Development of Smart Emotion Recognition System for Companion Robot

Kuo-Ho Su, *IEEE Member*, Chung-Hsien Kuo, *IEEE Member* and Ya-Tang Feng

***Abstract*—The first stage of this study is to establish a convolutional neural network based image recognition model to identify the face image of "Happiness", "Anger", and "Sadness". To further improve the model's recognition accuracy, the second model is established via adding the collected physiological data such as the heartbeat and body temperature, to form a physiological data-assisted psychological recognition system. The Googlenet is selected as the first model and the long short-term memory (LSTM) and backpropagation neural network (BPNN) are chosen as the second model in this study. By importing the captured image and the detected physiological data, one of following six emotions—Happiness, Anger, Fear, Sadness, Surprise, and Disgust can be assessed. Some simulation results are provided in the first stage. The second stage of this study is to make the system lightweight furtherly, both of the models' formats are converted and embedded into Raspberry Pi smart control board. Finally, the control board, camera and physiological sensors are equipped into a companion robot to implement the smart object.**

Index Terms*—Convolutional neural network, Long short-term memory, Emotion recognition system, Companion robot, Microcontroller

I. INTRODUCTION

Emotion recognition is a category of artificial intelligence (AI), which specializes in detecting and identifying human emotions. The facial action coding system (FACS) published by psychologists P. Ekman and W. V. Friesen was the most common way to objectively classify facial movements [1, 2]. The current deep learning technology and the amount of accumulated data have changed the previous disadvantages, greatly increasing the direction and speed of emotion recognition development [3]. On the other hand, the methods of physiological response emotion identification including physiological response identification such as heartbeat, respiration, body temperature, sweat volume, blood oxygen content, and changes in exhaled carbon oxide can promote the system accuracy [4]. The proposed recognition system can not only be used in companion robot but also can be extended to other applications.

The process of emotion recognition based on model training and classification can be divided into four steps [5]. Step 1: Train the model via the collected database; Step 2: Detect the face range, some facial detection technologies can be utilized, e.g. a cuboid shape is adopted to frame the face and the facial motion key points (there are usually 34 points, and the changes of these points can be further interpret the facial action). Step 3: Extract the facial features and the movements. Step 4: Emotion classification and intensity interpretation, e.g., using a scale of 1-100 to rate the degree to which the outer eyebrows raised. The emotion intensity interpretation is based on different facial action compositions corresponding to the different emotional expressions [6]. The emotional intensity interpretation usually focuses on six core emotions of anger, disgust, fear, happiness, sadness and surprise. The method was proposed by P. Ekman [1]. Some researchers also added the additional emotions, such as neutral (calm), and contempt, etc. The intensity in this study, each emotion is rated on a scale of 0-1 and then output the emotion classification. AI can also be used to identify the gender or whether wear the glasses to adjust its emotional interpretation of facial movements.

Face recognition is based on image acquisition and face image capturing for analysis and comparison. It is often used as a computing technology for identity authentication and facial expression recognition. This research compares following three convolutional neural networks (CNNs) models, including Googlenet, Alexnet, and VGG-16 [7, 8]. After comparisons, the Googlenet is selected as the first model. On the other hand, the physiological signals or responses can also be used to recognize the emotion, such as heartbeat, respiration, body temperature, sweat volume, blood oxygen content, and changes in exhaled carbon oxides. In this study, the combination of psychological and physiological data model is proposed and initially tested with recurrent neural network (RNN). Since the time axis that RNN can store is limited, the detection results will be distorted if the time limit is exceeded [9]. Therefore, the long short-term memory (LSTM) and backpropagation neural network (BPNN) are used in this study finally. Because LSTM passes the input gate, forget gate and then outputs the result. The gate introduces the sigmoid function and combines the tanh function to add a summation operation to reduce the possibility of gradient disappearance and gradient explosion [9]. The developed model can increase the function of prediction; however, LSTM cannot be used on Raspberry Pi due to the file format transformation problem. Therefore, the BPNN is finally selected as the physiological data-assisted psychological model in this study [10].

Companion robot or interactive machine can be trained to possess human or other biological behavior and thinking ability. After embedding some deep learning technologies into the controller, the interactive function of robot will become more powerful with the environments, such as companion robot and automatic vehicle. Some electronic devices, such as radar, camera, LIDAR module, vehicle-to-everything communicator, microcontroller and visual processing unit (VPU) are utilize to sense and process the signals. Combining the deep learning

This research supported by the Ministry of Science and Technology of Taiwan, R.O.C. through its grant MOST 110-2221-E-034-015-.

