A Children's Psychological and Mental Health Detection Model by Drawing Analysis based on Computer Vision and Deep Learning

Amal Alshahrani
College of Computing, Umm Al-Qura University, Saudi Arabia
amshahrani@uqu.edu.sa (corresponding author)

Manar Mohammed Almatrafi
College of Computing, Umm Al-Qura University, Saudi Arabia
manaralmatrafi@outlook.sa

Jenan Ibrahim Mustafa
College of Computing, Umm Al-Qura University, Saudi Arabia
jenanaljurishi4@gmail.com

Layan Saad Albaqami
College of Computing, Umm Al-Qura University, Saudi Arabia
layanalbaqami@outlook.com

Raneem Abdulrahman Aljabri
College of Computing, Umm Al-Qura University, Saudi Arabia
I.raneemaljabri@gmail.com

Received: 12 May 2024 | Revised: 31 May 2024 | Accepted: 4 June 2024

ABSTRACT

Nowadays, children face different changes and challenges from an early age, which can have long-lasting impacts on them. Many children struggle to express or explain their feelings and thoughts properly. Due to that fact, psychological and mental health specialists found a way to detect mental issues by observing and analyzing different signs in children's drawings. Yet, this process remains complex and time-consuming. This study proposes a solution by employing artificial intelligence to analyze children's drawings and provide diagnosis rates with high accuracy. While prior research has focused on detecting psychological and mental issues through questionnaires, only one study has explored analyzing emotions in children's drawings by detecting positive and negative feelings. A notable gap is the limited diagnosis of specific mental issues, along with the promising accuracy of the detection results. In this study, different versions of YOLO were trained on a dataset of 500 drawings, split into 80% for training, 10% for validation, and 10% for testing. Each drawing was annotated with one or more emotional labels: happy, sad, anxiety, anger, and aggression. YOLOv8-cls, YOLOv9, and ResNet50 were used for object detection and classification, achieving accuracies of 94%, 95.1%, and 70.3%, respectively. YOLOv9 and ResNet50 results were obtained at high epoch numbers with large model sizes of 5.26 MB and 94.3 MB. YOLOv8-cls achieved the most satisfying result, reaching a high accuracy of 94% after 10 epochs with a compact model size of 2.83 MB, effectively meeting the study's goals.

Keywords-psychology; mental health; drawings; CNN; artificial intelligence; YOLO; deep learning; computer vision

I. INTRODUCTION

Many children around the world are affected by mental health issues that may have long-lasting effects on their lives, both academically and personally. Meanwhile, ignoring or not noticing these issues could negatively influence children's development process and limit their abilities, especially for children who cannot express themselves and explain their feelings clearly. Therefore, psychology specialists found drawing to be a powerful means of communication with children, as drawing tests reflect an individual's personality, attitudes, and experiences [1]. By drawing, children can express their thoughts and feelings and describe their point of view to the world [2]. This concept was of interest to experts and psychological professionals, who conducted many studies in this field and identified clinical semantics [3] for analyzing children's drawings and detecting their psychological meanings in mental health centers. However, it is a complex, time-consuming, and costly process for specialists and parents. Moreover, this procedure contains several steps before and after the drawing analysis stage, such as asking specific questions by psychologists to parents and analyzing their responses manually [4].

The work on this project started by identifying user needs by interviewing several psychology specialists: Razan Al-Hamdani, Rula Muhanna, Anwar Al-Hadi, and children's drawing analysts: Muhammad Al-Tamami and Maha Abdullah. They have been asked about using drawings to express children's feelings and the possibility of analyzing psychological and mental health from these drawings. They praised the importance of drawing as a necessary step to uncover the feelings of children and adolescents that cannot be expressed by words. Since the analysis and discovery are done in a long manual process, they assured that the model would be a big help in saving time and effort. On the other hand, a questionnaire was developed to ask parents about their need for the model, and the results showed their interest in and desire for the idea. Due to that, the need to use technology to develop and simplify this process was observed; hence, the present study came up with an Artificial Intelligence (AI) model to solve these issues. In the same context, the ESRA model [2] analyzed drawings with limited information by classifying them into only two categories: positive and negative emotions, without any other details, with an accuracy range of 55-79%. This model is designed to assist psychologists by analyzing children's drawings and collecting initial information about the child. It also aims to assist parents and help them check their children's mental health by developing this model into an application as future work to allow parents to upload pictures of their children's drawings and obtain accurate results from home. The results are provided by a trained AI model that can detect the meanings of the drawings and their symbols to produce correct results depending on the clinical semantics. There is also an intention to provide a way to communicate with trusted psychologists in this field to confirm the condition and provide a suitable treatment method after the analysis stage.

