Facial Emotion Recognition using CNN

Erfan Salehi, Ibtissam Saadi, Douglas W. Cunningham,
Department of Graphical Systems, Univ. BTU Cottbus-Senftenberg, Cottbus, Germany

INTRODUCTION

This project reimplements a CNN-based facial emotion recognition model using the FER2013 dataset, which is notably imbalanced, especially with very few samples for the "Disgust" class. To address this, the method applies inverse-frequency class-weighted loss, assigning higher weights to underrepresented classes during training while maintaining learning on frequent classes. The aim is to evaluate whether this weighting approach leads to measurable improvements in real-time emotion recognition accuracy.

RESEARCH QUESTION

Does using inverse-frequency class weights in the loss function improve facial emotion recognition accuracy compared to training without ?

METHODOLOGY

- FER2013 has 35,887 grayscale 48×48 images, resized, normalized, and augmented with flips and rotations.
- **Class weights use inverse frequency:**
Weight_k = Total Samples / (Num Classes × Samples_k) to emphasize rare classes.
- A CNN with 3 conv blocks uses ReLU, BatchNorm, and Dropout for effective feature learning.
- Weighted CrossEntropyLoss combined with Adam optimizer (lr=0.0001) handles imbalance and trains efficiently.
- Trained for 100 epochs with checkpointing : evaluated by accuracy, per-class accuracy, and confusion matrix.
- Best model saved

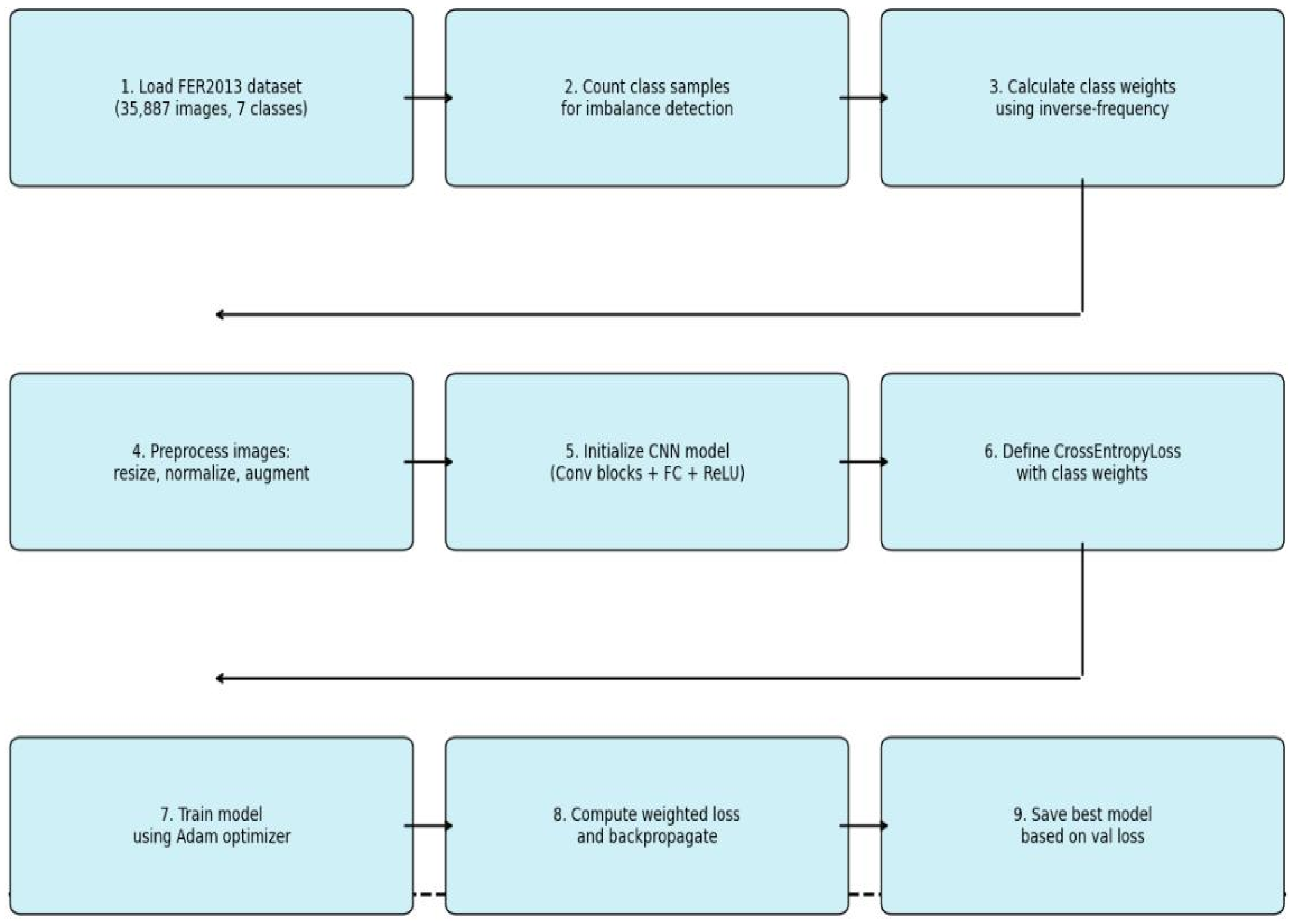

Figure 1: Step-by-step workflow of the facial emotion recognition method using class-weighted loss on the FER2013 dataset.

Figure 2. Class Weight Distribution Using Inverse Frequency Method

RESULT

This experiment compares **CNN** performance on the FER2013 dataset with and without **inverse class weighting**.

Boosting Rare Class Accuracy: Using class weights significantly improved both overall and per-class accuracy, with notable gains for underrepresented classes.. For example, the rare 'Disgust' class saw its precision improve from **41.9%** to **45.5%** indicating better recognition after addressing imbalance.

