Orthogonalization-Guide Feature Fusion Network for Multimodal 2D+3D Facial Expression Recognition

Shisong Lin [1], Mengchao Bai [1], Feng Liu [1], Linlin Shen[1*] and Yicong Zhou [2]

[1]school of Computer Science & Software Engineering, Shenzhen University, China

[2]School of Science and Technology, Macau University, Macau, China

➤ Abstract

We present an efficient orthogonalization-guided feature fusion network, namely OGF2Net, to fuse the features extracted from 2D and 3D faces for facial expression recognition. While 2D texture maps are fed into a 2D feature extraction pipeline (FE2DNet), the 3D attribute maps are input of the 3D feature extraction pipeline (FE3DNet). To reduce the redundancies among features extracted from 2D and 3D streams, we design an orthogonal loss-guided feature fusion network to orthogonalize the features before fusing them. Experimental results show that the proposed method significantly outperforms the state-of-the-art algorithms on both the BU-3DFE and Bosphorus databases.
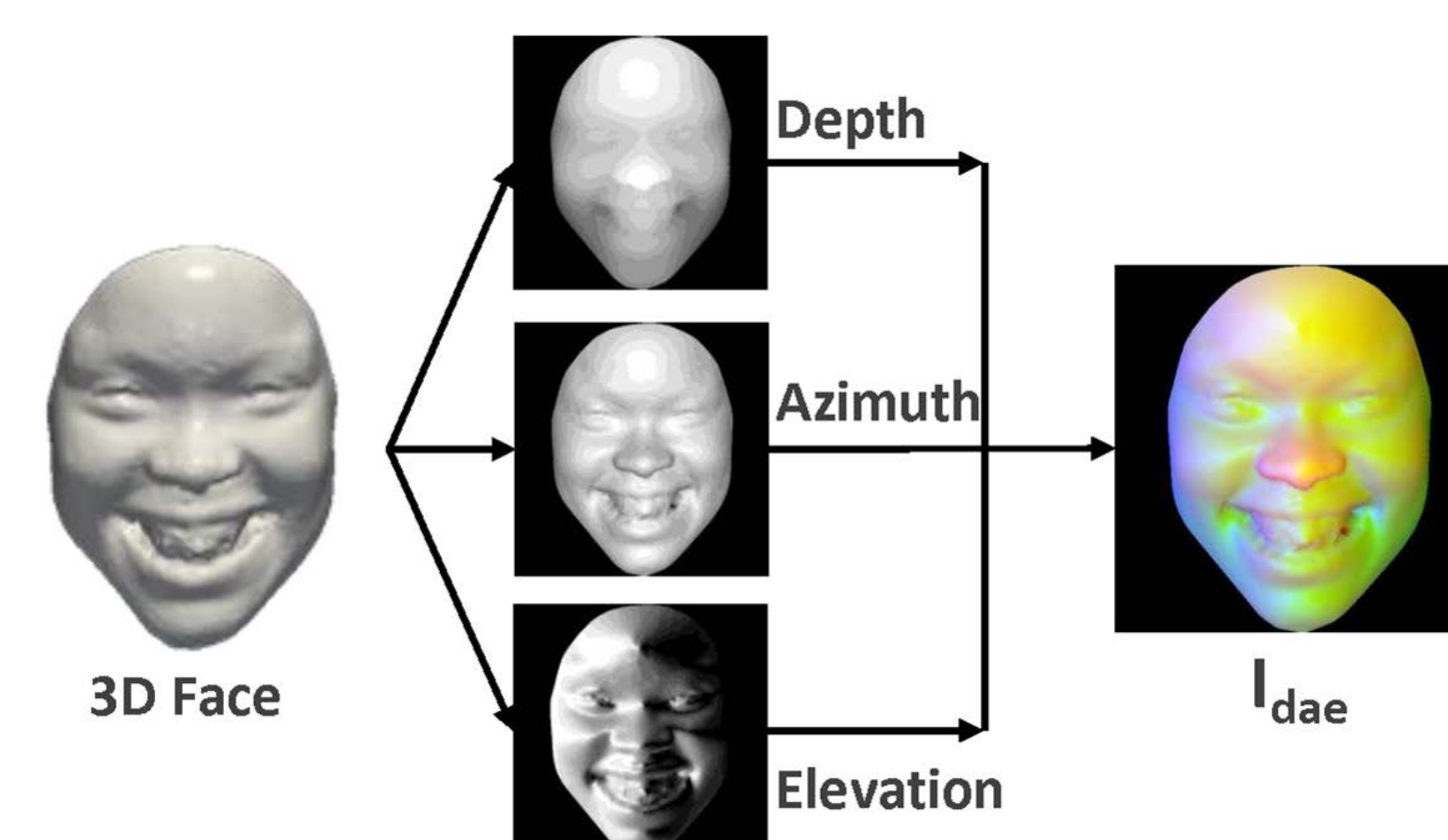
➤ Motivation

