ICCCM
The 10th International Conference on Computer and Communications Management
July 29-31, 2022 | Okayama University, Japan

Deep Learning based Micro-expression Recognition

Wenjuan Gong

China University of Petroleum (East China)

中國石油大學（华东）
CHINA UNIVERSITY OF PETROLEUM

What are micro-expressions and what is MER for?

- The universal **facial expressions** include seven emotions – anger, contempt, disgust, fear, joy, sadness, and surprise. Facial expressions typically last between **0.5 to 4 seconds**.
- **Microexpressions**, however, are expressions that go on and off the face in a fraction of a second, sometimes as fast as **1/30 of a second**.
- Microexpression recognition (MER) is widely applied in:

Psychology Diagnosis

Teaching Assessment

Business Negotiations

Criminal Investigation

Roadmap

- A survey on deep learning based micro-expression recognition

- Meta-Learning Based Multi-Model Fusion Network

Deep learning-based microexpression recognition: a survey

Wenjuan Gong[1] · Zhihong An[1] · Noha M. Elfiky[2]

Received: 9 July 2021 / Accepted: 28 February 2022 / Published online: 1 April 2022

Contributions

- **Deep learning based micro-expression recognition: a survey**

1. We studied related works on deep learning based micro-expression recongnition

2. We fused several public datasets to form a composite dataset

3. We provided a baseline method on the composite dataset

Introduction

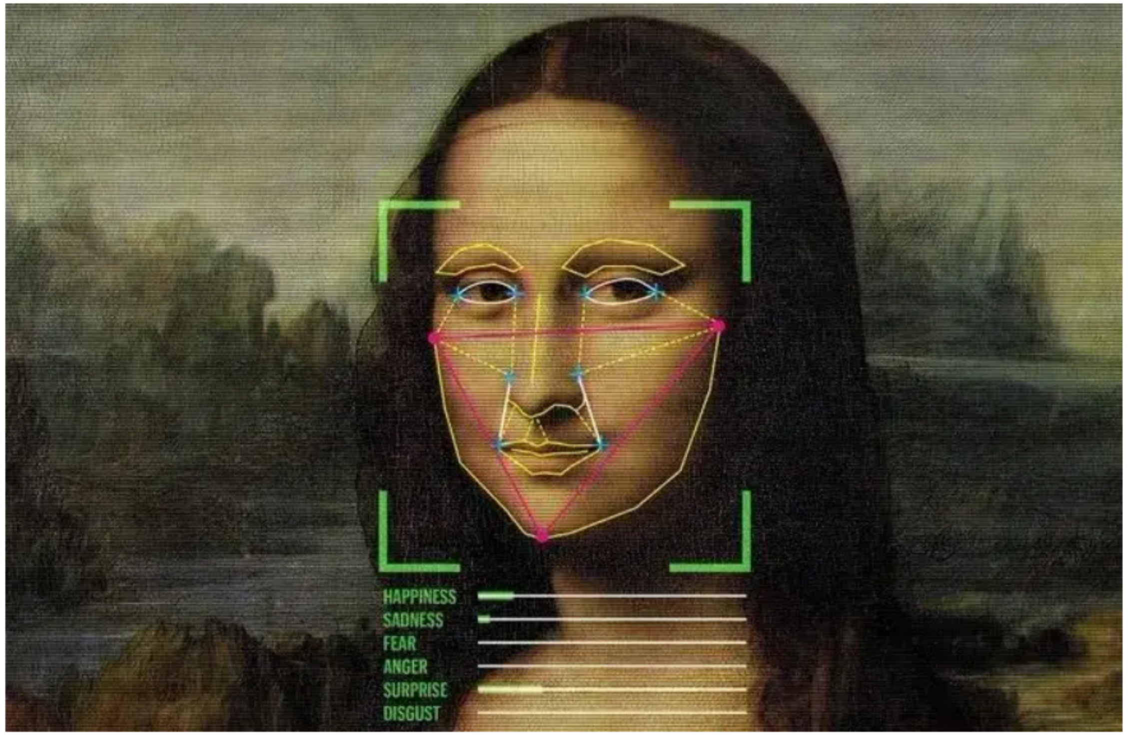
Micro-expressions are special facial expressions. Compared with ordinary expressions, micro-expressions mainly have the following characteristics:

1. The duration is short, usually only 1/25s~1/3s;

2. The action intensity is low and difficult to detect;

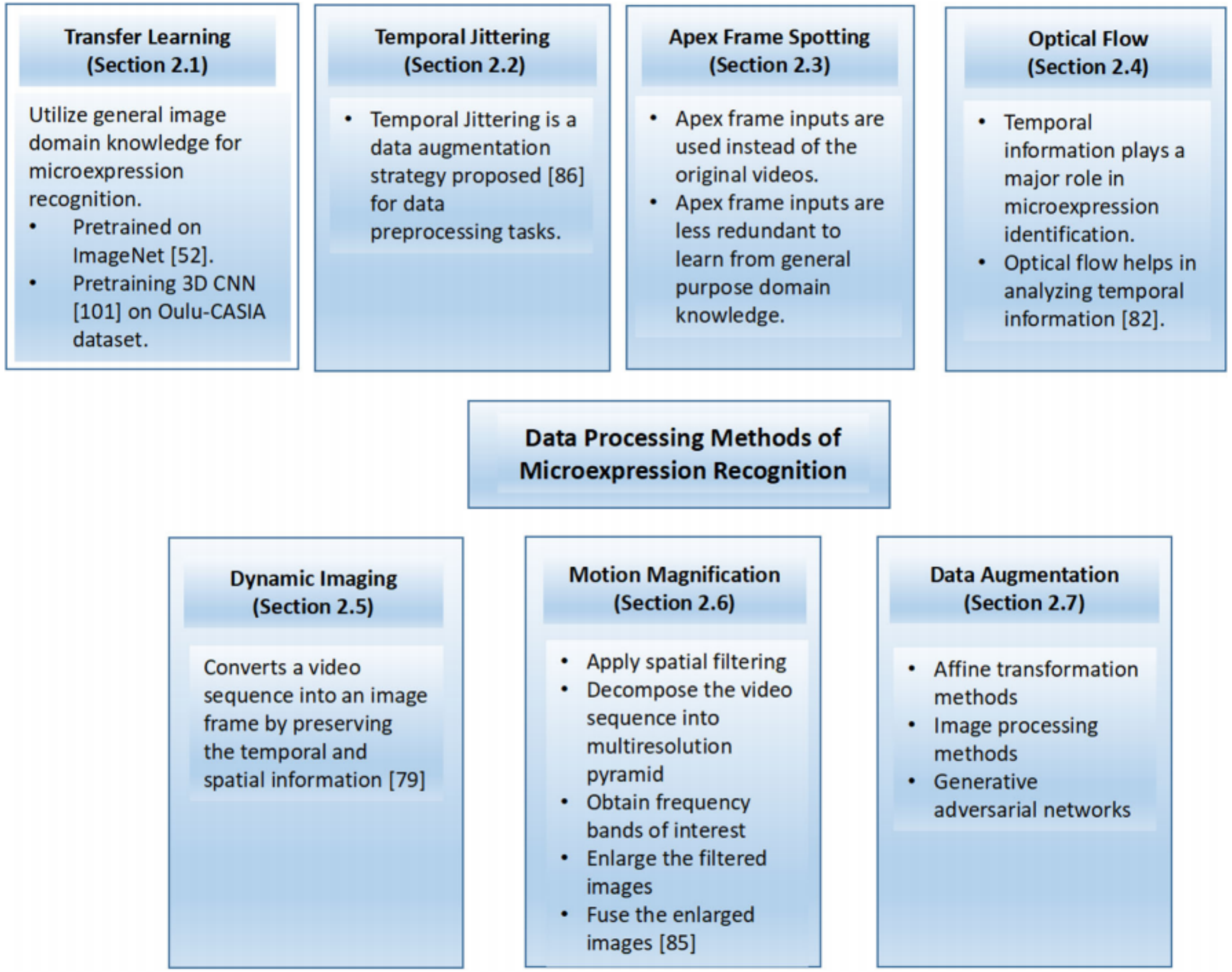
3. Produced in an unconscious state, usually difficult to disguise;

