Enhancing Robot Localization Accuracy through Neural Networks and Boosting Techniques

[1]Rob Argent and [2]Derek Stewart
Faculty of Sciences, University of Tizi Ouzou, Algeria.
[1]robarg5585@gmail.com, [2]stewartd@ummto.dz

Correspondence should be addressed to Derek Stewart : stewartd@ummto.dz

Article Info
Journal of Robotics Spectrum (https://anapub.co.ke/journals/jrs/jrs.html)
Doi: https://doi.org/10.53759/9852/JRS202301015
Received 10 October 2023; Revised from 15 November 2023; Accepted 18 December 2023.
Available online 26 December 2023.

Abstract – Robot localization is the computational procedure used to determine the exact spatial coordinates of a mobile robot relative to its surrounding environment. The acquisition of localization is an essential ability for an autonomous robot, since it plays a fundamental role in enabling the robot to accurately ascertain its own location. The understanding of the robot's spatial coordinates is an essential need for the robot to make well-informed decisions on its subsequent actions. In a normal scenario of robot localization, a map of the surrounding region is available, and the robot is equipped with sensors that facilitate the examination of the environment and the monitoring of its own motion. The difficulty of localization thereafter becomes the job of calculating the precise orientations and position of the robot inside the map via the use of data gathered from these sensors. To adequately handle the existence of noisy observations, it is essential for robot localization algorithms to possess the capacity to not only provide an assessment of the robot's position, but also to quantify the degree of uncertainty associated with this estimation of location. The main purpose of this study is to formulate a methodology that combines neural networks with boosting techniques to enhance the effectiveness of robot localization. The suggested methodology entails the selection and validation of neural network topologies, the extraction of pertinent features, and the use of boosting methods to augment classification performance. The objective is to get accurate and dependable robot localization via the use of these strategies.

Keywords – Robot Localization, Convolutional Neural Networks, Long Short-Term Memory, Long-Term Recurrent Convolutional Network, Support Vector Machine.

I. INTRODUCTION

The localization of mobile robots, which involves determining the location and orientation of the robot inside its working environment, is a crucial task for achieving autonomous navigation capabilities. In the context of interior applications such as the manufacturing sector and hospitals, it is reasonable to assume that the motion of the robot frame is planar. Under this assumption, the pose of the robot may be characterized by the position coordinates $x(t)$ and $y(t)$ of a specific point on the robot, as well as the absolute orientation angle $\varphi(t)$. Over the last several decades, many distinct methods have been put out to address the challenge of robot localization. Burghardt and Lanillos [1] categorized these approaches into two broad categories: relative localization and absolute localization.

Relative localization involves the utilization of internal measures to ascertain the progression of the robot's posture via the process of time integration of the kinematic equations. The primary approaches used in this category include inertial navigation, which involves the acquisition of acceleration and angular velocity data from various sites on the robot; and odometry, which relies on the measurement of wheel rotation angles. Due to its affordability and frequent updating capabilities, odometry has become an extensively used method. Nevertheless, the inevitable unlimited increase in time integration errors as the robot covers greater distances is a notable nuisance. The investigation of the spread of odometry errors is a significant area of study in the sector of mobile robots. Numerous studies have been done to address this issue, focusing on either the calibration of systematic errors, or the modeling of non-systematic random errors.

There remains a limited number of mobile robots that has the capability to really achieve autonomous placement and navigation. Autonomous line inspection is the predominant method used in the industrial sector for mobile robot navigation. Angerer, Strassmair, Staehr, Roettenbacher, and Robertson [2] present a mobile robot designed for large-scale automatic production or logistics storage lines. This robot is capable of transporting shelves filled with goods to various locations, under the control of a centralized warehouse control system. Additionally, the robot can strategically position the best-selling goods at the front of the shelves and determine the most efficient path for its movement. The use of Simultaneous

Localization and Mapping (SLAM) technology in [3] enables autos to receive map updates, ensuring unobstructed navigation over global regions and facilitating destination identification. The topic of SLAM based features is explored in [4], where the algorithm's dependability is assessed by simulation tests. This study focuses on the exploration of literature [5] that examines the application of SLAM techniques in the context of laser range and monocular vision for autonomous navigation of mobile robots within interior environments with unknown characteristics.

Kutila, Pyykönen, Ritter, Sawade, and Schäufele [6] have used laser rangefinders as a means of gathering data. By using this technology, they have successfully achieved autonomous navigation of mobile robots inside interior environments, without the need for any pre-established parameters. FastSLAM is a method that combines recursive techniques to approximate the robot pose and compute the whole posterior distribution of map landmarks. This enables the building of a map and localization of the robot. Zhang, Qin, Ma, and You [7] propose an optimal resampling approach that facilitates the movement of the sub-plasma towards the superior particles. The proposed technique effectively mitigates the deterioration of particle swarm and has significant importance in maintaining the variety within the swarm. FastSLAM is a commonly used algorithm in complicated environments because to its efficient computational speed and robust resistance to interference, which is facilitated by the independence of data association.

Computer vision-based positioning technology typically utilizes the picture frames captured by the camera in order to compute its own stance. The method for feature detection will extract descriptors and feature points for every each picture frame. The determination of the translation vector and rotation matrix between the camera and the object coordinate system is possible under specific constraints. Additionally, visual properties can be extracted based on the corresponding relationship between the positions of these feature descriptors and points in the image. This relationship is established through the matching process between two images. Ultimately, targeted location identification may be achieved via the use of precise algorithms.