Kuo-Ho Su is with the Graduate Institute of Digital Mechatronic Technology, Chinese Culture University, Taipei, Taiwan (e-mail: sgh@faculty.pccu.edu.tw)

Chung-Hsien Kuo is with the Department of Mechanical Engineering, National Taiwan University, Taipei, Taiwan (e-mail: chunghsien@ntu.edu.tw)

Ya-Tang Feng is with the Graduate Institute of Digital Mechatronic Technology, Chinese Culture University, Taipei, Taiwan(e-mail: B0101384@ulive.pccu.edu.tw)

architecture and above semiconductor devices to form the control board or platform are the foundation of intelligent robot development. Companion robot can not only replace human's role, but also improve the efficiency of production capacity in industry. This research aims to realize emotional recognition in an intelligent way, to improve the skill of recognition, and to use the Raspberry Pi as the control board. The emotion recognition ability of pure convolutional neural network is powerful but slightly insufficient. In this study, the proposed data-assisted framework is more powerful by adding some physiological sensors. The behavior of the companion robot is manipulated by the established models and Raspberry Pi microcontroller directly. Because the lightweight and practicality of the Raspberry Pi, the proposed architecture can greatly reduce the volume of the control box and unnecessary waste of connections.

II. Recognition System Design

In this study, two models are used for emotional recognition, the psychological data model for facial recognition and the physiological+psychological data recognition model, respectively. After testing and evaluation by simulation and experiment, Googlenet and BPNN are finally selected as the recognition models because of their efficiency and conversion. In the first stage, both of these two models are trained by Matlab 2021a and Colab on the computer side, respectively.

2.1 Model Establishment of Facial Recognition

The first stage of this study is to establish a facial recognition model, Googlenet, Alexnet, and VGG-16 are selected to test and compare their computation efficiency and accuracy. Among these models, Googlenet is proposed by Google in 2015. It possesses 5 identification characteristics as follows. (1) Face verification: Identify whether it is the same person; (2) Face identification: Verify their identities or names; (3) Face cluster: Categorize persons with similar conditions; (4) Face search: Search for persons with similar conditions; (5) Face tracking: Track specific faces. On the other hand, Alexnet achieved the lowest Top-5 error rate with 15.3% in the Challenge of ImageNet Large-Scale Visual Recognition on 30th September, 2012. The main conclusion of the original paper is that the depth of the model is crucial to improve performance, especially the GPU used in Alexnet, though the computational cost is much higher than other types, it is worth. On other hand, VGG is the acronym Visual Geometry Group of University of Oxford. The main feature of this recognition is used more hidden layers and a large number of images for training to increase the accuracy to 90%. There are 16 layers of VGG-16. (13 convolutional layers and 3 fully connected layers) [5, 7].

Table 1. Comparison of different facial recognition systems

	Alexnet	Googlenet	VGG-16
Forward time(ms)	18	39	120
Energy(kJ)	12	35	112
Momery(MB)	420	590	1550
Accuracy(%)	58	70	72

According to the evaluation benchmark, some simulated comparisons are carried out in this study. The results of the Forward time(ms), Energy(kJ), Memory(MB), and Accuracy(%) of above three models are tabulated in Table 1. Comparing the simulated results, Googlenet is finally selected as the first model of emotion recognition system in this study.

2.2 Model Establishment of Physiological Data-assisted Mental Data

The model of physiological data-assisted mental data is developed with RNN firstly. However, there exists exponential explosion of weights, vanishing gradients and recursive problem, make it difficult to capture the correlations of long-term temporal and the detection results would be distorted if the time limit is exceeded. The diagram of RNN architecture is shown in Fig. 1.

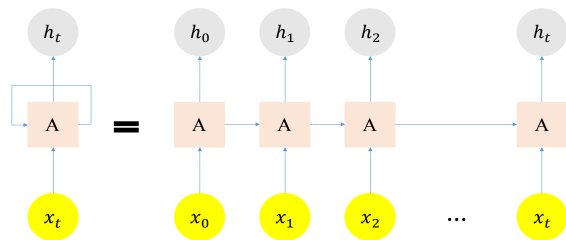

Fig. 1. Diagram of RNN architecture

Therefore, the model is replaced by the LSTM [11, 12, 13]. Because LSTM combines the tanh function into the sigmoid function through the input gate, forget gate, and output gate, adding the sum operation, and reducing the probability of the vanishing and exploding gradient. The output gate and shadow state can be described as follows.

$$o_t = \sigma(W_o\,[h_{t-1}, X_t] + b_o) \quad (1)$$
$$h_t = o_t * \tanh(C_t) \quad (2)$$

where C_t is the memory cell. The developed model can add the function of predictive; however the LSTM cannot be used in the Raspberry Pi due to the file conversion problem.

