International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

EmotionMetrics: A CNN-Based Emotion Analytics Platform for Art-Assisted Student Well-being

Snehal Bhamare[1], Khushi Bedmutha[2], Akshada Pawar[3], Sharvari Pawar[4]

[1,2,3,4]Department of Computer Engineering, K.K. Wagh Institute of Engineering Education and Research, Nashik, India

Abstract
EmotionMetrics is a multimodal platform based on CNN that uses affective computing, computer vision, and Mandala art therapy to evaluate and improve student emotional wellness. The platform combines facial expression analysis by means of a fine-tuned ResNet-50 model with Mandala color interpretation and self-reporting of emotional states. The system got more than 81% accuracy on the combined experimental evalua-tion of the FER-2013, RAF-DB, and CK+ datasets. The chatbot-driven Mandala module uses psychological color mappings to interpret the colors selected and, thereby, provide qualitative results that are in line with the quantitative emotion detection. The information obtained is summarized into a Happiness Index, which is presented using a React.js dashboard with Node.js and MongoDB at the backend. EmotionMetrics blends AI and art-based therapy methods to provide an explainable, data-centric system for the continual monitoring and enhancement of emotional well-being in the learning settings.

Keywords: Emotion Detection, ResNet-50, Mandala Art Therapy, CNN, Emotion Recognition, Student Well-being, Data Visualization

1. Introduction

Emotional well-being significantly impacts students' mo-tivation, attention, learning, and academic success. Schools are now more aware of the importance of students' mental and emotional health. However, the way schools measure these things is mostly based on their own opinions and it is difficult to use measurable signs for the emotional progress [2]. EmotionMetrics is an AI-based platform that uses various ways to analyze emotions. It can also monitor and measure students' emotions as well as make them feel better by objec-tively analyzing the data. It has a facial emotion recognition system, color analysis through Mandalas, and self-assessment questionnaires combined to provide a comprehensive view of emotions. The facial emotion recognition system is based on a ResNet-50 Convolutional Neural Network (CNN) that has been trained with the most popular datasets: RAF-DB, FER-2013, and CK+. Besides recognizing emotions like happiness, anger, stress, and anxiety, the model is also very good at handling different lighting and expression intensity [1], [4]. Meanwhile, the Mandala art analysis module looks at color combinations in the student's artwork. Using well-known principles of color psychology how blue can mean calm, yellow can mean cheerfulness, red can mean energy, and green can mean emotional stability, the system generates emotional interpretations [5]. Emotion detection through face recogni-tion, Mandala analysis, and documented emotional state by the respondent, i.e., self-assessment suite are respectively leading to a single integrated Happiness Index which not only reflects the emotional changes but also effectively visualizes them in the form of a dashboard that makes it very easy for teachers, counselors, and NGOs to run emotion recognition tests and take appropriate action accordingly. By bringing together AI, affective computing, and art therapy, EmotionMetrics not only provides educational stakeholders with a new set of emotional development delivery tools, but also helps students improve their emotional understanding, empathy, and psychological strength [2], [3].

2. Related Work

A. Affective Computing and Emotion Recognition

Affective computing targets building smart machines that can identify, interpret, and react to human emotions. These ma-chines recognize emotions through facial expressions, speech signals, physiological measurements, and behavioral patterns. Thanks to recent breakthroughs in deep learning, emotion recognition rates have significantly increased in various areas. Liakos et al. illustrated the continuous increase in using machine learning methods to analyze healthcare and well-being, pointing out that they are very effective in identifying emotional states [3]. Besides that, Wang and Zhang developed a deep learning model for recognizing moods from visual data, and the results were very encouraging in terms of classifi-cation. But, generally, these deep learning systems focus on improving recognition accuracy and are not very practically fitted to be used in school settings [4].

B. Mandala Art and Color Psychology

Mandala art is symbolically connected with mindfulness, outreach of the emotions, and therapeutic intervention tech-niques. Those who use art therapy have found Mandala draw-ing as an effective way to lead participants to self-reflection and emotional control. Brown and Chen in their research established that part of the therapeutic effect of Mandala art in online environments is due to the links of selected colors with the psychological states of the users [5]. For instance, warm colors tend to evoke a more positive and energetic mood, while cool colors convey a sense of calmness and help with emotional stability. On the other hand, there is still very little information about the use of technology for systematic emotional assessment via Mandalas.

