Comparison of YOLO-v8 and YOLO-v10 in Detecting Human Facial Emotions

Guilliano Rasyid[1]*, Joko Sutopo[2]

[12]Universitas Teknologi Yogyakarta

DOI: https://doi.org/10.47134/ijat.v2i1.3395
*Correspondence: Guilliano Rasyid
Email: rasyidguilliano@gmail.com

Received: 02-11-2024
Accepted: 10-12-2024
Published: 05-01-2025

Abstract: This study evaluates the performance of YOLOv8 and YOLOv10 in recognizing human facial emotions. Both state-of-the-art object detection models were trained on a diverse dataset of facial expressions. While YOLOv10 demonstrated superior performance in certain metrics, it required significantly more training time compared to YOLOv8. Both models exhibited effective learning, as evidenced by the steady decrease in training loss. However, both models encountered challenges in accurately recognizing subtle emotions, such as disgust and contempt. To enhance the accuracy and robustness of facial emotion recognition systems, future research should prioritize improving data quality, exploring advanced model architectures, and optimizing hyperparameters.

Keywords: YOLO-v8, YOLO-v10, Facial Emotion Recognition

Introduction

Detecting facial emotions has become a pivotal area of research, particularly in fields like human-computer interaction, security, and healthcare. State-of-the-art object detection algorithms, such as the YOLO (You Only Look Once) series, have significantly advanced the accuracy and efficiency of emotion recognition systems. YOLO's ability to process images in real-time makes it an ideal choice for applications demanding immediate feedback on emotional states(Hasan, 2023; RamaKrishna, 2024)

In recent years, the YOLO framework has evolved rapidly, with YOLOv8 and YOLOv10 representing the latest advancements. These models have pushed the boundaries of object detection and recognition, offering superior performance in terms of accuracy and speed. By incorporating innovative techniques like advanced backbone networks, attention mechanisms, and improved loss functions, YOLOv10 has demonstrated significant improvements over its predecessor, YOLOv8(Amor et al., 2023; Chaitanya, 2023).

This study aims to delve into the capabilities of YOLOv8 and YOLOv10 in the context of facial emotion detection. We will conduct a comparative analysis to evaluate their performance on various datasets and under different lighting conditions. By understanding the strengths and weaknesses of these models, we can identify potential areas for future research and development in the field of emotion recognition(Hasan & Lazem, 2023; Ting et al., 2024).

Methodology

This study utilized a facial emotion dataset sourced from RoboFlow(Emotions dectetion, 2024). The dataset comprises 9400 images categorized into eight distinct emotions: anger, contempt, disgust, fear, happiness, neutral, sadness and surprise in Figure 1.

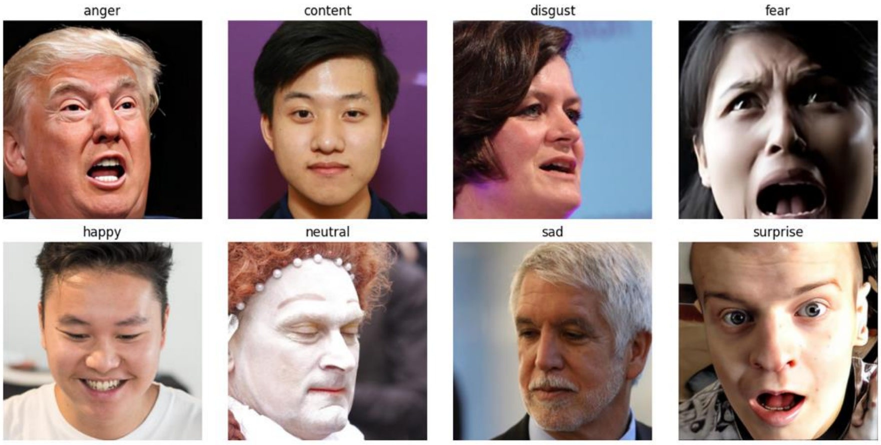


Figure 1. Dataset Categorize

To ensure robust model training and evaluation, the dataset was divided into three subsets: a training set of 3000 images, a validation set of 1000 images, and a testing set of 500 images in Figure 2.

