Machine Learning for Enhancing Dementia Screening in Ageing Deaf Signers of British Sign Language

Xing Liang, Bencie Woll, Epaminondas Kapetanios, Anastasia Angelopoulou, Reda Al batat
IoT and Security Research Group, University of Greenwich, UK
Cognitive Computing Research Lab, University of Westminster, UK
Deafness Cognition and Language Research Centre, University College London, UK

LREC 2020

OUTLINE

- Introduction
- Related Work
- Methodology
- Experiments and Analysis
- Conclusions
- Progress Report

Introduction

- British Sign Language (BSL), like other sign languages, uses movements of the hands, body and face for linguistic expression.
- Diagnosis of dementia is subject to the quality of cognitive tests and BSL interpreters alike.
- Hence, the Deaf community currently receives unequal access to diagnosis and care for acquired neurological impairments, with consequent poorer outcomes and increased care costs.

Introduction

- we propose a methodological approach to initial screening that comprises several stages.
- The first stage of research focuses on analysing the motion patterns of the sign space envelope in terms of sign trajectory and sign speed.
- The second stage involves the extraction of the facial expressions of deaf signers.
- the further stage of research implements both VGG16 and ResNet-50 networks using transfer learning to incrementally **identify and improve recognition rates for Mild Cognitive Impairment** (MCI).

Related Work

- Computer vision have been applied to the classification of MR imaging, CT scan imaging, FDG-PET scan imaging or the combined imaging of above, by comparing patients with early stage disease to healthy controls, to distinguish different types or stages of disease and accelerated features of ageing.[1][2]
- In terms of dementia diagnosis, there have been increasing applications of various machine learning approaches, most commonly with imaging data for diagnosis and disease progression.[3][4]

[1] Spasova, S., Passamonti, L., Duggento, A., Li`o, P., Toschi, N., and ADNI. (2019). A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to alzheimer's disease. In: NeuroImage, 189:276–287.
[2] Lu, D., Popuri, K., Ding, G. W., Balachandar, R., Beg, M., and ADNI. (2018). Multimodal and multiscale deep neural networks for the early diagnosis of alzheimer's disease using structural mr and fdg-pet images. In: Scientific Reports, 8(1):5697.
[3] Negin, F., Rodriguez, P., Koperski, M., Kerboua, A., Gonz`alez, J., Bourgeois, J., Chapoulie, E., Robert, P., and Bremond, F. (2018). Praxis: Towards automatic cognitive assessment using gesture. In: Expert Systems with Applications, 106:21–35.
[4] Iizuka, T., Fukasawa, M., and Kameyama, M. (2019). Deep-learning-based imaging-classification identified cingulate island sign in dementia with lewy bodies. In: Scientific Reports, 9(8944).

