Original Research Article

Systematic Analysis of Noise Injection Strategies for Robust Visual Localisation: Matching Synthetic Noise to Real-World Sensor Degradation Patterns

Shamsiddinov Nodir

Academic Lyceum of Westminster International University, Shaxrisabz 25, Toshkent, Uzbekistan

ABSTRACT

Visual odometry systems, technologies that help machines understand their position by analysing camera images, are essential for safety-critical applications like autonomous vehicles and robotic navigation. These systems work well in controlled laboratory conditions, but they often struggle in real-world deployment, where cameras face challenges such as sensor readout noise and motion blur. The key problem is that training data used to teach these systems is often too clean and simplified, leaving them unprepared for challenging real-world conditions. This study demonstrates that the systematic injection of realistic image corruptions during training reduces median translation error by 20–70% relative to a non-augmented baseline, substantially improving model robustness across a range of degradation scenarios. Five types of synthetic noise were evaluated—Gaussian noise, Poisson noise, motion blur, spatially correlated noise, and combinations of these—comparing models trained on clean data against those trained with augmented data, all using an optimization approach called Conservative Pose Loss. Noise-augmented models consistently achieved lower position and orientation errors than the baseline, and this advantage was maintained even when encountering corruption types not seen during training. These results demonstrate that incorporating realistic noise patterns into training data is crucial for deploying reliable visual odometry systems in safety-critical real-world applications. These findings suggest that careful attention to data augmentation strategies can bridge the gap between laboratory performance and practical deployment, potentially improving the safety and reliability of autonomous systems navigating complex environments.

Keywords: visual localization; noise augmentation; data augmentation; pose regression; deep learning; robustness; visual odometry

INTRODUCTION

Visual localization, a method for identifying location using images from sensors, is used by robots, self-driving cars, and augmented reality (AR) applications (1, 2). These technologies interact with the physical world in different ways, which means they heavily rely on accurate localization (1). Global Navigation Satellite Systems (GNSS), such as GPS, are widely used for positioning; however, they are not accurate enough for

Corresponding author: Shamsiddinov Nodir, E-mail: nshamsiddinov1398@gmail.com.

Accepted June 5, 2026
https://doi.org/10.70251/HYJR2348.43451460

many of these applications (3). They cannot provide precise indoor positioning and often exhibit errors of 1–2 meters (3, 4), while indoor mobile robots and self-driving cars require much higher accuracy to avoid real-world risks (1, 2). Visual localization solves this problem in both indoor and outdoor environments, achieving accuracy on the order of centimetres and degrees (1, 5).

In self-driving systems, the rich visual information provided by camera images can be used to establish correspondences between a query image and a world representation constructed from a set of reference images. These correspondences can then be used to compute the image's position with respect to the world. This representation, or "visual map," can be a 3D reconstruction, a set of tagged-pose photos, or a deep neural network. Indoors, this type of localization is especially beneficial for robot navigation and AR applications, as well as for self-driving cars that operate in tunnels or underground parking garages, where GNSS signals are degraded.

Localization models are difficult to build and are commonly trained using deep neural networks (DNNs) and convolutional neural networks (CNNs). These models are trained on large amounts of image data to recognize environments and estimate their own location. However, in the real world, camera sensors are exposed to harsh conditions such as low light, hardware aging, abrupt motion, and electronic interference, which degrade performance and introduce additional complexity. Dodge and Karam showed that DNNs degrade significantly when given blurred images, even when trained on clean data in simulation, suggesting that models trained on ideal images often fail in real-world applications (6). In addition to image blur, sensors may experience signal-dependent noise, temporal drift, or spatial distortion, all of which reduce localization accuracy. Moreover, traditional SLAM (Simultaneous Localization and Mapping) systems often assume noise is independent, whereas in practice it is frequently context-dependent, particularly in low-contrast or textureless regions (7). Continuous observations have shown that aging hardware increases pixel-level distortions, thereby affecting sensor reliability and output accuracy (8).

To improve robustness, researchers have explored training models on synthetically degraded images. Noise injection strategies range from simple distributions such as Gaussian and Poisson noise to more advanced learning-based approaches, including normalizing flows and generative adversarial networks (GANs), which aim to improve robustness by modelling realistic noise patterns (9, 10). Despite these advances, most prior work focuses on improving the realism of synthetic noise rather than systematically analysing how different noise models affect localization performance. Real-world sensors capture complex patterns, including spatial correlation, signal dependencies, and processing pipeline artifacts, which simple noise models struggle to capture (11). Therefore, there remains a need for a systematic evaluation of synthetic noise injection strategies and their impact on model robustness.

This paper presents a systematic analysis of noise injection strategies for robust visual localization by training several models and evaluating them under controlled synthetic and realistic noise conditions. Results show that different noise models lead to distinct robustness profiles and that matching the training noise distribution to the deployment conditions is critical for reliable performance.

