

ДОДАТОК А
ГРАФІЧНИЙ МАТЕРІАЛ КВАЛІФІКАЦІЙНОЇ РОБОТИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
РАДІОЕЛЕКТРОНІКИ
КАФЕДРА ЕОМ

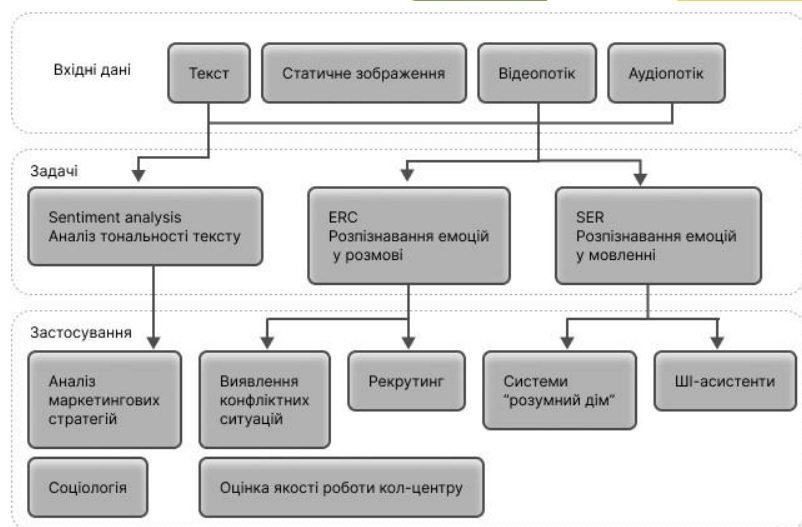
Кваліфікаційна робота
Другий рівень (магістр)

Методи визначення
емоційного стану спікера

Автор: Уваров Г.О., ст. гр. СПМ-22-4

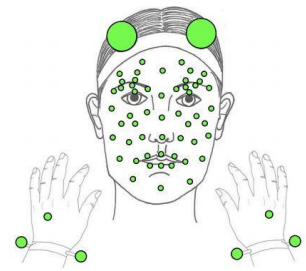
Керівник: Барковська О.Ю., доц. каф. ЕОМ