The present study examines the methodology used in determining the precise location of a robotic entity via the utilization of binary classification algorithms. The proposed methodology put forward by the authors involves the integration of motion and sensor data with simulation in order to provide annotated training data. Two techniques are used in this study: one way relies on the robot`s distribution probability stance, while the other approach utilizes statistical characteristics obtained via perception and particle filter-based pose estimation. The authors also investigate the process of selecting and validating neural network topologies for classification tasks, such as convolutional neural networks and recurrent networks. The researchers find advantageous characteristics for enhancing performance and use AdaBoost and Support Vector Machines as classification algorithms. The evaluation and comparison of the suggested technique are conducted to assess its efficacy in precisely localizing the robot. The subsequent sections of the article have been structured in the following manner: Section II presents a discussion of related works on the selection and validation of neural networks for the purpose of localization classification. Section III presents a detailed evaluation of the results on localization accuracy estimation approach. Lastly, Section IV draws a conclusion to the paper, and proposes future research works.

II. RELATED WORKS

The literature review section examines the methodology used in the selection and validation of neural networks for the purpose of localization classification. The proposed approach by the authors involves using a fusion of CNN to extract features and recurrent networks to learn the temporal evolution of the particle collection. This integrated model is referred to as a Long-Term Recurrent Convolutional Network (LRCN). Three distinct network architectures, namely and Long-term Recurrent Convolutional Networks (LRCN) (Figure 3), Convolutional neural networks (CNNs) (**Fig 2**), and Long Short-Term Memory (LSTM) (**Fig 1**), are developed and trained in order to assess the efficacy of various network topologies.

According to Krizhevsky, Sutskever, and Hinton [8], Convolutional neural networks (CNNs) consist of linked layers of neurons, including pooling, convolutional, and fully connected layers. CNNs are mathematical models that are specifically built to handle data with many dimensions. These networks have the ability to learn simpler patterns at shallower layers and gradually shift to more complex patterns as the depth of the network increases. The use of numerous hidden layers in deep neural networks effectively addresses the challenge posed by the exponential growth of parameters. CNN has two prominent characteristics, namely weight sharing and local connection. Weight sharing refers to the practice of using same weights across all nodes inside a given layer. The concept of local connection refers to a scenario where individual nodes exclusively receive input from a limited number of local values inside an array. Additionally, each output generated by these nodes is specifically associated with certain segments of the input vector, as seen in **Fig 2.**

Javanmardi and Liu [9] argue that the human cognitive capacity exhibits persistence, a characteristic that is not well replicated by conventional neural networks, hence presenting a significant limitation. For instance, consider the scenario where one aims to categorize the many sorts of occurrences transpiring at each juncture of a film. The manner in which conventional neural networks use inference from preceding events to influence subsequent occurrences remains ambiguous. The challenge at hand may be effectively addressed via the use of recurrent neural networks. RNNs are characterized by their ability to retain and propagate information during loops, leading to significant achievements in several study domains such as voice recognition, language modeling, and translate on.

According to Kumar, Goomer, and Singh [10], LSTM models are a distinct variant of Recurrent Neural Networks (RNNs) that has the ability to acquire and retain long-term dependencies. They have exceptional proficiency in a diverse range of issues and have attained widespread adoption.The fundamental component of the LSTM model is the cell state, represented by the horizontal line seen at the top of **Fig 1.** This cell state serves as the conduit via which the initial information is propagated without alteration. The gates may be described as a transfer mechanism that selectively incorporates information. They are composed of layers of point-by-point multiplication operations and sigmoid neural networks. The sigmoid layer produces values ranging from 0 to 1, which in turn decide the percentage of information that is allowed to flow through. The architecture of the LSTM neural network is shown in **Fig 1,** while the constituent departments and their respective tasks are presented below.

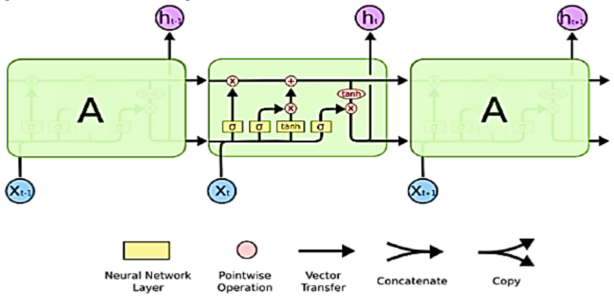

Fig 1. LSTM Structure Schematic

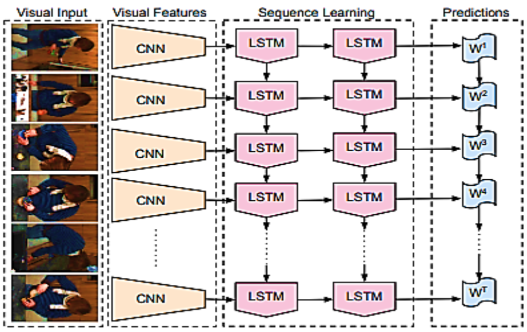

Fig 2. Long-term Recurrent Convolutional Networks (LRCN) architecture

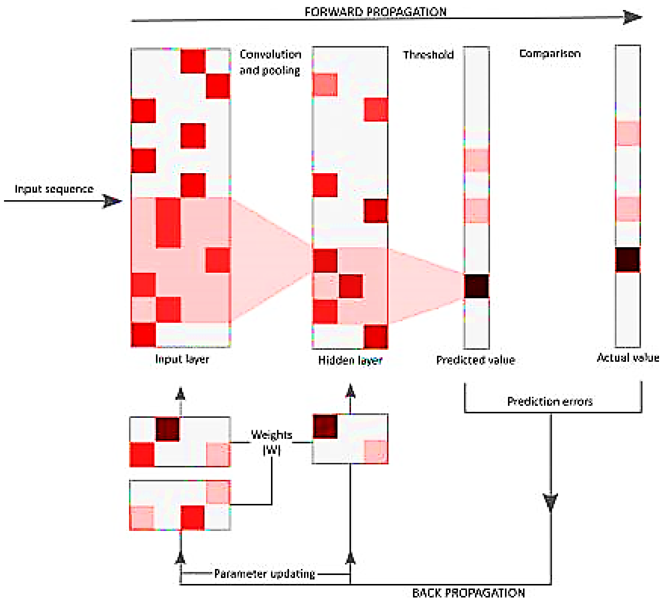

Fig 3. The CNN architecture

According to Li [11], the identification and characterization of visual content, such as photos and videos, provide a key obstacle in the field of computer vision. Supervised convolutional models have made significant advancements in the field of image identification tasks, and there have been recent proposals for various extensions aimed at processing video data. Ideally, a video model should possess the capability to handle input sequences of varying lengths and accommodate different length outputs. This includes the production of comprehensive sentence descriptions that surpass the typical one-versus-all prediction tests. Kim, Kim, and Lee [12] introduce a unique architecture called LRCNs for visual description and identification. LRCNs combine convolutional layers with long-range temporal recursion, and are designed to be end-to-end trainable. **Fig 3** presents an illustration of the proposed architecture.

