

RESEARCH

Open Access



A noise robust automatic radiolocation animal tracking system

Liang Wang^{1,2*} , Foivos Diakogiannis^{3,4}, Scott Mills⁵, Nigel Bajema⁵, Ian Atkinson⁵, Greg J. Bishop-Hurley² and Ed Charmley¹

Abstract

Agriculture is becoming increasingly reliant upon accurate data from sensor arrays, with localization an emerging application in the livestock industry. Ground-based time difference of arrival (TDoA) radio location methods have the advantage of being lightweight and exhibit higher energy efficiency than methods reliant upon Global Navigation Satellite Systems (GNSS). Such methods can employ small primary battery cells, rather than rechargeable cells, and still deliver a multi-year deployment. In this paper, we present a novel deep learning algorithm adapted from a one-dimensional U-Net implementing a convolutional neural network (CNN) model, originally developed for the task of semantic segmentation. The presented model (ResU-Net-1d) both converts TDoA sequences directly to positions and reduces positional errors introduced by sources such as multipathing. We have evaluated the model using simulated animal movements in the form of TDoA position sequences in combination with real-world distributions of TDoA error. These animal tracks were simulated at various step intervals to mimic potential TDoA transmission intervals. We compare ResU-Net-1d to a Kalman filter to evaluate the performance of our algorithm to a more traditional noise reduction approach. On average, for simulated tracks having added noise with a standard deviation of 50 m, the described approach was able to reduce localization error by between 66.3% and 73.6%. The Kalman filter only achieved a reduction of between 8.0% and 22.5%. For a scenario with larger added noise having a standard deviation of 100 m, the described approach was able to reduce average localization error by between 76.2% and 81.9%. The Kalman filter only achieved a reduction of between 31.0% and 39.1%. Results indicate that this novel 1D CNN U-Net like encoder/decoder for TDoA location error correction outperforms the Kalman filter. It is able to reduce average localization errors to between 16 and 34 m across all simulated experimental treatments while the uncorrected average TDoA error ranged from 55 to 188 m.

Keywords: Radiolocation, Machine learning, Encoder/decoder, Convolutional neural network

Introduction

The development and implementation of precision farming practices are enabled by location-aware platforms. Such platforms can track assets across holdings enabling more efficient management strategies.

Animal tracking has increasingly become an active area of both research and applied innovation. Trade-offs

exist between the types of geolocation employed including price, precision, accuracy, power consumption, and the frequency of position updates. Geolocation systems include angle of arrival (AoA), Doppler approaches, power on arrival (PoA), time of arrival (ToA) and time difference of arrival (TDoA).

The oldest forms of radio-tracking animals used AoA to triangulate the position of individuals fitted with a radio transmitter with the first publications appearing in the 1960s, such as work conducted on the early summer activities of porcupines [16]. Modern versions of this

*Correspondence: l.wang@csiro.au

² Agriculture and Food, CSIRO, 4067 Brisbane, Australia

Full list of author information is available at the end of the article



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

technique estimate the AoA by comparing the amplitude variation of an antenna array at a signal receiver and typically six or more antennas are used for this purpose. These techniques can yield very low power use depending upon the duty cycle and strength of the transmission.

Satellite tracking of wildlife has a long history starting with Craighead Jr et al. [6] tracking elk via the Nimbus meteorological satellites using a bulky 11.3 kg collar in April of 1970. However, it was the creation of the ARGOS system in 1978 [5], utilizing the Doppler shift of a carrier frequency over successive transmissions, that initiated the first generation of relatively accessible animal satellite tracking devices. Early deployments using ARGOS to track wildlife included basking sharks in June of 1982 [21] and wandering albatrosses in 1989 [11].

The ARGOS system employs a repetition period between two consecutive payloads, of between 45 and 200 s for as little as 360 ms (PPT-A3 Argos Specification), to estimate the location of a platform transmitter terminal. The accuracy of this system is based upon seven location qualities ranging from 150 m to tens of kilometers, with a low power consumption using as little as 0.8 J per location estimate (assuming two ARTIC R2 transmissions plus amplification to 0.5 W).

Power of arrival (PoA) localization methods rely upon the received signal strength at a minimum of two receivers. Implementation of PoA in LPWANs (low-power wide area networks) over long distances is not practical due to the inverse square law and signal attenuation, where the received signal strength quickly becomes too weak for the receiver to meaningfully differentiate small changes in received signal strength. These techniques are most suitable for factory-scale localization.

Time of arrival (ToA) localization methods use the time of reception of signals received by a roaming device from multiple transmitters of a known location; the most ubiquitous implementation of this technology is Global Navigation Satellite Systems (GNSS). Efficient implementations of GNSS networks, such as those using the Ublox Zoe-M8B, require around 1.8 J to acquire a location from a cold start; this signal then needs to be retransmitted adding additional energy overhead. For many animal tracking systems using GNSS, it is the single largest power drain on the system. The advantage of GNSS based systems is the lack of ground-based infrastructure; however, they require expensive space-based infrastructure. GNSS yield excellent spatial fidelity of around 2.5 m, with some systems achieving cm accuracy. Modern approaches to ToA use LPWAN to unload the on-device processing to remote services over communication

protocols such as LoRaWAN, for example, Kolmostar's¹ JEDI-200 module.

Time difference of arrival (TDoA) localization methods require no transmitter based location processing and only short transmission bursts. Locations using this method are estimated by examining the time difference of a transmitted signal arriving at multiple fixed time-synchronized receivers. In this paper, we have chosen to examine Taggle's² proprietary TDoA localization system, each location requires only 0.12 J of energy making it suitable for tracking solutions using non-rechargeable batteries, allowing for a greater number of location transmissions per energy consumed. Taggle's radio transmitter operates in a frequency range of 912–927 MHz using direct sequence spread spectrum modulation at a power output of 14 dBm. Its high capacity receivers can accommodate at least 14,000 device transmission per hour. Along with a message header, single transmissions can hold 12–15.5 bytes of user data with a transmission duration of around 300 ms. The theoretical TDoA localization accuracy of the Taggle's system is approximate:

$$\text{accuracy} = \frac{\text{speed of light}}{4 \times \text{bandwidth}} \approx 5 \text{ m.} \quad (1)$$

However, the clock synchronization of the receivers is ± 20 nanoseconds resulting in a potential error of 12 m. Menzies et al. [17] conducted a small-scale trial with Taggle's localization system using twelve Taggle tags on a plot, the size of which was approximately 5 ha. They found that the positions had a mean precision of ± 22 m with an SD of 49 m. Ground-based TDoA location systems can experience substantial noise due to multipathing where the time differences are exaggerated due to signal paths that are not line-of-sight, this is the kind of error we seek to address in this study. The examined localization method uses TDoA among four fixed receivers to estimate the origin of the transmission; the two most basic analytical methods are the Taylor series method [18] and Chan method [3] that can localize objects with minimal error in the absence of multipathing. To address noise in TDoA location estimates, machine learning-based methods are introduced for localization in an arbitrarily complex system. Current approaches to this problem fall into two main methodologies.

The first uses fingerprint references and machine learning models to derive insights about the geometrical structure of the environment which can provide information about TDoA error and improve localization accuracy.

¹ Kolmostar 48531 Warm Springs Blvd, Suite 407, Fremont, CA 94539 (www.kolmostar.com).