As 2D and 3D data present different views of the same face, the features extracted from them can be both complementary and redundant. The redundancy may have a negative effect on the FER, so it is necessary to reduce the redundancy before feature fusion.
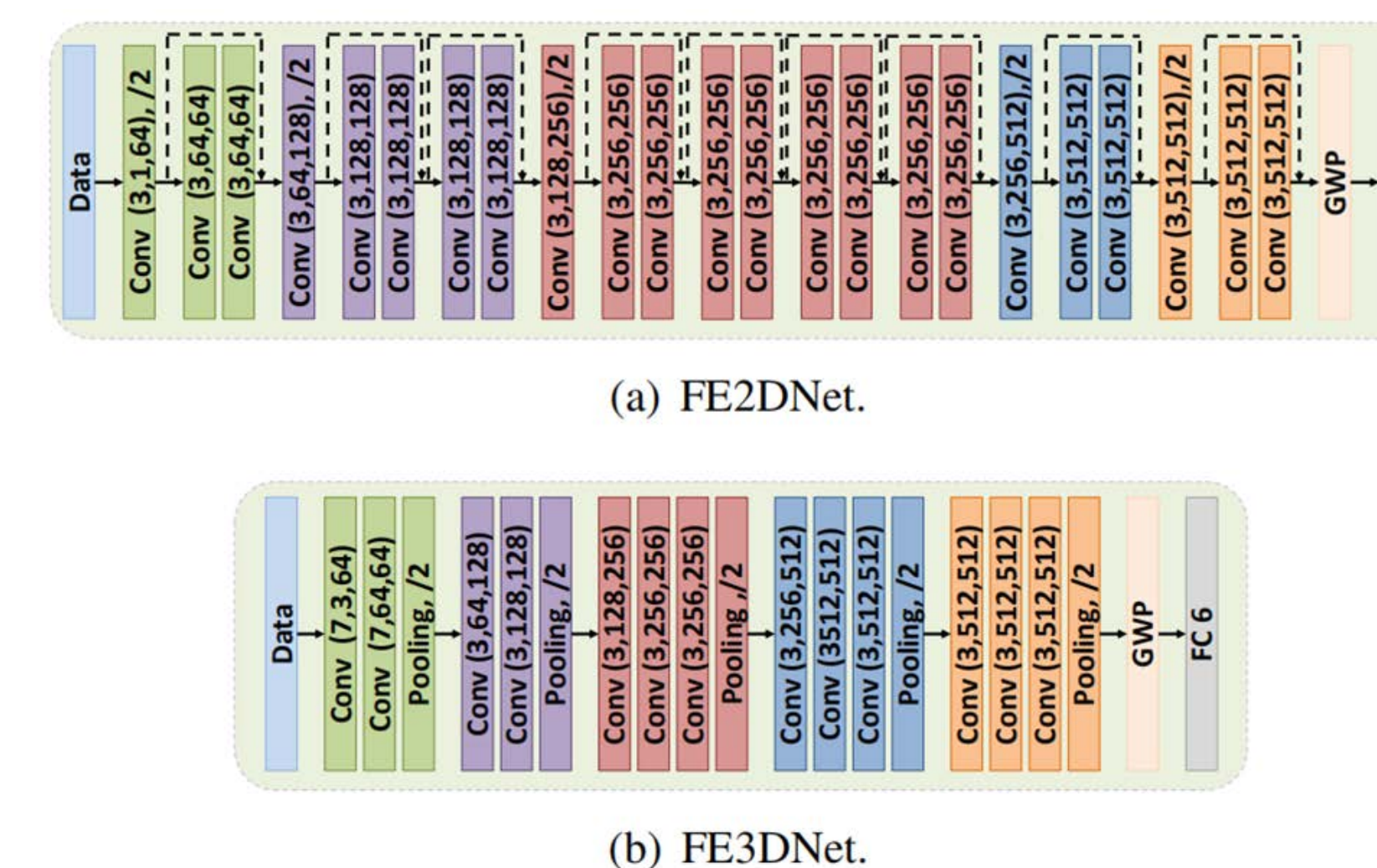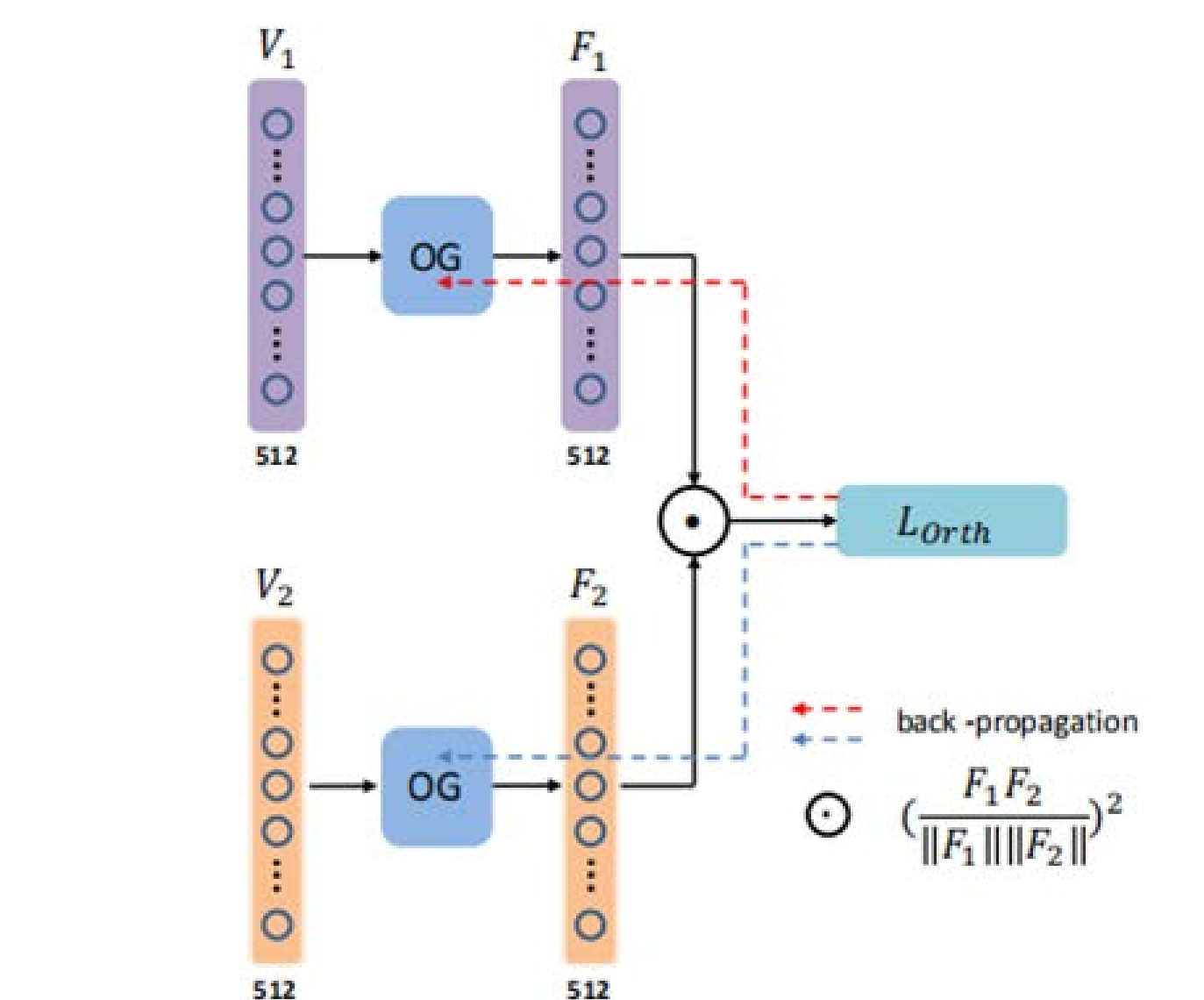
➤ Method