4. The analysis of micro-expression usually needs to be in the video, while the ordinary expression can be analyzed in the image.



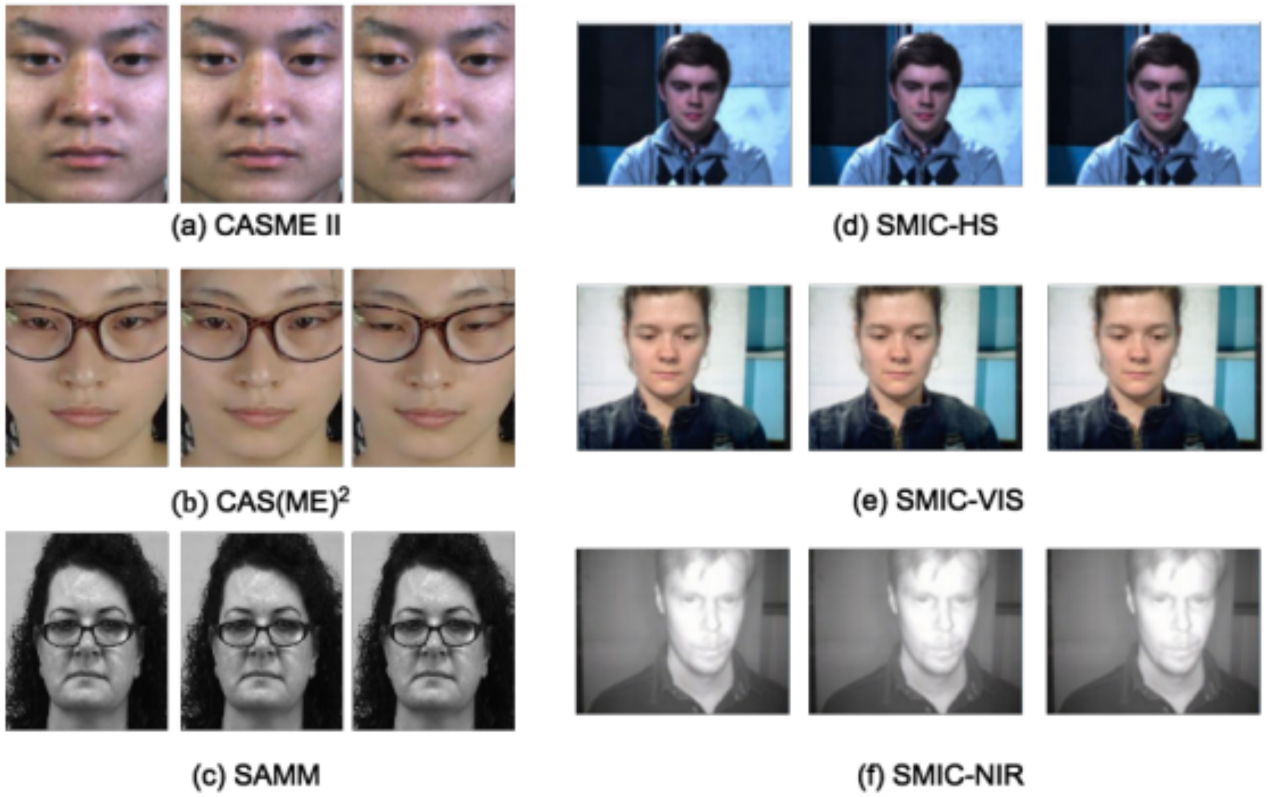
Data preprocessing methods

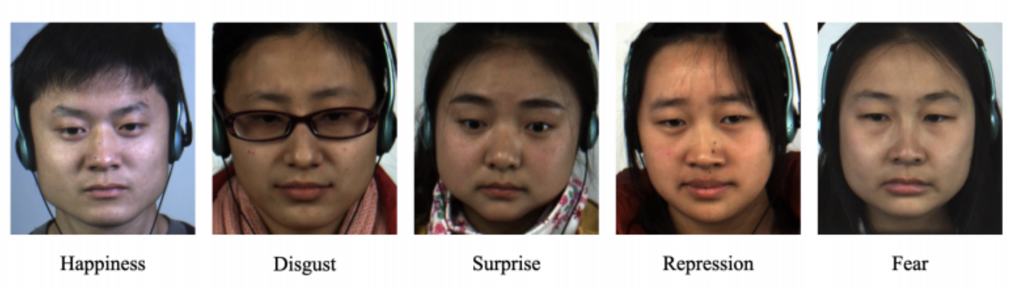
➢ Due to the short duration and low action intensity of micro-expressions, it is difficult to extract features, so appropriate data preprocessing and feature extraction are required.



Micro-expression Datasets

➢ Due to the difficulties in data collection and identification of micro-expressions, there are few existing micro-expression datasets, which makes the application of deep learning in micro-expression recognition difficult.





CASME II dataset

SAMM dataset

Micro-expression Recognition Methods

01 Spatial deep learning methods
- ➢ Convolutional neural network-based method
- ➢ Multi-stream convolutional neural networks based methods

02 Temporal deep learning methods

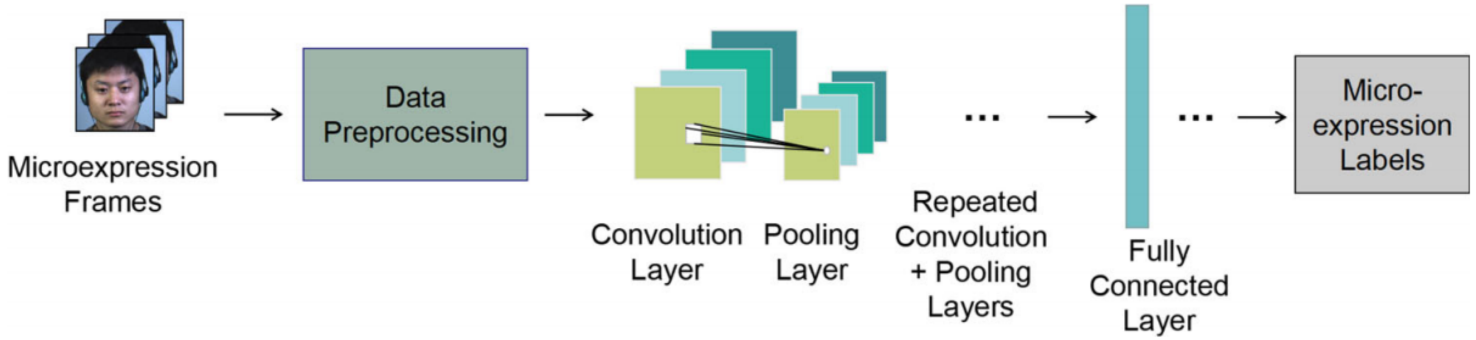
03 Spatial-temporal deep learning methods
- ➢ 3D convolutional neural network-based methods
- ➢ Combined spatial-temporal methods

