From Crayons to Code:
AI-Driven Insights into a Child's Mental Health Through Drawings

Moomal Farhad, Mohammad Mehedy Masud
Aisha Alnaqbi, Rawan Mubarak, Aaisha Aladawi, Sara Alnaqbi
College of Information Technology, United Arab Emirates University
PO Box 15551, Al Ain, UAE
{moomalfarhad, m.masud, 202005456, 202007209, 202013426, 202005842} @uaeu.ac.ae

Abstract

Children's mental health is crucial for their development, but it's often overlooked, leading to psychological issues. Many children struggle to express their thoughts and feelings effectively. To address this issue, we have proposed a novel approach to analyze children's drawings for psychological screening using artificial intelligence. Specifically, we're focusing on the 'draw a person' (DAP) test, where a child's drawing is used to identify potential indicators of their mental and emotional state. Thus, we are introducing an AI-powered technique to automate the psychological screening process for children using the DAP test, which a human professional would traditionally conduct. The screening tool would suggest whether the child needs or doesn't need further psychological referral. We have collected a dataset consisting of children's drawings and labeled them by experts as either 'need' or 'no need', indicating whether the child needs or does not need a referral. We have proposed two alternative approaches for the screening process. The first approach consists of extracting features from the drawings following expert guidelines and training a classification model using the features to classify the drawing as either 'need' or 'no need'. We also propose an out-of-the-box technique applying prompt engineering on state-of-the-art LLMs to automatically extract features from the images. The second approach involves training an image classification model using the drawings. Both approaches are challenged by the issue of class imbalance, as most of the drawings correspond to the 'no-need' class. To address this challenge, we introduce Siamese++, a novel Siamese network for image classification, which uses feature embedding and an adaptive distance threshold for classification, instead of the nearest neighbor classification employed by traditional Siamese. Our proposed method achieves a high F1 score (up to 88%) even with a large class imbalance and without the need for any image augmentation. Thus, we have proposed an innovative interdisciplinary integration of AI with psychology and developed novel techniques to solve the real-world problem of psychological screening.

1 Introduction

Mental health disorders among children are a significant and growing concern worldwide. According to the World Health Organization, an estimated 10-20% of children and adolescents globally experience mental health disorders, with conditions like anxiety, depression, and behavioral issues (World-Health-Organization 2021). Early identification and intervention are critical in mitigating the long-term effects of mental health disorders. Research shows that untreated mental health issues in childhood can lead to more severe problems in adolescence and adulthood, including chronic mental illness, academic difficulties, and impaired social functioning (Kessler et al. 2005). However, identifying these issues early can be challenging, particularly in younger children who may not have the verbal skills to articulate their emotions. This gap necessitates alternative methods of assessment that can provide insights into a child's emotional state.

One effective way to understand children's emotions is by analyzing their drawings. Drawing is a natural and accessible form of expression for children and often reflects their inner world. Research indicates that children's drawings can provide valuable insights into their emotional and psychological well-being (Malchiodi 1998). Specific patterns, colors, and themes in drawings have been associated with anxiety, depression, and other emotional issues. Therefore, analyzing these drawings can be a valuable tool for the early detection of psychological distress in children.

Based on the above observation, researchers and clinicians have increasingly turned to well-defined and structured tools for screening psychological disorders in children through their drawings. Two such assessment tools are the Draw-A-Person test (DAP) (Laak, De Goede, and Aleva 2005) and the House-Tree-Person (HTP) test (Becker-Weidman 2017). In the DAP test, a child is asked to draw a person or three people (a male, a female, and self). In the HTP test, a child is asked to draw a house, a tree, and a person. These two tests have emerged as valuable instruments for evaluating cognitive and developmental aspects, as well as for screening psychological disorders (Jurovatý and Démuthová 2023).

In this study, we target the DAP test for psychological screening. We collected a substantial dataset of drawings from young children and had them labeled by expert psychologists based on their emotional content. These labels categorized the drawings into those indicating a need for further psychological referral (denoted as the 'need' class) and those that did not need referral (denoted as the 'no need' class). The expert-labeled data serve as the foundation for

---

our subsequent analysis and classification.