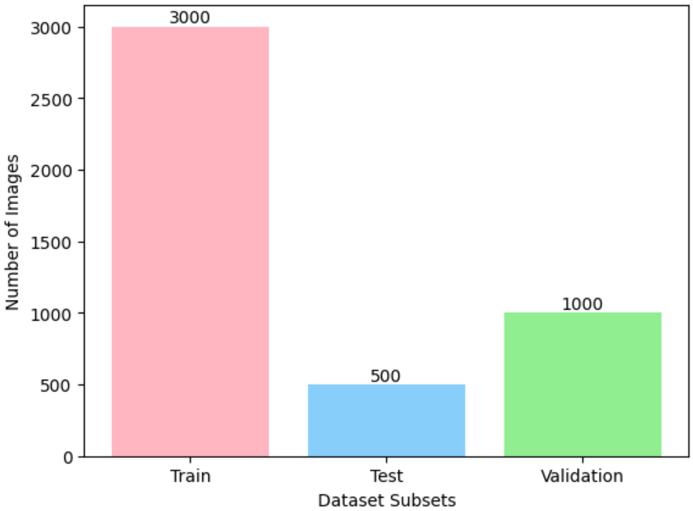
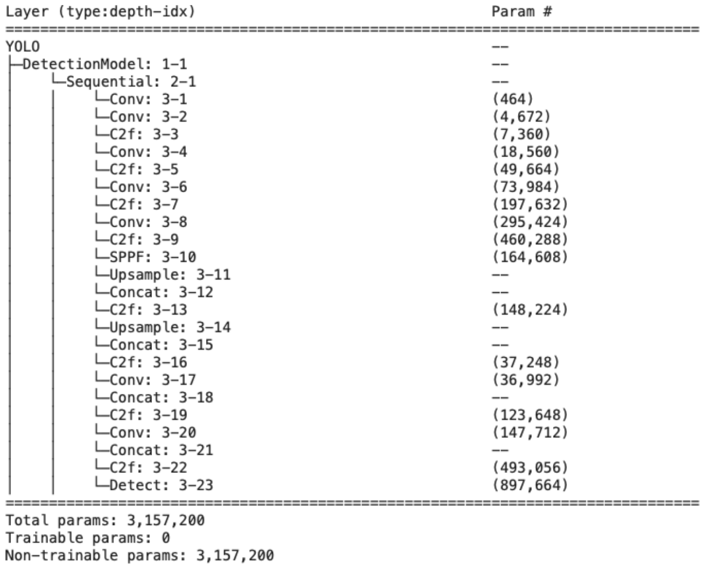


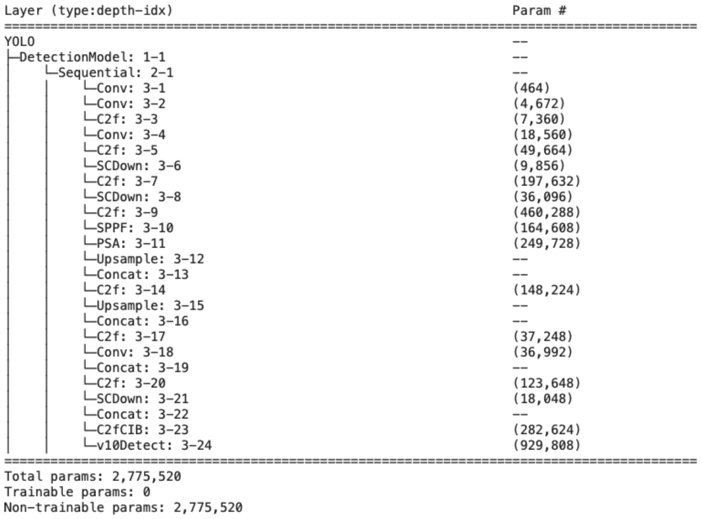
Figure 2. Dataset Subsets Plot

The You Only Look Once (YOLO) family of algorithms, specifically YOLOv8 and YOLOv10, were employed for facial emotion detection. These state-of-the-art object detection models excel in real-time performance and accuracy, making them ideal for applications like human-computer interaction and behavioural analysis(Alshammari & Alshammari, 2024). YOLOv8 and YOLOv10 are known for their ability to simultaneously predict multiple bounding boxes and their respective class probabilities within a single forward pass of the neural network. This allows for efficient and rapid processing of images and videos(Parambil et al., 2024).



```
Layer (type:depth-idx)                                       Param #
================================================================================
YOLO                                                         --
├─DetectionModel: 1-1                                        --
│    └─Sequential: 2-1                                       --
│    │    └─Conv: 3-1                                        (464)
│    │    └─Conv: 3-2                                        (4,672)
│    │    └─C2f: 3-3                                         (7,360)
│    │    └─Conv: 3-4                                        (18,560)
│    │    └─C2f: 3-5                                         (49,664)
│    │    └─Conv: 3-6                                        (73,984)
│    │    └─C2f: 3-7                                         (197,632)
│    │    └─Conv: 3-8                                        (295,424)
│    │    └─C2f: 3-9                                         (460,288)
│    │    └─SPPF: 3-10                                       (164,608)
│    │    └─Upsample: 3-11                                   --
│    │    └─Concat: 3-12                                     --
│    │    └─C2f: 3-13                                        (148,224)
│    │    └─Upsample: 3-14                                   --
│    │    └─Concat: 3-15                                     --
│    │    └─C2f: 3-16                                        (37,248)
│    │    └─Conv: 3-17                                       (36,992)
│    │    └─Concat: 3-18                                     --
│    │    └─C2f: 3-19                                        (123,648)
│    │    └─Conv: 3-20                                       (147,712)
│    │    └─Concat: 3-21                                     --
│    │    └─C2f: 3-22                                        (493,056)
│    │    └─Detect: 3-23                                     (897,664)
================================================================================
Total params: 3,157,200
Trainable params: 0
Non-trainable params: 3,157,200
```