04 Spatial-contextual deep learning methods
- ➢ Capsule neural network-based methods
- ➢ Graph convolutional networks based methods

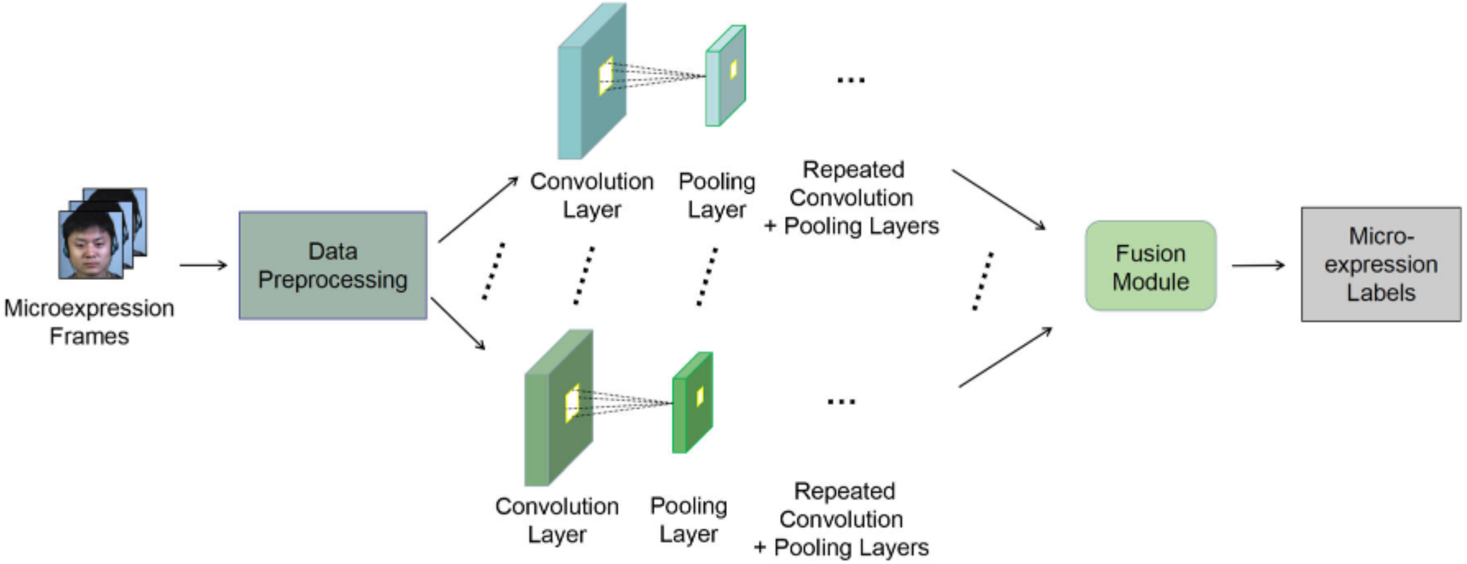
05 Hybrid approaches combining traditional and deep learning methods

Spatial deep learning methods

➢ The Structure of Convolutional Neural Network based Methods



➢ The structure of multistream convolutional neural network-based methods



Micro-expression Recognition Methods

01 Spatial deep learning methods
- ➢ Convolutional neural network-based method
- ➢ Multi-stream convolutional neural networks based methods

02 Temporal deep learning methods

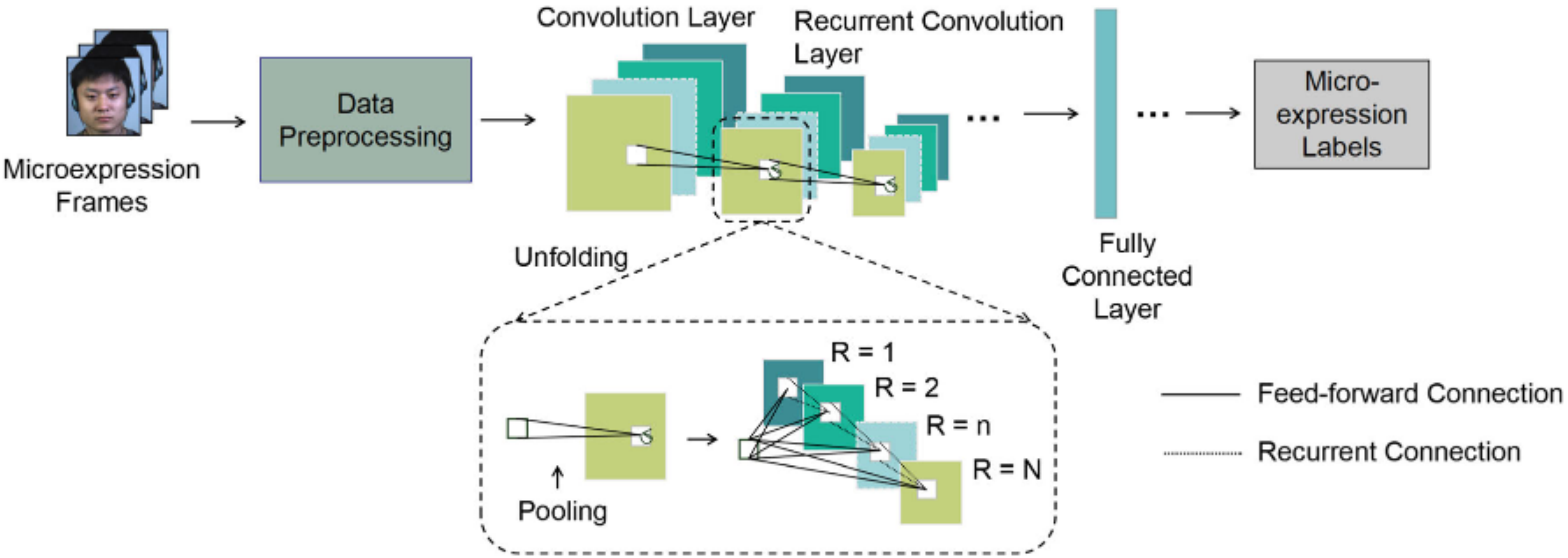
03 Spatial-temporal deep learning methods
- ➢ 3D convolutional neural network-based methods
- ➢ Combined spatial-temporal methods

