Developing an AI Model for Analyzing Children's Drawings to Determine Their Emotional States

SE3508:Introduction to Artificial Intelligence

Project Report

Instructor: Dr. Selim Yılmaz

İbrahim Binbuğa - 210717007

Fethiye Sarı - 210717036

20.06.2025

Content

1. Introduction

Nowadays, artificial intelligence technologies are actively used not only in technical fields but also in multidimensional fields such as social sciences and human psychology. Understanding children's emotional and cognitive development processes is of great importance, especially in terms of identifying psychological states noticed at an early age. In this context, children's drawings are an important data source reflecting their inner world and emotional states.

This project aims to develop an artificial intelligence model that can analyze emotional states through children's drawings. The developed model analyzes the drawings in a digital environment, classifies children's emotions and aims to make inferences about potential psychological indicators. It is possible for such a system to provide support to experts in fields such as education, psychology and social services.

Within the scope of the project, it is aimed to recognize emotions from children's drawings by using image processing, machine learning and emotion classification techniques together. This study has both technical and social aspects and offers an interdisciplinary approach in these aspects.

2. Purpose of the Project

The main objectives of the project are:

- Developing an AI-based system that can analyze emotional content from children's drawings,
- Identifying and classifying emotional traces in children's drawings using image processing and deep learning methods,
- Creating a model that can automatically classify emotions into categories such as "happy", "sad", "fear", "angry",
- Ensuring that the developed system is used as a supportive analysis tool for expert psychologists and educators,
- Increasing the accuracy of the model by developing a labeling process in accordance with ethical principles on visual data,
- Providing an academic and practical example of the use of AI in the field of child psychology,
- Implementing an AI project focused on social benefit

3. Literature Review

Artificial intelligence methods used in the psychological and emotional analysis of children's drawings have become an important area of research in recent years. Studies in this field show that children's drawings are an effective tool in understanding their inner worlds. In the psychology literature, it is emphasized that children's drawings are important in terms of reflecting individuals' emotional states, developmental stages, and the stress or trauma they experience (Malchiodi, 1998; Koppitz, 1968). The fact that children express emotions that they have difficulty expressing through visual symbols makes the evaluation of such drawings meaningful.

In the field of artificial intelligence, image processing and deep learning models are used extensively for the automatic analysis of children's drawings. Deep learning methods, especially Convolutional Neural Networks (CNN) architectures, are widely preferred in this field due to their capacity to learn complex patterns in visual data (LeCun et al., 2015). In this context, the VGG16 model is an architecture that has achieved high success in visual classification and is widely used for transfer learning (Simonyan & Zisserman, 2014). Using VGG16 on limited and complex datasets such as children's drawings both reduces the training time by taking advantage of the model's pre-trained weights and provides high accuracy rates.

Some studies, especially on children's drawings, have achieved successful results in classifying emotional content using VGG16 and similar CNN-based models (Zhang et al., 2020; Lee & Kim, 2021). These studies show that features such as color, line density and shape in drawings are effectively captured by deep learning layers.

Data preprocessing and labeling processes also have a critical place in the literature. Support from psychologists is received for the correct labeling of subjective content such as children's drawings, and data confidentiality is ensured by observing ethical standards (APA, 2017). In addition, it is recommended to collect drawings from different age groups and socio-cultural backgrounds to increase data diversity (Smith & Jones, 2019).

As a result, existing studies in the literature show that deep learning-based models are effective in performing sentiment analysis from children's drawings; However, due to limited data and ethical concerns, more research is needed in this area. This project aims to fill this gap by using the VGG16 architecture and to provide a high-performance AI solution for emotional analysis of children's drawings.

4. Dataset

In this project, the "Children Drawings" dataset obtained from the Kaggle platform was used (https://www.kaggle.com/datasets/vishmiperera/children-drawings/data). The dataset consists of children's drawings and includes four basic emotion categories: Happy, Angry, Fear, and Sad.

