FACIAL EMOTION RECOGNITION SYSTEM

QIU HONGBO

NOR SAMSIAH SANI

Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia

ABSTRAK

Facial Emotion Recognition (FER) technology is playing an increasingly important role in today's rapidly evolving field of intelligent interaction, especially in high-risk occupations that have stringent requirements on emotional states, such as medical emergencies, airline piloting, and security monitoring. It is able to analyze human face expressions to determine various emotional changes of individuals in real time, thus helping systems or decision makers to take timely countermeasures. For example, in the cockpit, FER can detect whether the pilot's attention is declining due to fatigue or stress, and issue a timely warning to change driving to prevent potential accidents; or in public places, it can remind security personnel to go forward in time based on recognizing panic and anger in the crowd. This technology not only enhances the safety and intelligence of the working scene, but also makes the human-machine interaction more humanized, truly realizing the goal of using technology for people.

INTRODUCTION

Facial Emotion Recognition (FER) Systems: Enhancing Human-Computer Interaction for High-Risk Occupations

Facial Emotion Recognition (FER) technology plays a critical role in the increasingly important human-computer interaction, especially in high-risk occupations that require real-time emotion monitoring. In order to meet diverse user needs, this project develops a comprehensive FER system with three core functions aimed at providing flexible and efficient emotion recognition services:

Local Image Recognition: Users can upload local images, and the system will analyze and recognize facial expressions in the images and output corresponding emotion labels. This function is suitable for analyzing the emotions of historical pictures, such as investigation and research, archive management and other scenarios, which is convenient for users to review and understand the past emotional state.

Video Recognition: Users can upload video files, and the system will analyze facial expressions in the video frame by frame, and provide a timeline of emotional changes and analysis of emotional trends. This feature is suitable for monitoring long-term emotional changes, such as training effectiveness evaluation, customer experience analysis and other scenarios, to help users understand the dynamic changes of emotions.

Instant Camera Recognition: Users can capture facial expressions in real-time via the camera, and the system will immediately recognize the emotion and provide real-time emotional feedback. This feature is suitable for scenarios that require quick response and real-time monitoring of emotions, such as medical diagnosis, psychological counseling, and game interaction. Especially in high-risk occupations, such as drivers, security personnel, psychologists, etc., real-time emotion monitoring can help identify potential risk factors and take timely interventions, thus improving work efficiency and safety.

With these three major functions, our FER system can provide users with a comprehensive, flexible, and efficient emotion recognition solution that enhances human-computer interaction and plays a key role in high-risk occupations, providing them with an efficient and intelligent user experience.

The main goal of this project is to develop a Facial Emotion Recognition (FER) system that is both applicable to real-life scenarios and capable of reading data. This is designed to enable the system to accurately recognize and classify facial emotions in a variety of contexts. To achieve this goal, the project will focus on the following key aspects: firstly, building a powerful FER system that is capable of efficiently processing dynamic facial expressions from video and live cameras, and detecting rapid changes in emotions in a timely manner. Secondly, the project will thoroughly evaluate the generalisation performance of the model by conducting rigorous tests and practices on data-enhanced diverse datasets to ensure its reliability in real-world applications. Finally, to facilitate user understanding and application, the project will also develop a visual FER system that not only provides intuitive image outputs to show the recognised facial expressions, but also presents more detailed emotion information in the form of data outputs. By integrating these techniques and efforts, the project aims to build a reliable, accurate, and easy-to-use FER model, which will provide a solid foundation for real-world applications of FER technology.

STUDY METHODOLOGY

The approach used to develop this Facial Emotion Recognition (FER) system is the incremental development model. This model is a flexible approach to building systems that allows for incremental additions and refinements to the system functionality during the development process so that it can respond quickly to changes in user requirements. The incremental development model emphasises the continuous delivery of working versions and iterative improvements based on user feedback. Here are some of the processes included in the incremental development model.

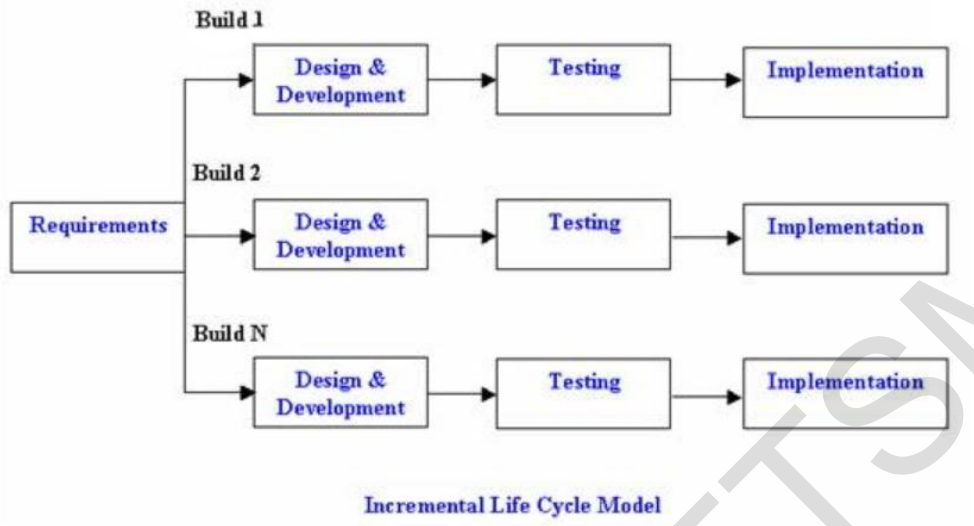

Figure 1.1 Methodology Diagram

Requirements Analysis Phase

The incremental development model begins with the requirements analysis phase. In this phase, the most important thing is to define the core objectives and application scenarios of the FER system being developed. For example, will it be used for real-time sentiment monitoring or for offline data analysis? Define the initial requirements and create a list of core features, e.g., basic emotion recognition, facial expression analysis, support for different data sources (images, videos, cameras).

Design phase

Based on the information gathered during the requirements analysis phase, the design of the FER system can be determined. In this phase, system architecture design, algorithm selection, and identification of interfaces and software GUI libraries for key modules are required.