Figure 3. YOLO-v8 Architecture



```
Layer (type:depth-idx)                                       Param #
================================================================================
YOLO                                                         --
├─DetectionModel: 1-1                                        --
│    └─Sequential: 2-1                                       --
│    │    └─Conv: 3-1                                        (464)
│    │    └─Conv: 3-2                                        (4,672)
│    │    └─C2f: 3-3                                         (7,360)
│    │    └─Conv: 3-4                                        (18,560)
│    │    └─C2f: 3-5                                         (49,664)
│    │    └─SCDown: 3-6                                      (9,856)
│    │    └─C2f: 3-7                                         (197,632)
│    │    └─SCDown: 3-8                                      (36,096)
│    │    └─C2f: 3-9                                         (460,288)
│    │    └─SPPF: 3-10                                       (164,608)
│    │    └─PSA: 3-11                                        (249,728)
│    │    └─Upsample: 3-12                                   --
│    │    └─Concat: 3-13                                     --
│    │    └─C2f: 3-14                                        (148,224)
│    │    └─Upsample: 3-15                                   --
│    │    └─Concat: 3-16                                     --
│    │    └─C2f: 3-17                                        (37,248)
│    │    └─Conv: 3-18                                       (36,992)
│    │    └─Concat: 3-19                                     --
│    │    └─C2f: 3-20                                        (123,648)
│    │    └─SCDown: 3-21                                     (18,048)
│    │    └─Concat: 3-22                                     --
│    │    └─C2fCIB: 3-23                                     (282,624)
│    │    └─v10Detect: 3-24                                  (929,808)
================================================================================
Total params: 2,775,520
Trainable params: 0
Non-trainable params: 2,775,520
```

Figure 4. YOLO-v10 Architecture

YOLOv8, depicted in Figure 3, and YOLOv10, shown in Figure 4, differ in their architectural designs, with YOLOv10 incorporating advancements in various components to achieve superior performance. YOLOv10 utilizes more sophisticated backbone networks, which are responsible for extracting meaningful features from the input images. These backbone networks often employ deeper and wider architectures, allowing for the capture of intricate details that are crucial for accurate emotion recognition(Vanamoju et al., 2024).

Additionally, YOLOv10 features enhanced neck architectures, which play a vital role in fusing information from different layers of the network. These neck architectures enable the model to effectively integrate low-level and high-level features, resulting in more robust and informative representations. Finally, YOLOv10 incorporates refined head designs, which are responsible for generating the final predictions, including bounding box coordinates and class probabilities. These head designs often employ attention mechanisms or other advanced techniques to improve the accuracy and localization of the detected faces and their associated emotions(Huang et al., 2024).

Overall, the combination of these architectural enhancements in YOLOv10 allows for more precise facial emotion detection, even in challenging conditions such as variations in lighting, pose, and occlusion. This makes YOLOv10 a powerful tool for a wide range of applications that require real-time and accurate emotion analysis(Aina et al., 2024).

To initiate our facial emotion classification project, we utilized Google Colab, a versatile platform well-suited for intricate machine learning tasks. The dataset, sourced from RoboFlow(Emotions dectetion, 2024) a renowned repository for labeled visual data, comprised images categorized into eight distinct emotional states. To ensure rigorous model training and evaluation, the dataset was divided into training, testing, and validation sets.

For the implementation of YOLOv8 and YOLOv10 we meticulously prepared the dataset by resizing images to a uniform 640 pixels and setting a batch size of 64. The models were trained for 10 epochs using the AdamW optimizer with a learning rate of 0.000833.

To assess the performance of YOLOv8 and YOLOv10 in recognizing facial emotions, we conducted a comparative analysis within the Google Colab environment. By leveraging these powerful models, we aimed to develop a robust system capable of accurately interpreting human emotions.

Result and Discussion

To evaluate the performance of YOLOv8 and YOLOv10 in recognizing human facial emotions, we trained both models on a dataset containing images of faces with various emotional expressions. The training process was monitored, and the training time for each epoch was recorded. Table 1 presents the training time for each epoch of both YOLOv8 and YOLOv10 models.

Table 1: Training Time Comparison of YOLO-v8 and YOLO-v10

Epoch	YOLO-v8	YOLO-v10
1	73.243	113.021
2	133.975	179.049
3	196.404	243.724
4	258.319	308.991
5	319.239	373.397
6	380.211	439.513
7	443.07	503.073
8	504.359	569.97
9	565.596	633.573
10	627.879	699.329

As shown in Table 1, YOLOv10 consistently takes longer to train than YOLOv8 for each epoch. This difference in training time could be attributed to the increased complexity of YOLOv10's architecture or the larger number of parameters it has to learn.

During the training process, we monitored the training loss values for three key components: box loss, classification loss, and detection loss (dfl loss). Table 2 and 3 present the loss values for each epoch of YOLOv8 and YOLOv10, respectively.