04 Spatial-contextual deep learning methods
- ➢ Capsule neural network-based methods
- ➢ Graph convolutional networks based methods

05 Hybrid approaches combining traditional and deep learning methods

Temporal deep learning methods

➢ The structure of recurrent convolutional network-based methods



Micro-expression Recognition Methods

01 Spatial deep learning methods

- ➢ Convolutional neural network-based method
- ➢ Multi-stream convolutional neural networks based methods

02 Temporal deep learning methods

03 Spatial-temporal deep learning methods

- ➢ 3D convolutional neural network-based methods
- ➢ Combined spatial-temporal methods

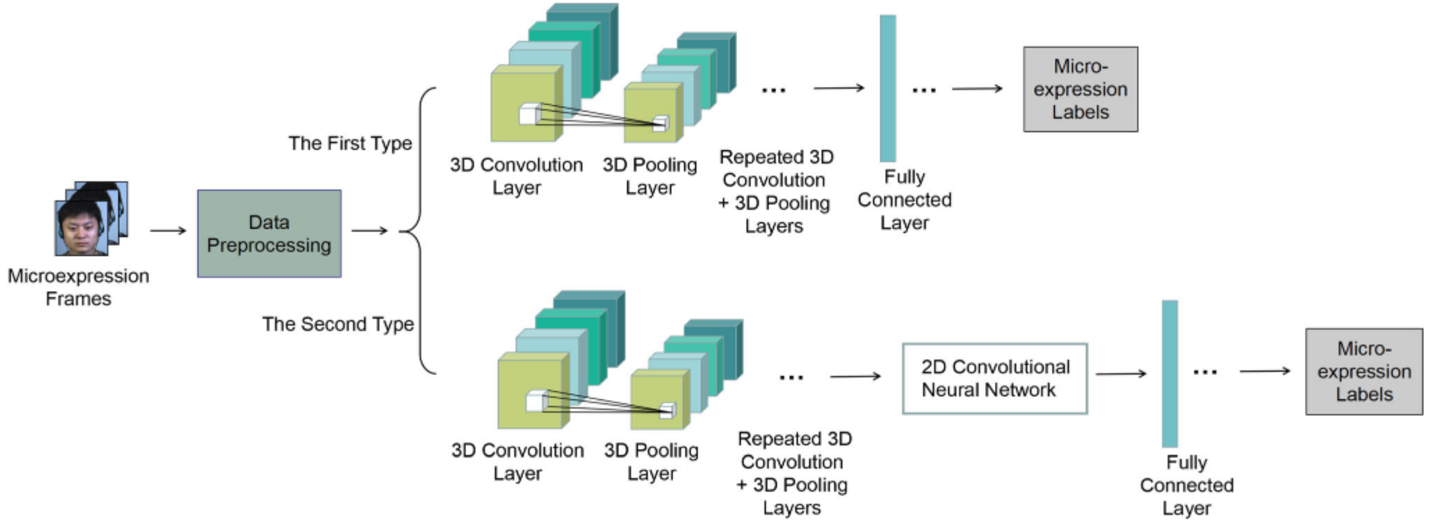
04 Spatial-contextual deep learning methods

- ➢ Capsule neural network-based methods
- ➢ Graph convolutional networks based methods

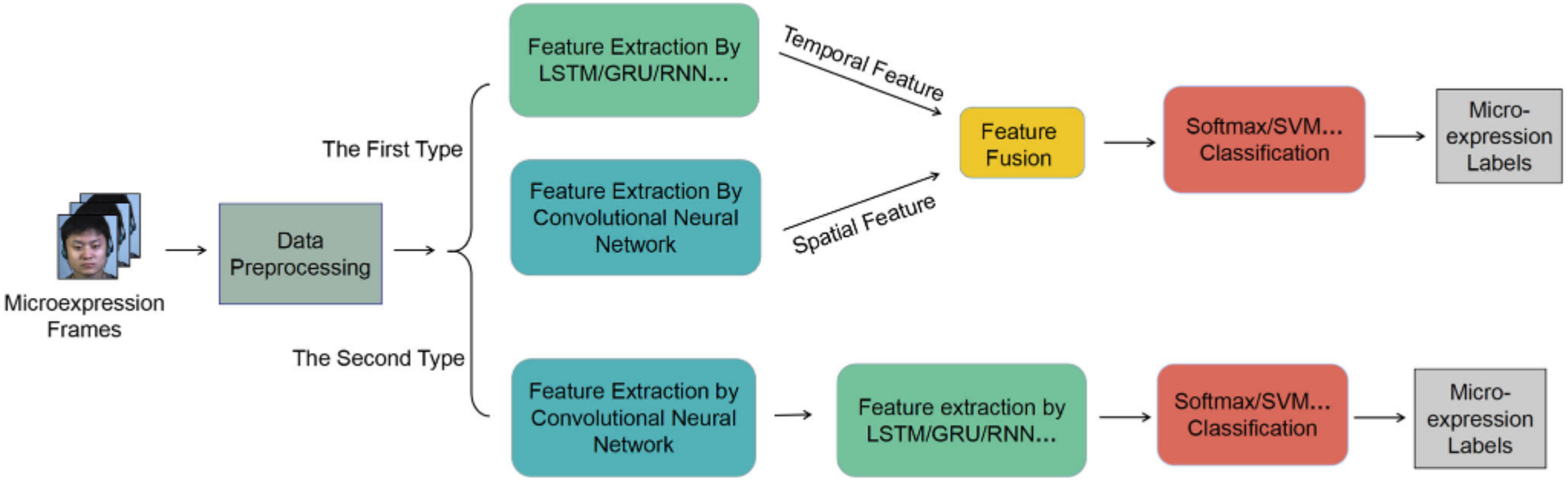
05 Hybrid approaches combining traditional and deep learning methods

Spatial-temporal deep learning methods

➢ The structure of 3D convolutional neural network-based methods



➢ The structure of combined spatial-temporal methods



Micro-expression Recognition Methods

01 Spatial deep learning methods
- ➢ Convolutional neural network-based method
- ➢ Multi-stream convolutional neural networks based methods

