Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2018) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

The 3rd International workshop on Internet of Things: Scope and Opportunities (IOTSO2022)
October 26-28, 2022, Leuven, Belgium

LSTM Filter for Smart Agriculture

Junwhan Kim[a], Byunggu Yu[a], Sabine O'Hara[b]

[a]*Computer Science and Information Technology, University of the District of Columbia, Washington, DC, USA*
[b]*College of Agriculture, Urban Sustainability and Environnmental Sciences, University of the District of Columbia, DC, USA*

Abstract

Enormous amounts of data are generated each day by sensor devices. In agriculture, these devices continuously monitor numerous environmental properties in the fields of aquaponics, hydroponics, and soil-based food production. Data stream mining is the process of extracting data from continuous, rapidly sampled data sources. The data accuracy that can be achieved in data stream mining is highly dependent on the algorithm chosen to suppress noise. For threshold-based automation, an actuator can be activated when the value of sensor data is above a permissible threshold. Noise from sensors may activate the actuator. Several statistical and machine learning-based noise- suppression algorithms have been proposed in the literature. The proposed LSTM (Long Short-Term Memory) filter performs better noise suppression than other traditional filters – Kalman and moving average filters. The LSTM filter is installed in our threshold-based aquaponics automation to maximize sustainable food production at minimum cost.

Keywords: LSTM; Smart Agriculture; IoT Sensor

1. INTRODUCTION

Aquaponics is a field that combines aquaculture with hydroponics and creates systems in which nutrients from a recirculating fish rearing system serve as support for plant culture. Urban aquaponics research will generate novel and innovative contributions toward solving the crisis of urban food insecurity, which remains one of our society's greatest challenges. About 37% of households with children in the District of Columbia are currently unable to afford to purchase enough food to feed their families [1]. The main contribution of urban aquaponics research is to find ways to improve the state of nutrition-associated disorders, including obesity, diabetes, and hypertension. These con-

* Corresponding author.
E-mail address: junwhan.kim@udc.edu (J. Kim), byu@udc.edu (B. Yu), sabine.ohara@udc.edu (S. O'Hara)

ditions have reached pandemic levels in the US, and are particularly prevalent among those living in urban settings and neighborhoods known as "food deserts" that have few to no grocery stores [2]. The research into urban aquaponics automation will also address concerns related to the food miles currently required to bring resources to urban populations, an issue that is directly related to environmental pollution [3]. Eleven percent of the greenhouse gas emissions associated with the US food supply chain are related to transportation [4]; global estimates suggest that agriculture is responsible for 25% of all CO2, 65% of all methane, and 90% of all nitrous oxide emissions [5].

As smart farming, farming automation is increasing. However, farming requires agriculture knowledge and hands-on experiences. Specifically, knowledge and experience for both animals (i.e., fish) and plants are mandated to maintain aquaponics. Due to such labor and knowledge intensive work, aquaponics business may be less profitable than hydroponics and aquaculture [7]. The techniques of machine learning identify the patterns and relationship that may otherwise be hidden. Artificial intelligence will learn the practical knowledge of animal and plant experts, minimizing maintenance cost.

To maximize sustainable food production at minimum cost, we have developed next-generation aquaponics automation based on the most recent neural network models, discussed in Section 3. The Internet-of-Things (IoT) is begining to impact a wide arrary of sectors and industires including agriculture to reduce inefficiencies and to improve the performance [12]. Our aquaponics automation based on the capabilities offered by IoT consists of agrecultural sensors devices, actuators, wireless modules, and data management.

Enormous amounts of data are generated each day by the sensor devices. These devices continuously monitor numerous environmental properties of the aquaponics system. Data stream mining is the process of extracting data from continuous, rapidly-sampled data sources. The data accuracy that can be achieved in data stream mining is highly dependent on the algorithm chosen to suppress noise. Several statistical and machine learning-based noise-suppression algorithms have been proposed in the literature [11]. However, the longer-term dependencies of the data streams from agriculture sensors have not been considered. Motivated by the effectiveness of machine learning on time series data, we propose a new machine learning-based noise suppression approach, called an LSTM (Long Short-Term Memory) filter.

The main contribution is that an LSTM model is applied for noise suppression of agriculture sensors in a large-scale aquaponics system. We show that the LSTM filter performs better than Kalman and moving average filters.

The rest of this paper is organized as follows. We overview related efforts in Section 2. In Sections 3 and 4, we explain our proposed method and empirical results, respectively. The paper is concluded with Section 5.

2. RELATED WORK

Agricultural productivity has significantly increased through computers and sensors-based automation [9]. Many methods have been proposed to minimize sensor specific noise, but these rely on the type of data. This section presents existing machine learning-based filtering methods for different data types.