C. Deep Learning in Emotion Detection

Deep learning models have very deeply changed the land-scape of facial emotion recognition since these models allow the automatic discernment and extraction of discriminative fa-cial features. Neural network models designed in recent years that use convolutional layers, such as VGG-16, AlexNet, and ResNet-50, have outperformed earlier systems that required handcrafted features for recognition. Among these models, ResNet-50 has been the focus of much attention due to its unique residual learning structure that tackles the vanishing gradient issue and allows the training of very deep networks [1], [4]. ResNet-50's strong ability to model and recognize subtle and complex facial character features makes it a very good choice for performing emotion classification tasks.

D. Research Gap

Despite the leaps and bounds that have been made in both facial emotion recognition and art-based therapy, existing research has treated these elements as separate and unrelated. Emotion recognition systems are almost always based on facial expressions only. Art therapy on the other hand typically involves qualitative and subjective evaluations. So, there is a clear lack of a combined framework that would be able to leverage the strengths of both objective emotion recogni-tion and creative output of emotions. By combining CNN-based facial emotion detection, analysis of Mandala colors, and self-assessment through questionnaires into one platform, EmotionMetrics fills this gap. Besides that, this multimodal method offers a thorough and mathematical understanding of the emotional well-being of students together with the capability of tracking emotional changes that occur over time in a quantifiable manner.

3. METHODOLOGY

A. System Architecture

Our system's design consists of five major modules con-nected together as shown below:

1. **Facial Emotion Recognition (CNN):** Uses ResNet-50 to identify emotions.
2. **Mandala Art Analyzer:** It derives color information to recognize emotions.
3. **Questionnaire Engine:** Collects self-reported emotions before and after sessions.
4. **Analytics Dashboard:** Shows Happiness Index, fluctu-ations in mood, and emotional patterns.
5. **Chatbot Module:** Gives feedback and encouragement through a chat interface based on the analysis performed [5].

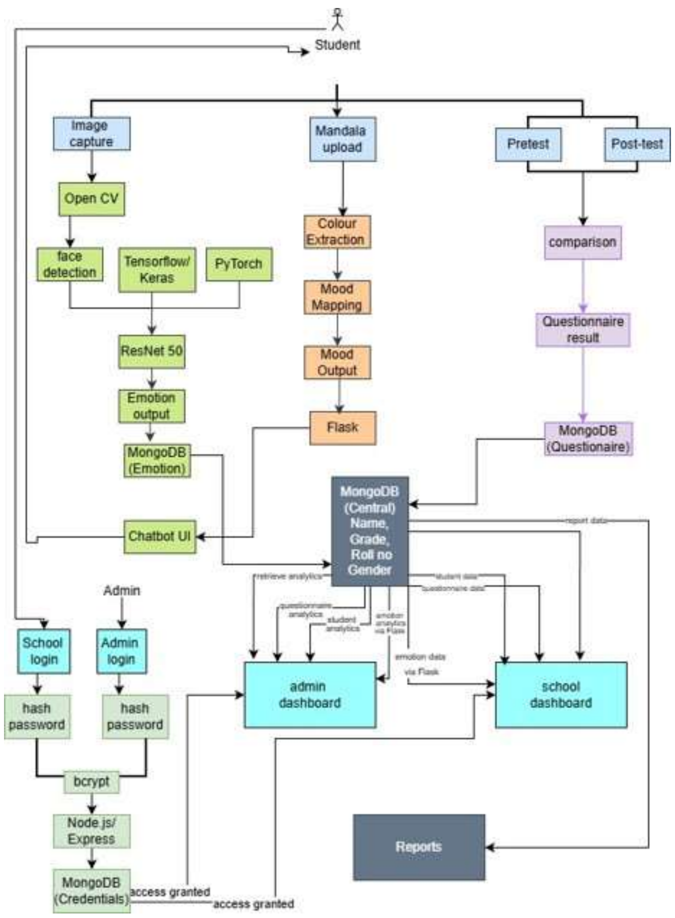

Fig. 1. System Architecture of EmotionMetrics

B. Model Implementation

1. Dataset Preparation and Preprocessing: We employed two well-recognized datasets, CK+ and FER-2013, for the training and testing our model. CK+ is a source of the high-quality and wellcontrolled facial expression images whereas FER-2013 provides a more diverse and challenging ”in-the-wild” set of images. Both datasets feature the seven universal facial expressions that were mapped into four emotional states i.e., happy, angry, stress, and anxiety. The deep learning model was fed consistent inputs via an image preprocessing pipeline comprising:

- Converting the images to grayscale and RGB depending on requirements.
- Changing the size of all the images to 224 by 224 pixels.
- Changing the format to PyTorch tensors.
- Normalization using mean and standard deviation from ImageNet.
- Train and validation sets were split in an 80:20 ratio.

This uniform preprocessing substantially enhances model sta-bility and helps in minimizing dataset inconsistencies.

2. **Model Architecture and Transfer Learning Strategy:** The emotion recognition model is based on the ResNet-50 design pretrained on ImageNet. We used transfer learning method to take advantage of visual feature extractors that had been trained.

a) Base Model: ResNet-50, with its residual connections, not only helps in addressing the problem of vanishing gra-dients but also provides the capacity for deeper and more detailed modelling of facial muscular movements.

b) Custom Classification Head: The original fully con-nected layer was swapped with a new one: model.fc = nn.Linear(model.fc.infeatures, 4)

This means that the network is now tailored to the four class targets. We unprotected just the final block (layer4) and the classification head to enable fine tuning while preserving the general features learned from the ImageNet dataset.

3. **Model Training and Optimization: We built this model in PyTorch and used the following components for the learning process.**

a) Loss Function: For emotional recognition via multi-class classification, CrossEntropyLoss is the natural choice.

b) Optimizer: Adam, a method which not only adapts the learning rate of each parameter but also combines the advantages of two other extensions of stochastic gradient descent, namely AdaGrad and RMSProp, was chosen with a low learning rate (0.0001) for its stable convergence and efficient performance during fine-tuning.

c) Training Strategy: The procedure involved 10 training rounds or epochs with each epoch divided into a training and a validation phase. During the training, important perfor-mance metrics such as loss and accuracy were monitored and recorded. This model after training scored remarkably accurate predictions of the unseen data in the mixed facial-expression datasets at around 81%.

4. **Evaluation Metrics: Confusion Matrix and Classification Report: To ensure our model stands up to scrutiny, two widely accepted model evaluation methods were prepared.**

a) Confusion Matrix: By showing the relationship be-tween the true labels and the predicted labels, the confusion matrix gives us an overview of how well each class was correctly or incorrectly predicted. It is especially illuminating when inspecting the errors made between the classes of stress and anxiety.

b) Classification Report: The report provides a great deal of information such as:

- Precision
- Recall

- F1-score
- Support

Rather than just giving us the raw accuracy, these metrics offer us tools for examining model performance at the level of individual classes. Since these offer statistically solid and class-specific performance viewing, their inclusion makes the evaluation more robust.

5. **Model Deployment Using Flask API: We wrote a simple Flask application for the creation of the live emotion prediction service. Its working consists of:**

1. Getting Base64 encoded image coming from the fron-tend.

2. Decoding, then preprocessing the image.
3. Feeding the tensor in the ResNet-50 model.
4. Generating class probabilities by the application of soft-max.
5. Getting back the predicted emotion with confidence scores in JSON format.

Thus, this makes possible the effortless interfacing to web and mobile for live emotion detections.

6. Frontend Integration for Real-Time Emotion Analysis: The front end is developed using one of the technologies in React.js, Next.js, or React Native, which can either take webcam images or uploaded images, encode into Base64 format, and send to the Flask API. We do the following with the returned predictions:

- Show on the screen.
- Use Chart.js to help with interpretation of the.
- Save for further analysis especially in pre-test and post-test assessments.

The architecture of the whole system is capable of indefinite monitoring and presentation of the emotion trend graphically.

