Thread Depth research results

Kolobok team

June 2025

1 Introduction

Thread depth is a key quality parameter for fasteners. This sprint we upgraded our vision-based measurement system to achieve sub-millimetre accuracy under varied lighting, surface-finish, and occlusion conditions while still running in real-time.

2 Previous work

Last sprint we built a baseline thread-depth pipeline.

2.1 Neural Depth Estimation models

Neural depth networks such as MiDaS failed to resolve the fine geometry of threads (Fig. 1).

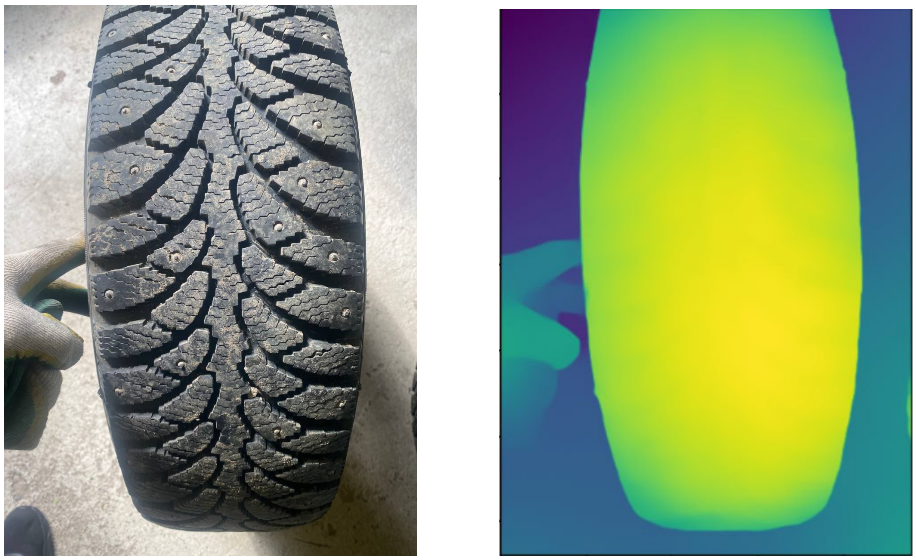

Figure 1: Example of input (left) and corresponding OCR output (right)

2.2 Regression models

A lightweight regression on handcrafted thread features proved faster and more interpretable than deep networks. We enhanced feature extraction, tuned a precision-oriented loss, and validated across thread types; this remains our production baseline.

3 Methodology & Evaluation

Experimental setup:

- Model: GoogLeNet (constant to avoid variance across multiple different models)
- Metrics: MAE, 0.9 quantile of error distribution, Fraction predictions within 1 mm from GT

To enhance the model we explored:

- **Data augmentation** (Random affine, Random noise, Random brightness):
- **Edge detection with Canny**:
- **Edge detection with Sobel**:
- **Contrast CLAHE**:
- **Regularization (Weight decay, gradient clipping)**

4 Results

We evaluated each technique cumulatively: every new experiment started from the best pipeline obtained so far, so reported gains already include all previously accepted novelties.

The baseline model (GoogLeNet without enhancements) achieved the following performance:

- MAE: 0.85 mm
- 0.9 quantile of error: 1.91 mm
- Fraction of predictions within 1 mm from GT: 68.3%

After applying the proposed enhancements, we observed the following improvements:

4.1 Data Augmentation

Applying random affine transformations, Gaussian noise ($\sigma = 0.05$), and brightness adjustments ($\pm 20\%$) increased the model's robustness:

- MAE: 0.81 mm ($\downarrow$ 0.04 mm)
- 0.9 quantile of error: 1.84 mm ($\downarrow$ 0.07 mm)
- Fraction within 1 mm: 70.5% ($\uparrow$ 2.2%)

4.2 Edge Detection Preprocessing

We extracted edge maps and concatenated them to the three RGB channels (giving the network a 4-channel input). Canny and Sobel were both benchmarked, but only Sobel was kept in the final pipeline:

Method	MAE (mm)	0.9 Quantile (mm)	Within 1 mm (%)
Canny ($\sigma = 1.0$)	0.81	1.79	71.04
Sobel (3×3 kernel)	0.81	1.77	71.11

Table 1: Edge detection comparison (**Sobel chosen**)

4.3 Contrast Enhancement (CLAHE) – *rejected*

CLAHE histograms (clip limit = 2.0, tile grid = 8×8) were concatenated as an additional channel, but this hurt accuracy so the technique was dropped:

- MAE: 0.82 mm ($\uparrow$ 0.01 mm)
- 0.9 quantile: 1.87 mm ($\uparrow$ 0.03 mm)
- Fraction within 1 mm: 69.88% ($\downarrow$ 0.17%)

As the metrics worsened, CLAHE was not included in later stages.

4.4 Regularization Techniques

Combining weight decay ($\lambda = 0.01$) and gradient clipping ($max\_norm = 1.0$) provided additional improvements:

- MAE: 0.78 mm ($\downarrow$ 0.07 mm)
- 0.9 quantile: 1.68 mm ($\downarrow$ 0.11 mm)
- Fraction within 1 mm: 72.05% ($\uparrow$ 0.94%)

4.5 Final results

The final pipeline therefore includes Data augmentation + Sobel edge channel + Regularisation. Its performance is:

- MAE: 0.78 mm ($\downarrow$ 0.07 mm)
- 0.9 quantile: 1.68 mm ($\downarrow$ 0.11 mm)
- Fraction within 1 mm: 72.05% ($\uparrow$ 0.94%)

5 Conclusion

Our study shows:

- **Used**: Data augmentation – robustifies the model.
- **Used**: Sobel edge channel – emphasises thread boundaries.
- **Not used**: CLAHE – discarded owing to accuracy drop.
- **Used**: Regularisation – mitigates over-fitting.

All reported improvements are relative to the best previously accepted configuration; cumulatively they deliver a **MAE of 0.78 mm** and 72 % within 1 mm.

Future work could explore:

- Other network architectures
- Attention mechanisms to focus on thread regions