Finally the BPNN is selected as the model of physiological data-assisted mental data in this study. Its structure is described as follows. The BPNN belongs to supervised learning; it learns the relationship between input and output by the collected training data [14, 15, 16]. It contains input layer, hidden layers and output layer, where the number of hidden layers and neurons will affect the connection weight and activation function. Some learning equations are shown in Eqs. (3-7) and the proposed architecture is shown in Fig. 2.

activation = sum (weight(i) * input(i)) + bias (3)
output = 1 / (1 + e^(-activation)) (4)
derivative = output * (1.0 - output) (5)
error = (output - expected) *f(derivative(output)) (6)
new weight = weight - learning_rate * error * input (7)

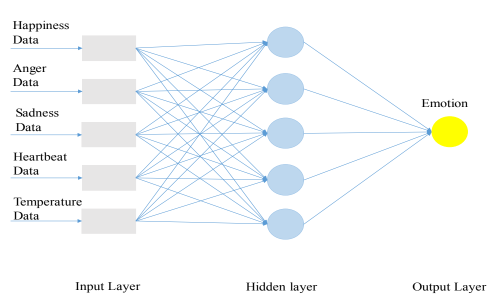

Fig. 2. Diagram of proposed BPNN architecture

III. Implementation With Smart Controller

The goal of this study is to implement a smart emotion recognition system. So the Raspberry Pi 4 and VPU are chosen as the main controller as they can greatly reduce the volume of the control box and unnecessary waste of connections with lightweight and intelligence. The main functions of Raspberry Pi 4 are to control camera and sensors, operate the models of Googlenet and BPNN, and drive the motion patterns of the robots. The camera is used to capture the face image, the AMG8833 infrared thermal imager is to detect the body temperature, and the MAX30100 to detect the pulse. The image accelerating operations are performed by the VPU (the Intel Neural Compute Stick 2 is adopted in this study). Finally integrating all the sensors, camera and control board into the companion robot to implement desired object [17, 18]. The program flow is shown in Fig. 3. The detail contents are described in the following subsections.

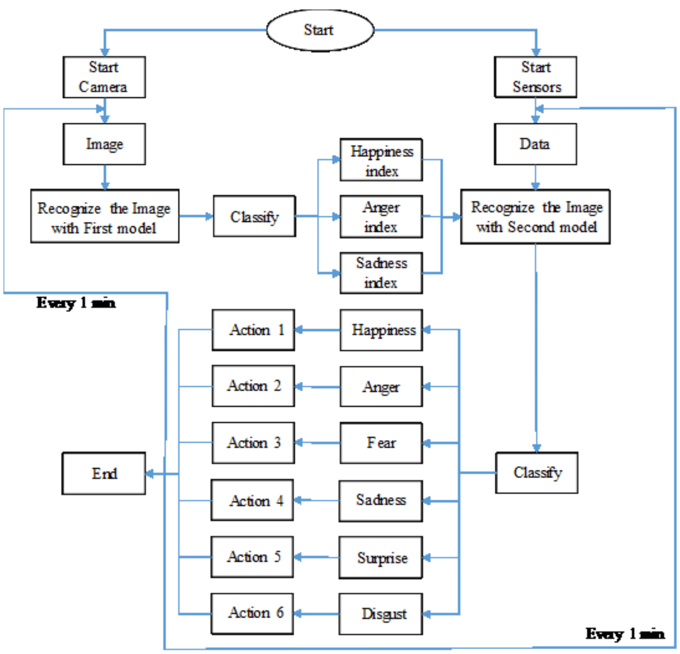

Fig. 3. Program flowchart of Raspberry Pi

3.1 Raspberry Pi Microcontroller

The microcontroller used in this study is the fourth generation of Raspberry Pi, a tiny single-board computer developed by the Raspberry Pi Foundation in U.K., compared to previous generation specifications, the latest fourth generation has been improved processors, and the graphics processors, supported 4K, USB3.2, the speed of Ethernet network 1Gbite/s, bluetooth5.0, Mini HDMI 2.0 terminal etc., therefore, the fourth generation is chosen as the main controller shown in Fig. 4.

Fig. 4. Raspberry Pi 4B 8G [17]

3.2 The Sensors of Physiological Data

In this research, the physiological data used to identify emotions are the heartbeat and body temperature. These two physiological data can provide the data for the emotion recognition and can strengthen the identification accuracy. The type of infrared thermal imager used in this study is AMG8833, which possesses a 8x8 infrared thermal sensor array. When connected to a microcontroller, it will return a set of 64 individual IR temperature readings via I2C. The range of the sensing temperature measurement is 0˚C to 80˚C (32˚F to 176˚ F) and the precision is ±2.5˚C(4.5˚F). The maximum frame rate is 10Hz, it is great for creating the own body detector or mini thermal camera.

The type of the pulse oximeter used in this research is MAX30100, which is a non-invasive blood oxygen saturation and pulse contact sensor. It combines two light-emitting diodes, and a photodetector, optimizes optics and the processing of low noise analog signal to detect pulse oximetry and heart rate signals. The pulse oximetry (SpO2) and the pulse (equivalent to a heartbeat) can be estimated simply by pressing a finger against the sensor.