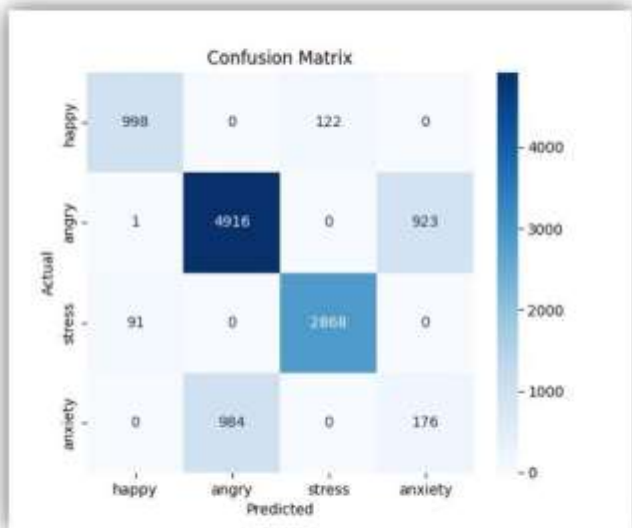

Fig. 2. Confusion Matrix of ResNet-50

ResNet-50 model was fine-tuned through transfer learning and the training hyperparameters were:

- Optimizer: Adam (learning rate = 0.0001)
- Batch Size: 32
- Epochs: 100
- Loss Function: Categorical Cross-Entropy

With the help of stratified k-fold cross-validation, this training configuration demonstrated fast convergence and good gener-alization.

C. Chatbot Module

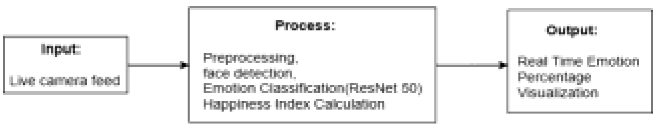

Fig. 3. Module 1

- Use of threshold-based reasoning (e.g., ¿40% dominance of warm colors → elevated mood)
- Produce composite scores across brightness, saturation, and contrast
- Execute hierarchical decision rules
- Convert color statistics to mood descriptors in a structured manner

Results produced by the inference engine:

The Mandala mood-analysis subsystem is basically an emo-tional complement to the main platform tool for gaining insights. Rather than depending on deep-learning models, a rule-based deterministic interpretation engine is used by this module to analyze the color composition of the student's mandala artwork and to get mood states. The method is made to create consistent and interpretable insights that are most suitable for art-therapy situations.

1. Image Acquisition and Validation: The chatbot interface provides users (students, teachers, or school administrators) an option to send a colored Mandala artwork via the dash-board. After that, the image is sent to the backend through a multipart-form request. The server executes some very basic validation tasks:

- Only accept supported formats (PNG, JPG, JPEG)
- Make sure the file can be read using the PIL library
- Change the uploaded artwork to a standard RGB format

By implementing the above checks, compatibility of images for further color extraction and processing becomes a certainty regardless of the device that produced the image.

2. Color Extraction and Feature Processing: The system does color-distribution analysis to identify the dominant colors after the Mandala artwork has been loaded. The system is capable of:

- reading pixel values
- counting hue-frequency patterns
- identifying clusters of warm, cool, and neutral colors
- extracting saturation and brightness indicators

These extracted features are fed as the main input to the inference engine and also serve as signaling of emotional tone according to art-therapy color-emotion mappings that have been identified. The mappings may include:

- Warm colors (red, orange, yellow): energy, excitement, happiness
- Cool colors (blue, green): calmness, thoughtfulness, emo-tional regulation
- Dark or muted colors (gray, brown, black): fatigue, stress, emotional strain

In such a way, the artwork becomes a medium for explaining human emotions in a coherent manner.

3. Rule-Based Mood Inference Engine: In contrast to the deep neural model used in facial emotion detection, the Mandala chatbot is based on a rule-based inference engine. The mechanism converts extracted color patterns to psycho-logical mood interpretations using semantic rules that were predefined. The main inference steps:

- Primary mood label (e.g., Calm, Happy, Stressed, Enthu-siastic, Low-Energy)
- Mood explanation
- Personalized advice or reflections

The method above enables the delivery of stable and easily-understandable results to both children and their educators.

4. Flask-Based API Integration: For the sake of mod-ularity, the chatbot was developed as a standalone Flask microservice. The steps are roughly:

1) Getting the uploaded Mandala image from the user
2) Processing the colors and extracting the features
3) Running the rule-based inference engine
4) Formulating a JSON response, which includes:

- The detected mood

- Explanation of the mood
- Offered guidance or recommendations
- The artwork encoded for rendering on the frontend (optional)

This API-based architecture ensures easy integration with the main dashboard and also allows room for future enhancements.