After collecting and labeling the data, we proposed innovative automated artificial intelligence (AI) techniques, using cutting-edge machine learning (ML) and AI technology, to automate the process of identifying children who may need psychological support. We introduced two different approaches along this line, which are summarized below.

The first approach involves identifying key features in the drawings following expert guidelines and manually extracting them. Additionally, we propose an improvised technique for automatically extracting these features from the images by applying prompt engineering to state-of-the-art LLM, GPT4o. These feature vectors are then used to train traditional ML models, such as random forests, and classify the images into 'need' or 'no need' classes. The second approach involves training an image classification model from the drawings, such as a convolutional neural network (CNN). However, both approaches face the issue of class imbalance, as most of the drawings correspond to the 'no-need' class. For the feature-based approach, we use Synthetic Minority Oversampling (SMOTE) (Chawla et al. 2002). For the image classification approach, we introduce Siamese++, a novel Siamese network for image classification, which uses feature embedding and adaptive cutoff distance for classification instead of the nearest neighbor classification employed by traditional Siamese. The proposed approach adaptively adjusts the decision boundary, ensuring more precise classification, especially when there is a significant disparity between the positive and negative classes. Our proposed method achieves a high F1 score (up to 88%) even with a large class imbalance and without the need for any image augmentation.

The contributions of this study are highlighted below.

- We have **collected a dataset** consisting of children's drawings (DAP) and had the drawings **labeled by a psychology expert** as either 'need' or 'no need', indicating whether the child needs a psychological referral or not.
- We have **identified key features** in children's drawings (i.e., drawing a person) using expert guidance and extracted those features manually from the collected data.
- We have proposed an improved technique to extract the same set of features automatically from the drawings through **prompt engineering**. We have developed the code for prompting a Large Language Model (LLM) using API calls, and we get the desired features extracted by the LLM. The use of LLM technology to extract features from images not only demonstrates the potential of these advanced techniques in automating complex feature extraction tasks from images but also highlights their broader applicability in fields such as medical diagnostics and psychology, where nuanced and complex data interpretation is crucial.
- We have **developed Siamese++, a novel few-shot Siamese** learning technique to address the extreme class imbalance in the dataset. The model utilizes an adaptive cutoff distance threshold to adjust the decision boundary, ensuring a more balanced and accurate classification. We have demonstrated the superiority of Siamese++ over the traditional Siamese approach, where classification is done using the nearest neighbor.
- The proposed technique presents a **scalable solution** that applies machine learning for psychological screening of children, supporting children's mental health in schools, clinics, and other settings.
- We have demonstrated a framework using **innovative interdisciplinary AI integration** for psychological screening.

In summary, this study provides a novel approach to the early and automated identification of emotional and psychological issues in children. It bridges the gap between psychological expertise and technological innovation, demonstrating the application of AI to complex, nuanced tasks. The use of machine learning enhances the efficiency of the analysis and increases the accuracy and consistency of the assessments, which are critical in clinical settings. The details of the proposed work are presented in the following sections.

2 Related Work

In this section, we discuss the work related to the application of various machine learning models in healthcare, focusing on methods for automating psychological assessments and diagnosing mental and physical health conditions through drawing analysis.

The research paper (Pan et al. 2022) focuses on enhancing objectivity in the House-Tree-Person (HTP) test by applying the Support Vector Machine (SVM) algorithm. In a similar vein, another study (Lee, Kim, and Kim 2024) explores the automation of mental evaluations through deep learning, tackling the subjectivity and time constraints associated with traditional drawing-based assessments. The authors propose a system utilizing object detection algorithms applied to the House-Tree-Person (HTP) test, leveraging a dataset of 1,100 drawings from the Goyang City Datathon.

Recent studies demonstrate the effectiveness of artificial intelligence in analyzing children's drawings to address emotional challenges. For instance, (Alshahrani et al. 2024) trained various YOLO and ResNet50 models on 500 annotated drawings, achieving high diagnostic accuracy with YOLOv8-cls at 94%, showcasing its potential for detecting emotional states. Similarly, (Baird et al. 2022) applied digitized drawing analysis to Syrian refugee children, using LASSO regression to find significant correlations between drawing features and psychological disorders such as anxiety, depression, and PTSD. This approach offers a rapid, cost-effective, and non-invasive method for mental health assessments in crisis settings.