3.3 Camera and Accelerated Computing Processing Unit

The used camera module of Raspberry Pi is the second generation camera module launched by Raspberry Pi officially, it owns 5 million pixels and 160-degree wide-angle camera with IMX219PQ image sensor, and support 1080p 30fps video shooting. The used accelerated computing processing unit is the Intel Neural Compute Stick 2 (Intel NCS2), with plug-and-play features, can perform functions on WIN10, Ubuntu or macOS, and owns the prototype for low-cost edge devices such as Raspberry Pi.

3.4 Action Modes of Companion Robot

After the emotion is recognized by the Raspberry Pi and VPU through the established model, the companion robot will make an appropriate action corresponding to the owner's emotion. The companion robot possesses four sets of the links, the action mode of one set of links is determined by the swing angle of two servo motors, the motion equation is shown as follows.

$$\cos C = (a * a + b * b - c * c)/(2 * a * b) \tag{8}$$

where C is one angle of the one link; a, b, and c are the sides corresponding to angles A, B, and C. The schematic diagram of connecting rod triangle is shown in Fig. 5.

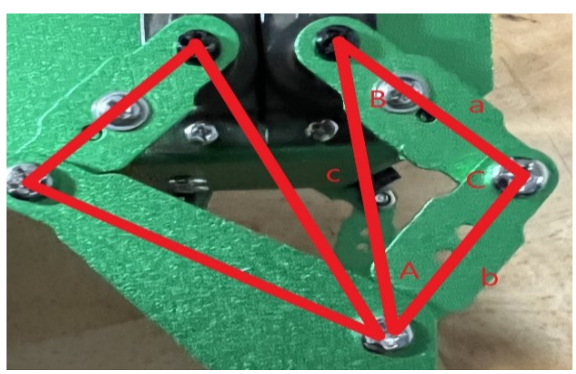

Fig. 5. Schematic diagram of connecting rod triangle

3.5 Description of Whole Electronic Control System

In this study, the robot's action is controlled by the Raspberry Pi microcontroller, Arduino Uno, and 8 servo motors. The power of camera, body temperature sensor, heartbeat sensor and VPU are supplied by the power module. The power control is responsible for the energy source of the whole system to realize the concept of portable and using anytime. The system is equipped with 2 SJ 18650 rechargeable batteries, one battery specification with 3.7V and 5600mAh, which can be used by the system for long time. The module is "Power Management HAT", the 40 PIN GPIO is designed consistent with the Raspberry Pi controller; DC 7-28V jack for external battery power which is built-in PCF8523 RTC clock chip to provide precise time; and CP2102 serial port chip can be used for serial port communication or downloaded programs to ATmega328P, it is built-in custom buttons, the microcontroller power-on, safe shutdown, or custom functions are provided, power supply anti-reverse connection, anti-current inversion and other circuits protection for working more stable and safer. With the voltage and current detection circuit, the working voltage and current of the Raspberry Pi can be monitored in time. The schematic of electronic control system is shown in Fig. 6.

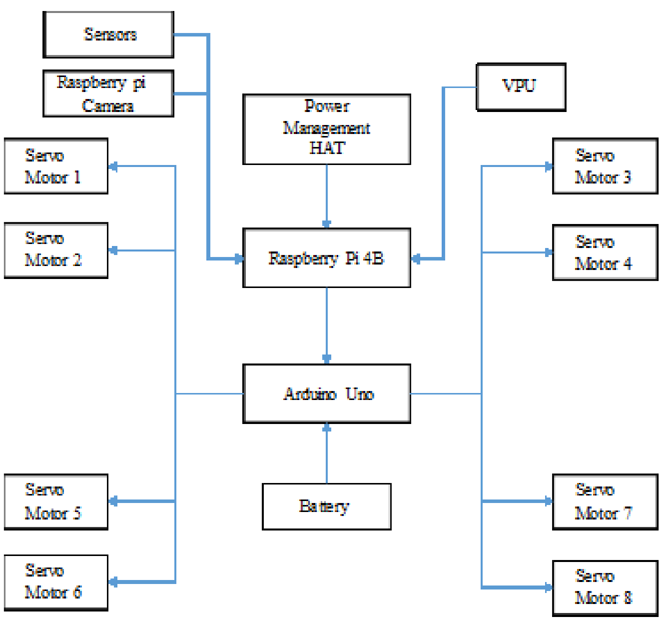

Fig. 6. Schematic of whole electronic control system

IV. MODELLING RESULTS

Some modelling results of the recognition are provided in this section. The first and the second model are trained in Matlab 2021a and Colab, respectively.