Jiang et. al. proposed LSTM based adaptive filtering [8] by learning the weight variations from the weight sequence directly. The trained networks are used to regulate the weights generated by conventional filtering schemes for reduced prediction error of hyperspectral images. This filter models not only the correlations between pixels from different spectral bands, but also the temporal dependencies of the filtering weights. This approach has not been evaluated with data stream from real sensors.

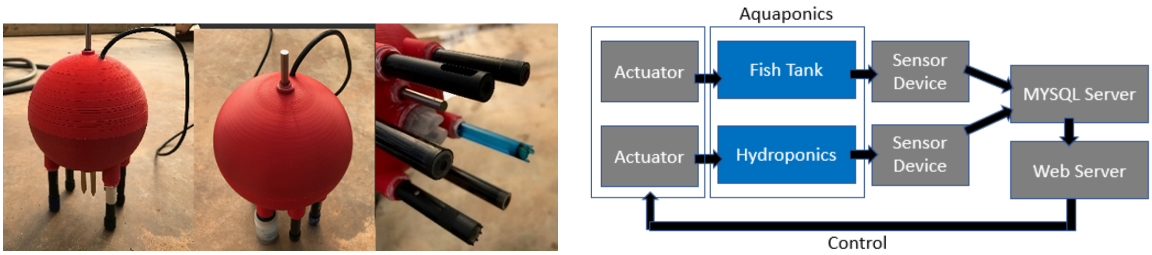
As another recent research activity, a new target tracking filter using a deep learning has been proposed by Cui et. al. [10]. The filter is developed in a unified neural network including feedforward, recurrent, and attention neural network structures to resolve tracking problems. The proposed filter has been simulated to estimate a target state.

Yu and Kim proposed an LSTM anomaly filter [11] performing better and detects anomalies more effectively than can be achieved with other statistical filters. Anomaly detection is an approach that can be used to profile the normal runtime behavior of computer programs and thus to detect intrusions and errors as anomalous deviations from the observed normal baseline. However, normal but unobserved behavior can trigger false positives. The LSTM anomaly filter detects anomalies with zero false positives when filtering noisy behavior observations in an unsupervised manner.

To the best of our knowledge, none of machine learning based filters are applied for agriculture sensors and large scale aquaponics systems.

3. LSTM FILTER IN AQUAPONICS

We have developed two sensor devices, including one for aquaculture and one for hydroponics. The aquaponics system has been controlled based on the data collected from these devices for sustainable agriculture automation. The major goal is to develop aquaponics automation technology that will revolutionize the sustainable production of fresh natural foods in and/or near urban centers in the US. The secondary goal is to determine how to apply this automation technology for use by a broad spectrum of sustainable food production sectors, including aquaculture, hydroponics, and traditional farming. This automation technology has been designed to maximize sustainable food production at minimum cost by addressing two of the most cost-intensive challenges currently faced by this sector: (1) the development of a high-quality labor force and prevention of loss due to human errors, and (2) continuous application of up-to-date science and knowledge. Figure 1(a) includes photographs of the sensor devices for use in agriculture automation. Two sensor devices containing Raspberry Pi and agriculture sensors have been installed and are currently used to collect data from the aquaponics system. The sensor devices managed by Amazon IoT services generate data for every minute. A web server controls actuators by filtering the data as illustrated in figure 1(b).



(a) Sensor Devices (b) Overal Achitecture

Fig. 1: Aquaponics System

The aquaponics automation system consists of several components including, for example, the capacity to turn on the water heater in the fish tanks when the temperature goes below a lower threshold and likewise, the capacity to turn the heaters off when the temperature goes above an upper threshold. Other examples include the ability to divert water flow from the fish tanks to the vegetable beds when the water level or root moisture is below a threshold value and to redirect the water pump so that it moves water between the fish tank and the microbial chamber when the ammonia level is above a permissible threshold. To minimize false alerts in the threshold-based agriculture autoamtion, how much the noise of data stream from sensors is suppressed is our main research challenge. In this paper, we consider three approaches – MA, Kalman and LSTM filters.

A form of an MA model order q, i.e., MA(q), can be written as follows: $x_i = \mu + \sum_{i=0}^{q} \theta_i \epsilon_{i-1}$, where μ is the mathematical expectation of the process; θ is weights; ϵ_i is the Gaussian white noise with mean zero.

The state vector for the Kalman filter is represented by x_i and y_i represents unobserved data. The forms of this filter can be written as follow: $x_{i+1} = A_i x_i + B_i u_i + G_i w_i$, and $y_i = H_i x_i + v_i$. x_i represents the system state vector, y_i represents the system observation vector, and u_i represents the system control vector. A_i shows the transision matrix, H_i shows the observation matrix, B_i and G_i are dimensional matrices. w_i and v_i represent zero-mean, white noisy error terms.

