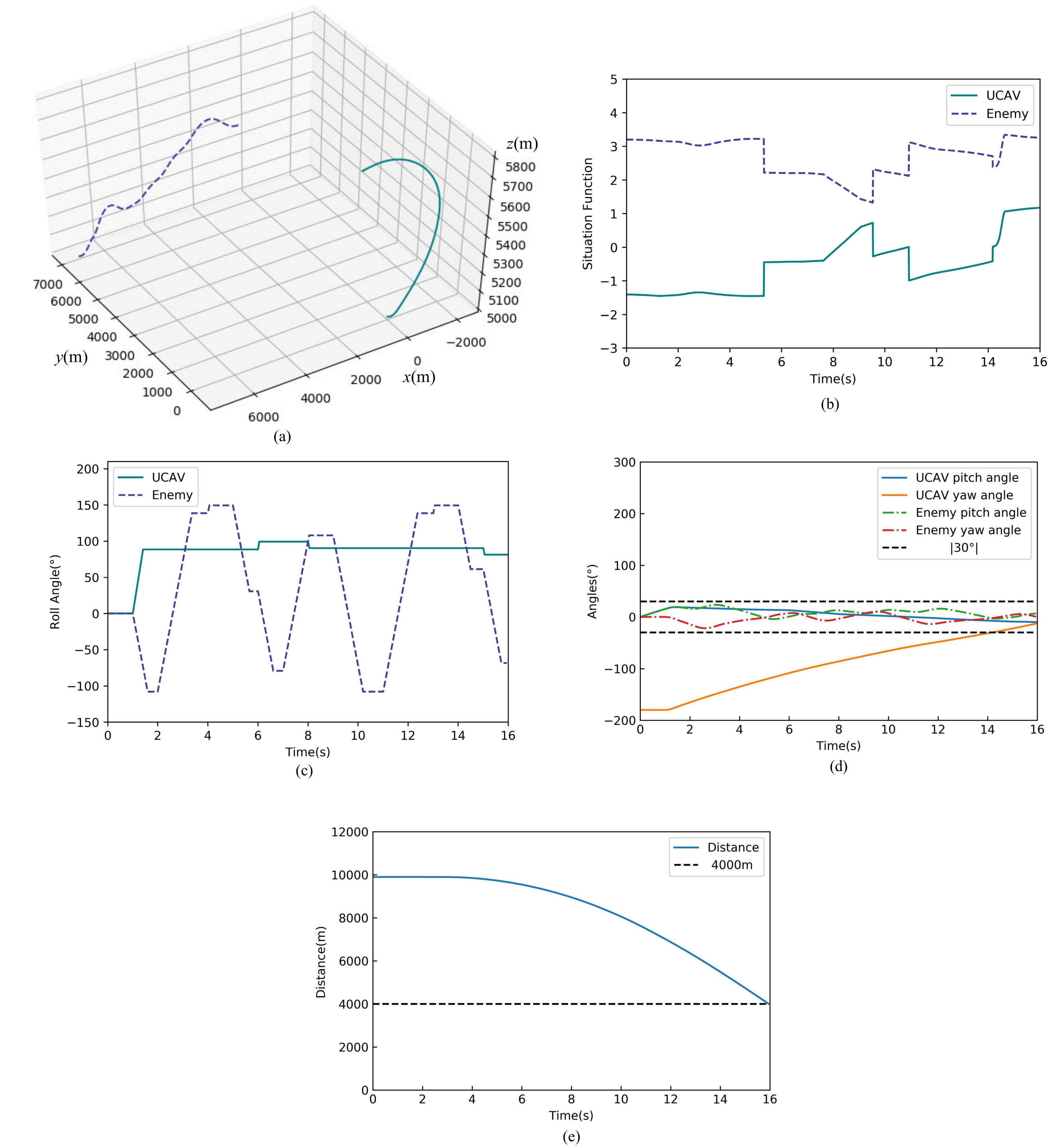

FIGURE 11. The results of an air combat confrontation under the fourth condition. (a) Air combat trajectories. (b) Simulation function value. (c) Change of the roll angle. (d) Projections of the included angles in the horizontal and vertical planes. (e) Distance between the two sides.

