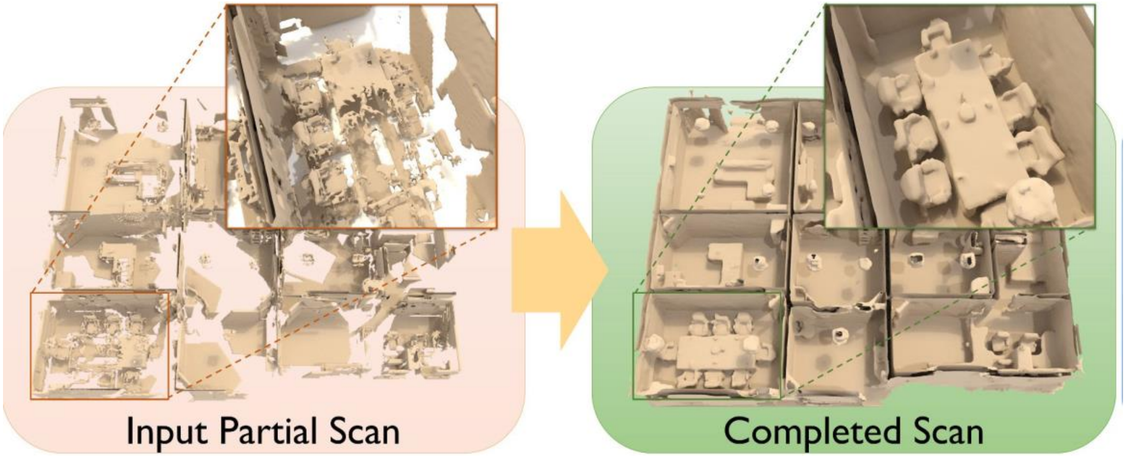
Scalable uncertainty quantification for scene completion

Visio Impulse Ltd & National Physical Laboratory

M. Dziemian, J. Venton, A. Thompson, F. Bazyari, A. Forbes

Background

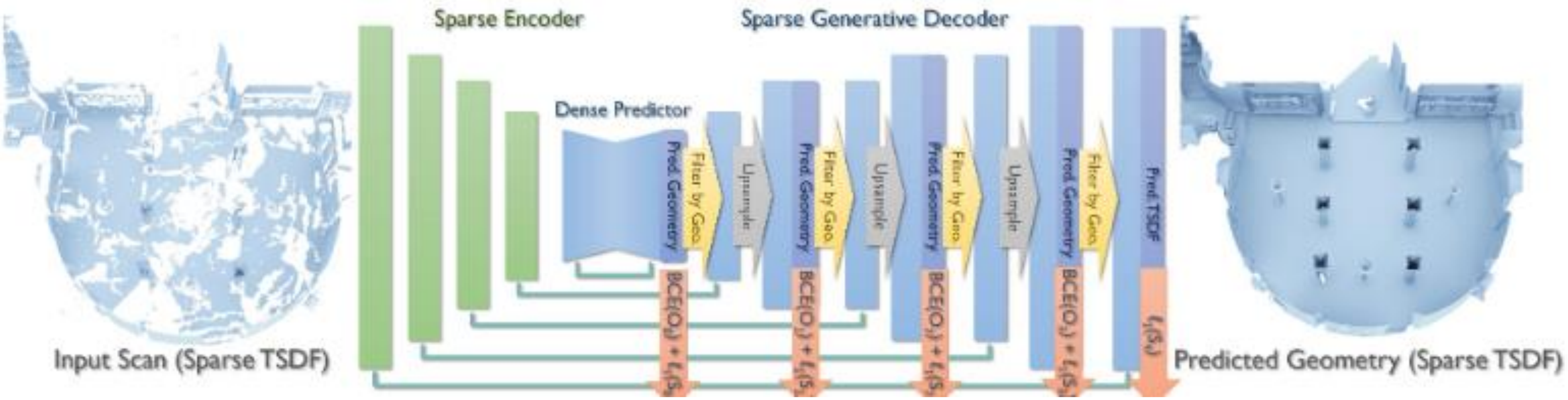
- Generating complete 3D scans can be expensive and time consuming
- Machine learning models can use partial scans to generate a completed version
- Issue: These generated scans can have mispredictions and unknown uncertainties which in high stakes situations aren't acceptable
- Issue: Training & fine-tuning machine learning models on 3D data requires a lot of computing resources (200x200x200 = 8,000,000 data points, 512x512x3 ≈ 800,000 data points)
- Solution: Visualise the uncertainty in the completed scans
- Solution: Reduce the training data amount by finding redundant examples



