Mathematical Sciences and Informatics Journal
Vol. 4, No. 1, May 2023, pp. 42-48
http://www.mijuitmjournal.com DOI : 10.24191/mij.v4i1.20625

Detection of Malaysian Sign Language with Single Shot Detector Algorithm

Nurfarah Idayu Mohamad Fauzi
College of Computer, Informatics and Media, Universiti Teknologi MARA, Perak Branch Tapah Campus, Perak, Malaysia
2019685818@student.uitm.edu.my

Shahirah Mohamed Hatim
College of Computer, Informatics and Media, Universiti Teknologi MARA, Perak Branch Tapah Campus, Perak, Malaysia
shahirah88@uitm.edu.my

Zalikha Zulkifli
College of Computer, Informatics and Media, Universiti Teknologi MARA, Perak Branch Tapah Campus, Perak, Malaysia
zalik764@uitm.edu.my

Article Info

Article history:

Received Feb 16, 2023
Revised Apr 01, 2023
Accepted Apr 24, 2023

Keywords:

Malaysian Sign Language
Hearing impairment
Image processing
Single Shot Detector
TensorFlow Lite

ABSTRACT

Sign language is a nonverbal communication that relies on facial expressions, postures, and gestures to facilitate communication between individuals who are deaf or have a hearing impairment. Despite its importance, many people do not recognize or understand sign language, which creates communication barriers for individuals with disabilities. Malaysia faces a lack of interest among its people in learning sign language, which may be attributed to various factors such as limited awareness, resource constraints, or a perception that sign language is not relevant or necessary. To address this issue, the research introduced a simple mobile application for that could potentially increase interest and awareness in sign language and promote greater inclusivity for individuals with disabilities. Single Shot Detector (SSD) algorithm was implemented to perform the Malaysian Sign Language object detection in the application. To facilitate the training of a custom TensorFlow Lite model, the project leveraged the TensorFlow eLite Model Maker library. The outcome of the research indicated a detection accuracy of 75.2%, which is significant as it demonstrates the potential for the developed model to serve as an effective Malaysia Sign Language detector. The framework used in this project can serve as a useful reference for future developers seeking to create similar custom models. Moreover, the promising results of the research indicate the potential for mobile applications utilizing the developed model to significantly enhance communication and inclusivity for individuals with hearing impairments in Malaysia.

Corresponding Author:
Shahirah Mohamed Hatim
Computing Sciences Studies, College of Computer, Informatics and Media, Universiti Teknologi MARA Perak Branch, Tapah Campus, Perak Malaysia.
shahirah88@uitm.edu.my

1. Introduction

Effective communication is crucial in our day-to-day lives, as it involves the exchange of information between a sender and a receiver. The successful transfer of the message from the sender to the receiver is essential to ensure mutual understanding. While most people can communicate effortlessly, individuals with hearing disabilities face significant challenges in a world designed primarily for those who can hear and speak. Disabilities such as profound deafness and mutism can lead to social exclusion and limit a person's ability to perform to their full potential. Empowering individuals with disabilities is crucial in promoting inclusivity in an unequal world.

PENERBIT PRESS
UNIVERSITI TEKNOLOGI MARA

Sign language is a widely utilized mode of communication by people with hearing disabilities to engage in daily interactions. Nevertheless, it is not mandatory for us to acquire knowledge about it, resulting in some individuals not even recognizing it. This creates significant difficulties for people with disabilities to communicate with others effectively[1]. Considering numerous countries have distinct sign languages, there are many sign languages in the world. The sign language was created by people who have the hearing ailment due to their native culture and it was strongly inspired and translated from the spoken language [2]. Sign language consists of two parts, namely hand posture and gesture. Hand posture is conveyed by the position and arrangement of the fingers, while gesture is expressed by the movement trajectory of the hand [3]. Moreover, sign language has five fundamental parameters, including hand form, hand orientation, movement, placement, and emotion, along with non-verbal cues [4]. To accurately convey a sign word, all of these parameters must be executed correctly.

Based on statistics, it is estimated that there are over 9 billion people worldwide who are hearing-impaired, with around 2.8 million individuals having this disability in Malaysia alone [5]. However, if people continue to overlook or underestimate the importance of sign language as a communication medium, it can lead to misunderstandings and difficulties in communication. For those who are not familiar with the language, it can be a significant barrier. This can ultimately result in a loss of social connections for individuals with special needs [6]. The contribution of this paper is the framework for developing detector using the Single Shot Detector (SSD) algorithm and TensorFlow Lite Model Maker library, which can serve as a communication tool for individuals with hearing impairments and bridge the gap between them and the hearing society.

