D1.1 Progress Submission

Practical Machine Learning and Deep Learning

Blueprint Reconstruction and Feature Discovery

Interim Report

Team JAXAXAX

Nikita Zagainov (n.zagainov@innopolis.university)
Nikita Tsukanov (n.tsukanov@innopolis.university)
Said Kadirov (s.kadirov@innopolis.university)
Tetkin Dmitry (d.tetkin@innopolis.university)

September 17, 2025

Abstract

We aim to reconstruct blueprint-style sketches from 3D geometry and extract high-level blueprint features. Our pipeline consists of (1) a diffusion model that maps line-art / sketch renderings to surface normal maps and (2) a feature-understanding model that detects and classifies blueprint features to infer applicable design principles. We prepared a synthetic paired dataset (C++-generated sketches and aligned normal maps) and are commencing training of the diffusion component.[1]

1 Project Topic
(Short Description, Value, Target Users)

Final goal. Reconstruct a clean, blueprint-like sketch from 3D models and derive as much semantic information as possible (e.g., edges, holes, fillets/chamfers, symmetries), producing both an enhanced 2D representation and a surface normal map.

Why it matters. Blueprint extraction accelerates reverse engineering, technical documentation, and quality control workflows. It reduces manual drafting effort and helps downstream CAD/CAM and inspection steps by providing structured, machine-readable features.

Target users. CAD engineers, mechanical designers, architects, technical illustrators, and education/research teams dealing with legacy parts or incomplete drawings.

2 Repositories and Resources

- **Main repository (public):** github.com/touch-topnotch/sketch-vision
- **Research paper (SOTA overview):** research.pdf
- **C++ converters for dataset generation:** github.com/touch-topnotch/sketch-tool

3 Current Status

- Synthetic paired dataset created: automatically generated C++ sketches and aligned normal maps for the same models/poses.
- Data preprocessing scripts in progress; dataset organization for training/validation/test splits.

[1] Project repositories: sketch-vision and sketch-tool.

- Next step: start training the diffusion model for sketch $\rightarrow$ normal map translation.

4 Dataset

We use an in-house **synthetic dataset** built with `sketch-tool`, containing pairs of: (a) C++-generated sketches produced via algorithmic line extraction and (b) corresponding normal maps rendered from the same geometry and viewpoint. The dataset will be split into train/val/test with standardized resolution and augmentation (random noise, line-width variation, small geometric jitter) to improve robustness.

5 Method

5.1 Model A: Sketch $\rightarrow$ Normal Map (Diffusion)

We train a conditional diffusion model to produce per-pixel surface normals from input sketches. The model learns to denoise towards the ground-truth normal distribution while conditioned on the sketch; we explore latent diffusion backbones and control-conditioned variants to better follow line geometry.

5.2 Model B: Feature Understanding

A second model detects blueprint features (e.g., edges, holes, fillets/chamfers, symmetries) and infers design principles they adhere to. We will evaluate both CNN/Transformer backbones for detection/classification and geometric post-processing (e.g., line grouping, Hough-based primitives) for consistent vector outputs.

6 Research of Competitors / SOTA (with references)

For sketch-to-geometry and image-to-image translation, diffusion and adversarial methods set strong baselines, often with structural control:

- **Latent Diffusion / Stable Diffusion** for efficient high-resolution generation [1].
- **ControlNet** for conditioning generation on control signals like edges/poses [2].
- **Palette (Diffusion)** for general image-to-image translation [3].
- **Pix2Pix** for paired translation tasks [4].
- Classical feature detectors and line extractors for blueprints: **Canny** [5] and **LSD (Line Segment Detector)** [6] as strong geometric baselines; **SIFT** [7] for local invariant features.

A consolidated list of our reviewed approaches and notes is maintained in our research paper: `research.pdf` (*link to our SOTA solutions*).

7 Success Criteria and Metrics

Primary criteria:

- **Feature understanding**: mean F1-score over target feature classes (holes, fillets/chamfers, edges, symmetries).
- **Normal-map accuracy**: percentage of pixels with angular error $< 15°$; also report MAE and SSIM.

Targets for milestone 1:

- Feature understanding F1-score $\geq$ **0.85** on the test split.
- Normal-map accuracy (% pixels $< 15°$) $\geq$ **90%**; SSIM $\geq$ **0.90**.

8 Work Distribution

- **Nikita Tsukanov**: Research of SOTA/competitors, methodology design, model selection.
- **Nikita Zagainov**: He got in the way and messed up commits.
- **Tetkin Dmitry**: Implementation of C++ OBJ→sketch algorithms and dataset generation tools.
- **Said Kadirov**: Data preprocessing, dataset curation, pipeline glue code.

9 Risks and Mitigations (Brief)

- *Domain gap between synthetic and real blueprints*: use augmentations and, if possible, small real fine-tuning set.
- *Ambiguous lines in sketches*: introduce control signals (e.g., edge maps, depth hints) and post-processing constraints.
- *Generalization to unseen geometries*: diversify parametric CAD primitives and compositions during synthetic generation.

10 How to Reproduce (Short)

1. Clone `sketch-tool` and generate paired (sketch, normal) data as per README.
2. Clone `sketch-vision` and follow training instructions (diffusion for sketch→normal, feature model thereafter).
3. Evaluate with our metric scripts to obtain F1, angular error, and SSIM on the test split.

Acknowledgments

We thank Innopolis University for computing resources and guidance.

References

[1] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer. "High-Resolution Image Synthesis with Latent Diffusion Models." In *CVPR*, 2022.

[2] Lvmin Zhang, Maneesh Agrawala. "Adding Conditional Control to Text-to-Image Diffusion Models." 2023. (ControlNet preprint).

[3] Chitwan Saharia et al. "Palette: Image-to-Image Diffusion Models." In *SIGGRAPH*, 2022.

[4] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. "Image-to-Image Translation with Conditional Adversarial Networks." In *CVPR*, 2017.

[5] John Canny. "A Computational Approach to Edge Detection." *IEEE TPAMI*, 1986.

[6] Rafael G. von Gioi, Jérémie Jakubowicz, Jean-Michel Morel, Gregory Randall. "LSD: A Fast Line Segment Detector with a False Detection Control." *IEEE TPAMI*, 2010. (Initial tech report, 2008).

[7] David G. Lowe. "Distinctive Image Features from Scale-Invariant Keypoints." *IJCV*, 2004.