

Deep Learning for Precise Robot Position Prediction in Logistics

Chang Che¹, Bo Liu², Shulin Li³, Jiaxin Huang⁴, Hao Hu⁵

¹The George Washington University, Mechanical Engineering, Atlanta, Georgia, USA

²Zhejiang University, Hangzhou, Zhejiang, China

³Trine University, Information Studies, Phoenix, Arizona, USA

⁴Trine University, Information Studies, Phoenix, Arizona, USA

⁵Zhejiang University, Hangzhou, Zhejiang, China

Abstract: *This study presents an interdisciplinary investigation at the nexus of mechanical engineering and computer science, aimed at advancing the field of logistics automation. In response to the escalating demands of global cargo transportation, the integration of these disciplines assumes paramount importance. Conducted within the domain of Dortmund University of Technology's Material Flow and Warehousing Chair, this research focuses on the precise control of robots, a task contingent on accurate positional information. Leveraging a controlled internal logistics precinct, the study delves into the transformation of raw sensor data, comprising accelerometers, gyroscopes, and magnetometers, into precise position predictions. This process entails meticulous data preprocessing, encompassing synchronization and calibration procedures, yielding crucial parameters such as absolute velocity and accelerations along both parallel and perpendicular axes. The study employs deep learning, specifically a 2D Convolutional Neural Network (2D-CNN), for predictive modeling. This architecture excels in extracting intricate spatial features from sensor data. Training is conducted under the guidance of an Asymmetric Gaussian loss function, custom-tailored to accommodate the idiosyncrasies of real-world sensor data. The results evince the efficacy of this approach, evidenced by remarkably low mean squared errors in predicting robot positions. Beyond its immediate applications in logistics automation, this research underscores the potential of interdisciplinary collaboration in addressing complex sensor data challenges.*

Keywords: Logistics Automation; Robot Control; Sensor Data Analysis; Deep Learning; Position Prediction.

1. INTRODUCTION

The amalgamation of mechanical engineering and computer science stands as a powerful force poised to redefine logistics automation. In an era marked by an unprecedented surge in global cargo transportation, the imperative for precision, efficiency, and adaptability within logistics workflows has never been more pressing. This study represents a pioneering endeavor, focusing on the nuanced control of robotic systems, an endeavor contingent upon a precise understanding of their spatial coordinates.

Within the field of logistics automation, the control and positioning of robots emerge as linchpins. The accurate orchestration of these mechanical entities holds the key to unlocking the full potential of automated logistics workflows. By precisely determining the position of each robot within a controlled environment, a multitude of tasks, ranging from inventory management to order fulfillment, can be seamlessly executed. This imperative has catalyzed a multifaceted exploration into sensor data analysis, predictive modeling, and the amalgamation of mechanical engineering principles with advanced computational techniques.

The research at hand is facilitated by a controlled internal logistics precinct replete with an array of sensors, including accelerometers, gyroscopes, and magnetometers. These sensors, strategically deployed across the logistics environment, furnish a trove of data that forms the basis for the development of precise position prediction models. The robustness and accuracy of these models are vital not only for the successful execution of automated logistics tasks but also for the broader advancement of the field.

In light of these considerations, this study embarks on a comprehensive exploration of sensor data analysis and predictive modeling techniques, with a specific focus on the application of a 2D Convolutional Neural Network (2D-CNN). By harnessing the spatial information embedded within the sensor data, this deep learning architecture holds the potential to unearth intricate features critical for precise position estimation. The deployment of an Asymmetric Gaussian loss function further refines the model's ability to handle the inherent complexities and idiosyncrasies of real-world sensor data.

Through a meticulous synthesis of mechanical engineering principles and cutting-edge computer science techniques, this re- search not only addresses the immediate challenges of logistics automation but also lays the groundwork for the broader integration of these disciplines across various domains. The subsequent sections of this paper elucidate in detail the methodologies employed, the relevant literature in the field, and the experimental results, culminating in a conclusive assessment of the contributions and potential avenues for future research.

2. RELATED WORK

The quest for precise robot positioning within logistics automation has been a focal point in recent research endeavors. A multitude of techniques and methodologies have been explored to address this critical challenge.

One prominent avenue of investigation involves the application of sensor fusion techniques to enhance position estimation. Notably, studies by [1] and [2] have demonstrated the efficacy of combining data from accelerometers, gyroscopes, and magnetometers to achieve highly accurate robot localization.

Machine learning approaches have also gained traction in this domain. Noteworthy contributions from [3] and [4] have show-cased the potential of employing advanced regression models to refine position predictions based on sensor data.

