Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105
Volume **9**, Number **4**, November **2024**

Application of Improved MobileNet2 SSD Algorithm in Embedded Systems

Benzheng Fan*, Zhoushu Lv

School of Mechanical and Electrical Engineering
Zhumadian Vocational and Technical College, Zhumadian 463000, China
ei202202@163.com, 15093577936@163.com

*Corresponding author: Benzheng Fan
Received September 20, 2023, revised December 22, 2023, accepted April 10, 2024.

Abstract. *Since the 21st century, with the continuous advancement of national industrialization, embedded systems have been widely used in various industries. Due to their highly customisable and adaptable characteristics, embedded systems can fulfil the diverse demands of various industries for intelligence, automation and portability. They offer significant practicality for detecting patient attention. However, embedded systems based on traditional algorithms have many limitations. To solve the problems of low accuracy and reliability in traditional embedded systems, this study builds on the MobileNetV2 Single Shot Multibox Detector algorithm and uses EfficientNet to improve the model's feature representation and detection performance. The multi-scale training and prediction strategies are also introduced and finally applied to the embedded systems to design an attention detection model. The test results showed that after 240 iterations, the training error of the improved MobileNetV2 Single Shot Multibox Detector model tended to converge, which was 70.83% faster than the standard MobileNetV2 Single Shot Multibox Detector. Finally, it converged to 0.05 and the error decreased by 0.30. The embedded attention detection system in this study exhibits greater recognition accuracy and consumes less time, indicating potential application in the field of attention detection.*
Keywords: PSO MobileNet2 SSD; Embedded system; EfficientNet; Attention detection; Howland Current Source

1. **Introduction.** The identification of attention-related illnesses in children is increasingly gaining attention from the public. These diseases, such as attention deficit and hyperactivity disorder (ADHD), have a negative impact on children's learning and development. The rapid development of the Internet of Things and mobile computing has brought new opportunities for detecting children's attention, among which embedded systems are highly favored in this field [1]. MobileNetV2 Single Shot Multibox Detector (MobileNet2 SSD) is a commonly used algorithm in embedded systems. Although it is a lightweight object detection algorithm, there are still many limitations at this stage.

Firstly, there are certain limitations to the detection performance of the MobileNet2 SSD algorithm, especially in the detection of small-sized targets, which often face difficulties [2]. Secondly, the resources and computing power of embedded systems are limited, and it is necessary to improve the speed and efficiency of MobileNet2 SSD while maintaining accuracy [3]. To address the detection challenges and low efficiency of the conventional MobileNet2 SSD algorithm, this study presents an enhanced and innovative attention detection model. This study endeavours to enhance the precision of attention detection to facilitate a greater number of individuals with attention deficit disorder in

seeking appropriate medical care in a timely manner. The innovation of this study lies in: (1) To enhance the precision and sensitivity in detecting children's attention, the current study initially enhances MobileNet2 SSD's network structure. (2) EfficientNet has been introduced to improve the feature representation and detection performance of the model. (3) By introducing multi-scale training and prediction strategies, and selecting the loss function Focal Loss that is more suitable for object detection tasks, the improved MobileNet2 SSD algorithm (IMobileNet2 SSD) is obtained. Finally, the algorithm is applied to the embedded systems and an attention detection model is designed.

This paper mainly consists of four parts. The second part is a review of the current research status of domestic and foreign experts and scholars on children's attention detection. The third part establishes an improved embedded attention detection model for MobileNet2 SSD. The fourth part conducts comparative experiments and efficiency verification on the optimization effect of the model.

2. **Related Work.** In the past few decades, embedded systems have been widely used, and embedded systems with stronger performance have gradually received attention from multiple enterprises and researchers. Li proposed a fuzzy logic C-means (FLC) clustering algorithm to address the difficulty of compensating for fiber non-linearity using digital signal processing (DSP). This technology could greatly reduce signal damages [4]. Yang et al. designed a catalyst preparation prediction model using machine learning methods to address the issue of long experimental steps and cycles typically required in traditional catalyst development processes. This algorithm had good performance in the field of polyolefin catalysts [5]. Huu designed a feature extraction scheme based on MobileNetV2 and combined it with a single trigger detector (STD) network to address the issue of low system accuracy in gestures and actions in smart homes. This model had an accuracy rate of over 90% and was suitable for practical applications [6]. Arora et al. designed a deep Convolutional neural network (DCNN) for face detection to solve the problem of computer automatically detecting whether people wear masks. It used feature detection and extraction techniques to identify whether a person is wearing a mask, which has high accuracy and is easy to deploy in embedded systems [7]. Li et al. proposed a fruit recognition method that combines BM2DK PCANet and support vector machine (SVM) classifier to improve the intelligent mechanization of orchards. The recognition rate of this method was 11.84% higher than that of PCANet [8]. Ahn et al. developed an automatic generation method for textile circuits using machine vision to address the issue of inaccurate capture of fabric images by image acquisition devices. This method improved the accuracy of fabric image recognition by 8.9% [9]. The application of machine learning algorithms in daily life is becoming increasingly diverse, providing a certain reference for the improvement of embedded systems. Sahoo et al. designed a distracted driving detection model using lightweight DCNN to solve the problem of distracted driving among traffic participants, with a classification accuracy of 99.93% [10]. Carreon and other researchers designed a normal timing device that used the execution path of embedded applications to address the issue of insufficient active security measures to provide protection for embedded systems. This method had significant advantages in terms of effectiveness and accuracy [11]. Raji et al. designed an embedded system multi-objective task scheduling framework to address the increased manufacturing process changes caused by active technological expansion. This framework considered the uncertainly of parameters in hardware and software components. According to experiments, the execution of this framework improved by 28.6% [12]. Thangavel et al. proposed an embedded unsupervised system to reduce conflicts between humans and wildlife in the Anamalai Tiger Reserve (ATR), which is used to recognize and classify animals based on their sound signals. This system could indicate

animal movement approaching forest boundary areas [13]. Krupa et al. designed a sparse and low memory optimization algorithm to solve the problem of model predictive control of tracking formula in embedded systems, and used the expansion of alternating direction multiplier method. This algorithm increased the domain of attraction and ensured the feasibility of recursion even in the case of sudden changes in reference [14]. Envelope et al. designed an embedded system with Multilayer Perceptron (feedforward)-Artificial Neural Network (MLP-ANN) and Field Programmable Gate Array (FPGA) to analyze potential triggering factors of epilepsy and formulate suitable medical interventions. The accuracy rate of seizure classification of this system was high [15].