Development Phase

In the development phase, an incremental functional version of the system is developed based on the identified requirements and design. Each incremental version contains a portion of the core functionality and can be run and tested independently. For example, the image recognition function can be developed first, and then the video recognition and real-time camera recognition functions can be added gradually.

Testing phase

When the functionality of each incremental version is available, it is fully tested. This phase includes unit testing, integration testing and system testing to ensure the quality and stability of the incremental version. Also random data from the internet that the machine has not seen before is used for model performance testing. Once testing is complete, it can be released to users for experience and feedback.

Iteration and Improvement

Based on user feedback, the system is continuously improved and optimised. In the next incremental release, new features can be added, existing features can be improved, or bugs can be fixed. the incremental development model encourages continuous improvement throughout the development process to ultimately meet the needs of users.

DECISIONS AND CONVERSATIONS

The client software based on face recognition has been developed, and according to the minimalist style of the interface, users can easily grasp the operation of the FER system.The main page of FER can see its three main functions: 1. image recognition 2. real-time recognition 3. video recognition . The upper left corner of the main page is a design to synchronize the display of the real time, which will be more important in the real time processing tasks

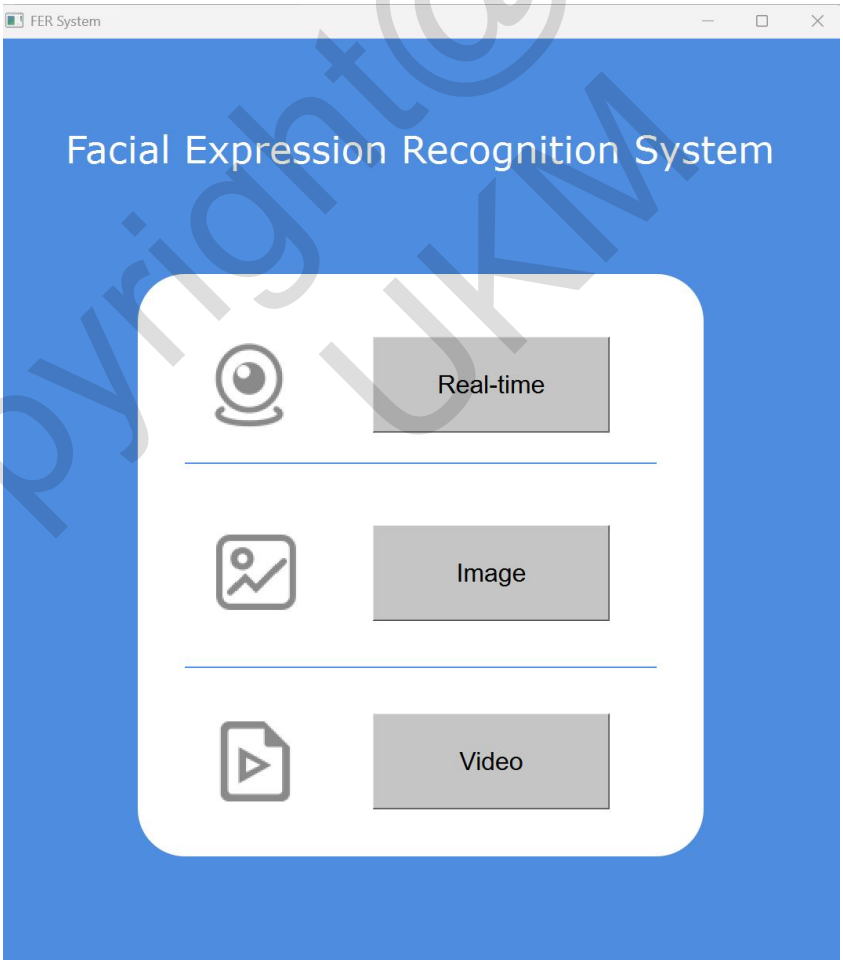

Figure 1.2 Design UI

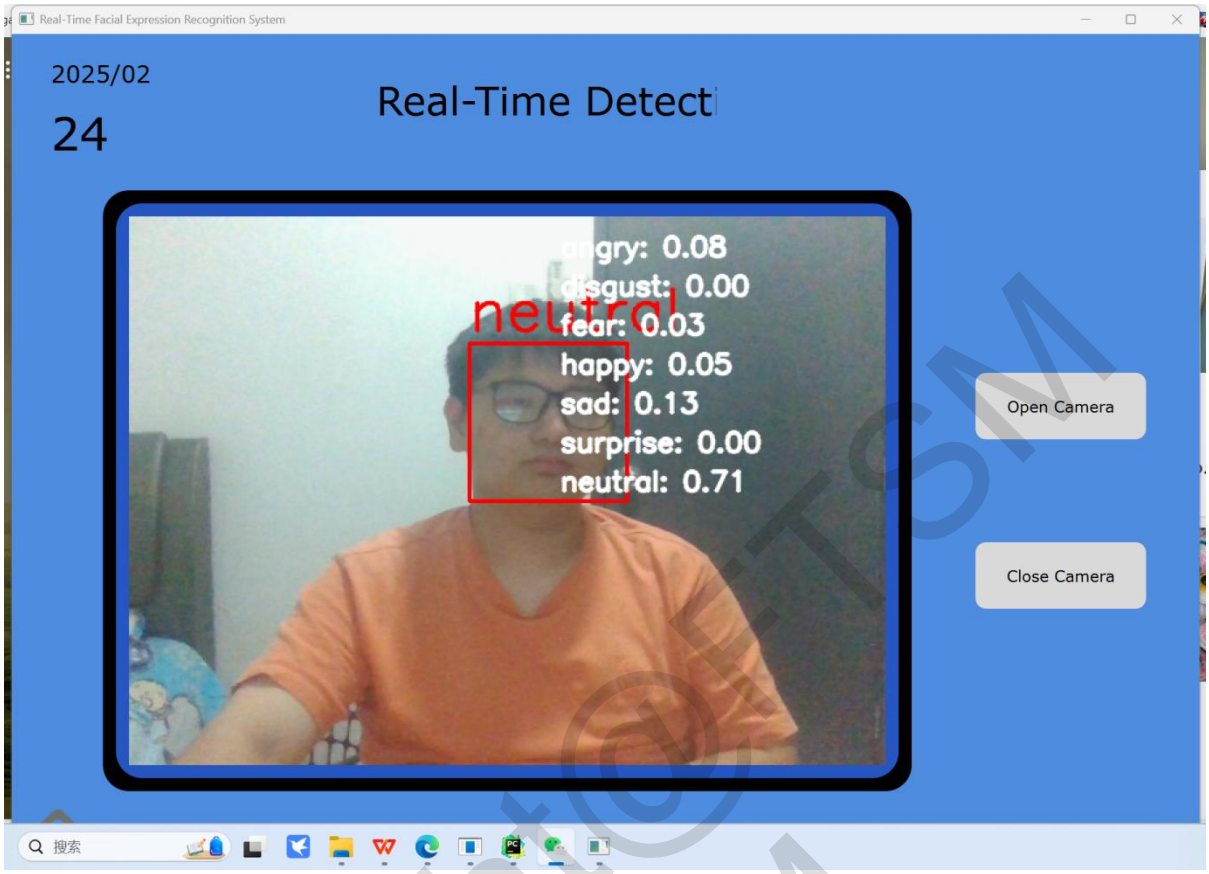

Figure 1.3 Real-Time in Camera Detect

Using Picture 1.3 as an example, there are two buttons on the interface for the user to operate. The user is free to turn the camera on and off because the camera is detected in real time. When the camera is turned on, the user needs to pass the permission to turn on the camera by himself, this is to avoid letting the user think that the system can turn on the camera arbitrarily.

When the user opens the camera, the output screen is the real-time camera plus the faces and expressions recognized by the machine are automatically classified. The user is able to know the result of the face emotion recognized through the output. And this function page has a time in the upper left corner and a back to home page marker in the lower right corner. This is to make it easier for users to go back to the homepage and switch to other functions.

Incidentally, the output screen has weight values for each of the six expressions. Although in practical use, it is less necessary for the user to know the values of the other weights. But providing the weighting data lets the user know that the model's inferences are not made out of thin air, but are well-founded. In addition, the benefit of providing the weights for the extra expressions is to help the user analyze more complex mixed expressions.

