|  |  |
| --- | --- |
| 研究所/中心 | Computer Vision Institute |
| 成果（论文）名（英文） | Surrogate network-based sparseness hyper-parameter optimization for deep expression recognition |
| 年份 | 2021 |
| 刊物名（英文） | Pattern Recognition |
| 作者（多个作者按顺序用“,”隔开）（英文） | Weicheng Xie, Wenting Chen, Linlin Shen a, Jinming Duan, Meng Yang |
| 全部作者单位（英文） | School of Computer Science & Software Engineering, Shenzhen University, PR China, Shenzhen Institute of Artificial Intelligence and Robotics for Society, PR China, Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen University, PR China, School of Computer Science, University of Birmingham, UK, School of Data and Computer Science, Sun Yat-sen University, PR China |

Surrogate network-based sparseness hyper-parameter optimization for deep expression recognition

Pattern Recognition, 2021

Weicheng Xie a , b , c , Wenting Chen a , b , c , Linlin Shen a , b , c , ∗, Jinming Duan d , Meng Yang e

a School of Computer Science & Software Engineering, Shenzhen University, PR China

b Shenzhen Institute of Artificial Intelligence and Robotics for Society, PR China

c Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen University, PR China

d School of Computer Science, University of Birmingham, UK e School of Data and Computer Science, Sun Yat-sen University, PR China