Even though the UCAV only won enemy once in the four air combat simulations, we can still conclude that our decision-making method is better, because the UCAV still met the conditions of missile launching when it was at a disadvantage in the three defeated simulations.

VII. CONCLUSION

In this study, we develop a novel decision-making method using a deep neural network and a novel maneuver library containing 72 maneuvers. A new decision objective function consisting of the pitch angle and the yaw angle is proposed.

The deep neural network aims at predicting the pitch angles and yaw angles of the 72 maneuvers if they were selected and used for simulating. This prediction system is trained with offline simulation data, and the test results show that the network performs well and accurately captures the tendency of the state development within a tolerant prediction error. Particularly, the neural network prediction system provides a reliable method to predict the aircraft movement after a long period without increasing the operational time.

We designed three disadvantageous conditions and one mutually beneficial condition. Our decision-making method can win the air combat in mutually beneficial circumstances and can meet missile launching conditions when in adverse circumstances. Therefore, the proposed method can solve the decision making problem effectively and competitively compared with the method in [12]. The decision-making time of the deep neural network method is reduced by 0.01 s compared with the method in [12]. However, the limitation in our method is that the neural network cannot be improved once the training has been accomplished. It is necessary to develop a novel neural network that can be improved automatically and a novel maneuver library containing more maneuvers.

REFERENCES

[1] S. Goudarzi, N. Kama, M. H. Anisi, S. Zeadally, and S. Mumtaz, "Data collection using unmanned aerial vehicles for Internet of Things platforms," *Comput. Electr. Eng.*, vol. 75, pp. 1–15, May 2019.

[2] F. Austin, G. Carbone, M. Falco, H. Hinz, and M. Lewis, "Game theory for automated maneuvering during air-to-air combat," *J. Guid., Control, Dyn.*, vol. 13, no. 6, pp. 1143–1149, Nov. 1990.

[3] H. Park, B.-Y. Lee, M.-J. Tahk, and D.-W. Yoo, "Differential game based air combat maneuver generation using scoring function matrix," *Int. J. Aeronaut. Space Sci.*, vol. 17, no. 2, pp. 204–213, Jun. 2016.

[4] J. Poropudas and K. Virtanen, "Game-theoretic validation and analysis of air combat simulation models," *IEEE Trans. Syst., Man, Cybern. A, Syst. Hum.*, vol. 40, no. 5, pp. 1057–1070, Sep. 2010.

[5] T.-Y. Sun, S.-J. Tsai, Y.-N. Lee, S.-M. Yang, and S.-H. Ting, "The study on intelligent advanced fighter air combat decision support system," in *Proc. IEEE Int. Conf. Inf. Reuse Integr.*, Sep. 2006, pp. 39–44.

[6] K. Virtanen, J. Karelahti, and T. Raivio, "Modeling air combat by a moving horizon influence diagram game," *J. Guid., Control, Dyn.*, vol. 29, no. 5, pp. 1080–1091, Sep. 2006.

[7] K. Virtanen, T. Raivio, and R. P. Hamalainen, "Modeling Pilot's sequential maneuvering decisions by a multistage influence diagram," *J. Guid., Control, Dyn.*, vol. 27, no. 4, pp. 665–677, Jul. 2004.

[8] H. Changqiang, D. Kangsheng, H. Hanqiao, T. Shangqin, and Z. Zhuoran, "Autonomous air combat maneuver decision using Bayesian inference and moving horizon optimization," *JSEE*, vol. 29, no. 1, pp. 86–97, Feb. 2018.