Leonard and Durrant-Whyte [13] further identify advantageous characteristics for the boosting phase, whereby features taken from the particle collection are used to enhance the precision of calculating the robot`s localization state. The researchers use a localization scoring methodology, whereby they assess the correspondence between the robot`s map and the scan data. They proceed to document the observed characteristics while the robot navigates in a simulated environment in a random manner. Following the generation of a substantial quantity of feature samples, the researchers proceed to conduct an individual analysis of each feature. Subsequently, they partition the recordings into distinct localized and delocalized data sets in order to ascertain the degree of divergence between these two classes. The use of the Kullback-Leibler divergence, also known as KL-divergence, is employed by the authors to determine the dissimilarity between the probability distributions of the negative and positive instances. This approach aids in the identification and selection of features that possess strong discriminatory capabilities. A threshold is established for the KL-divergence, and features with a value beyond this barrier are chosen as informative features for the purpose of boosting.

According to Leyrit, Chateau, Tournayre, and Lapresté [14], the Adaboost method is a machine learning technique used for the purpose of face recognition. It utilizes eigenvalues as a means of extracting features. AdaBoost, commonly known as the adaptive boost algorithm, is a machine learning technique. In order to cultivate a proficient learner, it is vital to use several rounds inside the AdaBoost algorithm. The AdaBoost algorithm constructs a robust learner by progressively incorporating weak learners. In order to enhance the efficacy of a classifier via the use of multiple classifiers throughout the training process, a novel approach involves the incorporation of a new weak learner, with the adjustment of a weighting vector. This vector is particularly crafted to give priority to instances that were misclassified in earlier iterations. Facial recognition analysis has been widely used in various applications. Based on the results of literature review, many facial recognition algorithms have been developed. The AdaBoost technique, on the other hand, is a well-known simple technique that can be used to improve detection accuracy.

Several network topologies such as recurrent networks were examined and evaluated for their effectiveness using accuracy as a performance metric. Additionally, we aimed at unearthing key features that could potentially allow better localization estimates. This was done by employing localization score method and carrying out analysis of KL-divergence between localized and non-localized data sets. In order to come up with a classifier that learns how to do classification by itself, the AdaBoost algorithm combined weak classifiers together into a strong one. Support vector machines (SVMs) are also considered in this section as classifiers and their performances are compared against those of AdaBoost via binary classification results derived from neural networks.

III. LOCALIZATION ACCURACY ESTIMATION APPROACH

The primary objective of high-precision localization is to achieve centimeter or even millimeter level accuracy in localization. The technique called PL-ICP, which is a modified ICP (Iterative Closest Point) algorithm was introduced by Bouaziz, Tagliasacchi, and Pauly [15]. This approach utilizes point-to-line correspondence technique for proper alignment between scanned data. In their research paper, Saarinen et al. [16] proposed the new way of localizing NDT-MCL that uses the NDT model to describe both sensor and the map information efficiently. However, this method can only localize to an accuracy of around 5cm which is inadequate for precision requirements within industrial environments. In their study, Rizzo, Seco, Espelosín, Lera, and Villarroel [17] provide a novel approach to robot localization that leverages both 3D structural information and 2D mapping techniques. While the use of 3D matching offers several benefits and effectively reduces space and processing demands, it is important to note that equipping mobile robots with 3D LiDAR technology will result in a significant rise in hardware expenses.

There is a subset of researchers that have started the exploration of including artificial landmarks as a means to enhance the mobile robots localization accuracy. One instance of a suggested visual system using 2D bar code type tags, namely AprilTag, was put out by [18]. This system enables a robot to acquire a six DOF (degrees of freedom) position by analyzing a single picture. However, the visual system is very susceptible to variations in lighting conditions, hence compromising the overall reliability of the localization system. In [19] enhanced the precision and reliability of the overall system of localization by integrating 2D laser technology with reflectors. This integration resulted in an angle error of less than 0.2 and a position error of less than 1 cm. Petrovskaya and Ng [20] in their study introduced an innovative way of improving the accuracy of mobile manipulator localization. They combined two laser rangefinders with several artificial landmarks for this purpose, which allowed them to obtain high precision in localization. The methodology used depends on trilateration, which states that when a distance is measured between an object and three known locations, the position of that item can be determined. They employed an adaptive unscented Kalman filter approach to increase the accuracy of localization.