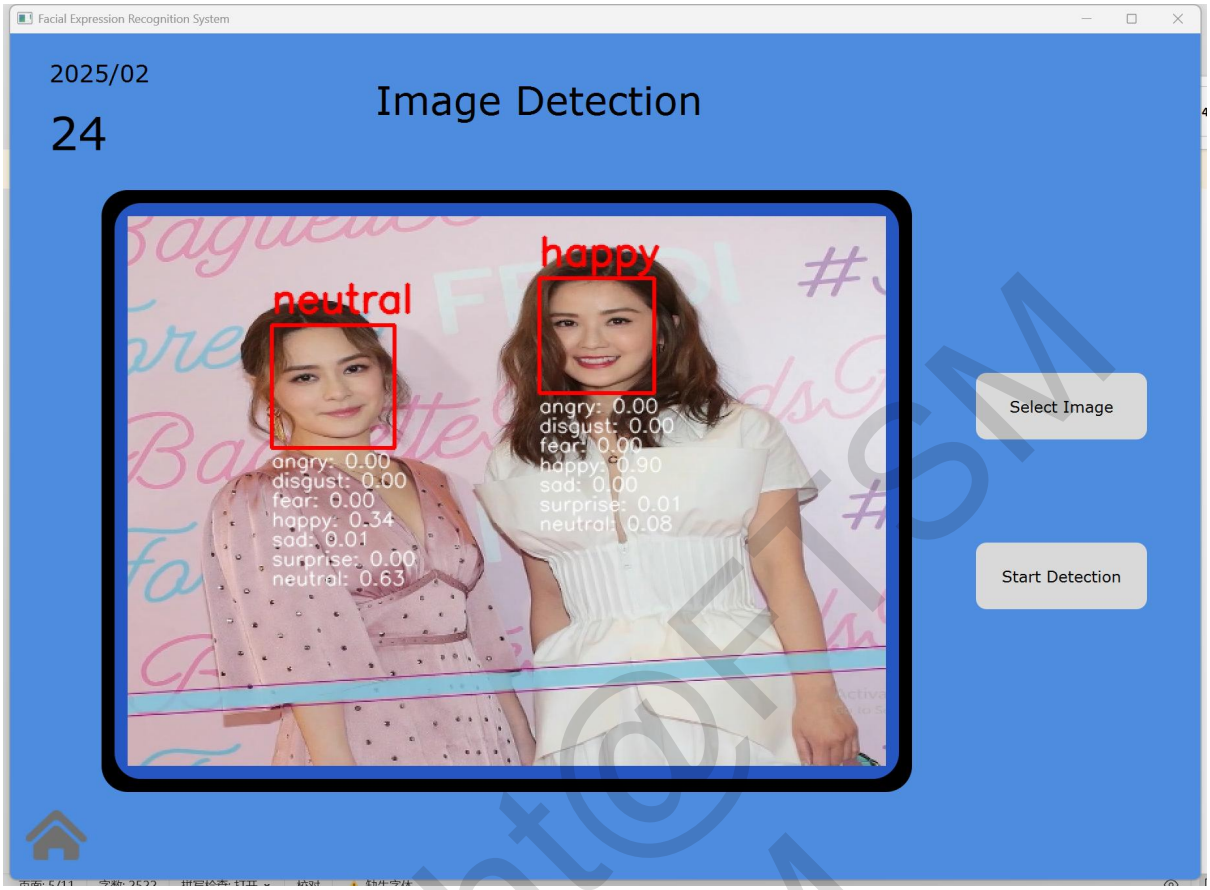

Figure 1.4 Image in Camera Detect

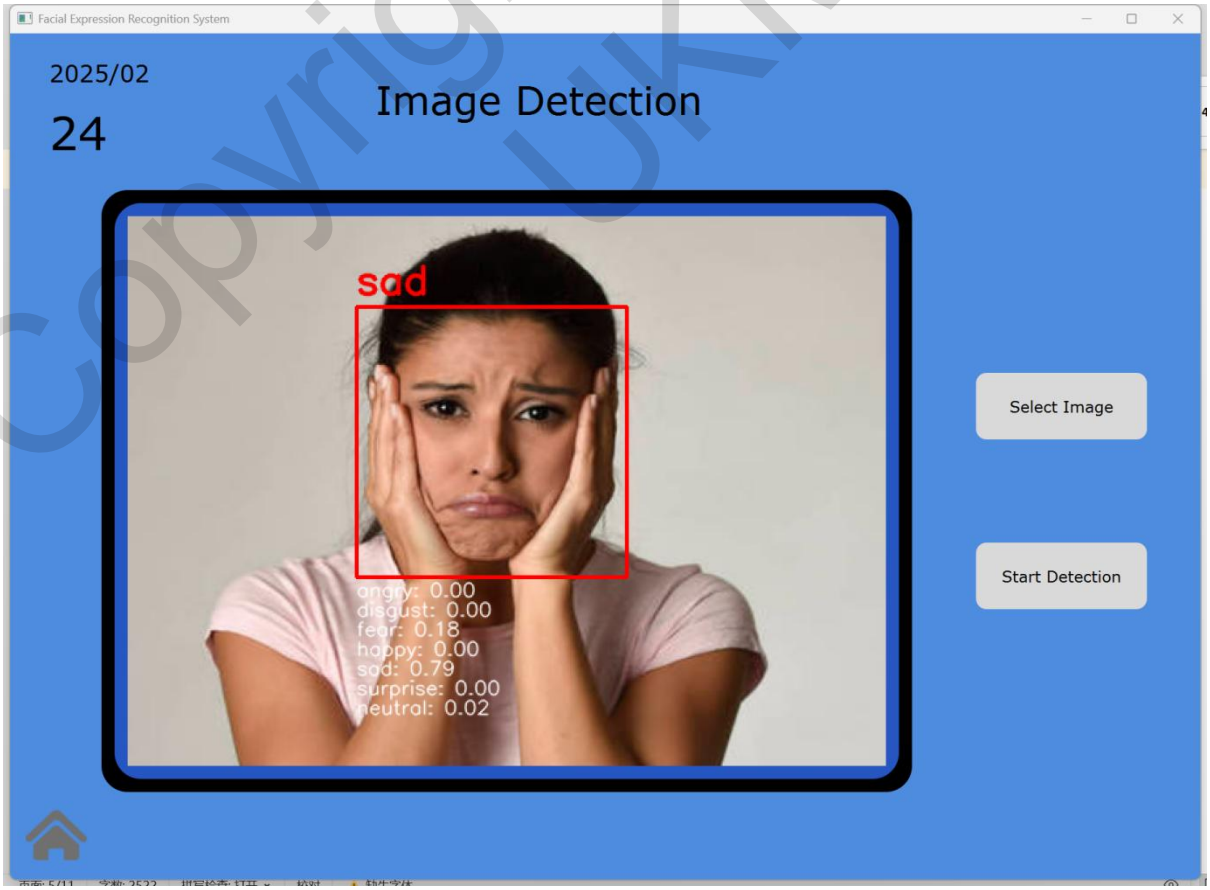

Figure 1.5 Image in Detection

The Face Expression Recognition (FER) system in this project clearly demonstrates its core functionality in image expression recognition through Figure 1.4 and Figure 1.5. Figure 1.4 vividly demonstrates the system's ability to handle complex scenarios: in an image containing multiple people, the machine successfully detects and frames each face, while accurately recognizing the weights of the expressions and other expressions presented by each face.

This fully demonstrates that the present FER system has a strong multi-target face expression recognition capability, and is able to cope with real scenarios where multiple people are present at the same time. In contrast, Figure 1.5 may demonstrate the recognition effect of this system under different lighting conditions and backgrounds, and the recognition effect on people of different age groups. The robustness and accuracy of the system under different conditions are further verified. By comparing these two figures, we can get a comprehensive understanding of the powerful function and wide applicability of this FER system in image expression recognition.