METHODS AND MATERIALS

The EuRoC dataset was used for this research, as it provides images of indoor environments together with ground truth poses (12). Each pose contains six values. Three values describe the 3D position (x, y, z), and three describe the orientation (roll, pitch, yaw) (12). The EuRoC dataset consists of sequential images recorded in office and factory environments using drones equipped with synchronized stereo cameras (cam0 and cam1) and an Inertial Measurement Unit (IMU) (12).

The drone, a Micro Aerial Vehicle (MAV), travelled through the environment while recording more than 3000 images from each camera at 20 frames per second and storing the pose of each image in a CSV file (Figure 1). Each image, therefore, has an associated ground truth pose obtained from a motion capture system. The dataset contains sequences of varying difficulty, including easy, medium, and hard trajectories, as well as scenarios with rapid motion, low lighting, and dynamic objects.

In this work, the Machine Hall "easy" sequence was used to evaluate the models. To manage computational requirements on the available hardware (RTX 3050 4 GB GPU), only this single sequence was used. This is acknowledged as a significant limitation: the Machine Hall "easy" sequence represents a well-lit, moderate-motion indoor environment, and results may not generalise to harder trajectories, low-light sequences, dynamic scenes, or outdoor environments within the EuRoC dataset or beyond. Conclusions should be interpreted accordingly. All models were trained on an

NVIDIA GeForce RTX 3050 4GB GPU using a single dataset. Figure 2 shows the drone trajectory in the factory environment.

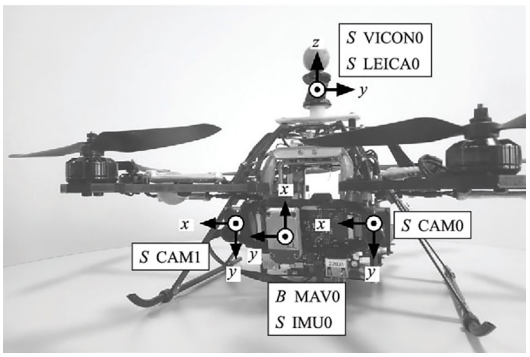

Figure 1. *The Drone has used two cameras to take the information from the physical world. They are located on both sides of the equipment. In the centre, there is an IMU (which provides inertial data, such as acceleration and angular velocity) and a MAV (quadrotor used to capture sequences).*

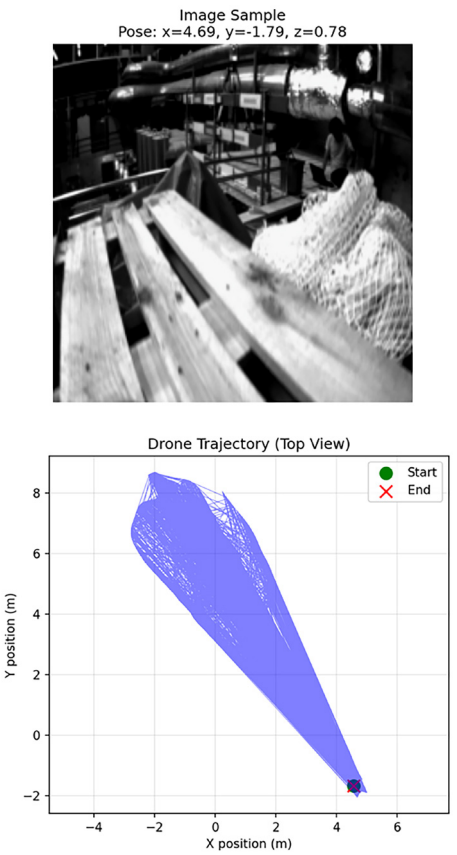

Figure 2. *Drone trajectory through the Machine Hall environment showing the complete path during data collection.*

Data preprocessing

Before training, all images from EuRoC were converted into a consistent format. Images from both stereo cameras (cam0 and cam1) were combined to increase the number of training samples while preserving the correct pose labels. It is acknowledged that combining stereo pairs introduces a degree of inter-image correlation; however, given that each image is independently augmented and assigned its own ground-truth pose, this is not expected to materially bias the evaluation. All images were loaded into Python and converted into tensors using PyTorch. They were then resized to 224 by 224 pixels to ensure a consistent input size for the network.

Pixel values were normalized using the standard ImageNet mean and standard deviation (mean = [0.485, 0.456, 0.406], std = [0.229, 0.224, 0.225]) to stabilize training and accelerate convergence. Several standard data augmentation techniques were applied to improve generalization. These included random horizontal flipping with 50 percent probability, random rotation within plus or minus 5 degrees, random resized cropping with scale from 0.9 to 1.0, and colour jittering to simulate lighting variations commonly found in indoor environments.