In the proposed methodology, it is possible to represent both the process model and the observation model as linear when certain conditions are satisfied:

$$X_{k+1} = Fx_k + Gu_k + w_k \tag{1}$$

And

$$z_{k+1} = \mathrm{H}x_{k+1} + v_{k+1} \tag{2}$$

where X_{k+1}, X_k represent the state of systems at time denoted by $k+1$, k, F represent a matrix of system transition, G represents the control gain U_k, while W_k represents the Gaussian process noise with zero-mean $\overrightarrow{W_k} N(Q,O), H$ represents the matrix of observation, while v_{k+1} represents the Gaussian observation noise with zero-mean $\overrightarrow{v_{k+1}} N(R,O)$. In this scenario, the state estimations could be obtained based on the applications of the following Kalman filter equation considering the initial state x_0 considering the $\overrightarrow{x_0} N\left(\widetilde{x_0}, P_0\right)$ Gaussian distribution. Prediction is obtained using the following model:

$$z_{k+1} = \mathrm{F}x_k + \mathrm{G}u_k \tag{3}$$

$$\overline{H}_{k+1} = \mathrm{F}P_k F^T + Q \tag{4}$$

where the covariance S innovation ($z_{k+1} - \overline{H}_{k+1}$ is known as innovation) and the K Kalman are achieved by:

$$s = H\overline{P}_{k+1}H^T + R \tag{5}$$

$$K = \overline{P}_{k+1}H^T S^{-1} \tag{6}$$

In most experimental settings, the localization accuracy typically ranges around 8 mm. Nevertheless, the efficacy of the suggested localization approach is limited to small experimental scenarios. Additionally, the localization of the robot necessitates the simultaneous detection of more than three reflectors, hence introducing complexities in deploying the localization system. To assess the accuracy of the robot's location estimate, we use the methodology shown in **Fig 4**. In our application, we are faced with the task of addressing a binary classification issue, namely determining if the robot is accurately delocalized or located. The robot is said to be delocalized when the estimated orientation and location deviate from the real orientation and position by a magnitude above a predetermined threshold.

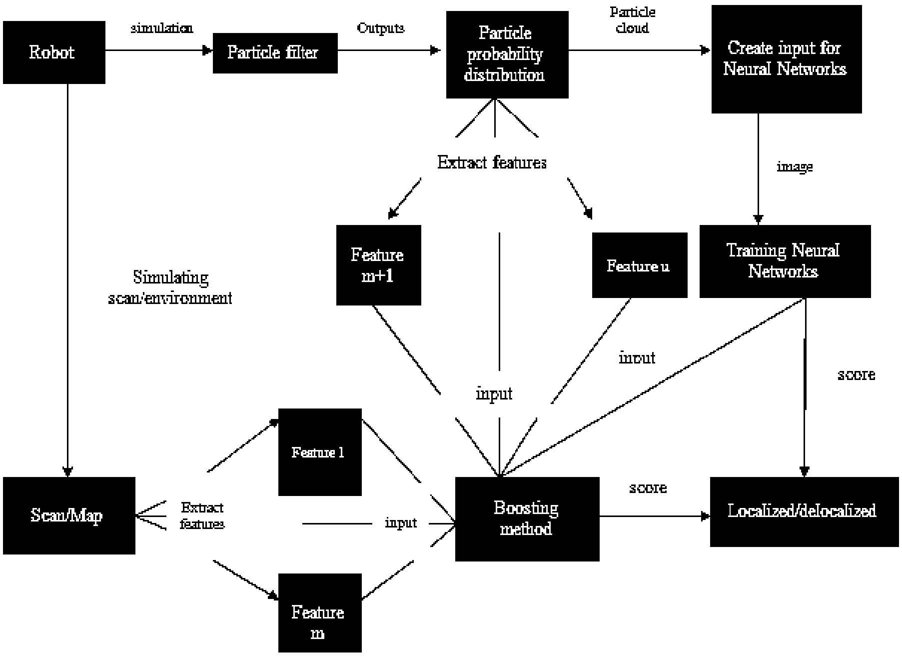

Fig 4. The General Methodology for Robot Localization Accuracy Estimation

To get accurate information on the position and orientation of the robot, as well as to induce delocalization, we use simulation techniques to mimic the robot's sensing and movement capabilities, such as laser scans, along with the virtual representation of the habitat. The recorded data from perception and particle filter-based posture estimation, namely particles, are turned into labeled training data for further classification procedure. In the categorization process, we use two methodologies that are also utilized in conjunction. The first approach is oriented on the perspective that the robot`s posture probability distribution, which is represented by the particle set, and its usual increase in uncertainty include valuable data about the localization accuracy.

Therefore, we conduct training on recurrent and non-recurrent neural networks to perform classification tasks using the particle dataset. The second technique is based on the premise that statistical characteristics, which are essentially weak classifiers, may be extracted from the perception and set of the particle. These features are then used in a boosting algorithm to train a strong classifier. This approach is motivated by the fact that the particle set effectively reflects a distribution within the habitat. By using statistical techniques to analyze the distribution and correlating the obtained scan data with the surrounding environment, it is possible to uncover insights on the localization status. In order to assess the potential of a particular feature, the KL-divergence is computed between a delocalized set and a localized set. Ultimately, the use of the classification derived from the trained neural network (NN) may be included alongside the weak classifiers during the boosting phase, with the aim of assessing if the NN output can provide additional enhancements to the classifier.

In the subsequent sub-sections, we will elaborate on the further stages of the suggested procedure: (1) the creation of training data, (2) the evaluation and selection of NN topologies, (3) the selection and validation of characteristics, and (4) the training of classifiers.

Creation of training data

This section provides an overview of the process by which the training data for the NN is created. **Fig 5** depicts the whole process of creating a sample of training and ascertaining the accurate label. In order to generate a representative sample, it is necessary to get the particle set derived from the filter of the particle. This particle set is then used to construct a training picture denoted as x. A particle is a representation of a possible pose, including both location and orientation, accompanied with a weight that signifies its significance. The particle may be denoted as $\langle x, y, \theta, w \rangle$. The efficacy of the filter of the particle is contingent upon the magnitude of the particle ensemble, necessitating the selection of an adequate quantity of particles.

In this study, a particle set size of $M = 1000$ was used. The last step involves transforming the collection of particles into a binary picture, since this representation effectively captures the spatial arrangement of the particles and is compatible with contemporary machine learning methodologies. The picture has black pixels that correspond to individual particles inside the particle cloud. The picture is subsequently annotated by comparing the precise posture $\langle xGT, yGT, \theta GT \rangle$ obtained from the simulation with the estimated pose $\langle xPF, yPF, \theta PF \rangle$ derived from the particle distribution. The robots' localization state may be classified and labeled using a preset distance threshold α, β between them. This allows for the creation of a training sample label y.