Regarding related work, many papers introduced several methods for detecting children's mental health disorders by drawing analysis. One study proposed mental health disorder prediction by a Decision Support System (DSS) based on AI that aimed to find solutions for problems with mental health assessment tools that require a long time and have low participation rates. The results displayed that the DSS is able to automatically diagnose mental disorders, with an accuracy level of 89% [5]. Other studies presented Machine Learning (ML) algorithms for mental health behavioral modeling. Authors in [6] addressed the challenge of predicting mental disorders in different groups, such as high school students, college students, and working professionals. The common notion between these categories is the stress and depression they go through in their daily lives. The study aimed to identify people with mental illness by using ML algorithms. At first, they prepared a questionnaire supported by experts. Then, based on the recipient's score, his mental health status was predicted. In addition, the authors deployed clustering techniques to identify groups and potential classes for further processing, as well as classification algorithms, such as SVM, KNN, and Random Forest (RF) for finding mental health illnesses. The research was conducted on two target groups in the ages of 18 to 21 years and 22 to 26 years and revealed that SVM, KNN, ensemble (bagging), and RF performed best with 90% accuracy. Authors in [7] proposed the use of Convolutional Neural Networks (CNNs). YOLO (You Only Look Once) itself is a CNN that is well-known for detecting patterns and objects in images [8]. The core of the YOLO algorithm lies in the model's small size and fast calculation speed. Its structure is straightforward. It can directly output the position and category of the bounding box through the neural network. YOLO is fast because it only needs to put the picture into the network to get the final detection result [9]. Using YOLO has shown promising results and high accuracy in the context of object classification. In [10], the YOLO v5 model achieved a 93% mean Average Precision (mAP) rate, which was significantly higher compared to the Faster R-CNN and EfficientDet models. In [11], YOLOv8 substantially enhanced classification results, achieving an impressive accuracy rate of 99.5%. Meanwhile, in [12], utilizing a CNN model for image recognition achieved 89% accuracy when trained on a children's sketch dataset.

Furthermore, this research aims to assist psychologists and parents in caring for children's mental health by developing an AI model to detect one or more of the following emotional states: sadness, anger and aggressiveness, anxiety, and happiness. This study compared RestNet-50, YOLOv8, and YOLOv9 models on the same dataset using the following steps:

- Review previous research in the field of analyzing children's drawings using ML and Deep Learning (DL).
- Highlight the most efficient object classification and detection algorithms and methods.
- Collect and create a dataset of 500 drawings from several sources collected manually by the authors.
- Preprocess and adjust the size of the dataset images to 640×640 using roboflow [13].
- Add labels (classes) by annotating the images in roboflow.
- Train and test the models (RestNet50, YOLOv9, and Yolov8) to classify and detect drawings.
- Evaluate the accuracy and compare the results of each model.

This paper compared three ML models to determine the best approach for detecting mental health issues from children's drawings. Accordingly, the YOLOv8 model exhibited the highest accuracy of top1_acc = 0.94 and top5_acc =1, both at 10 epochs. Top1_acc presents conventional accuracy with the highest probability as the model must predict the correct expected answer, while Top5_acc considers a classification correct if any of the five predictions match the expected label [14].

II. DATA COLLECTION

The dataset serves as the foundation for training the model and was collected after conducting comprehensive research and several interviews with psychologists to meet the requirements with high accuracy. No data set was found that included all the requirements, for this reason, children's drawings were collected from different places in Mecca's community by conducting interviews with several children, in addition to drawings offered from the interviewed psychologists. This study analyzed children's drawings using AI, in the same context with [2]. However, the ESRA model [2] analyzes drawings employing object detection technology and classifies them into only two categories: positive or negative feelings. In this paper, a CNN model was applied to categorize drawings, and three pre-trained models were used to train the model: ResNet18, ResNet34, and ResNet50. The ESRA application paper conducted four experiments utilizing different combinations of datasets, consisting of 102 drawings from a local school and 521 drawings from Google and Instagram [2]. The DL model trained implementing the Fastai library in Python classified graphics into positive or negative emotions with an accuracy of 55% and 79%.