02 Temporal deep learning methods

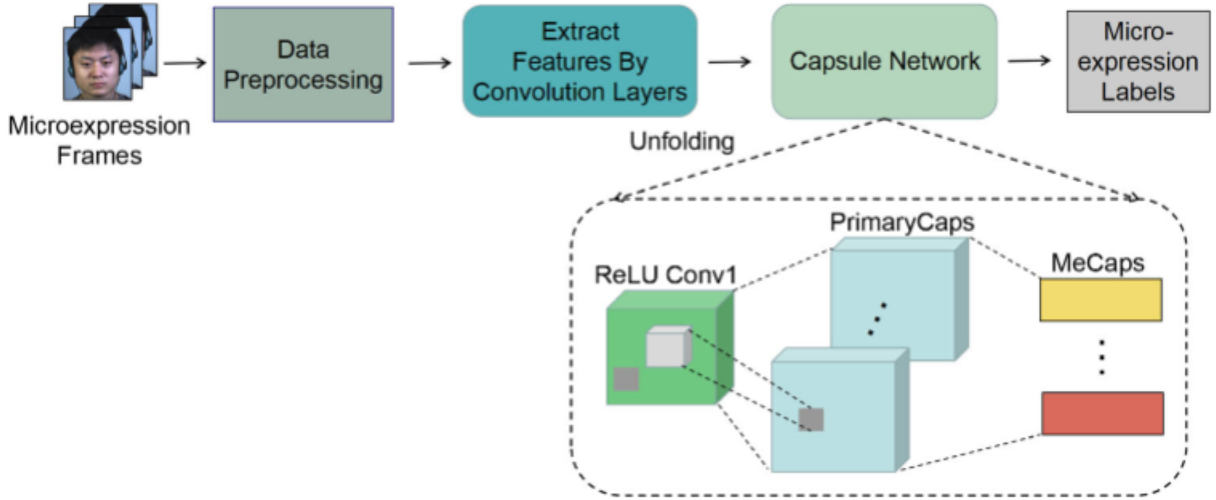
03 Spatial-temporal deep learning methods
- ➢ 3D convolutional neural network-based methods
- ➢ Combined spatial-temporal methods

04 Spatial-contextual deep learning methods
- ➢ Capsule neural network-based methods
- ➢ Graph convolutional networks based methods

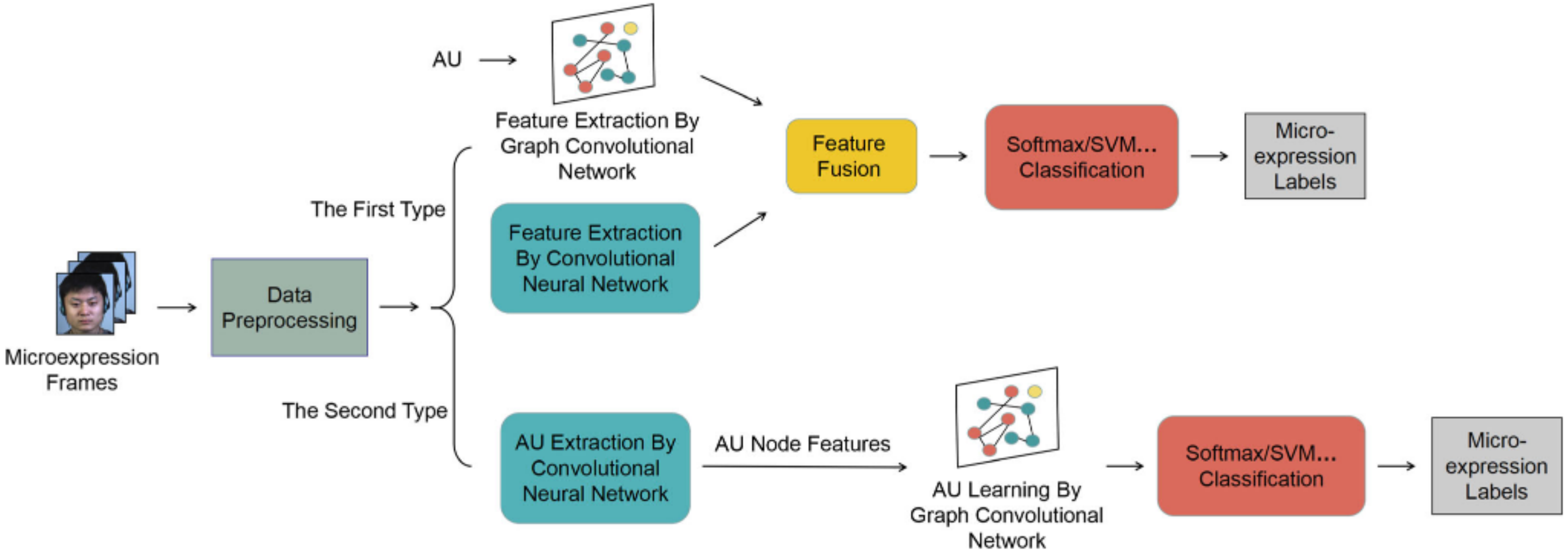
05 Hybrid approaches combining traditional and deep learning methods

Spatial-contextual deep learning methods

➢ The structure of capsule neural network-based methods



➢ The structure of graph convolutional network-based methods



Micro-expression Recognition Methods

01 Spatial deep learning methods

- ➢ Convolutional neural network-based method
- ➢ Multi-stream convolutional neural networks based methods

02 Temporal deep learning methods

03 Spatial-temporal deep learning methods

- ➢ 3D convolutional neural network-based methods
- ➢ Combined spatial-temporal methods

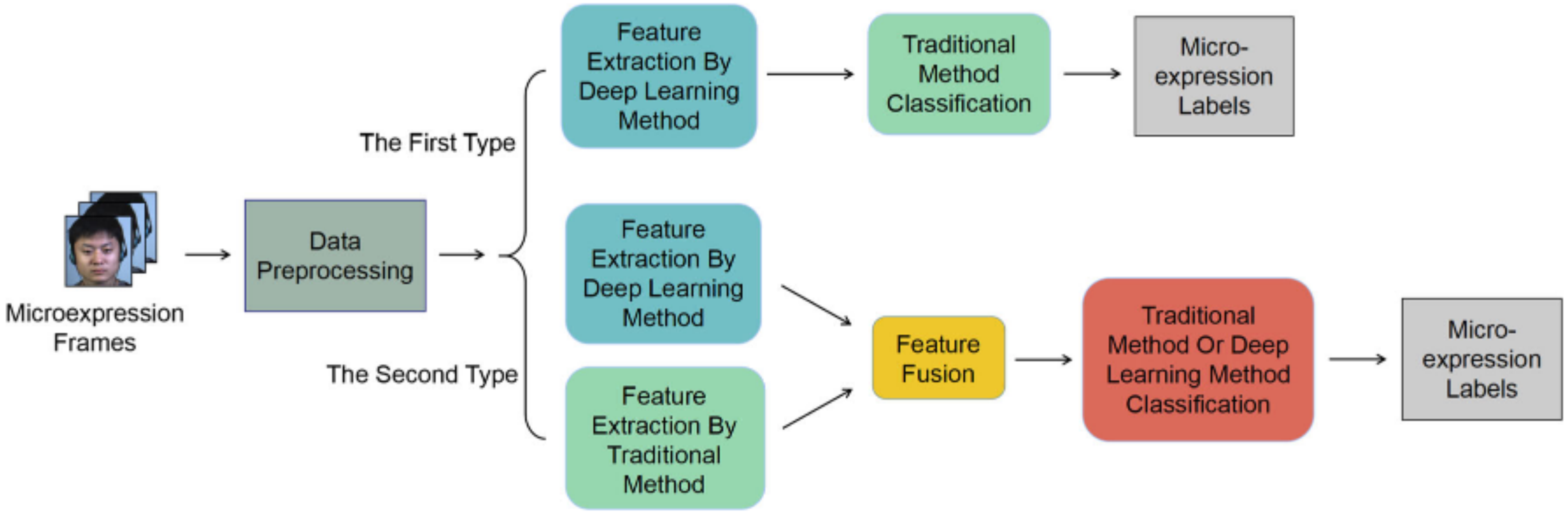
04 Spatial-contextual deep learning methods