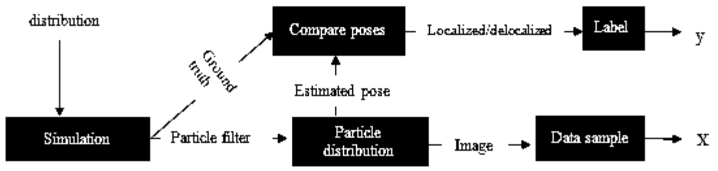

Fig 5. Generation of Data Labelling and Sampling for Robot Training

$$y = \begin{matrix} 0, & if \\ 1, & else \end{matrix} \sqrt{(x_{GT} - x_{PF})^2 + (yGT - yPF)^2} < \alpha \wedge \|\theta_{GT} - \theta_{PF}\| < \beta \quad (7)$$

Typically, the particles have a distribution that spans a substantial state space, including the whole of the environment map. Given the variation in size and the localized accumulation of particles on this map, it is not essential to retain the whole map as a sample of data. To prioritize pertinent regions of the distribution of particle, the picture is centered on the mean of the locations shown in the collection of particles. Furthermore, the direction of the image is aligned with the average orientation seen in the particles. Subsequently, the distribution's area is partitioned and resized to a predetermined standard dimension. Through the examination of the particle's distribution, an estimate of the ideal area size was derived. The procedure included partitioning the region into a square area of dimensions $s \times s$ meters, such that when charting the particle ensemble, it was seen that in 95% of instances, all particles were contained inside the designated area.

To ascertain the most favorable side, a dataset was compiled, including the relevant information pertaining to the particle cloud. The result yielded a cutting region of s = 1.5m. The use of a focused representation has the benefit of enabling the trained network to be applicable to contexts with varying sizes and shapes. The region that has been trimmed is then resized to form an image of dimensions $p \times p$, where p represents the pixels total number along each axis. The magnitude of p is contingent upon the specific characteristics encapsulated by the particle distribution, as well as the intended efficacy of the neural network training process. Through the examination of various values for p using pre-training on tiny NN, it was determined that a size of $p = 36$ pixels yielded the most favorable outcomes while maintaining satisfactory training performance. The selection of these settings was based on the outcomes of runtime and accuracy.

Evaluation and selection of neural network topologies

Once a training set for neural networks has been established, it becomes necessary to choose a suitable network structure, including the kind and layout, that can be effectively used for localization categorization. Therefore, it is necessary to train and test several potential network architectures.

Determining the optimal network type and structure for calculating the quality of a robot's localization using information from filters of the particle as a priori input presents challenges. This is mostly because to the ambiguity around the specific information that a neural network extracts in this context. One potential network architecture that might be used for the purpose of pattern recognition is a convolutional neural network (CNN). An alternative approach involves the use of a recurrent neural network, which has the capacity to acquire knowledge pertaining to the temporal evolution of the particle ensemble.

In order to comprehensively assess a diverse array of options, both sorts of networks are taken into consideration. The suggested methodology entails the utilization of a CNN to perform feature extraction, in conjunction with a recurrent network structure to enable the acquisition of temporal transformations. To optimize performance, a LSTM architecture is employed to train a recurrent neural network (RNN). Moreover, a hybrid model is trained by combining both CNN and LSTM components. The acronym LRCN is employed to denote this concept. The network structure is employed for evaluating the potential integration of the benefits provided by the preceding types. Three separate network architectures are constructed and trained for each of the primary network types, specifically LCRN, LSTM, and CNN. The three layouts created are primitive, middle and advanced layouts. The reason behind the use of these three layouts is that one may examine the consequences of a basic network configuration on under fitting as well as the influence of an intricate architecture on over fitting.

To determine the effectiveness of a network architecture, it is crucial to construct a quantifiable metric that enables the evaluation of different network configurations in terms of localization monitoring. Ghorpade, Zennaro, and Chaudhari [21] propose accuracy as a suitable metric for evaluation. Accuracy is a metric that quantitatively assesses the extent to which a given dataset is accurately classified. When conducting an analysis of a classification task that involves two distinct classes, it is feasible to assess the accuracy by partitioning the predictions of a given dataset into a binary classification. Given a binary classification problem, each class can be assigned either a positive or negative label. For example, the information sample which is being authenticated is inserted in the network architecture where its output is recorded. The resulting output is thereafter allocated to one of the two groups based on the outcome.

The accuracy of class assignment may be assessed by relating the given class label with the predicted class label. In the event that both the label and the result are congruent, the resultant outcome is deemed to be true. A data sample is considered a genuine positive when it is classed as positive and the predicted class is likewise positive. Conversely, when the sample of data is defined as an adverse and the outcome is also adverse, it is characterized as a genuine adverse. If the two categories are not congruent, the resulting conclusion is said to be incorrect. In cases when a data sample is categorized as negative despite an anticipated positive outcome, it is referred to as a false negative. Conversely, when a sample of data is labeled as optimistic but the forecast is negative, it is termed a false positive. The correctness of the data set may be determined by calculating the frequency of the findings for the given cases, as:

$$acc = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \tag{8}$$

where t_x shows various correct classifications and f_x represents the incorrect ones with x either n (negative) or p (positive)

Validation and selection of features

As the volume of data grows exponentially, there is a progressive decline in the quality of data needed for processing by various algorithms such as Pattern Recognition, Image Processing, Machine Learning, and Data Mining. Bellman refers to this situation as the "Curse of Dimensionality." The presence of noisy, irrelevant, and redundant data becomes more prevalent as the dimensionality of the data increases. Which intern contributes to the phenomenon of overfitting in the model and thereafter leads to an increase in the error rate of the learning algorithm. In order to address these issues, the use of "Dimensionality Reduction" methods is employed, which is an integral component of the preprocessing phase. Feature Extraction (FE) and Feature Selection (FS) are frequently used techniques for reducing dimensionality in various applications. Feature selection (FS) is a technique used to eliminate noisy, redundant, and unnecessary data. Consequently, there is an enhancement in performance. The research provides evidence that feature selection has the ability to enhance the predictive accuracy, scalability, and generalization capacity of classifiers. In the field of knowledge discovery, feature selection (FS) plays a crucial role in mitigating computational complexity, storage requirements, and associated costs. **Fig 6** illustrates the many steps of the FS process, which will now be elucidated. The performance of the system is contingent upon the choice made at each individual level.