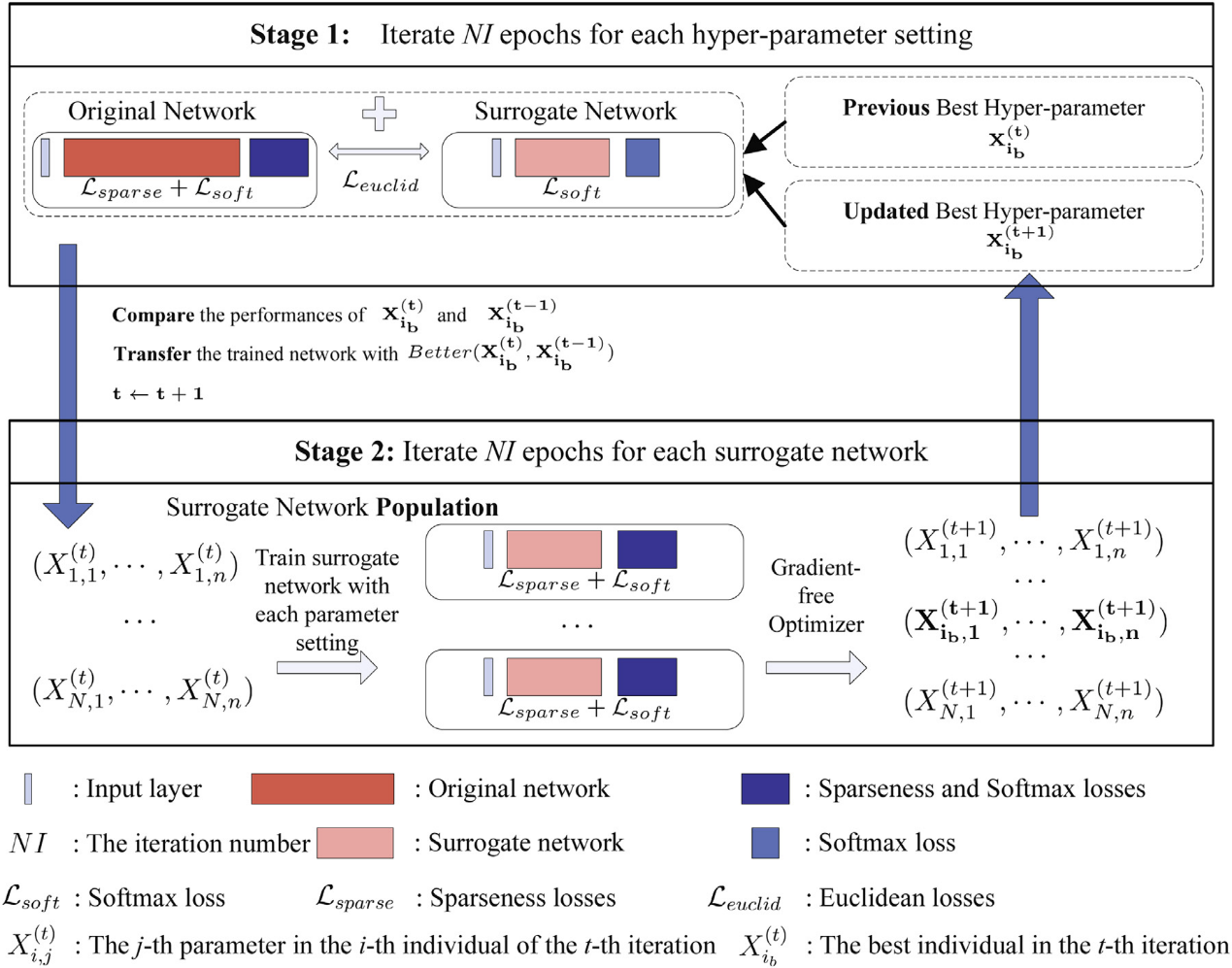


Fig. 1. The iterative framework of the proposed hyper-parameter optimization. The original network is trained based on the sparseness and softmax loss, where four Euclidean losses are employed for surrogate network fitting. The hyper-parameters are optimized with the gradient-free optimizers in the second stage and the candidate hyper-parameter setting X (t+1) i b is acquired for the training of fusion network in the next iteration.

ABSTRACT

For facial expression recognition, the sparseness constraints of the features or weights can improve the generalization ability of a deep network. However, the optimization of the hyper-parameters in fusing dif- ferent sparseness strategies demands much computation, when the traditional gradient-based algorithms are used. In this work, an iterative framework with surrogate network is proposed for the optimization of hyper-parameters in fusing different sparseness strategies. In each iteration, a network with signifi- cantly smaller model complexity is fitted to the original large network based on four Euclidean losses, where the hyper-parameters are optimized with heuristic optimizers. Since the surrogate network uses the same deep metrics and embeds the same hyper-parameters as the original network, the optimized hyper-parameters are then used for the training of the original deep network in the next iteration. While the performance of the proposed algorithm is justified with a tiny model, i.e. LeNet on the FER2013 database, our approach achieved competitive performances on six publicly available expression datasets, i.e., FER2013, CK+, Oulu-CASIA, MMI, AFEW and AffectNet.

