

**Academic Activities**

|  |  |  |
| --- | --- | --- |
| **Student**  **Name** | **:** | **Mingchen Wang** |
| **Student No** | **:** | **2230025907** |
| **Program** | **:** | **Academic Activities** |
| **Supervisor** | **:** | **Liu Xin** |

**第三屆IEEE數字孿生和平行智能國際會議***Abstract*

In order to better help able-bodied people to communicate with hearing-impaired or speech-impaired people and build a barrier-free society. In this paper, an End-to-end Lip-Reading Recognition Architecture (ELRA) based on multimodal fusion is constructed to realize the Chinese lip-reading video translation function. Experimental results show that the proposed ELRA is applied to the CMLR dataset and achieves a character error rate of 8.0%. Compared with previous lip recognition models, it shows good performance in fusing image features and audio features.

*Keywords—End-to-end lip-reading recognition; Multi-modal fusion; Hearing-impaired*

1. Introduction

**未來Disease X的預測預警與國際聯邦學習體系構建**