2. Literature Review

Sign language is one of the ways to communicate. It uses hand gestures to deliver meaningful messages. It entails the simultaneous use of hand forms, hand orientation and movement, arm or body movement and face emotions. It is the most prevalent means of communication for those who are deaf or having hearing impairment [7]. In addition, Sign language is a highly organized nonverbal language that use both manual and non-manual communication methods. Manual communication consists of hand or arm gestures and orientation expressing symbolic meaning, whereas non-manual communication consists primarily of facial expression, head movement, body posture and orientation, which assist to amplify the meaning of the manual signals. Furthermore, sign language is made up of static signs, which are primarily alphabets and dynamic signals, which include certain gestures [8].

As referred to the World Health Organization (WHO), more than 5% of the global population or 430 million people require rehabilitation to address their "disabling" hearing loss. Over 700 million people or one out of every 10 people are expected to suffer debilitating hearing loss by 2050 [9]. Therefore, hearing loss is one of the most prevalent impairments that can have a negative impact on people's quality of life. It can lead to social isolation, decreased opportunities for education and employment, and even depression. Despite the availability of modern technologies and hearing aids, the cost of these devices can be high, and they are not always accessible to everyone, especially in low- and middle-income countries. Therefore, developing low-cost and reliable systems to assist people with hearing loss, such as sign language recognition systems, can have a significant impact on their quality of life. Sign languages around the world have their own distinct characteristics, varying in vocabulary, grammar, and syntax. Like spoken languages, sign languages also have their unique linguistic features shaped by the cultural and historical context of their users. In Malaysia, the Malaysian Sign Language (Bahasa Isyarat Malaysia or BIM) is the primary sign language used by the deaf community, with its own set of distinct features and signs developed over time [10].

Digital image processing plays a crucial role in sign language detection as it involves the analysis and interpretation of visual information contained in sign language gestures. Image processing techniques are used to extract features and patterns from sign language images or videos, which are then used to train machine learning models for sign language detection. These models rely on computer vision algorithms to identify and recognize the hand shapes, movements, and facial expressions that are characteristic of different sign language gestures. Digital image processing can be classified into four categories which are image enhancement, image restoration, image analysis and image compression. There are several Image processing techniques including Principal Component Analysis (PCA) [11], Convolutional Neural Network (CNN) [12] You Only Look Once (YOLO) [13], and Single Shot Detector (SSD) [14],[15]. This research employed the SSD technique to bridge the gap in image processing analysis of Malaysian Sign Language. SSD is a real-time object detection algorithm that capable of detecting objects and classifying them in a single forward pass of a neural network, making it efficient and fast for real-time processing of images and video streams.

3. Methodology

This section explains on the methodology of this research in order to develop the detection system. The descriptions of processing the images and the system architecture are provided. The architecture is elaborated in detailed, where all the steps of development are illustrated.

3.1 Image Processing

In this research, the SSD approach was utilized, which is a widely-used deep learning architecture for object detection tasks. The SSD approach is characterized by its ability to perform object detection and classification simultaneously using a single neural network, which makes it a popular choice for real-time applications. The approach achieves high accuracy by utilizing multiple convolutional layers to extract features at various scales and resolutions, allowing it to identify objects of different shapes and sizes. A fixed-size collection of bounding boxes and scores are generated, followed by a non-maximum suppression phase to provide the final detection process. The base network, which forms the first layers of the network, is based on a standard architecture used for high-quality image categorization.

To streamline the training process, the researchers utilized the TensorFlow Lite library with the EfficientDet-Lite module to extract feature maps. The model was trained over 20 epochs, meaning that it processed the entire training dataset 20 times. During training, the validation accuracy was monitored and the training was stopped once the validation loss stopped decreasing to prevent overfitting. The batch size used was 16. Once the model was trained, it was exported for deployment in the mobile application developed in Android Studio.

