30th National & 8th International Iranian Conference on
Biomedical Engineering
November 30 & December 1, 2023 - Sharif University of Technology – Department of Mechanical Engineering

Leveraging Self-Supervised Learning for Accurate Facial Keypoint Detection in Thermal Images

Poorya Aghaomidi[1], Sadaf Aram[2], Zahra Bahmani[3]
[1, 2, 3] Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran
[1] poorya_aghaomidi@modares.ac.ir, [2] sadaf_aram@modares.ac.ir, [3] z.bahmani@modares.ac.ir

Abstract

In the field of computer vision, precise keypoint detection is crucial for applications like face recognition and facial expression analysis. Thermal imaging offers untapped potential for mental state detection. We present a self-supervised approach for accurate facial keypoint detection in thermal images. Key contributions include three pretext tasks, high accuracies (100%, 97.92%, and 2.37% mean absolute percentage error), and a remarkable Normalized Mean Error (NME) of 1.05, demonstrating the efficacy of self-supervised learning in thermal facial analysis.

Introduction

Facial keypoint detection is essential in computer vision. Thermal imaging captures physiological indexes reflecting mental states. Accurate tracking is needed to extract this information. Our approach leverages self-supervised learning to address data scarcity and explores thermal facial keypoint detection's potential.

Proposed method

Dataset
We use a dataset by Kopaczka et al. from 2018. It consists of thermal images capturing facial expressions of 94 subjects under various conditions, providing manual annotations of 68 facial keypoints.

Self-supervised Pretraining
We pretrain a Convolutional Neural Network (CNN) with rotation prediction, subject classification, and 5-point keypoint detection tasks. This equips the model with knowledge about thermal facial images.

Downstream Task
We fine-tune the CNN for keypoint detection. The model efficiently detects facial keypoints in thermal videos with improved accuracy.

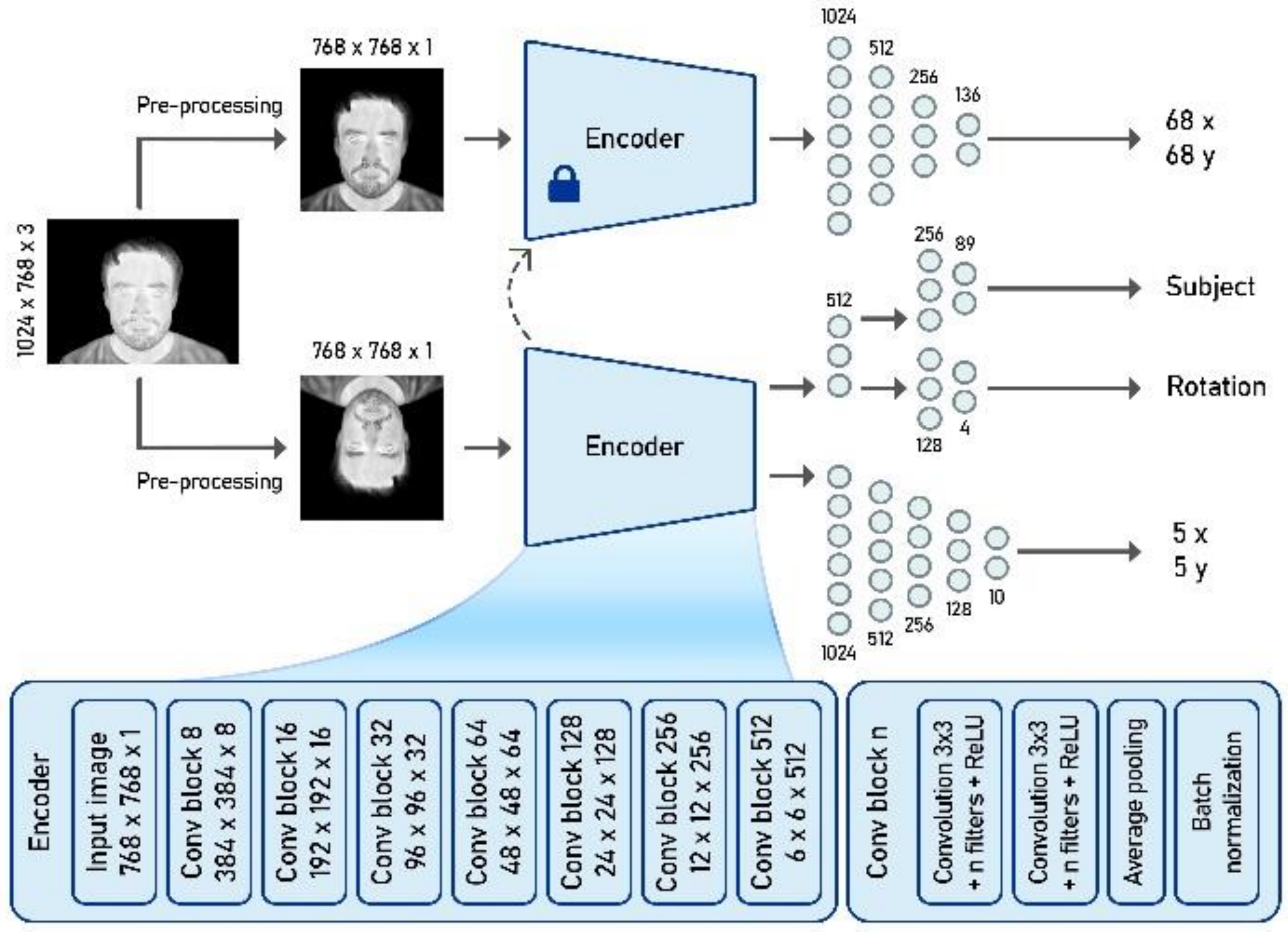

Fig. 1 architecture of the self-supervised thermal facial keypoint detection model

Facial Keypoint Detection Evaluation
We evaluate performance using Normalized Mean Error (NME), measuring the average Euclidean distance between predicted and ground truth coordinates, normalized by the interocular distance.