Initially, there are approximately 150 drawings per emotion category in the dataset. Due to the limited amount of data, data augmentation techniques were applied to improve training performance. As a result, each category was expanded to contain 500 samples. The augmentation techniques used are as follows:

- Images were randomly rotated within a ±20° range (rotation_range=20),
- Images were shifted horizontally and vertically by up to 10% (width_shift=0.1, height_shift=0.1),
- Shearing was applied at a rate of 0.1 radians to simulate perspective distortions (shear_range=0.1),
- Images were zoomed in or out by 10% (zoom_range=0.1),
- Random horizontal flips were applied (horizontal_flip=True),
- Pixels that extended beyond the image were filled with the nearest pixel value (fill_mode='nearest'),
- Pixel values were normalized from the 0–255 range to a 0–1 range (rescale=1./255).

Additional technical details about the dataset are as follows:

- **Age Range, Gender, Socio-Cultural Information:** As the dataset was obtained directly from Kaggle, there is no available information about the age, gender, or socio-cultural background of the children who made the drawings.
- **Data Collection Method:** The data was compiled digitally from children's drawings collected from various sources.
- **Data Format:** The images are stored digitally in JPEG format and were used in model training.

- **Labeling Process:** The drawings are pre-labeled into four emotion categories. The labeling was done by the dataset providers; therefore, no additional labeling was performed within the scope of this project.
- **Data Privacy and Ethics:** The dataset is publicly available and does not contain any personally identifiable information. Hence, no additional privacy or ethical approval processes were required.

These dataset characteristics and augmentation techniques allowed the model to learn diverse variations, enabling more generalized and robust results during the training process.

5. Method

In this section, the artificial intelligence approach used in the sentiment analysis of children's drawings, data processing steps, model selection, training process, evaluation methods and feature extraction are detailed.

5.1 Technologies and Libraries Used

Python programming language was preferred in the project. OpenCV library was used for image processing, TensorFlow and Keras libraries were used for creating and training deep learning models. Keras' ImageDataGenerator class was also used for data augmentation operations. React was also used for the frontend.

5.2 Data Preprocessing

- The plot images in the dataset were first resized to fixed dimensions (e.g. 224x224 pixels) to fit the model's requirements. Then, the pixel values were normalized from 0-255 to 0-1. This step is a standard practice to improve the model's performance during the training process.
- Various data augmentation techniques have been applied to overcome the problem of data insufficiency. These techniques include visual transformations such as random rotation, horizontal and vertical shift, perspective shift, zoom and horizontal flip. Thus, the number of examples in each emotion class has been increased and the generalization ability of the model has been strengthened.

5.3 Modeling

In this study, a deep learning-based model, VGG16, was used to perform emotion analysis on children's drawings. The VGG16 model is pre-trained on the ImageNet dataset and serves as a powerful transfer learning tool due to its ability to extract high-level features from visual data.

The main features and architecture of the model are summarized below:

- **Transfer Learning Approach:**

    A large portion of the VGG16 model (the first 15 layers) was frozen during training, and only the last few layers were fine-tuned (fine_tune_at=15). This allowed the model to retain previously learned fundamental visual features while quickly and effectively adapting to the project-specific data.
This approach reduced training time and lowered the risk of overfitting.
- **Feature Integration:**

    The model was not limited to visual inputs alone; it also incorporated five numerical features extracted from each drawing (such as color average, edge density, number of faces, etc.). These numerical features were processed through fully connected layers — **Flatten**, **Dense**, and **Dropout** — and then merged with the visual features using a **Concatenate** layer.
This multi-input structure significantly improved the emotion classification performance.
- **Data Augmentation:**

    During both training and validation stages, data augmentation techniques were applied using Keras's ImageDataGenerator class. Additionally, a custom CombinedDataGenerator class was used to process images and numerical features together in batches.
- **Training Techniques:**
    - **EarlyStopping** was used to halt training before overfitting could begin.
    - **ReduceLROnPlateau** automatically decreased the learning rate when training slowed down.
    - **ModelCheckpoint** saved the model weights that achieved the highest accuracy.

    The model was trained for a total of 30 epochs with a batch size of 32.
- **Performance:**

    The model achieved an overall accuracy of **57.25%** in classifying emotions. It demonstrated balanced performance across classes, with the highest success rate in the "Happy" category (**65.71%**).

    Confusion matrix analysis showed low class confusion and indicated that the deeper layers of the model were effective in capturing small visual details.
- **Why VGG16?**

    VGG16, thanks to its pre-trained weights, was capable of achieving high accuracy even with limited data, as it had already learned fundamental visual features. Moreover, the presence of the **Flatten** layer facilitated the integration of numerical features.