3.2 System Architecture

The system architecture for real-time object detection using SSD is presented in Figure 1. The architecture comprises two models: the training model and the detection model. The training model is composed of a labelled and trained dataset, while the detection model is used to test the detection of Malaysian Sign Language. The dataset was obtained from Kaggle and labelled using *labelImg*, which is a graphical image annotation tool that allows users to label object bounding boxes in images. It provides an intuitive interface that enables users to draw bounding boxes around objects of interest and assign labels to them. This tool is commonly used in computer vision tasks such as object detection, segmentation, and tracking. It is an open-source tool and is available for free on various platforms, including Windows, Linux, and macOS.

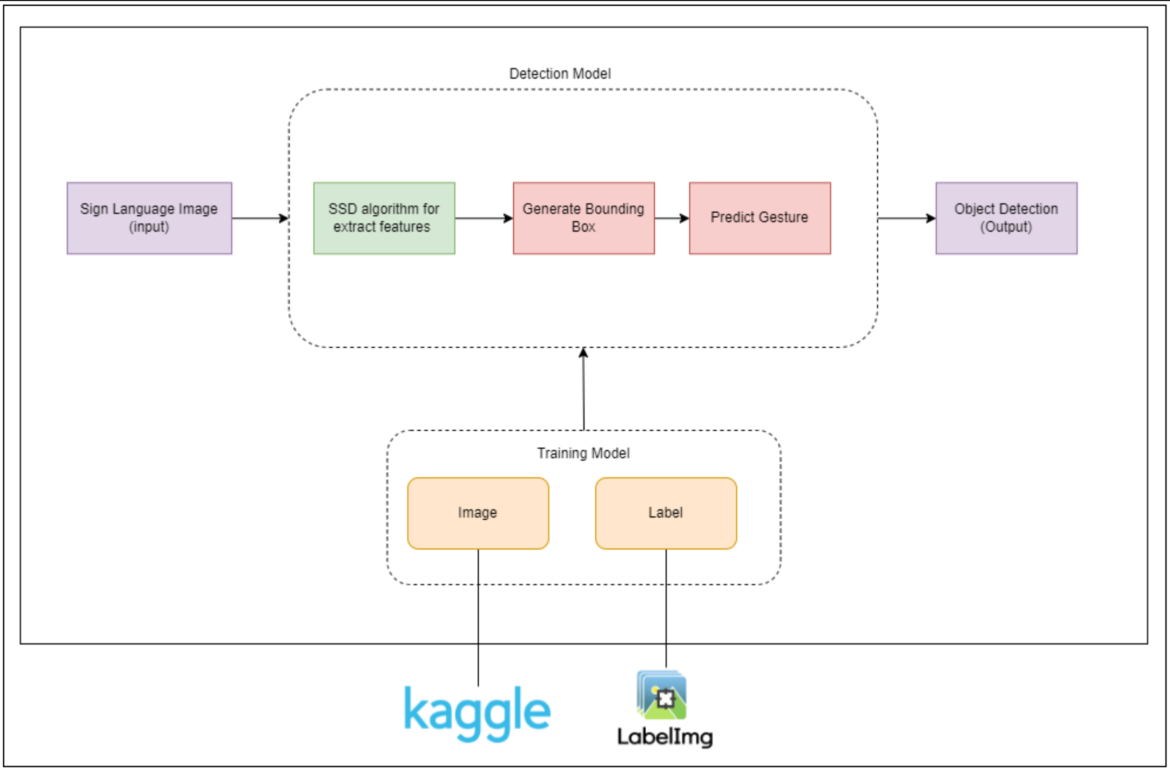

Figure 1. System architecture of the Malaysian Sign Language detection

The system starts by receiving real-time input of Malaysian Sign Language images. Then, it enters the training phase, during which the system uses the SSD algorithm and EfficientDet-Lite0 SSD model to extract features from the input image. Once the model has been trained, the Sign Language Detection stage can be initiated. In order to recognize the object, the TensorFlow Object Detection API is employed. The extracted features from the captured photos are sent to the TensorFlow module, which compares them to the real-time images in the frame. If any of these features are detected, a bounding box is generated around the gesture, and a prediction is made. It is crucial to correctly label the image to ensure an accurate prediction. Thus, if the application recognizes the sign language, the meaning and accuracy of the sign language will be displayed.