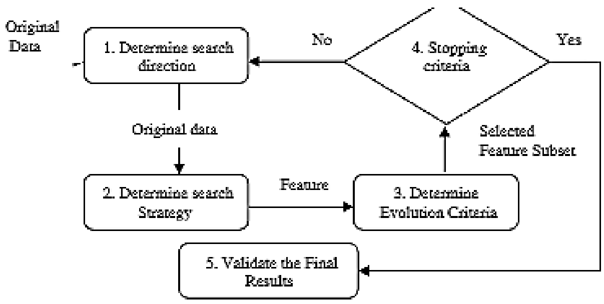

Fig 6. Stages of the Feature Selection Process

After the selection of an adequate network topology, it is necessary to identify important properties for the boosting stage. The objective is to enhance the accuracy of calculating the robot`s localization state by using characteristics taken from the particle set and incorporating them into an improving phase. Within the organizational setting of the firm where the project was conducted, a method known as localization scoring was used as a preliminary step prior to the implementation process. The proposed localization score methodology utilizes extracted characteristics obtained by relating the scan information with the map of the robots. An application of the Hough line transform [22] was conducted on the scan and map points, followed by a comparison between them. The study aimed to examine the potential of incorporating more data on the particle set, namely the balance point, into the current localization scoring process. By analyzing the characteristics, the researchers sought to see whether they may provide valuable insights into the evaluation of the robot's localization.

In order to ascertain the significance of certain aspects, the robot was subjected to random navigation inside a simulated environment, during which the corresponding features were meticulously documented. After creating around 100,000 property samples, each property is analyzed independently. The recordings have been divided into two distinct sets of data: a localized XQ dataset and a delocalized XP dataset. The sets are thereafter shown as a separation distribution over k bins in order to ascertain the discrepancy between the two classes. Nevertheless, the use of continuous variables provides a more comprehensive understanding of the robots' position estimation. Consequently, the initial continuous values were reintroduced for the purpose of enhancing the boosting process subsequent to the selection of pertinent characteristics. The determination of the bin size, denoted as k, is accomplished by examining the preexisting dataset and assessing the highest and lowest values within it. Subsequently, a bin size was decided based on the appropriate unit of measurement.

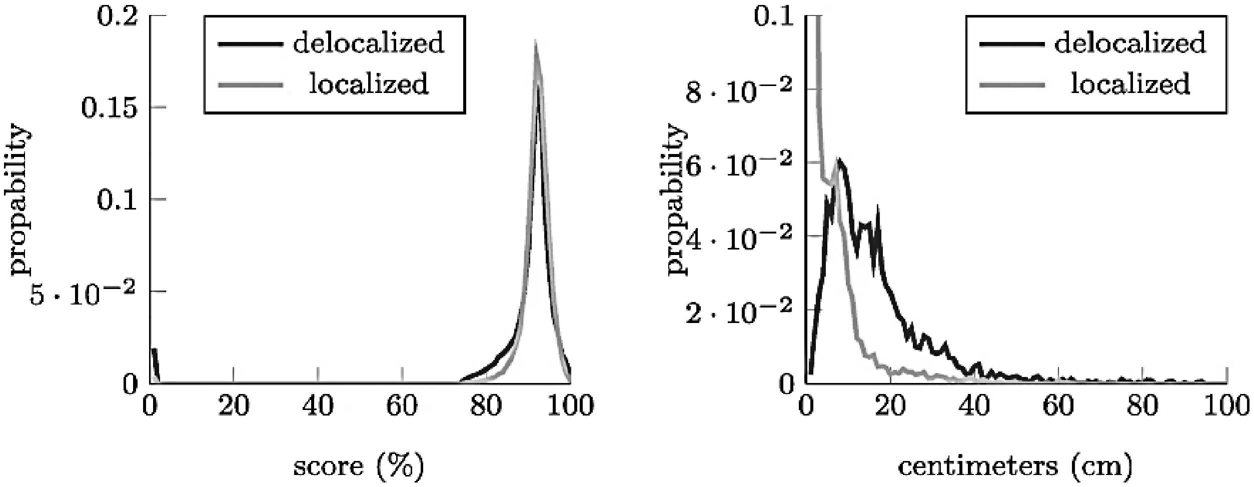

Fig 7. Property Values for the Delocalized and Localized Situations

Fig 7 presents the distributions probability of the property values for the delocalized and localized situations examined for features 18 (left, non-discriminating) and 31 (right, discriminating). The bin size for meters was set at 1 centimeter, while the bin size for percentage was set at 1 percent, and the bin size for degree was set at 1 degree. Each bin is comprised of many samples that are included inside it. The discrete probability distribution P may be mathematically expressed as the ratio of the samples falling into a certain bin to the samples total number, denoted as n.

$$P(i) = \frac{|bin(i)|}{n} \tag{9}$$

This is the same case for the Q distribution. With two distributions that are discrete (Q, P) for delocalized and localized collections one can effective calculate the divergence of Kullback-Leibler using the equation below:

$$D(P||Q = KL(P, Q) = \log\frac{P(i)}{Q(i)}.\sum_{i}^{k} P(i) \tag{10}$$

Where k represents the number of bins. KL-divergence is illustrated as $0 \rightarrow P(i) = \forall i{:}\, Q(i) = 0$ Applies. If $P(i) = 0$ the contributions of the i^{th} bin is 0. It is illustration that $D(P||Q) \geq 0$ for the various distributions and $D(P||Q) \geq 0$ if $Q = P$.