Огляд проблемної області

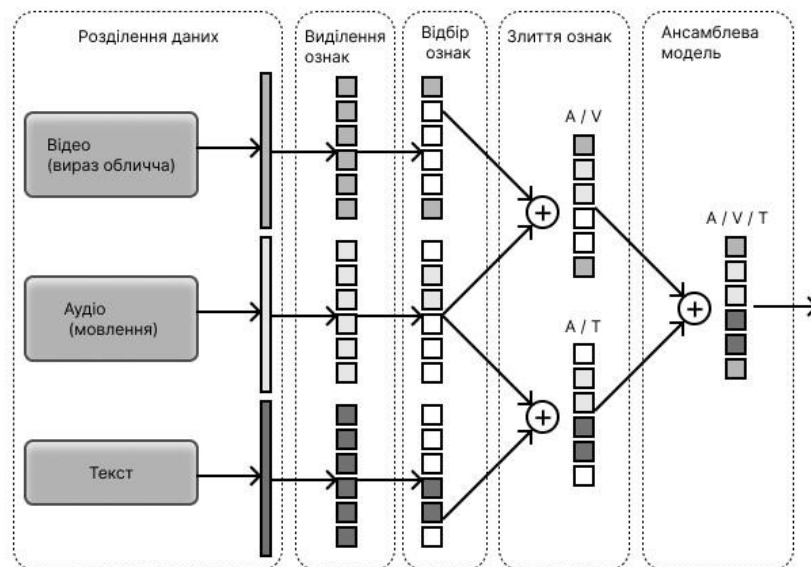


Датасет IEMOCAP

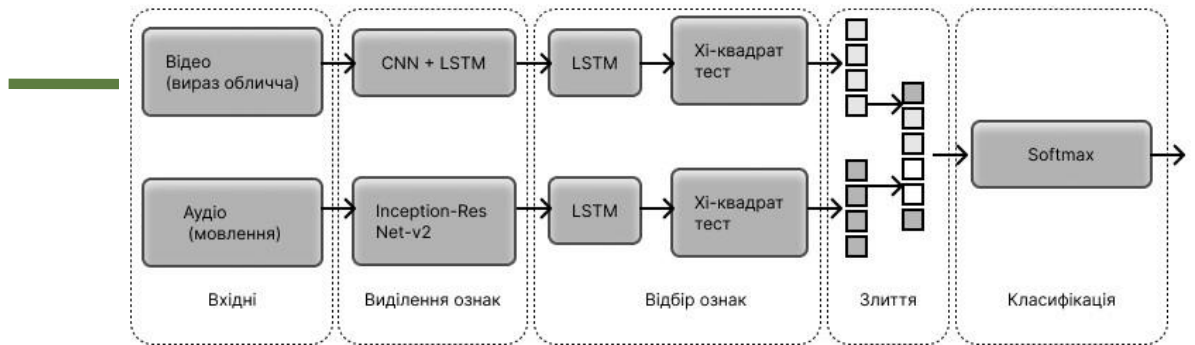
Датасет	Датапоінтів	Модальність	Кількість мовців	Кількість емоцій	мова
IEMOCAP	302	Відео	10	9	Анг.
CREMAD	7442	Аудіо	91	6	Анг.
RAVDESS	7356	Відео	24	7	Анг.
TESS	2800	Аудіо	2	7	Анг.
EMODB	535	Аудіо	10	7	Нім.



Запропонований метод тримодального аналізу емоцій мовця

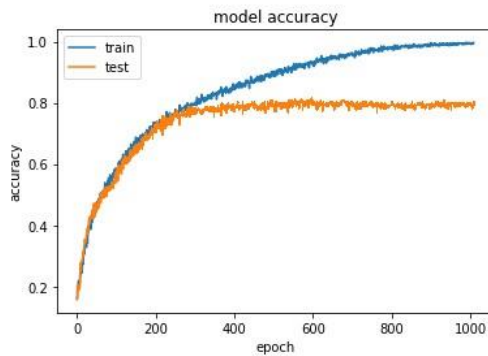


Метод злиття аудіо-відео



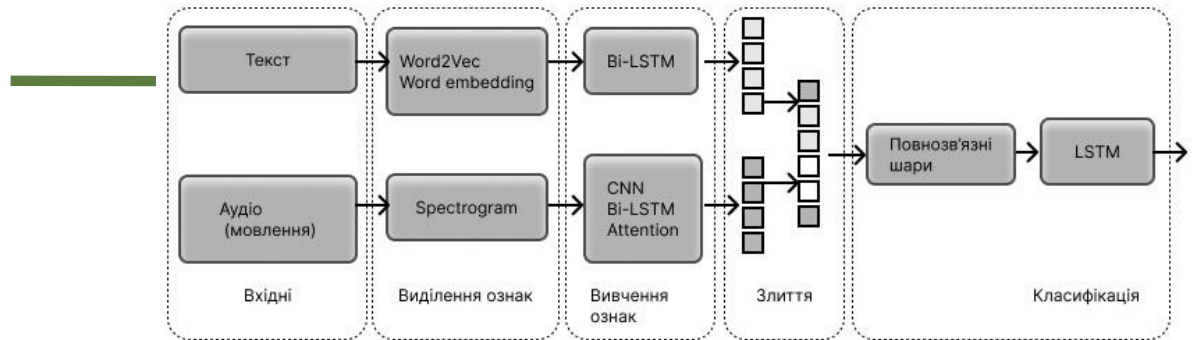
A/V	Злість	Щасття	Нейтр.	Сум
Злість	71.89	2.02	11.55	14.17
Щасття	4.73	77.92	2.74	14.29
Нейтр.	4.48	2.22	80.93	12.05
Сум	2.91	9.87	12.16	74.62

Метод злиття аудіо-відео



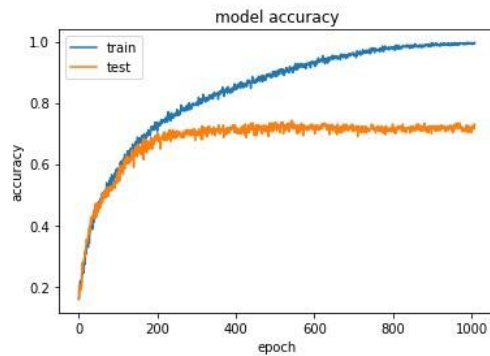
A/V	Злість	Щасття	Нейтр.	Сум
Злість	71.89	2.02	11.55	14.17
Щасття	4.73	77.92	2.74	14.29
Нейтр.	4.48	2.22	80.93	12.05
Сум	2.91	9.87	12.16	74.62

