St. Pölten University of Applied Sciences

/ medien & digitale technologien

/fh/// st.pölten

A Machine-Learning Based Approach for 2D Character Animation

<Subtitle>

Bachelor Thesis

Bachelor Course on Creative Computing
at St. Pölten University of Applied Sciences

Submitted by:
Georg Becker
01228308

Advisor: <Pre-Title> <FirstName> <LastName>, <Pos-Title>

<Place>, <DD>.<MM>.<YYYY>

Declaration

I assure that

- I have written this work independently, have not used other sources and aids than those indicated and have not made use of any other unauthorized assistance.

- I have not yet submitted this topic to an assessor in Austria or abroad for assessment or in any form as an examination paper.

- this work corresponds to the work assessed by the assessor.

Date: _____ Signature: _____

Abstract

Introduction: Warum behandeln wir das Thema

Purpose: Welches Problem soll gelöst werden

Method: Wie wurde die Problemlösung gemacht

Product: Was war das Ergebnis

Conclusion: Was sind die Folgerungen / Schlussfolgerungen aus den gewonnen Erkenntnissen

keine Referenzen und Zitate

Contents

1 Introduction

2 Method

Literature review

I reviewed previous work, focusing on two areas. I explored already available methods for creating animations from sketches by performing skeleton classification and reviewed previous work dealing with the classification of sketched objects.

Related work

Eitz et al. (2012) collected a dataset of 20,000 sketches and divided them into 250 categories of 80 images each. Humans recognized on average 73.1% of these sketches correctly. This dataset is used in my work to train and validate the classifier to choose which animation is the most appropriate to show.

Huang et al. (2022) proposes a pipeline to create rigged and animated characters from a single image. Their solution aims for a holistic approach, requiring no user intervention, to assist non-professional users in creating animated characters. The proposed pipeline performs contour extraction with salient object detection and extrudes a 3D mesh from geometry generated by applying constrained Delaunay to the contours. Afterwards, a skeleton is estimated using a mean curve method and an animation is transferred onto the skeleton. In our work, we want to follow a similar philosophy of no user interaction and hope to improve the believability of the animated results by not only classifying the skeleton type but also the subject class of the input sketch.

Training classification models

We used a subset of the dataset provided by Eitz et al. (2012) and Sangkloy et al. (2016) to train our classification models. Only the classes "cat" and "dog" were taken as training data for our models. To train and evaluate our models we used the scikit-learn library introduced by Pedregosa et al. (2011).

kNN classifier

We trained a kNN classifier with the pixel values of the input images. Before using the values to train the model, we resized the images to a size of 64 times 64 pixels, and flattened the array to get a feature vector with 12288 entries, ranging from 0 to 255 in value. To find the best-performing k, we performed a grid search with cross-validation on 3 folds leading to k = 5 as the model with the highest accuracy at 61.8798%.

SVM classifier

We trained an SVM classifier with a total of 1544 labeled images of sketches of cats and dogs. Before training the model we resized the images to a size of 64 times 64 pixels, and flattened the array to get a feature vector with 4480 entries. The images were imported as grayscale images. The SVM classifier performed with an accuracy of 53.7578%.

CNN classifier

Using the same preprocessing as the kNN classifier I created a neural network classifier using pytorch Paszke et al. (2019). This is the network's setup:

```
class MyNN(nn.Module):
    def __init__(self):
        super().__init__()
        self.flatten = nn.Flatten()
        self.linear_relu_stack = nn.Sequential(
            nn.Linear(64*64*3, 16*16),
            nn.ReLU(),
            nn.Linear(16*16, 16*16),
            nn.ReLU(),
            nn.Linear(16*16, 2),
            nn.Sigmoid()
        )
    def forward(self, x, **kwargs):
        x = self.flatten(x)
        logits = self.linear_relu_stack(x)
        return logits
```

Performing a grid search with cross-validation on 2 folds on this network leads to a network with a loss rate of 0.001 and 1 epoch as the best-performing setup with an accuracy of

51.25%. This indicates to me, that the model is prone to overfitting and we probably need to augment or expand the training data.

Implementing the pipeline

For this work, we reimplemented the pipeline proposed by Korpitsch (2023), and adapted the code where needed.

Sketchdetection

To repeat the steps introduced by Korpitsch (2023) we collected the dataset provided by Sarvadevabhatla et al. (2017). In Smith et al. (2023), a Mask-R-CNN, as described by He et al. (2018), is used to detect the bounding boxes of figures drawn by children.

3 Results / Ergebnisse

Presenting found literature in a useful way

3.1 First Section

Ich bin Text, Text, Text[1]

3.1.1 First Subsection

[1] http://mfg.fhstp.ac.at

4 Discussion / Diskussion

Comparison of presented technologies/methods/projects

Kritische Diskussion / Vergleich der Ansätze

Welche Methoden werden zumeist genutzt, warum?

Überblick / Zusammenfassung der gefundenen Literatur in einer sinnvollen Kategorisierung / Charakterisierung

5 Conclusion / Fazit

Was kann man daraus lernen?

Was fehlt?

Ideen für zukünftige Forschung

Bibliography

Eitz, M., Hays, J., & Alexa, M. (2012). How do humans sketch objects? *ACM Trans. Graph. (Proc. SIGGRAPH)*, *31*(4), 44:1–44:10.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2018). Mask r-cnn.

Huang, Z., Han, R., Huang, J., Yin, H., Qin, Z., & Wang, Z. (2022). Automatically generate rigged character from single image. *ACM Multimedia Asia*. https://doi.org/10.1145/3469877.3490565

Korpitsch, T. (2023). *Semantic-aware animation of hand-drawn characters* (Master's thesis). Research Unit of Computer Graphics, Institute of Visual Computing and Human-Centered Technology, Faculty of Informatics, TU Wien. Favoritenstrasse 9-11/E193-02, A-1040 Vienna, Austria. https://www.cg.tuwien.ac.at/research/publications/2023/korpitsch-2023-sao/

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., ... Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In *Advances in neural information processing systems 32* (pp. 8024–8035). Curran Associates, Inc. http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Sangkloy, P., Burnell, N., Ham, C., & Hays, J. (2016). The sketchy database: Learning to retrieve badly drawn bunnies. *ACM Trans. Graph.*, *35*(4). https://doi.org/10.1145/2897824.2925954

Sarvadevabhatla, R. K., Dwivedi, I., Biswas, A., Manocha, S., & Babu, R. V. (2017). Sketchparse : Towards rich descriptions for poorly drawn sketches using multi-task hierarchical deep networks.

Smith, H. J., Zheng, Q., Li, Y., Jain, S., & Hodgins, J. K. (2023). A method for animating children's drawings of the human figure.

List of Figures

List of Tables

Listings

Appendices

A Appendix

LoHrem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.LoHrem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. LoHrem

B Appendix

LoHrem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.LoHrem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. LoHrem