SEMESTER 1, SESSION 2025/2026
WQF7006 COMPUTER VISION AND IMAGE PROCESSING
GROUP ASSIGNMENT
CASE STUDY: MALAYSIAN SIGN LANGUAGE

BACKGROUND:

Sign language understanding has gained significant attention with the growing advancements in sign language translation technologies. This rise reflects a broader commitment to **inclusive communication** and highlights the vital role of **Artificial Intelligence (AI)** in enabling seamless interaction for the deaf and hard-of-hearing community. However, for **Malaysian Sign Language (MSL)**, translation research and applications remain in their early stages. Key challenges include **limited datasets**, **insufficient linguistic resources**, and **low-resource neural translation models**, which make developing an effective MSL recognition or translation system particularly complex.

PROJECT OVERVIEW:

In this project, each group will **simulate a multidisciplinary AI development team** within a **technology company**. Students will assume different stakeholder roles: for example and not limited to, **AI Engineer, Software Developer, UX/UI Designer, Project Manager, Accessibility & Ethics Officer,** and **Communication Officer**. Each team should collaborate to design a prototype or conceptual framework for an **AI-based Malaysian Sign Language System**.

The project encourages students to apply both technical knowledge and communication skills to design a socially impactful solution. The emphasis is on team collaboration, professional communication, and integration of diverse expertise within a simulated real-world environment.

REQUIREMENTS:

To address the above issues, you are required to complete following tasks with your group members.

1. **Team Roles**
   a. Assign specific roles to each team member. Each member will present their part of the group project based on their assigned role.
   b. Roles may include (but are not limited to): Project Manager, AI Engineer, Developer, UX/UI Designer, Accessibility & Ethics Officer, and Communication Officer.
2. **Data Collection**
   a. All students are required to participate in the **data collection session**, tentatively scheduled as follows:
      OCC 1: Tuesday, 11 November 2025, 6.00 p.m.
      OCC 2: Thursday, 13 November 2025, 6.00 p.m.
      OCC 3: Saturday, 15 November 2025, 11.00 a.m.
      (Schedule is subject to change and will be confirmed during class.)
   b. All activities in this group assignment are conducted **for academic purposes only**.
   c. During the session, you will record **sign language interpretations** under the instructor's guidance. Recording stations will be provided to store the videos.
   d. Refer to **Appendix I** for the list of glosses and sign language interpretations that will be taught and recorded during the session. The dataset collected will later be used to **train a deep learning model**.
   e. Ensure that **all group members** participate in the data collection process.
3. **Data pre-processing**

a. You are expected to store the sign language videos according to the glosses in respective folder. You may edit the videos if necessary. The instructor will also guide you on this.
b. Group leaders are expected to ensure the sign language videos for all group members are stored properly into a cloud drive, e.g., Google Drive, and provide the shareable link in the **Group Report**. Refer *Submission*.

4. **Deep learning implementation**
   a. You will be provided Python codes to train your own deep learning model as the baseline reference with the dataset that we have collected. Please find the Python code to train the MSL model via this link: https://colab.research.google.com/drive/1foQCf7TtO95Ost6piizMTq8Eo_7H2AII?usp=sharing
   b. For the dataset that we have collected from the Data Collection Session, you may find the compiled folder via this link: https://drive.google.com/drive/folders/1mt36VY2LZRMjpT89__bh_PGsJE3q4RRN?usp=drive_link . This dataset should consist of **90 folders** which is equivalent to 90 glosses.
   c. For the model training, you can discuss with your group member how to improve the baseline code or retain it for the subsequent steps. If your group has sufficient computing resources, you may try to train all glosses. Otherwise, each group should CHOOSE and TRAIN a**t least ten (10) glosses**. All groups must explain the model training settings in the report.
   d. After you have successfully trained the model, you must demonstrate that the proposed model can translate sign language interpretations when given a new input.
5. **Application Mock-up/Prototype or User Practicality Test**
   a. Extra marks will be awarded to groups that develop a **mock-up or prototype application (web-based or mobile-based)**.
   b. Additional marks will also be given for demonstrating or conducting the user **practicality and societal impact** of the application. This may include discussions on and not limited to:
      i. Target users
      ii. Revenue or sustainability model
      iii. Educational or social impact
      iv. Broader implications in accessibility and inclusion