² Taggle Systems Pty Ltd. Sydney, N.S.W. 2000, Australia (taggle.com).

de Sousa and Thomã Electronic [7] applied the Random Forest algorithm embedded in a machine learning framework to extract a reference dataset of TDoA fingerprints in outdoor scenarios. In the experiment, four TDoA sensors were deployed in an area of 2 km² in the City of Ilmenau in Germany, representative of a typical suburban scenario with small buildings and spaced streets. The empirical cumulative density function (CDF) used in the experiment showed 210 m of error for 65% of location estimates, compared to 300 m for the raw TDoA calculations. Similarly, Alonso-González et al. [1] implemented a neural network model to estimate the positions for TDoAs using an indoor fingerprint approach to predict a transmitters positions in a 3D environment. They tested their model in a 4 × 4 × 3 m room; their experimental results indicated a substantial improvement in accuracy, with a best average error of 390 μm.

The second approach applies denoising neural networks to reduce the TDoA error to improve localization accuracy. Wu et al. [28] proposed a radial basis function (RBF) neural network to improve localization accuracy. They tested the model on simulated TDoAs within a 2D 500 × 500 m space with seven receivers, the root mean square error (RMSE) of the localization was 17 m, while the Chan algorithm leads to an ≈ 30 m RMSE. Zhang et al. [30] presented a novel localization algorithm based upon a neural network ensemble to estimate the positions of objects in indoor multipathing environments. The ensemble method was tested on the simulated TDoA in a 2D 60 × 60 cm space. The best RMSE of localization was < 1 cm and the ensemble method also showed better generalization and stability than a single neural network.

These approaches demonstrate the utility of machine learning models for localization in reducing the initial TDoA error or for correcting location estimates from noisy TDoAs. In this work, we present a novel denoising 1D convolutional neural network. The denoising encoder/decoder we propose has many similarities with a denoising autoencoder; however, the autoencoder lacks the skip connections we employ. Denoising autoencoders are an extension of simple autoencoders and were originally invented to reduce the risk of overfitting [2, 27]. Denoising autoencoders can be applied to remove the effect of stochastic noise to inputs, for example, to clean the noise from corrupted images. Convolutional layers emphasize local features of structured data, such as those evident in images or sequence data. The information from these local features help the model to reduce noise from TDoA sequences and their associated movement sequences. Diakogiannis [8] proposed a novel deep learning framework for semantic segmentation of remotely sensed data; this framework consisted of stacked CNN layers in a U-Net-like backbone [24]. We propose a

denoising encoder/decoder algorithm based on this framework with 1D CNN layers. The algorithm (ResU-net-1d) is a deep learning approach for TDoA localization error correction, using noisy TDoA tracks to correct for multipathing. The performance of this algorithm was tested on simulated animal track TDoA sequences with added noise derived from real-world data. The results show that this algorithm can recover animal tracks from noisy TDoAs. We then compared our approach with a Kalman filter to see how our approach compared to this widely employed strategy for reducing statistical noise in time series data.

The remainder of the article is organized as follows. In Sect. 2, we discuss the problem of TDoA localization in terrestrial systems. Section 3 describes the model architecture and the methodology of the experiment. Section 4 describes the animal movement simulation and the method for generating the TDoA data. The final Sect. 5 presents the performance of the developed algorithm in comparison to the non-corrected TDoA location estimates.

Problem overview and formulation

Classical TDoA localization methods assume that radio signals travel without obstruction in line-of-sight with localization solutions based on solving the following hyperbolic equations:

$$\text{TDoA}_{ij} = \text{ToA}_i - \text{ToA}_j, \quad (2)$$

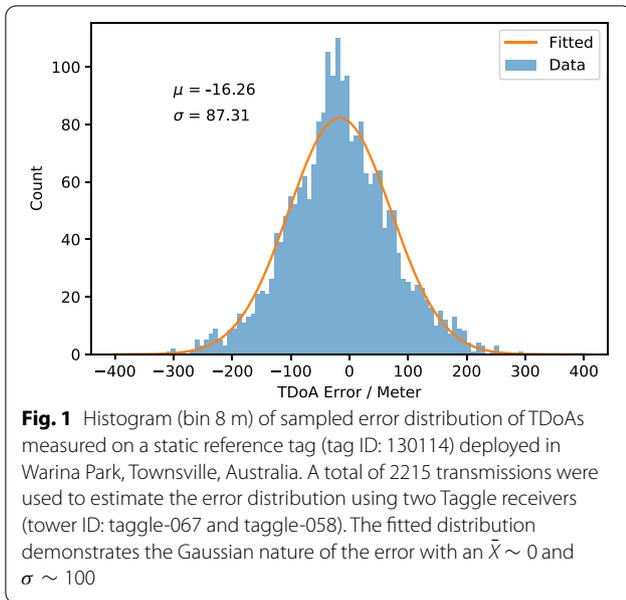
where ToA_i is the time of arrival to the i th receiver, defined as

$$\text{ToA}_i = \frac{1}{c} \sqrt{(x - X_i)^2 + (y - Y_i)^2}. \quad (3)$$

Here, X_i and Y_i are the coordinates of the i th receiver station, while x and y are the coordinates of the transmission origin track, and c is the speed of light in air.

The height of receivers and transmitters in TDoA networks can vary adding systematic error; however, in most practical cases, the introduced error is negligible. For instance, the error introduced by a 100 m elevation across 1 km is < 5 m. Of most concern is the occlusion of line-of-sight between transmitters and receivers due to vegetation, topography, or man-made objects. Such obstructions can lead to increased path length, and hence the time of arrival, between a transmitter and receiver. This multipathing phenomenon can add significant error to TDoA localization estimates.

In multipathing scenarios, signals always travel along a longer path than line-of-sight, therefore the error of ToAs is always positive, but the error for TDoAs, according to the Eq. (2) can be either positive or negative. Since



the real path of each signal traveled corresponds to a unique multipathing scenario, it is not possible to predict or model the error of TDoA from a single transmission.

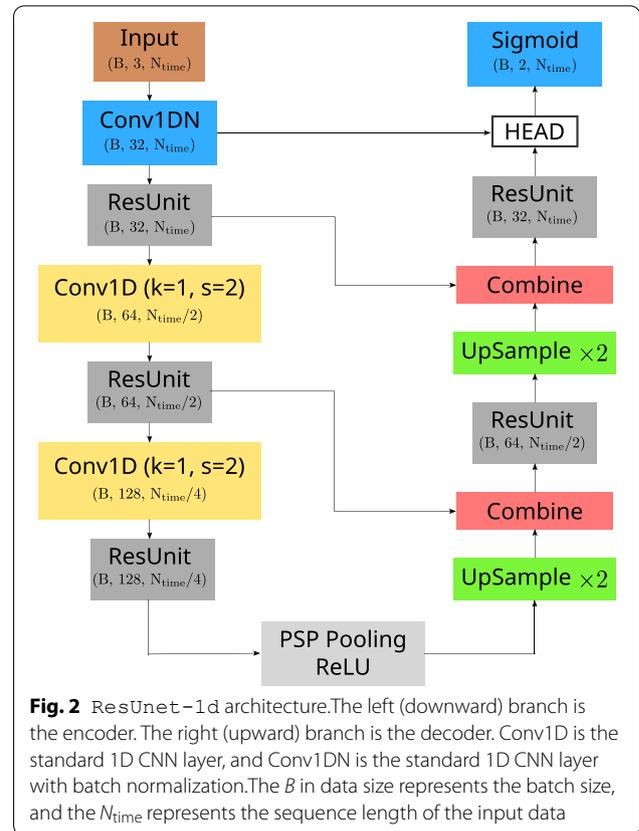
To get an estimate of a real-world TDoA error distribution, we placed a single static transmitter, tag ID is 130114, in Warina Park, Townsville Australia (Lat. -19.279, Long. 146.771) and made 2215 localization transmissions at two-minute intervals (approximately 3 days). The TDoA was examined by looking at two receivers in Townsville’s Taggle network, towers taggle-058 and taggle-067. The resulting error distribution in meters is shown in Fig. 1, using a bin size of 8 m. The noise is normally distributed and the problem can be framed as reducing Gaussian noise from TDoA measurements.