Table 2: YOLO-v8 Training Loss Value

Box Loss	Classification Loss	Detection Loss
0.89071	3.67579	1.64507
0.51206	2.66688	1.23607
0.51847	2.21863	1.22811
0.50604	1.85809	1.21575
0.47892	1.62388	1.18286
0.44708	1.45225	1.15633
0.41874	1.35986	1.12565
0.39857	1.24374	1.10776
0.37487	1.15594	1.07741
0.359881	1.11257	1.0668

Table 3: YOLO-v10 Training Loss Value

Box Loss	Classification Loss	Detection Loss
1.47398	14.8684	3.05956
1.0362	11.2693	2.48401
1.07833	8.61814	2.50643
1.03774	6.09714	2.45479
0.97179	4.65911	2.38637
0.91387	3.89408	2.32878
0.84026	3.44994	2.258
0.79867	3.12425	2.20719
0.7557	2.8799	2.17484
0.71273	2.70736	2.13855

Both models showed a steady decrease in training loss, indicating effective learning. YOLOv8 generally had slightly lower training loss values across all three components compared to YOLOv10, suggesting faster convergence or better fitting to the training data.

During the training process, we monitored several key performance metrics, including precision, recall, and mean Average Precision (mAP) at different Intersection over Union (IoU) thresholds. Table 4 and 5 present the loss values for each epoch of YOLOv8 and YOLOv10, respectively.

Table 4: YOLO-v8 Performance Metrics

Precision	Recall	mAP50	mAP50-95
0.00396	0.99606	0.18323	0.13536
0.26226	0.4265	0.30504	0.24788
0.24702	0.65025	0.34495	0.2628
0.31104	0.61529	0.38285	0.30137
0.3334	0.70579	0.4319	0.36542
0.37698	0.70078	0.48881	0.41916
0.40537	0.66932	0.52427	0.46189
0.50975	0.65275	0.57718	0.51678
0.51436	0.64211	0.58887	0.52895
0.54422	0.62356	0.60551	0.54819

Table 5: YOLO-v10 Performance Metrics

Precision	Recall	mAP50	mAP50-95
0.00471	0.99419	0.13615	0.11619
0.18776	0.40647	0.13128	0.10851
0.23033	0.2342	0.19026	0.15806
0.2125	0.40462	0.26384	0.22124
0.29865	0.55304	0.32439	0.2725
0.33815	0.55223	0.39248	0.34889
0.36581	0.56641	0.41456	0.36684
0.40815	0.6006	0.4859	0.43133
0.47875	0.60664	0.53983	0.48581
0.47701	0.59619	0.53068	0.47929

Both models showed an increasing trend in precision and recall, indicating improved accuracy in detecting and classifying facial emotions. YOLOv8 generally maintained higher precision and recall values. Both models exhibited an increasing trend in mAP values, indicating overall performance improvement. YOLOv8 consistently outperformed YOLOv10 in terms of mAP, suggesting better accuracy in detecting objects with various overlap levels.

During the validation phase, we monitored the loss values for three key components: box loss, classification loss, and detection loss (dfl loss). Tables 6 and 7 present the validation loss values for each epoch of YOLOv8 and YOLOv10, respectively.

Table 6: YOLO-v8 Validation Loss Value

Epoch	Box Loss	Classification Loss	Detection Loss
1	0.72184	2.57462	1.49967
2	0.63951	2.07803	1.36846
3	0.75551	1.79613	1.52415
4	0.735	1.70214	1.52661
5	0.61364	1.57605	1.30868
6	0.5821	1.32357	1.2491
7	0.52364	1.30406	1.19483
8	0.497	1.21049	1.14795
9	0.47531	1.14969	1.12509
10	0.46127	1.11145	1.11085

Table 7: YOLO-v10 Validation Loss Value

Epoch	Box Loss	Classification Loss	Detection Loss
1	0.98175	6.12706	2.48984
2	1.28094	5.41928	2.91543
3	1.33216	4.18497	3.04428
4	1.32826	3.63148	2.9884
5	1.27995	3.78083	2.75739
6	1.07291	2.78775	2.436
7	1.028	2.73553	2.33725
8	1.00881	2.47448	2.31562
9	0.93744	2.29808	2.23458
10	0.91608	2.27498	2.21389

By analysing the validation loss values, we can observe that YOLOv8 generally exhibits lower loss values across all three components compared to YOLOv10. This suggests that YOLOv8 generalizes better to unseen data and has a lower tendency to overfit the training data.

To evaluate the training process of YOLOv8 and YOLOv10 in recognizing human facial emotions, we monitored the learning rate (LR) for three parameter groups (pg0, pg1, pg2) during training. The LR is a crucial hyperparameter that controls the step size of the optimization algorithm. Tables 8 and 9 present the learning rate values for each parameter group of YOLOv8 and YOLOv10, respectively.