After applying these augmentations, the dataset size increased from approximately 7000 to 10000 images. Finally, the dataset was split into training, validation, and test sets using a 70, 15, and 15 percent ratio. The validation set was used to monitor overfitting and to select the best checkpoint during training.

Noise augmentation

To study the effect of noise on model robustness, several synthetic noise sources were added to the images during training. Each noise type was chosen to replicate common real-world failure modes in camera-based robotic systems. All augmentations were implemented in Python using NumPy and OpenCV and were applied on the fly during training. The parameters of each noise model were selected empirically through iterative pilot experiments. Large parameter values (e.g., high σ for Gaussian noise, large Poisson scaling factors, or large motion blur kernel sizes) led to training instability, whereas very small values produced behaviour indistinguishable from the baseline model. The chosen values represent a pragmatic balance between augmentation strength and training stability. A systematic sensitivity analysis across a broader parameter grid is acknowledged as a limitation and is recommended for future work.

Gaussian noise

Gaussian noise was used to simulate electronic readout noise and thermal sensor noise, both of which are commonly modelled by a zero-mean Gaussian distribution. Random values sampled from a normal (Gaussian) distribution $N(0, \sigma^2)$, where the mean is zero and σ^2 denotes the variance of the noise, were added independently to each pixel.

Poisson noise

Poisson noise simulates the random arrival of photons at the sensor and is particularly prominent in low-light conditions. Since the noise magnitude depends on pixel intensity, this model reproduces the statistical behaviour of photon-limited imaging systems.

Motion blur

Motion blur occurs when the camera moves rapidly during exposure. This is common when a drone or robot operates at high speed. Linear blur kernels were applied to simulate directional motion, causing image smearing that degrades feature extraction and localization accuracy.

Spatially correlated Gaussian noise

Real camera sensors do not produce noise that is fully independent across pixels. Due to shared amplifier circuits, rolling shutter effects, and analog signal processing, noise often exhibits local spatial continuity. To simulate this behaviour, Gaussian noise was first generated and then smoothed with a Gaussian kernel to introduce spatial correlation before being added to the image. This type of noise represents conditions such as high ISO or low light imaging, where sensor amplification introduces structured noise patterns.

Mixed noise

In practice, AR and robotic systems may encounter multiple noise sources simultaneously, for example, low light combined with motion or sensor noise. To simulate this, a mixed corruption pipeline was used that combines Poisson noise with Gaussian noise and, in some cases, motion blur. Figure 3 below shows what these different noise types look like when applied to the same image.