Метод злиття аудіо-текст



А/Т	Злість	Щасття	Нейтр	Сум
Злість	74.92	3.58	3.83	16.81
Щасття	1.92	73.25	2.08	22.03
Нейтр	2.81	2.51	80.85	13.39
Сум	2.41	9.62	6.68	80.82

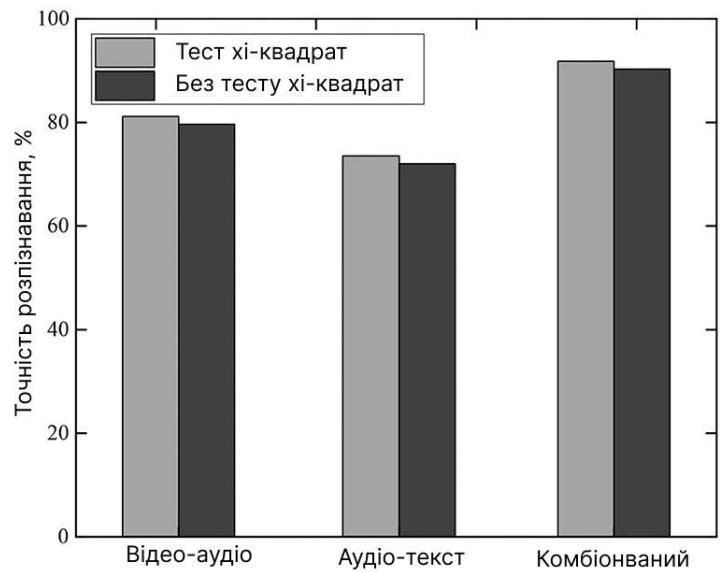
Метод злиття аудіо-текст



А/Т	Злість	Щасття	Нейтр.	Сум
Злість	74.92	3.58	3.83	16.81
Щасття	1.92	73.25	2.08	22.03
Нейтр.	2.81	2.51	80.85	13.39
Сум	2.41	9.62	6.68	80.82

Тримодальний метод

A/V/T	Злість	Щасття	Нейтр	Сум
Злість	79.62	1.26	2.1	16.51
Щасття	1.62	82.9	0.3	14.9
Нейтр	4.07	1.48	80.94	13.22
Сум	1.91	9.81	7.03	80.88



ВИСНОВКИ

Результати експериментів підтверджують, що система успішно виконує аналіз мультимодальних даних і класифікує виражені емоції з високою точністю.

Максимальна точність була досягнута комбінованим тримодальним методом з відбором значущих ознак тестом хі-квадрат, навченому на 100 епохах на даних з ІЕМОСАР. Показало тестову точність 82,9%.

Визначено, що аналіз аудіо та динамічно створеної транскрипції мовлення ефективно доповнює аналіз відео і підвищує точність у комбінованому методі. Використання мел-частотних коефіцієнтів, спектрограм мовлення та інших ознак витягнутих з аудіопотоку як вхідних даних дозволяє системам машинного навчання точніше виявляти емоційні стани спікера, аніж робота з одновимірним аудіопотоком.

ISSN 2710-0766
DOI 10.31891/CSIT

THE INTERNATIONAL SCIENTIFIC JOURNAL

**COMPUTER SYSTEMS
AND INFORMATION
TECHNOLOGIES**



No 1-2024

Стаття у фаховому виданні

Olesia BARKOVSKA, Oleksandr BILICHENKO, Heorhii UVAROV, Tymur MAKUSHENKO Improved rendering method of skeletal animation on control points base. *Computer systems and information technologies* 2024, (1), 71-81.

МІЖНАРОДНИЙ НАУКОВИЙ ЖУРНАЛ

**КОМП'ЮТЕРНІ СИСТЕМИ
ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ**

2024