1) Overview
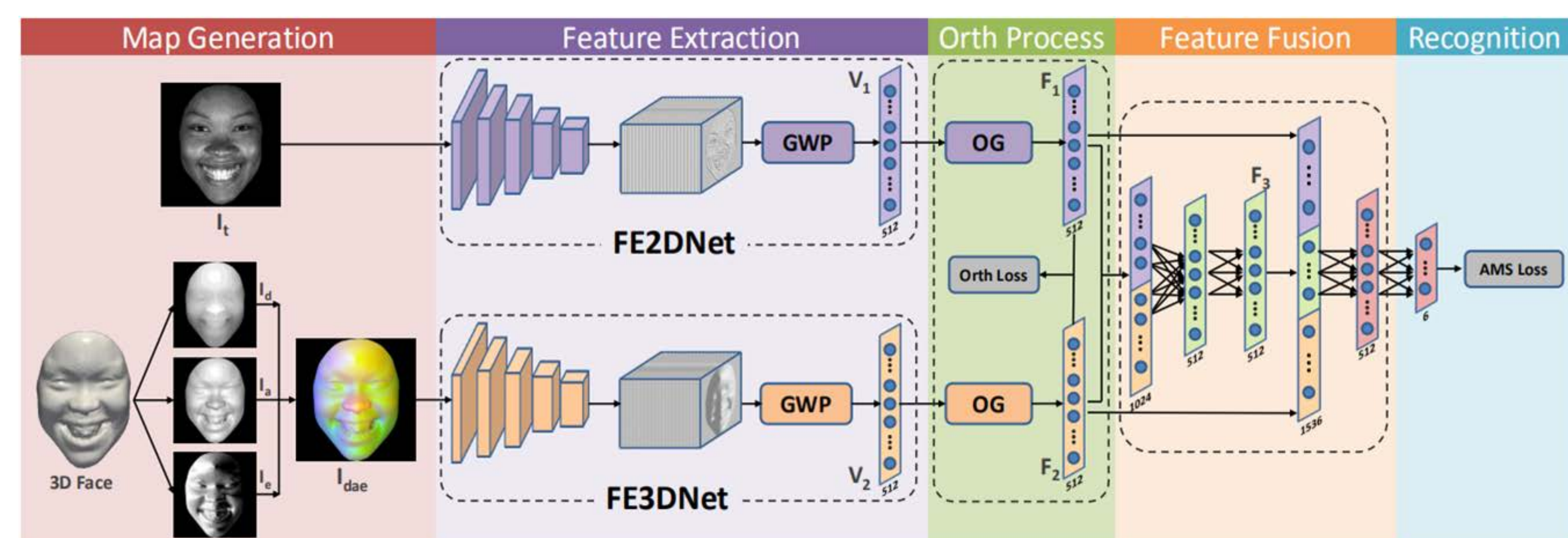
2) 3D Attribute Map Generation