A basic LSTM unit consisting of a self-connected memory cell with multiplicative gates as shown in Figure 2. As a special recurrent neural network (RNN) with an input, hidden and output layers, LSTM has been utilized for long-range dependencies. The hidden layer of LSTM at time t, the ouput – $c^{t-1,l}$ and $h^{t-1,l}$ of the previous layer at $t-1$ come in the layer at t as inputs. LSTM controls a cell status $c^{t-1,l}$ that indicates an accumulated state information. The cell state is updated or cleared by several operations. If this state is cleared, the past cell status is forgetten by $f^{t,l}$. If updated, $c^{t,l}$ – one of the outputs at t will be propagated to the final state. The cell state is prevented from vanishing or exploding gradient, which is a problem of the traditional RNN, resulting in more learning capacity.

Neural network models in the LSTM filter are trained beforehand on a set of training data to perform noise suppression. In general, they can be described as a mapping from an input feature vector x_i to the output vector

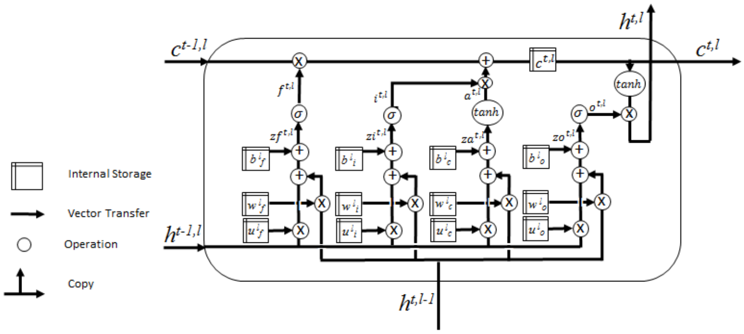


Fig. 2: The detailed Schematic of LSTM

$u_l = f(x_l, h_{l-1}, \theta)$, where the trainable parameter denotes θ, and hidden states in l, h_{l-1} for preceding data are used to model temporal context in LSTM.

The root mean square error is used to assess the quality of model prediction. It meansures the difference between the true and predicted value. The form of the root mean square error is as follows: $\sqrt{\frac{1}{N}\sum_1^N (x_i - \hat{x}_i)^2}$, where N is the total number of observations, and $\sum_1^N (x_i - \hat{x}_i)$ represents the squire of the sum of the differentces of the true value and the prediction of the model.

4. EVALUATION

Kalman [6] is a widely used, studied and developed filter, which occupies a dominant position in the field of noise suppression. The moving aveage (MA) filter is a simple Low Pass finite impulse response filter commonly used for regulating an array of sampled data. In this section, we compare the proposed LSTM with Kalman and MA filters. For fair comparison, the same period of sample data streams in our sensor devices were used.

Figure 3 illustrates three ammonium charts. Ammonia plays a significant role in an aquaponics system. Fish produces waste that is full of ammonia. Bacteria convert them into nitrites and then nitrates necessary for plant growth. Figure 3(a) shows the normal case of ammonium. Ammonia water drains to a bacterial chamber when the level of ammonium reaches a certain point. Figures 3(b) and 3(c) show ammonium charts at this moment. Figure 3(c) presents when the drain stops.

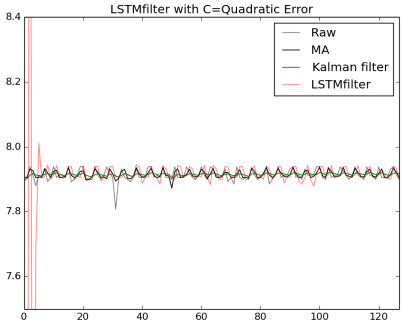
Figure 4 illustrates a different type of noise because of the movement of fishes. The nitrogen cycle is an essential biological process in aquaponics. Nitrates are a less toxic form of nitrogen, which is food for vegetable. Figure 5 shows two noise cases in nitrates. To show the effectiveness of the LSTM filter, Table 1 presents the comparison of the mean squire error.

Table 1: Comparison on Mean Square Error

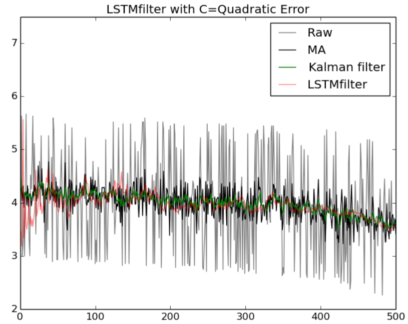
Filters	Moving Average	Kalman	LSTM
Ammonium	25.3	12.4	5.1
Nitrates	10.8	4.2	3.5

5. CONCLUSIONS AND FUTURE REMARKS

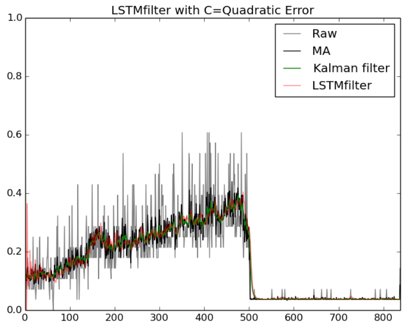
We introduced an LSTM filter for noise suppression of agriculture sensors and compared it with traditional filtering algorithms – moving average and Kalman filters. The LSTM filter is applied for large-scale aquaponics automation, improving sustainable food production. The accuracy of sensors and actuators' interactions is essential for knowledge-based automation.



