Imagining Hidden Supporting Objects using Volumetric Conditional GANs and Differentiable Stability Scores

Hector Basevi[1,2] and Aleš Leonardis[1]

[1]School of Computer Science,
University of Birmingham

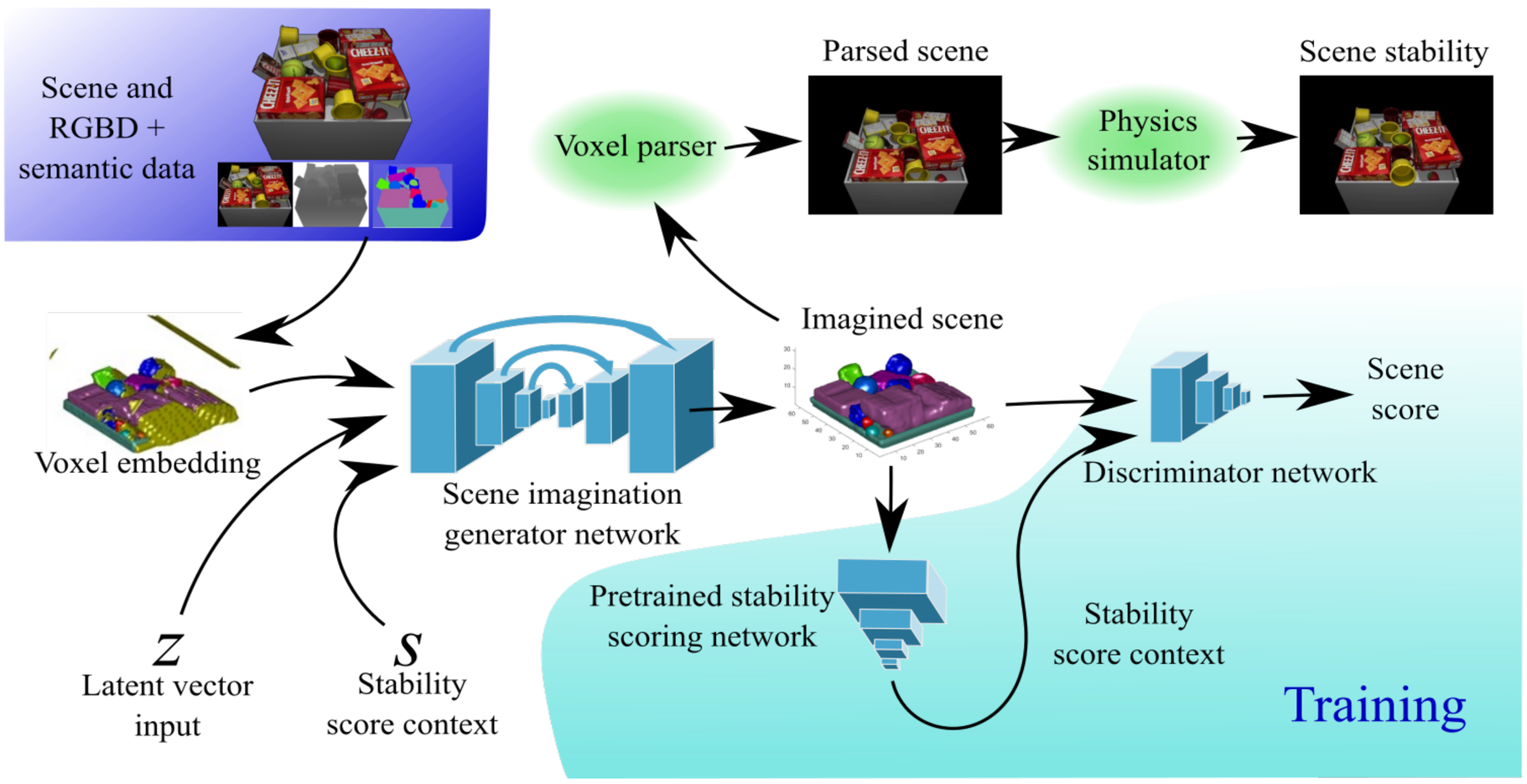
[2]h.r.a.basevi@bham.ac.uk

Abstract

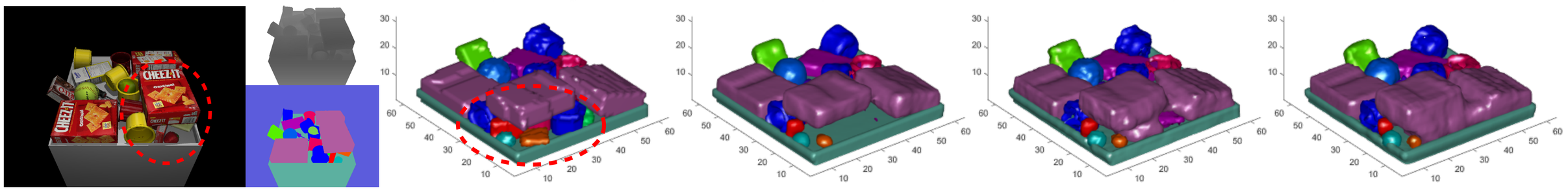
Objects supporting the physical stability of an unstructured heap of items are often heavily or *completely* occluded by the objects that they are supporting. Identifying plausible supporting object candidates and their poses from visual information is challenging because there may be many candidates and it is not practical to exhaustively verify each one using physical simulation. We present a generative system which predicts the complete volumetric structure of a heap of objects from visible depth and semantic information. We leverage 3D conditional Wasserstein generative adversarial networks to perform this task and inject differentiable context about physical stability from a second network trained to score the physical stability of object heaps. We demonstrate that our system is capable of generating physically stable heaps from visual information, and that the use of both generative models *and* context about physical stability are crucial in replicating the true distribution of hidden objects. We train and evaluate our system using a novel simulation-based dataset which we also present in this work.

Method

- Our scenario and datasets contain complex physical behaviour, high occlusion, and hidden objects, generated using physical simulation.
- Our conditional Wasserstein generative adversarial network is conditioned on visible depth and semantic information.
- Semantic voxel representations make shape, and object overlap, explicit and accessible.
- Our separately trained stability scoring network provides context for generation and discrimination.
- Our separate stability dataset teaches the effects of object perturbations and removal.
- Our semantic voxel representation can be parsed back into individual objects and poses via an ICP-based parser.



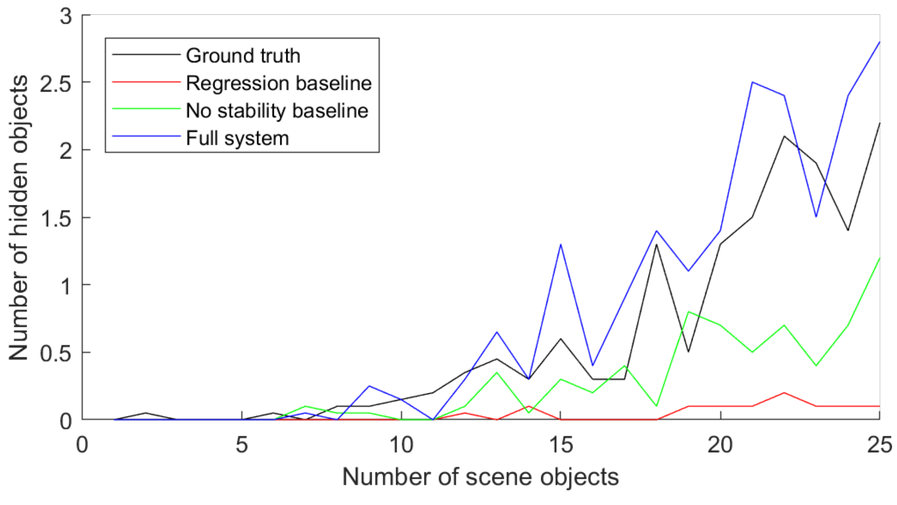
