Generating Synthetic CAD Models for Deep Learning

Collaborators: George Weinert, Michael Ward, Mark Zelinski, Alan Noun, Max Klein

DSSI DATA SCIENCE SUMMER INSTITUTE

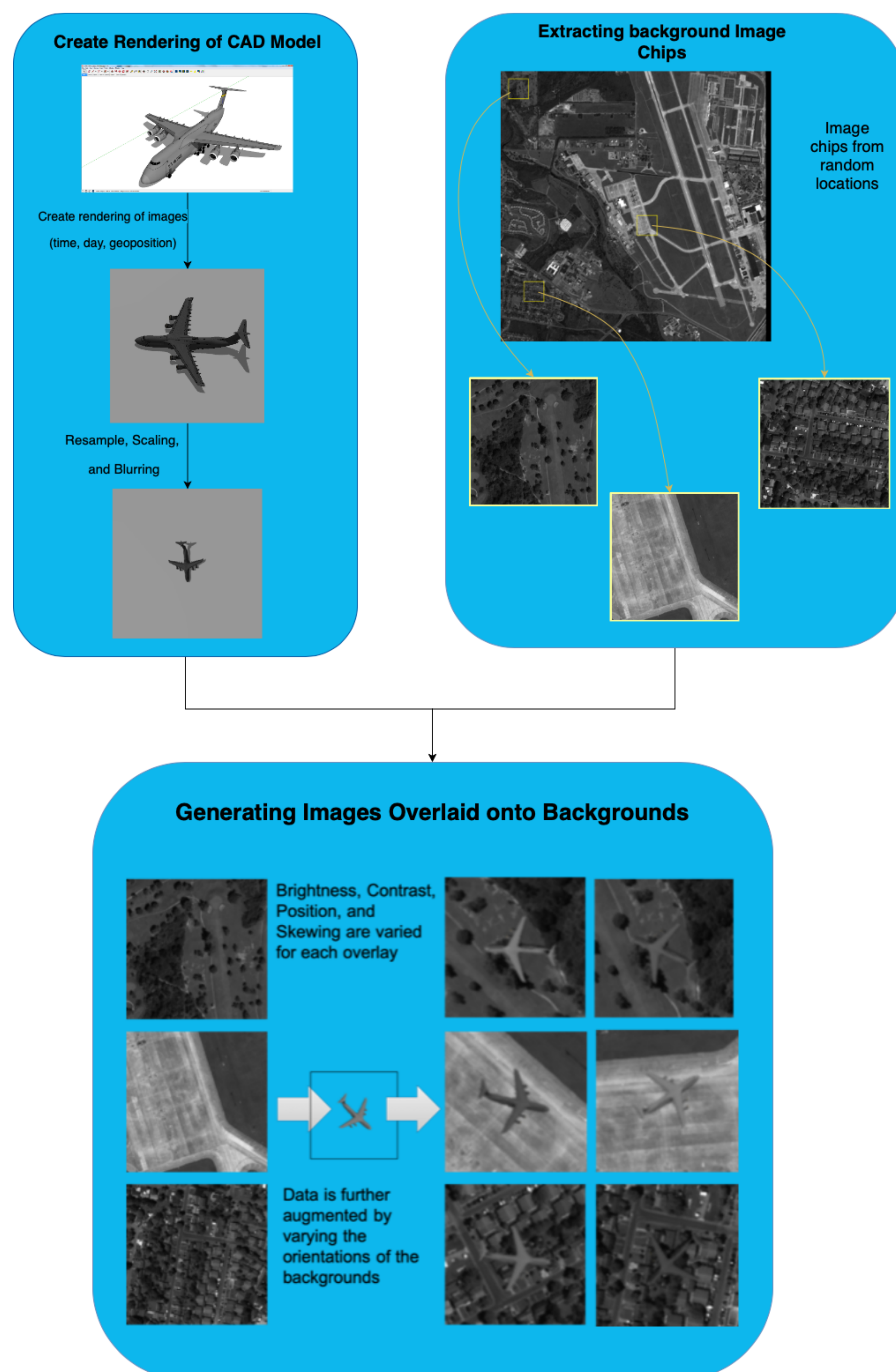Lawrence Livermore National Laboratory

Context

- Need to classify objects in overhead imagery, e.g. Aircraft, satellite
- CAD models are rarely seen in existing imagery
- Models need synthetic imagery with realism