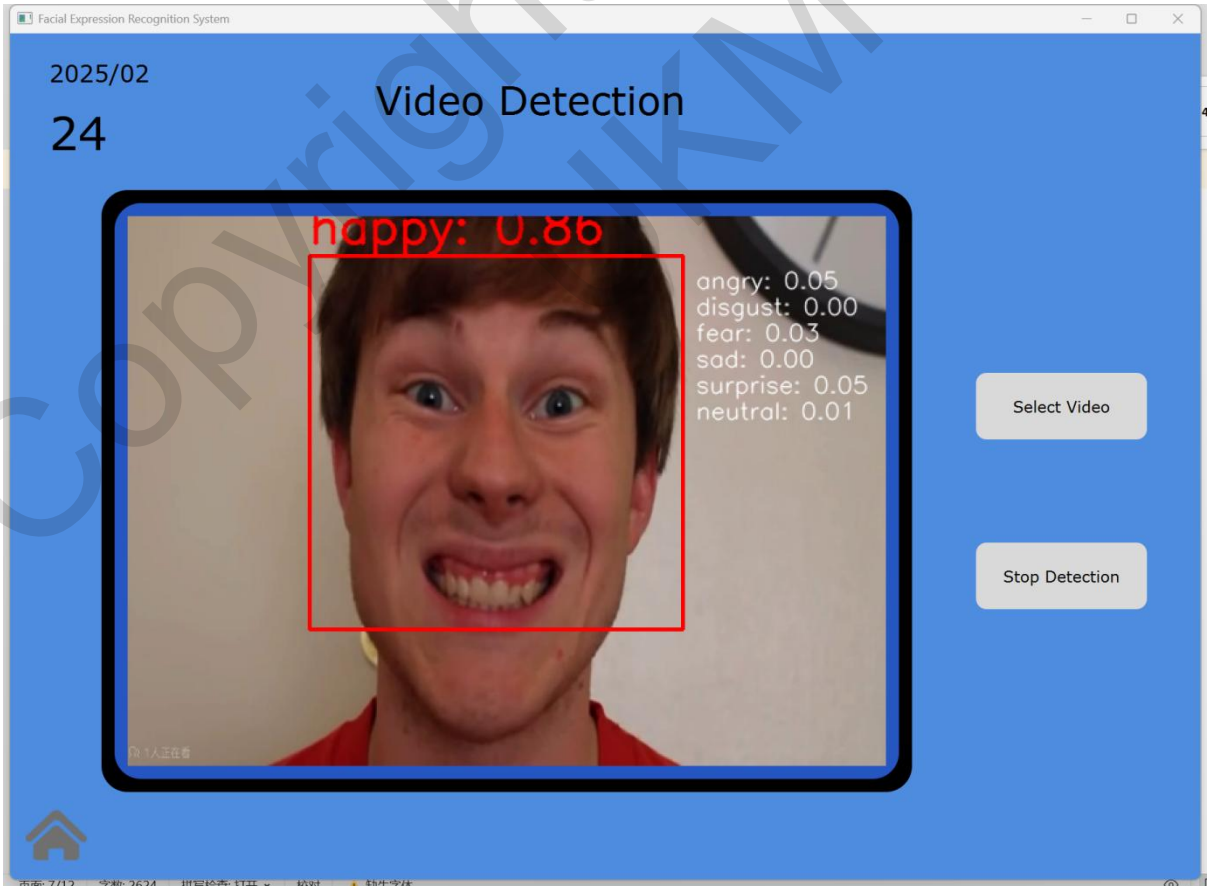

Figure 1.6 Video Detection

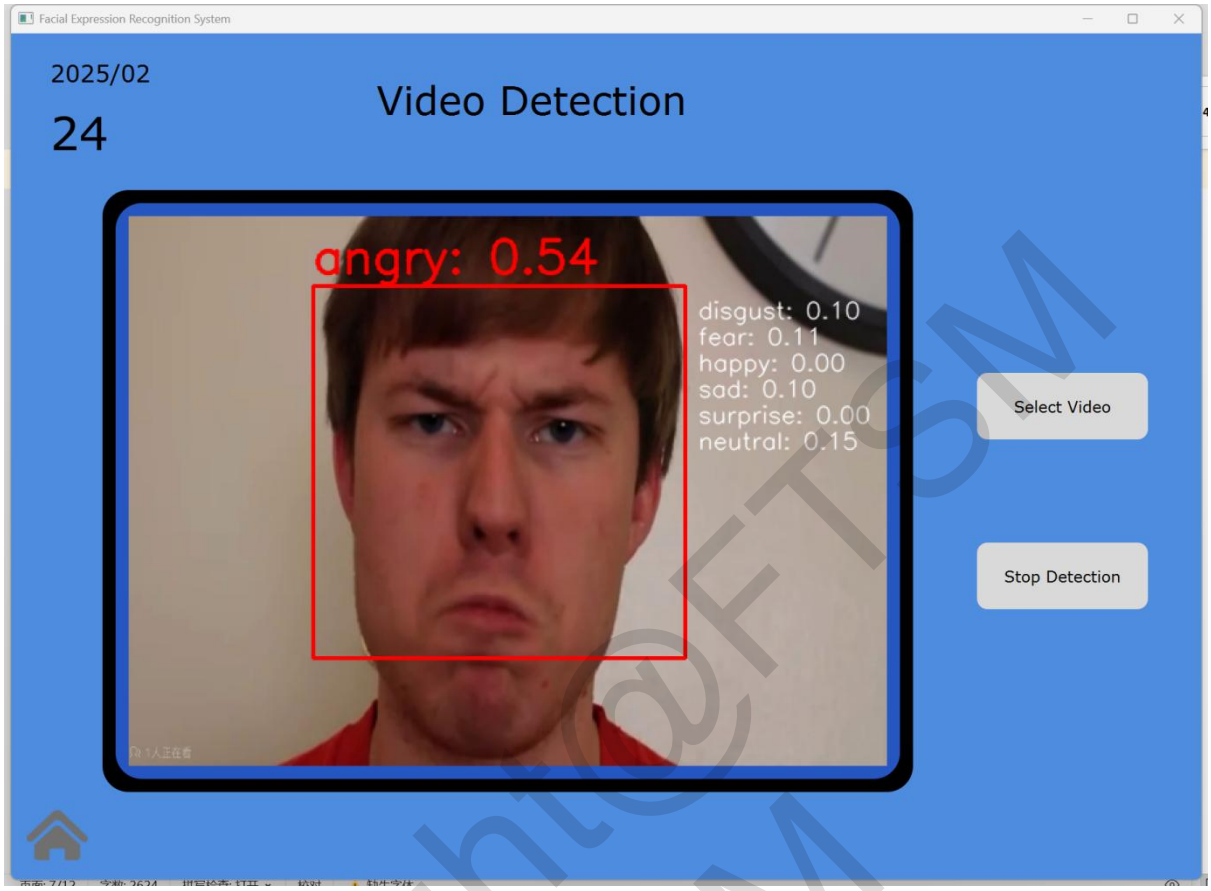

Figure 1.7 Video Detection

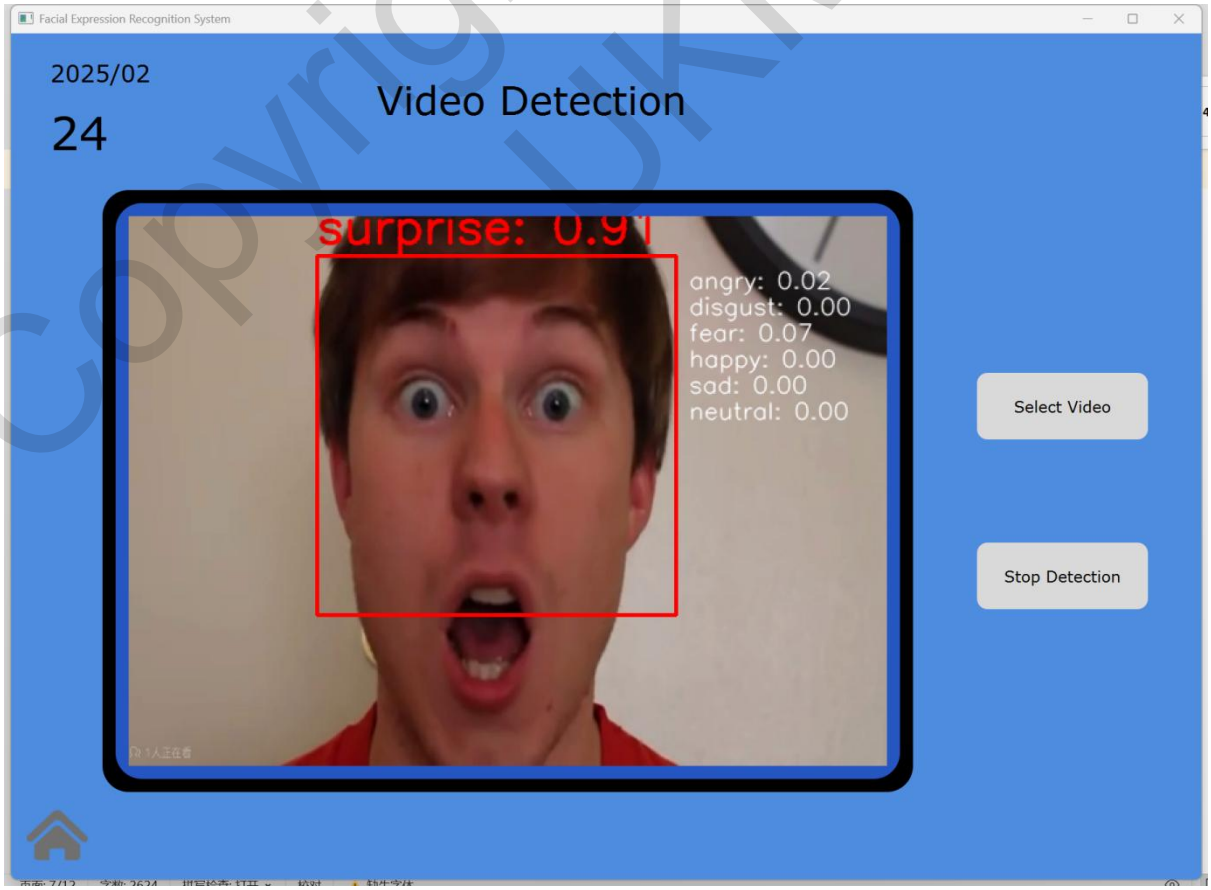

Figure 1.8 Video Detection

Video recognition in this system is a feature not found in common FER systems. Video recognition is a feature that needs to recognize frequent changes in the emotions of faces. It captures subtle dynamic changes in emotions, which is not possible with static image analysis. This feature is predictive of potential violent conflicts in public places. Sharply capturing those hidden risk signals allows security personnel to gain insight and intervene in a timely manner to nip potential threats in the bud and build a stronger line of defense for public safety.