System design: A scene imagination generator network takes a volumetric embedding of visual semantic information, and produces an output volumetric embedding of a complete scene including occluded regions. An additional random latent vector input allows the generator to sample from the output distribution, and a stability score context conditions both the generator and discriminator on expected stability. The stability score is predicted by a stability scoring network trained on a separate dataset which includes multiple modified versions of each scene.



Input data | **Ground truth** | **Regression** | **Generative without stability context** | **Generative with stability context**

Scene generation above: The bottom portion of the scene (circled in red) is occluded and contains supporting objects. Each system, trained with different objectives, produces a different prediction for this region, and only our full system produces a physically stable scene.

Scene generation below: The ground truth data contains on average several hidden objects for scenes composed of many objects and only our full system closely replicates this distribution of hidden objects.



Results

- Regression fills occluded regions with empty space.
- Generation without stability context produces some imagined objects, which are often unstable.
- Stability context results in similar numbers of hidden objects to ground truth data, and more stable scenes.
- Stability context reduces sample diversity.

Conclusions

- We have presented a novel system and datasets for scene imagination.
- Stability context influences the number of imagined hidden objects *and* the scene stability.
- Voxel-based computation is effective but expensive.
- Future work could incorporate a differentiable mechanism for conversion between object instance and voxel representations.

We acknowledge MoD/Dstl and EPSRC for providing the grant to support the UK academics involvement in a Department of Defense funded MURI project through EPSRC grant EP/N019415/1. The research in this paper was supported in part by the Engineering and Physical Sciences Research Council (grant number EP/S032487/1).

https://hectorbasevi.github.io/imagining-hidden-supporting-objects