The study stressed the need for further development to improve the accuracy of the AI model, with the addition of more feelings and emotions. On the other hand, the introduced model analyzes children's drawings using classification image technology and groups them based on the most common feelings among children, such as happiness, sadness, anxiety, anger, and aggression. All models were trained on a dataset of 500 children's drawings. The data set was divided into 80% for training, 10% for testing, and 10% for validation subsets, making a total of 400 drawings for training, 50 for testing, and 50 for validation. The next step was to use Roboflow to annotate by dividing the images into four classes, i.e. Happiness, Sadness, Anxiety, Anger and Aggression. The dataset contains 331 images for happiness, 68 images for anxiety, 33 images for sadness, and 68 images for anger and aggression. Among these drawings, psychologists submitted 321 drawings, and 179 drawings were collected from children in Mecca, between the ages of 6 and 15 years. Before collecting the drawings, a comprehensive examination of the children's health background and psychological condition was carried out, as this information is essential for accurate analysis. The drawings were analyzed with the help of psychologists based on clinical indicators of children's drawings. The analysis focused on elements such as shapes (geometric shapes) and symbolic elements (clouds, stars, flowers, nature). For example, clouds, especially rainy ones, indicate the feeling of anxiety in the child. The drawings were examined and classified based on facial and body features, with focusing on the mouth, teeth, and hands, to identify anger and aggression. Sadness was identified through the use of colors such as black with sad expressions, while happiness was identified through elements like nature, houses, orbs, and facial expressions shown in bright colors. The emotions in each image have been validated by expert psychologists using specific criteria and clinical markers of emotional expression in children's drawings.

(a) (b) (c) (d)

Fig. 1. Samples from the dataset: (a) Anger and Aggression, (b) Happiness, (c) Sadness, (d) Anxiety.

III. RESEARCH METHODS

A. ResNet50 Model

Residual Network (ResNet50) is an adaptable and scalable DL model for computer vision applications. A ResNet model was trained, hoping to achieve high accuracy and speed in object detection within images.

B. YOLO Algorithms

YOLO algorithms are known for their high accuracy and speed. Given these advantages, a comprehensive exploration was initiated by training multiple versions of the YOLO algorithm on the proposed dataset to evaluate the diverse results it provides. The advent of YOLO has brought about fundamental changes in the field of object detection by introducing real-time object recognition with exceptional accuracy [15].

1) YOLOv9

Object detection methods are usually three-stage processes: proposed region selection in the image using bounding boxes; feature extraction; and then the trained classifier is put into service to perform classification [16]. YOLOv9 depends on YOLOv8's strengths and introduces new features for even better object detection and classification, such as Programmable Gradient Information (PGI), This technique allows for the dynamic adjustment of gradients during training, which enables the model to prioritize learning from specific types of errors, leading to more targeted improvements in accuracy. The second feature that improved the YOLOv9 model is the Generalized Efficient Layer Aggregation Network (GELAN), which refines the feature extraction process by efficiently combining features from various network layers. It contributes to both speed and accuracy by providing a richer feature set for object classification.

2) YOLOv8

YOLOv8 has a good balance between accurate object detection and fast inference. It achieves this through the anchor-free design where YOLOv8 predicts bounding boxes directly. This makes the model simpler and reduces parameter number resulting in faster processing. YOLOv8 also has its own classification model "yolov8-cls". The basic task of this model is to classify directly and by utilizing the features of CNNs. It also focuses on informative regions within an image, allocating more resources to areas with higher object presence. This improves accuracy for object detection, especially the small ones.

Table I depicts the details of the YOLOv8 and YOLOv9 classification models.

TABLE I. PERFORMANCE METRICS FOR YOLO MODELS

Model	Size (px)	mAP 50-95	mAP 50	Params (M)	FLOPs (B)
YOLOv9c	640	53.0	70.2	25.5	102.8
YOLOv8n-cls	640	53.0	70.2	25.5	102.8

C. Training Methodology

This study aims to analyze children's drawings with the highest level of performance while maintaining the minimum size suitable for use in mobile applications as future work. This is performed by deploying CNNs, especially the recent developments in the YOLO versions, which support CNNs. Resnet50 was used to execute the same task, and this study compared all the results to determine the best performance. YOLO offers different sizes for its models (nano, small, medium, large, and extra-large). In this study, YOLOv8n and YOLOv9c were implemented, which effectively improve the architecture and achieve comparable accuracy. The evaluation has been extended to include resnet50. These models were trained on the same dataset, basic training settings include batch size, learning rate, momentum, and weight decay. These models were trained using a set of hyperparameters including varying durations from 10 to 150 epochs, a learning rate of 0.01, batch size of 8 for YOLOv9c, 16 for YOLOv8n ResNet50 with a training dataset of 500 images, a test set of 50 images, and a validation set of 50 images. Table II portrays the training settings.