ScanComplete: Large-Scale Scene Completion and Semantic Segmentation for 3D Scans, Dai et al

3D scan completion models

- Main idea: Input is encoded into a much smaller embedding space. This embedding is then used to decode the embedding and produce a complete scan
- Model used: Sparse Generative Neural Network (SG-NN)
- Input & output data type: Truncated space distance field (TSDF)
- A TSDF can be converted to a mesh for visualisation (marching cubes algorithm)



SG-NN: Sparse Generative Neural Networks for Self-Supervised Scene Completion of RGB-D Scans (Dai, Diller and Niessner, 2020)

Truncated Signed Distance Field

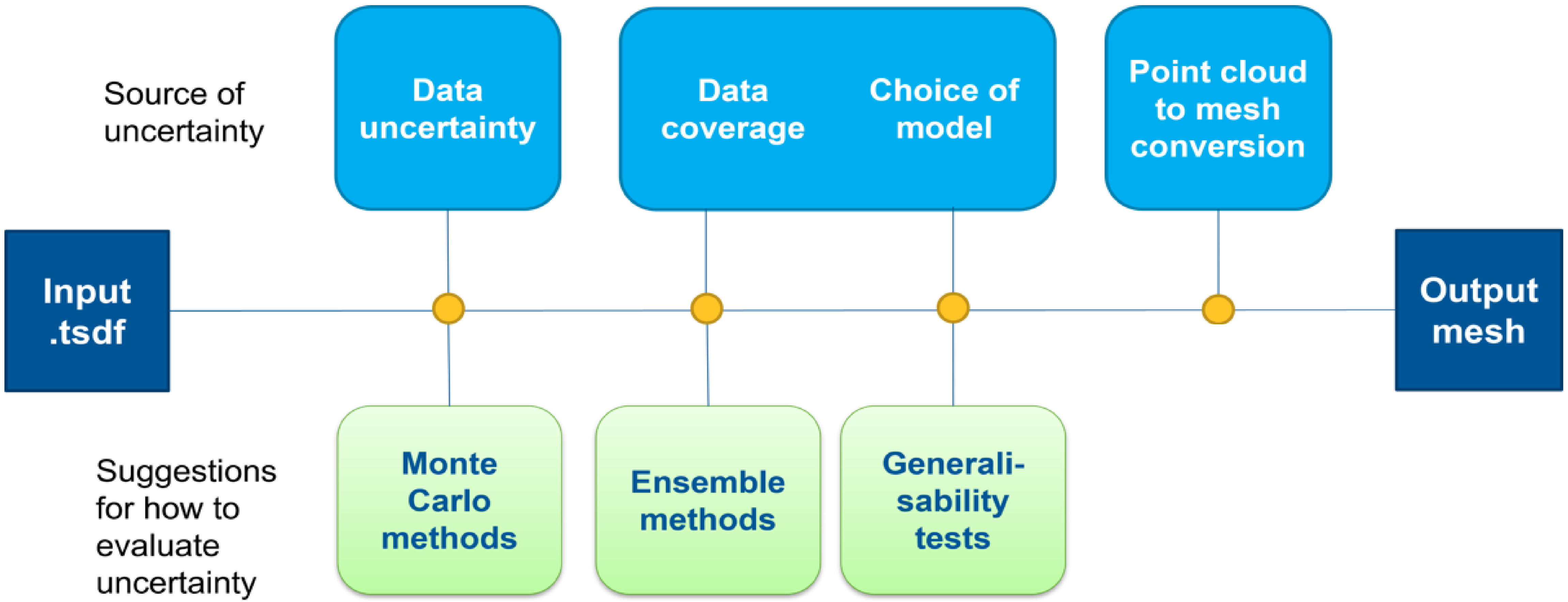
-0.9	-0.3	0.0	0.2	1	1	1	1	1
-1	-0.9	-0.2	0.0	0.2	1	1	1	1
-1	-0.9	-0.3	0.0	0.1	0.9	1	1	1
-1	-0.8	-0.3	0.0	0.2	0.8	1	1	1
-1	-0.9	-0.4	-0.1	0.1	0.8	0.9	1	1
-1	-0.7	-0.3	0.0	0.3	0.6	1	1	1
-1	-0.7	-0.4	0.0	0.2	0.7	0.8	1	1
-0.9	-0.7	-0.2	0.0	0.2	0.8	0.9	1	1
-0.1	0.0	0.0	0.1	0.3	1	1	1	1
0.5	0.3	0.2	0.4	0.8	1	1	1	1