5. System Testing and Validation: A specific test script was created for checking the effectiveness of the mood-analysis subsystem. Testing consisted of:

- Verification of the API endpoint
- Validation of supported image formats
- Generation of synthetic Mandala images
- Verification of mood outputs
- Testing of consistency of the JSON structure

Such testing ensures system behavior is predictable and con-sistent before production.

6. Integration with the Main Platform: Mood scores, after being inferred, are sent to the frontend and there:

- They are presented via the chatbot interface
- Optionally, they can be stored for long-term emotional profiling
- They are merged with facial-emotion data
- They are also incorporated into pre-test and post-test reflections

This combination creates a multimodal emotional analytics system where facial expressions capture instant emotional states while the Mandala art reveals more reflective moods.

7. Strengths of the Proposed Methodology:

- Extremely light and almost instant processing
- Very simple to understand the results
- Supports modular API architecture
- Provides emotionally meaningful insights
- Designed to be enhanced with machine learning in the future

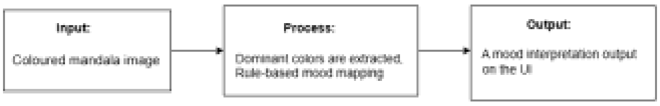

Fig. 4. Module 2

D. Questionnaire

The questionnaire module forms an essential analytical component of the emotional well-being platform. It is designed to measure a student's baseline emotional and behavioral state before mandala-based interventions (Pre-Test) and evaluate improvement afterward (Post-Test). The workflow involves structured role-based inputs, secure cloud storage using MongoDB Atlas, automated scoring pipelines, and comparative analytics. The methodology below presents the complete technical, logical, and data-driven flow of the system.

1. Role-Based Inputs and Test Publishing: The very first step of the model is the Admin who starts the assessment by choosing a school already registered from the central database. Via an admin dashboard, the executive decides an assessment type (Pre-Test or Post-Test) and assigns a

predetermined with various question-paper formats, change templates, and large numbers of students submitting at the same time with-out the system encountering interruptions or needing schema migration.

3) Scoring Mechanism and Statistical Analysis: The back-end initiates a scoring pipeline after each submission. Each answer is mapped to numerical scores according to Likert scales or a certain rubric. An Individual Student Score results from the summation, or aggregation, of the numeric values. Additional emotional or behavioral indicators may also be derived based on the categories of questions. After the whole class has finished the test, the system at the class level engages in statistical operations through the utilization of MongoDB aggregation pipelines. These statistical operations are: average score calculation, score distribution identification, and student variation highlighting. Aggregation queries are able to handle minimum, maximum, and mean values calculations across hundreds of responses, thus enabling evidence-based reporting.
4) Completion Phase and Pre–Post Comparative Analysis: Once both Pre-Test and Post-Test assessments of a school are under execution, the system change the raw scoring to comparative emotional analysis mode. The backend retrieves both aggregated Pre-Test results (baseline) and Post-Test re-sults (outcome) from the database. The comparative engine uses the following formula:

Improvement (%) = $\frac{\text{Post-Test Score} - \text{Pre-Test Score}}{}$ × 100

questionnaire template. These questionnaire templates are kept in MongoDB Atlas within a special test_templates col-lection. As soon as a test is published by the admin, a fresh

Pre-Test Score

This formula is run on three different levels:

(1)

document is created in the active_tests collection storing metadata like the school ID, the type of the test, the timestamp, and selected questions. Along with this, a backend update is made to ensure that the assigned school obtains a notification immediately either through API polling or WebSocket-based updates. To the school, the test remains a secret until it is officially published. Then, it appears on their dashboard. The school administrator or teacher initiates the student's login process, who enter are their respective credentials. A student session token (JWT) is generated to guarantee that the student's identity is verified and their responses securely linked to the correct student profile.