In summary, there is adequate theoretical and implementation basis for employing MobilenetV2 SSD in embedded systems. Various studies have shown that it has broad application prospects, but there have been no relevant studies attempting to combine the two to solve the problem of attention detection. To enhance the detection performance and overcome the low detection efficiency challenges of the traditional MobileNet2-SSD algorithm, this study introduces EfficientNet in a ground-breaking manner to elevate the model's feature representation capabilities. By incorporating both multi-scale training and prediction strategies and selecting the Focal Loss loss function that is appropriate for object detection tasks, the MobileNet2-SSD algorithm can be enhanced. This study will apply the improved algorithm to embedded attention detection systems to further promote the development of the attention detection industry.

3. **Solution design of embedded attention detection model.**

This chapter is mainly divided into two sections. The first section focuses on improving MobileNet2 SSD and establishing a model for the IMobileNet2 SSD algorithm. In the second section, the algorithm is applied to embedded systems, and the processing and analysis results are fed back to other modules through Bluetooth for subsequent processing.

3.1. **Attention detection model based on IMobileNet2-SSD algorithm.**

MobileNet2 SSD is an object detection model based on the combination of MobileNetV2 and Single Shot MultiBox Detector (SSD) algorithm. MobileNetV2 is a lightweight CNN suitable for real-time image recognition tasks in resource constrained environments such as mobile devices. SSD is a commonly used object detection algorithm that can simultaneously detect multiple targets of different categories in an image, and provide their positions and confidence levels. MobileNet2 SSD combines its lightweight characteristics with SSD's multi-scale feature extraction and prediction technology, making the model have high speed and accuracy. It is widely used in real-time target detection tasks in mobile devices, embedded systems and edge computing, such as face detection, vehicle detection, etc. Although MobileNet2 SSD is a lightweight model, it may perform poorly in terms of detection accuracy and recall compared to some more complex object detection algorithms. In addition, MobileNet2 SSD may encounter difficulties in detecting small-sized objects. Due to the network structure and design of feature extractors, it may not be able to capture the detailed features of small targets, resulting in a decrease in detection accuracy. In response to the limitations of the MobileNet2 SSD algorithm in terms of detection performance and difficulty in detecting small target sizes, the following improvements have been made in this study. MobileNet2 SSD uses MobileNet as the feature extractor, but its accuracy is relatively low and its adaptability to large-sized inputs is weak. Therefore, it is recommended to utilize a stronger foundational network, such as EfficientNet, to enhance the model's capability of representing features and detecting performance. Firstly, it is necessary to integrate the various components of EfficientNet into the architecture of MobileNet2 SSD. In MobileNet2 SSD, MobileNet V2 plays a role in

feature extraction and requires EfficientNet to replace this function. This means removing the original MobileNet V2 layer and replacing it with a series of layers from EfficientNet. Secondly, the original MobileNet2 SSD uses a special structure such as Inverted Residual Block. At this time, it is necessary to substitute it with the MBConv layer in the EfficientNet version. Afterward, it is imperative to modify the pertinent parameters, like step size and the number of output channels to align it with the configuration of EfficientNet. The activation function is changed to Swish instead of ReLU6. The EfficientNet structure is shown in Figure 1.

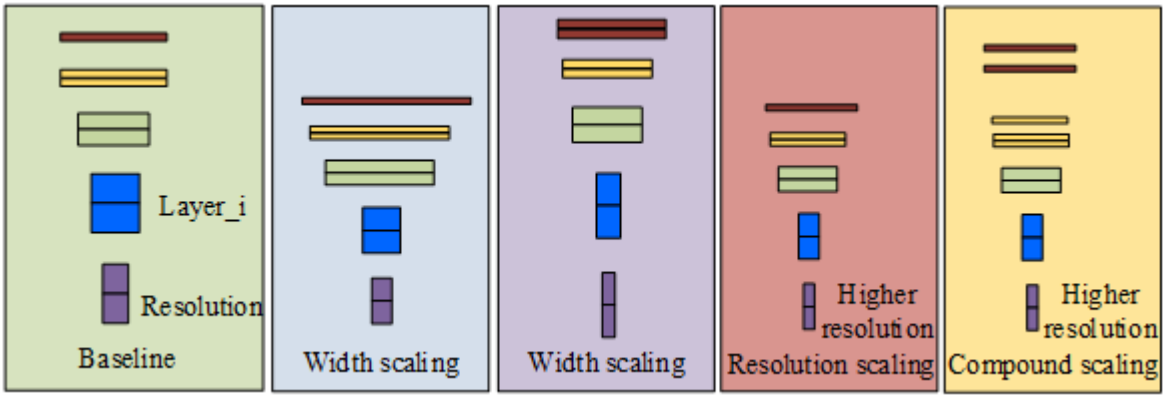

Figure 1. EfficientNet structure diagram

EfficientNet is a series of CNN models proposed in 2019. The design goal of these models is to achieve state-of-the-art performance in image classification tasks while also possessing efficient computing power. The key idea behind EfficientNet is to scale the dimensions of the baseline network architecture in a systematic manner. The process begins with a simple initial model and then expands it by utilizing composite coefficients to adjust the network's depth, width, and resolution. This method can strike a balance between model size and accuracy. The EfficientNet model utilizes various techniques to improve efficiency, including the use of deep separable convolution combinations to reduce parameters and computational complexity, as well as the use of mobile reverse bottleneck structures to effectively capture more complex patterns. The model also adopts a technique called compound scaling, which uniformly scales the depth, width, and resolution of the network. This method ensures that the network learns representations on inputs of different sizes and can better generalize. The EfficientNet model has achieved excellent results on various benchmark datasets, such as ImageNet, with fewer parameters and computational resources compared to previous state-of-the-art models [16]. They are also widely used in various computer vision tasks such as Transfer learning and feature extraction. Additionally, to enhance the model's performance, one can improve the network structure by increasing the depth and width of the network, adding more convolutional layers, and attention mechanisms [17]. In addition, multi-scale training and prediction are required, and relevant strategies are introduced to improve the detection ability of small targets by training and detecting images at different scales. In multi-scale image training, data augmentation methods are used to generate training samples of different scales. The original image is randomly scaled and cropped to generate multi-scale training images, and all generated images and their corresponding labels (i.e. the labels of the original image) are used as the training set. During the training process, the model will use these images of different scales to learn the features of the target. Meanwhile, multiple branches are introduced in the output layer of the model, each responsible for object detection at

different scales. This can improve the model's detection ability for targets of different scales. When conducting multi-scale training and prediction, it is necessary to ensure that features between different scales can be effectively fused and matched. The study uses a feature pyramid network to ensure this. Finally, the loss function Focal Loss, which is more suitable for target detection tasks, is used to optimize the detection results and improve the processing ability of small targets. The resulting IMobileNet2 SSD is shown in Figure 2.