https://itnext.io/understanding-real-time-3d-reconstruction-and-kinectfusion-33d61d1cd402

Uncertainty in machine learning

- Data:
  - 3D laser scanners are only able to locate the boundaries of objects up to some uncertainty
- Model:
  - The extent to which training data is representative of all possible inputs will impact model uncertainty
  - There can be uncertainty about:
    - The choice of model architecture
    - The parameters of the model

Uncertainty in machine learning



Our method to show uncertainty

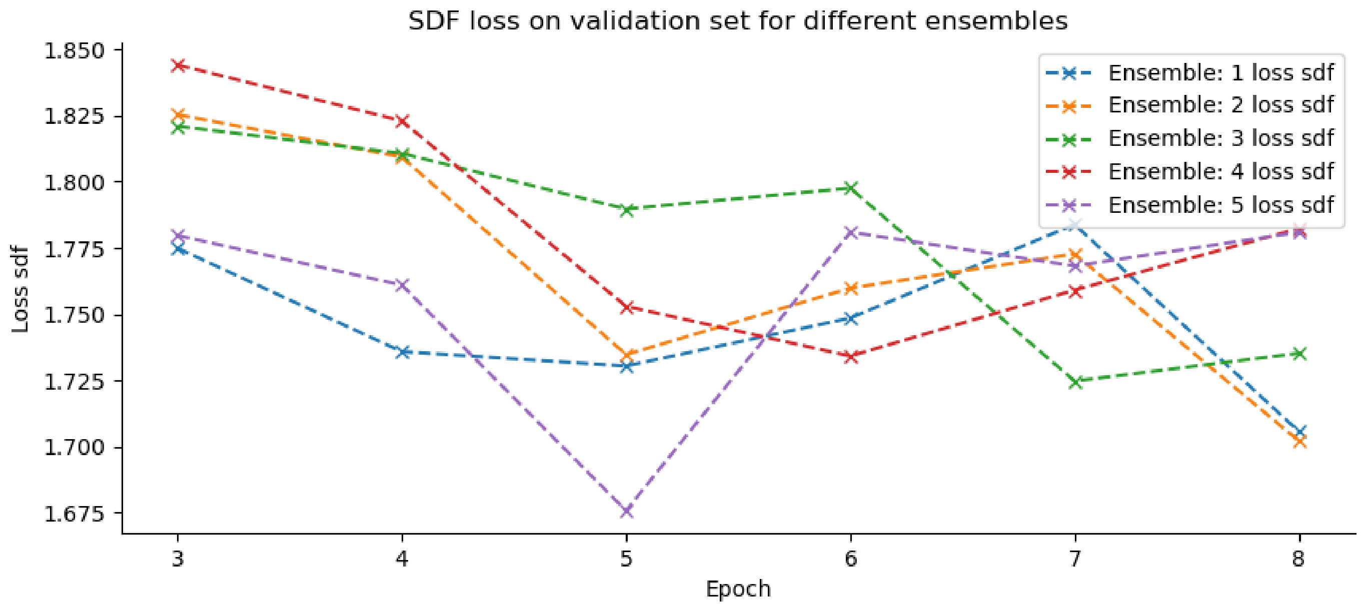
Visio Impulse

NPL National Physical Laboratory

- Aim: visualise the uncertainty for each prediction/ inference
- Chosen method: Ensembling