Moreover, studies like (Kotsavasiloglou et al. 2020) and (Ferdib-Al-Islam and Akter 2020) have investigated hand-drawn drawings for diagnosing Parkinson's disease. The research by (Kotsavasiloglou et al. 2020) utilized a pen-and-tablet device to analyze hand movement and muscle coordination, achieving a 91% classification accuracy through features such as pen tip velocity and trajectory entropy. This approach shows potential for telemedicine and monitoring disease symptoms using accessible technology.

In the broader context of advanced models, recent research on Siamese models has yielded promising results across various domains. For instance, one paper introduces a holistic approach for matching face sketches using a Siamese model, which is designed to be both simple and effective (Shukla et al. 2020). This method specifically addresses the computational complexity typically encountered in feature-based schemes, offering a streamlined solution for recognizing face sketches. Another related work discusses a Siamese Graph Convolution Network (GCN) tailored for face sketch recognition, utilizing a graph structure to boost recognition accuracy (Fan, Sun, and Rosin 2021). This model achieves an F1 score of 0.73 and an accuracy of nearly 74%, proving its effectiveness even under challenging conditions like aging effects in sketches. In a different application, Siamese networks are also leveraged in the development of Siamese-EEGNet, a model that employs a CNN to analyze time-frequency features from Electroencephalogram (EEG) signals, using both the original EEG data and their Power Spectral Density (PSD) (Yan et al. 2022). Together, these studies highlight the versatility and effectiveness of Siamese models in various complex recognition tasks.

3 Methodology

Our methodology consists of five key phases: data collection, preprocessing, feature extraction, model training, and evaluation. Each phase is detailed below.

3.1 Data Collection

We have gathered children's drawing data from a publicly available dataset on Kaggle, specifically the "Draw-a-Person hand-drawn sketches" dataset (Rakhmanov 2021). This dataset includes around 1,000 images drawn by primary school students aged 4 to 11, ordered according to their cognitive age as provided by professionals. Currently, we are also in the process of collecting data from local schools.

3.2 Data Labeling

We label each drawing as 'need' or 'no need,' based on whether the child requires expert consultation. This labeling is done in collaboration with psychology experts. Thus, 915 drawings were labeled as 'no need' and 37 were labeled as 'need'.

3.3 Feature Extraction

The feature extraction phase involves extracting three sets of features mentioned below.

Manually Extracted Features: These features consist of specific characteristics manually identified from the drawings based on expert guidance. They include visual characteristics (e.g. size, absence, presence) of the head, mouth, teeth, nose, ears, eyes, belly button, sex organ, blood, arms, hands, talons, legs, feet, lines, color, color choice, placement and weapon. These features have been suggested by the psychology expert who labeled the same data. According to the experts, these features are vital in identifying key patterns in the drawing indicating the child's psychological state.

LLM Extracted Features: We have used prompt engineering to extract the same set of features mentioned above from LLM. For this, we have coded the prompts and called appropriate APIs to invoke GPT4o to extract the features from the images automatically. The LLM analyzes the drawings and generates features based on its understanding of visual patterns and relationships. We have extracted features this way to examine the potential of using LLM in interpreting complex data such as children's drawing in a human-like efficacy. Our experimental results prove that LLM-extracted features are comparable and sometimes more accurate than manually extracted features. Thus, we have introduced a novel and efficient way of extracting features from images.

3.4 Model Training and Evaluation

We have trained and evaluated two types of machine learning models. The first type is the traditional ML models, trained with the extracted features mentioned in Section 3.3, such as Random Forest (RF), SVM, and XGBoost (XGB). The second type comprises CNN-based deep learning models, including VGG and ResNet, which were trained on raw image data. Both types of models aim to classify drawings into two categories: 'need' and 'no need', as introduced before. Note that the class distribution is highly imbalanced (915 to 37), which poses a major challenge in training the classification models. For the traditional ML models, we address this challenge using synthetic minority oversampling (SMOTE) (Chawla et al. 2002). For the CNN-based models, we address the issue by proposing a novel Siamese approach called Siamese++, which is described in Section 4.