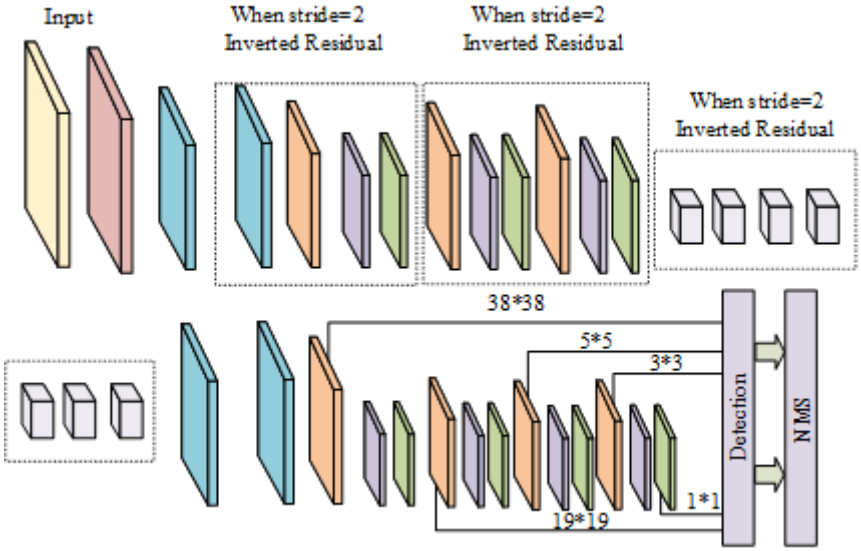

Figure 2. IMobileNet2 SSD algorithm flowchart

This model has strong feature representation ability and higher accuracy. The EfficientNet structure uses deeper networks and convolutional layers, which can learn more deep features and help improve the model's performance. EfficientNet incorporates a technique known as Composite Scaling that can balance trade-offs between network size and accuracy, finally enhancing the model's generalization performance across diverse input sizes. The ratio is selected as a feature of attention. Attention features are calculated using IMobileNet2 SSD, and left prefrontal lobe electroencephalo-graph (EEG) signals $\frac{\beta}{\theta}$ are collected, with the ratio used as the attention feature. The 256 EEG data collected each time are used to conduct a Fast Fourier Transform (FFT) to obtain attention features. The EEG sampling rate of the system is 128 Hz, and the frequency resolution (FR) of the FFT result calculated based on the FR formula is 0.5 Hz. Figure 3 shows the algorithm flow.

In the program, the power spectrum of EEG signal is calculated by FFT, and the attention feature is obtained. FFT is a calculation method derived from Discrete Fourier transform (DFT). The DFT formula for $n = 0, 1, \ldots, N-1$ is Equation (1).

$$x(k) = \sum_{n=0}^{N-1} x(n) W_N^{kn} \tag{1}$$

In Equation (1), there is $k = 0, 1, \ldots, N-1$, and $x(n)$ represents a finite length discrete signal. When N is an integer power of 2, DFT can be decomposed into two points of DFT, and $x(n)$ can be decomposed into the sum of two sequences, as shown in Equation (2).

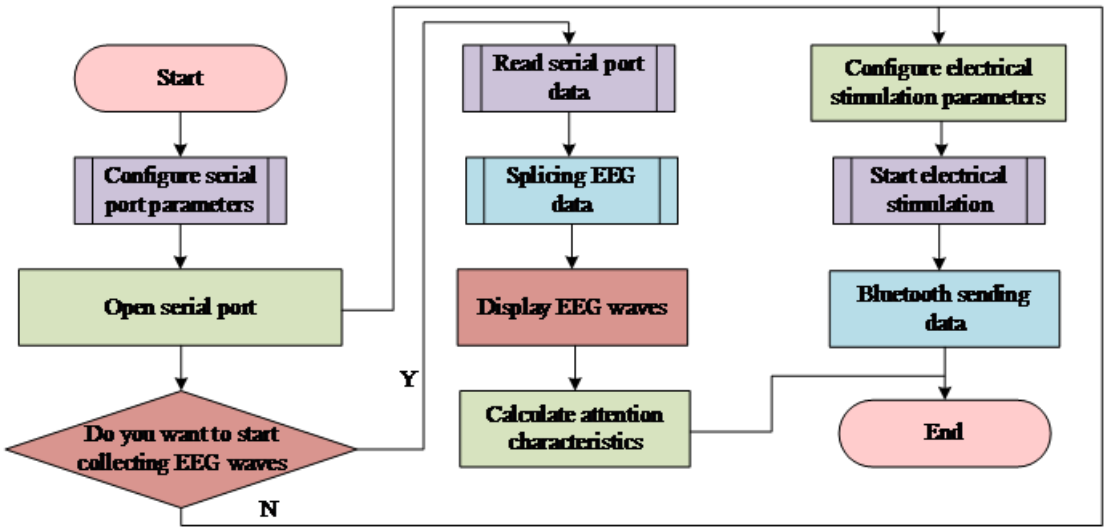

Figure 3. The flow of attention detection algorithm based on IMobileNet2-SSD

$$x(n) = x_1(n) + x_2(n) \tag{2}$$

In Equation (2), $x_1(n)$ and $x_2(n)$ are even and odd sequences. Due to the fact that $W_N^{2kn} = e^{-\frac{2\pi jkn}{N}}$, and $x(n)$ can be represented by transformations, as shown in Equation (3).

$$x(k) = \sum_{n=0}^{\frac{N}{2}-1} x_1(n) W_N^{2kn} + W_N^k \sum_{n=0}^{\frac{N}{2}-1} x_2(n) W_N^{2kn} \tag{3}$$

In Equation (3), $x_1(k)$ represents point $N/2$ of $x_1(n)$, and $x_2(k)$ represents point $N/2$ of $x_2(n)$. According to the symmetry of the Twiddle factor, $x(k)$ can be transformed, as shown in Equation (4).

$$\begin{cases} x(k) = x_1(k) + W_N^k x_2(k) \\ x\left(k + \frac{N}{2}\right) = x_1(k) - W_N^k x_2(k) \end{cases} \tag{4}$$

In Equation (4), $x_1(k)$ and $x_2(k)$ both have a period of $N/2$. By analogy, after $m-1$ decompositions, the N-point DFT is finally decomposed into $N/2$ 2-point DFTs.