3) Feature Extraction

(a) FE2DNet.

(b) FE3DNet.

4) Feature Orthogonalization

➤ Results

● The recognition rates on BU-3DFE

Approach	Feature	Accuracy(%)	
		P1	P2
Yang et al.(2015) [44]	Hand-crafted	84.80	82.73
Zhen et al.(2015) [15]	Hand-crafted	84.50	83.20
Li et al.(2015) [20]	Hand-crafted	86.32	-
Li et al.(2017) [10]	Deep learning	86.86	-
Chen et al.(2018) [18]	Deep learning	86.67	85.96
Wei et al.(2018) [45]	Deep learning	88.03	-
Jan et al.(2018) [22]	Deep learning	88.54	-
Zhu et al.(2019) [21]	Deep learning	88.35	87.06
FE3DNet	Deep learning	85.20	85.13
FE2DNet	Deep learning	86.58	86.69
OGF2Net	Deep learning	**89.05**	**89.07**

➤ Ablation Study

Model	Pre-trained Model	Vectorization Approach			Early Fusion	Concatenation of F_1 and F_2	Orth Loss	Parameters (MB)	Accuracy (%)
		Flatten	GAP	GWP					
FE2DNet	✓	✓						96.63	85.58
	✓		✓					71.63	86.16
				✓				-	84.86
				✓	✓			-	85.01
	✓			✓				71.69	**86.58**
FE3DNet	✓	✓						81.83	85.08
	✓		✓					56.83	83.01
				✓				-	81.09
				✓	✓			-	81.58
	✓			✓				56.88	**85.20**
OGF2Net	✓	✓				✓	✓	186.48	88.54
	✓		✓			✓	✓	136.48	88.59
	✓			✓		✓		-	86.79
	✓			✓				-	88.33
	✓			✓	✓			-	88.47
	✓			✓		✓	✓	136.59	**89.05**

➤ Conclusion

1) A novel orthogonalization-guided feature fusion network, namely OGF2Net, is proposed to extract and fuse complementary features for 2D+3D FER.

2) The orthogonal loss is proposed to reduce the correlation and redundancy among the features learned by FE2DNet and FE3DNet.

3) The global weighted pooling (GWP) module id utilized to vectorize the feature maps in FE2DNet and FE3DNet.