- Possible options:
    - **Use the same model with the same parameters but change the training data**
    - Change the model parameters such as the model size and keep the training data the same
    - Use different models/ architectures and keep the training data the same

Ensembles training results

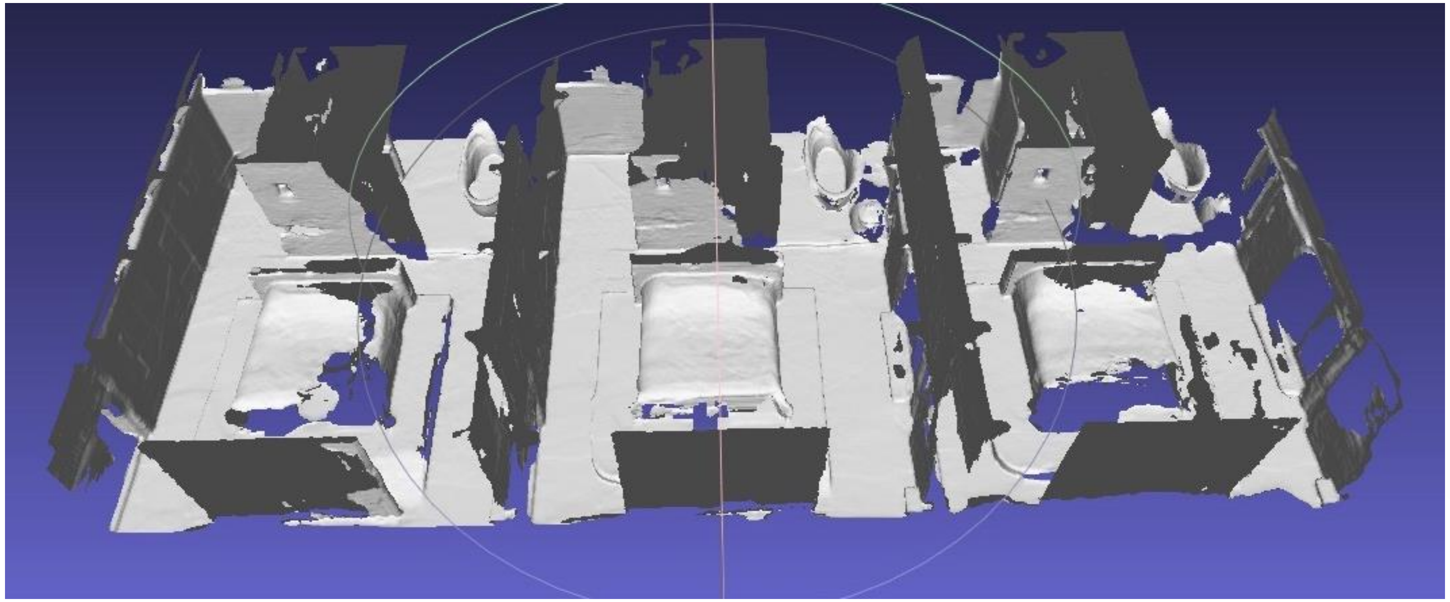
- To train 5 different models, we randomly split the training data into 5 folds (18k examples each) and kept the same validation set for each model.
- The model is trained progressively, meaning that higher resolution outputs are introduced iteratively after N iterations.
- This shows the loss for the highest resolution level on the same validation set.