4. Results and Discussion

To assess the precision of the sign language recognition system, accuracy tests were conducted on five different classes of Malaysian Sign Language gestures from A to E. The evaluation metric used to assess the performance of the Malaysian Sign Language detection system in this research was the Average Precision (AP) of the Common Objects in Context (COCO) metric. AP is a popular evaluation metric used in object detection tasks. It measures the precision of an algorithm in detecting objects of interest within an image, as well as the accuracy of the predicted bounding boxes around those objects. AP is calculated by computing the area under the precision-recall curve for each object class. A higher AP value indicates better detection and classification performance. As depicted in Figure 1, AP is calculated across all categories (A-E objects). In this case, the system achieved an AP of 0.73 for class A, 0.75 for class B, 0.72 for class C, 0.75 for class D, and 0.80 for class E. This suggests that the system performs relatively well in detecting and recognizing sign language gestures, especially for class E, which achieved the highest AP value.

The Average Precision with Interpolation (APl) is a variation of the Average Precision (AP) metric, which considers the precision values for multiple recall levels. It is a more conservative estimate of the model's performance as it uses interpolated precision values at various recall thresholds. In this research, the APl for the five sign language classes ranged from 0.72 to 0.80, with an overall APl of 0.756. This indicates that the model has a relatively high precision across all classes and can accurately detect and recognize Malaysian Sign Language gestures in real-time.

```
3/3 [==============================] - 19s 3s/step

{'AP': 0.7523802,
 'AP50': 1.0,
 'AP75': 0.969436,
 'AP_/A': 0.7363927,
 'AP_/B': 0.7588859,
 'AP_/C': 0.7200163,
 'AP_/D': 0.74517655,
 'AP_/E': 0.80142975,
 'APl': 0.7566435,
 'APm': 0.726973,
 'APs': -1.0,
 'ARl': 0.81615347,
 'ARm': 0.77127975,
 'ARmax1': 0.77933335,
 'ARmax10': 0.8146667,
 'ARmax100': 0.8146667,
 'ARs': -1.0}
```

Figure 2. OCOC evaluation result

Furthermore, Mean Average Precision (mAP) is a metric that measures the overall performance of an object detection algorithm. As seen in Figure 2, the mAP of the Malaysian Sign Language detection system was found to be 0.752, indicating a reasonably high level of precision in detecting and recognizing sign language gestures. The interfaces of mobile application for recognizing Malaysian Sign Language on objects A, B, and C using SSD Real-Time Object Detection are shown in Figure 3.