3.2. **Embedded attention detection model.** The first step of the study fully demonstrates the feasibility of applying the MobileNet2 SSD algorithm in attention detection systems. In addition, it is necessary to combine the algorithm with embedded systems. Firstly, Emotiv Epoc+ is used for EEG signal acquisition. This head mounted electrode device has high resolution and comfort, and can collect high-quality EEG signals. Using the BioSemiActiveTwo signal conditioning circuit to collect simulated EEG signals, it can receive analog signals from electrodes and convert them into signals suitable for Analog to Digital Converter (ADC) processing. At this time, the EEG signals need to be amplified and filtered through the signal conditioning circuit to improve signal quality. Then, the EEG signal is converted into a digital signal through an ADC [18]. Finally, the obtained digital signal is sent via Bluetooth to the embedded platform of the Bluetooth receiver to complete the collection of EEG signals. The Raspberry Pi serves as the embedded platform for the research experiment, possessing adequate processing power and storage capacity to operate the MobileNet2 SSD algorithm. Additionally, it is furnished with

interfaces that permit reception of digital EEG signals transmitted via Bluetooth and execution of attention detection algorithms for data processing and display of outcomes. Then the parameters of the electrical stimulation are sent to the portable collection and control module through Bluetooth. Finally, the Howland current source generates the corresponding electrical stimulation to regulate attention. In this process, electrical stimulation regulation, as a means of attention regulation, requires operation within safety constraints. The stimulation parameters shall be kept within 1-2 mA to avoid discomfort or harm to the user. The workflow of the whole system is shown in Figure 4.

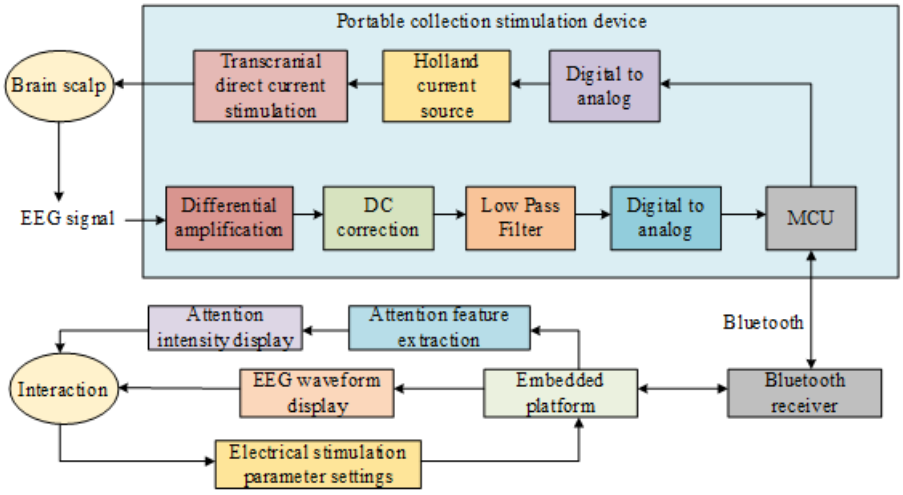

Figure 4. The workflow of an embedded attention detection model

Figure 4 shows the workflow of an embedded attention detection and regulation system. The ADC and Howland current sources (HCS) are selected to generate a constant current to power the Transcranial Direct Current Stimulation (tDCS) circuit in this system. HCS is a circuit configuration used to generate a stable constant current output. It was designed by American engineer Charles G Howland in 1962. The HCS uses a combination of differential amplifiers and feedback networks, which can achieve the required constant current output by adjusting the input voltage. It is typically used in precision measurements, sensor drives, and other applications that require precise current control. HCS can be applied to tDCS [19, 20]. In the tDCS, electrodes are placed on the scalp and a constant DC current is transmitted to the brain region to simulate, regulate, or affect neural activity, as shown in Figure 5.

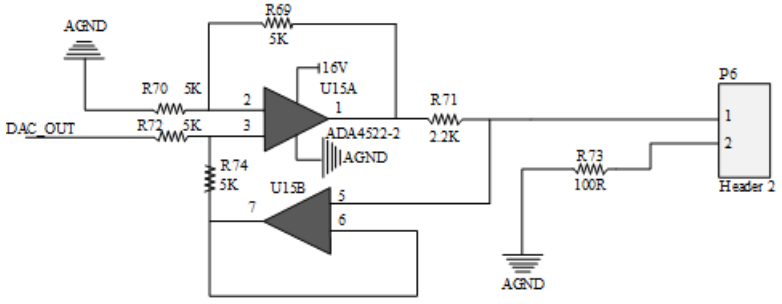

Figure 5. HCS structure diagram

HCS is a commonly used circuit used to provide a voltage controlled constant current source. It is widely used in fields such as medical and industrial automation. This current source is particularly suitable for electrical stimulation of the human scalp, as the impedance of the scalp is usually large and not fixed. HCS was invented by Brad Howland and is based on a current source of an operational amplifier. According to the principle of virtual short and virtual break of operational amplifiers, this circuit maintains the characteristic of equal input voltage of operational amplifiers (Op-amp Input), as shown in Equation (5).

$$V_+ = V_- \tag{5}$$

In Equation (5), V_+ and V_- represent the positive and negative voltage. The current flowing into Op-amp Input is shown in Equation (6).

$$I_+ = I_- \tag{6}$$

In Equation (6), I_+ is the positive voltage and I_- is the negative voltage. Further, the negative equation can be obtained as shown in Equation (7).

$$\frac{V_G - V_-}{R_{70}} = \frac{V_- - V_s}{R_{69}} \tag{7}$$

In Equation (7), R_{70} and R_{69} are the resistances of the 70th and 69th resistors, and the positive electrode equation in Equation (8) can be obtained.

$$\frac{V_{IN} - V_+}{R_2} = \frac{V_+ - V_{OUT}}{R_4} \tag{8}$$

In Equation (8), V_{IN} and V_{OUT} represent the input and output voltages, respectively. The ratio between R_{69} and R_{70} can be obtained by changing Equation (7), as shown in Equation (9).

$$\frac{R_{69}}{R_{70}} = \frac{V_- - V_s}{V_G - V_-} \tag{9}$$

In Equation (9), $V_G = 0$. When the ratio between R_{69} and R_{70} is equal to n, the expression for n can be obtained, as shown in Equation (10).