4 Proposed Siamese++ Framework Based on Adaptive Distance Threshold

To address the significant class imbalance in the dataset, we propose a novel Siamese network, which we call 'Siamese++", designed to perform few-shot learning for image classification. The overall architecture of Siamese++ is given below.

4.1 Overall Architecture

A traditional Siamese neural network (Bromley et al. 1994) is designed to address the challenge of training an effective classifier when there is a significant class imbalance. The Siamese network consists of two identical networks that process a pair of images and calculate the distance between their outputs. In these networks, feature vectors are learned using CNNs applied to labeled pairs of matching and non-matching images (Melekhov, Kannala, and Rahtu 2016). The twin networks share the same parameters, ensuring that similar images are mapped to nearby points in the feature space since each network computes an identical function. The similarity between these feature vectors is then evaluated using a distance metric, and the image is classified based on a similarity score using the nearest-neighbor approach.

We also face the challenge of high class imbalance, which is aggravated by the fact that same-class drawings can be significantly different. For the traditional Siamese network, this

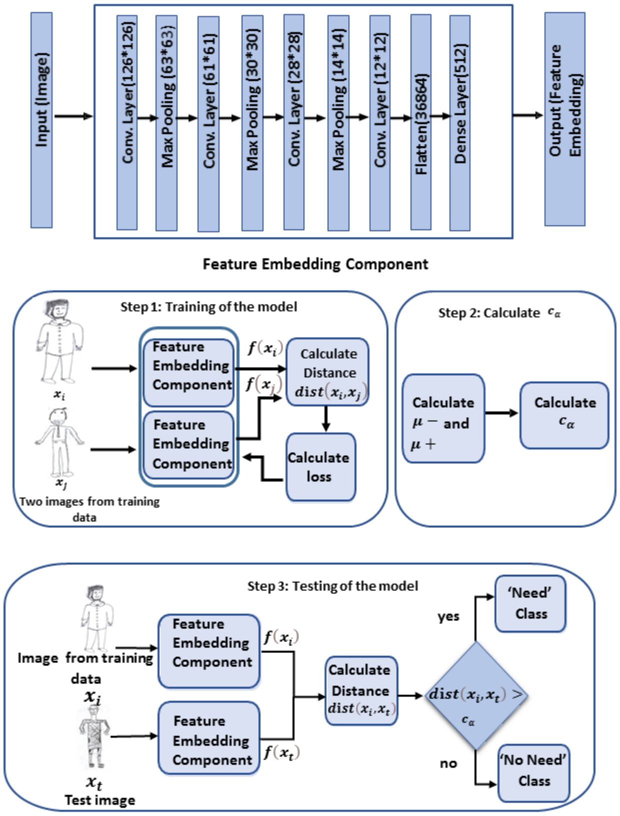

Figure 1: Training and testing of Siamese++ model

makes learning the distance metric, significantly difficult, and the nearest neighbor classification produces a higher error. Therefore, we propose Siamese++, which effectively addresses both these challenges. Figure 1 illustrates the process of Siamese++ designed to classify children's drawings into 'Need' or 'No Need' for expert consultation. The model comprises a feature embedding component with a series of convolutional and pooling layers, followed by a dense layer that outputs feature embeddings. The steps are detailed below.

Step 1: During training, two images from the training data, comprising both 'Need' and 'No-Need' class samples are passed through the network to generate their feature embeddings. The distance between these embeddings is then calculated, and a loss function is optimized based on these distances. This step is crucial for creating positive and negative pairs to effectively train the model.

Step 2: In this step, the mean distances for positive and negative pairs are used to calculate an adaptive distance threshold. This threshold, derived from the training data, helps prevent erroneous classification decisions based solely on the nearest-neighbor approach, especially in cases of significant class imbalance. Relying solely on the nearest-neighbor method could lead to incorrect classification, as the drawings belonging to the same class can be significantly different. However, the adaptive distance threshold reduces the chance of incorrect classification by considering the intra-class and inter-class similarities as a whole. Details about calculating the adaptive distance threshold are provided in the subsequent subsection.