Three different emotions of the same person are reflected in Figure 1.6, 1.7, and 1.8. The person's emotions change three times in just five seconds. Without the video emotion recognition feature, it would be difficult for the user to capture such rapid changes.

Usability Testing

Usability testing is a process that involves final testing by a representative group of users and interested parties to ensure that the FER system provides the required functionality before it is released to the public. The purpose of usability testing is to evaluate the usability of the system, collect quantitative data and assess user satisfaction. Up to this point, the demo version and demo video of this project were sent to 27 UKM students in order to guarantee functionality. In total, 27 questionnaire reports were received for this project. We believe that these questionnaire reports must be able to reflect some potential problems of the FER system.

The questionnaire provided eight questions to the users, and in this report I have shown only the important and useful graphical charts.

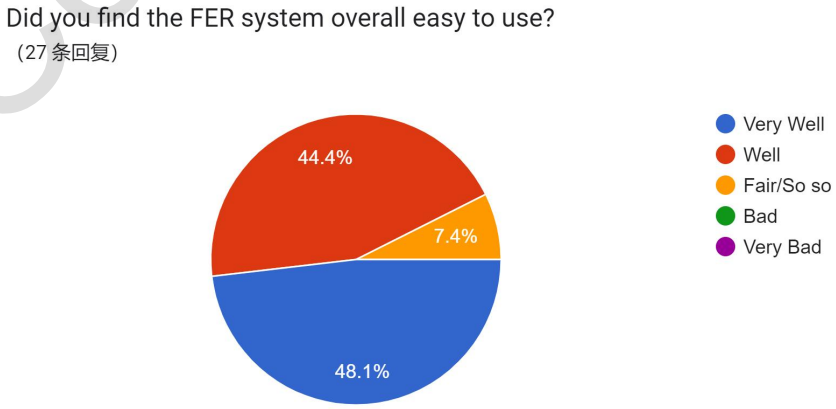

Figure 1.9 Survey

Are you satisfied with the accuracy of the system when recognizing emojis?
（27 条回复）

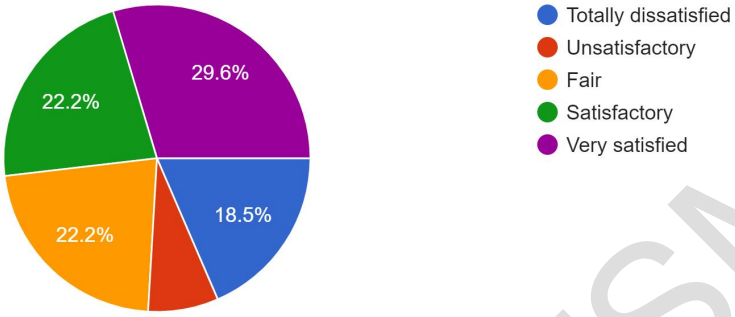

Figure 1.10 Survey

Do you think the functionality of the FER system useful?
（27 条回复）

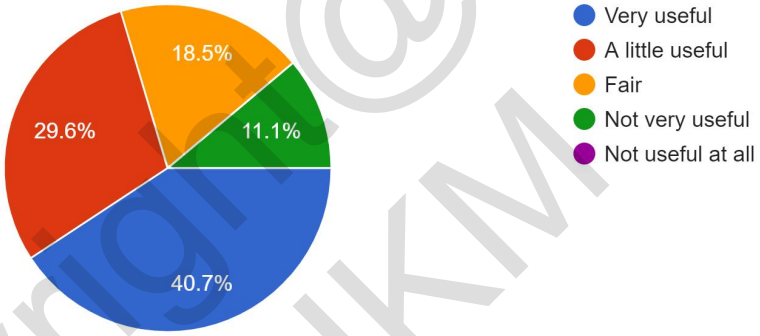

Figure 1.11 Survey

What do you think the FER system could do better? (Multiple choices allowed)
（27 条回复）