$$n = \frac{V_- - V_s}{V_G - V_-} = \frac{V_+ - V_{OUT}}{V_I - V_+} \tag{10}$$

In Equation (10), V_I represents the input voltage. According to Ohm's law, V_s can be obtained, as shown in Equation (11).

$$V_s = I_{OUT} \times (R_1 + R_x) \tag{11}$$

In Equation (11), R_x is the impedance of the human body. According to Ohm's law, the expression V_{out} can be obtained, as shown in Equation (12).

$$V_{OUT} = I_{OUT} \times R_x \tag{12}$$

In Equation (12), I_{OUT} represents the output current. Assuming that $V_+ = V_- = V_a$ substitutes Equation (12) into Equation (11), the expression for V_a can be obtained, as shown in Equation (13).

$$V_a = \frac{I_{OUT} \times (R_K + R_{71})}{1 + n} \tag{13}$$

In Equation (13), n represents the ratio between R_{69} and R_{70}, and then substituting Equations (13) and (12) into (10) to obtain the output current formula of HCS, as shown in Equation (14).

$$I_{OUT} = \frac{nV_I}{R_{71}} \tag{14}$$

In Equation (14), as long as the resistance values of R_{69}, R_{70}, R_2, and R_4 are the same, the current output by HCS is only related to the resistance R_{71}. Therefore, by selecting an appropriate resistor ER, the DAC can be controlled by the main control module to modify the output voltage and thus the output DC power. The front-end analog circuit is composed of a front-end amplification circuit, a DC correction circuit, a low-pass filtering circuit, and a voltage rise circuit. In this circuit, the reference signal and the active signal are first subtracted and amplified by the front-end amplification circuit to enhance the signal strength, as displayed in Figure 6.

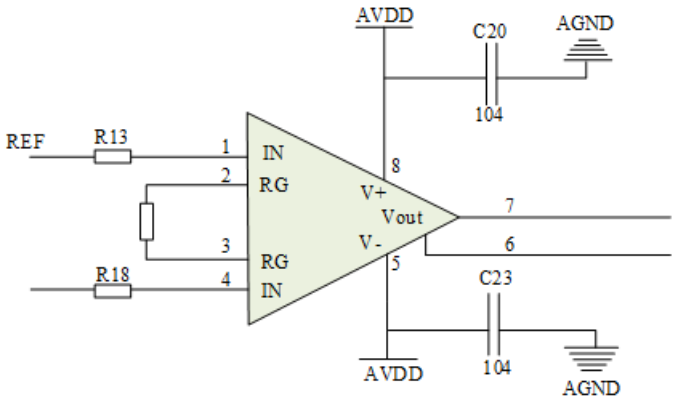

Figure 6. Front end analog circuit structure diagram

Next, the signal is subjected to DC bias correction and high-frequency noise filtering through a DC correction circuit and low-pass filtering circuit. Due to the fact that ADC uses a single power supply and can only process positive voltage signals, it is necessary to raise the corrected filtered signal voltage above the 0 scale to ensure that ADC can correctly collect negative voltage signals. To achieve this goal, a voltage rise circuit is used to perform gain processing on the signal. Finally, the raised signal is differentially inputted with the raised voltage and enters the analog-to-digital converter for analog-to-digital conversion, converting the analog signal into a digital signal for subsequent DSP or storage. Then, the digital signal is processed by MCU and interacted with the embedded platform through Bluetooth receiver.

4. **Algorithm performance testing and experimental verification of attention detection model.** This chapter mainly verifies the ability of the proposed attention detection model. The first section mainly compares and analyzes IMobileNet2 SSD with other algorithms to verify its comparative advantages, and uses different datasets to verify the model's generalization ability. The second section mainly conducts simulation experiments to verify the efficiency of the model in practical environments.

4.1. **Comparative analysis and generalization capability verification of MobileNet2 SSD algorithm.** This study established an attention detection system based on IMobileNet2-SSD, which solved the problem of lacking objective means for attention

detection. To evaluate the optimization ability of IMobileNet2 SSD on the MobileNet2 SSD algorithm, the experiment used Python 3.8 on the Windows 10 platform and used the Common Objects in Context dataset (COCO) to perform 600 iterations on the traditional MobileNet2 SSD and IMobileNet2 SSD models, respectively. Figure 7 shows the relationship between its training error and the number of iterations.

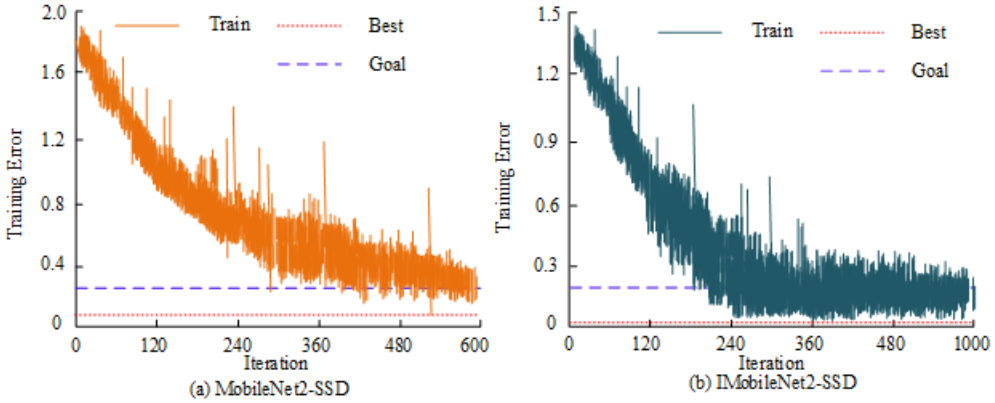

Figure 7. Comparison of MobileNet2-SSD and IMobileNet2-SSD Training Error

Figure 7 shows the training error trend of MobileNet2 SSD algorithm and IMobileNet2 SSD algorithm on the training dataset. The error reduction rate of IMobileNet2 SSD is not significantly different from MobileNet2 SSD in the first 120 iterations. But when the number of iterations between the two reaches the range of [120240], the error of IMobileNet2-SSD rapidly decreases. After 240 iterations, its training error has converged and finally converged to 0.05. The training error of MobileNet2-SSD only converged after 410 iterations, and finally converged to 0.35. It can be proven that the IMobileNet2-SSD model has a faster convergence rate and a lower error during final convergence. To ensure the effectiveness and reliability of the model, it is necessary to conduct generalization ability verification. The experiment used the Salient Object Prediction dataset (SALICON) to train the above two algorithms. Table 1 shows the composition of the SALICON parameter

Table 1. Basic parameters of the action dataset

Parameter type	Parameter scale
Natural scene image	10000
Resolution of images	480*640–640*480
Gazing at data	10000
Fixation Maps	9800
Training set	5000
Validation set	2500
Test set	2500

In Table 1, SALICON is a relatively large dataset that contains a large amount of attention data. Therefore, appropriate preprocessing and sampling are necessary when using this dataset for analysis. In addition, to conduct a horizontal comparative analysis of the application effect of IMobileNet2 SSD, the experiment introduced MobileNetV1-SSD, ShuffleNetV2 SSD (channel shuffling SSD, ShuffleNetV2 SSD), and NASNet SSD (Neural Architecture Search SSD, NASNet SSD) to compare with IMobileNet2 SSD. The

results of training the four algorithms using SALICON and COCO are shown in Figure 8.