TABLE II. TRAINING SETTINGS FOR THE MODELS

Model	Hyperparametere	Value
YOLOv8c	Epochs	10 – 50
	Learning rate	0.01
	Batch size	16
YOLOv9c	Epochs	10 -150
	Learning rate	0.01
	Batch size	8
ResNet50	Epochs	10-50
	Learning rate	0.01
	Batch size	32

D. Training Environment

For the model training environment, advanced computational power, like GPUs, was needed. Due to that, Google Colab was chosen to run the Python code on its cloud platform. The free version of Google Colab gives access to a GPU T4 graphics card with 12 GB of VRAM, which proved sufficient for a significant portion of this study's training needs.

E. Evaluation

Several metrics, such as precision (P), recall (R), and mAP, were applied to assess the models' performance and ability on the detection task. The accuracy of YOLOv8n-cls for classification was evaluated with the Top-N approach.

$$P = \frac{TP}{TP+FP} \quad (1)$$

$$R = \frac{TP}{TP+FN} \quad (2)$$

These metrics provide comprehensive knowledge about the accuracy and reliability of the model in analyzing children's drawings by discovering psychological connotations through the annotation of the drawing. These metrics were calculated using a confusion matrix consisting of four parts:

- True Positive (TP): The factual value was positive, and the model predicted a positive value.
- True Negative (TN): The factual value was negative, and the model predicted a negative value.
- False Positive (FP): Type I Error: The predicted value was falsely prognosticated. The factual value was negative, but the model predicted a positive value.
- False Negative (FN): Type II Error: The predicted value was falsely prognosticated. The factual value was positive, but the model predicted a negative value.

IV. RESULTS AND DISCUSSION

After conducting several experiments and training the models, a comparative analysis was performed to evaluate the models and determine the optimal performance based on accuracy and model size. Table III illustrates the YOLOv9c results.

TABLE III. YOLOV9C RESULTS

Epoch number	mAP50
10	0.53
20	0.617
30	0.745
50	0.818
100	0.949

A. YOLOv9c

The YOLOv9c model shows a noticeable increase in accuracy at epoch 100. The model was re-evaluated for 150 epochs. Table IV displays the YOLOv9c results for evaluation metrics comparison. Table V presents the evaluation details for epoch 150.

TABLE IV. EVALUATION METRICS COMPARISON FOR YOLOV9C

Epoch	P (%)	R (%)	mAP 50 (%)
10	57	0.53	53
20	75	0.617	61.7
30	85.5	0.745	74.5
50	72.6	0.818	81.8
100	90.7	88.6	94.9
150	88.9	91.1	95.1

TABLE V. EVALUATION METRICS COMPARISON FOR YOLOV9C IN EPOCH 150

Class	P (%)	R (%)	mAP 50 (%)	mAP 50-95 (%)
All	88.9	91.1%	95.1	72.6
Anger & Aggression	95.6	97.3	61.7	59.7
Anxiety	91.8	86.2	74.5	75.7
Happiness	90.9	81	81.8	68.2
Sadness	77.2	100	94.9	86.8

The results provided in the previous Tables are based on the confusion matrix. Figures 2-6 manifest the confusion matrix, the recall-confidence curve, the precision-recall curve, and the precision-confidence curve – F1 confidence curve for YOLOv9c.

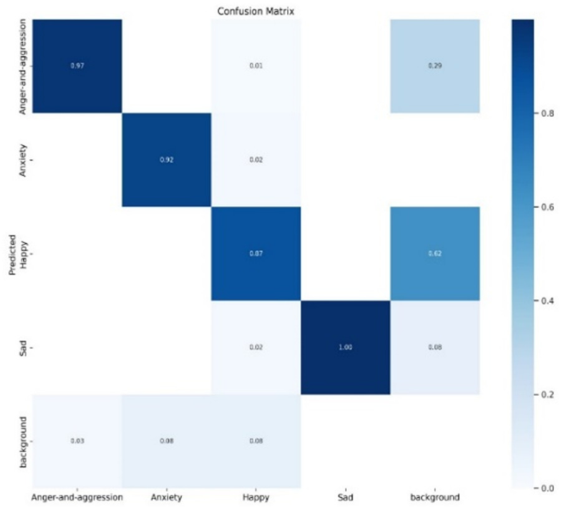


Fig. 2. Confusion matrix for YOLOv9c.