RUBRIC (PRESENTATION 20% + REPORT 10%):

Criteria	5 – Excellent	4 – Good	3 – Satisfactory	2 – Needs Improvement	1 – Poor	Weight
Teamwork and Cooperation	Team collaborates seamlessly with clear communication, consistent participation, and mutual support. All members submit consent forms before the deadline.	Team works well overall, with minor coordination or communication issues. Most members submit consent forms on time.	Team completes tasks but shows uneven participation or occasional miscommunication. Some members submit consent forms before the deadline.	Team faces frequent communication or coordination problems that affect progress. Several members submit consent forms late.	Team shows ineffective collaboration and poor communication. Some members fail to submit consent forms.	**2%**
Data Pre-processing and Handling	Data is thoroughly cleaned, transformed, and organized using clear and consistent protocols. Excellent attention to detail.	Data is mostly cleaned and well-structured, with minor inconsistencies or documentation gaps.	Data is generally cleaned and organized, but includes noticeable gaps or unclear structure.	Data pre-processing is incomplete or inconsistent, showing limited understanding of requirements.	Data pre-processing is poorly done, with major issues in cleaning, labeling, or organization.	**5%**
Deep Learning Implementation	Model is fully implemented and demonstrates	Model is correctly implemented and functional with	Model is implemented and produces results, but lacks optimization	Model is partially implemented or not fully functional.	Model implementation is incomplete, non-	**10%**

SEMESTER 1, SESSION 2025/2026
WQF7006 COMPUTER VISION AND IMAGE PROCESSING
GROUP ASSIGNMENT
CASE STUDY: MALAYSIAN SIGN LANGUAGE

Criteria	5 – Excellent	4 – Good	3 – Satisfactory	2 – Needs Improvement	1 – Poor	Weight
	excellent understanding of deep learning concepts. Includes improvements or optimization beyond baseline, with clear performance explanation and results.	minor issues. Some improvement or analysis beyond baseline is shown.	or thorough explanation.	Limited testing or unclear demonstration.	functional, or lacks understanding of key concepts.	
Application Mock-up / Prototype	A creative, polished, and functional mock-up or prototype demonstrating clear usability, target users, and real-world impact. Practical aspects (impact, education, accessibility, etc.) are discussed in depth.	A well-designed mock-up showing good functionality and clear purpose. Some practical aspects discussed.	A basic but functional mock-up is provided, though lacks detail or user-focused discussion.	Mock-up is minimal or incomplete with limited explanation of practicality or usability.	No mock-up or prototype presented.	**3%**
Report and Documentation	Report is clear, complete, and well-organized. All sections (introduction, methodology, results, discussion, conclusion) are cohesive and well-justified. Excellent formatting and references.	Report is well-written and includes all main sections. Minor lapses in structure or depth. Documentation supports project effectively.	Report covers essential sections but lacks depth or clarity in some areas. Acceptable formatting and organization.	Report is incomplete or poorly structured with unclear explanations or missing parts.	Report is disorganized, lacks key content, or demonstrates minimal effort.	**10%**

PRESENTATION:
Group presentation is scheduled in **Week 12**. All members must attend for the presentation.