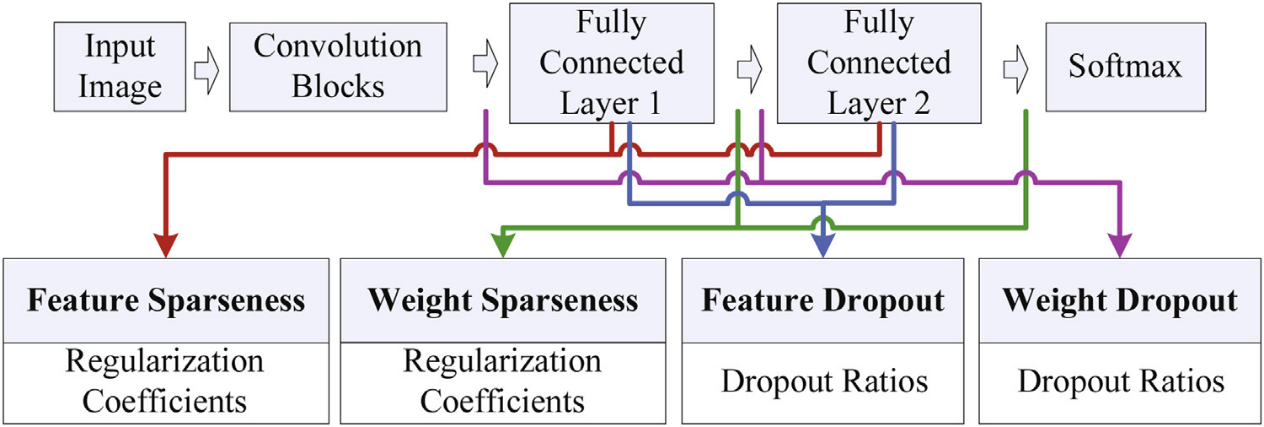


Fig. 2. Four sparseness strategies and their hyper-parameters

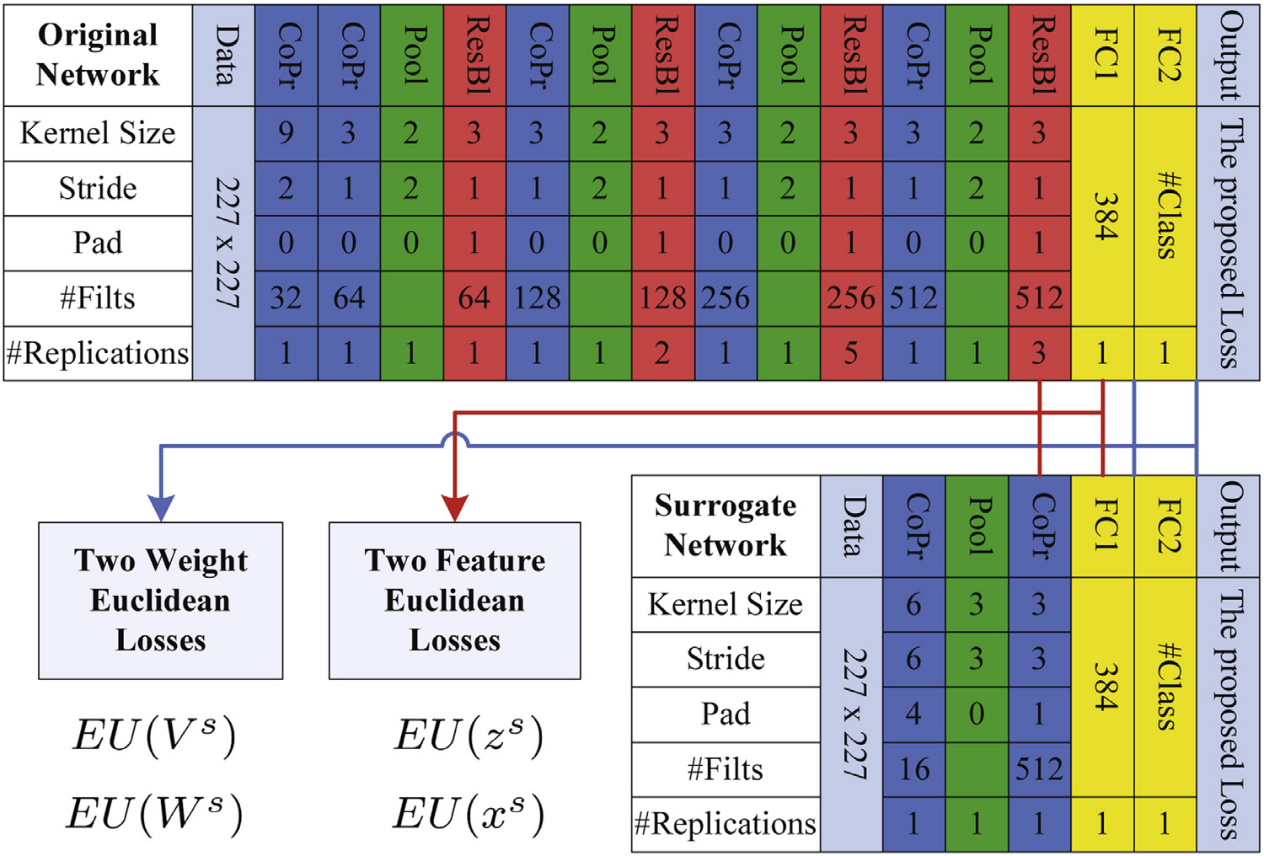


Fig. 3. The configuration of ‘original+surrogate’ network. CoPr denotes the convolutions followed by the PReLU activation function. Pool is the MaxPooling function. ResBl is a residual block with output ResOutput = PoolOutput + CoPr(CoPr(PoolOut put )) . # Replicat ions denotes the times that the same block is replicated. # F ilts denotes the number of feature maps. # class denotes the number of expression classes.



Fig. 4. Example images of the FER2013, CK+, Oulu-CASIA, MMI, AFEW and AffectNet databases.

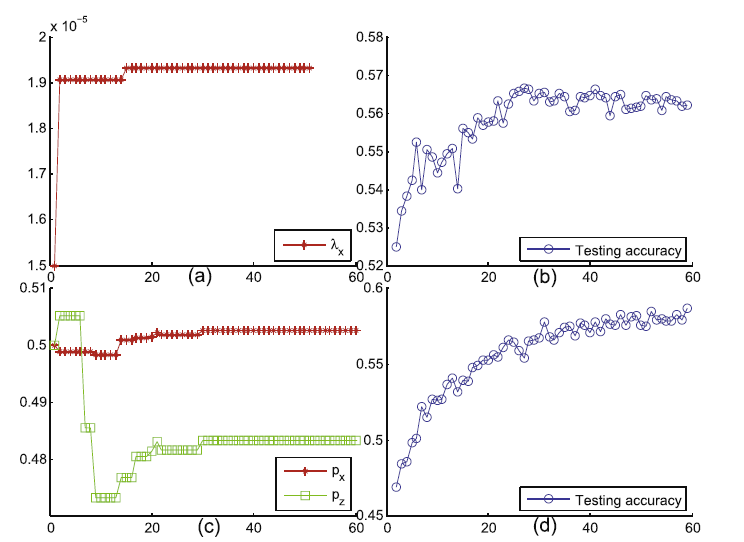


Fig. 5. The evolution curves of the hyper-parameters and the testing accuracies against the iteration epochs for PSO optimizer on the FER2013 database.