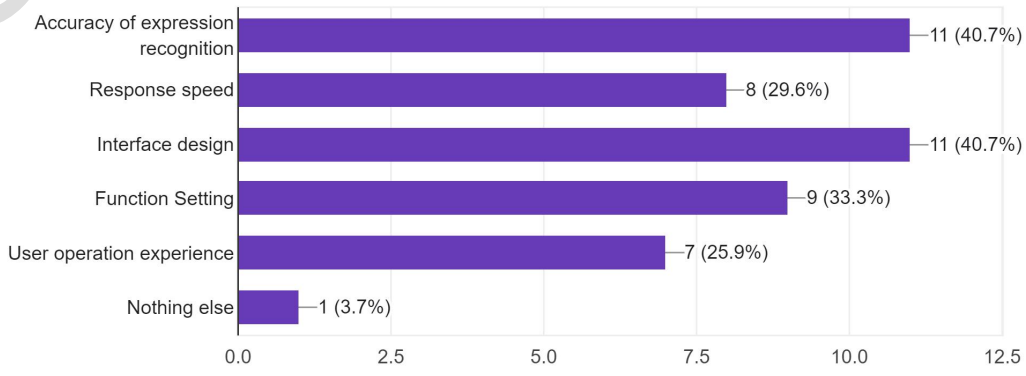

Figure 1.12 Survey

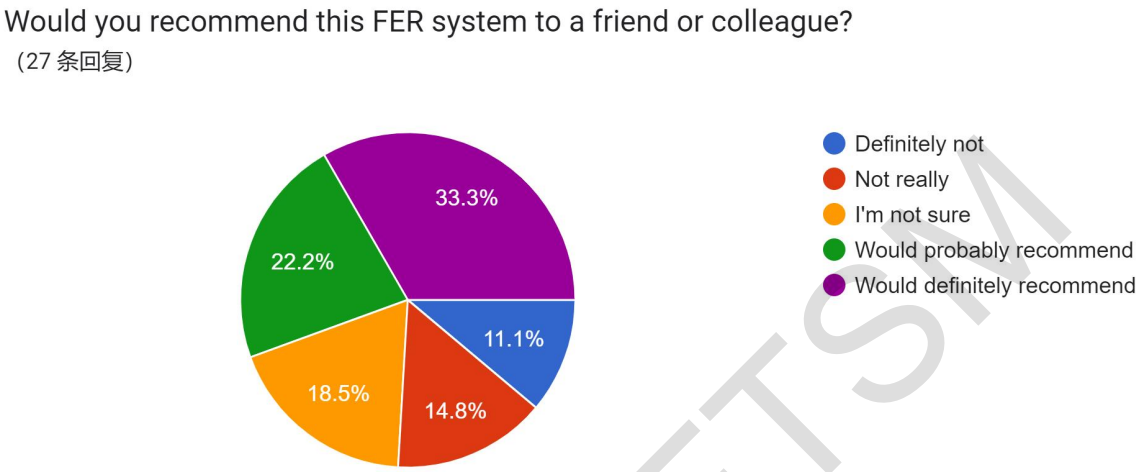

Figure 1.13 Survey

Although the results of the questionnaire survey show that the functionality and overall performance of the FER (Facial Expression Recognition) system largely meets the general needs of the public and has gained wide acceptance at the user level, a key contradiction is worth pondering over: the willingness of users to recommend the FER system to their close friends and relatives is surprisingly low. This phenomenon is like the tip of the iceberg, revealing the deep limitations of the FER system in practical application scenarios - it may be able to fulfill specific tasks better, but it still has obvious deficiencies in comprehensively covering daily life and providing a seamless emotional interaction experience. This means that the FER system still has a lot of room for improvement in terms of “versatility” and “scene adaptability”, and it needs to further expand its functional boundaries in order to truly integrate into users' daily lives and become an indispensable intelligent assistant.

On the other hand, users' high praise for the FER system's user interface (UI) confirms the correctness of the original system design principles. The simple, intuitive and easy-to-understand UI design has successfully lowered the threshold for users, making the FER system acceptable and usable by a wider range of people. Behind this, it reflects the pro-people nature of the design concept of the FER system. the FER system is essentially a deep learning application based on the convolutional neural network (CNN) model, and its internal operation mechanism is undoubtedly complex for ordinary users. However, the excellent UI design successfully hides this complexity so that users can easily get started and enjoy the convenience of the technology without having any specialized IT technical background.

Taken together, the results of the questionnaire survey present a complex and interesting picture: the FER system performs well in specific areas and has a friendly user interface, but it still needs to be improved in a wider range of application scenarios.

Improvement Reserve

In order for the FER (Facial Expression Recognition) system to truly move from a demo version to a practical one, and to show robustness in real-world complex environments, future improvements need to be centered around three core goals:

First, and most critically, is to dramatically improve the accuracy and generalization ability of the model. Existing FER systems usually achieve good recognition results under ideal conditions, such as well-lit, frontal faces. However, real-life scenarios are far more complex and variable than laboratory environments. People may be wearing masks, glasses, or looking sideways or down, or even have their faces partially obscured. These "imperfect" situations pose serious challenges to the recognition of existing FER systems. Therefore, future FER systems must have a strong "anti-interference" ability, and be able to maintain high accuracy in recognizing emotions under various complex and non-ideal conditions. This requires researchers to explore more advanced algorithmic models, such as the introduction of more powerful deep learning architecture, the use of transfer learning techniques, the use of adversarial training strategies, etc., so that the FER system in the face of occlusion, changes in illumination, gesture differences, etc., but still be able to "eyes of fire", accurately recognize the emotions of the face.

Secondly, enriching the daily application functions of the system is the key to the popularization of FER systems. At present, most FER systems are mainly focused on recognizing several basic emotions, such as joy, anger, sadness, happiness, surprise, fear, disgust and so on. However, the human emotional world is much more complex than these seven basic emotions. In our daily lives, we also experience more subtle emotions such as anxiety, frustration, doubt, anticipation, embarrassment, and guilt. Future FER systems will need to be able to capture and recognize these more subtle and richer categories of emotions to better understand and respond to user needs. In addition, FER systems should also expand their application areas, which should not be limited to academic research or security monitoring, but should go deeper into all aspects of daily life, such as smart home, online education, healthcare, customer service, etc., to provide people with more personalized and intelligent services.