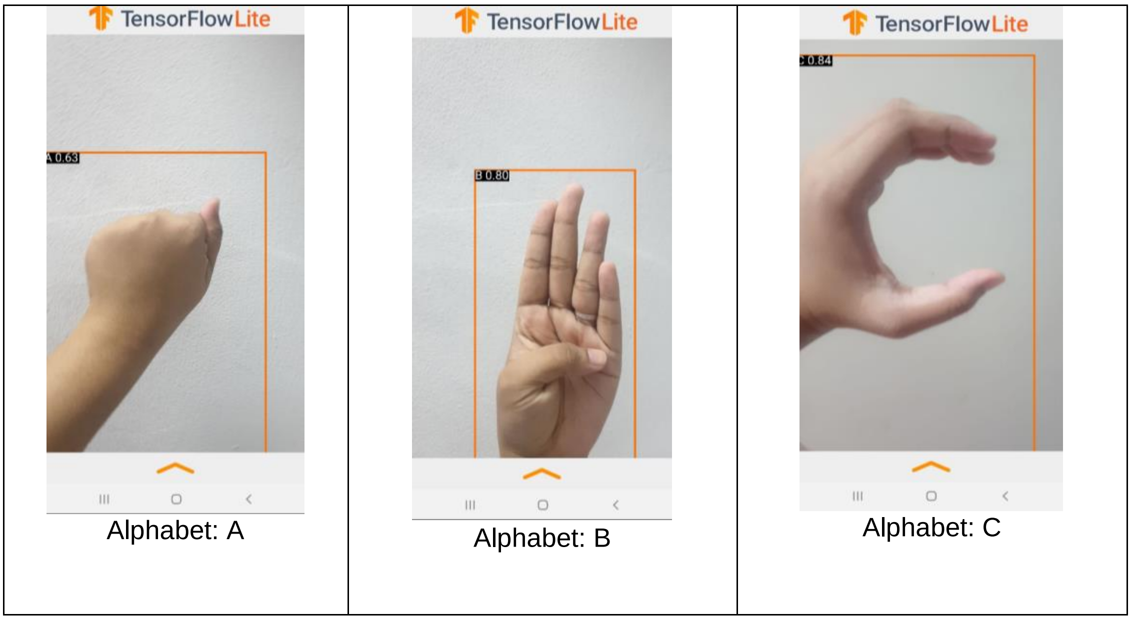

Figure 3. Application Interface

5. Conclusion

The purpose of developing the Malaysian Sign Language detector is to translate sign languages into a comprehensible format for users. This will not only benefit individuals with hearing impairments but also those who are unfamiliar with sign language to communicate seamlessly. The application has achieved a commendable detection accuracy rate of 75.2% in detecting sign language. Additionally, as the application can identify gestures, it has enormous potential for labeling gestures in any language, facilitating communication between users irrespective of language barriers. The implication of the application is that it has the potential to bridge communication gaps and promote inclusivity in society. By facilitating communication between individuals who are deaf or hard of hearing and those who are not familiar with sign language, the application can contribute to creating a more accessible and inclusive environment. Additionally, the gesture recognition feature has the potential to be useful in various industries, such as education, healthcare, and customer service, where language barriers often impede effective communication. Overall, the development of such applications can have a positive impact on society by promoting communication and inclusivity.

Acknowledgements

The authors express their gratitude to the Universiti Teknologi MARA (UiTM), Perak Branch for their support and recognition of this research.

Conflict of Interest

The authors declare no conflict of interest in the subject matter or materials discussed in this manuscript.

References

[1] P. R. Sanmitra, V. V. S. Sowmya, and K. Lalithanjana, "Machine Learning Based Real Time Sign Language Detection," vol. 4, no. 6, pp. 137–141, 2021.

[2] M. A. M. M. Asri, Z. Ahmad, I. A. Mohtar, and S. Ibrahim, "A real time detection algorithm based on YOLOv3," Int. J. Recent Technol. Eng., vol. 8, no. 2 Special Issue 11, pp. 651–656, 2019, doi: 10.35940/ijrte.B1102.0982S1119.

[3] C. Yi, L. Zhou, Z. Wang, Z. Sun, and C. Tan, "Long-range Hand Gesture Recognition with Joint SSD Network," 2018 IEEE Int. Conf. Robot. Biomimetics, ROBIO 2018, pp. 1959–1963, 2018, doi: 10.1109/ROBIO.2018.8665302.

[4] R. Rastgoo, K. Kiani, and S. Escalera, "Sign Language Recognition: A Deep Survey," Expert Syst. Appl., vol. 164, no. August 2020, p. 113794, 2021, doi: 10.1016/j.eswa.2020.113794.

[5] A. Z. Shukor, M. F. Miskon, M. H. Jamaluddin, F. Bin Ali Ibrahim, M. F. Asyraf, and M. B. Bin Bahar, "A New Data Glove Approach for Detection," Procedia Comput. Sci., vol. 76, no. Iris, pp. 60–67, 2015, doi: 10.1016/j.procs.2015.12.276.

[6] R. Rastgoo, K. Kiani, and S. Escalera, "Hand sign language recognition using multi-view hand skeleton," Expert Syst. Appl., vol. 150, p. 113336, 2020, doi: 10.1016/j.eswa.2020.113336.

[7] N. Saquib and A. Rahman, "Application of machine learning techniques for real-time sign language detection using wearable sensors," MMSys 2020 - Proc. 2020 Multimed. Syst. Conf., pp. 178–189, 2020, doi: 10.1145/3339825.3391869.

[8] R. Akmeliawati et al., "Assistive technology for relieving communication lumber between hearing/speech impaired and hearing people," J. Eng., vol. 2014, no. 6, pp. 312–323, 2014, doi: 10.1049/joe.2014.0039.

[9] K. Neumann, S. Chadha, G. Tavartkiladze, X. Bu, and K. R. White, "Newborn and infant hearing screening facing globally growing numbers of people suffering from disabling hearing loss," Int. J. Neonatal Screen., vol. 5, no. 1, p. 7, 2019.

[10] D. Kamarudin, D. Kamarudin, and Y. Hussain, "Hearing impaired student achievement on the bahasa melayu subject: are these tests applicable?," J. Pendidik. Bitara UPSI, vol. 12, pp. 59–67, 2019.