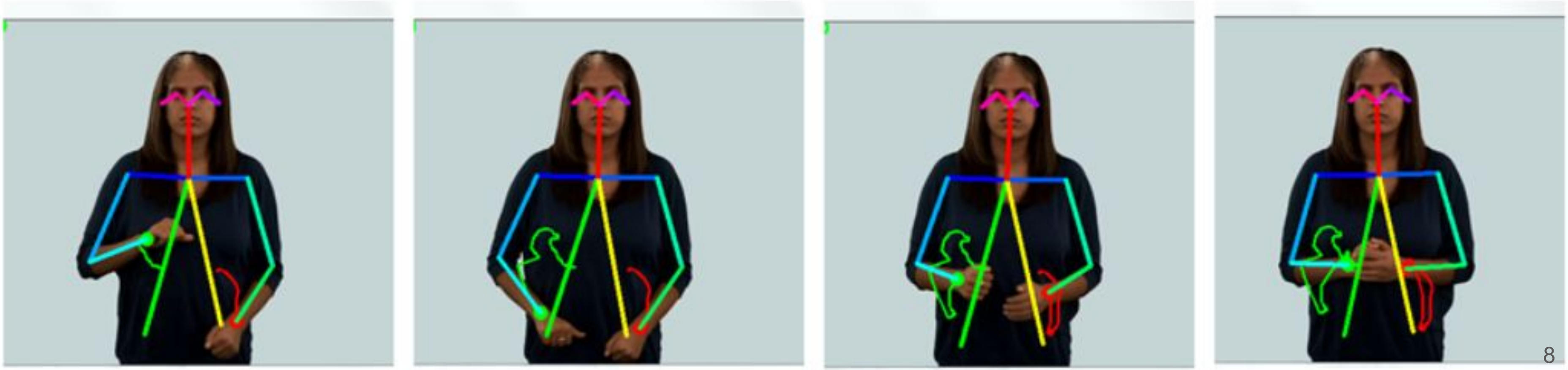
Methodology

Dataset

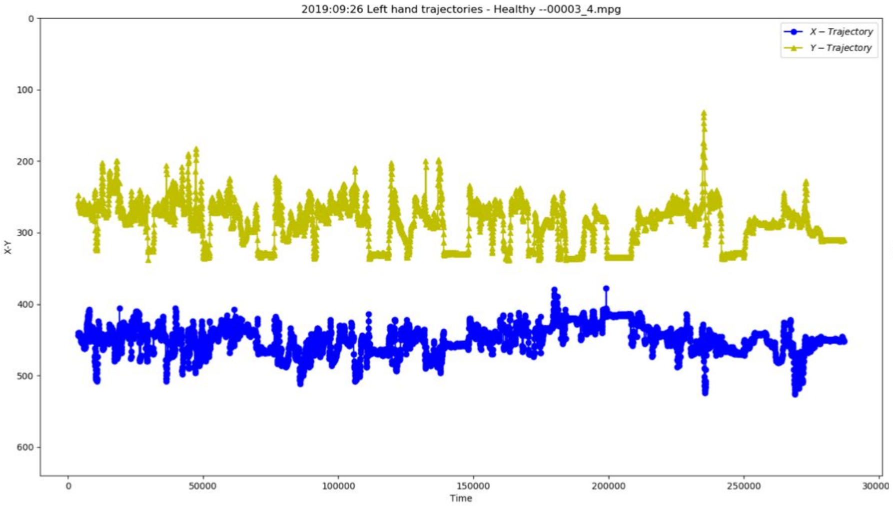
- Deafness Cognition and Language Research Centre (DCAL) at UCL, a collection of 2D video clips of 250 Deaf signers of BSL.
- From the video recordings, we selected 40 case studies of signers (20M, 20F); 21 are signers considered to be healthy cases based on their scores on the British Sign Language Cognitive Screen (BSL-CS); 9 are signers identified as having Mild Cognitive Impairment (MCI) on the basis of the BSL-CS; and 10 are signers diagnosed with MCI through clinical assessment.
- we segmented each into 4-5 short video clips - 4 minutes in length. In this way, we were able to increase the size of the dataset from 40 to 162 clips. Of the 162, 79 have MCI, and 83 are cognitively healthy.

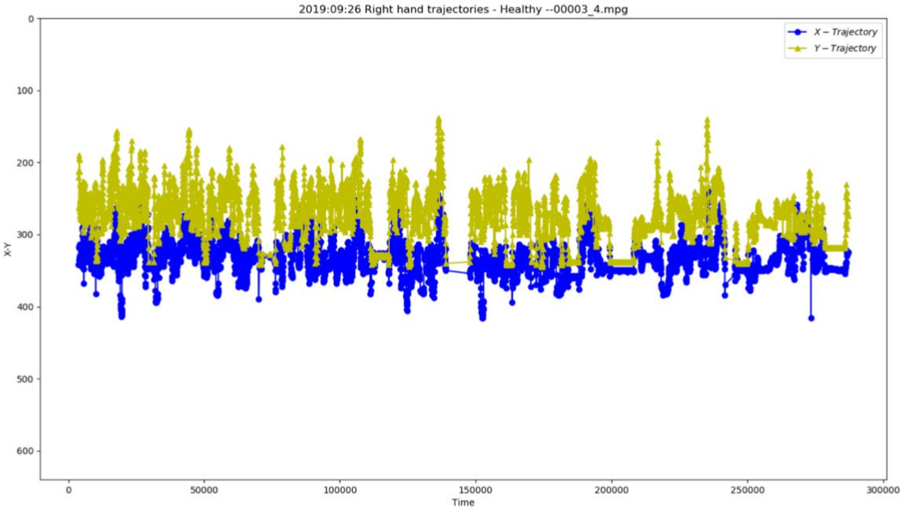
Real-time Hand Trajectory Tracking Model