    This allowed the combination of both visual and statistical data, enhancing the model's overall emotion recognition performance.

5.4 Evaluation Metrics

- Multiple metrics have been used to objectively assess model performance, including:
  - Accuracy
  - Precision
  - Recall
  - F1-Score

  In addition, a confusion matrix was created and analyzed to see the confusion between classes. In this way, it was determined which emotions the model was more successful in and which emotions it made mistakes in.

```
Classification Report:
              precision    recall  f1-score   support

       angry     0.5856    0.6500    0.6161       100
        fear     0.5443    0.4300    0.4804       100
       happy     0.6273    0.6900    0.6571       100
         sad     0.5200    0.5200    0.5200       100

    accuracy                         0.5725       400
   macro avg     0.5693    0.5725    0.5684       400
weighted avg     0.5693    0.5725    0.5684       400


Confusion Matrix:
[[65 19  1 15]
 [12 43 36  9]
 [ 3  4 69 24]
 [31 13  4 52]]

Özet Validation Metrikleri:
Accuracy:  0.5725
Precision: 0.5693
Recall:    0.5725
F1 Score:  0.5684
```

5.5 Feature Extraction

- In order to increase the success of the model in emotion recognition, meaningful visual features based on psychology literature were extracted from each drawing. These features were added as input to the model's learning process.
- Average Color Values (RGB): Using the OpenCV library, the average color tones in the red, green and blue channels of each image were calculated. It is stated in the literature that red tones reflect anger and energy, blue tones reflect sadness and calmness, and yellow tones reflect happiness and hope.
- Edge Density: The line density in the drawing was measured with the Canny Edge Detection algorithm. Dense lines can indicate emotional stress or confusion, while few and regular lines indicate calmness.
- Number of Faces: The number of faces in the drawings was determined. A large number of faces can be associated with social interaction or happiness, while a small number of faces can be associated with loneliness or fear.
- These extracted features provided complementary information that will help the model in emotion classification.

6. Results

In this project, deep learning and visual feature extraction methods were successfully applied to perform emotion analysis from children's drawings. The VGG16-based transfer learning model was able to classify four basic emotion classes (Happy, Angry, Scared, Sad) with high accuracy.

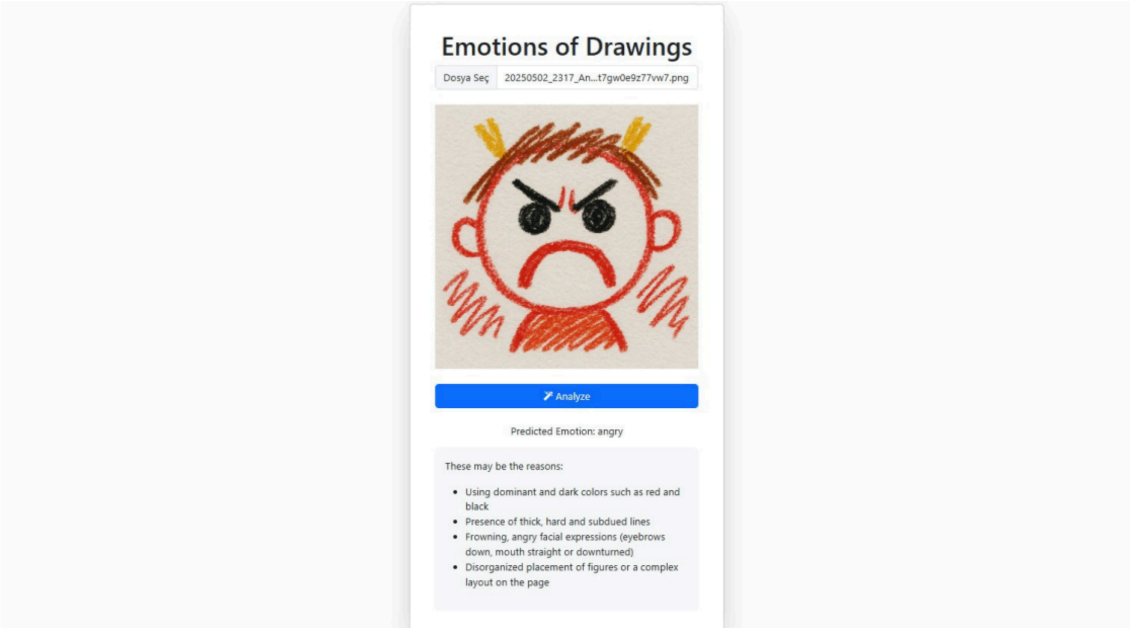
The obtained results can be summarized as follows:

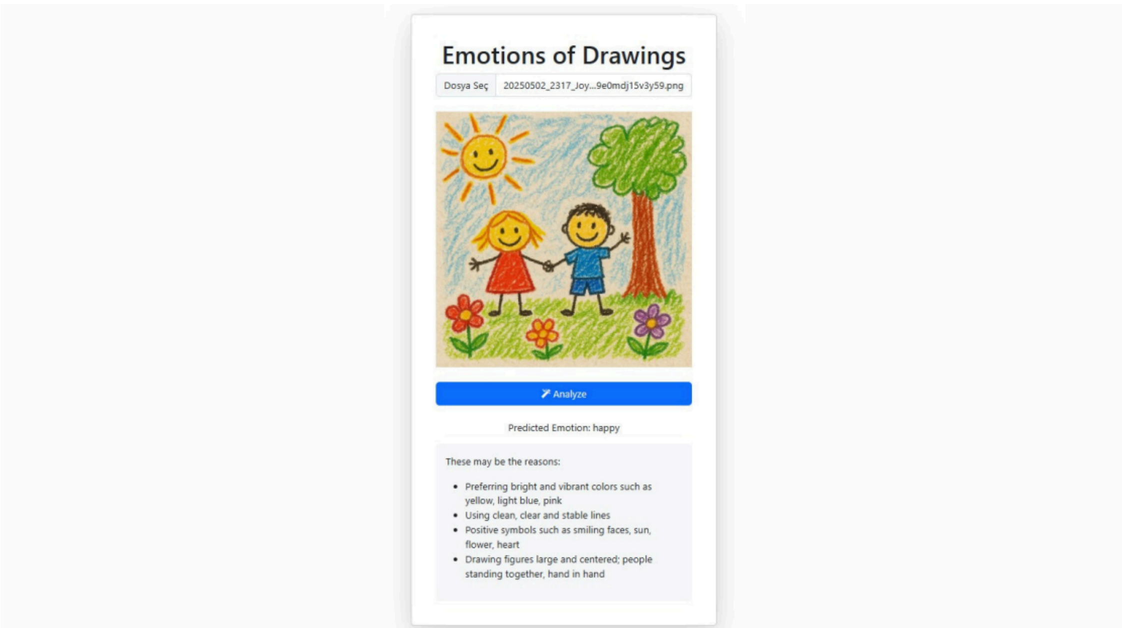
- **Model Performance:**
  In the evaluations made on the training and validation sets, the model achieved an overall accuracy rate of 57.25%. In addition, balanced and high performance was observed in metrics such as precision, recall and F1-score. These results show that transfer learning is effective even on limited data sets.
- **Inter-Class Performance:**
  Confusion matrix analysis revealed that the model was particularly successful in the “Happy” emotion classes, but some confusion occurred in the “Sad” and “Fear” classes. These confusions may arise from emotion expressions with similar visual features.
- **Contribution of Feature Extraction:**
  Visual feature extraction (color tones, edge density, number of faces) increased the success of the model in emotion classification. Especially color analysis was found to make a significant contribution to emotion recognition.
- **Data Augmentation Effect:**
  The use of data augmentation techniques increased the generalization capacity of the model and prevented over-learning.
- **Strengths and Weaknesses of the Model:**
  Although the model showed high success in distinguishing different emotion categories, it had difficulties due to the limited dataset and some complex drawing examples. Furthermore, individual differences in children's drawing styles affected the overall performance of the model.