4.1 Recognition Results of Model 1

The model 1 is the facial recognition. After comparison, the Googlenet is selected as the model 1 of emotion recognition system in this study. The Googlenet model is established in Matlab 2021a on the PC side. Because it will be used on the non-Matlab environment, it needs the assistance of OpenVINO software. Firstly, the model needs to be converted to ONNX format. After the conversion, the OpenVINO optimizer is further used to convert the ONNX format into xml and bin format via mo.py provided by Intel. This study uses 600 face photos of happiness, anger and sadness to train the Googlenet model. The training results are shown in Fig. 7. Before starting to identify the photos, run the setupvars.bat of the bin folder by CMD to initialize OpenVINO, then put the converted xml and bin, the photos, the model of Googlenet and Python into the same folder and execute the commands. The results are shown in Fig. 8.

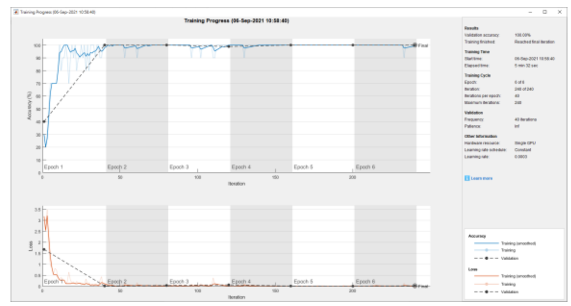

Fig. 7. Training result of model 1

Fig. 8. Recognition result of Model 1 after OpenVINO

4.2 Recognition Results of LSTM model

According to "Bodily maps of emotions", the changes of body temperature and heartbeat rate for each mood [19], the average data of the physiological states under various emotions are tested in this study, and the result is shown in Table 2.

Table 2. Change of heartbeat, temperature under each mood

	Heartbeat(s/m)	Temperature (□)
Happiness	+1~+5	+0.03~+0.04
Surprise	+2~+6	-0.01~-0.02
Sadness	+7~+12	+0.01~+0.02
Fear	+8~+12	-0.01~-0.02
Angry	+8~+11	+0.15~+0.16
Disgust	-1~-5	-0.03~-0.04

The training dataset1 of model 1 is cited from previous work: Smart face identification system and its application on robot [16]. This model can generate the facial results of happiness, anger, or sadness. Furthermore, according to the relationship between emotion and physiological state (Table 2), the training dataset2 is formed by test and experiment. The total amount of data is 3000 pieces in dataset2 and it is used to train model 2. The contents of dataset2 contain five inputs and six outputs. The dataset2 can be classified into 6 types and each type contains

500 pieces. The training result under Matlab 2021a is shown in Fig. 9. Then the XTest (test data of mental and physical state) and YTest (test data of six emotions) data are imported, the recognition results are shown in Figs. 10-11. In order to further verify the stability of the integrated data, another development platform Colab is used in this study. The recognition results are shown in Figs. 12-14.