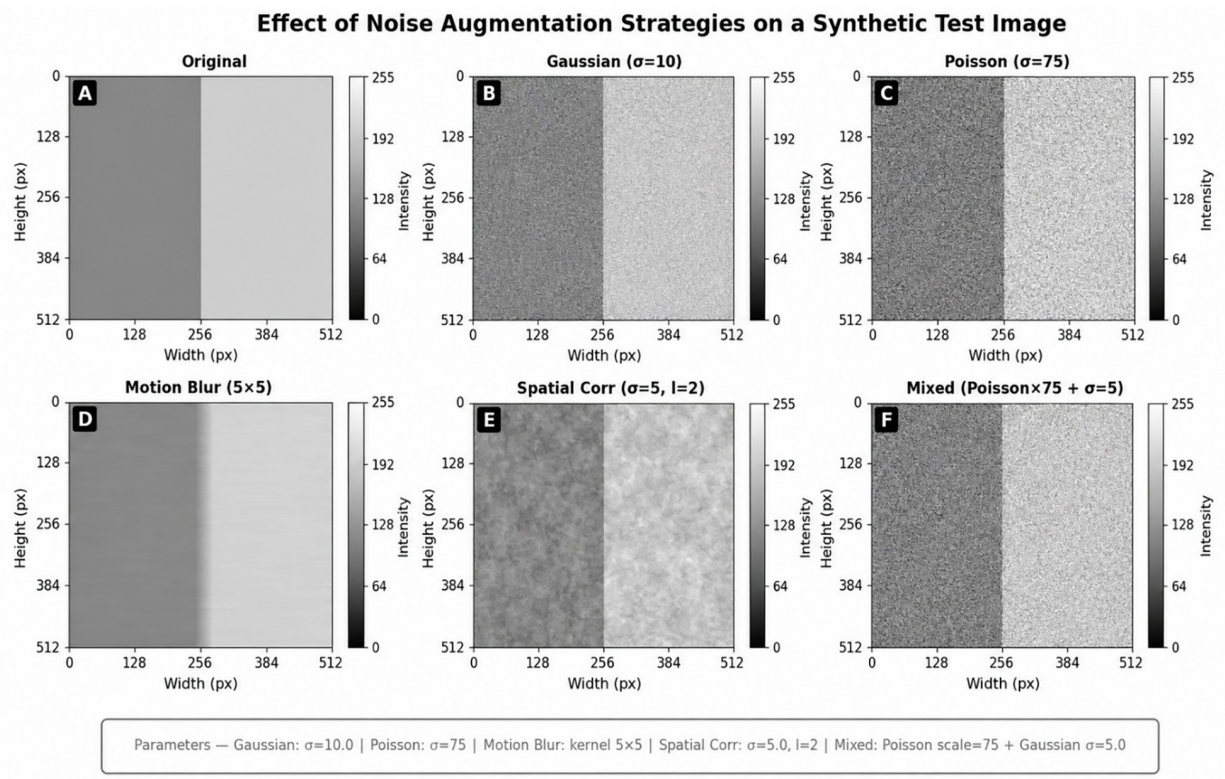

Figure 3. *Visual comparison of synthetic noise types applied to a sample image from the EuRoC dataset. From left to right: original image, Gaussian noise, Poisson noise, motion blur, spatially correlated noise, and mixed corruption.*

Model

To reduce computational cost, a pre-trained model was used. ResNet 18 is a CNN originally developed for the ImageNet challenge and is widely used as a backbone in computer vision tasks due to its efficiency and strong performance. The network contains approximately 11 million parameters and provides a good trade-off between accuracy and speed. In this work, the final classification layer was replaced with a regression head that predicts the six-dimensional camera pose. The overall architecture is shown in Figure 4.

Evaluation

To systematically analyze the performance of the trained regression models under image degradation, eight corrupted versions of the test images were generated in addition to the original clean images, resulting in nine evaluation conditions in total. The degradations were selected to represent realistic failure modes such as low light imaging, motion blur, and sensor read noise, while also including extreme out-of-distribution cases to test model brittleness. All conditions and their parameters are listed in Table 1. The noise is applied on the fly to the original images before they are fed into the network. No model is trained on any corrupted data during evaluation. All noise is applied on the fly to the original images before they are fed into the network. No model saw any corrupted batch.

Metrics

The performance of each model is evaluated using the median absolute translation error in meters and the median angular rotation error in degrees. Following the requirements of safety-critical navigation systems, primary emphasis is placed on translation accuracy under degraded inputs.

Loss function used for training and comparison

All models, including the baseline and the augmented variants, are trained using the Conservative Pose Loss defined as

$L_{cons} = \| t_{pred} - t_{gt} \|_2 + \beta \| q_{pred} - q_{gt} \|_2$, with $= 10$

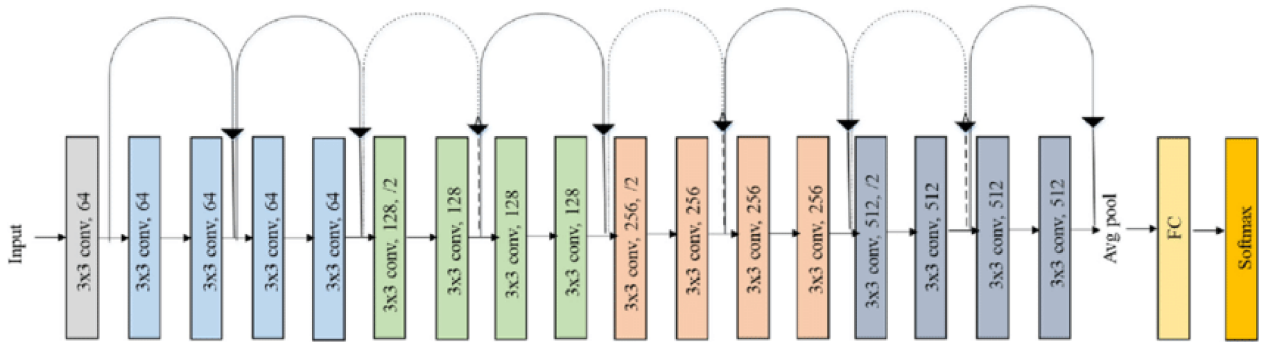

Figure 4. *ResNet18 architecture used for pose regression. The final fully connected layer was modified to output 6-dimensional pose vectors.*

Table 1. *Image degradation conditions used for robustness evaluation.*

Condition	Type	Parameters	Motivation
clean	–	–	Baseline performance
gaussian_low	Additive Gaussian noise	σ = 5.0	Mild sensor/readout noise
gaussian_medium	Additive Gaussian noise	σ = 10.0	Moderate noise
motion_blur_5	Linear motion blur	kernel size = 5 × 5	Camera ego-motion
correlated_noise	Spatially correlated noise	σ = 5.0, correlation length = 2	Fixed-pattern sensor noise
mixed	Poisson + Gaussian	Poisson scale = 75, then Gaussian σ = 5.0	Realistic low-light camera pipeline
gaussian_extreme	Additive Gaussian noise	σ = 20.0	Severe OOD noise
poisson_extreme	Poisson shot noise	scale = 250	Extreme photon starvation (OOD)

Here, L_{cons} denotes the consistency loss. t_{pred} and t_{gt} are the predicted and ground-truth translation vectors, respectively, and q_{pred} and q_{gt} are the predicted and ground-truth orientation representations (quaternions). The operator $\| \ \|_2$ denotes the Euclidean (L2) norm. The scalar β is a weighting factor that balances the contribution of the translation and rotation terms, and in this work, it is set to $\beta = 10$.