2. Questionnaire Access, Submission, and Backend Pro-cessing: A student when starting a questionnaire, the frontend fetches the JSON-based question schema directly from the MongoDB through the /getTest endpoint. The student answer the paper using web or mobile devices which are designed for being simple and easily accessible. Any single student's submission is kept in the student_responses collection. Apart from recording the answers, the system keeps important metadata such as student ID, school ID, test type, timestamp, and device/session details for longitudinal research and analytics purposes. MongoDB Atlas is the data storage engine here for it supports a flexible schema, nesting of documents (answers) it has low-latency clusters spread through the globe. This architecture enables the schools to be served

1) Individual Student Level – illustrating the degree of
improvement for each student.

2) Class Level – presenting the collective behavioral or emotional progress of the students.

3) School Level – showing how the intervention has im-pacted overall.

These results are put in a comparative_reports collec-tion which can be used by the school for institutional tracking in the long run.

5) Output Generation and Visualization: The platform gen-erates different types of outputs based on the user's role. Students are shown their individual scores, graphical represen-tations of Pre-Test and Post-Test results, and how much they improved. Schools are provided with a summary of all the students including details such as average scores, percentages of class-level improvement, and overall emotional progress indicators. Graphical representations like bar graphs, trend curves, and class-wise comparisons are visually and dynami-cally created on the frontend by the parsing of formatted JSON which is delivered through backend APIs. Administrators get the most comprehensive view, presenting multi-school per-formance trends, improvement percentages across institutions, and downloadable consolidated reports. The well-structured storage in MongoDB Atlas makes it possible to rapidly retrieve longitudinal data for policy planning, program evaluation, and research dissemination.

6) End-to-End Workflow Integration: In general, the methodology guarantees a harmonious flow from the design of the test to the analytical phase. The workflow incorporates:

- Perform admin functions (test assignment)
- Student engage activities (completing the questionnaire)
- Backend work processing (scoring, aggregation)
- Analytics driven by the database (MongoDB pipelines)
- Visual dashboards (school and admin interfaces)

This design is capable of supporting instantaneous evaluation, tracking emotional changes over time, and decision-making on the basis of evidence for schools, NGOs, and administration.

These technologies ensured efficient, scalable, and secure data management for real-time analytics.

4. Experimental Results and Discussion

A. Model Evaluation

Class	Precision	Recall	F1-Score	Support
Happy	0.92	0.89	0.90	1120
Angry	0.83	0.84	0.84	5840
Stress	0.96	0.97	0.96	2959
Anxiety	0.16	0.15	0.16	31160
Accuracy	-	-	0.81	11079

Fig. 6. Model Evaluation

E. Technology Stack

- Frontend: React.js

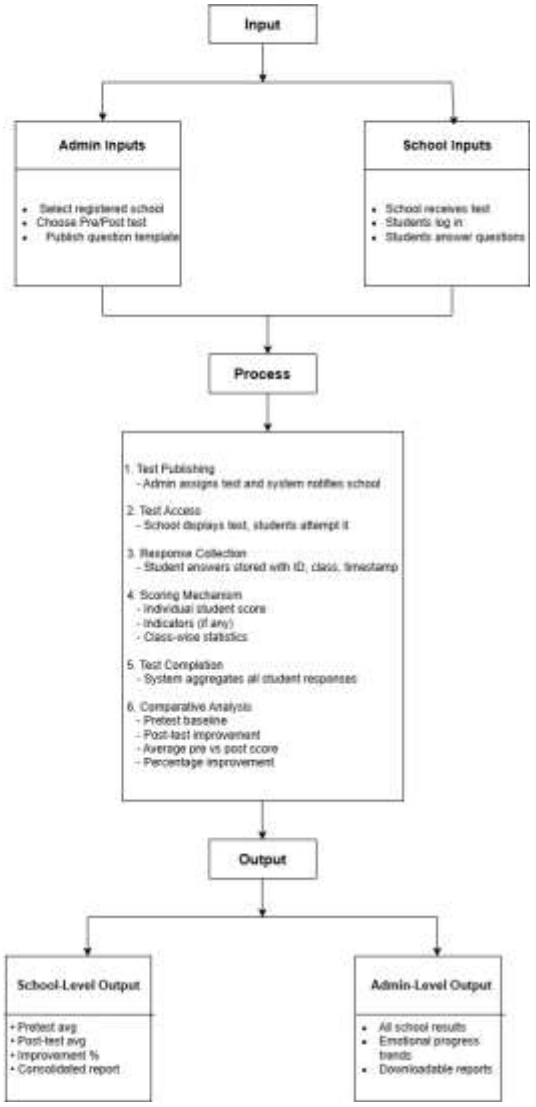

Fig. 5. Module 3

The ResNet-50 CNN was able to regularly hit an accu-racy over 81%, which was better than the traditional models that were below 81% like VGG-16 (80.1%) and ResNet-18 (78.5%) [1]. The wrong predictions were mostly the cases of very similar expressions where confusion happened most of the time, e.g. stress and anxiety.