Model framework

This section gives an overview of the architecture of the model ResUnet-1d (Sect. 3.1) which reduces the localization errors of animal tracks. The ResUnet-1d model combined two tasks, converting the TDoAs to positions and localization denoising. Section 3.2 introduces the process of model training and the use of the trained model.

Architecture

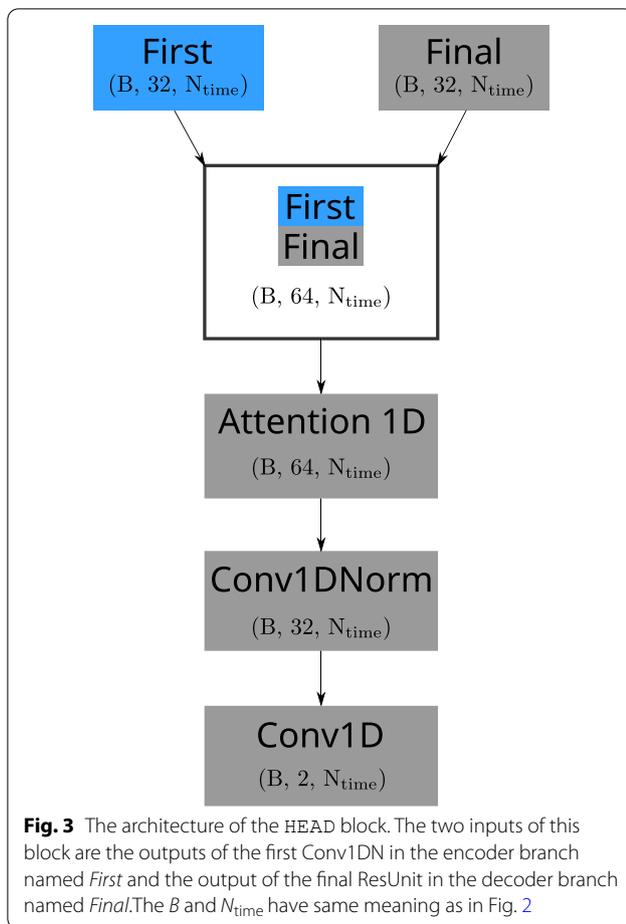
In this study, we implement a modified 1D version of the ResUnet-a model [8] that is designed for semantic segmentation of mono-temporal very high-resolution aerial images. ResUnet-a uses a UNet encoder/decoder backbone and residual building blocks with atrous convolutions for feature extraction. In the middle and at the end of the network, a pyramid scene parsing pooling layer is



implemented. The network implements a conditioned multitasking approach, estimating the semantic classes, their boundaries, and their distance transforms.

This model was chosen as it has a few advantages for the problem of TDoA positioning and localization error reduction. First, the U-Net backbone architecture is recognized in the field of computer vision for achieving state-of-art image denoising [13, 14]. Second, the residual connections [9] allow for the efficient gradient propagation in deep architectures, thus guaranteeing fast convergence and improved performance. Heinrich et al. [10] integrated ResNet into the fully convolutional neural network (FCN) with U-Net architectures for Low-Dose Computerized Tomography (CT) image denoising, showing that U-Net combined with ResNet yields the most promising result with an enhanced peak signal–noise ratio. Therefore, we have migrated this successful framework from the domain of image denoising and applied it to our TDoA animal tracking problem. The architecture of the framework is shown in Fig. 2.

In the proposed model, ResUnet-1d, we introduce several changes to the original ResUnet-a that make it suitable for our application. First, the TDoAs and the positions are multi-channel one-dimension time-series data, the 2D convolution layers in the model are replaced



by 1D convolutions. Second, compared with semantic segmentation tasks, the time series denoising tasks should be simpler in both the input data format and the difficulty of the tasks. Therefore, in ResUnet-1d, the encoder part only consists of three ResBlock—a building blocks followed by a 1D PSPPooling layer. Each feature extraction unit is a standard residual unit (we did not use atrous convolutions). This shallower model can potentially prevent overfitting issues while reducing the computational burden. Lastly, as PSPPooling does not perform well on regression problems [8], the last PSPPooling layer is replaced by an attention block which is embedded in the HEAD block for increased performance, details of this block are illustrated in Fig. 3.

Methodology

As the received TDoAs of each transmitter denote time-series data, we needed to segment the continuous time-series data into fixed-length sequences. The proposed model expects 256 records in one piece of track data. But it is a free parameter determined by the specific task, the only requirement is that the length of the training time

series tracks should match the length of the tracks on which the model will be applied.

The input data are a sequence of noised TDoAs, with the shape $256 \times N_t$, where N_t is 3, which is the number of the TDoAs at each transmission. The simulated ground truth animal tracks, that were used to generate the noisy TDoAs are the ground truth values that the algorithm is trying to recover. The shape of the output track is 256×2 , as each position only has two coordinates, x , and y . The output of the model can predict denoised tracks using noised TDoAs as an input, hence reducing TDoA multipathing error.

In principle, the proposed model could be trained by real-world animal tracks and TDoAs collected from the paddock. However, in practise, training of the deep learning model requires thousands of data samples, which may not be feasible for most real-world applications of this problem. Furthermore, it is unlikely that an animal will be fitted with both a GNSS and TDoA tracking system outside of a research setting. We, therefore, propose the application of this model be proceeded by two initial steps, deployment of static tags to estimate the TDoA error distribution followed by the simulation of TDoAs tracks. We will introduce the simulation methods in the next section. The simulated data can be used to train the model and the trained model will be able to reduce the localization error.

Kalman filter

We compared our results against those obtained with a Kalman filter to observe the efficacy of our approach to this widely employed method for noise reduction in time series data. A Kalman filter is a recursive algorithm to estimate the state of a dynamic system having certain types of random behavior and demonstrates the capability of noise reduction [16]. We applied this approach to reduce localization noise.