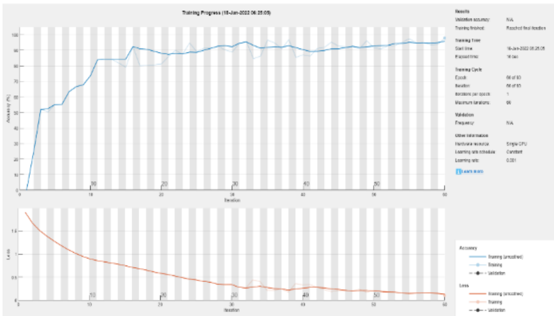

Fig. 9. Training result of LSTM model under Matlab 2021a

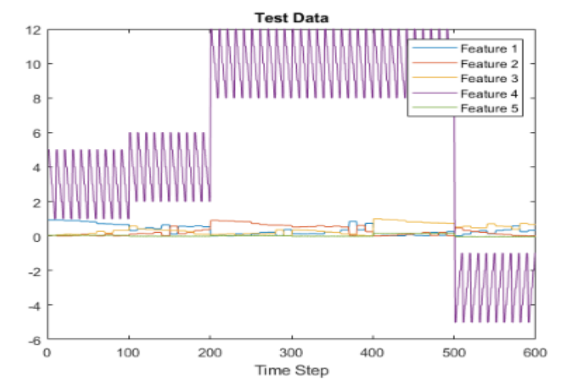

Fig. 10. Test data distribution under Matlab 2021a

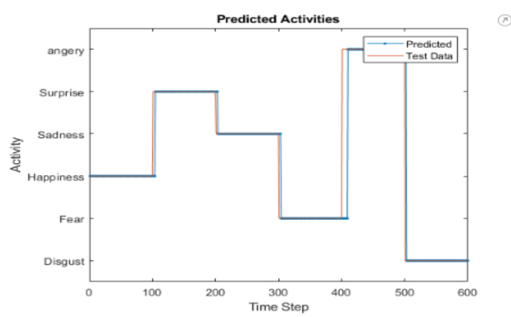

Fig. 11. Recognition result of LSTM model under Matlab 2021a

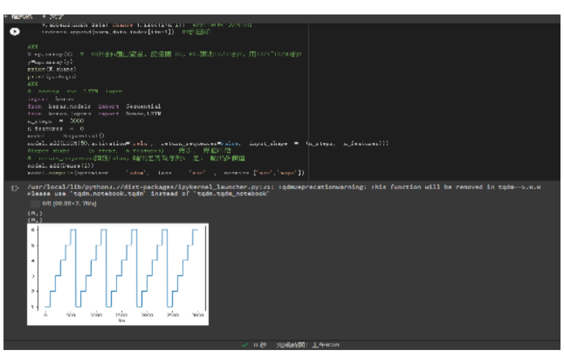

Fig.12. Test data distribution under Colab

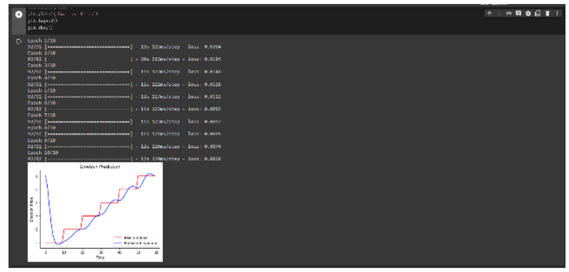

Fig.13. Recognition result of LSTM model under Colab Epoch=10

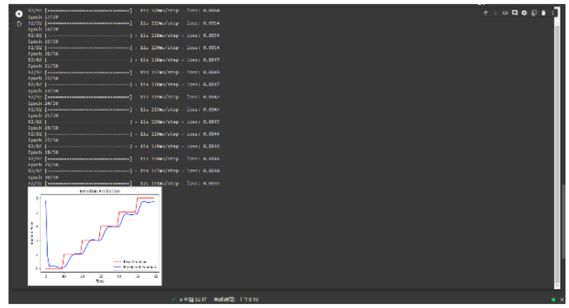

Fig.14. Recognition result of LSTM model under Colab Epoch=100

4.3 Recognition Results of BPNN model

In this study, the proposed BPNN contains 1 input layer (5 neurons) and 1 output layer (6 neurons) and 1 hidden layer (5 neurons). The training epoch is set as 500. Then 3000 pieces of data are converted into CSV as training data. After training (about 4 minutes and 34 seconds), the test data is imported to test. The recognition results are shown in Fig. 15. From the simulation result, the recognition accuracy could be 100%.

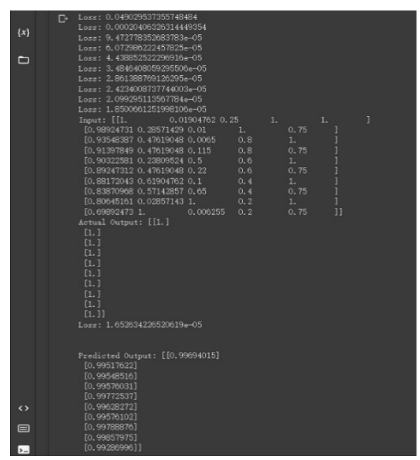

Fig.15. Recognition result of BPNN model under Colab Epoch=500

V. Implementation Results

After integrating the control board, VPU, camera, and sensors into the companion robot, the companion robot prototype is implemented as shown in Fig. 16. The .xml and .bin file (converted from ONNX file via model optimizer of OpenVINO) can be run on the Raspberry Pi. However, the network type of LSTM is SeriesNetwok, the current conversion command does not support SeriesNetwok. So LSTM is replaced by BPNN to construct model 2. Raspberry Pi 4 are used to control camera and sensors, operate the models of Googlenet and BPNN, and drive the motion patterns of the robots. The relationship of the robot link is shown in Fig. 17 and the rotation angles of the servo motors are listed in Table 3.