Step 3: During testing (Step 3), a test image is compared against a reference image from the training set (No-Need class). The distance between their feature embeddings is computed, and the image is classified as 'Need' or 'No Need' based on whether the distance exceeds the adaptive distance threshold.

4.2 Adaptive distance threshold

In this section, we will describe the adaptive distance threshold calculation, its significance, and its application.

Definitions and notations

Let x_i be an input image to the Siamese++. The network computes a feature embedding $f(x_i)$ of the input image. For a pair of input images x_i and x_j, the network computes the distance between their embedding $dist(x_i, x_j)$. Also, let T^- be all the negative (i.e., 'no-need') class instances in the training data, and let T^+ be all the positive (i.e., 'need') class instances in the training data.

Definition 1 (Image-pair distance $dist(x_i, x_j)$) *The image distance between a pair of images x_i, x_j is the distance between their feature embeddings $f(x_i)$ and $f(x_j)$ using a distance metric, and denoted using the function $dist(x_i, x_j)$.*

Definition 2 (Mean negative-pair distance μ^-) *The mean negative-pair distance is the mean of all negative class image-pair distances, denoted by μ^-, as described by the following equation:*

$$\mu^- = \frac{1}{N(N-1)} \sum_{\forall x_i \in T^-} \sum_{\forall x_j \in T^-}^{x_j \neq x_i} dist(x_i, x_j) \tag{1}$$

where N is the number of negative image class instances in training data, i.e., $N = |T^-|$

Definition 3 (Mean opposite-pair distance μ^+) *The mean opposite-pair distance is the mean of all positive-negative pair distances, denoted by μ^+, as described by the following equation:*

$$\mu^+ = \frac{1}{M(N)} \sum_{\forall x_i \in T^+} \sum_{\forall x_j \in T^-} dist(x_i, x_j) \tag{2}$$

where M is the number of negative image class instances in training data, i.e., $M = |T^+|$, and N has the usual meaning introduced earlier.

Definition 4 (Cutoff distance C_α) *The cutoff distance is defined as a distance threshold, parameterized by a variable α, described by the following equation:*

$$C_\alpha = \alpha\mu^+ + (1-\alpha)\mu^- \tag{3}$$

Using C_α for Classification: C_α is used as the threshold distance between a test instance and a negative training instance. If the distance is greater than the cutoff, the test instance will be identified as positive, and vice-versa. Therefore, in summary, we use C_α as follows.

- Step 1: During testing, we calculate the distance between the test image's embedding and the negative class images in the training data.
- Step 2: If the calculated distance (step 1) is greater than C_α, classify the image as positive (i.e., 'Need'); otherwise, classify it as negative (i.e., 'No-need').

Significance of parameter α: The parameter α controls the decision boundary, which can be thought of as a hypersphere around the negative instances, thereby influencing the class prediction. We can explain its influence in the following three scenarios.

- ($\alpha \to 0$): This is an extreme case where the cutoff distance tends to encompass the smallest possible hypersphere, meaning the most restrictive decision boundary for the negative class. Therefore, more test instances tend to be classified as positive, which can be advantageous when the positive class images are not very distinct from the negative images.
- ($\alpha \to 1$): This is the other extreme case where the cutoff distance tends to encompass the largest possible hypersphere, meaning the least restrictive decision boundary for the negative class. Therefore, more test instances tend to be classified as negative, which can be advantageous when the positive class images are significantly distinct from the negative images.
- ($0 < \alpha < 1$): This is a balanced scenario compared to the extreme cases, providing a flexible and adaptive cutoff criterion. Our experimental study finds the middle value (0.5) showing the best performance.

4.3 Why Siamese++

Siamese++ presents an innovative and adaptive framework for image classification with a highly imbalanced class distribution. This novel approach is superior to the traditional Siamese network, as we have proven empirically. The main advantages are as follows.