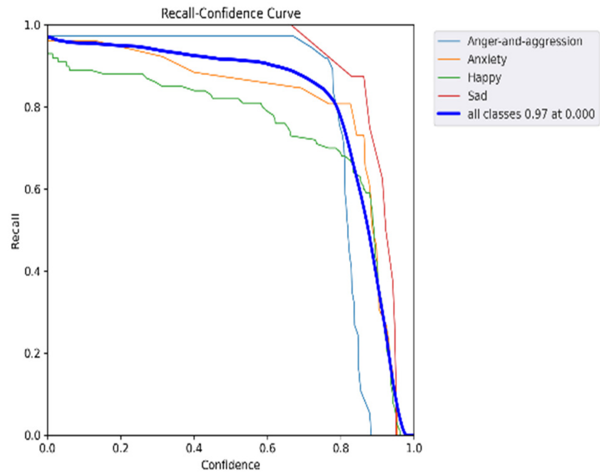


Fig. 3. Recall-confidence curve for YOLOv9c.

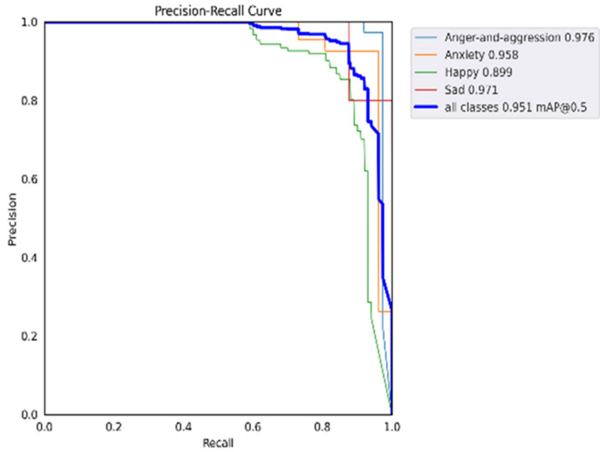


Fig. 4. Precision- recall curve for YOLOv9c.

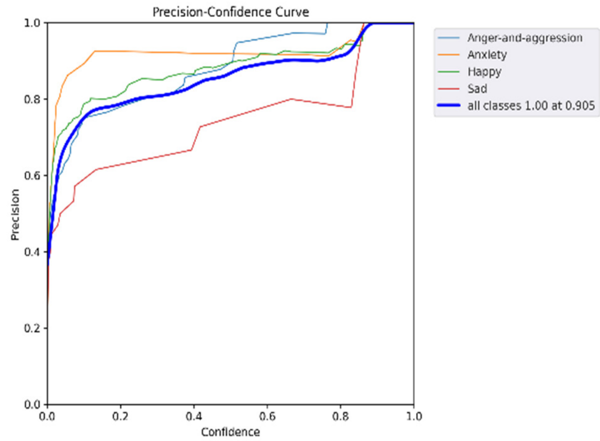


Fig. 5. Precision- confidence curve for YOLOv9c.

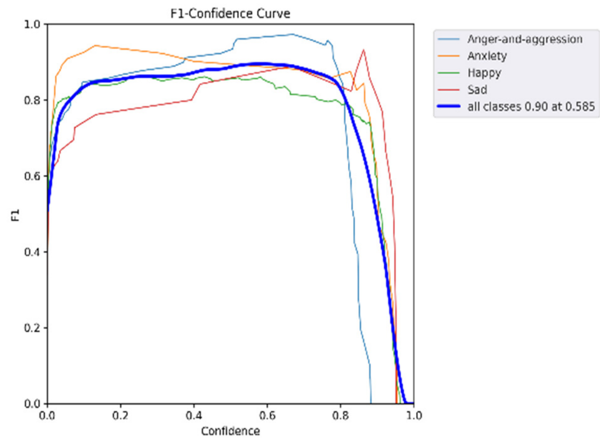


Fig. 6. F1 confidence curve for YOLOv9c.

B. ResNet50

The ResNet50 model, demonstrates the highest accuracy of 70.3% at epoch 24 (Table VI). Figures 7 and 8 display both model accuracy and model loss for the ResNet50 model.