- The real-time hand movement trajectory tracking model is developed based on the **OpenPose Mobilenet Thin model**.
- only 14 upper body parts in the image are outputted from the tracking model. The hand movement trajectory is obtained via <u>wrist</u> joint motion trajectories.



Real-time Hand Trajectory Tracking Model

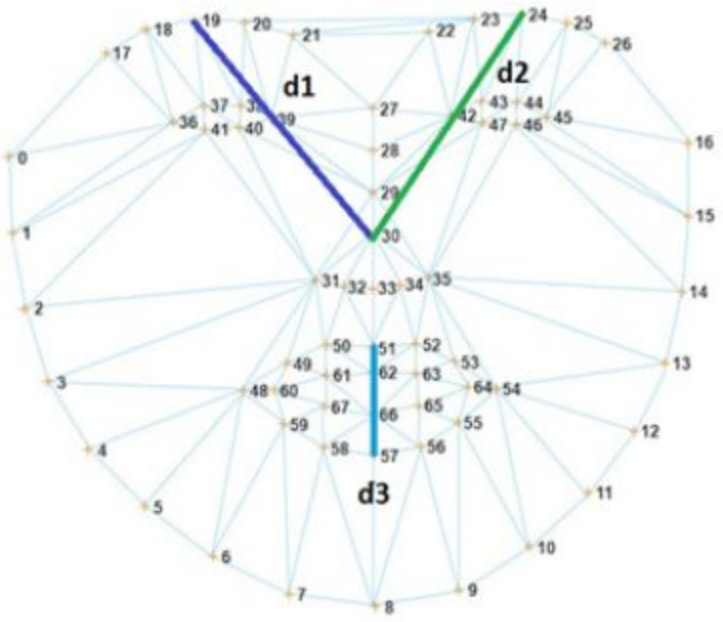




- A spiky trajectory indicates more changes within a shorter period, thus faster hand movement.

Real-time Facial Analysis Model

- The facial analysis model was implemented based on a facial landmark detector inside the **Dlib** library.
- The facial analysis model extracts subtle facial muscle movement by calculating the average Euclidean distance differences between the nose and right brow as d1, nose and left brow as d2, and upper and lower lips as d3 for a given signer over a sequence of video frames.