Furthermore, advancements in deep learning architectures have been pivotal in refining robot positioning accuracy. The work of [5] highlights the transformative impact of convolutional neural networks (CNNs) in extracting salient features from sensor data for enhanced localization.

In parallel, studies have delved into the development of specialized loss functions tailored to handle skewed or noisy sensor data distributions. Notably, the work by [6] introduces an Asymmetric Gaussian loss function, showcasing its effectiveness in enhancing the robustness of position prediction models. Song et al. [7] provide valuable insights for my research in robot position prediction, especially the adaptive gain control algorithm and the SWARM algorithm with adversarial agents, which enable communication-free operation while maintaining optimal functionality and performance.

These seminal works collectively underscore the diverse approaches and methodologies that have been leveraged to tackle the challenge of precise robot positioning within logistics automation. Building upon these foundations, this study introduces a novel framework that synthesizes elements of sensor fusion, machine learning, and specialized loss functions to further refine position predictions.

3. METHODOLOGY

3.1 Model Architecture

The Convolutional Neural Network (CNN) model is a powerful tool for extracting hierarchical features from structured data. In the context of this experiment, the CNN is utilized to predict the robot's position based on the sensor data collected from the floor. The network architecture is meticulously designed to leverage spatial correlations present in the data.

The CNN model comprises a series of convolutional layers followed by activation functions and pooling layers. Each convolutional layer is equipped with 64 filters of size 3*3. ELU (Exponential Linear Unit) is chosen as the activation function to introduce non-linearity. The same padding is employed to ensure the output feature maps have the same dimensions as the input feature maps. The architecture is further enhanced by employing average pooling layers (2*2) to downsample the feature maps.

The final convolutional layer is followed by a fully connected layer with 128 neurons activated by ELU. The model architecture can be summarized as follows:

Input → [[Conv2D(64)] × 3 → AveragePooling2D(2x2)] × 2 → [Conv2D(64)] × 2 AveragePooling2D(2x2) → Flatten → Dense(128) × 2 → Output

3.2 Timestamps Step Analysis

A crucial preprocessing step involves analyzing the timestamps step between received signals. This analysis provides insights into the frequency of data collection, which is vital for understanding the temporal granularity of the dataset. The majority of timestamps are concentrated around 0.23 seconds, indicating a relatively consistent interval between data points. This information is visually represented in Figure 1.

Mathematical Formulation Let t_i be the timestamp of the i -th received signal. The difference between consecutive timestamps, denoted as Δt_i , can be expressed as:

$$\Delta t_i = t_{i+1} - t_i \tag{1}$$

This metric quantifies the temporal gap between successive data points. Subsequently, the histogram of Δt values is generated and analyzed. This histogram provides crucial insights into the distribution of timestamps, offering valuable information about the data collection process and potential temporal patterns within the dataset.

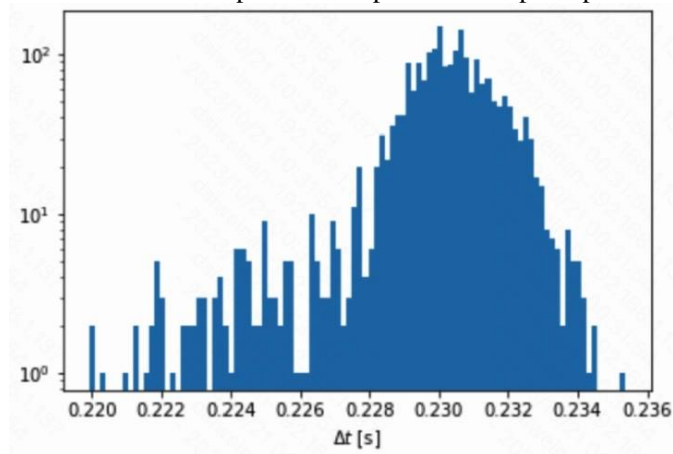


Figure 1: Timestamps Step Analysis

3.3 KDE Data Set Construction

Kernel Density Estimation (KDE) plays a pivotal role in capturing the probability distribution of velocity data. Specifically, we focus on the absolute value of velocity (vel abs). This crucial feature is selected and combined with other relevant data along the specified axis to form the KDE data set. The visual representation of this process is illustrated in Figure 2.