TABLE VI. RESNET50 RESULTS

Epoch number	Accuracy
10	69.29%
24	70.3%
25	69%
50	68%

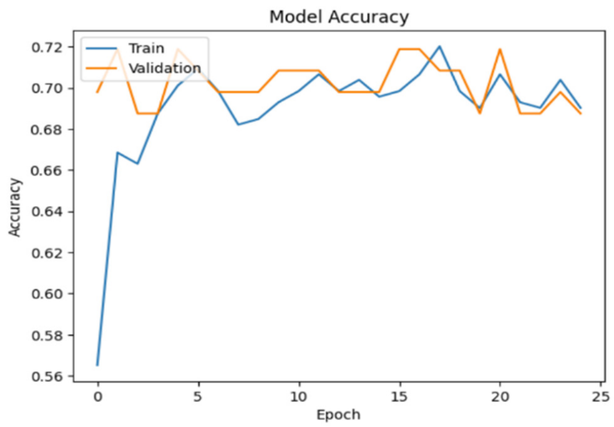


Fig. 7. Model accuracy for ResNet50.

Fig. 8. Model loss for ResNet50.

C. *YOLOv8n-cls*

Table VI depicts the accuracy results of YOLOv8n-cls. It can be observed that the YOLOv8n-cls model showed better results than the previous models.

TABLE VII. YOLOV8N-CLS RESULTS

Epoch number	top 1_acc
10	94%
15	92%
20	94%
25	92%
50	94%

Figures 9 to 12 exhibit the normalized confusion matrix, accuracy, and loss for both training and validation datasets if the YOLOv8n-cls model.

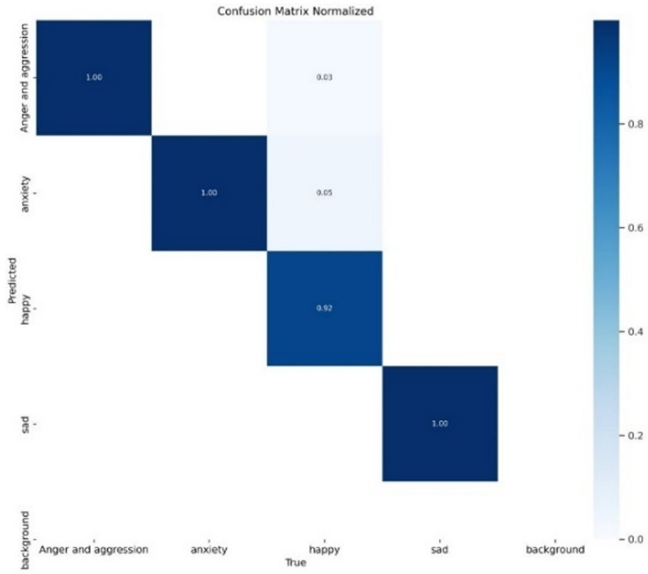


Fig. 9. Normalized confusion matrix for YOLOv8n-cls.

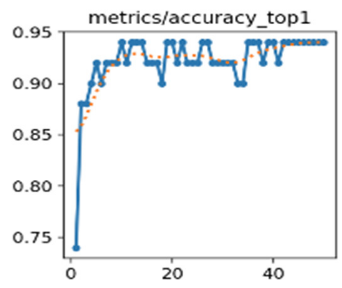


Fig. 10. Top1 accuracy matrix.

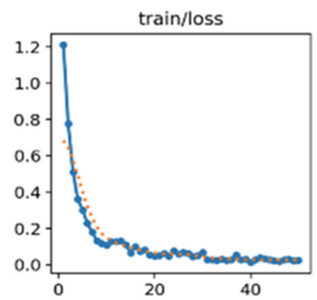


Fig. 11. Train loss of YOLOv8n-cls.

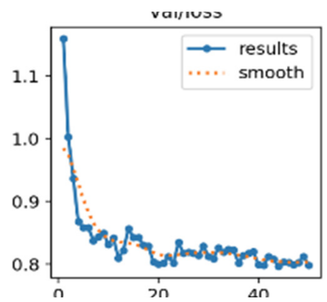


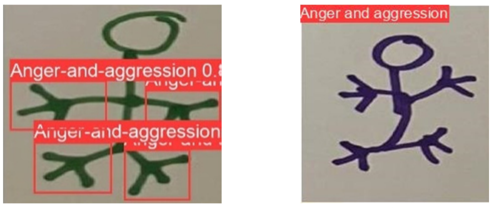
Fig. 12. Validation loss for YOLOv8n-cls.