Table 8: YOLO-v8 Learning Rates

Pg0	Pg1	Pg2
0.000271759	0.000271759	0.000271759
0.000536115	0.000536115	0.000536115
0.000748902	0.000748902	0.000748902
0.000663029	0.000663029	0.000663029
0.000548084	0.000548084	0.000548084
0.000420665	0.000420665	0.000420665
0.000293246	0.000293246	0.000293246
0.000178301	0.000178301	0.000178301
8.71E-05	8.71E-05	8.71E-05
2.85E-05	2.85E-05	2.85E-05

Table 9: YOLO-v10 Learning Rates

Pg0	Pg1	Pg2
0.000271759	0.000271759	0.000271759
0.000536115	0.000536115	0.000536115
0.000748902	0.000748902	0.000748902
0.000663029	0.000663029	0.000663029
0.000548084	0.000548084	0.000548084
0.000420665	0.000420665	0.000420665
0.000293246	0.000293246	0.000293246
0.000178301	0.000178301	0.000178301
8.71E-05	8.71E-05	8.71E-05
2.85E-05	2.85E-05	2.85E-05

By analysing the learning rate schedules, we can observe that both YOLOv8 and YOLOv10 follow a similar learning rate decay strategy. The learning rate is initially increased to accelerate the training process, and then gradually decreased to fine-tune the model. This strategy helps to balance exploration and exploitation during training.

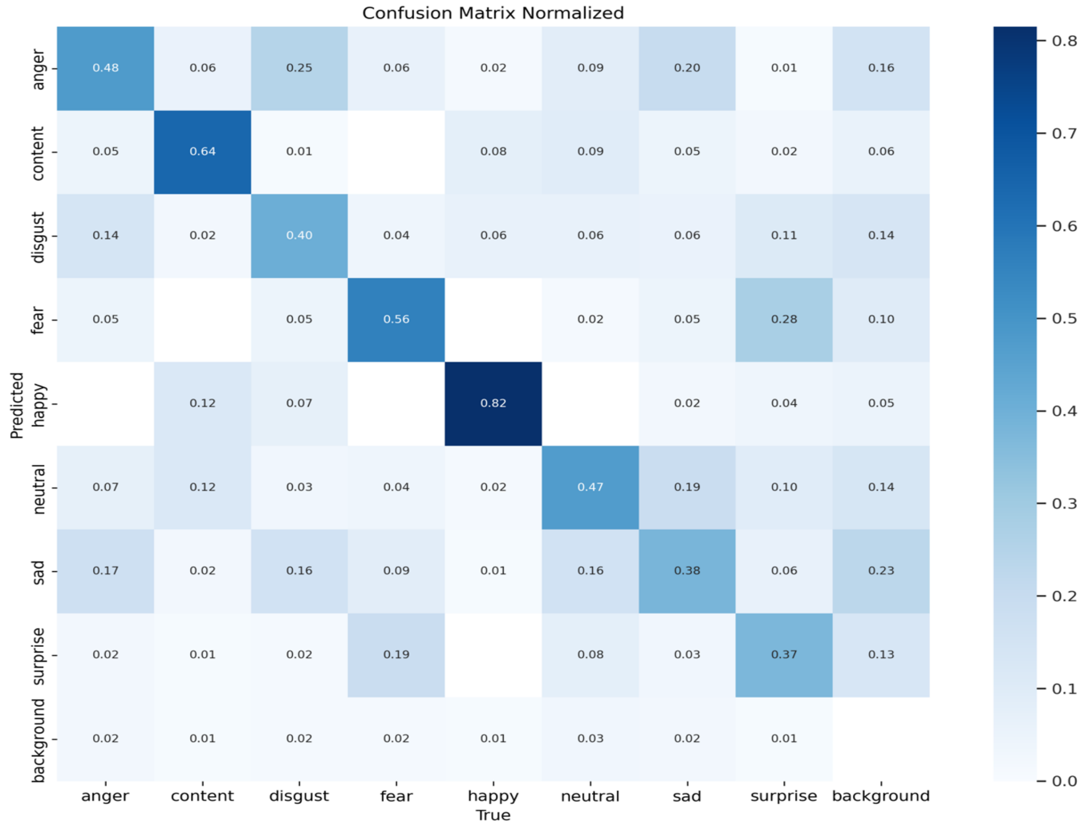


Figure 5. YOLO-v8 Normalized Confusion Matrix

The normalized confusion matrix in Figure 5 provides a visual representation of the model's predictions against the true labels. The diagonal elements represent correct predictions, while off-diagonal elements indicate misclassifications.

Based on the matrix, YOLOv8 seems to perform reasonably well in recognizing several emotions like happy, neutral, and disgust. However, it struggles with emotions like anger, content, and surprise.

Specific Observations:

a. Happy: The model is quite accurate at recognizing happy expressions, with a high diagonal value and relatively low off-diagonal values.

b. Neutral: The model also performs well in identifying neutral expressions.

c. Disgust: The model shows decent performance in recognizing disgust, although there's room for improvement.

d. Anger, Content, and Surprise: The model struggles with these emotions, as evidenced by the lower diagonal values and higher off-diagonal values in the corresponding rows and columns. This indicates that the model often misclassifies these emotions.