The Kalman filter describes the system by the state vector $x_t^T = (x, y, v_x, v_y)$ and updates the state vector and error covariance matrix P_t in each iteration. To update the state vector and error covariance matrix, we need the state transmission matrix F , measurement noise covariance matrix Q , system noise covariance matrix R , and the measurement vector z_t . In this case, the state transmission matrix is:

$$F = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4)$$

The measurement noise covariance matrix Q and system noise covariance matrix R are estimated from the dataset. The measurement vector z_t is the localization position

estimated from the raw TDoAs at each time t . The update of the state vector and the error covariance matrix is given by:

$$\bar{x}_t = Fx_{t-1}, \quad (5)$$

$$\bar{P}_t = FP_{t-1}F^T + Q, \quad (6)$$

$$K_t = \bar{P}_tH^T(H\bar{P}_tH^T + R)^{-1}, \quad (7)$$

$$x_t = \bar{x}_t + K_t(z_t - H\bar{x}_t), \quad (8)$$

$$P_t = (I - K_tH)\bar{P}_t, \quad (9)$$

where I is the identity matrix, \bar{x}_t is the priori estimate of the state vector, and K_t is the Kalman filter gain. In each time step, the Kalman filter estimates the state vector by combining the prior estimate of the state vector and the measurement vector with the Kalman filter gain. In this work, we implement the Kalman filter using an open source Kalman filter module, `pykalman`, on Github.

Data simulation and preprocessing

Within the field of animal movement behavior, the modeling of movement data is implemented in many ways. Quaglietta and Porto [22] introduced an algorithm, `SimRiv`, to simulate individual-based, spatially explicit movements in river networks and heterogeneous landscapes. In this study, we simulate a cows' movement using this approach on a totally homogeneous landscape.

Animal track modeling

Animal movements are considered to be Brownian motion and multistate. The main states of the movement include random walking, correlated random walking, and rest. The random walk state is a random movement state, in which the direction of the steps is completely independent. The correlated random walk means the direction taken in one step by an individual animal should be correlated with the direction of the previous step [23, 26]. The correlation, which is the turning angle concentration parameter of the wrapped normal distribution, has a value between $[0, 1]$, where 0 means there is no correlation between two steps (yielding a random walk state), and 1 means the direction does not change. In this study, we use 0.98 (as chosen by Quaglietta and Porto [22]) as the value of correlation. The resting state corresponds to a state where the individual animal does not change position.

Following Quaglietta and Porto [22], we assume the cows are Lévy-like walkers who alternate between random walks and correlated random walks. This multi-state

movement simulation required specifying the probabilities of transition between the random walk state and correlated random walk state [19]. We used a transition matrix to define the probabilities of transition between states. The transition matrix is a square matrix where all values are probabilities, and the element at row i column j defines the probability of the individual changing from state i to state j . The transition matrix in this study is as same as the example in Quaglietta and Porto [22], which is

$$\begin{pmatrix} 0.995 & 0.005 \\ 0.01 & 0.99 \end{pmatrix}. \quad (10)$$

The step length, in meters, for both states are set to a random number from the uniform distribution $U(0, 1)$.

TDoA simulation

In the simulation, the coordinates of the receivers are: $(-2000, 2000)$ m, $(-2000, -2000)$ m, $(2000, -2000)$ m, and $(2000, 2000)$ m, an area of 16,000 ha. To mimic real-world animal movements at different time scales the simulations recorded the positions of the track at different step intervals, N_s . We converted the recorded positions into ToA using Eq. (3). The TDoAs were obtained by the substitution of two ToAs using Eq. (2). The input TDoAs in our model are TDoA₁₂, TDoA₁₃, and TDoA₁₄. For computational simplicity, we multiplied the derived times by the speed of light, c for all the input TDoAs to get distances. We added Gaussian noise $N(0, \sigma)$, based upon the static tag observation, into the simulated TDoAs to generate noised TDoAs, where σ is the standard deviation of the TDoA error distribution. In this work, we evaluate our model's ability to reduce noise for TDoAs exhibiting error standard deviations of either 50 m or 100 m.

Training data simulation and preprocessing

For the training simulation dataset, we first generated multiple animal movement tracks within a virtual paddock. Each track started from a random position within the paddock and picked a direction for the first step with equal probabilities. The subsequent steps were generated with the method discussed in Sect. 4. As discussed in Sect. 4.1, we assumed the device recorded the position of the animal every N_s steps. The number of steps in each simulated track sequence was $N_s \times 256$. We then down-sample the track by extracting the first position in every N_s steps. The down-sampling mimics receivers only recording the animal's positions at set time intervals, the length of all tracks after down-sampling is 256. We saved these down-sampled tracks as the target outputs of the model. When the tracks were generated, we used the method detailed in Sect. 4.1 to calculate the TDoA

sequence of each track and then add random Gaussian noise.

The resultant noised TDoAs (model input) and the corresponding ground truth positions (model output) pairs were split into a training dataset and a testing dataset in the fraction of 8 : 2. Meeting the requirements of the ResUnet-1d, the values of the model inputs and the model outputs should be in the range of [0, 1], we re-scaled the input and output data in both training and testing datasets through min-max normalization. As all data used in this work are simulated, we do not need to tackle the issue of missing. However, the occurrence of missing data in time-series data is very common, we will discuss this issue in Sect. 5.3.

Evaluation

In this section, the predictive accuracy of the ResUnet-1d is evaluated by using simulated animal movement data described in the previous section. The localization results from our model are compared with the results of the analytical method implemented in Menzies et al. [17].

Design of experiment

In this work, we implemented the ResUnet-1d model to reduce localization error, the main source of this error in real-world TDoA deployments is multipathing. We evaluated our model on the simulated Lévy-like tracks with different recording steps N_s and two different TDoA error standard deviations σ . We aimed to investigate if ResUnet-1d could reduce the localization error significantly and how the ResUnet-1d model performance varied with step interval N_s and the original TDoA error's standard deviation σ .

As observed from the static tag in Townsville, the error distribution of TDoAs is Gaussian, and the standard deviation of the error in the range of ~ 100 m. We use this value to simulate a urban-like environment. However, we suspect that the urban environment has increased multipathing issues due to large metallic moving objects, such as vehicles, and other effects of the built environment. In more remote agricultural areas, these aforementioned obstructions are greatly reduced, thus we can hypothesize that the standard deviation of the TDoA error in these locations is likely to be lower. To mimic this expectation, we chose to halve the standard deviation of the error to 50 m to emulate a more rural setting. The choice of this value for the SD is corroborated by Menzies et al. [17].

In our simulation, the step size was selected from a uniform distribution $U(0, 1)$ m to mimic continuous movement. As discussed in Sect. 4.2, we down sampled the track positions by recording only one position in every N_s positions, where $N_s \in [10, 20, 40, 60, 80, 100, 200, 300, 500]$

Table 1 Hyperparameters of the ResUnet-1d

Hyperparameters	Value
Depth	3
Number of filters	32
Batch size	256
Learning rate	10^{-4}
Optimizer	Adam
Loss Function	L1Loss

which is equivalent to a time interval ranging from 0.6 – 30 min when considering the average speed of a cow [20, 25]. The values of this time interval would be much higher for grazing individuals.

The optimized hyperparameters of ResUnet-1d are summarized in Table 1. For all models, we used the Adam [12] optimizer, with a learning rate of 10^{-4} . We chose the L1Loss loss function to obtain the best training performance for ResUnet-1d. Our model was built and trained using the MXNet deep learning library [4], under the GLUON API. Each of the models was trained on 8000 simulated tracks with a batch size of 256 on a single NVIDIA Tesla P100 GPU using the CSIRO's HPC facilities. We used 2000 simulated tracks to test the performance of our model.