The use of the Kullback-Leibler divergence that measures the dissimilarity between the distributions of probability of property values for negative and positive situations, enables the identification of highly discriminative features. A greater value of KL-divergence indicates a greater dissimilarity between the distributions, hence implying a higher level of informativeness in the underlying feature. Based on this criterion, potential characteristics may be chosen for the process of boosting. A threshold of $D(P||Q) \geq 0.1$ was used as the criteria for feature selection. Property values for a discriminative and non-discriminative property are shown as a probability distribution in Figure 7.

Training of classifiers

Support vector machines

The Support Vector Machine (SVM) is a widely used and effective supervised statistical machine learning approach. It was introduced to the computer science society in the 1990s is mostly employed for solving classification issues. The adaptability of this method may be attributed to its ability to learn nonlinear decision surfaces and exhibit strong performance even when faced with a substantial number of predictors, even in scenarios when the number of instances is limited. The SVM has gained significant interest due to its suitability for addressing various problems, including but not limited to credit rating analysis, speech recognition, junk mail classification, text categorization, faulty card detection, face detection, image recognition, and diabetes and cancer classification. The majority of the foundational work for the SVM was established by Vladimir Vapnik during his doctoral studies in the Soviet Union in the 1960s [23].

The SVM approach operates by transforming the observations, or data points, into spatial representations. This transformation involves mapping the original observations from various categories, or classes, in a manner that maximizes the separation between them with a clear gap. The process of making predictions for new observations involves mapping these observations into a shared space, and then assigning them to a specific category based on their position relative to the gap [25]. Regarding the methodology using AdaBoost, we use the aforementioned chosen features from the preceding section and utilize the classification of binary outcome obtained from the NN as classifiers for the SVM.

Adaptive boosting

The AdaBoost algorithm [24] is a machine learning technique that leverages information to perform classification tasks. Given that a solitary property alone is an inadequate classifier for accurately determining the state of localization, the approach proposed involves amalgamating numerous weak classifiers that individually lack sufficient information about a specific class. This amalgamation results in a robust classifier capable of effectively identifying classes. Generally, the AdaBoost algorithm $\langle x_i y_i \rangle, 1 \leq i \leq N \ with \ y_i \in (+1, -1). x_i$ and $x_i \in \mathbb{R}^K$ considers as the vector input that has K features, which are utilized for training; and y_i considered as the required label that is either $+1$ or -1.

There are several iterations of boosting algorithms that have a same structure. The present study employs the conventional discrete Ada Boosting method, which is designed for binary classification tasks. The algorithm utilizes a collection of inputs of size N and assigns weights to each input sample, where the weight w_i is initialized as 1/N. Next, the algorithm calculates a febble classifier $f_m(x)$, the scaling factor cm, and the weighted training error ϵm. Subsequently, the weights are augmented for samples of input that have been inaccurately categorized. Following this procedure, the weights are normalized and the process of identifying a new feeble classifier is repeated M times. The conclusion of the study involves the identification of a resultant classifier, denoted as $F(x)$, which utilizes the sign of the weighted input set sum. Subsequently, this may be used to approximate the condition of the input information, specifically in our scenario, the localization condition.

The suggested methodology involves using a suitable subset of characteristics identified in the preceding section, together with the same training trajectory examples used throughout the neural network training process. Furthermore, we consider the trained NN binary classification as an extra weak classifier.

IV. CONCLUSION AND FUTURE RESEARCH

The suggested methodology for estimating the location of robots using neural networks, boosting techniques, and support vector machines exhibits encouraging outcomes. By integrating these methodologies, the localization status of the robot may be precisely categorized as either well-localized or delocalized. The procedure encompasses the selection and validation of neural network designs, the identification of advantageous features for enhancement, and the use of AdaBoost and support vector machines (SVM) as classifiers. The neural networks undergo training by using labeled training data that is created via the process of perception and particle filter-based posture estimation. The collected attributes from the collection of particles are used to enhance the precision of calculating the localization state of the robot. The AdaBoost technique is used to aggregate a collection of weak classifiers, which may include trained neural networks, in order to construct a robust classifier. SVMs are often used in order to determine the most appropriate hyperplane for the purpose of categorization. The outcomes of the neural networks' binary categorization are used as classifiers for the SVM.

The potential for future development in this study lies in the refinement and enhancement of the neural network designs and the process of selecting features. Various combinations of neural network architectures and feature sets may be investigated in order to boost the efficacy of classification. Furthermore, it is possible to assess the efficacy of alternative boosting algorithms and versions of support vector machines (SVM) in order to ascertain if they provide superior outcomes. In addition, the suggested methodology has the potential to be used in practical situations and validated using physical robot systems. The integration of the system with a robot's perception and mobility control systems enables the provision of real-time localization state estimate. The evaluation of the system's performance may be conducted across several contexts and situations in order to gauge its resilience and dependability. In general, the integration of neural networks, boosting, and support vector machines offers significant promise in achieving precise and dependable robot position prediction. Through more study and development, this particular technique has the potential to make significant contributions to the field of robotics and autonomous systems.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests.

References

[1]. D. Burghardt and P. Lanillos, "Robot localization and navigation through predictive processing using LIDAR," in Communications in computer and information science, 2021, pp. 857–864. doi: 10.1007/978-3-030-93736-2_61.

[2]. S. Angerer, C. Strassmair, M. Staehr, M. Roettenbacher, and N. Robertson, "Give me a hand — The potential of mobile assistive robots in automotive logistics and assembly applications," 2012 IEEE International Conference on Technologies for Practical Robot Applications (TePRA), Apr. 2012, doi: 10.1109/tepra.2012.6215663.