- ➢ Capsule neural network-based methods
- ➢ Graph convolutional networks based methods

05 Hybrid approaches combining traditional and deep learning methods

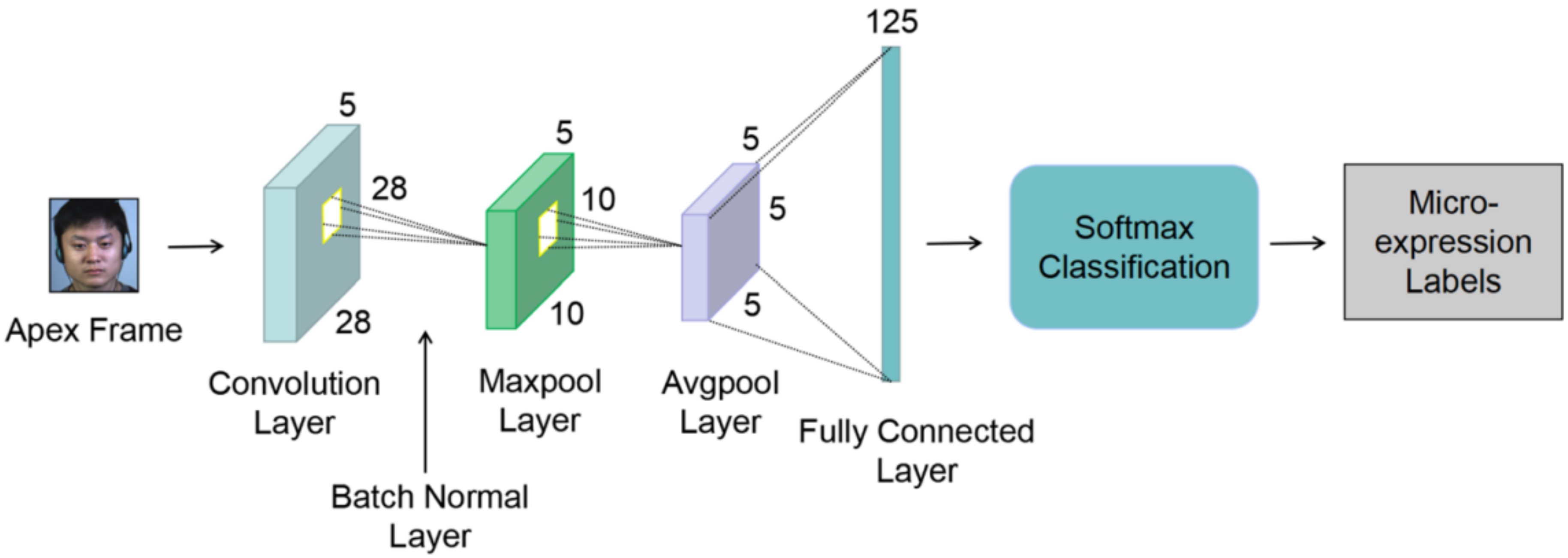
Hybrid approaches combining traditional and deep learning methods

➢ The structure of hybrid approaches combining traditional and deep learning methods



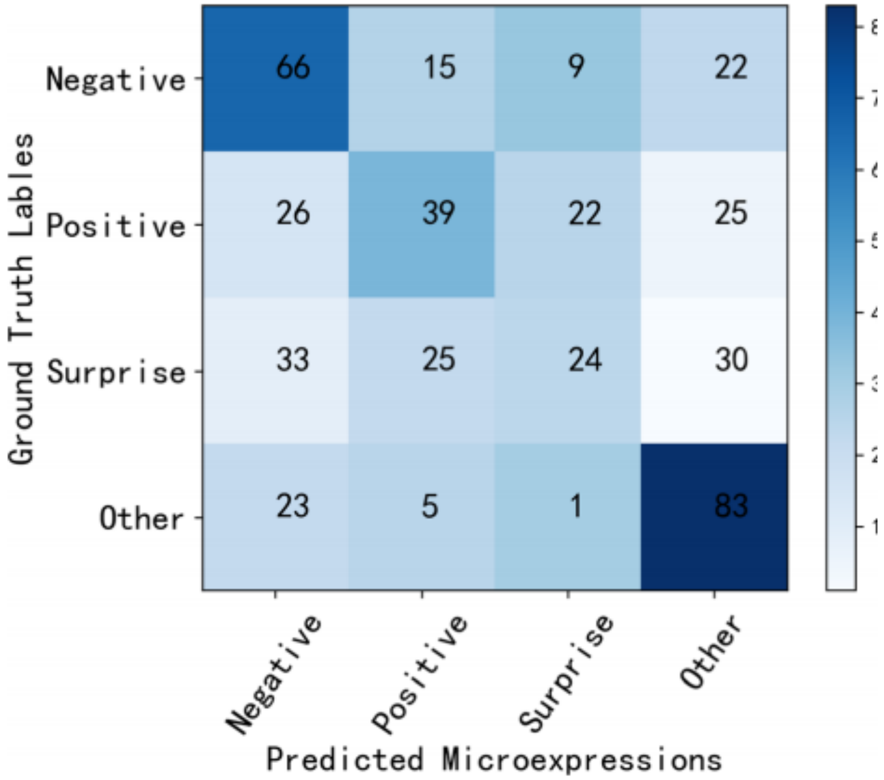
The baseline method

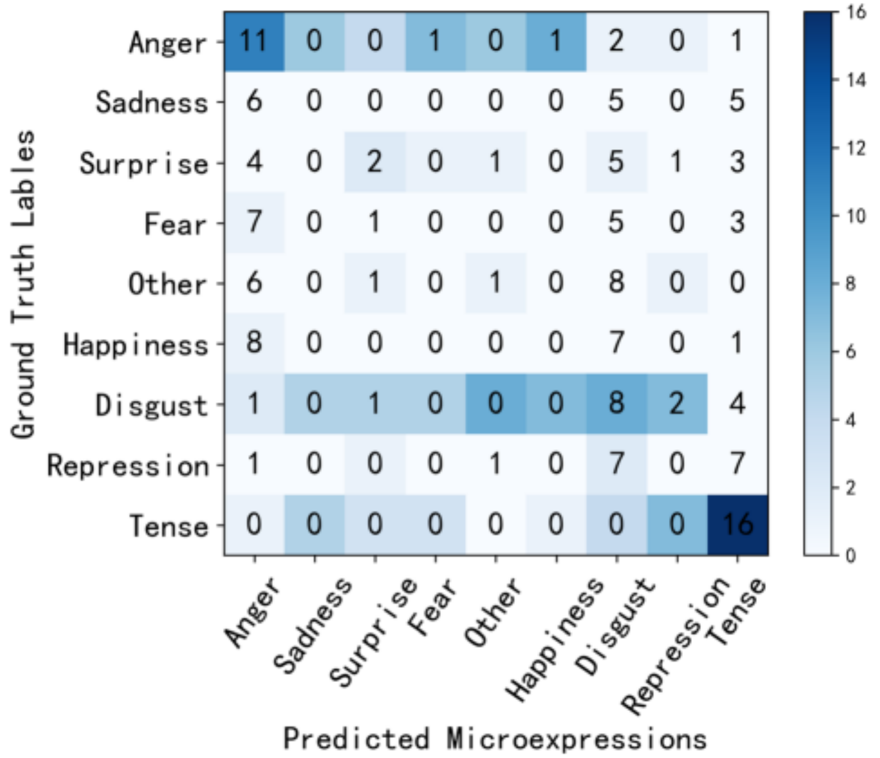
➢ Five datasets (the CASME database, the CASME II database, the CAS(ME)^2 database, the SMIC database, and the SAMM dataset) were fused to form two new datasets (the 4-class fused dataset and the 9-class fused dataset), and use these as the baseline model input.