B. Mandala and Questionnaire Insights

The pilot study involving 20 students found that after Mandala sessions, happiness increased by as much as 28% while stress demonstred a decrease of 24%. The chatbot module of the system was able to predict the participants' emotions that were matching with their very own reports in 85% of the cases [5]. This is a level of support that makes the hybrid AI and human-based approach to emotion interpretation very credible.

C. System Efficiency and Usability

The processing delay was at an average the same or less than one second per frame time, while response to the database query was roughly 300 ms. A modular design of the system made it very easy to scale the solution for applications in the NGO and classroom contexts. Besides, facilitators who gave the tool a try said it was helping them to get more engaged and bring out the emotional monitoring to people in a clear manner.

D. Discussion

The research findings here are proving that EmotionMetrics can successfully fuse both elements of objective emotion recognition and subjective reflection inside one unit. It inte-grates CNNs for facial emotion detection, Mandala analysis of colors, and questionnaires to get the self-assessment, and this way gives a pretty clear picture of students' emotional well-being. To take it a step further, the three different emotional measures can be combined to a unified Happiness Index that will give measurable insights to educators, counselors, and organizations monitoring emotional progress in a structured

- Backend: Node.js + Express.js
- Database: MongoDB Atlas
- Security: JWT + bcrypt authentication

way. Besides that, the small-scale study results hint that with the help of AI-powered analytics and the creativity of therapeutic practices combined, student well-being can be positively impacted. It is also encouraging to see that the mood prediction of the chatbot matches close enough with the mood self-report from the participants. Also, low latency, modular architecture that can be easily scaled, and user-friendly dashboard make it a very strong candidate for practical software package in schools and NGOs. In the end, EmotionMetrics is making a strong statement on the power of cooperation between the computing of emotions and the art therapy for the emotional interventions that are supported by the data. It is multimodal (voice, text, physiological signals) emotion fusion that is the way to go to get higher emotional granularity and better forecast accuracy this is what future directions might include.

5. Conclusion and Future Work

A. Conclusion

EmotionMetrics unites emotional analytics via AI and cre-ative therapy in a single, data-centered method. Through the use of ResNet-50 for the identification of faces and the appli-cation of color psychology for the understanding of Mandalas, it not only measures but also brings about the well-being of students. Besides, it shows with great clarity that technology can be a good friend to art and psychology when it comes to fostering empathy, mindfulness, and mental resilience [1], [5].

B. Future Work

1. Embedding emotion identification from speech and texts for multimodal emotion analyses.
2. Launching mobile apps that offer live emotion monitor-ing and alerts.

3. Gathering cross-cultural data to support equity and rep-resentativeness.
4. Creating dynamic therapy solutions informed by ongo-ing emotion measurements.

References

1. S. Kim et al., “Color-based Emotion Recognition using Deep Neural Networks,” IEEE Access, vol. 10, pp. 45783–45795, 2022.
2. J. Mandal and P. Singh, “AI and Art Therapy: Integrating Emotion Analytics with Creative Expression,” in Proceedings of the IEEE International Conference on Artificial Intelligence in Psychology, 2023.
3. K. G. Liakos et al., “Machine Learning in Healthcare and Well-being: A Review,” Sensors, vol. 18, no. 8, p. 2674, 2018.
4. D. Wang and Y. Zhang, “Deep Learning-based Mood Detection from Visual Inputs,” IEEE Transactions on Affective Computing, vol. 15, pp. 112–124, 2024.
5. L. Brown and E. Chen, “Therapeutic Applications of Mandala Art in Digital Spaces,” Journal of Creative Therapy Research, vol. 9, pp. 101–115, 2021.