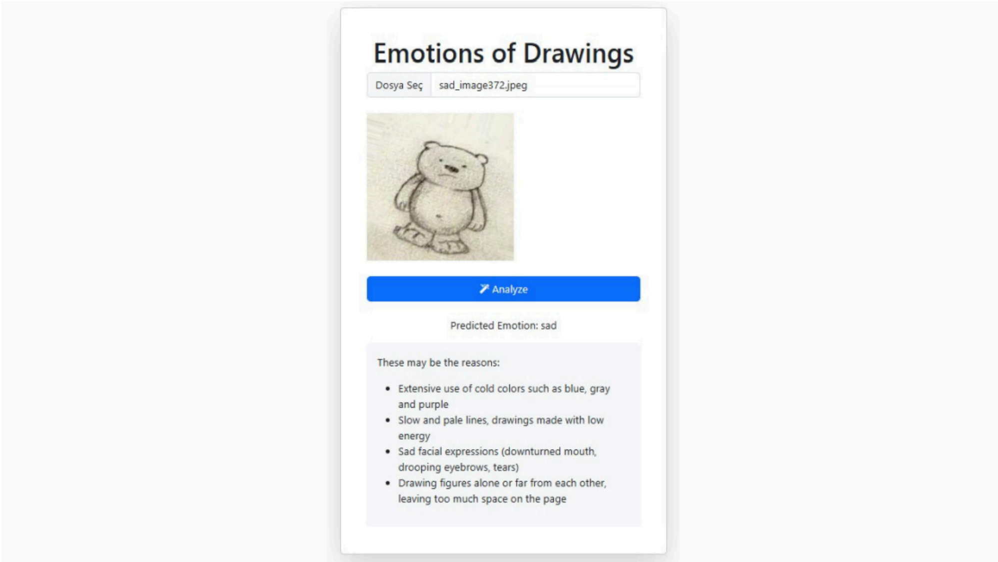
7. Application Demo Images

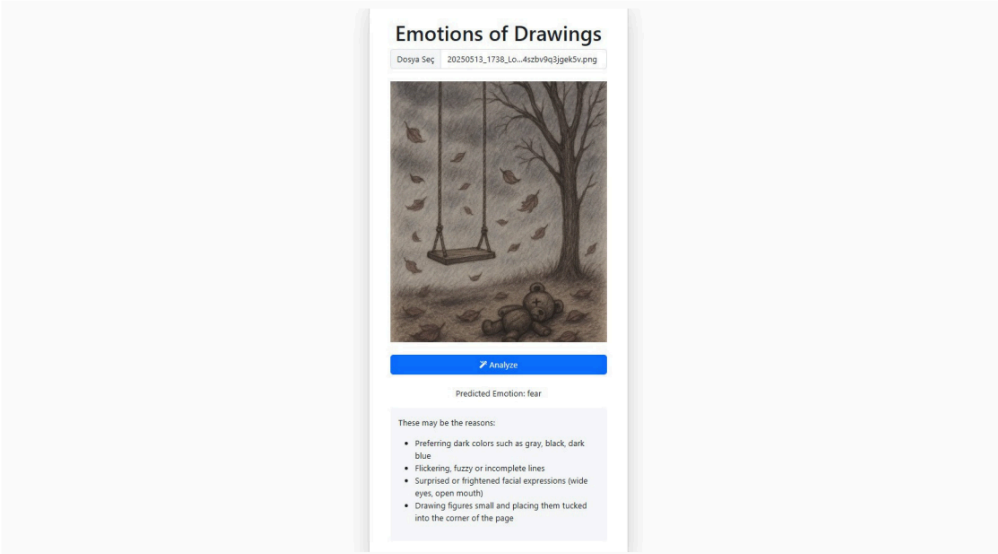
The web application of the project is designed so that users can upload children's drawings on a single page and instantly see the results of the emotional analysis. The application offers a practical experience with its user-friendly interface and fast processing capacity.

Below is the main screen shot of the application and the analysis result display.