The baseline method

Evaluation Method	4-class Fused Dataset	9-class Fused Dataset
LOVO	47.32%	26.39%



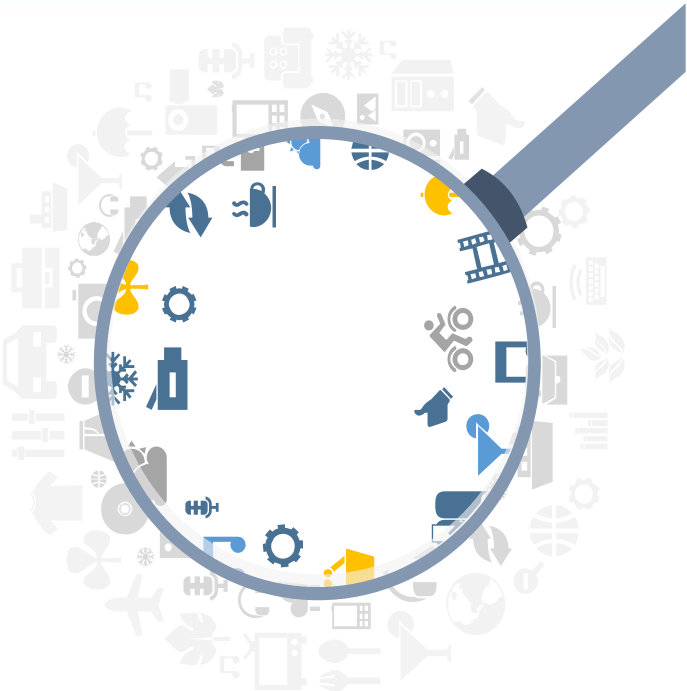


Possible future works

01 Micro-expression dataset collection

02 Optimizing network structure

03 Multimodal information for microexpression recognition



Meta-MMFNet: Meta-Learning Based Multi-Model Fusion Network for Micro-Expression Recognition

WENJUAN GONG and YUE ZHANG, China University of Petroleum (East China), China
WEI WANG, Institute of Automation Chinese Academy of Sciences, China
PENG CHENG, Institute of High Performance Computing, A*STAR, Singapore
JORDI GONZÀLEZ, Computer Vision Center, Autonomous University of Barcelona, Spain

Contributions

- **Meta-Learning Based Multi-Model Fusion Network**

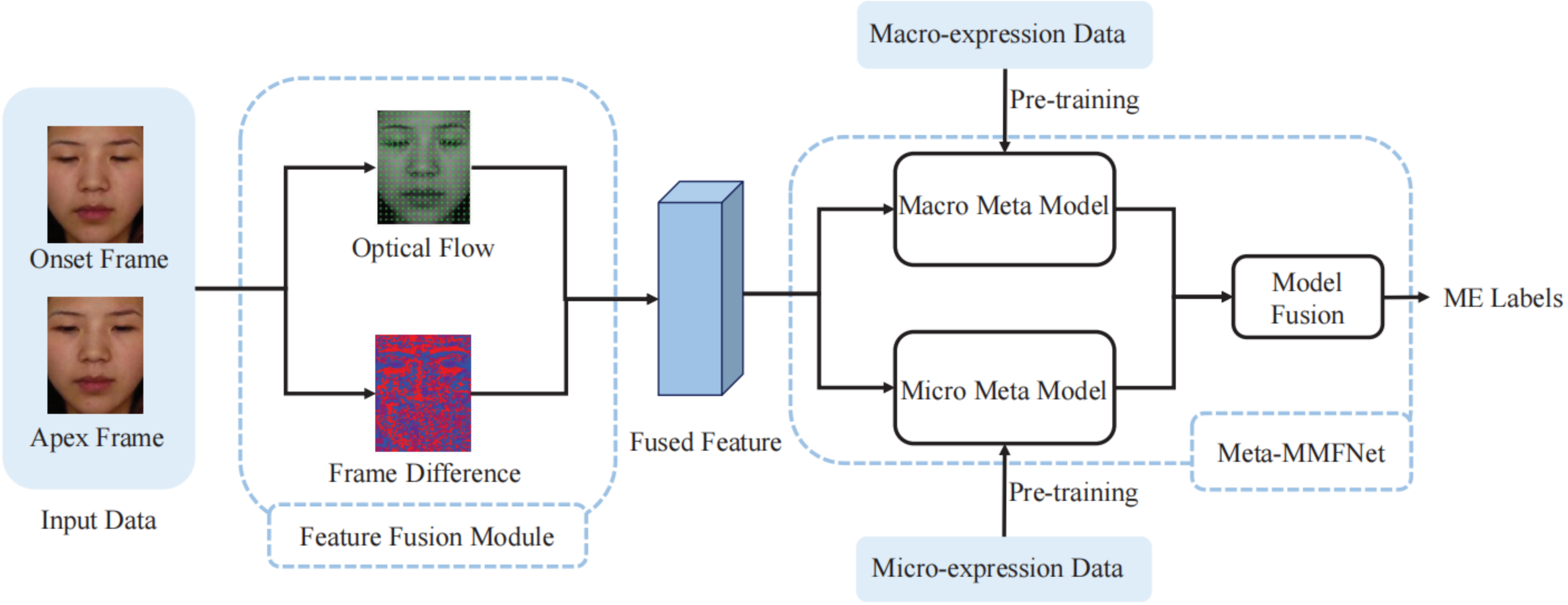
1. Meta-MMFNet fuses optical flow and frame difference features