Possible Reasons for Misclassifications:

a. Data Quality: The quality and quantity of the training data can significantly impact the model's performance. If the dataset lacks sufficient examples of certain emotions, the model may struggle to learn their distinguishing features.

b. Class Imbalance: If the dataset is imbalanced, with some emotions being underrepresented, the model may be biased towards the majority classes.

c. Model Complexity**:** The complexity of the model architecture might not be sufficient to capture the subtle nuances between similar emotions.

d. Hyperparameter Tuning**:** The choice of hyperparameters, such as learning rate and batch size, can influence the model's performance.

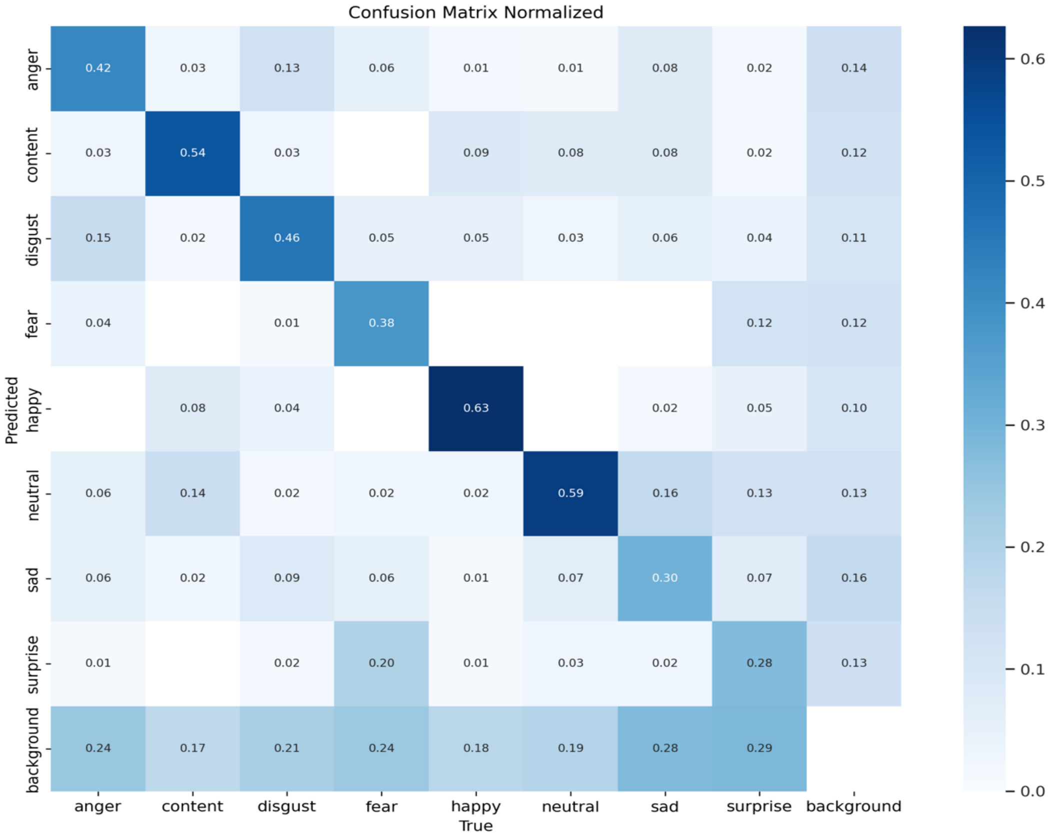


Figure 6. YOLO-v8 Normalized Confusion Matrix

Similar to the YOLOv8 confusion matrix, the matrix in Figure 6 also provides a visual representation of the model's predictions against the true labels. The diagonal elements indicate correct predictions, while off-diagonal elements represent misclassifications.

YOLOv10 shows a similar performance to YOLOv8, with some strengths and weaknesses. It performs well in recognizing happy, neutral, and disgust expressions. However, it struggles with anger, content, and surprise, as seen by the lower diagonal values and higher off-diagonal values in the corresponding rows and columns.

Specific Observations:

a. Happy**:** The model is quite accurate at recognizing happy expressions, with a high diagonal value and relatively low off-diagonal values.

b. Neutral**:** The model also performs well in identifying neutral expressions.

c. Disgust**:** The model shows decent performance in recognizing disgust, although there's room for improvement.

d. Anger**,** Content**,** and Surprise**:** The model struggles with these emotions, as evidenced by the lower diagonal values and higher off-diagonal values in the corresponding rows and columns. This indicates that the model often misclassifies these emotions.

The reasons for misclassification are similar to those mentioned for YOLOv8:

a. Data Quality: The quality and quantity of the training data can significantly impact the model's performance.
b. Class Imbalance: If the dataset is imbalanced, the model may be biased towards the majority classes.
c. Model Complexity: The complexity of the model architecture might not be sufficient to capture the subtle nuances between similar emotions.
d. Hyperparameter Tuning: The choice of hyperparameters can influence the model's performance.

