Application of 3D graphic synthetic dataset generation for the means of image alpha matting using deep neural networks

Machine Learning

Maksym Gontar, max.gontar@gmail.com

Ukrainian Catholic University
APPLIED SCIENCES FACULTY

Idea:

To use synthetic dataset, created using 3D graphics modeling, for training deep neural network model, which aim is to solve task of image alpha matting (foreground object masking)

Goal

Use synthetic 3D graphics dataset to train deep neural network model for image processing

Approach

Use Python scripting API in Blender (3D graphics editor) to render 3D object model from different angles with transparent background

Use ImageMagic to extract object masks and combine rendered images with real world photos as a background for object

Results

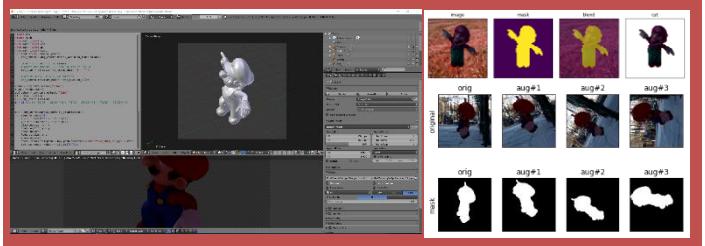
Trained UNet-based model shows considerable precision on validation part of the same dataset, and some precision on unseen objects

Remaining Work

Use more varieties for data generation:

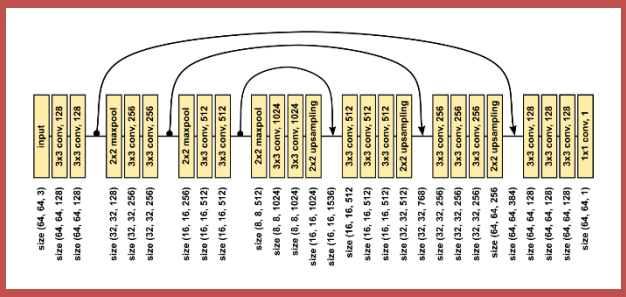
- **different 3D objects, materials, lightning**
- **Randomly generated 3D models**
- **more different background photos**

Dataset



To generate the dataset for this project Blender (open source 3D graphic software) was used. A ready to use 3D model Mario with creative commons license was picked, render was set up with transparent background to PNG and a python script was used to render the scene while rotating camera around the object - a sequence of 12 images with step of 30 degrees rotation in horizontal plane.

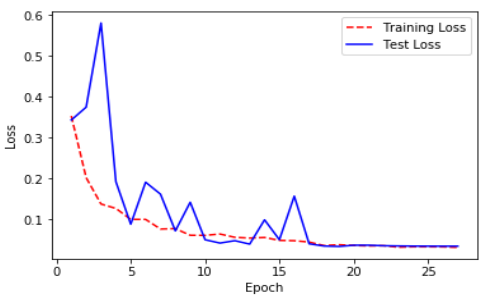
Model



Residual convolution autoencoder UNet-based model, implemented by one of Kaggle competitors of Carvana Challenge, adapted for 64x64x3 input. Effectiveness of the UNet based architecture for object masking task can be explained by the fact that low level feature extraction layers are contributing the most to output of the model - and here low level features like the edges are important.

Evaluation

Loss plot of the model training



Loss function: combination of binary cross-entropy and Dice loss. Training set up: 100 epochs, with early stopping condition on minimum loss delta. Model converged with 0.9797 accuracy after 27 epochs. Prediction object mask with the same dataset shows high precision. Also, the model was evaluated with several samples from Alpha Matting Evaluation Website. Some samples shows high precision as well, some shows poor precision, some fails to predict object mask whatsoever.