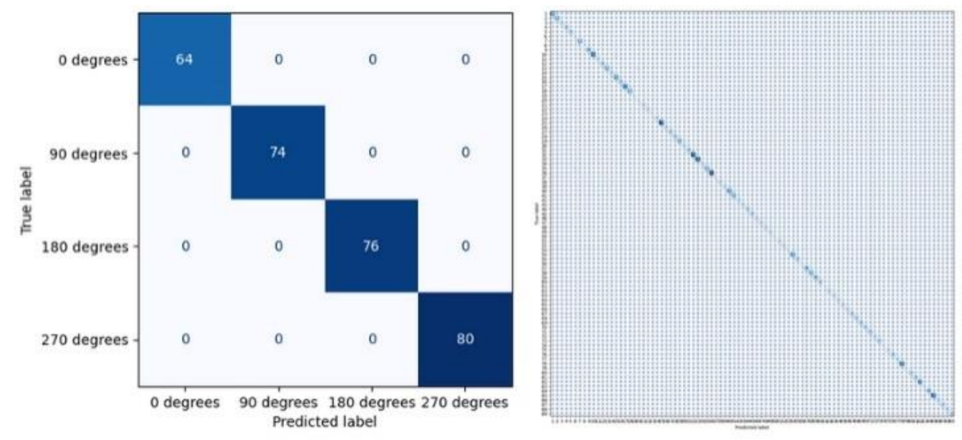

Fig. 2 The confusion matrix visually displays the outcomes of rotation classification (left) and subject classification (right).

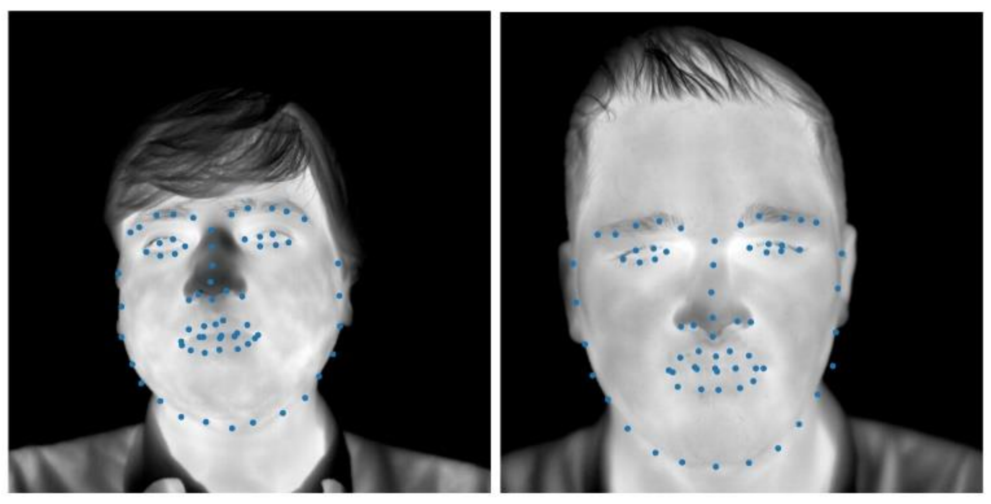

Fig. 3 an illustrative example of the outcomes achieved through our proposed downstream task.

- Pretext task results: CNN achieves 100% accuracy in rotation prediction, 97.92% accuracy in subject classification, and a mean absolute percentage error of 2.37% in 5-point keypoint detection.
- Downstream keypoint detection results: Self-supervised approach outperforms fully supervised learning and other methods, achieving an NME of 1.05.

Table 1 comparison of our method with other state of the art methods.

Method	NME	Dataset
AAM	5.20	TFL
DML	4.31	TFL
RCPR	3.87	AFLW
CCL	2.45	AFW
KEPLER	2.98	AFLW
OurOur Method	1.05	TFL

Conclusions

Our study pioneers self-supervised learning in thermal facial keypoint detection, overcoming data scarcity and revealing thermal imaging's potential. Self-supervised learning enhances keypoint detection accuracy, with implications in diverse fields. This approach offers valuable insights for future advancements in facial analysis and real-world applications.

References

1. M. Kopaczka, R. Kolk and D. Merhof, "A fully annotated thermal face database and its application for thermal facial expression recognition," 2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), pp. 1-6, 2018.
2. Kopaczka, M., Kolk, R., Schock, J., Burkhard, F., & Merhof, D., "A thermal infrared face database with facial landmarks and emotion labels," IEEE Transactions on Instrumentation and Measurement, pp. 1389-1401, 2018.
3. Han, S. Colaco and D. S., "Facial Keypoint Detection with Convolutional Neural Networks," 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Fukuoka, Japan, pp. 671-674, 2020.
4. Tian, L. Jing and Y., "Self-Supervised Visual Feature Learning With Deep Neural Networks: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 11, pp. 4037-4058, 2021.
5. W. -T. Chu and Y. -H. Liu, "Thermal Facial Landmark Detection by Deep Multi-Task Learning," 2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP), Kuala Lumpur, Malaysia, pp. 1-6, 2019.