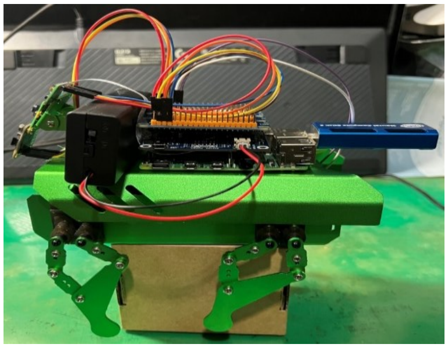

Fig. 16. Implemented prototype

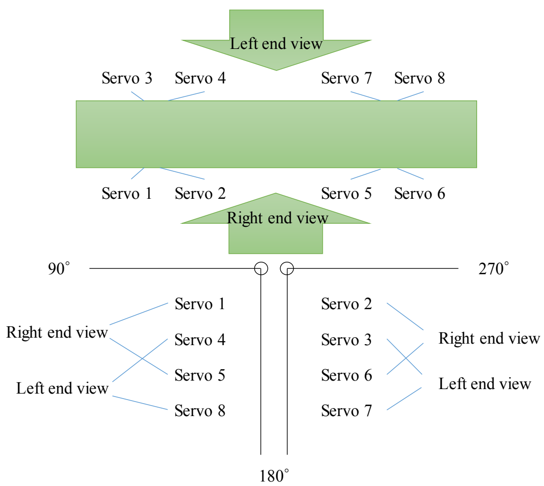

Fig. 17. Relationship diagram of the robot link

Table 3. Rotation angles of servo motors

	Servo1 /servo2	Servo3 /servo4	Servo5 /servo6	Servo7 /servo8
Stand	135° /225°	225° /135°	135° /225°	225° /135°
Happiness	180° /225°	180° /135°	135° /180°	225° /180°
Surprise	90° /240°	180° /120°	180° /210°	90° /210°
Sadness	120° /240°	240° /120°	120° /240°	240° /120°
Fear	90° /180°	270° /180°	180° /270°	180° /90°
Anger	90° /225°	180° / 135°	180° /225°	180° /225°
Disgust	135° /225°	270° /180°	180° /225°	180° /135°

After converting the Googlenet model created by Matlab 2021a and storing the BPNN model created by Colab as *.py for Raspberry Pi, the execution results of the two models on the Raspberry Pi are as follows. The same 30 photos are used to test the accuracy of Raspberry Pi smart board. The results are summarized in Table 4. From Table 4, it is evident that there exist two recognition errors in 3rd and 14th row, respectively. So the rate of the recognition accuracy is about 93.3%.

Table 4 Recognition results in smart board

	Happiness	Anger	Sadness	Heartbeat(s/m)	Temperature(℃)	Result
1	0.67160	0.32511	0.00328	1.00000	0.04000	Happiness
2	0.43102	0.41664	0.15234	5.00000	-0.01000	Surprise
3	0.13527	0.12161	0.74313	3.00000	0.04000	Happiness
4	0.53129	0.09931	0.36941	2.00000	-0.02000	Surprise
5	0.71204	0.17580	0.11215	2.00000	0.04000	Happiness
6	0.89215	0.10652	0.00133	1.00000	0.04000	Happiness
7	0.15235	0.32451	0.52314	3.00000	-0.01000	Surprise
8	0.72514	0.27361	0.00125	2.00000	0.04000	Happiness
9	0.42534	0.35749	0.21716	1.00000	0.03000	Happiness
10	0.54101	0.24468	0.21431	4.00000	-0.01000	Surprise
11	0.46121	0.42664	0.11215	7.00000	0.02000	Sadness
12	0.90125	0.09643	0.00232	9.00000	0.01000	Sadness
13	0.14130	0.32051	0.53819	9.00000	-0.01000	Fear
14	0.00022	0.88455	0.11523	8.00000	0.01000	Sadness
15	0.37814	0.00866	0.61320	10.00000	-0.02000	Fear
17	0.21413	0.11285	0.67302	10.00000	0.01000	Sadness
18	0.58316	0.30430	0.11254	11.00000	-0.02000	Fear
19	0.13548	0.34113	0.52339	8.00000	0.01000	Sadness
20	0.02543	0.10034	0.87423	10.00000	0.02000	Sadness
21	0.31512	0.57233	0.11255	12.00000	0.16000	Anger
22	0.21040	0.76411	0.02549	8.00000	0.15000	Anger
23	0.58214	0.39432	0.02355	-2.00000	-0.03000	Disgust
24	0.01254	0.77333	0.21413	11.00000	0.15000	Anger
25	0.25310	0.73436	0.01254	-5.00000	-0.03000	Disgust
26	0.08465	0.81682	0.09852	9.00000	0.16000	Anger
27	0.25879	0.73988	0.01325	8.00000	0.15000	Anger
28	0.35841	0.55617	0.08542	-1.00000	-0.04000	Disgust
29	0.03152	0.81616	0.15232	11.00000	0.15000	Anger
30	0.11525	0.56886	0.31589	10.00000	0.16000	Anger

V. CONCLUSION

The goal of this study is to develop and implement a smart emotion recognition system and extend its application to companion robot. Considering the smart and lightweight requirement, both of the neural models are embedded into a small Raspberry Pi control board. Some difficulties are encountered in the developing process, e.g. the file conversion, accuracy, and runtime. In the future, some new modules (e.g. Nvidia Jetson Xavier NX) and novel skeleton-driven posture estimator or other deep learning-based point cloud registration methods will be utilized to promote its accuracy and speed.

ACKNOWLEDGMENT

This work is financially supported by the Ministry of Science and Technology, Taiwan, R.O.C., under Grant MOST 111-2221-E-034-006-.