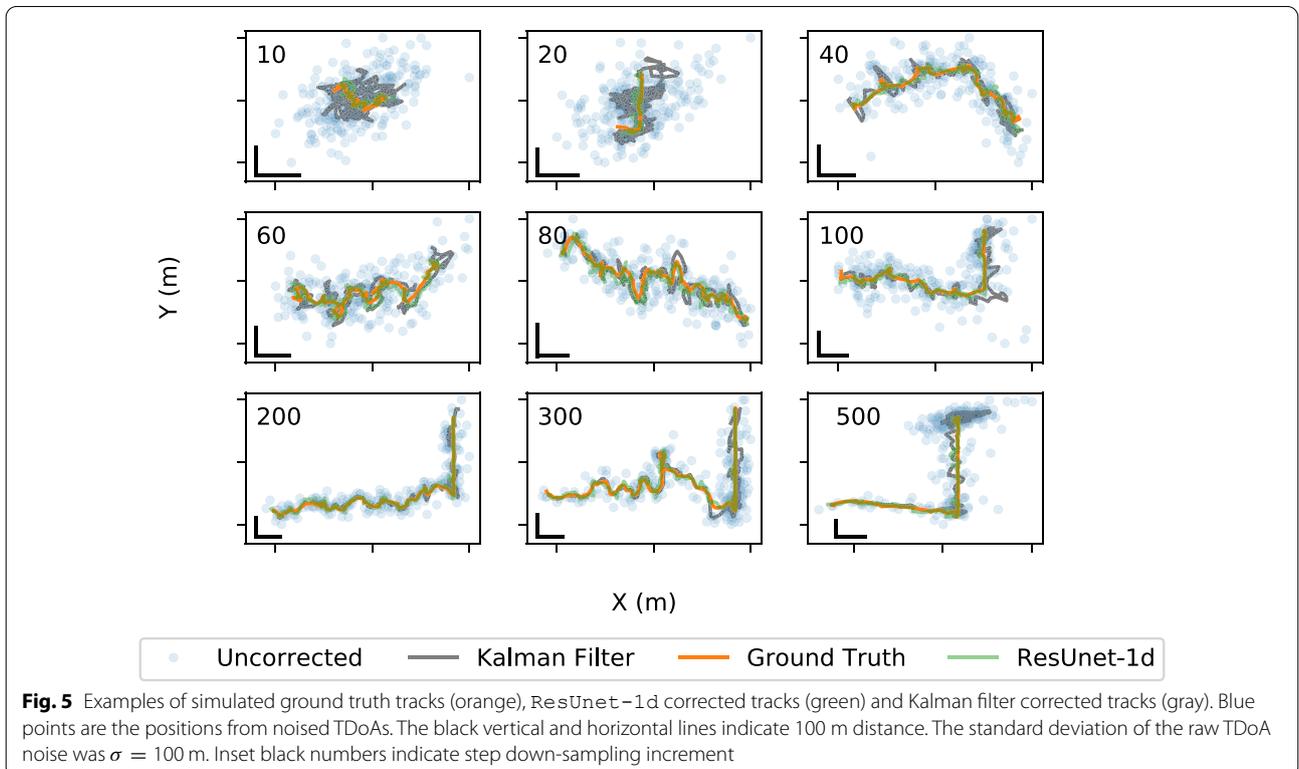
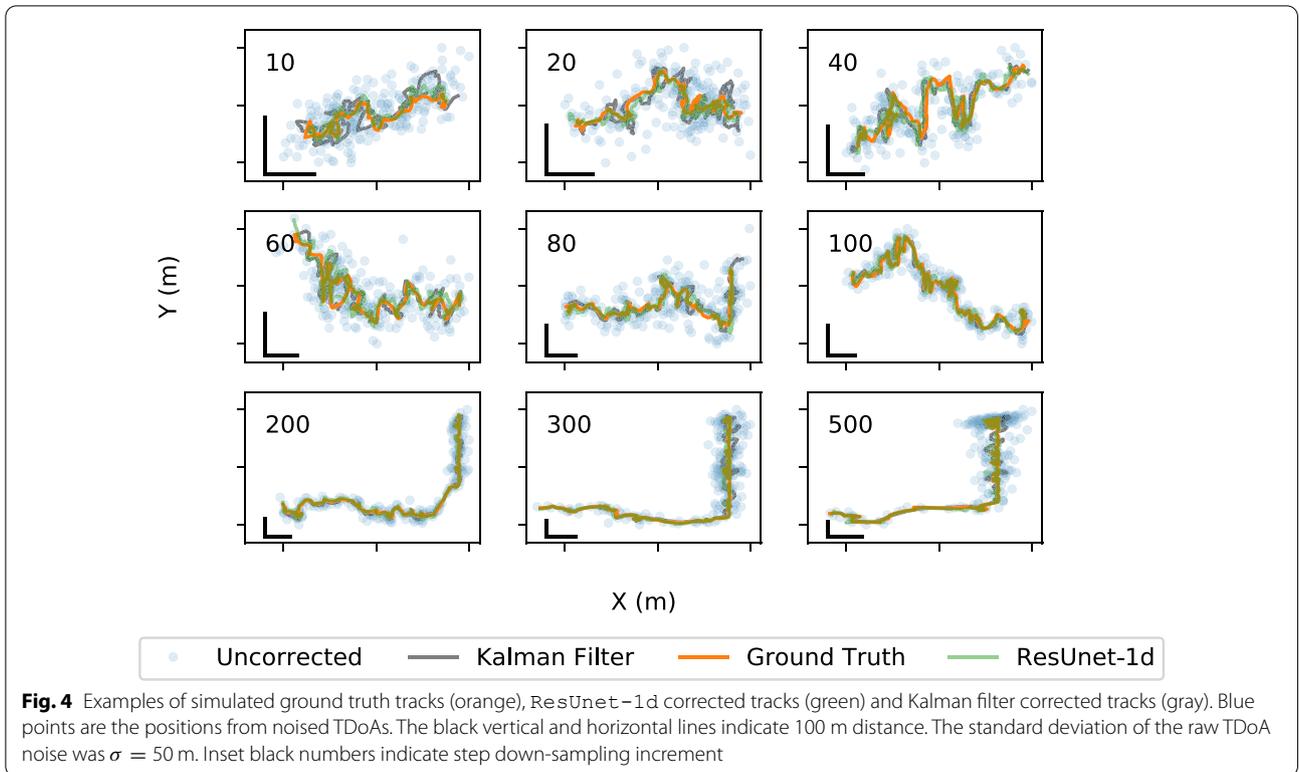
Performance of ResUnet-1d

Figures 4 and 5 illustrate simulated tracks, the orange lines are the ground truth movement track generated from the animal movement simulations and the faded blue points are the measured positions calculated with noised TDoAs. The green lines represent the recovered tracks by the ResUnet-1d model, they reproduce the shape of the ground truth tracks and recovered most of their features. The gray lines are the tracks recovered by the Kalman filter and also exhibit noise reduction. However, the tracks recovered by the Kalman filter have larger errors than those recovered by our ResUnet-1d model.