Generating Training Data

Create Rendering of CAD Model

Create rendering of images
(time, day, geoposition)

Resample, Scaling, and Blurring

Extracting background Image Chips

Image chips from random locations

Generating Images Overlaid onto Backgrounds

Brightness, Contrast, Position, and Skewing are varied for each overlay

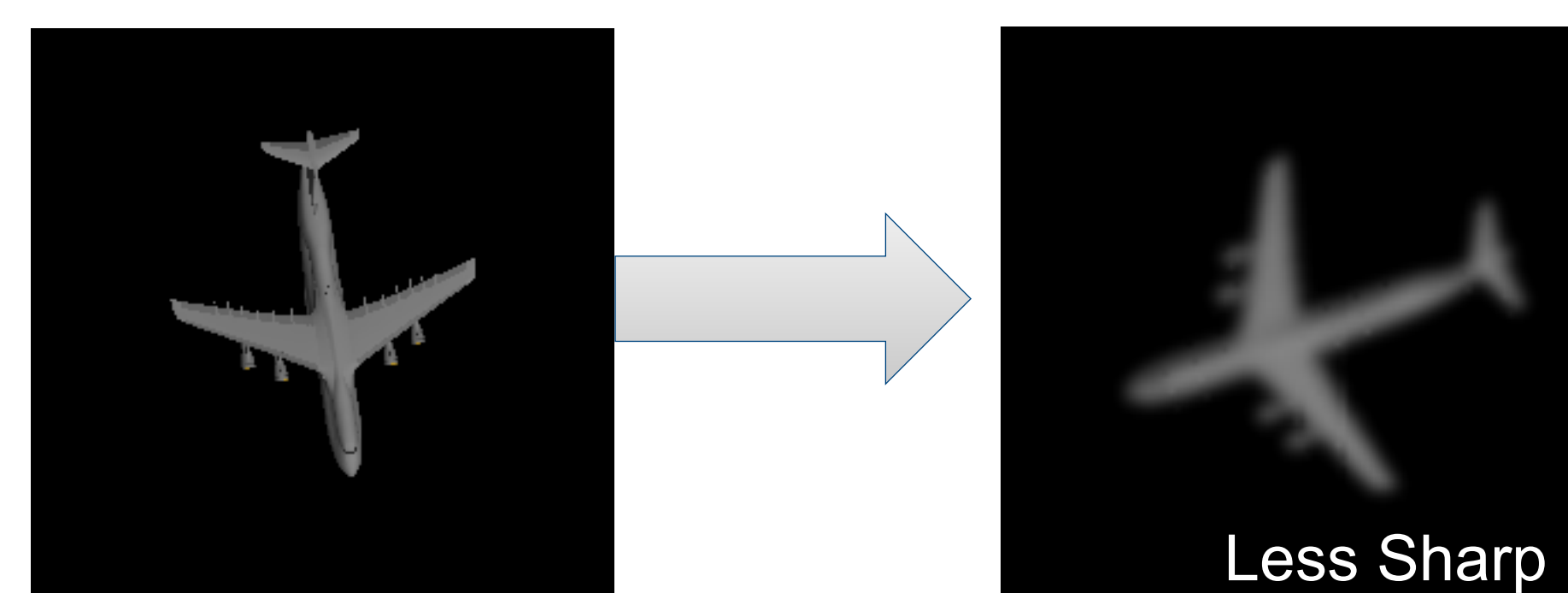Data is further augmented by varying the orientations of the backgrounds

