Project Plan

Visual Computing and Games Technology (MSc)

Tell Me mAI Story

from

Vanessa Skowronsky

Sven Ludwig

Hoan Vu

2023-11-09

1 Work Packages

1.1 Functional Area - Processing and Storage of Children's Drawings

This section covers the topic of ingesting new previously unseen children's drawings, i.e. when a child has created a drawing and first presents it to our AI.

1.1.1 Work Package - Basic dialog for uploading a new, previously unseen drawing

Upon starting a new session, the tool initiates with a warm welcome greeting to users (children). It will then explain the concept and present to children a catalog of animals to choose from and draw on their paper. After the user uploads their drawings (supported image formats are JPEG and PNG), one of these scenarios could happen:

- If there is no detectable object, tell the user to draw an animal chosen from the catalog. (note that the term object here refers to any recognized object from the drawing, it could be anything from a shape to an animal that is from the catalog or not)

- If there are detected objects, based on the classification score for each object, verify with the user if the classification is correct (by showing the detected animal region and the predicted animal name)

The tool provides a friendly farewell when the interaction session concludes.

1.1.2 Work Package - Curation of the training dataset of drawings

First batch of dataset is collected from a kindergarten. To prepare the dataset for training, some preprocessing steps are required:

- Resize the image to a fixed size

- Annotating bounding boxes coordinates and labels of animals in a csv file

- Split the dataset into training and test set

For a wider range of recognition ability and better accuracy, we could take use of public dataset. Some available datasets are: Google Quick, Draw! dataset (hand drawn doodles) and TU-Berlin Human Sketch Recognition.

1.1.3 Work Package - Experiments with preprocessing of drawings

1.1.4 Work Package - Preprocessing pipeline of drawings

1.2 Functional Area - Recognition of animals

This section covers the solution to the problem of recognizing animals in children's drawings, that is recognizing that an animal is present and classifying the animal.

1.2.1 Work Package - Experiments with a first CNN for recognizing animals

Implement a first CNN for recognizing animals in terms of a binary classification (animal present vs. animal not present) and evaluate its performance. Frameworks could be tested are: PyTorch, TensorFlow, Keras, etc.

1.2.2 Work Package - CNN for recognizing animals in terms of a multi-class classification

Implement a CNN for recognizing animals in terms of a multi-class classification (recognizing the animal and classifying it) and evaluate its performance.

1.2.3 Work Package - CNN for recognizing multiple animals in a single drawing

Implement an advanced CNN model (e.g. Faster R-CNN, Mask R-CNN) for recognizing multiple animals in a single drawing, localize the animal with bounding boxes and classify each one of them. Evaluate its performance.

1.3 Functional Area - Advanced recognition of animals (optional)

This section is optional. We want to look at alternative advanced models for recognizing animals in children's drawings, in particular we want to consider recent state-of-the-art models and concepts that have been developed by researchers and companies in the years 2020-2023.

1.3.1 Work Package - Theoretical consideration of alternative models for animal recognition

- ConvNet: arXiv link

- CoTNet: arXiv link

- VAN: arXiv link

- MAE: arXiv link

- ViT: arXiv link

- Faster R-CNN arXiv link

- YOLO (fightweight, real time recognition but might be not as accuracy as Faster RCNN) arXiv link

- EfficientDet (lightweight, suitable for small size dataset) arXiv link

1.4 Functional Area - Annotating animals with digits to denote their importance

This section covers the topic of annotating animals with digits to provide the model a clue about their importance in the story. An animal annotated with a 9 would be a main character, while an animal annotated with a 1 would just appear somewhere in the background of the story.

1.5 Functional Area - Annotating animals with markers to denote their relations

Markers are used to indicate which characters should meet with and talk to which other characters during the story.

1.5.1 Work Package - Evaluate different marker types and choose the marker type to use

1.6 Functional Area - Annotating animals with activity cards to denote their actions

Pre-defined activity cards can be placed onto the picture to indicate which characters should take part in which activities (e.g. eating, drinking, jumping, going for a walk) during the story. (The activity cards are considered a potential feature, still to be prioritized with respect to other features and details.)

1.6.1 Work Package - Sketch and design various activity cards

1.6.2 Work Package - Experiment with various styles in terms of art

1.6.3 Work Package - Create the activity cards digitally

1.6.4 Work Package - Define and make first use of a process for producing physical activity cards

1.7 Functional Area - Generating a children's story based on a children's drawing in text form

1.8 Functional Area - Telling a previously generated children's story using elegant text-to-speech