Balanced Emphasis Boosts Angry and Fear Recognition:
Angry and Fear showed strong improvement from **48.2%** and **47.7%** to **58.2%** and **59.1%** due to having a moderate amount of data combined with increased classweights.This balance gave them enough emphasis during training without overpowering the model, helping it generalize these emotions more effectively.

Boost Accuracy Across All Classes : Despite the focus on boosting performance for low-data classes, the class weighting strategy also led to gains across nearly all classes, improving total accuracy from **59.3%** to **65.0%** . This is because class weights redistributed learning emphasis proportionally, ensuring more balanced training.

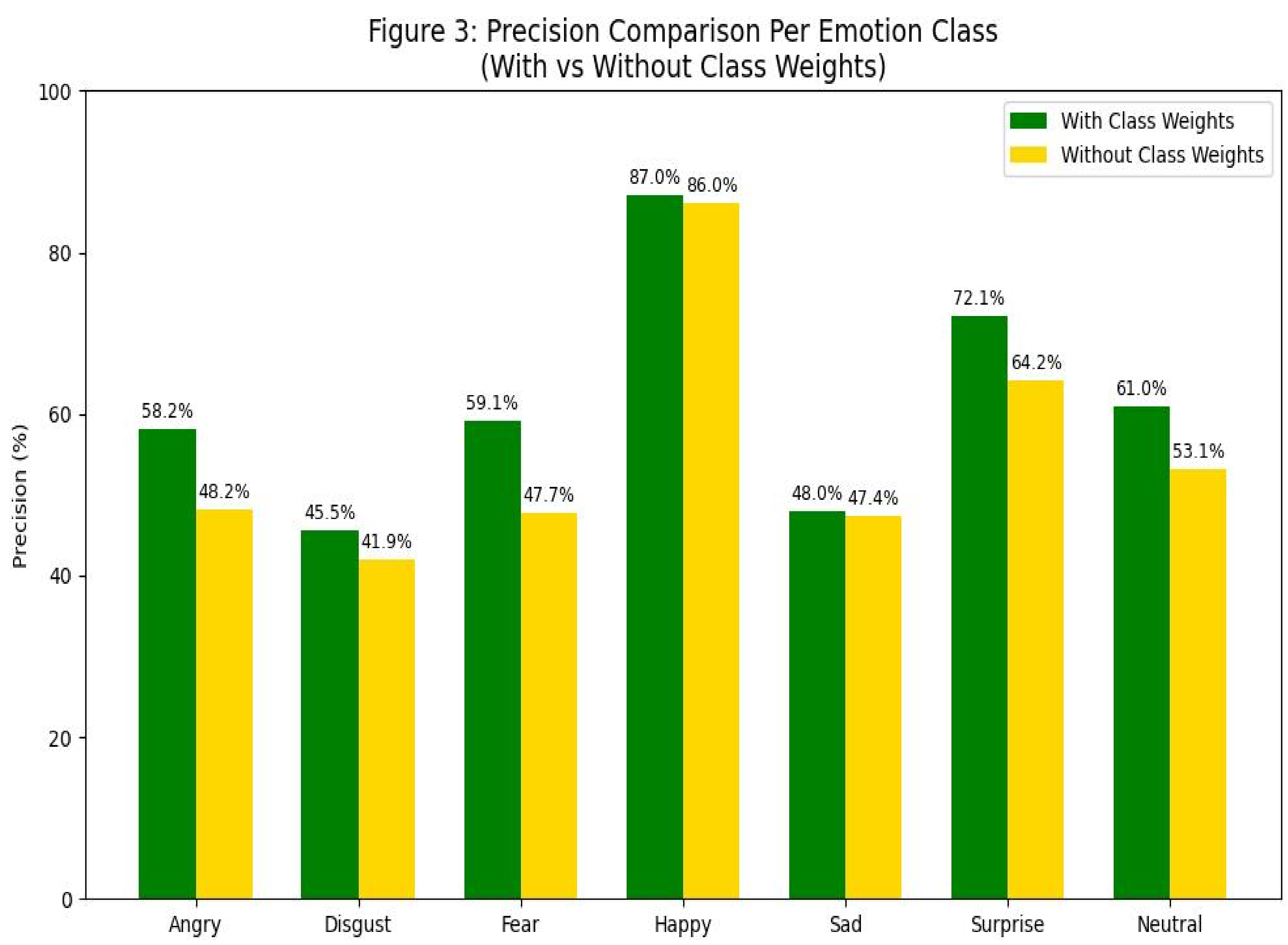

Figure 3 : precision comparison per Emotion Class (With and Without class Weights)

CONCLUSION

This study evaluated the impact of inverse class weighting on CNN performance for facial emotion recognition using the FER2013 dataset. The results demonstrate that applying class weights significantly boosts both overall and per-class accuracy, especially for underrepresented emotions such as Disgust, Angry, and Fear. This confirms the effectiveness of class balancing in improving model generalization across all emotion categories.

Key takeaways:
Inverse class weighting improves recognition accuracy for rare classes without compromising overall performance.
Moderately represented classes benefit from balanced emphasis, enhancing generalization.
This method is effective in datasets with class imbalance, offering a practical solution for more equitable and accurate emotion recognition.

References :
[1] Ngo, Q., & Yoon, S. (2020). Facial expression recognition based on weighted-cluster loss and deep transfer learning using a highly imbalanced dataset. Sensors, 20(9), 2639.https://doi.org/10.3390/s20092639
[2] Mehendale, Ninad. "Facial emotion recognition using convolutional neural networks (FERC)." SN Applied Sciences 2.3 (2020): 446.

contact : salehierf@b-tu.de