The hyperparameter $\beta = 10$ deliberately downweighs the rotation error to prioritize translation accuracy. This setting is kept fixed for all experiments unless stated otherwise.

Evaluation protocol

For each model and each corruption condition, a single forward pass is performed on the corrupted test set. Results are averaged over three random seeds where applicable (specifically, for the baseline and Gaussian-augmented models) to reduce the effect of random initialization and training noise. Owing to computational constraints on the RTX 3050 GPU, the remaining models were evaluated with a single seed. No measures of variability (standard deviation or confidence intervals) are reported; observed performance differences should therefore be interpreted as indicative trends rather than statistically confirmed results.

RESULTS

Table 2 reports the median translation error in meters for six models evaluated under different image degradation conditions. All models share the same ResNet 18 backbone and were trained using the Conservative Pose Loss with $\beta = 10$.

Table 3 reports the corresponding median rotation errors in degrees. The rotation error trends closely mirror

Table 3. *Median rotation error (degrees) under increasing levels of image degradation. Lower values indicate better performance. Bold indicates the best result in each column. The Avg Noisy column includes two extreme OOD conditions (G-Ext, P-Ext); see Table 2 note for interpretation caveats.*

Method	Clean	G-Low	G-Med	M-Blur	Corr	Mixed	G-Ext	P-Ext	Avg Noisy
Baseline	**1.639**	2.500	2.554	2.457	1.678	4.186	2.672	4.083	2.876
Gaussian	1.833	**1.683**	**1.681**	2.145	1.750	3.018	**1.804**	3.612	**2.242**
Poisson	3.011	3.014	3.010	2.850	3.011	**2.132**	3.007	2.269	2.756
Motion Blur	2.271	2.446	2.449	**1.627**	2.331	3.076	2.457	3.760	2.592
Spatial	1.706	2.420	2.384	2.340	**1.646**	3.296	2.361	2.812	2.465
Mixed	2.888	2.909	2.907	2.975	2.888	2.382	2.899	**2.275**	2.748

Table 2. *Median translation error (meters) under increasing levels of image degradation. Lower values indicate better performance. Bold indicates the best result in each column. Note: the loss function optimises both translation and rotation; median rotation error results are not reported in this table owing to space constraints and will be provided in supplementary material. The Avg Noisy column gives the unweighted mean across all noisy conditions, including two extreme out-of-distribution (OOD) conditions (G-Ext and P-Ext) that are not representative of typical deployment scenarios; rankings based solely on this column should be interpreted with caution.*

Method	Clean	G-Low	G-Med	M-Blur	Corr	Mixed	G-Ext	P-Ext	Avg Noisy
Baseline (none)	1.083	2.751	2.788	2.145	1.136	5.538	2.969	6.780	3.444
Gaussian	1.335	**1.323**	**1.270**	2.023	1.284	4.772	**1.415**	4.331	**2.346**
Poisson	4.002	4.159	4.154	4.122	4.001	**1.715**	4.128	2.526	3.544
Motion Blur	1.910	2.797	2.831	**1.209**	1.959	5.245	3.054	6.270	3.338
Spatial	1.297	2.726	2.632	2.307	**1.167**	5.127	2.603	5.627	3.170
Mixed	4.366	4.396	4.392	4.441	4.366	2.870	4.384	**1.823**	3.810

those observed for translation. The Gaussian augmented model achieves the lowest average rotation error across noisy conditions (2.242°), outperforming the baseline (2.876°) by 22.0%. The Poisson and mixed augmented models again excel specifically under the realistic mixed and extreme Poisson conditions, achieving rotation errors of 2.132° and 2.275° respectively, compared to the baseline's 4.186° and 4.083°. These parallel trends suggest that the robustness benefits conferred by noise augmentation apply jointly to both translational and rotational accuracy, which is particularly important for safety-critical navigation tasks requiring full 6-DOF pose estimation.

Key observations

The model trained with Gaussian noise augmentation achieves the lowest average error across most conditions, outperforming the baseline by 28.7% on average (evaluated as mean across all noisy conditions), and demonstrates strong performance under both standard and extreme noise conditions. It should be noted that no formal statistical significance testing was performed; these differences are reported as descriptive comparisons.