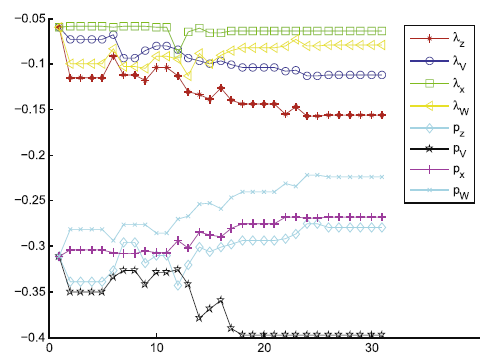


Fig. 6. The evolution curves of the eight hyper-parameters of LeNet optimized with PSO for the FER2013 database. Transformation functions of log 10 (·) / 100 and log 10 (·) are employed for the demonstration of the regularization coefficients and dropout ratios, respectively.

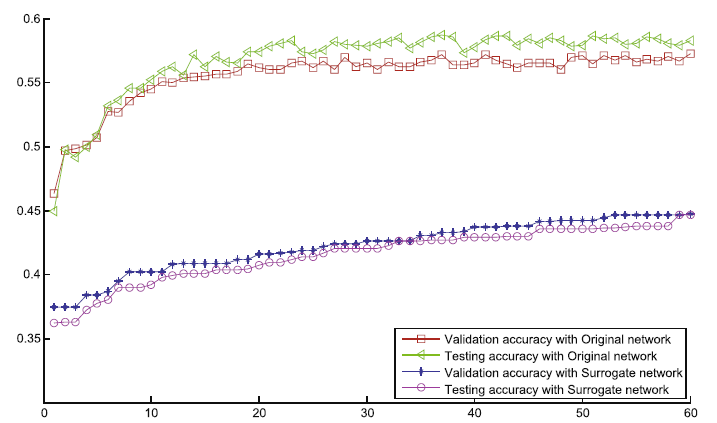


Fig. 7. The evolution curves of the validation and testing accuracies of the original LeNet and its surrogate network on the FER2013 database.

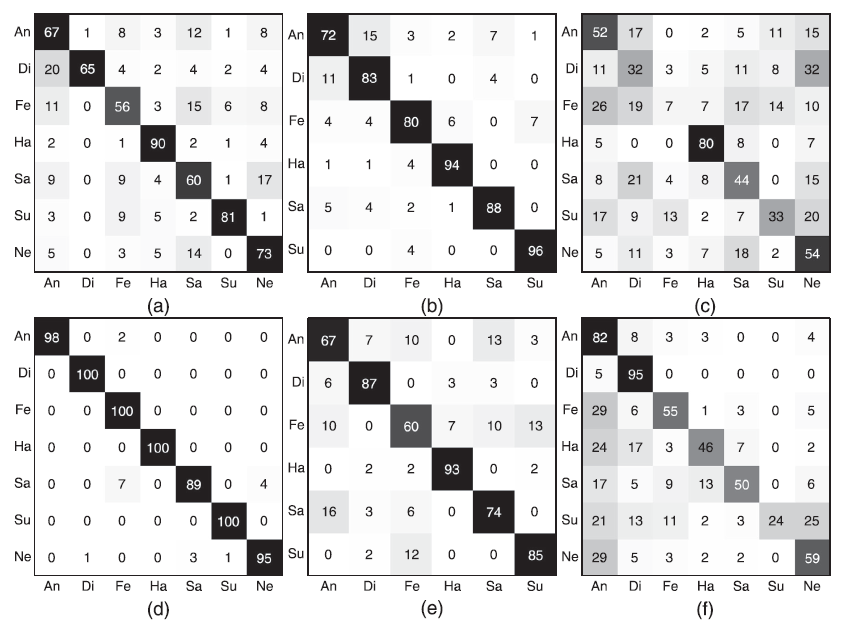


Fig. 8. The confusion matrices (%) of the proposed algorithm for the FER2013 (a), Oulu-CASIA (b), AFEW (c), CK+ (d), MMI (e) and AffectNet (f) databases.

ACKNOWLEDGMENTS

The authors thank the anonymous reviewers for their helpful comments and constructive suggestions. The authors thank Dr. Il- ija Ilievski in National University of Singapore for providing the source codes of the RBF-based surrogate [13]. The work was sup- ported by Natural Science Foundation of China under grants nos. 61602315 , 61672357 andU1713214 , the ScienceandTechnology Project of Guangdong Province under grant nos. 2020A1515010707 and 2018A050501014 , the Science and Technology Innovation Com- mission of Shenzhen under grant no. JCYJ20190808165203670.

Bibtex

@article{xie2021surrogate,

title={Surrogate network-based sparseness hyper-parameter optimization for deep expression recognition},

author={Xie, Weicheng and Chen, Wenting and Shen, Linlin and Duan, Jinming and Yang, Meng},

journal={Pattern Recognition},

volume={111},

pages={107701},

year={2021},

publisher={Elsevier}

}