D. A Sample of the Testing Output

Figures 13-16 somisclose of the testing results in terms of object detection (ResNet) or classification (YOLO).

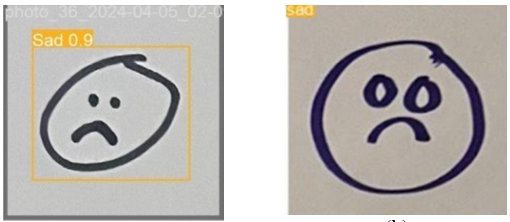
(a) (b)

Fig. 13. Detection of happiness using (a) ResNet and (b) YOLOv8n-cls.



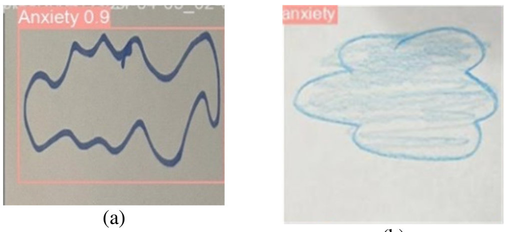
(a) (b)

Fig. 14. Detection of anger and aggression using (a) ResNet and (b) YOLOv8n-cls.



(a) (b)

Fig. 15. Detection of sadness using (a) ResNet and (b) YOLOv8n-cls.



(a) (b)

Fig. 16. Detection of anxiety using (a) ResNet and (b) YOLOv8n-cls.

E. Model Size

This study focuses on the highest level of performance while maintaining the minimum size suitable for use in mobile applications. Table VIII shows the comparison between the trained models in terms of size.

TABLE VIII. MODEL SIZE

Model	Model size (MB)
YOLOv8-cls	2.83
YOLOv9c	5.26
ResNet50	94.3

F. Result Comparison

Table IX illustrates a general comparison between the models in this study and the ESRA model [2]. The results in this study demonstrated the superiority of the proposed Yolo 8 classification model over all other models.

TABLE IX. RESULT COMPARISON

Model	Model	Type	Accuracy
ESRA [2]	ResNet34	Object Detection	0.76
	ResNet50		0.79
Proposed	ResNet50		0.703
	Yolov9c		0.951
	Yolov8n-cls	Classification	0.94

G. Discussion

The results indicate that YOLOv8n-cls meets the requirements in terms of overall performance, given the importance of achieving accurate classification with a smaller model size. Yolov9n reached an accuracy of 95.1% at epoch 150 and an accuracy of 94.9 % at epoch 100, with precision of 88.9% at epoch 150 for all classes. In [2], the highest precision was only 74% for the ResNet50 model. The ResNet50 model reached 70.3% at epoch 24. Therefore, the classification made a big difference in terms of performance and accuracy, as its accuracy was 94% even at epoch 10. The YOLOv8-cls model fully met the requirements in terms of detecting the feeling of the drawing by classification. The small size of the model stands out from the rest of the other models.

V. CONCLUSION AND FUTURE WORK

While previous research has primarily focused on questionnaire-based approaches and detected two categories of feelings with low accuracy, this study proposes a novel method utilizing computer vision and deep learning techniques to analyze children's drawings for emotional indicators. Taking advantage of computer vision and deep learning techniques, this research presented a system dedicated to detecting several emotions in drawings through object detection or classification. A comparison between the methods and models used was also provided. A dataset was created, containing 500 images, each image was classified into one category: happiness, sadness, anxiety, and anger or aggression. Experiments and evaluations were conducted, including YOLOv8n for classification and YOLOv9c and ResNet50 for object detection. Among these models, the YOLOv8n-cls algorithm reached the best performance, achieving a top-1 accuracy of 94% at epoch 10 with a smaller model size. Therefore, this model was chosen as the most suitable due to its accuracy and size efficiency, making it ideal for mobile applications. Moving on, the YOLOv9 object detection model achieved a performance of 95.1% at mAP50 in epoch 150, but it is less preferred due to its bigger size and longer training time. As for future work, the authors of this study plan to expand the training dataset to include more diagnoses, such as sexual harassment, family relationship problems, and other issues to improve the scope of the model. In conclusion, the proposed approach is scalable, with the ability to evolve and adapt to analyze more sentiments on a large scale. Furthermore, this system can be seamlessly integrated in a mobile application, thus expanding its usability and effectiveness. Finally, this technological advance could make a huge impact, driving faster recognition of emotions in children's drawings.