To evaluate the performance of the model quantitatively, we measure the root mean square error (RMSE) of noised tracks and recovered tracks to the ground truth tracks in the following way

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x})^2 + (y_i - \hat{y})^2}, \tag{11}$$

where the (x, y) represents the position of noised tracks or recovered tracks, and (\hat{x}, \hat{y}) represents the position of ground truth tracks. The probability density function (PDF) of the errors is illustrated in Figs. 6 and 7. In both figures, the distribution of the error of the ResUnet-1d



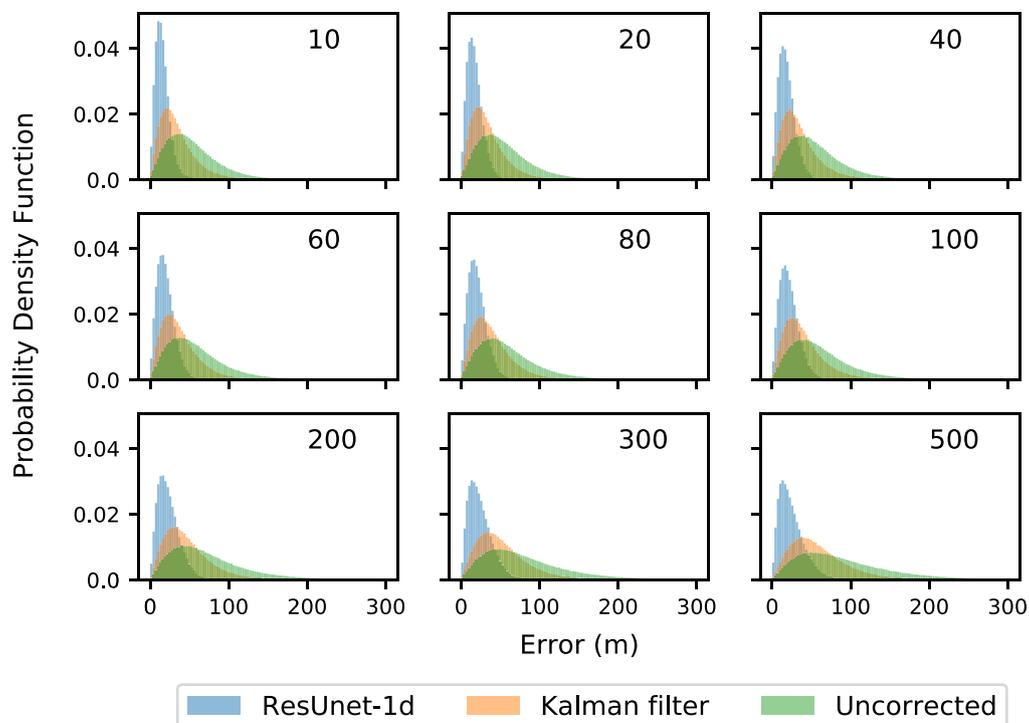


Fig. 6 Probability density function for localization errors after noise with a standard deviation of 50 m had been applied to simulated tracks. Results are displayed for uncorrected TDOAs, ResUnet-1d recovered tracks and Kalman filter recovered tracks. Inset black numbers indicate step down-sampling increment

recovered tracks is significantly narrowed and of a lower value compared with the distribution of the original error of the noised tracks. Comparison of the distribution of the error of the ResUnet-1d recovered tracks and Kalman filter recovered tracks confirm what Figs. 4 and 5 illustrate, the distribution of the error from the tracks recovered by ResUnet-1d are clearly narrower and lower than those from Kalman filter. Therefore, we demonstrate that our ResUnet-1d model can effectively reduce localization error introduced by random Gaussian errors in TDoA position estimates, and can outperform a Kalman filter.

Figure 8 compares the RMSE of uncorrected and corrected tracks from the two simulations with different noise levels. The RMSE of the two recovered tracks by ResUnet-1d are significantly lower than the RMSE of the two noised tracks and the tracks recovered by Kalman filter. When the original localization error was lower, orange lines, the efficacy of the recovered tracks dropped slightly. On average, for simulated tracks having added noise with a standard deviation of 50 m, the ResUnet-1d approach was able to reduce localization error by between 66.3% and 73.6%. The Kalman filter only achieved a reduction of between 8.0% and 22.5%. For a scenario with larger added noise having a standard

deviation of 100 m the ResUnet-1d approach was able to reduce average localization error by between 76.2% and 81.9%. The Kalman filter only achieved a reduction of between 31.0% and 39.1%.

Results indicate that ResUnet-1d is able to reduce average localization errors to between 16 and 34 m across all simulated experimental treatments while the corresponding uncorrected average TDoA location error ranged from 55 to 188 m, and in the case of the Kalman filter between 48 and 115 m. Our ResUnet-1d approach was robust to the down sampling (step) interval; however, the Kalman filter error correction tended to improve at greater step intervals but never approached the levels observed by ResUnet-1d. On average across the two noise treatments ResUnet-1d exhibited a 54% (noise $\sigma = 50$ m) and 44% (noise $\sigma = 100$ m) difference in localization error correction.

Discussion

Conventional analytic methods for calculating locations based upon TDOAs work well in the absence of multipathing; in real-world settings, multipathing degrades the performance of these systems. Previous machine learning methods have shown a reduction in localization noise but have only considered a single static position.

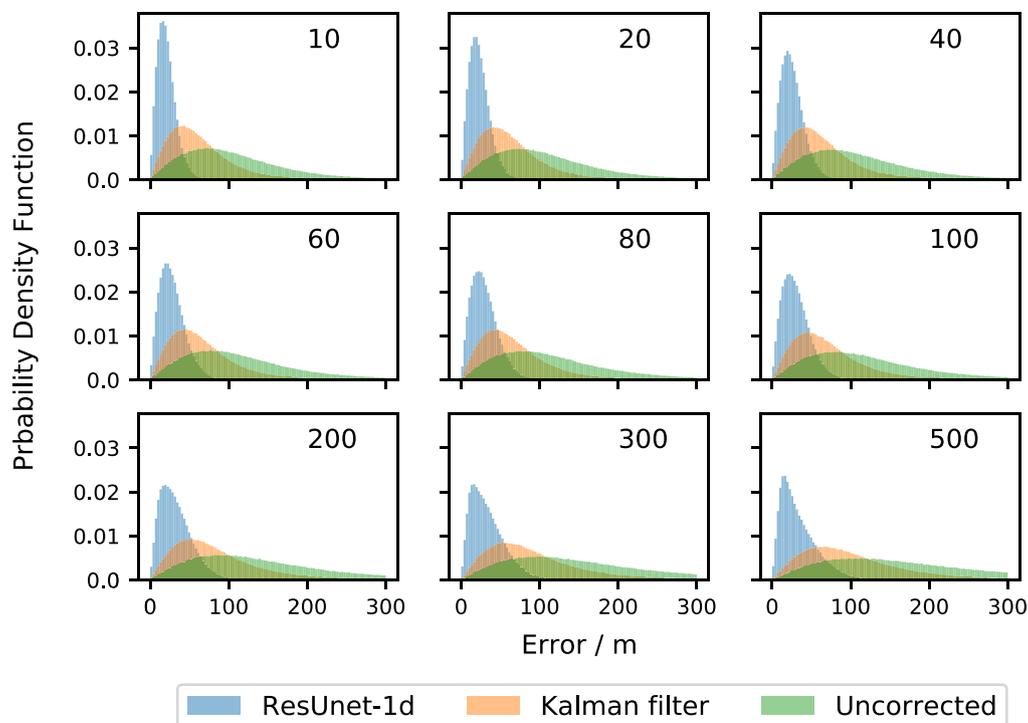


Fig. 7 Probability density function for localization errors after noise with a standard deviation of 100 m had been applied to simulated tracks. Results are displayed for uncorrected TDOAs, ResUnet-1d recovered tracks and Kalman filter recovered tracks. Inset black numbers indicate step down-sampling increment

