

# Synthetic Dataset Study: Influence of Generated Images, Augmentation and CLAHE

Kolobok team

June 2025

## 1 Introduction

The increasing demand for precise tread-depth estimation in both automotive and industrial settings has driven research towards leveraging advanced data augmentation techniques. Accurate measurements are critical for safety and maintenance applications, yet the collection of extensive, high-quality real-world data remains resource-intensive and time-consuming. Consequently, synthetic datasets generated via domain randomisation have emerged as a compelling alternative, offering flexible control over environmental factors and enabling large-scale data production without prohibitive cost.

Recent advancements in computer vision have demonstrated that combining synthetic imagery with conventional augmentation methods can enhance model robustness and generalisation. However, the relative contributions of synthetic images, standard augmentations, and contrast enhancement techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) have not been thoroughly compared in a unified framework. To address this gap, we propose a comprehensive experimental study that isolates the effects of each augmentation strategy.

In this study, we systematically evaluate three orthogonal strategies to enrich the tread-depth training data:

1. **Synthetic images:** generated with our domain-randomisation pipeline to simulate diverse lighting, texture, and scene configurations.
2. **Standard augmentations:** including random affine transformations, noise injection, and brightness adjustments to emulate natural variability.
3. **CLAHE:** contrast pre-processing applied to normalize lighting conditions and emphasize local texture details.

By enabling each strategy independently, we obtain  $2^3 = 8$  distinct experimental conditions. The performance of each condition is quantified as the fraction of predictions within 1 mm of the ground truth, as recorded in `results.csv`.

## 2 Results

The following sections present detailed outcomes from our experiments, including per-backbone analysis, aggregated synthetic data effects, and an ablation study on the production model.

### 2.1 Raw results per backbone

Table 1 summarizes the fraction of predictions within 1 mm of the ground truth for three reference backbone architectures under all eight augmentation configurations. Metrics are reported for Swin\_v2\_s, GoogLeNet, and DenseNet201 to illustrate model-specific sensitivity to each augmentation.

Table 1: Raw experiment results for reference backbones — fraction of predictions within 1 mm.

Model	Score	Aug	CLAHE	Synthetic
swin_v2_s	0.7816	yes	yes	yes
swin_v2_s	0.7793	yes	no	yes
swin_v2_s	0.7632	no	yes	yes
swin_v2_s	0.8023	no	no	yes
swin_v2_s	0.7977	no	no	yes
swin_v2_s	0.7839	yes	yes	no
swin_v2_s	0.7816	yes	no	no
swin_v2_s	0.7816	no	yes	no
swin_v2_s	0.7724	no	no	no
googlenet	0.7494	yes	yes	yes
googlenet	0.7448	yes	no	yes
googlenet	0.7586	no	yes	yes
googlenet	0.7494	no	no	yes
googlenet	0.7425	yes	yes	no
googlenet	0.7402	yes	no	no
googlenet	0.7540	no	yes	no
googlenet	0.7494	no	no	no
densenet201	0.7471	yes	yes	yes
densenet201	0.7724	yes	no	yes
densenet201	0.7747	no	yes	yes
densenet201	0.7747	no	no	yes
densenet201	0.7816	yes	yes	no
densenet201	0.7655	yes	no	no
densenet201	0.7563	no	yes	no
densenet201	0.7678	no	no	no

Analysis of these raw results reveals that model performance varies substantially across both architecture and augmentation choices, with some chains of augmentations yielding notable improvements over the baseline.

### 2.2 Effect of synthetic data (mean over backbones)

Table 2 reports the average score across all reference backbones, comparing conditions with and without synthetic data.

Table 2: Mean score aggregated by presence of synthetic data (reference backbones).

Synthetic	Mean score
no	0.7647
yes	0.7689

The inclusion of synthetic images yields a modest but consistent gain of approximately 0.0042 in the mean fraction of accurate predictions.

### 2.3 Production model: EfficientNet-B7

Table 3 presents an ablation study on the production-grade EfficientNet-B7 backbone, highlighting the benefits of synthetic data without CLAHE.

Table 3: EfficientNet-B7 ablation results.

Model	Score	Aug	CLAHE	Synthetic
effnet_b7	0.8161	no	no	yes
effnet_b7	0.8000	no	no	no

The best single run achieves a fraction of 0.8161 for predictions within 1 mm of the ground truth when training EfficientNet-B7 with synthetic images and without CLAHE.

## 3 Discussion

Across all experiments, **synthetic images** yielded a reliable and measurable improvement in performance. This gain, although modest in magnitude, was observed consistently for all reference backbones and the production model, underscoring the value of domain-randomised data in supplementing real-world examples.

Standard augmentations such as random affine transformations and brightness modulation contributed additional robustness to minor variations but did not surpass the impact of synthetic data when applied alone. Nevertheless, their additive effect in combination with synthetic imagery suggests that conventional augmentations remain a useful tool for smoothing the decision boundary in low-resource regimes.

Contrary to expectations, **CLAHE** contrast enhancement did not yield further gains when applied alongside other strategies. The interactions between CLAHE and synthetic images were negligible, leading us to omit CLAHE from the final production pipeline to reduce computational overhead without sacrificing accuracy.

## 4 Conclusion

This study demonstrates that integrating domain-randomised synthetic images into the training data is a low-effort yet effective method to enhance tread-depth prediction accuracy. While standard augmentations offer complementary benefits, the omission of CLAHE simplifies the preprocessing pipeline without detrimental effects.

Future work will investigate the impact of additional synthetic modalities, such as 3D-based texture variation, and explore the transferability of these findings to real-world camera systems and other regression tasks.