Group presentation sequence can be accessed via this link: https://365umedumy.sharepoint.com/:x:/s/WQF7006CVIP/IQBBPPNKD4fqTJtGRcqEWw9_AWNvigbFVq9h_8KYPWM6U44?e=puuCtG

GROUP SUBMISSION:
Kindly submit your presentation slides and report completing all requirements and tasks and following the rubrics mentioned above, by **Week 15, 1 February 2026, 12.00am. Late submission will not be entertained.**

SEMESTER 1, SESSION 2025/2026
WQF7006 COMPUTER VISION AND IMAGE PROCESSING
GROUP ASSIGNMENT
CASE STUDY: MALAYSIAN SIGN LANGUAGE

APPENDIX I

Refer this link for sign language video reference:
https://drive.google.com/drive/folders/1b44aycPedjL5ZYj2fWt2c7ZNEUdVpQ5k?usp=drive_link

Module	Gloss	OCC 1	OCC 1	OCC 2	OCC 3	OCC 3
1	hi	Everyone				
6	nasi	1	8			4
6	lemak	1	8			4
2	baik (II)	1	8			4
3 & 4	anak perempuan	1	8			4
3 & 4	bapa saudara	1	8			4
3 & 4	emak saudara	1	8			4
3 & 4	anak lelaki	1	8			4
1	assalamualaikum	1	8			4
9	marah (II)	1	8			4
6	nasi_lemak	1	8			4
1	apa_khabar	2	9			5
8	tolong_2	2	9			5
9	beli_2	2	9			5
8	mohon_2	2	9			5
6	pukul	2	9			5
10	mari	2	9			5
8	boleh	2	9			5
9	beli	2	9			5
8	sudah	2	9			5
8	tanya	2	9			5
9	marah	3	10			6
10	curi	3	10			6
10	kesakitan	3	10			6
5	ribut	3	10			6
6	kereta	3	10			6
8	ada	3	10			6
8	buang	3	10			6
9	main	3	10			6
10	pergi_2	3	10			6
8	dapat	3	10			6
8	minum	4	11			7
6	teh tarik	4	11			7
8	berlari	4	11			7
9	jangan	4	11			7
10	berjalan	4	11			7
6	bola	4	11			7
7	pen	4	11			7
8	dari	4	11			7
8	mohon	4	11			7
8	tolong	4	11			7
10	pergi	5		1		8
10	tidur	5		1		8
6	bas	5		1		8
7	sekolah	5		1		8
2	panas	5		1		8
2	pandai_2	5		1		8
3	emak	5		1		8
7	arah	5		1		8
9	buat	5		1		8
9	pinjam	5		1		8

SEMESTER 1, SESSION 2025/2026
WQF7006 COMPUTER VISION AND IMAGE PROCESSING
GROUP ASSIGNMENT
CASE STUDY: MALAYSIAN SIGN LANGUAGE

3	ayah	6		2		9
6	teksi	6		2		9
9	bawa	6		2		9
9	jumpa	6		2		9
9	lupa	6		2		9
10	suka	6		2		9
2	jahat	6		2		9
3	bomba	6		2		9
3	lelaki	6		2		9
6	masa	6		2		9
7	makan	7		3		10
9	kacau	7		3		10
9	sampai	7		3		10
10	ambil	7		3		10
10	baca	7		3		10
2	baik	7		3		10
2	sejuk	7		3		10
5	masalah	7		3		10
6	jam	7		3		10
6	pensil	7		3		10
7	kafetaria			4	1	11
7	keretapi			4	1	11
10	mari_2			4	1	11
3	polis			4	1	11
5	tandas			4	1	11
7	hospital			4	1	11
2	panas_2			4	1	11
2	perlahan_2			4	1	11
3	perempuan			4	1	11
7	jalan			4	1	11
10	belajar				2	12
5	payung				2	12
5	hari				2	12
6	esok				2	12
7	gambar				2	12
2	bila				2	12
2	mana				2	12
2	pandai				2	12
2	siapa				2	12
3	abang				2	12
7	kedai				3	13
2	perlahan				3	13
2	berapa				3	13
3	keluarga				3	13
2	apa				3	13
3	kakak				3	13
2	bagaimana				3	13
5	hujan				3	13
10	hilang_habis				3	13
4	saudara				3	13
1	bahasa isyarat			ALL OCC 2 x2		