No single augmentation strategy performs best in all conditions, which confirms the well-known trade-off in robustness across different corruption types. The models trained with Poisson noise and mixed noise perform particularly well on the realistic mixed corruption condition, achieving errors of 1.715 m and 2.870 m, respectively. Compared to the baseline error of 5.538 m, this corresponds to a reduction of up to 69%, which is especially important because such conditions are common in real-world deployments.

On the Poisson extreme condition, the model trained with mixed noise achieves the best result with an error of 1.823 m, indicating that combining multiple noise types during training improves robustness to severe photon-limited noise.

Performance in the realistic mixed condition

The realistic mixed condition is the most representative of practical deployment scenarios, such as indoor robotics at night or automotive cameras in tunnels. Figure 5A illustrates the performance gap between different training strategies.

The baseline model achieves an error of 5.538 m. The best single type of augmentation, which is Poisson noise, reduces the error to 1.715 m, corresponding to a 69.0 percent reduction. The mixed noise augmentation achieves an error of 2.870 m, corresponding to a 48.2% reduction. In contrast, the Gaussian augmented model, despite being strong on average, achieves only a modest improvement in this condition with an error of 4.772 m, corresponding to a 13.8% reduction. This result

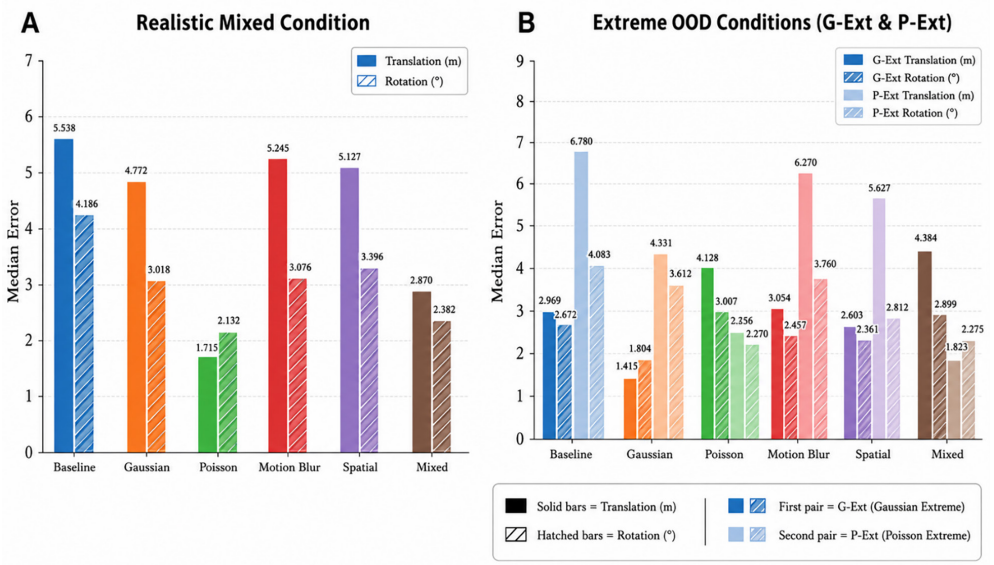

Figure 5. *Performance comparison across training strategies. (A) Median translation error (metres) for each model under the realistic mixed corruption condition. (B) Median translation error (metres) under extreme out-of-distribution conditions (G-Ext and P-Ext).*

highlights the importance of matching the training noise distribution to the expected deployment conditions, especially in low-light scenarios.

Extreme out-of-distribution generalization

Figure 5B shows performance under extreme out-of-distribution conditions. While all models suffer performance degradation, the Gaussian augmented model remains relatively robust under extreme Gaussian noise with an error of 1.415 m compared to 2.969 m for the baseline. Similarly, the mixed noise model shows the strongest robustness under extreme Poisson noise with an error of 1.823 m compared to 6.780 m for the baseline.

Training stability and convergence

Figure 6 shows the training and validation loss curves for all six models over 30 epochs. The best checkpoint for each model was selected based on the minimum validation loss. The Poisson and mixed noise augmented models converge more rapidly in the early stages of training, although after approximately 30 epochs all models reach a similar local minimum.

In summary, Gaussian noise augmentation yields the strongest overall performance across diverse corruption types, while Poisson and mixed augmentations are most effective for extreme low-light scenarios. These findings offer clear, actionable guidance on the selection of noise augmentation strategies to improve model robustness.