[9] J. Ren, X. Gao, J. Zheng, and Y. Zhang, "Mission decision-making for UCAV under dynamic environment," *J. Syst. Eng. Electron.*, vol. 32, no. 1, pp. 100–103, Jan. 2010.

[10] J. S. Mcgrew, J. P. How, and B. Williams, "Air combat strategy using approximate dynamic programming," *J. Guid. Control. Dyn.*, vol. 33, no. 5, pp. 1641–1654, Oct. 2010.

[11] J. S. Mcgrew, "Real-time maneuvering decisions for autonomous air combat," M.S. thesis, Dept. Electron. Eng., MIT, Massachusetts, CA, USA, 2008.

[12] Y. Wang, C. Huang, and C. Tang, "Research on unmanned combat aerial vehicle robust maneuvering decision under incomplete target information," *Adv. Mech. Eng.*, vol. 8, no. 10, Oct. 2016, Art. no. 168781401667438.

[13] N. Ernest, D. Carroll, and C. Schumacher, "Genetic fuzzy based artificial intelligence for unmanned combat aerial vehicle control in simulated air combat missions," *J. Defense Manage.*, vol. 6, no. 10, pp. 144–156, Apr. 2016.

[14] H. Duan, X. Wei, and Z. Dong, "Multiple UCAVs cooperative air combat simulation platform based on PSO, ACO, and game theory," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 28, no. 11, pp. 12–19, Nov. 2013.

[15] R. W. Holsapple, P. R. Chandler, and J. J. Baker, "Autonomous decision making with uncertainty for an urban intelligence, surveillance and reconnaissance (ISR) scenario," in *Proc. AIAA Guid., Navigat., Control Conf.*, Aug. 2008, pp. 1–14.

[16] Y. Bengio, "Learning deep architectures for AI," *Found. Trends Mach. Learn.*, vol. 2, pp. 1–127, Mar. 2009.

[17] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 25, Oct. 2012, pp. 1106–1114.

[18] G. E. Hinton, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[19] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[20] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. Van Den Driessche, T. Graepel, and D. Hassabis, "Mastering the game of Go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, Oct. 2017.

[21] C.-J. Hoel, K. Driggs-Campbell, K. Wolff, L. Laine, and M. J. Kochenderfer, "Combining planning and deep reinforcement learning in tactical decision making for autonomous driving," May 2019, *arXiv:1905.02680*. [Online]. Available: https://arxiv.org/abs/1905.02680

[22] P. Liu and Y. F. Ma, "A deep reinforcement learning based intelligent decision method for UCAV air combat," in *Proc. Int. Conf. Cloud Comput. Intell. Syst.*, Mar. 2017, pp. 274–286.

[23] X. B. Zhang, G. Q. Liu, and C. J. Yang, "Research on air confrontation maneuver decision-making method based on reinforcement learning," *Electronics*, vol. 7, no. 279, pp. 1–19 , Apr. 2018.

[24] G. T. Lee and C. O. Kim, "Amplifying the imitation effect for reinforcement learning of UCAV's mission execution," Jan. 2019, *arXiv:1901.05856*. [Online]. Available: https://arxiv.org/abs/1901.05856

[25] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, Beijing, China: Posts Telecom, 2017, pp. 126–149.

HONGPENG ZHANG was born in 1996. He received the bachelor's degree from Air Force Engineering University, Xi'an, China, in 2017, where he is currently pursuing the master's degree in unmanned aircraft combat system and technology. His current research interests are autonomous air combat for unmanned combat aerial vehicle and machine learning, including deep learning and its application on unmanned aerial vehicle combat systems.

CHANGQIANG HUANG was born in Jiangsu, in 1961. He received the Ph.D. degree in navigation, guidance and control from Northwestern Polytechnical University, Xi'an, China, in 2006. He is currently a Professor and a Doctoral Tutor with Air Force Engineering University. His current research interests are autonomous air combat for unmanned combat aerial vehicle and artificial intelligence, including knowledge extraction, big data application, and air combat simulation systems.