[11] J. Ma and Y. Yuan, "Dimension reduction of image deep feature using PCA," J. Vis. Commun. Image Represent., vol. 63, p. 102578, 2019.

[12] R. K. Karunanayake, W. G. M. Dananjaya, M. S. Y. Peiris, B. Gunatileka, S. Lokuliyana, and A. Kuruppu, "CURETO: skin diseases detection using image processing and CNN," in 2020 14th international conference on Innovations in Information Technology (IIT), 2020, pp. 1–6.
[13] T. Diwan, G. Anirudh, and J. V Tembhurne, "Object detection using YOLO: Challenges, architectural successors, datasets and applications," Multimed. Tools Appl., pp. 1–33, 2022.
[14] A. Kumar, Z. J. Zhang, and H. Lyu, "Object detection in real time based on improved single shot multi-box detector algorithm," EURASIP J. Wirel. Commun. Netw., vol. 2020, no. 1, pp. 1–18, 2020.
[15] W. Zhu, H. Zhang, J. Eastwood, X. Qi, J. Jia, and Y. Cao, "Concrete crack detection using lightweight attention feature fusion single shot multibox detector," Knowledge-Based Syst., vol. 261, p. 110216, 2023.

Biography of all authors

Picture	Biography	Authorship contribution
	Nurfarah Idayu Mohamad Fauzi is a final year student in Bachelor of Computer Sciences at the College of Computing, Informatics and Media, Universiti Perak Branch, Tapah Campus, Perak, Malaysia. Currently, she is undergoing internship program for her degree at Grass2Route Sdn. Bhd.	Design the research work, data collection, data analysis and interpretation, drafting article.
	**Shahirah Mohamed Hatim** is currently an academia at the College of Computing, Informatics and Media, Universiti Perak Branch, Tapah Campus, Perak, Malaysia. Her expertise covers Artificial Intelligence, Image Processing and Internet of Vehicles. She can be reached at shahirah88@uitm.edu.my.	Design and supervising the research work and camera ready.
	**Zalikha Zulkifli** is currently an academia at the College of Computing, Informatics and Media, Universiti Teknologi MARA Perak Branch, Tapah Campus (UiTM Perak). She has experience in computer science, emphasis on Artificial Intelligence, machine learning and image processing. She can be contacted at zalik764@uitm.edu.my.	Literature review and camera ready.

Pejabat Perpustakaan
Librarian Office

Universiti Teknologi MARA
Cawangan Perak
Kampus Seri Iskandar
32610 Bandar Baru Seri Iskandar,
Perak Darul Ridzuan, MALAYSIA
Tel: (+605) 374 2093/2453 Faks: (+605) 374 2299

UNIVERSITI TEKNOLOGI MARA

Surat kami : 700-KPK (PRP.UP.1/20/1)
Tarikh : 20 Januari 2023

Prof. Madya Dr. Nur Hisham Ibrahim
Rektor
Universiti Teknologi MARA
Cawangan Perak

Tuan,

PERMOHONAN KELULUSAN MEMUAT NAIK PENERBITAN UiTM CAWANGAN PERAK MELALUI REPOSITORI INSTITUSI UiTM (IR)

Perkara di atas adalah dirujuk.

2. Adalah dimaklumkan bahawa pihak kami ingin memohon kelulusan tuan untuk mengimbas (*digitize*) dan memuat naik semua jenis penerbitan di bawah UiTM Cawangan Perak melalui Repositori Institusi UiTM, PTAR.

3. Tujuan permohonan ini adalah bagi membolehkan akses yang lebih meluas oleh pengguna perpustakaan terhadap semua maklumat yang terkandung di dalam penerbitan melalui laman Web PTAR UiTM Cawangan Perak.

Kelulusan daripada pihak tuan dalam perkara ini amat dihargai.

Sekian, terima kasih.

"BERKHIDMAT UNTUK NEGARA"

Saya yang menjalankan amanah,

SITI BASRIYAH SHAIK BAHARUDIN
Timbalan Ketua Pustakawan

nar

Setuju.

27.1.2023

PROF. MADYA DR. NUR HISHAM IBRAHIM
REKTOR
UNIVERSITI TEKNOLOGI MARA
CAWANGAN PERAK
KAMPUS SERI ISKANDAR