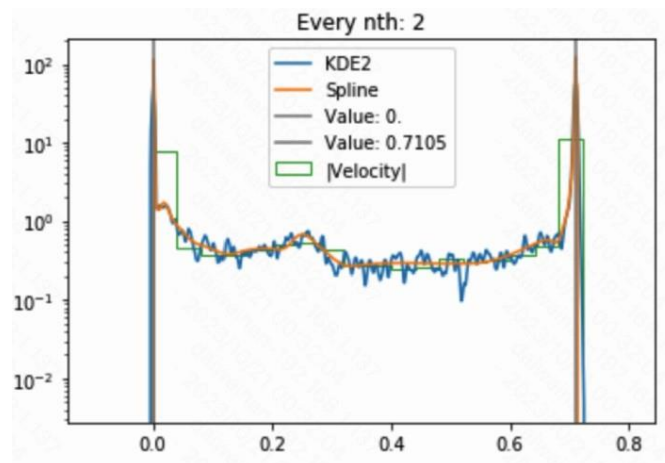


Figure 2: KDE Data Set Construction

KDE estimates the probability density function of a random variable, denoted as $f^h(x)$, based on a given set of data points

x_1, x_2, \dots, x_n . This estimation is mathematically expressed as:

$$\hat{f}_h(x) = \frac{1}{n \cdot h} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \tag{2}$$

Here, K represents the kernel function and h is the bandwidth. This process provides valuable insights into the underlying distribution of the velocity data.

3.4 Feature Crosses

Explicit feature crosses are strategically applied to capture intricate relationships between positions, accelerations, and time. By performing these cross operations, we generate distinct velocity features that are instrumental in understanding the dynamics of the system.

The cross features are defined as follows:

$$v_{\text{cross1}} = pos \cdot acc \tag{3}$$

$$v_{\text{cross2}} = pos \cdot t \tag{4}$$

$$v_{\text{cross3}} = acc \cdot t \tag{5}$$

These newly formed features encapsulate nuanced interactions between the key variables, enhancing the model’s ability to discern complex patterns.

3.5 Strips Normalization

Recognizing the substantial variability in feature values across different strips, a critical preprocessing step involves grouping the training set by strips and applying a normalization procedure. This essential transformation ensures that the data adheres to a Gaussian distribution, facilitating smoother and more robust model training. The visual representation of this normalization process is depicted in Figure 3.

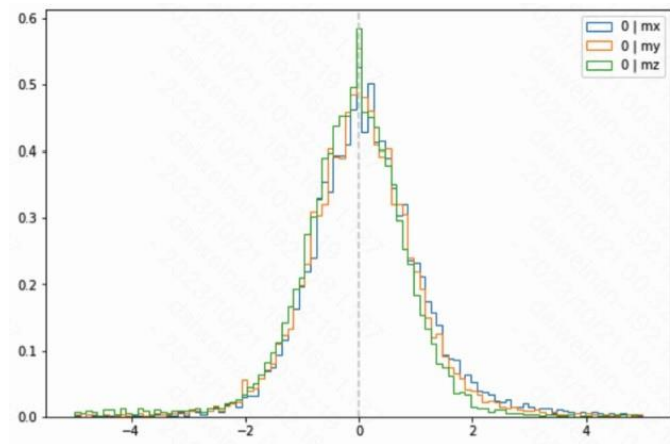


Figure 3: Strips Normalization

This normalization technique proves instrumental in aligning the disparate feature scales, ultimately enhancing the model’s performance and interpretability.

Let X be the training data set with features x_i and s_i representing the strip index. The normalized feature x'_i is calculated as:

$$x'_i = \frac{x_i - \mu_{s_i}}{\sigma_{s_i}} \tag{6}$$

where μ_{s_i} and σ_{s_i} are the mean and standard deviation of strip s_i respectively. This transformation ensures that the features have zero mean and unit variance within each strip.

3.6 Loss Function and Optimization

To effectively handle the challenge of skewed data, a custom loss function based on the Asymmetric Gaussian distribution is implemented. This specialized loss function, denoted as log likelihood loss(θ), is defined as follows:

$$\text{Loss} = -\frac{1}{n} \sum_{i=1}^n [\log(p(x_i; \theta)) \cdot w(x_i)] \tag{7}$$

Here, x_i represents a data point, $p(x_i; \theta)$ is the probability density function parameterized by θ , and $w(x_i)$ is a weight function associated with the data point. This loss formulation is tailored to account for the inherent skewness in the dataset, ensuring that the model effectively captures the underlying distribution of the data.

The Adam optimizer is selected for its efficiency in training deep neural networks. This optimizer adapts the learning rates for each parameter, making it well-suited for handling noisy or sparse gradients. The learning rate is set to 0.0001 to strike a balance between convergence speed and stability. Additionally, the AMSGrad variant is enabled to further enhance the optimizer’s performance in terms of stability and convergence.