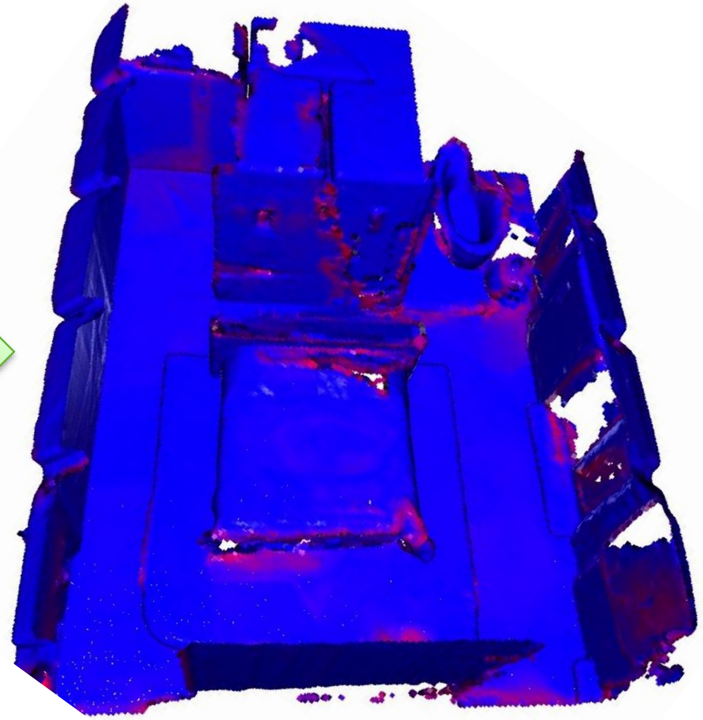


Representation of uncertainty and visualisation

- Apply each of the trained models on the partial scan, calculate mean and standard deviation values for the predicted TSDFs.
- Run marching cubes (TSDF to mesh) on the mean TSDF and convert this into a point cloud.
- To represent the uncertainty for each point cloud point, use regular grid interpolation from the standard deviation values. These interpolated values can then be displayed visually using a colour map representing the uncertainty.