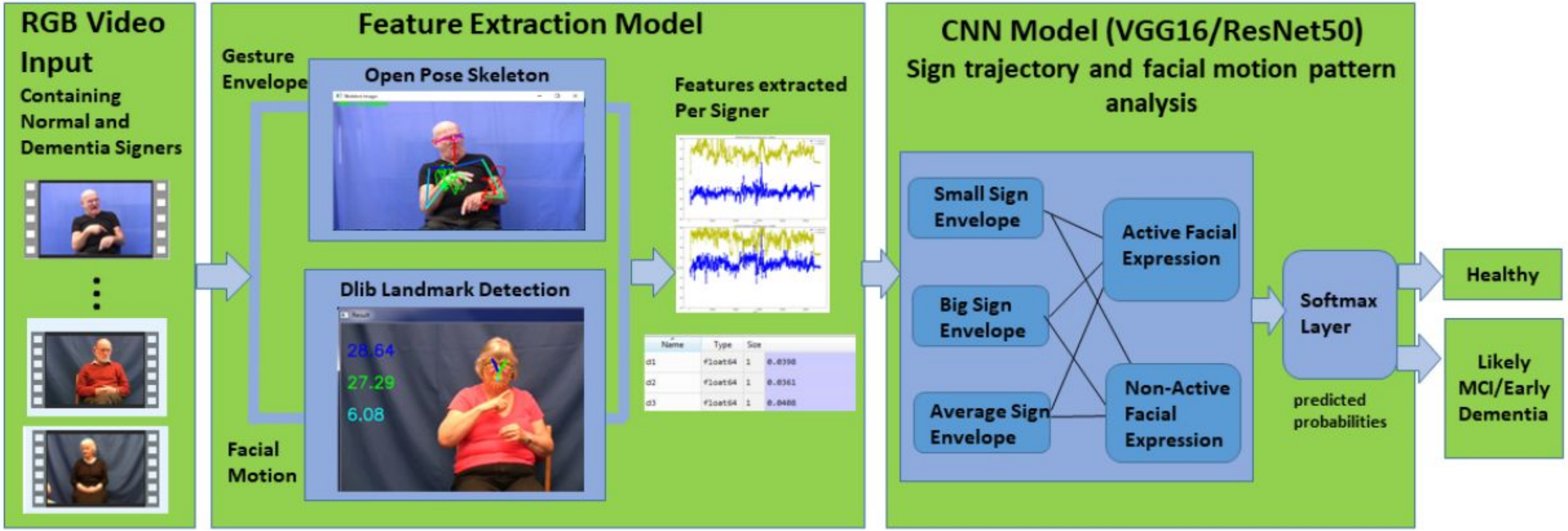
Real-time Facial Analysis Model

$$d1, d2, d3 = \frac{\sum_{t=1}^{T} |d^{t+1} - d^{t}|}{T}$$

where T = Total number of frames that facial landmarks are detected.

The vector [d1, d2, d3] is an indicator of a signer’s facial expression.

Name	Type	Size	Value
d1	float64	1	1.5543
d2	float64	1	1.1702
d3	float64	1	1.5754



MCI: Mild Cognitive Impairment

Experiments

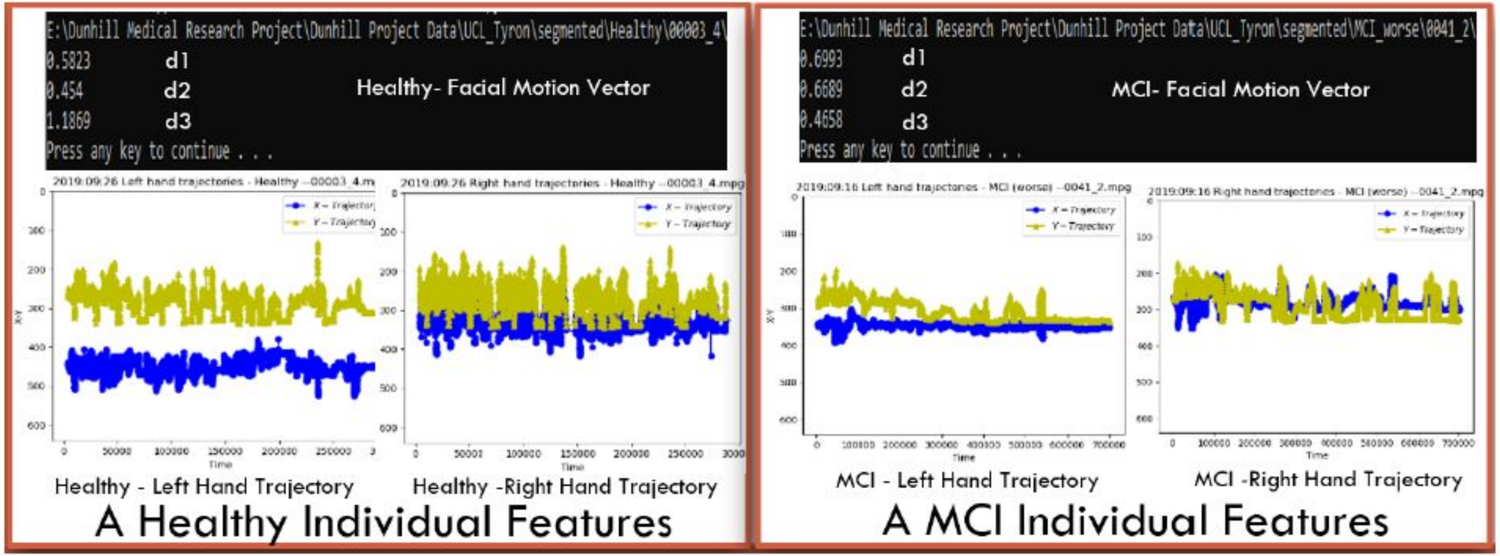
- In our approach, we have used VGG16 and ResNet-50 as the base models, with transfer learning to transfer the parameters pre-trained for the 1000 object detection task on the ImageNet dataset to recognise hand movement trajectory images for early Mild Cognitive Impairment (MCI) screening.
- Due to the very small dataset, we train ResNet-50 as a classifier alone and fine tune the VGG16 network.