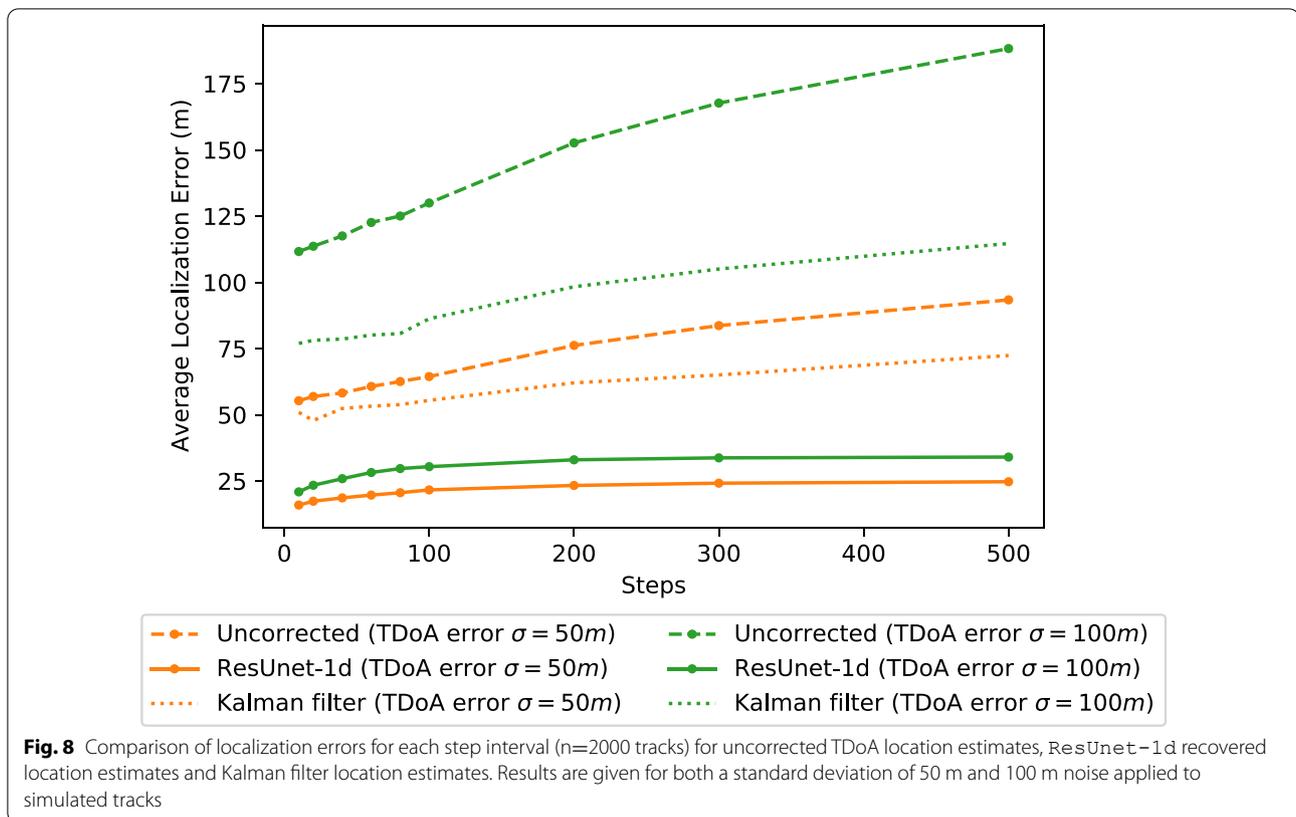
Due to the nature of track localization, the positions of neighboring points have information which can help contribute to a reduction in localization error. Therefore, a model which can combine information about neighboring points to make an enhanced prediction should provide a better solution to this problem. A Kalman filter is a conventional method to reduce noise in just such a scenario. The method proposed in this work implements a CNN layer to consider the connection between neighboring positions for animal movement tracks. The combination of the CNN layer and U-Net like encoder/decoder architecture enhances the ability of the model to reduce noise in the data. The complexity of the model we proposed here simultaneously considers and extracts hidden patterns from the random noise over a long portion of the track. A Kalman filter only looks at the position before the given position and assumes a linear function to make the prediction. By testing this algorithm on a correlated random walk simulation, we have made the assumption of a homogeneous paddock environment using SimRiv. It is recognized that SimRiv does not take into account the impacts of a heterogeneous landscape, where the willingness of an animal to cross a specific environment is not taken into account. We anticipate a

heterogeneous environment, such as creek lines, would elicit less random animal movements resulting in a higher correlation between neighboring positions. This kind of correlation should translate to increased performance when compared to animal tracks simulated under the assumption of a homogeneous environment.

The use of this method requires a sequence of the TDOAs without missing data. Missing data are likely to be common in TDoA localization networks. In real deployments, the signal transmission can not only be reflected by objects but also blocked by them. Moreover, network communication outages along with transmitter maintenance or failure will also lead to data gaps.

It would be possible to implement a simple interpolation preprocessing step to the proposed model to recover missing data. Two promising methods for enhance the interpolation are the implementation of a momentum term or a Taylor approximation. Zhang et al. [29] proposed a new sequence-to-sequence imputation model for recovering missing data in wireless sensor networks. This method could also be employed in the data preprocessing pipeline to address missing data.

We acknowledge that the simulated data in this paper represent an oversimplification of this problem; we expect real-world deployments to pose challenges such as



shifting error means and standard deviations for TDoA data spatially. This paper does, however, demonstrate the potential utility for this approach. We intend to take this simulated model and apply it to a working cattle station to see if the demonstrated gains hold true for a real-world scenario. This will include the deployment of multiple static nodes that will inform the TDoA stochasticity both spatially and temporally.

Conclusion

In this paper, we developed and investigated a 1D CNN-based U-Net like encoder/decoder model for denoising TDoA position estimates for animal tracking using simulated animal movement data (ResUnet-1d). We have demonstrated that our model can successfully recover simulated animal movement tracks from noised TDoA sequences, and reduce localization error by between 66.3% and 81.9%. As the results for ResUnet-1d tracks with different step intervals do not show a clear trend, it would be possible the algorithm to be implemented for animal tracks constructed from lower frequency transmissions, at down-sampling intervals greater than 500 m. Our model outperforms a Kalman filter for this TDoA noise reduction problem

and is more robust to changes in the down-sampling interval. These findings need to be validated in a working cattle station in conjunction with an assessment of missing data preprocessing.

Acknowledgements

The authors acknowledge the support of the Scientific Computing team of CSIRO. We would also like to acknowledge Gordon Foyster and Richard Keaney for their useful discussions and comments on this manuscript.

Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Author Contributions

LW and FD conceived of the presented idea. LW, FD, SM, and NB planned the experiments. LW carried out data simulation, deep learning model development, training and data interpretation. FD supported the development of deep learning model architecture. LW took the lead in writing the manuscript in consultation with FD, SM and NB. All authors discussed the results and contributed to the final manuscript. All authors read and approved the final manuscript.