Scene completion uncertainty: Representations



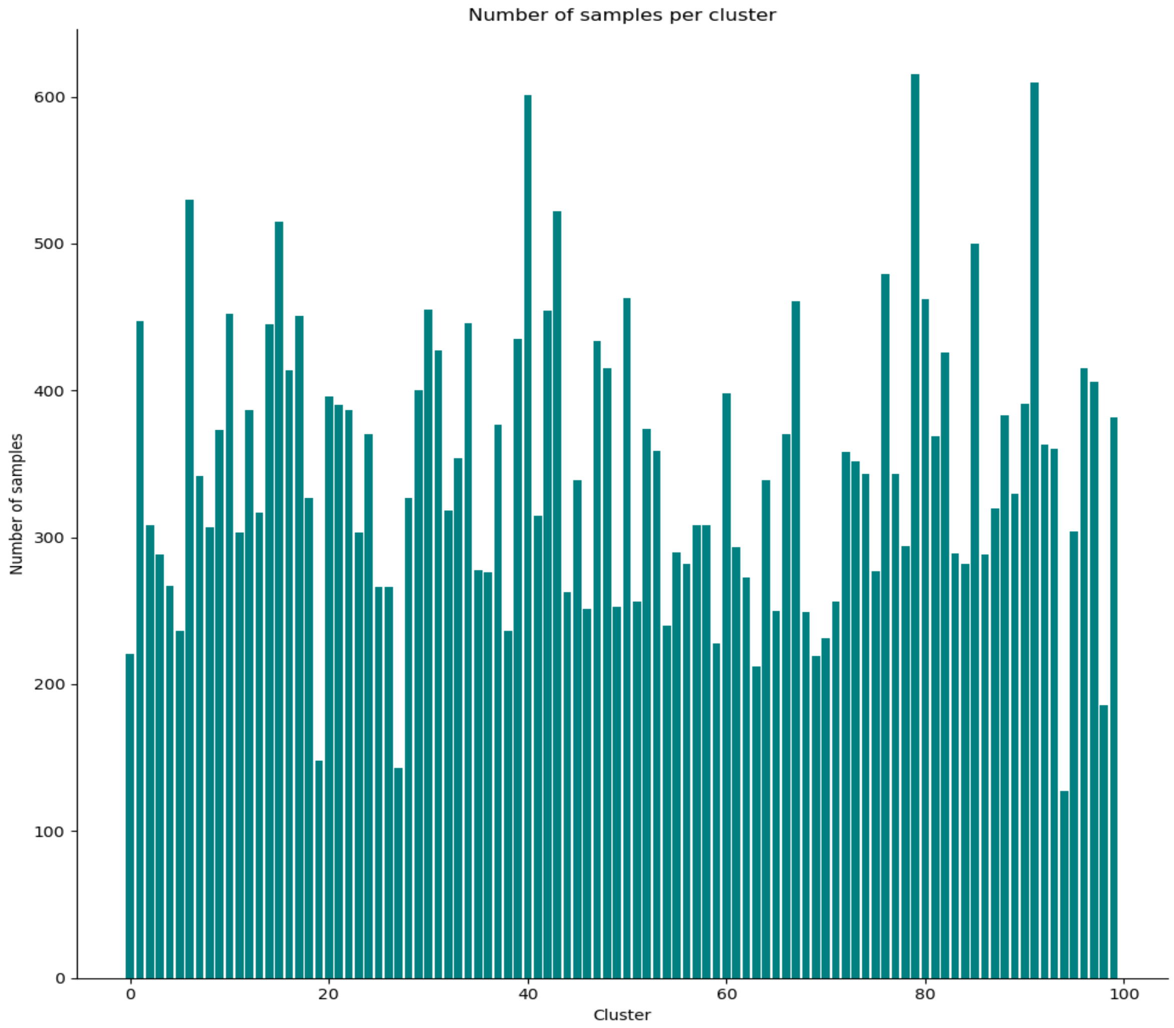


Challenge of scalability

- As mentioned previously, a 200x200x200 scan has 8,000,000 data points which is around 10x that of a 512x512x3 image.
- A lot of sections and objects in scans occur in many of the scans
- Hence, there might be a way to filter out the repeating/ redundant data
- Hypothesis: By sampling data which represents a wide variety of the dataset we can reduce the training size without losing too much accuracy.