Results Discussion

Table 1: Performance Evaluation over VGG16 and RestNet-50 for early MCI screening

Method	40 Participants 21 Healthy, 19 Early MCI			6 Participants 5 Healthy, 1 Early MCI	
	Train Result (129 segmented cases)	Test Result (33 segmented cases)		Validation Result (24 segmented cases)	
	ACC	ACC	ROC	ACC	ROC
VGG 16	87.5969%	**87.8788%**	**0.93**	**87.5%**	**0.96**
ResNet-50	69.7674%	69.6970%	0.72	66.6667%	0.73

Results Discussion



- Signers with MCI produced more static poses/pauses during signing.
- At the same time, the Euclidean distance d3 of healthy signers is larger than that of MCI signers, indicating active facial movements by healthy signers.

Conclusion

- We have outlined a methodological approach and developed a toolkit for an automatic dementia screening system for signers of BSL.
- The experiments show the effectiveness of our deep learning based approach for early stage dementia screening.
- The results are validated against cognitive assessment scores with a test set performance of 87.88%.

Progress Report

Datasets

- 24 videos
- Conversational Question Answering
- 臨床失智評估量表 CDR
- 0 健康 / 0.5 疑似輕微 / 1 輕度 / 2 中度
- CDR = 0 / 0.5 / 1 / 2 : 2 / 5 / 15 / 2 videos

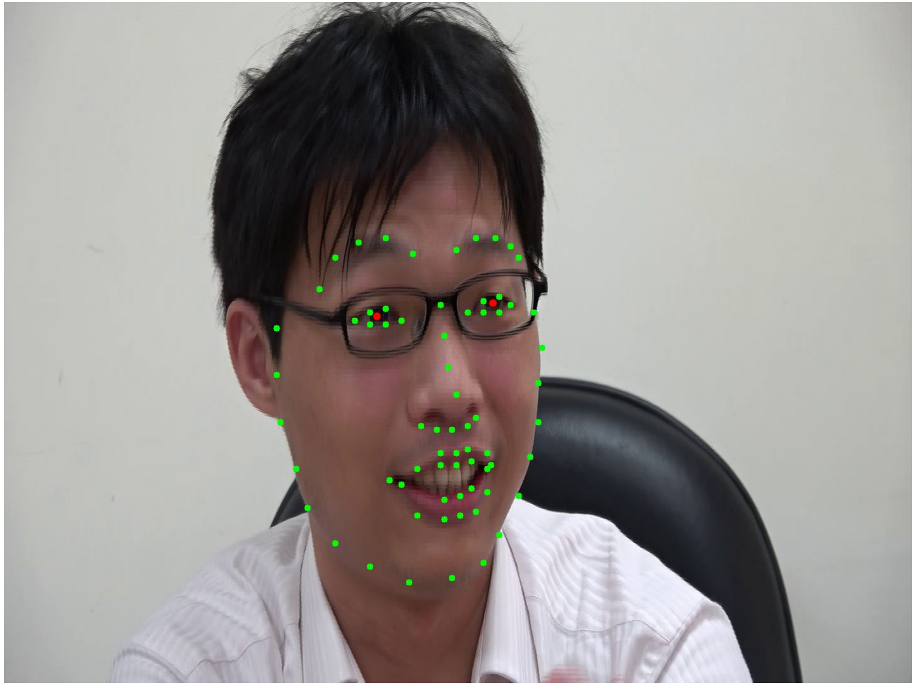
GOAL

- Input Videos → **Model** → CDR Score (0/0.5/1/2)