[3]. T.-W. Chong, X. Tang, C. Leng, M. Yogeswaran, O.-E. Ng, and Y. Z. Chong, "Sensor Technologies and Simultaneous Localization and Mapping (SLAM)," Procedia Computer Science, vol. 76, pp. 174–179, Jan. 2015, doi: 10.1016/j.procs.2015.12.336.
[4]. C. Cadena et al., "Past, present, and future of simultaneous localization and mapping: toward the Robust-Perception age," IEEE Transactions on Robotics, vol. 32, no. 6, pp. 1309–1332, Dec. 2016, doi: 10.1109/tro.2016.2624754.
[5]. G. Bresson, Z. Alsayed, Y. Li, and S. Glaser, "Simultaneous Localization and Mapping: A survey of current trends in Autonomous driving," IEEE Transactions on Intelligent Vehicles, vol. 2, no. 3, pp. 194–220, Sep. 2017, doi: 10.1109/tiv.2017.2749181.
[6]. M. Kutila, P. Pyykönen, W. Ritter, O. Sawade, and B. Schäufele, "Automotive LIDAR sensor development scenarios for harsh weather conditions," 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), Nov. 2016, doi: 10.1109/itsc.2016.7795565.
[7]. H. Zhang, S. Qin, J. Ma, and H. You, "Using residual resampling and sensitivity analysis to improve particle filter data assimilation accuracy," IEEE Geoscience and Remote Sensing Letters, vol. 10, no. 6, pp. 1404–1408, Nov. 2013, doi: 10.1109/lgrs.2013.2258888.
[8]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
[9]. E. Javanmardi and S. Liu, "Exploring the Human Cognitive Capacity in Understanding Systems: A Grey Systems Theory Perspective," Foundations of Science, vol. 25, no. 3, pp. 803–825, Jul. 2019, doi: 10.1007/s10699-019-09618-3.
[10]. J. Kumar, R. Goomer, and A. K. Singh, "Long Short Term Memory Recurrent Neural Network (LSTM-RNN) Based Workload Forecasting Model For Cloud Datacenters," Procedia Computer Science, vol. 125, pp. 676–682, Jan. 2018, doi: 10.1016/j.procs.2017.12.087.
[11]. S. Z. Li, Markov Random field modeling in Computer Vision. 1995. doi: 10.1007/978-4-431-66933-3.
[12]. J.-S. Kim, J. Kim, and G. Lee, "GPS data-based mobility mode inference model using long-term recurrent convolutional networks," Transportation Research Part C: Emerging Technologies, vol. 135, p. 103523, Feb. 2022, doi: 10.1016/j.trc.2021.103523.
[13]. J. J. Leonard and H. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," IEEE Transactions on Robotics and Automation, vol. 7, no. 3, pp. 376–382, Jun. 1991, doi: 10.1109/70.88147.
[14]. L. Leyrit, T. Chateau, C. Tournayre, and J.-T. Lapresté, "Association of AdaBoost and Kernel based machine learning methods for visual pedestrian recognition," 2008 IEEE Intelligent Vehicles Symposium, Jun. 2008, doi: 10.1109/ivs.2008.4621294.
[15]. S. Bouaziz, A. Tagliasacchi, and M. Pauly, "Sparse iterative closest point," Computer Graphics Forum, vol. 32, no. 5, pp. 113–123, Aug. 2013, doi: 10.1111/cgf.12178.
[16]. J. Saarinen, H. Andreasson, T. Stoyanov, and A. J. Lilienthal, "Normal distributions transform Monte-Carlo localization (NDT-MCL)," 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, Nov. 2013, doi: 10.1109/iros.2013.6696380.
[17]. C. Rizzo, T. Seco, J. Espelosín, F. Lera, and J. L. Villarroel, "An alternative approach for robot localization inside pipes using RF spatial fadings," Robotics and Autonomous Systems, vol. 136, p. 103702, Feb. 2021, doi: 10.1016/j.robot.2020.103702.
[18]. R. M. P, S. Ponnan, S. Shelly, Md. Z. Hussain, M. Ashraf, and A. Haldorai, "Autonomous navigation system based on a dynamic access control architecture for the internet of vehicles," Computers and Electrical Engineering, vol. 101, p. 108037, Jul. 2022, doi: 10.1016/j.compeleceng.2022.108037.
[19]. A. H and A. R, "Artificial Intelligence and Machine Learning for Enterprise Management," 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Nov. 2019, doi: 10.1109/icssit46314.2019.8987964.
[20]. A. Petrovskaya and A. Y. Ng, "Probabilistic mobile manipulation in dynamic environments, with application to opening doors," International Joint Conference on Artificial Intelligence, pp. 2178–2184, Jan. 2007, [Online]. Available: http://ai.stanford.edu/~ang/papers/ijcai07-openingdoors.pdf
[21]. S. N. Ghorpade, M. Zennaro, and B. S. Chaudhari, Optimal localization of internet of things nodes. Springer Nature, 2021.
[22]. D. Pao, H. F. Li, and R. Jayakumar, "Shapes recognition using the straight line Hough transform: theory and generalization," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, no. 11, pp. 1076–1089, Jan. 1992, doi: 10.1109/34.166622.
[23]. O. A. M. López, A. M. López, and J. Crossa, "Support vector machines and support vector regression," in Springer eBooks, 2022, pp. 337–378. doi: 10.1007/978-3-030-89010-0_9.
[24]. J. Zheng, D. Lin, Z. Gao, S. Wang, M. He, and J. Fan, "Deep learning assisted efficient AdABoost algorithm for breast cancer detection and early diagnosis," IEEE Access, vol. 8, pp. 96946–96954, Jan. 2020, doi: 10.1109/access.2020.2993536.
[25]. A. H, A. Ramu, and S. Murugan, "Cognitive Radio Communication and Applications for Urban Spaces," Computing and Communication Systems in Urban Development, pp. 161–183, 2019, doi: 10.1007/978-3-030-26013-2_8.