- **Utilization of both class information:** This method effectively utilizes information from both classes, making it suitable for a wide range of applications.
- **Adaptive:** The weighting factor α allows for fine-tuning, thereby making it adaptive to data variability.
- **Flexible:** The adaptive distance threshold offers a more flexible decision boundary compared to a fixed one. While the nearest neighbor approach may produce inaccurate results due to a limited number of training samples, using an adaptive distance threshold provides a more intuitive and robust method for classification.

This adaptive cutoff distance provides a way to create a threshold that incorporates both intra-class ('No-need') and inter-class ('No-need' vs. 'Need') distances, with the flexibility to adjust the influence of each.

5 Experiments and Results

In this subsection, we outline the experimental setup and analysis of the results achieved by Siamese++ and other competing approaches.

	Manual features			**LLM features**		
	Prec	**Recall**	**F1**	**Prec**	**Recall**	**F1**
RF	0.75	0.77	0.76	0.8	0.79	0.79
SVG	0.68	0.61	0.64	0.83	0.84	0.84
XGB	0.76	0.77	0.76	0.85	0.85	0.85

Table 1: Performance of machine learning classifiers with SMOTE using Manual and LLM features

5.1 Experimental Setup

For all experiments, the datasets were split into a 75:25 ratio for training and testing. Precision, recall, and F1 score were used as evaluation metrics, with all metrics reported considering 'need' as the positive class. The image classification models were trained for 30 epochs with a batch size of 10. The experiments are divided into the following sections:

Experiments with Manual and LLM Features

To evaluate the effectiveness of manual and LLM features, we applied machine learning algorithms such as RF, SVM, and XGBoost. To enhance the performance of these models, we utilized the SMOTE technique. The results of these experiments are detailed in the subsequent subsection.

Experiments with Image Data For the image dataset, we compared the performance of the following competing approaches.

Siamese++: the proposed novel Siamese++ network.

Transfer learning models: VGG16 and ResNet.

Siamese model (NN): The traditional Siamese model that uses the Nearest Neighbor (NN) classification for prediction.

We also compared Manhattan and Euclidean distances with varying numbers of positive training samples to assess their impacts on precision, recall, and F1 score.

5.2 Results

In this subsection, we have presented the results of the Siamese++ model and other competing approaches.

Performance of Machine Learning Models for Manual and LLM Features

Table 1 showcases the performance of three classifiers RF, SVM, and XGB using both Manual and LLM features with the SMOTE technique applied. Initially, with a class distribution of 687 'no need' and 27 'need', all classifiers exhibited poor performance with F1 scores of 0.00. After applying SMOTE to balance the dataset to 687 samples for both classes, significant improvements were observed. The F1 scores for RF and XGB with LLM features rose to 0.79 and 0.85, respectively, reflecting enhanced robustness and effectiveness. This indicates that SMOTE substantially addresses class imbalance, leading to more reliable and accurate model performance.

Classifier	**Precision**	**Recall**	**F1**
Transfer Learning Models	0.00	0.00	0.00
Siamese Model (NN)	0.80	0.76	0.78
Siamese++	0.85	0.89	0.87

Table 2: Performance of image-based classifiers

Metric	Manhattan distance			Euclidean distance		
	3	**5**	**10**	**3**	**5**	**10**
Precision	0.71	0.73	0.85	0.53	0.63	0.74
Recall	0.8	0.86	0.89	0.77	0.80	0.86
F1	0.75	0.79	0.87	0.62	0.70	0.795

Table 3: Comparison of Manhattan and Euclidean distances with varying positive training samples

Performance of Siamese++ for Image Dataset To conduct experiments for Siamese++ and the Siamese model (NN), we compared the testing image with three images from the training set. For Siamese++, we only used images from the No-Need class. However, for the Siamese model (NN), we randomly selected three images from both classes. Table 2 presents the performance of Siamese++ alongside other competing approaches, including Transfer Learning Models (VGG16, ResNet) and the Siamese Model (NN). Data augmentation was applied for transfer learning models using the following settings: rescaling, rotation (20 degrees), width and height shifts (0.2), shear (0.2), zoom (0.2), and horizontal flipping. Despite this, transfer learning models show the lowest performance, with all metrics at 0.00, indicating no successful predictions. The Siamese Model (NN) demonstrates improved results with a precision of 0.80, a recall of 0.76, and an F1 score of 0.78. Siamese++ outperforms both, with the highest precision of 0.85, a recall of 0.89, and an F1 score of 0.87, highlighting its superior capability in balancing precision and recall compared to the other models.