Data Augmentation

- Flip Horizontally → 48 videos

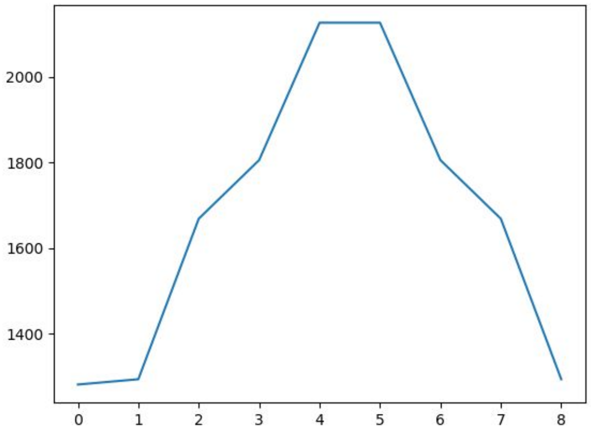
Facial Landmarks & Pupils Detection



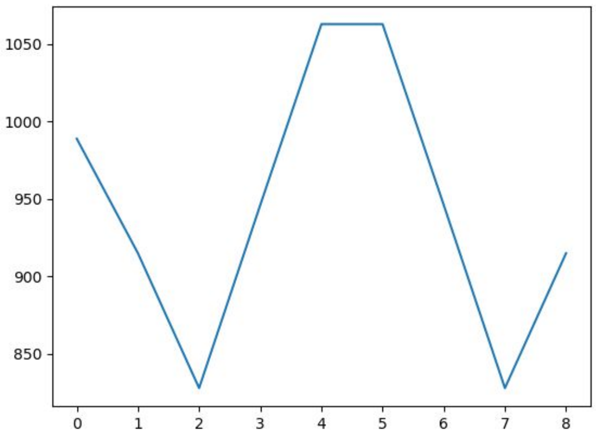
Short Time Fourier Transformation

- P(x,y) → dx, dy → STFT (10 frames) → superposition → 18-D vector

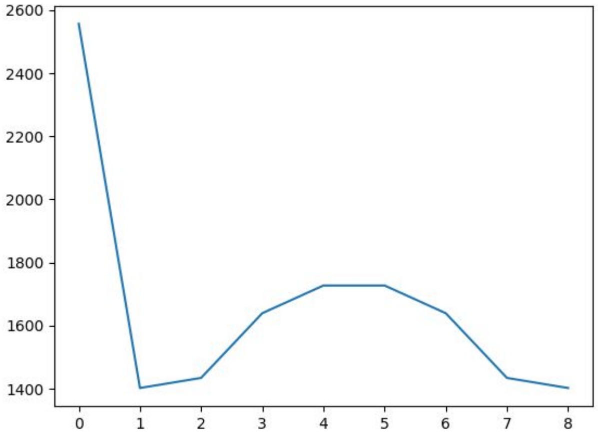
CDR = 0



CDR = 1



CDR = 2



Train SVM

- Pick 4 points (left pupil / right pupil / both corners of mouth) → 72-D vector
- Training set：22 videos / Testing set：20 videos
- Result
  - True = [0. 2. 2. 1. 2. 2. 1. 1. 2. 3. 0. 2. 2. 1. 2. 2. 1. 1. 2. 3.]
  - pred = [0. 1. 2. 1. 1. 3. 1. 3. 2. 2. 0. 0. 2. 1. 2. 3. 1. 3. 2. 2.]
  - **Accuracy：11/20**

Train SVM #2

- Pick 6 points (pupils / corners of mouth / corners of eyes) → 108-D vector
- Training set：22 videos / Testing set：20 videos
- Result
  - True = [0. 2. 2. 1. 2. 2. 1. 1. 2. 3. 0. 2. 2. 1. 2. 2. 1. 1. 2. 3.]
  - pred = [0. 1. 2. 1. 2. 3. 1. 3. 2. 2. 0. 0. 2. 1. 2. 3. 1. 3. 2. 2.]
  - **Accuracy：12/20**