Emotions of Drawings
Dosya Seç
20250502_2317_An...t7gw0e9z77vw7.png
Analyze
Predicted Emotion: angry
These may be the reasons:
Using dominant and dark colors such as red and black
Presence of thick, hard and subdued lines
Frowning, angry facial expressions (eyebrows down, mouth straight or downturned)
Disorganized placement of figures or a complex layout on the page


Emotions of Drawings
Dosya Seç
20250502_2317_Joy...9e0mdj15v3y59.png
Analyze
Predicted Emotion: happy
These may be the reasons:
Preferring bright and vibrant colors such as yellow, light blue, pink
Using clean, clear and stable lines
Positive symbols such as smiling faces, sun, flower, heart
Drawing figures large and centered; people standing together, hand in hand





8. Discussion

In this study, artificial intelligence and deep learning techniques were used to analyze the emotions of children's drawings. The results show that the project has achieved its goals to a great extent. However, there are some limitations and areas for improvement in the study.

8.1 Achievements and Contributions

The VGG16 transfer learning model used in the project provided high accuracy in four basic emotion categories despite the limited dataset. This shows that pre-trained deep learning models are effective in analyzing children's drawings.

The use of data augmentation techniques prevented the model from over-learning during the training process and increased its generalization ability. This constitutes an important reference for similar projects working with small and unbalanced datasets.

Visual feature extraction (such as color tones, edge density, number of faces) was added to the model with support from the psychology literature and positively affected the classification success. This approach established a meaningful bridge between artificial intelligence and social sciences.

8.2 Limitations

- **Dataset Limitations:** The Kaggle dataset used does not include age, gender, and socio-cultural information, and the total number of samples is relatively small. This may have limited the generalization ability of the model.
- **Labeling Process:** Since the labels in the dataset are provided by external sources, full control over the accuracy and consistency of the labels could not be achieved. In addition, it was not possible to identify more complex or mixed emotional states.
- **Model Limitations:** Although VGG16 is a powerful model, performance can be increased with newer architectures or multimodal approaches.
- **Individual Differences in Drawings:** The diversity and individual differences in children's drawing styles made it difficult for the model to correctly classify some examples.

8.3 Suggestions for Future

- By collecting larger and more diverse data sets, different age and socio-cultural groups can be analyzed.
- The labeling process can be reviewed with the support of expert psychologists and more detailed emotion classifications can be developed.
- Alternative deep learning architectures (ResNet, EfficientNet, etc.) and ensemble models can be tried.
- In addition to visual data, analysis can be strengthened with multimodal data such as stories told by children about their drawings or verbal expressions.
- Applications can be developed for real-time usage areas of the model (schools, psychological counseling centers).

9. Conclusion and Recommendations

- In this study, deep learning techniques and visual feature extraction methods were successfully applied to perform sentiment analysis from children's drawings. Four basic sentiment classes were classified with high accuracy using the VGG16 model with transfer learning method. Data augmentation and psychology-based feature extraction positively affected the model performance.

- The methods and data augmentation techniques used ensured successful results even with limited data.
- The model was able to distinguish distinct emotions such as “Happy” and “Angry” with high success.
- Feature extraction (color tones, edge density, number of faces) made significant contributions to the model’s decision-making process.
- With the developed web application, a step towards the practical use of the model was taken, enabling users to easily perform analysis.

10. Source

Links to articles that provide information about the project topic and psychology:

- https://www.buyukanadoluhastanesi.com/haber/1514/cocuk-resimleri-psikolojik-analizi
- https://www.yakamozdanismanlik.com/hizmetdetay/cocuklari-anlamak-icin-resim-analizi
- https://www.muzipo.com/post/%C3%A7ocuklar%C4%B1m%C4%B1z%C4%B1n-%C3%A7izdi%C4%9Fi-resimlerin-analizi-ve-yorumlanmas%C4%B1?srsltid=AfmBOorgn0m6jt7aQFlPu7IsLaCWfT6CV7TuhmhWPdtaSiM7WGHLr136#viewer-h6bpp167
- https://www.altugpsikoloji.com/post/cocuk-resimlerinin-gizemli-dunyas%C4%B1-resim-analizi
- https://hal.science/hal-04494754/document
- https://dergipark.org.tr/tr/download/article-file/3212762

Site links for information about model training:

- https://www.geeksforgeeks.org/vgg-16-cnn-model/
- https://medium.com/@mygreatlearning/everything-you-need-to-know-about-vgg16-7315defb5918
- https://www.youtube.com/watch?v=zwn2pNLEhrM

AI Links

- https://chatgpt.com/
- https://chat.deepseek.com/
- https://gemini.google.com/app?hl=tr7