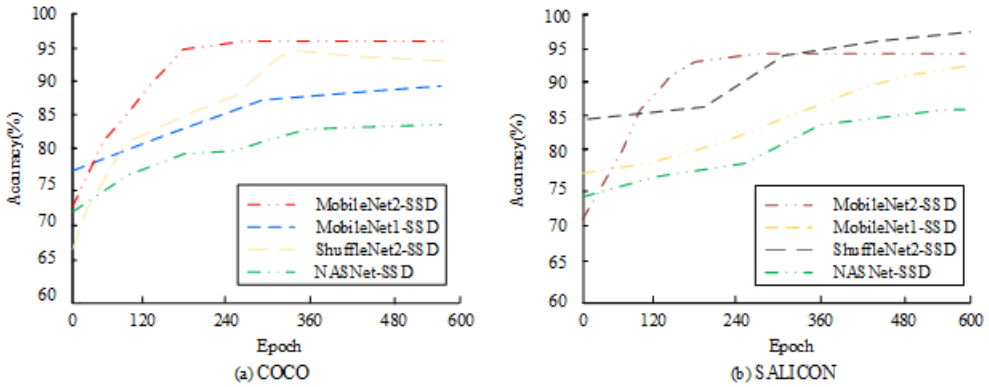

Figure 8. Comparison of MAWI and CAIDA data set

Figure 8 is the trend curve of the accuracy of various algorithms in the SALICON and COCO datasets as a function of the number of training rounds. In Figure 8(a), the IMobileNet2-SSD model performed the best, with accuracy tending to converge after 180 iterations, and the final accuracy converging to 95.3%. The other three models converged after 240 iterations, and the rate of convergence was slow. In Figure 8(b), the accuracy of all algorithm models varies to varying degrees when using SALICON. However, the convergence of IMobileNet2-SSD showed relatively minimal changes, with an accuracy decrease of only 1.1% and tending to converge after 170 iterations. However, MobileNet1-SSD showed the greatest change, with an accuracy increase of 9.8% during convergence and ultimately converging to 96.9%. The above results prove that IMobileNet2-SSD has the advantages of fast rate of convergence, high detection accuracy, good algorithm stability and strong generalization ability compared with the other three algorithms. Multi-scale training and prediction strategies were introduced to improve the detection ability of IMobileNet2-SSD to detect small targets. To further validate the functionality of the four algorithms in the embedded system, their performance was evaluated using ROC curve and AUC value as the criteria, with the SALICON dataset employed for testing. The experimental results are shown in Figure 9.

Among these four algorithms, the IMobileNet2 SSD algorithm performed the best, with an area under the ROC (AUC) curve of 0.945. This means that in the task of distinguishing positive and negative samples, the classifier performance of IMobileNet2 SSD algorithm is very excellent, even in cases of very high false positive rates, it can still maintain a high true positive rate. In contrast, the AUC value of the MobileNet1 SSD algorithm was 0.913, which was slightly lower than the IMobileNet2 SSD algorithm, but still showed good classifier performance. The AUC values of IMobileNet2 SSD algorithm and NASNet SSD algorithm were 0.878 and 0.813, respectively, indicating relatively poor performance.

4.2. **Experimental verification of attention detection model. To verify the.** effectiveness of IMobileNet2 SSD in practical application, five volunteers were invited to perform tasks such as GO/NOGO, annotating crosses, and resting. GO/NOGO task is a commonly used experimental task in Cognitive psychology, which is used to study and evaluate the ability of individuals in inhibition response and attention control. In this task, participants need to respond or not respond to specific stimuli, usually based on a predetermined rule or condition. GO/NOGO usually presents a series of stimuli, most

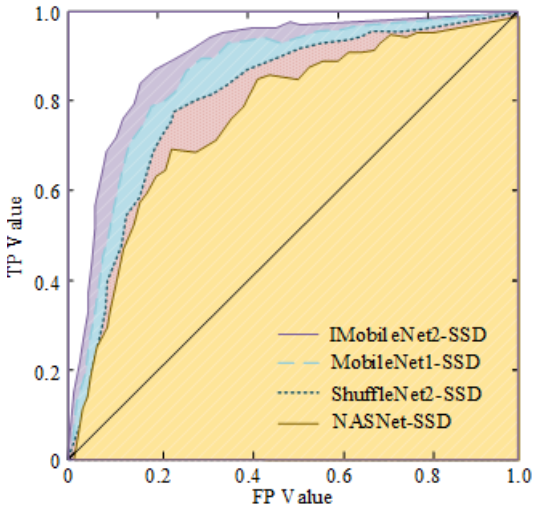

Figure 9. ROC curves of each algorithm

of which are "GO" stimuli, which require participants to respond. A few are "NOGO" stimuli, which require participants to suppress their reactions and not make any actions. Participants need to quickly and accurately distinguish and execute correct responses according to task requirements. By analyzing participants' performance in GO/NOGO tasks, researchers can evaluate their ability to suppress reactions and attentional control, as well as the frequency of error suppression. This task is commonly used to study ADHD, addiction, and other cognitive control related diseases, as well as to explore the mechanisms of attention and self-control in many other fields, as shown on Figure 10