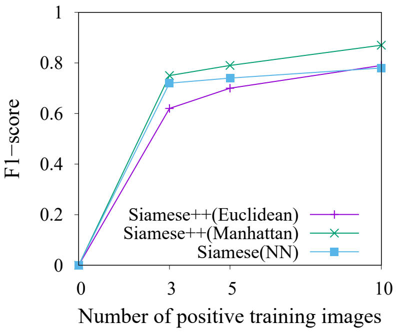

Figure 2: Impact of varying number positive class images in training data

Performance of Competing Approaches with Varying Number of Positive Training Images For few-shot learning experiments, we used 3, 5, and 10 images from the positive class during training to assess their impact on performance. For Manhattan distance, precision improves from 0.71 to 0.85, recall increases from 0.80 to 0.89, and the F1 score rises from 0.75 to 0.87 as the number of positive training samples increases, as shown in Table 3. In contrast, Euclidean distance shows lower precision and F1 scores compared to Manhattan distance, although recall values are somewhat competitive. Specifically, precision ranges from 0.53 to 0.74, recall ranges from 0.77 to 0.86, and the F1 score ranges from 0.62 to 0.795 with increasing training samples. Overall, Manhattan distance consistently outperforms Euclidean distance across all metrics as the number of positive samples grows.

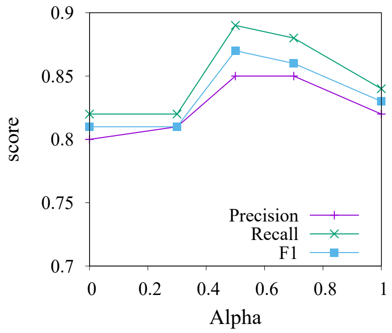

Figure 3: Sensitivity of parameter α on Precision, Recall, and F1 Score

Figure 2 illustrates the performance of three different model configurations in terms of F1 score as the number of images increases. All three configurations show an improvement in F1 score as the number of images increases. Initially, all models have low F1 scores with few images, but as the number of images rises, there is a rapid improvement in performance, with the scores leveling off after a certain point. The Siamese++ network with Manhattan distance achieves the highest F1 score across all image counts, followed by Siamese++ with Euclidean distance and the Siamese (NN).

Sensitivity of Parameter α Figure 3 shows the relationship between α values and the performance metrics: Precision, Recall, and F1 score. For this experiment, the value for μ^- is 0.46 and the value for μ^+ is 0.75. As α increases from 0 to 0.5, there is a noticeable improvement in all three metrics, with Recall peaking highest, followed by F1 score and Precision. After reaching the peak around α = 0.5, the performance metrics begin to decline as α continues to increase towards 1. This suggests that there is an optimal α value around 0.5 that provides a balanced decision boundary and a balanced value of precision and recall, leading to the highest F1 score.

6 Conclusion

This research demonstrates effective strategies for classifying children's drawings to determine the need for expert consultation. By employing both manually extracted features and those obtained through LLMs, we achieved high accuracy in distinguishing between 'need' and 'no need' categories. The introduction of the Siamese++ model, which addresses class imbalance through an adaptive distance threshold, outperformed traditional machine learning, deep learning, and Siamese approaches. This study highlights the effectiveness of integrating advanced feature extraction methods and innovative models to enhance diagnostic accuracy in psychological assessments based on visual data.

Acknowledgments

This work was funded by the United Arab Emirates University SURE+ Grant 12T078 (20-2024) and UAEU-ADU Joint Research Grant 12T041.

References

Alshahrani, A.; Mustafa, J. I.; Almatrafi, M. M.; and Aljabri, R. A. 2024. A Children's Psychological and Mental Health Detection Model by Drawing Analysis based on Computer Vision and Deep Learning. *Engineering, Technology and Applied Science Research*, 14(4): 15533–15540.