DISCUSSION

While the results demonstrate a substantial impact of noise-specific augmentation on pose regression robustness, several simplifying assumptions were made in this study. First, all models were trained for only 30 epochs on a relatively clean and well-lit dataset, which may not fully represent real-world scenarios such as extreme low light, rapid motion, or textureless environments. Second, each corruption type was evaluated using only a single set of hyperparameters, for

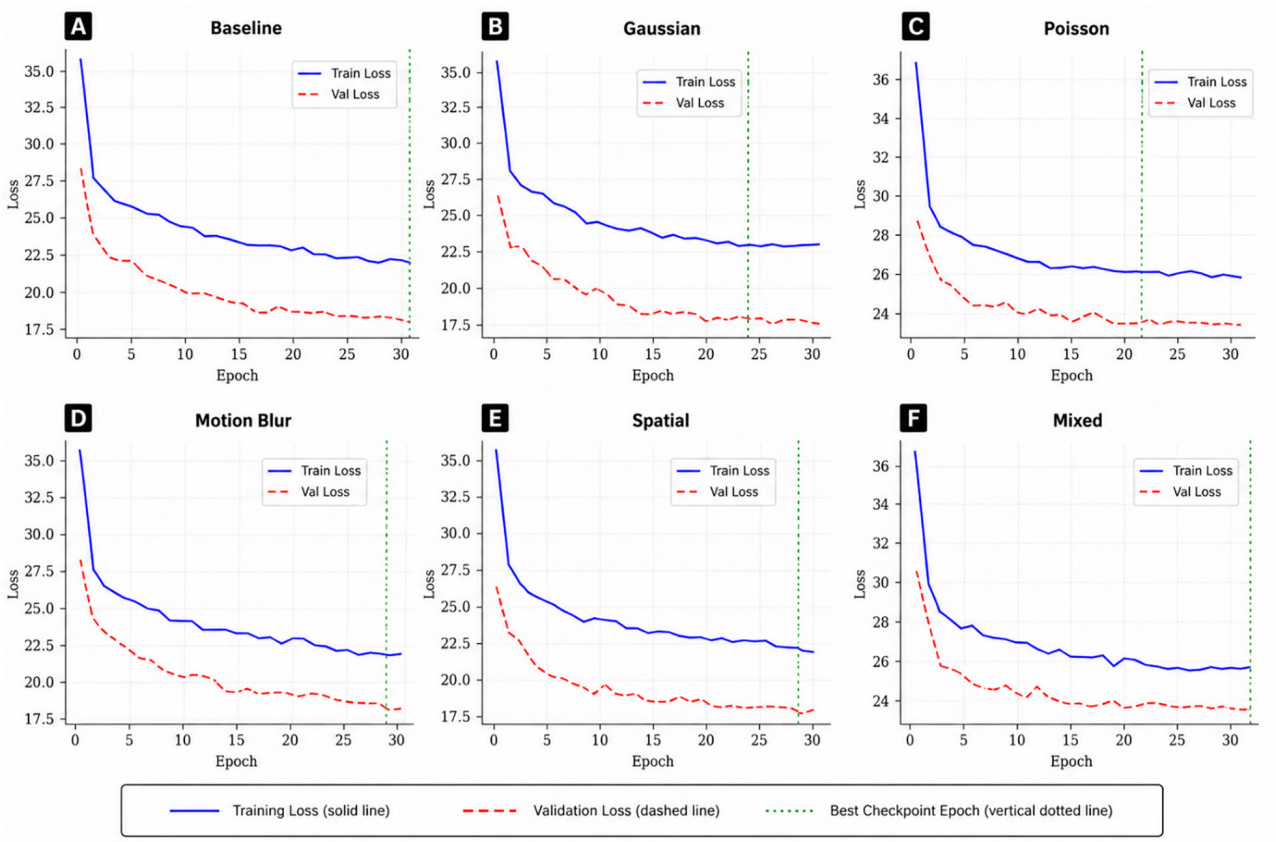

Figure 6. *Training and validation loss curves over 30 epochs for all six models. (A–F) Each sub-panel corresponds to one model: (A) Baseline, (B) Gaussian, (C) Poisson, (D) Motion Blur, (E) Spatial, (F) Mixed. The solid line represents training loss, the dashed line represents validation loss, and the green dotted vertical line marks the epoch of the best checkpoint.*

example, one value of σ for Gaussian noise, one scale for Poisson noise, and one kernel size for motion blur. This choice was made to ensure fair comparison under limited computational resources on a GeForce RTX 3050 GPU, but it likely underestimates the potential performance of more carefully tuned augmentation strategies.

A further limitation is the use of a single backbone architecture. All experiments were conducted with ResNet-18, and it is not known whether the observed robustness trends generalise to larger architectures or transformer-based models such as Vision Transformers (ViT). This should be acknowledged when drawing broader conclusions about the utility of noise augmentation for pose regression. Future work could address these limitations in several directions. One approach is to perform a broader hyperparameter sweep, for example, by randomizing noise strength per batch as in RandAugment or AugMix, in order to better cover the noise space. Another direction is to train for longer schedules using larger backbone networks and the full dataset to further improve clean data accuracy while preserving robustness. In addition, evaluation on real corrupted datasets, such as nighttime sequences from TUM VI, Oxford RobotCar, or low-light EuRoC subsets, would provide a more realistic assessment than synthetic corruptions alone. Finally, exploring adaptive or learned augmentation policies that select noise types based on training dynamics or domain priors could further improve performance.