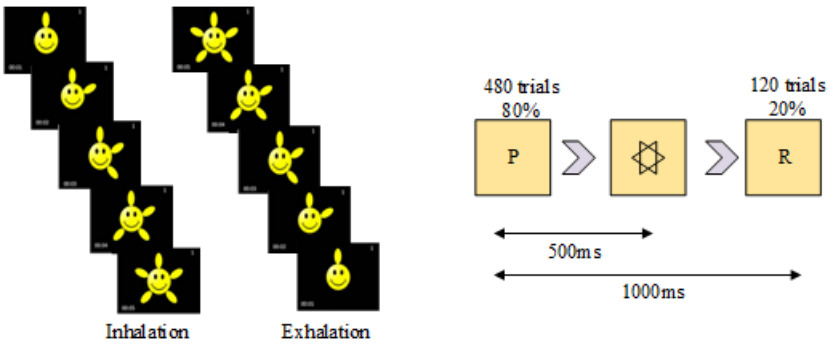

Figure 10. GO/NOGO task diagram

The overall process of the attention detection experiment is as follows. Firstly, volunteers are required to perform a GO/NOGO task that takes about five minutes to grow. After the task is completed, volunteers are required to rest for two minutes. Next, the volunteers need to focus their attention on the cross on the screen for three minutes. After the task is completed, the volunteers should rest for two minutes. Finally, rest for three minutes, during which the volunteers need to divert their attention. After the rest, the attention detection experiment is completed. After obtaining the attention feature data of the above three tasks, MATLAB is used to analyze the attention feature data. Based on the rhythmic components of signals with β and θ: the ratio E_β/E_θ of frequency band energy is used as a feature to measure attention. The average value of each task E_β, the average value of E_θ, and the average value of E_β/E_θ is calculated. Simultaneously, the

E_β/E_θ-waveform of each task for observation is drawn, and the experimental results are displayed in Figure 11.

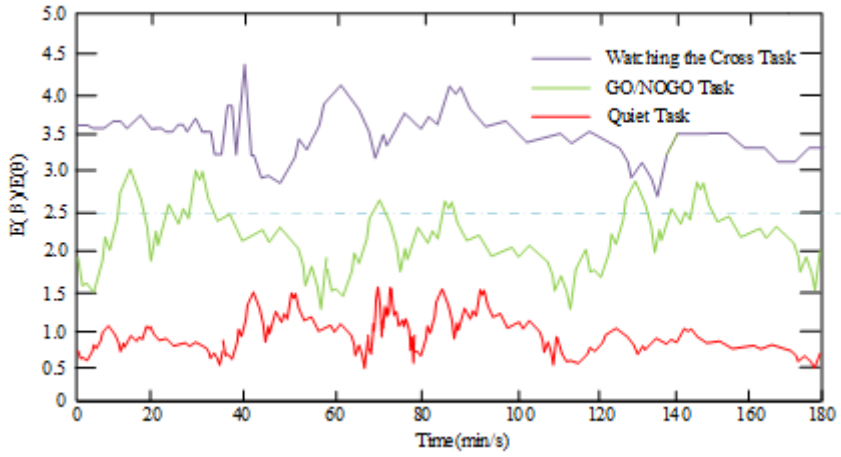

Figure 11. Waveform for each task

Figure 11 shows the E_β/E_θ waveform of each task. The x-axis in the figure represents the time axis, the y-axis represents the amplitude axis. The purple, green, and red curves represent the E_β/E_θ-waveform of the gaze on the cross, GO/NOGO, and resting tasks. Their curves fluctuate in the range of [2.5, 4.5], [1.5, 3.0], and [0.5, 1.5], respectively. This indicates the relaxed state of the volunteers in this task. It is not difficult to find that the volunteer's E_β/E_θ was significantly higher than the E_β/E_θ of the resting task when performing two focused attention tasks: gazing at the cross and GO/NOGO. These results indicated that the two tasks had a distinct effect on volunteer attention. It was possible that the task was required for increased brain activity in the volunteers, resulting in a higher brain wave amplitude. These findings facilitated the understanding of individual cognitive and attention control abilities under different tasks and provide an important reference for further research and application of brain waves. During the experiment, the results of the volunteers' attention regulation through electrical stimulation were also recorded, as shown in Figure 12.

Figure 12 shows the E_β/E_θ statistical chart of volunteers performing Schulte tasks before and after electrical stimulation. In Figure 12, before electrical stimulation, the E_β/E_θ ratio of volunteers performing Schulte tasks is between [1.1, 1.9]. After electrical stimulation, the E_β/E_θ ratio increased by more than 50%. This means that the system can increase the attention of volunteers by electrically stimulating the cerebral cortex when people are not paying attention. The systematic electrical stimulation scheme has a significant effect on increasing the attention of volunteers. In summary, by electrically stimulating the cerebral cortex, the system can increase the attention level of volunteers when they are not paying attention. This discovery is of great significance for further research and application of EEG stimulation technology to improve people's attention. Future research can further explore the parameter settings and stimulus location selection of electrical stimulation, optimize attention enhancement effects, and apply them to real life to help people in need of improving attention.

5. **Conclusion.** After in-depth research on traditional attention detection models, an embedded attention detection model IMobileNet2 SSD based on the improved MobileNet2 SSD algorithm was successfully designed. This model solved problems such as insufficient

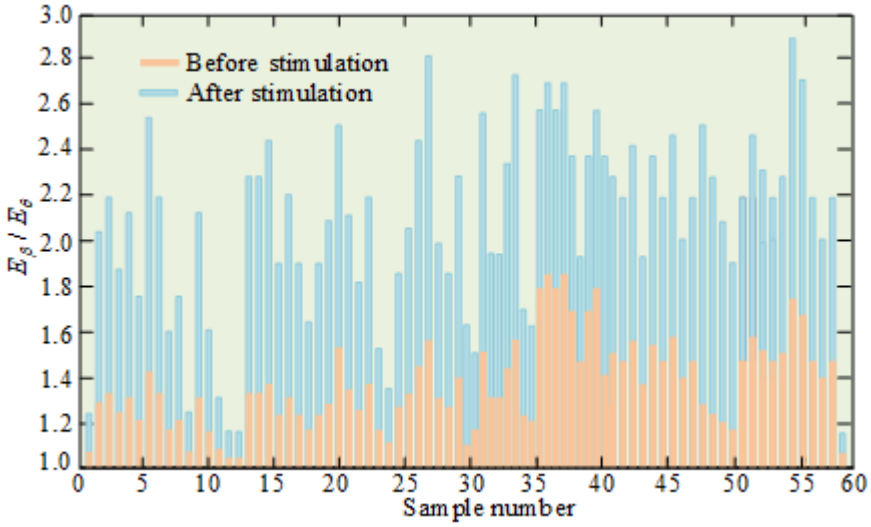

Figure 12. The results of attention regulation by electrical stimulation

accuracy, limited detection performance, and difficulty in detecting small-sized objects. To verify the performance of the model in practical applications, volunteers were invited to participate in experiments, display waveform diagrams, and provide data support. The MobileNetV1-SSD, ShuffleNetV2-SSD, and NASNet-SSD algorithms were introduced for horizontal comparison, and rigorous training tests were conducted using SALICON and COCO datasets. The results showed that IMobileNet2 SSD had excellent performance, with an AUC value of 0.945, indicating excellent classifier performance with a high true positive rate.