REFERENCES

[1] The Facial Action Coding System (FACS), https://www.paulekman.com/facial-action-coding-system/, Accessed, 21 March 2022.
[2] J. M. Garcia-Garcia, V. M. R. Penichet, M. D. Lozano, "Emotion detection: a technology review," in Proceedings of the 18th Int. Conf. on Human Computer Interaction, Sep. 2017, no. 8, pp. 1-8.
[3] S. Smys, R. Palanisamy, Á. Rocha, G. N. Beligiannis, "Computer networks and inventive communication technologies," in Proceedings of 3rd ICCNCT 2020, Coimbatore, India, 23–24 July 2020, pp. 185-196.
[4] Emotion recognition AI technology introduction, https://ai.iias.sinica.edu.tw/emotion-ai-tech-intro, Accessed, 06 May 2020.
[5] K. H. Su, Y. T. Feng, "Development of emotion recognition system and its application in interactive robot," in 2022 Artificial Intelligence Technology and Application (AITA2022), Taichung, Taiwan, 20 May 2022, pp. 765-770.
[6] C. C. Yang, "GSR based interactive robot with human emotional recognition model," Soochow University, School of big data management, Taipei, Taiwan, 2019.
[7] N. Thida, N. O. Aung, "University classroom attendance system using FaceNet and support vector machine," in 2019 Int. Conf. on Advanced Information Technologies (ICAIT2019), Yangon, Myanmar, Nov. 2019, pp. 171-176.
[8] R. Jozefowicz, W. Zaremba, I. Sutskever, "An empirical exploration of recurrent network architectures," New York University, New York, America, 2015.
[9] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," Physica D, Mar. 2020, vol. 404, article 132306.
[10] F. Arifin, "Study of optimal multi-shot molding of blu-ray objective lens with Taguchi method, BPNN-GA and BPNN-PSO," National Kaohsiung University of science and Technology, Department of Mechanical Engineering, Kaohsiung, Taiwan, 2019.
[11] D. Li, Y. Li, H. Hu, T. Z, C. J, "Heterogeneous platform-aware workload feature recognition for edge intelligence," Physical Communication, June 2022, vol. 52, article 101620.
[12] L. Nummenmaa, E. Glerean, Riitta Hari, J. K. Hietanen, "Bodily maps of emotions," PNAS, Jan. 2014, vol. 111, no. 2, pp. 646-651.
[13] S. Alhagry, A. A. Fahmy, R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," Int. Journal of Advanced Computer Science and Applica-tions(IJACSA), 2017, vol. 8, no. 10, doi: 10.14569/IJACSA.2017.081046.
[14] M. L. Qian, "What exactly is emotion? Unraveling the Mystery of Emotional Mechanisms and Obstacles from Modern Science," Yuan Liu, 2020.
[15] J. M. Chen, K. W. Liang, J. C. Wang, P. C. Chang, "Speech emotion recognition technology combining psychological characteristics and emotion label training," in TANET2019, Dec. 01, 2019, pp. 272 – 276.
[16] R. C. Wu, "Smart Face Identification System and Its Application on Robot," Chinese Culture University College of Engineering Department of Mechanical Engineering, Taipei, Tai-wan, 2021.
[17] Raspberry Pi: Infrared Camera, https://laplacetw.github.io/, Accessed, 19 April 2020.
[18] Raspberry Pi power control module (built-in Arduino, RTC), https://reurl.cc/p1eyOb, Accessed, Dec. 2020.
[19] Y. T. Feng, "Development of Emotion Recognition System and Its Application for Interactive Robots," Chinese Culture University College of Engineering Department of Mechanical Engineering, Taipei, Taiwan, 2022.

Kuo-Ho Su was born in Yunlin, Taiwan, R.O.C., in 1959. He received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from Tatung University, Taipei, Taiwan, R.O.C., in 1983, 1985, and 2005, respectively. Currently, he is a Professor with the Graduate Institute of Digital Mechatronic Technology, Chinese Culture University, Taipei, Taiwan, R.O.C. His research interests include the design and application of intelligent controller (fuzzy, neural network, genetic algorithm), robot system and embedded MCU system.

Chung-Hsien Kuo received the M.S. and Ph.D. degrees in Mechanical Engineering from National Taiwan University in 1995 and 1999, respectively. Currently, he is the Chairman and Professor of Department of Electrical Engineering, National Taiwan University of Science and Technology. He also is the Director of Taiwan TECH Industry 4.0 Center. His research areas include autonomous mobile robots, autonomous driving techniques, artificial intelligence, deep image learning, interactive cognitive learning, intelligent systems for Industry 4.0.

Ya-Tang Feng was born in Taipei, Taiwan, R.O.C., in 1998. He received the B.S. degree in Mechanical Engineering from Chinese Culture University in 2020. Currently, he is a master student with the Graduate Institute of Digital Mechatronic Technology, Chinese Culture University, Taipei, Taiwan, R.O.C.