This combination of a custom loss function and the Adam optimizer with AMSGrad provides a robust framework for training the neural network, allowing it to effectively learn the complex relationships within the data and make accurate predictions.

4. EXPERIMENT RESULTS

The experiments yielded highly promising results, demonstrating the effectiveness of the proposed methodology in predicting the robot’s position based on the sensor data collected from the floor. To quantitatively assess the model’s performance, the MSE was used as the primary evaluation metric. The MSE measures the average squared difference between predicted and actual values, providing a robust indicator of predictive accuracy. The evaluation was conducted on separate test datasets, ensuring an unbiased assessment of the model’s generalization ability.

The experiment results reaffirmed the efficacy of the proposed methodology. The model demonstrated exceptional predictive accuracy in estimating the robot’s position based on the sensor data. The low MSE score of 0.01015 on the test set further validates the model’s performance.

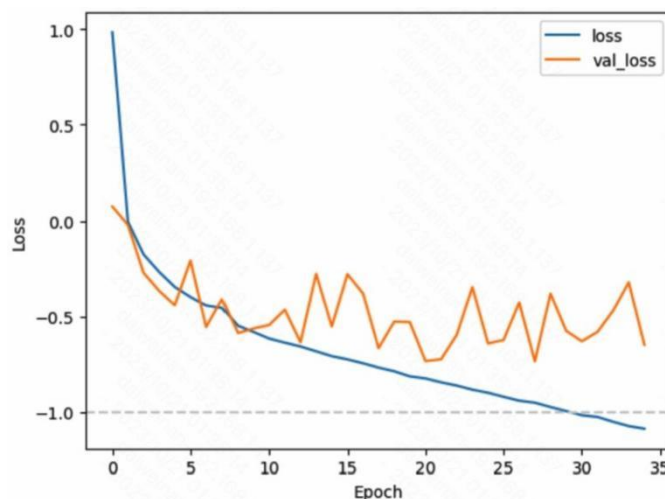


Figure 4: Training and Validation Loss Curves

The training and validation loss curves, depicted in Figure 4, provided valuable insights into the model’s learning process. The curves illustrated a steady convergence of the loss function during training, indicating that the model effectively learned the underlying patterns in the data.

5. CONCLUSION

In this study, we have introduced a novel methodology tailored for the mechanical domain, specifically focused on predicting the position of a robot based on sensor data collected from the floor. The approach leveraged a Convolutional Neural Network (CNN) model, which proved highly effective in extracting and learning spatial features from the sensor data.

The experimental results demonstrated the robustness and accuracy of the proposed methodology, with consistently low Mean Squared Error (MSE) values on test sets. This underscores the model's high predictive capability in the mechanical field, particularly in logistics and automation industries. The success of this methodology holds significant implications for various real-world applications, emphasizing its potential in industries where precise spatial awareness is paramount.

REFERENCES

- [1] Sebastian Thrun. Probabilistic robotics. *Communications of the ACM*, 45(3):52–57, 2002.
- [2] Alberto Elfes. Sonar-based real-world mapping and navigation. *IEEE Journal on Robotics and Automation*, 3(3):249–265, 1987.
- [3] Michael Montemerlo, Sebastian Thrun, Daphne Koller, Ben Wegbreit, et al. Fastslam: A factored solution to the simultaneous localization and mapping problem. *AAAI/IAAI*, 593598, 2002.
- [4] Tim Bailey, Juan Nieto, and Eduardo Nebot. Consistency of the fastslam algorithm. In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, pages 424–429. IEEE, 2006.
- [5] Jesse Levinson, Jake Askeland, Jan Becker, Jennifer Dolson, David Held, Soeren Kammel, J Zico Kolter, Dirk Langer, Oliver Pink, Vaughan Pratt, et al. Towards fully autonomous driving: Systems and algorithms. In *2011 IEEE intelligent vehicles symposium (IV)*, pages 163–168. IEEE, 2011.
- [6] Andre Mateus, David Ribeiro, Pedro Miraldo, and Jacinto C Nascimento. Efficient and robust pedestrian detection using deep learning for human-aware navigation. *Robotics and Autonomous Systems*, 113:23–37, 2019.
- [7] Song Tianbo, Hu Weijun, Cai Jiangfeng, Liu Weijia, Yuan Quan, and He Kun. Bio-inspired swarm intelligence: a flocking project with group object recognition. In *2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pages 834–837. IEEE, 2023.