Baird, S.; Panlilio, R.; Seager, J.; Smith, S.; and Wydick, B. 2022. Identifying psychological trauma among Syrian refugee children for early intervention: Analyzing digitized drawings using machine learning. *Journal of Development Economics*, 156: 102822.

Becker-Weidman, A. 2017. *House-Tree-Person Projective Drawing Test*, 1–3. Cham: Springer International Publishing. ISBN 978-3-319-28099-8.

Bromley, J.; Guyon, I.; LeCun, Y.; Säckinger, E.; and Shah, R. 1994. Signature verification using a "siamese" time delay neural network. In *Advances in Neural Information Processing Systems*, volume 6, 737–744. Morgan Kaufmann.

Chawla, N. V.; Bowyer, K. W.; Hall, L. O.; and Kegelmeyer, W. P. 2002. SMOTE: synthetic minority over-sampling technique. *J. Artif. Int. Res.*, 16(1): 321–357.

Fan, L.; Sun, X.; and Rosin, P. L. 2021. Siamese Graph Convolution Network for Face Sketch Recognition: An application using Graph structure for face photo-sketch recognition. In *2020 25th International Conference on Pattern Recognition (ICPR)*, 8008–8014.

Ferdib-Al-Islam; and Akter, L. 2020. Early Identification of Parkinson's Disease from Hand-drawn Images using Histogram of Oriented Gradients and Machine Learning Techniques. In *2020 Emerging Technology in Computing, Communication and Electronics (ETCCE)*, 1–6.

Jurovatý, P.; and Démuthová, S. 2023. Draw a Person: Intellectual Ability Test for Children, Adolescents, and Adults–The Potential for Screening Diagnostic of Selected Cognitive Abilities. *Global Journal of Arts Humanity and Social Sciences*, 3(3): 313–323.

Kessler, R. C.; Berglund, P.; Demler, O.; Jin, R.; Merikangas, K. R.; and Walters, E. E. 2005. Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6): 593–602. Accessed: 2024-08-21.

Kotsavasiloglou, C.; Kostikis, N.; Hristu-Varsakelis, D.; and Arnaoutoglou, M. 2020. Machine learning-based classification of simple drawing movements in Parkinson's disease. *Journal of Neuroscience Methods*, 333: 108555.

Laak, J. T.; De Goede, M.; and Aleva, P. V., A.and Rijswijk. 2005. The Draw-A-Person Test: an indicator of children's cognitive and socioemotional adaptation. *The Journal of Genetic Psychology*, 166(1): 77–93.

Lee, M.; Kim, Y.; and Kim, Y.-K. 2024. Generating psychological analysis tables for children's drawings using deep learning. *Data and Knowledge Engineering*, 149: 102266.

Malchiodi, C. A. 1998. *Understanding Children's Drawings*. New York: Guilford Press. ISBN 978-1572303720.

Melekhov, I.; Kannala, J.; and Rahtu, E. 2016. Siamese network features for image matching. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, 378–383.

Pan, T.; Zhao, X.; Liu, B.; and Liu, W. 2022. Automated Drawing Psychoanalysis via House-Tree-Person Test. In *2022 IEEE 34th International Conference on Tools with Artificial Intelligence (ICTAI)*, 1120–1125.

Rakhmanov, O. 2021. Draw-a-Person hand-drawn sketches. https://www.kaggle.com/datasets/lachin007/drawaperson-handdrawn-sketches-by-children. Accessed: (02 July 2024).

Shukla, G.; Desai, B.; Mehta, P.; and Karamchandani, S. 2020. Holistic Siamese Model Optimized for Aged Face-Sketch Similarity Detection. In *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, 763–768.

World-Health-Organization. 2021. Adolescent mental health. Accessed: 2024-08-21.

Yan, J.; Li, D.; Deng, J.; Wang, H.; Long, Z.; Sun, W.; Xue, W.; Zhou, Q.; and Liu, G. 2022. A Siamese Network Fusion Time-Frequency Domain Features for Mental Workload Level Identification. In *2022 4th International Conference on Intelligent Information Processing (IIP)*, 1–4.