REFERENCES

[1] M. Lee, Y. Kim, and Y.-K. Kim, "Generating psychological analysis tables for children's drawings using deep learning," *Data & Knowledge Engineering*, vol. 149, Jan. 2024, Art. no. 102266, https://doi.org/10.1016/j.datak.2023.102266.

[2] N. Ali, A. Abd-Alrazaq, Z. Shah, M. Alajlani, T. Alam, and M. Househ, "Artificial Intelligence-Based Mobile Application for Sensing Children Emotion Through Drawings," in *Advances in Informatics, Management and Technology in Healthcare*, IOS Press, 2022, pp. 118–121.

[3] Y. Yuan, J. Huang, X. Ma, and K. Yan, "Children's Drawing Psychological Analysis using Shallow Convolutional Neural Network," in *2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)*, Rhodes, Greece, Nov. 2020, pp. 692–698, https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics50389.2020.00120.

[4] M. O. Zeeshan, I. Siddiqi, and M. Moetesum, "Two-Step Fine-Tuned Convolutional Neural Networks for Multi-label Classification of Children's Drawings," in *Document Analysis and Recognition – ICDAR 2021*, 2021, pp. 321–334, https://doi.org/10.1007/978-3-030-86331-9_21.

[5] S. Tutun *et al.*, "An AI-based Decision Support System for Predicting Mental Health Disorders," *Information Systems Frontiers*, vol. 25, no. 3, pp. 1261–1276, Jun. 2023, https://doi.org/10.1007/s10796-022-10282-5.

[6] M. Srividya, S. Mohanavalli, and N. Bhalaji, "Behavioral Modeling for Mental Health using Machine Learning Algorithms, " *Journal of Medical Systems*, vol. 42, no. 5, Apr, 2018, https://doi.org/10.1007/s10916-018-0934-5.

[7] H. Lin and J. Yang, "Ensemble cross-stage partial attention network for image classification, " *IET Image Processing,* vol. 16, Jan, 2022, https://doi.org/10.1049/ipr2.12335.

[8] R. Taleti, "YOLO: You Only Look Once," *Towards Data Science*, Aug. 02, 2020. https://towardsdatascience.com/yolo-you-only-look-once-f05c054a06b4.

[9] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo Algorithm Developments," *Procedia Computer Science*, vol. 199, pp. 1066-1073, 2022, https://doi.org/10.1016/j.procs.2022.01.135.

[10] R. Rajamohanan and B. C. Latha, "An Optimized YOLO v5 Model for Tomato Leaf Disease Classification with Field Dataset," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12033–12038, Dec. 2023, https://doi.org/10.48084/etasr.6377.

[11] B. Xiao, M. Nguyen, W. Qi Yan, "Fruit ripeness identification using YOLOv8 model," *Multimedia Tools and Applications*, vol. 83, no. 9, pp. 28039–28056, Feb. 2023, https://doi.org/10.1007/s11042-023-16570-9.

[12] F. Shamrulismawi, "Sketch Recognition based on Deep Learning for Children's Psychological Development Monitoring," B.Sc. thesis, Universiti Teknologi PETRONAS, Seri Iskandar, Malaysia, 2021.

[13] M. Almatrafi, "anamelClassification Image Dataset." [Online]. Available: https://universe.roboflow.com/anamel/anamelclassification/dataset/1.

[14] A. T. Dang, "Accuracy and Loss: Things to Know about The Top 1 and Top 5 Accuracy," *Medium*, Jan. 07, 2021. https://towardsdatascience.com/accuracy-and-loss-things-to-know-about-the-top-1-and-top-5-accuracy-1d6beb8f6df3.

[15] T. Saidani, R. Ghodhbani, A. Alhomoud, A. Alshammari, H. Zayani, and M. B. Ammar, "Hardware Acceleration for Object Detection using YOLOv5 Deep Learning Algorithm on Xilinx Zynq FPGA Platform," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 13066–13071, Feb. 2024, https://doi.org/10.48084/etasr.6761.

[16] E. K. Ronoh, S. Mirau, and M. A. Dida, "Human-Wildlife Conflict Early Warning System Using the Internet of Things and Short Message Service," *Engineering, Technology & Applied Science Research*, vol. 12, no. 2, pp. 8273–8277, Apr. 2022, https://doi.org/10.48084/etasr.4662.