2. Meta-MMFNet fuses micro-expression pretrained model and macro-expression pretrained model

3. Meta-MMFNet achieves state-of-art performances

Meta-Learning Based Multi-Model Fusion Network

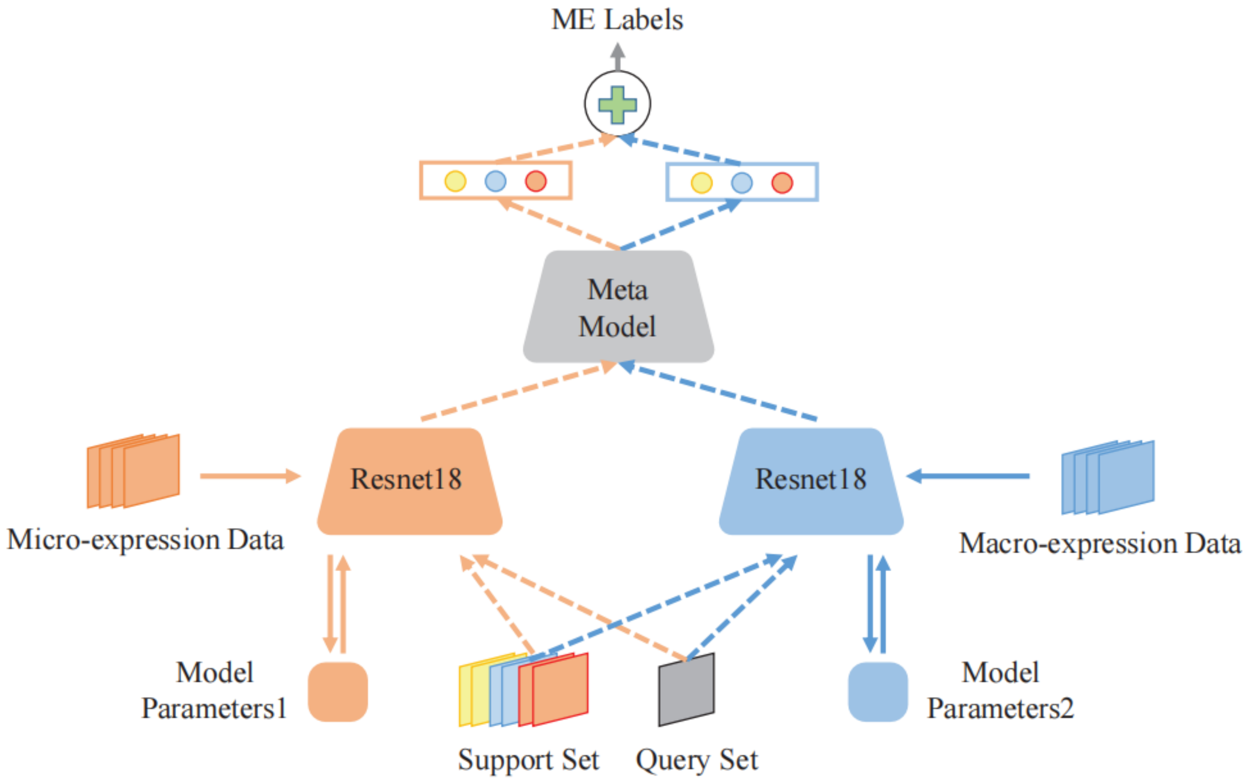
- The overall structure of the proposed Meta-learning based Multi-model Fusion Network (Meta-MMFNet) Method.



Meta-Learning Based Multi-Model Fusion Network

➢ The Network structure

- Meta-learning framework



Meta-Learning Based Multi-Model Fusion Network

➢ The Network structure

- Loss function:

$$w_c = \frac{1}{|S_c|} \sum_{x \in S_c} f_\theta(x),$$

$$p_i^c = p(y = c|x_i) = \frac{exp(< f_\theta(x_i), w_c >)}{\sum_{c'} exp(< f_\theta(x_i), w_{c'} >)},$$

$$Loss = -\sum_{i=1}^{n} Y_i log P_i,$$

$$d_{sum}^i = d_{micro}^i + \alpha d_{macro}^i, i \in \{1, 2, \dots, c\},$$

Experiments

Table 3. Ablation Study

Models	Dataset		
	CASME	CASME II	SMIC
Micro-model	0.6815	0.7733	0.6220
Macro-model	0.6051	0.7733	0.5244
Meta-MMFNet	0.6959	0.8095	0.6313

Table 4. Performance Comparisons between the Proposed Meta-MMFNet and the State-of-the-art Methods on the CASME Dataset

Methods	CASME Dataset				
	Disgust	Surprise	Repression	Tense	Overall
STCLQP [14]	0.64	0.50	0.53	0.58	0.5731
LBP-SIP [48]	-	-	-	-	0.3684
FHOFO [11]	-	-	-	-	0.6599
MER-RCNN [52]	-	-	-	-	0.632
STLBP-RIP [13]	0.5682	0.6	0.4211	**0.8136**	0.5906
DiSTLBP-RIP [13]	**0.7273**	0.6	0.5263	0.6667	0.6433
LGCcon [20]	0.57	0.8	0.21	0.77	0.6082
3DFCNN [17]	-	-	-	-	0.5444
Our Macro-model	0.625	0.8889	0.4242	0.6061	0.6051
Our Micro-model	0.7	0.8889	**0.5758**	0.6667	0.6815
Our Meta-MMFNet	0.7179	**0.9412**	0.5357	0.6875	**0.6959**

Table 5. Performance Comparisons between the Proposed Meta-MMFNet and the State-of-the-art Methods on the CASME II Dataset

Methods	CASME II Dataset				
	Surprise	Repression	Happiness	Disgust	Overall
DTCM [28]	-	-	-	-	0.7206
Our Macro-model	0.8929	**0.5926**	0.6563	0.8571	0.7733
Our Micro-model	**0.9643**	**0.5926**	0.5313	0.8889	0.7733
Our Meta-MMFNet	0.8929	**0.5926**	**0.6897**	**0.9206**	**0.8095**

Table 6. Performance Comparisons between the Proposed Meta-MMFNet and the State-of-the-art Methods on the SMIC Dataset

Methods	SMIC Dataset			
	Surprise	Positive	Negative	Overall
LBP-SIP [48]	-	-	-	0.4212
FDM [54]	0.53	**0.66**	0.41	0.5488
MER-RCNN [52]	-	-	-	0.571
FHOFO [11]	-	-	-	0.5122
Hierarchical STLBP-IP [61]	-	-	-	0.6078
FR [60]	-	-	-	0.579
3DFCNN [17]	-	-	-	0.5549
Our Macro-model	0.6512	0.5686	0.4143	0.5244
Our Micro-model	0.7442	0.5294	**0.6143**	0.6220
Our Meta-MMFNet	**0.7561**	0.5600	0.6087	**0.6313**

Experiments

Table 7. Composite Dataset Annotations

Datasets	Composite Database Evaluation (CDE) Dataset Categories		
	Surprise	Positive	Negative
CASME	Surprise	Happiness	Disgust,Repression,Tense,Sadness,Fear,Comtempt
CASME II	Surprise	Happiness	Disgust,Repression,Fear,Sadness
SMIC	Surprise	Positive	Negative

"positive," "negative," and "surprise." The relations between the original and composite dataset annotations are listed in Table 7.

Table 8. Experimental Results of the Composite Database Evaluation

Method	Surprise	Positive	Negative	Average Accuracy	Overall Accuracy
LBP-SIP [48]	-	-	-	0.3948	-
FHOFO [11]	-	-	-	0.5861	-
3DFCNN [17]	-	-	-	0.5497	-
Our Meta-MMFNet	-	-	-	0.7122	-
Our Meta-MMFNet_retrain	0.8736	0.5275	0.7855	-	**0.7526**

Thank You For Listening!

E-mail: wenjuangong@upc.edu.cn

MIGroup： http://wenjuangong.com/

中国石油大學（华东）
CHINA UNIVERSITY OF PETROLEUM