Funding

This document is the result of a research project funded by Advance Queensland Innovation Partnerships (AQIP)—Smart Ear Tag for Livestock, 2016.

Availability of data and materials

The datasets during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Agriculture and Food, CSIRO, 4814 Townsville, Australia. ²Agriculture and Food, CSIRO, 4067 Brisbane, Australia. ³ICRAR, The University of Western Australia, 6009 Perth, Australia. ⁴Data61, CSIRO, 6014 Perth, Australia. ⁵eResearch Centre, James Cook University, 4814 Townsville, Australia.

Received: 10 February 2021 Accepted: 18 June 2021

Published online: 01 August 2021

References

- Alonso-González I, Sánchez-Rodríguez D, Ley-Bosch C, Quintana-Suárez MA. Discrete indoor three-dimensional localization system based on neural networks using visible light communication. *Sensors*. 2018;18(4):1040. <https://doi.org/10.3390/s18041040>.
- Bengio Y, Delalleau O. Justifying and generalizing contrastive divergence. *Neural Computat*. 2009;21(6):1601–21. <https://doi.org/10.1162/neco.2008.11-07-647>.
- Chan YT, Ho K. Joint time-scale and tdoa estimation: analysis and fast approximation. *IEEE Trans Signal Process*. 2005;53(8):2625–34. <https://doi.org/10.1109/TSP.2005.850336>.
- Chen T, Li M, Li Y, Lin M, Wang N, Wang M, Xiao T, Xu B, Zhang C, Zhang Z. Mxnet: a flexible and efficient machine learning library for heterogeneous distributed systems. 2015. arXiv preprint [arXiv:1512.01274v1](https://arxiv.org/abs/1512.01274v1)
- Clark DD. Overview of the argos system. *Proc Oceans*. 1989;3:934–9. <https://doi.org/10.1109/OCEANS.1989.586711>.
- Craighead Jr FC, Craighead JJ, Cote CE, Buechner HK. Satellite and ground radiotracking of elk. NASA, Washington Animal Orientation and Navigation. 1972. <https://ntrs.nasa.gov/citations/19720017418>.
- de Sousa MN, Thomä RS. Enhanced localization systems with multipath fingerprints and machine learning. In: 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), IEEE, 2019; pp 1–6. <https://doi.org/10.1109/PIMRC.2019.8904120>.
- Diakogiannis F, Waldner F, Caccetta P, Wu C. Resunet-a: a deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J Photogram Remote Sens*. 2020;162:94–114. <https://doi.org/10.1016/j.isprsjprs.2020.01.013>.
- He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016; pp 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
- Heinrich MP, Stille M, Buzug TM. Residual u-net convolutional neural network architecture for low-dose ct denoising. *Curr Direct Biomed Eng*. 2018;4(1):297–300. <https://doi.org/10.1515/cdbme-2018-0072>.
- Jouventin P, Weimerskirch H. Satellite tracking of wandering albatrosses. *Nature*. 1990;343(6260):746–8. <https://doi.org/10.1038/343746a0>.
- Kingma DP, Ba J. Adam: a method for stochastic optimization. 2014. arXiv preprint [arXiv:1412.6980v9](https://arxiv.org/abs/1412.6980v9).
- Komatsu R, Gonsalves T. Effectiveness of u-net in denoising rgb images. *Comput Sci Inf Techn*. 2019; pp 1–10. <https://doi.org/10.5121/csit.2019.9201>.
- Liu D, Wen B, Liu X, Wang Z, Huang TS. When image denoising meets high-level vision tasks: a deep learning approach. 2017. arXiv preprint [arXiv:1706.04284](https://arxiv.org/abs/1706.04284)<https://doi.org/10.24963/ijcai.2018/117>.
- Marselli C, Daudet D, Amann HP, Pellandini F. Application of kalman filtering to noisereduction on microsensor signals. In: Proceedings du Colloque interdisciplinaire en instrumentation, C2I, 18-19 novembre 98, pp. 443-450, Ecole Normale Supérieure de Cachan, France, pp 443–450. 1998.
- Marshall WH, Gullion GW, Schwab RG. Early summer activities of porcupines as determined by radio-positioning techniques. *J Wildlife Manage*. 1962;26(1):75–9. <https://doi.org/10.2307/3798169>.
- Menzies D, Patison KP, Fox DR, Swain DL. A scoping study to assess the precision of an automated radiolocation animal tracking system. *Comput Electron Agricult*. 2016;124:175–83. <https://doi.org/10.1016/j.compag.2016.04.001>.
- Mo J, Deng Z, Jia B, Bian X. A pseudorange measurement scheme based on snapshot for base station positioning receivers. *Sensors*. 2017;17(12):2783. <https://doi.org/10.3390/s17122783>.
- Morales JM, Haydon DT, Frair J, Holsinger KE, Fryxell JM. Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*. 2004;85(9):2436–45. <https://doi.org/10.1890/03-0269>.
- O'Driscoll K, Schutz MM, Lossie A, Eicher S. The effect of floor surface on dairy cow immune function and locomotion score. *J Dairy Sci*. 2009;92(9):4249–61. <https://doi.org/10.3168/jds.2008-1906>.
- Priede I. A basking shark (*Cetorhinus maximus*) tracked by satellite together with simultaneous remote sensing. *Fish Res*. 1984;2:201–16. [https://doi.org/10.1016/0165-7836\(84\)90003-1](https://doi.org/10.1016/0165-7836(84)90003-1).
- Quaglietta L, Porto M. Simriv: an r package for mechanistic simulation of individual, spatially-explicit multistate movements in rivers, heterogeneous and homogeneous spaces incorporating landscape bias. *Movement Ecology*. 2019;7. <https://doi.org/10.1186/s40462-019-0154-8>.
- Reynolds AM. Mussels realize weierstrassian lévy walks as composite correlated random walks. *Sci Rep*. 2014;4:4409. <https://doi.org/10.1038/srep04409>.
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015; p 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.
- Swain D, Wark T, Bishop-Hurley G. Using high fix rate gps data to determine the relationships between fix rate, prediction errors and patch selection. *Ecol Model*. 2008;212(3):273–9. <https://doi.org/10.1016/j.ecolmodel.2007.10.027>.
- Turchin P. Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants. Sinauer Associates. 1998. <https://doi.org/10.1086/393125>.
- Vincent P, Larochelle H, Bengio Y, Manzagol PA. Extracting and composing robust features with denoising autoencoders. In: Proceedings of the 25th international conference on Machine learning, 2008;p 1096–1103. <https://doi.org/10.1145/1390156.1390294>.
- Wu L, Li F, Chen S. A improved wireless location algorithm in nlos environment. *Inform Technol J*. 2013;12(24):8563. <https://doi.org/10.3923/itj.2013.8563.8569>.
- Zhang YF, Thorburn PJ, Xiang W, Fitch P. Ssim—a deep learning approach for recovering missing time series sensor data. *IEEE Internet Things J*. 2019;6(4):6618–28. <https://doi.org/10.1109/JIOT.2019.2909038>.
- Zhang Z, Jiang F, Li B, Zhang B. A novel time difference of arrival localization algorithm using a neural network ensemble model. *Int J Distrib Sensor Netw*. 2018;14(11):1550147718815798. <https://doi.org/10.1177/1550147718815798>.