Combining multiple noise types in a more structured pipeline, for example, sequential Poisson, Gaussian, and motion blur with randomized strengths, or integrating geometry-aware auxiliary losses such as photometric or feature metric consistency, may further reduce the performance gap to classical geometry-based methods in extreme conditions. With greater computational resources, these extensions are likely to produce even more reliable and deployable learning-based pose estimation systems.

CONCLUSION

This study analysed the systematic effects of noise augmentation on the robustness of pose regression models. Using a controlled benchmark of synthetic image degradations, including Gaussian noise, Poisson shot noise, motion blur, spatially correlated noise, and realistic mixed pipelines, results demonstrate that targeted augmentations substantially improve translation accuracy under corrupted inputs, yielding error reductions of 20% to 70% compared to the non-augmented baseline.

The results reveal a clear trade-off in robustness. Gaussian noise augmentation provides the strongest overall performance across diverse corruption types, while Poisson-based and mixed augmentations are essential for realistic low-light scenarios that dominate many real-world deployments. The Conservative Pose Loss with $\beta = 10$ further supports this prioritization of translation accuracy without requiring any architectural changes.

These findings offer practical guidance for building more reliable visual odometry systems in safety-critical applications such as robotics and autonomous navigation. Future work could involve longer training on full-scale datasets and evaluation on real low-light video sequences. In addition, integrating geometric consistency losses may help further improve performance while maintaining the flexibility of learning-based approaches. These extensions would help narrow the gap between experimental benchmarks and real-world deployment.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

1. Lowry S, Sünderhauf N, Newman P, Leonard JJ, Cox D, Corke P, *et al*. Visual place recognition: a survey. *IEEE Trans Robot*. 2016; 32 (1): 1-19. https://doi.org/10.1109/TRO.2015.2496823
2. Cadena C, Carlone L, Carrillo H, Latif Y, Scaramuzza D, Neira J, *et al*. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. *IEEE Trans Robot.* 2016; 32 (6): 1309-32. https://doi.org/10.1109/TRO.2016.2624754
3. Groves PD. Principles of GNSS, inertial, and multisensor integrated navigation systems. 2nd ed. Norwood (MA): Artech House; 2013.
4. Zhuang Y, Shao W, Hua N, Ruan W, Gu F, Yang J, *et al*. A survey of positioning systems using visible LED lights. *IEEE Commun Surv Tutor.* 2018; 20 (3): 1963-97. https://doi.org/10.1109/COMST.2018.2806558
5. Sattler T, Maddern W, Toft C, Torii A, Hammarstrand L, Stenborg E, *et al*. Benchmarking 6DOF outdoor visual localization in changing conditions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2018 Jun 18-22; Salt Lake City, UT. Piscataway (NJ): IEEE. 2018; p.8601-10. https://doi.org/10.1109/

CVPR.2018.00897

6. Dodge S, Karam L. Understanding how image quality affects deep neural networks. ArXiv [Preprint]. 2016 [cited 2026-01-10]. Available from: https://arxiv.org/abs/1604.04004
7. Rabiee H, Biswas J. IV-SLAM: introspective vision for simultaneous localization and mapping. ArXiv [Preprint]. 2020 [cited 2026-01-10]. Available from: https://arxiv.org/abs/2008.02760
8. Hu T, Scargill T, Yang F, Chen Y, Lan G, Gorlatova M. SEESys: Online pose error estimation system for visual SLAM. In: SenSys '24: Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems; 2024 Nov 4-8; Hangzhou, China. New York (NY): ACM. 2024; p.322-35. Available from: https://pure.tudelft.nl/ws/portalfiles/portal/233033631/3666025.3699341.pdf [cited 2026 Jan 10]. https://doi.org/10.1145/3666025.3699341
9. Abdelhamed A, Lin S, Brown MS. Noise flow: noise modeling with normalizing flows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV); 2019 Oct 27-Nov 2; Seoul, Korea. Piscataway (NJ): IEEE. 2019; p.3165-74. https://doi.org/10.1109/ICCV.2019.00326
10. Henz J, Gastal ESL, Oliveira MM. Deep learning-based noise and blur synthesis for camera data augmentation. *Sensors*. 2019; 19 (5): 1120.
11. Fu H, Luo A, Gu J, Cheng M, Lin L. Modeling spatially correlated sensor noise in sRGB images. *IEEE Trans Image Process*. 2023; 32: 5462-75.
12. Burri M, Nikolic J, Gohl P, Schneider T, Rehder J, Omari S, *et al*. The EuRoC micro aerial vehicle datasets. *Int J Robot Res*. 2016; 35 (10): 1157-63. https://doi.org/10.1177/0278364915620033