Important Considerations

1. Realistic generation of images
2. Portable between operating systems
3. Convenient setup
4. Consistent implementation

Methodology

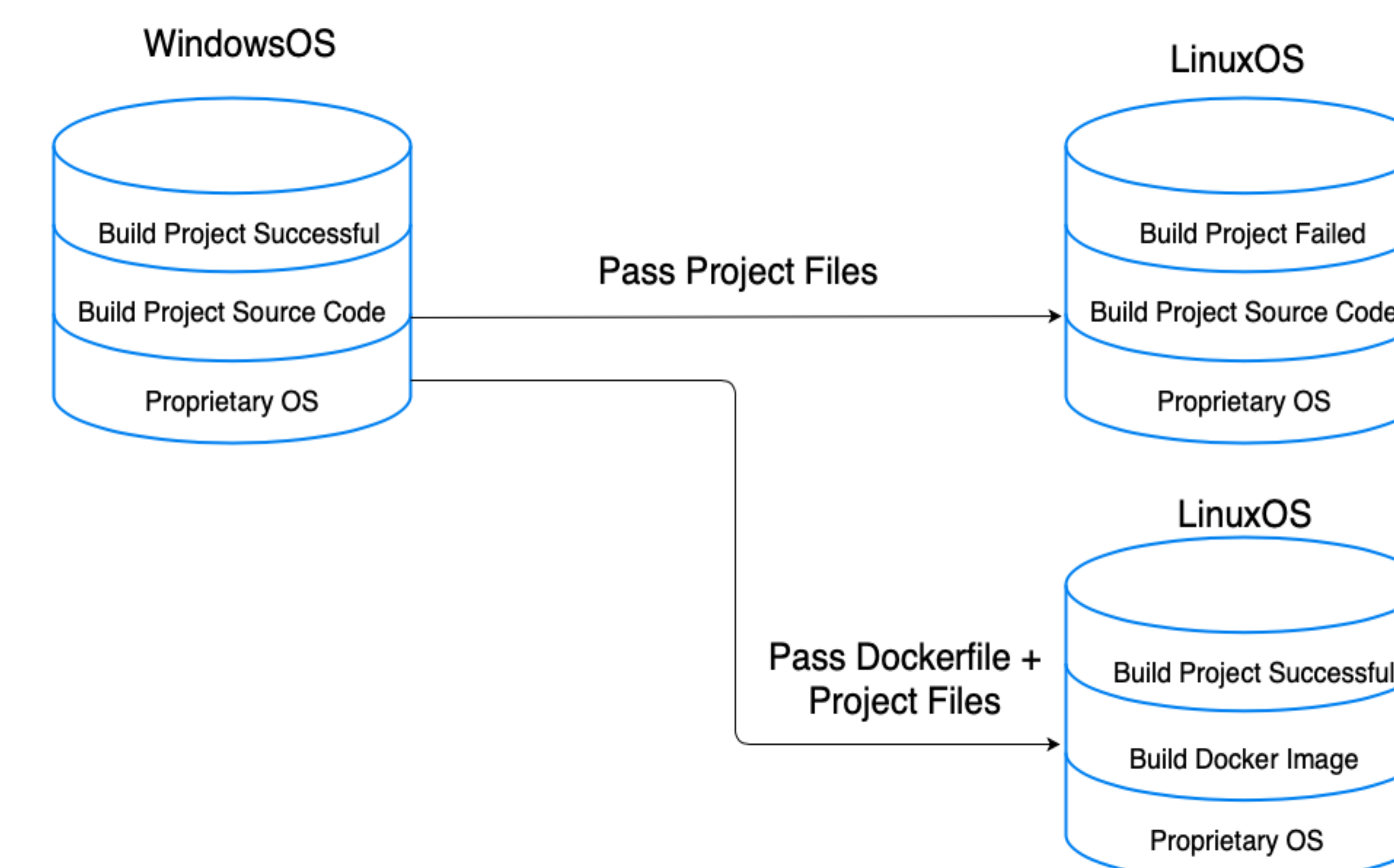- 2D Gaussian sampling kernel used to apply appropriate blurring to rendered CAD model.

Less Sharp

- Use Docker to solve considerations 2-4

Why Docker?

- The specifications of a project are all stored in a Docker file so using docker guarantees everything will function exactly the same on any system that supports Docker

- Easier to keep track of your application when sharing with team members because the environment is constant across all members.

- Dependencies within Docker file won't have any adverse effects on any configurations on your host machine so the project can build properly.

Implementation

WindowsOS

Build Project Successful

Build Project Source Code

Proprietary OS

Pass Project Files

LinuxOS

Build Project Failed

Build Project Source Code

Proprietary OS

Pass Dockerfile + Project Files

LinuxOS

Build Project Successful

Build Docker Image

Proprietary OS

Conclusion

- Our goal is to provide training data that will allow a deep neural network (DNN) to differentiate between background images without the object, and background images with the object

- Synthetic training data enables machine learning when few real examples exist.

- Docker turns the generation of synthetic training data for machine learning into a service

Future Work

- Diagnose features being used by applying deconvolutional networks

- Improve training process by providing examples of additional, unrelated, rendered objects to DNN, forcing it to learn a larger, more discriminative set of object features