Compared with other algorithms, IMobileNet2 SSD had significant advantages in computational efficiency and convergence speed, with only a decrease of 1.1%. It tended to stabilize after 170 iterations with a small change in accuracy. Its superiority was reflected in the classifier performance, computational efficiency, convergence speed, and stable performance on different datasets. It had higher practical value and broader application prospects. Compared with other methods, the superiority of the proposed IMobileNet2-SSD method was reflected in multiple aspects: Firstly, it had excellent performance in classifier performance, which could maintain a high true positive rate in complex practical application scenarios. Secondly, it had obvious advantages in computational efficiency and convergence speed, and could achieve a high accuracy convergence state in a shorter time. Finally, its performance on different datasets was relatively stable, demonstrating good model generalization ability. All of these made IMobileNet2 SSD have higher practical value and broader application prospects in practical applications.

In addition, this study validated the performance of attention detection models, but did not address how to apply these detection results to real-time feedback systems. Future work can explore how to combine this attention detection model with real-time feedback mechanisms for application in education, driving safety, or other scenarios that require attention regulation. Currently, the application of attention detection technology may involve personal privacy and ethical issues, especially in environments such as education and the workplace. Future research needs to address these aspects to ensure the rational and compliant application of technology.

REFERENCES

[1] M. U. Danjuma, B. Yusuf, and I. Yusuf, "Reliability, availability, maintainability, and dependability analysis of cold standby series-parallel system," Journal of Computational and Cognitive Engineering, vol. 1, no. 4, pp. 193-200, 2022.

[2] S. Nimrah, and S. Saifullah, "Context-Free word importance scores for attacking neural networks," Journal of Computational and Cognitive Engineering, vol. 1, no. 4, pp. 187-192, 2022.

[3] P. A. Ejegwa, and J. M. Agbetayo, "Similarity-distance decision-making technique and its applications via intuitionistic fuzzy pairs," Journal of Computational and Cognitive Engineering, vol. 2, no. 1, pp. 68-74, 2022.

[4] Y. Li, S. Liu, Y. Zhao, C. Lei, and J. Zhang, "Blind nonlinearity equalization by machine-learning-based clustering for QAM-based quantum noise stream cipher transmission," China Telecom, vol. 19, no. 8, pp. 127-137, 2022.

[5] W. Yang, and W. H. Sun, "Application and prospect of machine learning in polyolefin catalysts," Chinese Science Bulletin, vol. 67, no. 17, pp. 1870-1880, 2022.

[6] P. N. Huu, H. N. T. Thu, and Q. T. Minh, "Proposing a Recognition System of Gestures Using MobilenetV2 combining single shot detector network for smart-home applications," Journal of Electrical and Computer Engineering, vol. 2021, no. 3, pp. 1-18, 2021.

[7] M. Arora, S. Garg, and A. Srivani, "Face mask detection system using Mobilenetv2," International Journal of Engineering and Advanced Technology, vol. 10, no. 4, pp. 127-129, 2021.

[8] X. N. Li, H. Wu, X. Yang, P. Xue, and S. Tan, "Multiview machine vision research of fruits boxes handling robot based on the improved 2D kernel principal component analysis network," Journal of Robotics, vol. 12, no. 2021, pp. 1-13, 2021.

[9] J. Ahn, and S. Kim, "Automated textile circuit generation using machine vision and embroidery technique," Textile Research Journal, vol. 92, no. 11-12, pp. 1977-1986, 2022.

[10] G. K. Sahoo, S. K. Das, and P. Singh, "A deep learning-based distracted driving detection solution implemented on embedded system," Multimedia Tools and Applications, vol. 82, no. 8, pp. 11697-11720, 2023.

[11] N. A. Carreon, S. Lu, and R. Lysecky, "Probabilistic estimation of threat intrusion in embedded systems for runtime detection," ACM Transactions on Embedded Computing Systems (TECS), vol. 20, no. 2, pp. 1-27, 2021.

[12] M. Raji, and M. Nikseresht, "UMOTS: An uncertainty-aware multi-objective genetic algorithm-based static task scheduling for heterogeneous embedded systems," The Journal of Supercomputing, vol. 78, no. 1, pp. 279-314, 2022.

[13] S. Thangavel, and C. S. Shokkalingam, "The IoT based embedded system for the detection and discrimination of animals to avoid human–wildlife conflict," Journal of Ambient Intelligence and Humanized Computing, vol. 13, no. 6, pp. 3065-3081, 2022.

[14] P. Krupa, I. Alvarado, D. Limon, and T. Alamo, "Implementation of model predictive control for tracking in embedded systems using a sparse extended ADMM algorithm," IEEE Transactions on Control Systems Technology, vol. 30, no. 4, pp. 1798-1805, 2021.

[15] R. S. Envelope, N. B. Envelope, D. J. Envelope, and E. U. Envelope, "Towards efficient implementation of MLP-ANN classifier on the FPGA-based embedded system," IFAC-PapersOnLine, vol. 55, no. 4, pp. 207-212, 2022.

[16] H. Garg, B. Sharma, S. Shekhar, and R. Agarwal, "Spoofing detection system for e-health digital twin using EfficientNet Convolution Neural Network," Multimedia Tools and Applications, vol. 81, no. 19, pp. 26873-26888, 2022.

[17] Y. T. Liang, C. F. Li, H. L. Liu, J. H. Zhang, Y. B. Ji, and X. B. Li, "Cross layer bilinear fusion improvement of EfficientNet-B5 graptolite image classification method," Applied Science and Technology, vol. 49, no. 5, pp. 15-23, 2022.

[18] D. Jeong, and C. Yoo, "Voltage-controlled-oscillator based continuous-time Sigma-delta modulator analog-to-digital converter," Journal of the Institute of Electronics and Information Engineers, vol. 58, no. 4, pp. 32-39, 2021.

[19] C. R. Nettekoven, R. Jurdon, T. Nandi, N. Jenkinson, and C. J. Stagg, "Cerebellar anodal tDCS does not facilitate visuomotor adaptation or retention," Brain Stimulation, vol. 15, no. 6, pp. 1435-1438, 2022.

[20] X. Liu, C. Yu, H. H. Yu, Z. Chen, and D. Zhou, "The cognitive function effects of prefrontal tDCS for depression: A system review," Stress and Brain, vol. 1, no. 2.