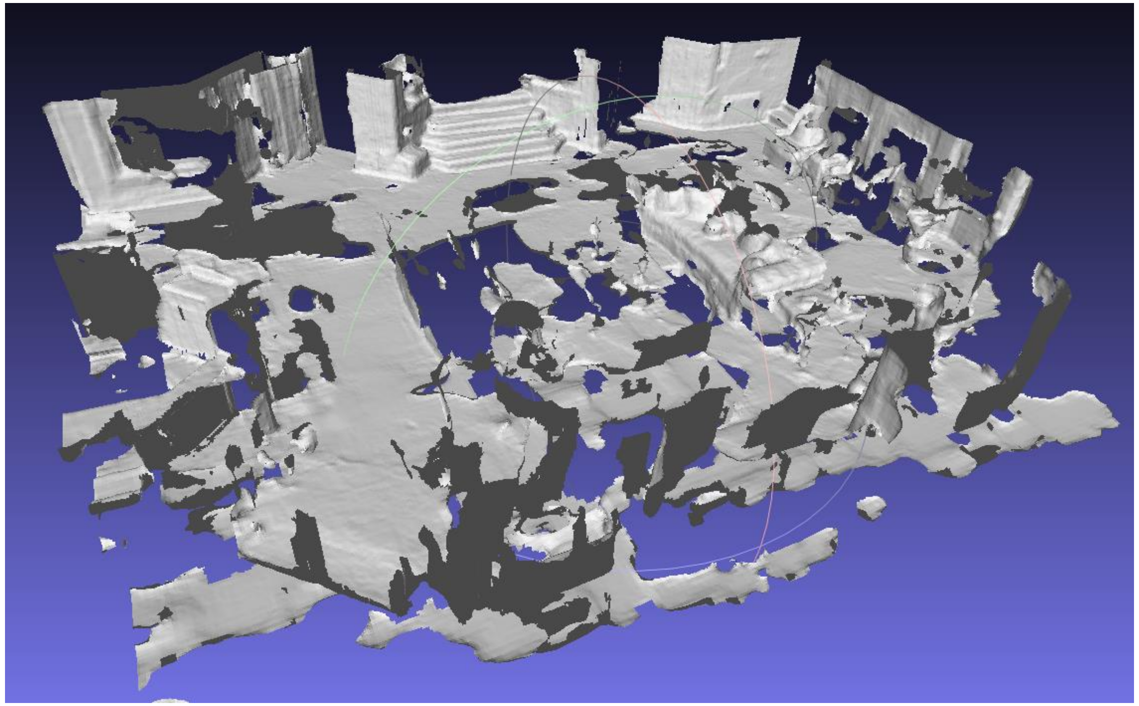
Our data distillation approach

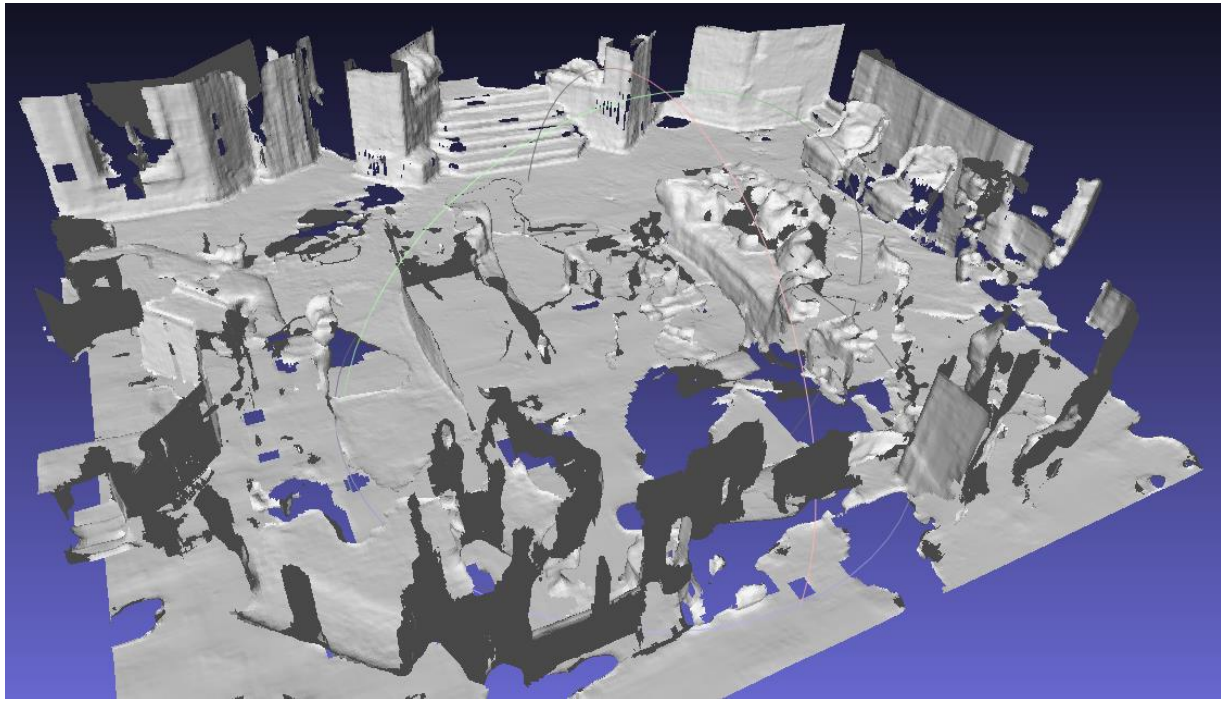
- Apply a pre-trained model to on 38k data points using a pre-trained model
- Use the middle embedding layer to extract embeddings for each example
- Run the K means clustering algorithm on the embeddings
- Randomly sample an equal number of examples from each cluster
- We set the number of clusters to 100 and chose 10 examples from each sample
- To compare, we also randomly sampled 1000 examples and trained on those.



