Voxel-Based 3D Object Generation from Single Images Using an Enhanced Deep Learning Architecture

A. Pocius, T. Blažauskas, E. Butkevičiūtė

Faculty of Informatics, Kaunas university of technology

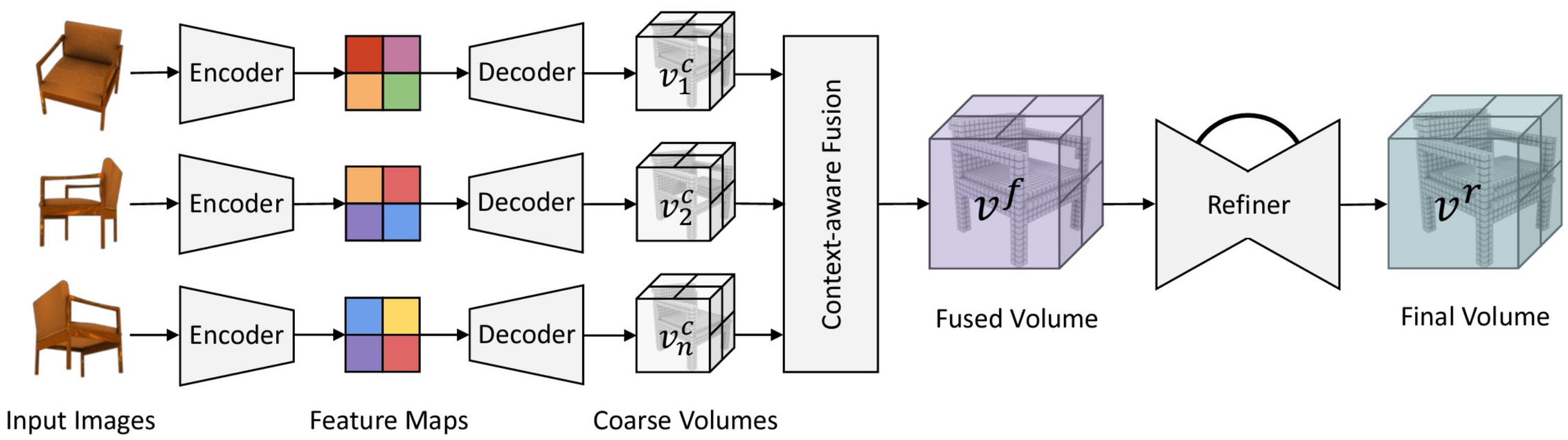
PROJECT DESCRIPTION

The core objective of this project is to refine a deep learning algorithm for generating 3D models from single-view images. This algorithm leverages the ShapeNetCore dataset and draws from advanced computer vision techniques to develop an optimized solution for voxel-based 3D object reconstruction. Inspired by existing models, the project addresses practical challenges to improve efficiency and accuracy.

ABSTRACT

Deep learning has significantly advanced 3D modeling, enabling accurate object generation from single-view images. This project aims to improve computer vision and graphics methodologies by optimizing deep learning algorithms to enhance efficiency in 3D object generation. Using the ShapeNetCore dataset, which includes over 51,000 unique models, a custom model was trained in PyTorch to create voxelized 3D objects [1]. Key challenges like geometric complexity and computational cost were addressed with optimized preprocessing and design. Evaluation with Chamfer Distance and IoU metrics yielded promising results, with an IoU score of 0.6549. This work contributes an adaptable and accurate solution with potential applications in design, virtual reality, and gaming.

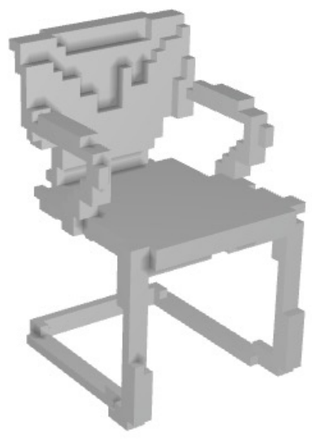
MODEL ARCHITECTURE [2]



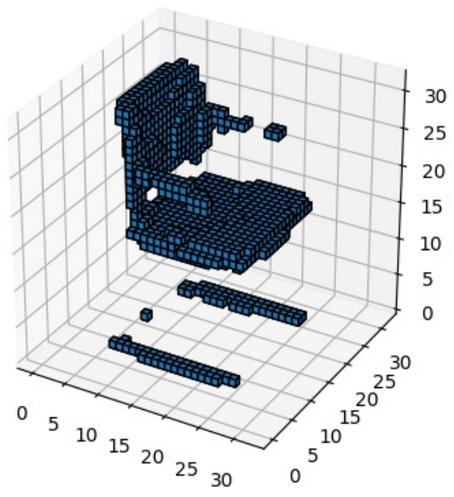
Results

Preliminary experimental results demonstrate that the proposed model effectively generates accurate 3D objects from single images, achieving an overall IoU score of 0.6549. These initial findings suggest that the model performs well across various object categories, however the model struggles with thin lines as they tend to not be rendered

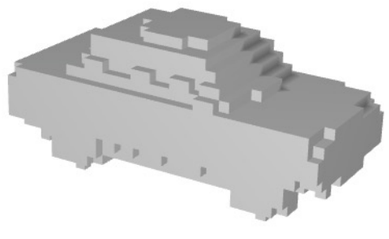
Input image

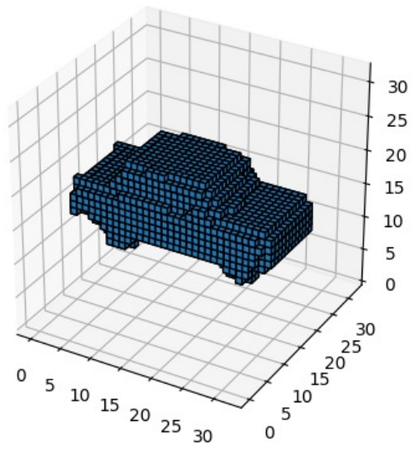


Voxel Ground truth



Predicted voxel model





Conclusions

This work contributes to the field of 3D object generation by presenting an optimized deep learning solution that enhances accuracy of reconstructed objects. The model's adaptability to various object categories and its potential applications in computer-aided design, virtual reality, and game development highlight its significance in advancing 3D modeling technologies.

REFERENCES

[1] Angel X. Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, Hao Su, Jianxiong Xiao, Li Yi, Fisher Yu, “ShapeNet: An Information-Rich 3D Model Repository“, 2015.

[2] Haozhe Xie, Hongxun Yao, Xiaoshuai Sun, Shangchen Zhou, Shengping Zhang, “Pix2Vox: Context-aware 3D Reconstruction from Single and Multi-view Images”, 2019.