To improve both YOLOv8 and YOLOv10's performance, similar strategies can be applied:

a. Data Augmentation: Increase the diversity of the training data by applying techniques like rotation, flipping, and colour jittering.
b. Class Balancing: Employ techniques like oversampling or undersampling to balance the class distribution.
c. Model Architecture: Experiment with different architectures, such as deeper or wider networks, or incorporating attention mechanisms.
d. Hyperparameter Tuning: Conduct a thorough hyperparameter search to find the optimal settings.
e. Transfer Learning: Utilize pre-trained models on larger datasets to initialize the weights of the model.

By addressing these factors, it is possible to improve YOLOv10's performance in recognizing human facial emotions, particularly for the challenging categories like anger, content, and surprise.

Both YOLOv8 and YOLOv10 show similar strengths and weaknesses in recognizing facial emotions. They both perform well on certain emotions like happy, neutral, and disgust, but struggle with others like anger, content, and surprise. Further analysis and experimentation are needed to determine if one model consistently outperforms the other.

Additional Considerations :

a. Evaluation Metrics: Consider using additional evaluation metrics like F1-score and ROC curves to gain a more comprehensive understanding of the model's performance.
b. Error Analysis: Analyse the specific misclassifications to identify patterns and potential areas for improvement.
c. Domain Adaptation: If the target data differs significantly from the training data, domain adaptation techniques can be employed to improve performance.

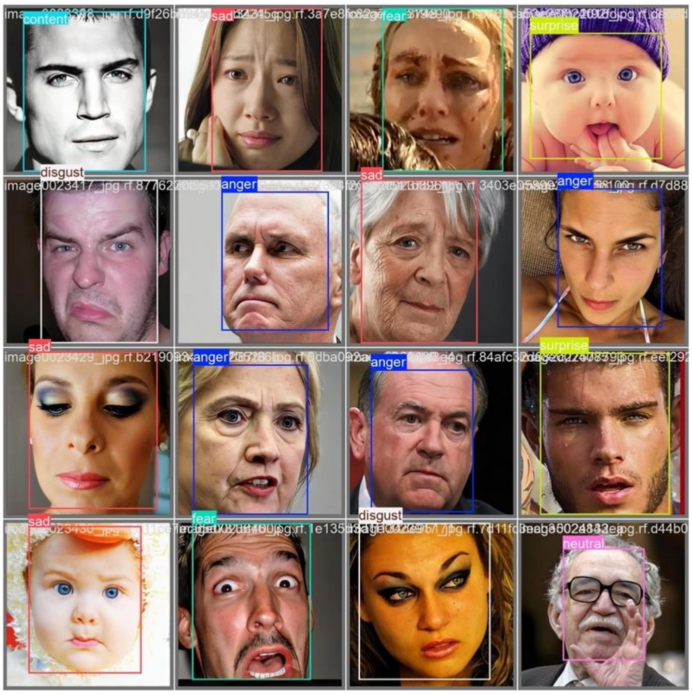


Figure 7. YOLO-v8 Prediction Result

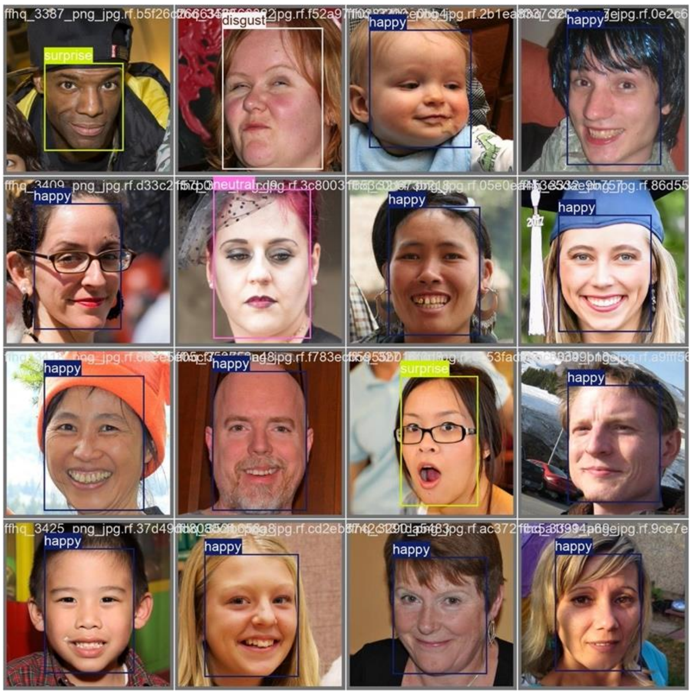


Figure 8. YOLO-v10 Prediction Result

The analysis reveals that both YOLOv8 and YOLOv10 are capable of recognizing human facial emotions with reasonable accuracy based on Figure 7 and 8. However, both models exhibit limitations in distinguishing certain emotions, particularly those with subtle differences.