Dataset distillation qualitative results

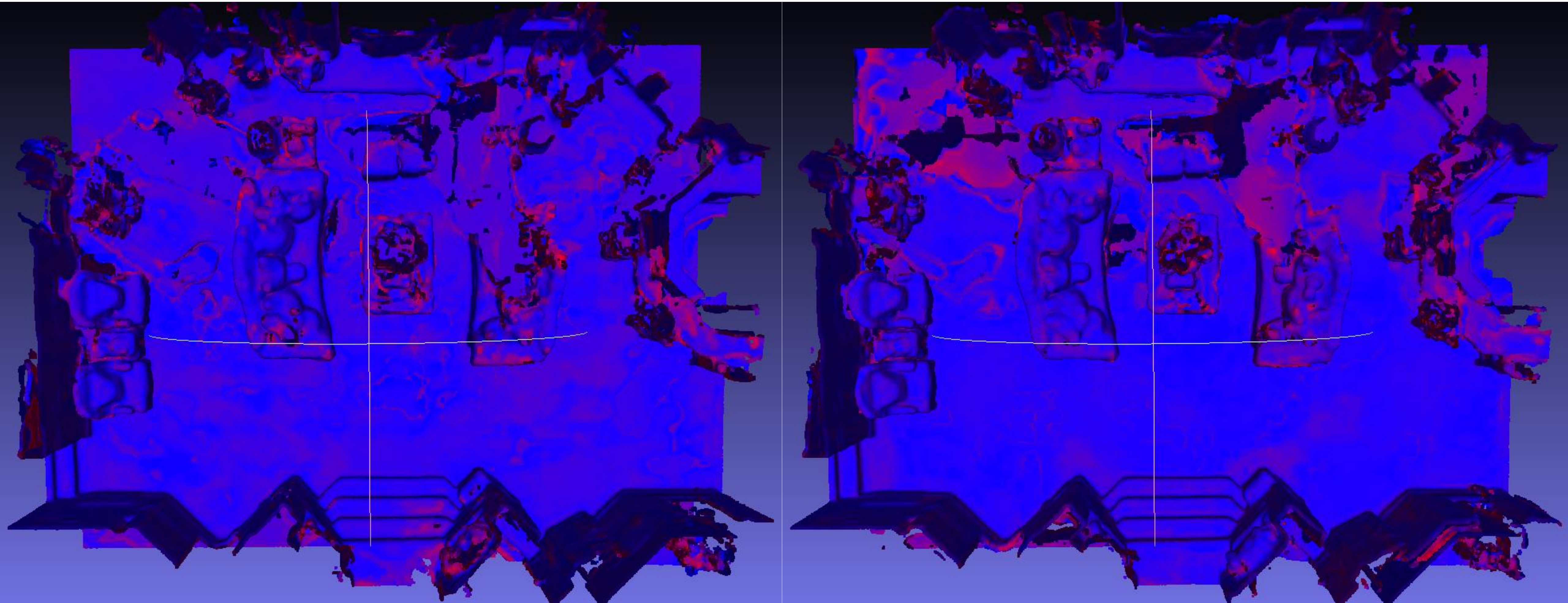
A model trained on 1000 random examples (left), a model trained on 1000 examples using our dataset distillation approach (right).





Dataset distillation qualitative results

Uncertainty for ensembles using random clustering (left) vs Uncertainty for ensembles using K means on embeddings (right)



Conclusion

- Showed how to use ensembles to measure and visualise uncertainty in 3D scan completion

- Showed how a pre-trained model can be used to sample data efficiently using Kmeans and embeddings

Questions?