Finally, deepening the interpretation of face information is a sure way for FER systems to realize a qualitative leap. In addition to recognizing basic emotions, future FER systems should also have the ability to mine deeper states from face information. For example, by analyzing the subtle changes in facial muscles and the color texture of the skin, it can determine whether a person is in a state of fatigue, nervousness, anxiety, etc., or even whether he or she is suffering from the early signs of certain diseases. This kind of deep-

level information interpretation will greatly expand the scope of application of FER systems, which will no longer just detect the emotions of the face, but also check the state of the face.

All in all, we hope that the future FER system will be a more powerful and practical emotion recognition partner. We believe it will profoundly change the way we interact with machines and bring more convenience and safety to our lives.

CONCLUSION

This research project has successfully developed a comprehensive facial expression recognition (FER) system, whose core objective is to significantly enhance the human-computer interaction experience, especially in high-risk occupational domains such as medical emergencies, flight pilots, and security surveillance, which require strict emotional states. The FER system integrates three core functional modules: local picture recognition, video clip recognition, and real-time camera recognition, so that it can flexibly respond to and meet the diverse needs of users in different scenarios. During the system development process, we adopted a rigorous incremental development model to ensure that the system functions are gradually improved, stable and reliable, and ultimately presented to the user a simple and friendly interface, intuitive and convenient operation of the FER system.

Although we are pleased to see that the core functions of the system and the well-designed user interface have been widely recognized by the user community through the feedback from the user questionnaire, we also notice a phenomenon worth pondering: the users' willingness to recommend the system to others is relatively low. This feedback suggests that there is still much room for improvement in terms of the versatility, universality, and scene adaptability of the FER system in a wider range of daily application scenarios.

Therefore, in order to move the FER system from the DEMO version to real-world applications, future improvements will focus on the following key directions: first, we will strive to further improve the model's recognition accuracy and generalization ability under complex conditions, especially in dealing with challenging situations such as facial occlusion, non-positive face images, and illumination variations, to ensure that the system can still stably produce Second, we will continue to enrich and expand the daily application functions of the FER system, so that it is no longer limited to the basic emotion category recognition, but can go deeper into a wider range of life scenarios that are closer to the actual needs of the users; finally, we will further deepen the ability to interpret facial information, and strive to extract deeper state information from facial features, such as early signs of disease, individual tension, etc. We will continue to develop the FER system in order to improve its ability to recognize emotions. early signs of disease, individual tension level, etc., so as to provide users with more valuable sentiment analysis services.

AWARDS

First of all, I would like to express my heartfelt gratitude to Prof. Dr. Nor Samsiah, the supervisor of the author of this study. Prof. Dr. Nor Samsiah is very knowledgeable and rigorous, and has given the authors meticulous care and guidance throughout the project. From the selection of the project topic and the construction of the research framework, to the selection of the research methodology and the refinement of the experimental design, to the analysis of the data and the writing of the thesis, every step of the process has been cohesive with the wisdom and hard work of the supervisor. Prof. Noor Samsia not only provided valuable academic advice, but also deeply influenced the authors with his rigorous attitude and persistent pursuit of academics, and became the guiding light on the authors' academic path. It is with the mentor's careful guidance and encouragement that this research project was successfully completed and this research report was finally presented.

It should also be emphasized that the authors would like to once again extend their sincerest gratitude to all those who have helped in this project. Your contributions will be remembered by the authors and will be used as a motivation to continue to work hard and move forward on the academic path in the future.

REFERENCE

Li, Q., cui, T., zhao, Y., & Wu, Y. (2023). Facial Recognition Technology: A comprehensive overview. Academic Journal of Computing & Information Science, 6(7), 15-26. DOI: 10.25236/AJCIS.2023.060703

Khan, M. A., Nanda, A. K., Almusallam, M. F., Rehman, M. U., Khan, M. A., Rehman, N. U., & Alshehri, A. D. (2023). A comprehensive survey on deep facial expression recognition: challenges, applications, and future guidelines. Alexandria Engineering Journal, 68, 817-840. https://doi.org/10.1016/j.aej.2023.01.017

Kopalidis, T., Solachidis, V., Vretos, N., & Daras, P. (2024). Advances in Facial Expression Recognition: A survey of 'Methods, Benchmarks, Models, and Datasets. Information, 15, 135.

Xiang, W. (n.d.). The Research and Analysis of Different Face Recognition Algorithms. WLSA Shanghai Academy, Shanghai, China.

Mehta, D., Siddiqui, M. F. H., & Javaid, A. Y. (n.d.). Facial Emotion Recognition: A survey and Real-world user Experiences in Mixed Reality.

Airport industry - News. (n.d.). What Are the Advantages of 'Automated passport control kiosks? Airport Industry News. Retrieved from https://airportindustry-news.com/what-are-the-advantages-of-automated-passport-control-kiosks/

James, W. (1884). What is an Emotion? Mind, 9(34), 188-2015.

Maithri, M., Raghavendra, U., Gudigar, A., et al. (2022). Automated Emotion Recognition: Current Trends and Future perspectives. Computational Methods and Programs in Biomedicine, 215, 106646.

PaddlePedia. (n.d.). Separable convolution. PaddlePedia Tutorials. Retrieved from https://paddlepedia.readthedocs.io/en/latest/tutorials/CNN/convolution_operator/separable_convolution.html

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 770-778. DOI: 10.1109/CVPR.2016.90

Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal covariate shift. arXiv preprint arXiv:1502.03167.

OpenCV Documentation. (n.d.). GaussianBlur. Retrieved from https://docs.opencv.org/3.4/d4/d13/tutorial_py_filtering.html

Acusio Bivona. (2021). Precision and Recall: What Are the Differences? Medium. Retrieved from https://medium.com/swlh/precision-and-recall-what-are-the-differences-bdd862d75e9

Zagoruyko, S., & Komodakis, N. (2020). Improved residual networks for image and video recognition. arXiv preprint arXiv:2004.04989. https://arxiv.org/abs/2004.04989

Qiu Hongbo (A184843)
Prof. Dr. Nor Samsiah
Fakulti Teknologi & Sains Maklumat
Universiti Kebangsaan Malaysia