By addressing the aspects, it is possible to further enhance the performance of both YOLOv8 and YOLOv10 in recognizing human facial emotions.

Conclusion

YOLOv8 and YOLOv10 demonstrated promising performance in human facial emotion recognition. Both models effectively learned to distinguish between different emotions, with YOLOv8 exhibiting slightly superior performance in terms of accuracy and generalization.

However, while both models achieved significant success, they still encountered challenges in recognizing certain emotions, particularly those with subtle differences, such as disgust and contempt. This limitation can be attributed to various factors, including the inherent complexity of facial expressions, the quality of the training data, and the limitations of the current model architectures.

To further improve the performance of facial emotion recognition systems, future research should focus on several key areas. Firstly, enhancing the quality and diversity of training data is crucial. By collecting a larger and more representative dataset, models can learn to recognize a wider range of emotional expressions, including nuanced and subtle variations. Secondly, exploring advanced model architectures, such as those based on transformer networks or graph neural networks, can potentially lead to significant performance gains. These architectures can capture complex dependencies between facial features and emotional states more effectively.

Additionally, optimizing hyperparameters through techniques like grid search, random search, or Bayesian optimization can fine-tune the models to achieve optimal performance. Furthermore, incorporating domain adaptation techniques can enable models to generalize well to new domains or datasets with different characteristics. This is particularly important for real-world applications, where facial expressions may vary across different cultures and environments.

Finally, real-time applications of facial emotion recognition systems hold immense potential. By developing efficient and accurate models that can process video streams in real-time, we can unlock a wide range of applications, such as human-computer interaction, mental health monitoring, and social robotics.

References

Aina, J., Akinniyi, O., Rahman, Md. M., Odero-Marah, V., & Khalifa, F. (2024). A Hybrid Learning-Architecture for Mental Disorder Detection Using Emotion Recognition. *IEEE Access*, *12*, 91410–91425. https://doi.org/10.1109/ACCESS.2024.3421376

Alshammari, A., & Alshammari, M. E. (2024). Emotional Facial Expression Detection using YOLOv8. *Engineering, Technology &amp; Applied Science Research*. https://api.semanticscholar.org/CorpusID:273295649

Amor, H. Ben, Bouallegue, S., Bohli, A., & Bouallègue, R. (2023). Development of a New Dynamic Approach for Facial Recognition and Emotion Detection. *2023 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 1–6. https://api.semanticscholar.org/CorpusID:263837618

Chaitanya, V. (2023). Understanding Human Emotions and Detecting Stress Levels using YOLO. *International Journal for Research in Applied Science and Engineering Technology*. https://api.semanticscholar.org/CorpusID:259738249

Hasan, M. A. (2023). Facial Human Emotion Recognition by Using YOLO Faces Detection Algorithm. *JOINCS (Journal of Informatics, Network, and Computer Science)*. https://api.semanticscholar.org/CorpusID:264045269

Hasan, M. A., & Lazem, A. H. (2023). *FACIAL HUMAN EMOTION RECOGNITION BY USING YOLO FACES DETECTION ALGORITHM*. https://api.semanticscholar.org/CorpusID:268325860

Huang, Y., Deng, W., & Xu, T. (2024). A Study of Potential Applications of Student Emotion Recognition in Primary and Secondary Classrooms. *Applied Sciences*. https://api.semanticscholar.org/CorpusID:274285872

Parambil, M. M. A., Ali, L., Swavaf, M., Bouktif, S., Gochoo, M., Aljassmi, H., & Alnajjar, F. S. (2024). Navigating the YOLO Landscape: A Comparative Study of Object Detection Models for Emotion Recognition. *IEEE Access*, *12*, 109427–109442. https://api.semanticscholar.org/CorpusID:271756461

RamaKrishna, N. (2024). Facial Emotion Detection: Applications, Advancements, and Implications in Human-Computer Interaction. *International Journal for Research in Applied Science and Engineering Technology*. https://api.semanticscholar.org/CorpusID:269337734

Emotions dectetion. (2024). Emotional Detection Dataset. In *Roboflow Universe*. Roboflow. https://universe.roboflow.com/emotions-dectection/human-face-emotions

Ting, Z., Yadav, A., Jiang, S. X., & Khan, A. (2024). Parameters Research of Facial Emotion Detection Algorithm Based on Machine Learning. *2024 Asia Pacific Conference on Innovation in Technology (APCIT)*, 1–6. https://api.semanticscholar.org/CorpusID:272724907

Vanamoju, S. V. M. D., Vineetha, M. V., Tekchandani, H., Joshi, P., Shukla, P. K., & Khanna, A. (2024). Facial Emotion Recognition using YOLO based Deep Learning Classifier. *2024 First International Conference on Electronics, Communication and Signal Processing (ICECSP)*, 1–5. https://api.semanticscholar.org/CorpusID:273226626