(a) Normal Case



(b) Noise Case



(c) Noise and Anomaly Case

Fig. 3: Ammonium Charts (ppm)

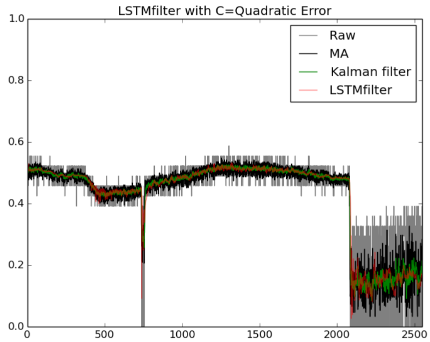


Fig. 4: Noise in Ammonium Chart (ppm)

Our threshold-based control system requires reliable readings from 30 field actuators and sensors. There are 230 combinations that represent the interplay between each set of parameters. For example, these combinations include the interplay between temperature. heating, and humidity as well as between ammonia, and oxygen dissolved via the actions of the nitrogen fixation cycle. Given the large number of possible combinations that might be considered, manual selection of combinations and the application of appropriate up-to-date knowledge is impractical and not

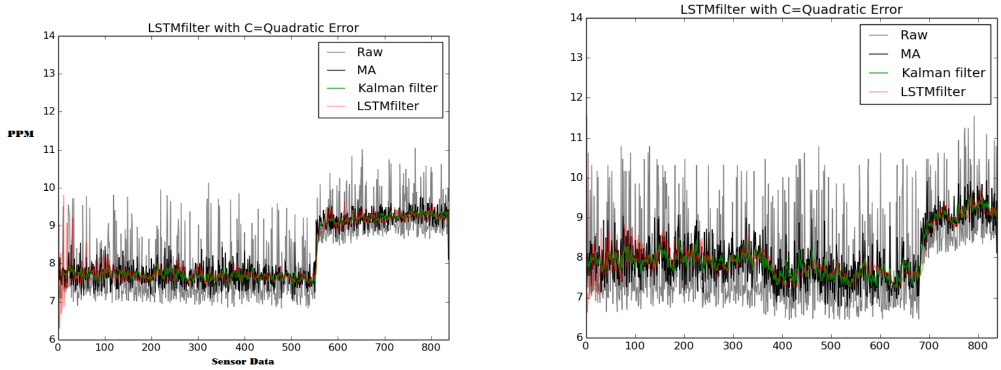


Fig. 5: Noise Case in Nitrates (ppm)

sufficiently dependable if one hopes to avoid the potential for human error. We plan to perform research focused on multi-stream neural networks to identify critical sensor-actuator interactions.

Acknowledgment

This work has been supported by US National Science Foundation CISE-MSI (RCBP-ED: CNS: Data Science and Engineering for Agriculture Automation) under grant 2131269.

References

[1] Sabine O'Hara, and Etienne C. Toussaint, "Food access in crisis: Food security and COVID-19," Ecological Economics, vol.180, 2021
[2] National Diabetes Education Program, "The Facts About Diabetes: A Leading Cause of Death in the U.S," , 2014
[3] Canning, P., A. Charles, S. Huang, K.R. Polenske, and A. Waters, "Energy Use in the U.S. Food System. United States Department of Agriculture," Economic Research Service, Report Number 94, 2010
[4] Food Carbon Footprint Calculator, Carbon Emissions, , 2006
[5] C. Weber, and S. Matthews, "Food-miles and the Relative Climate Impacts of Food Choices in the United States," Environmental Science and Technology, 4210, 3508-13, 2008
[6] Richard J. Meinhold, Nozer D. Singpurwalla, "Understanding the Kalman Filter," The American Statistician, Taylor and Francis, Pages 123-127, Mar. 2012.
[7] David C. Love, Jillian P. Fry, Ximin Li, Elizabeth S. Hill, Laura Genello, Ken Semmense, and Richard E.Thompson , "Commercial aquaponics production and profitability: Findings from an international survey," Aquaculture, vol. 435, January, 2015.
[8] Z. Jiang, W. D. Pan and H. Shen, "LSTM Based Adaptive Filtering for Reduced Prediction Errors of Hyperspectral Images," 2018 6th IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE), 2018, pp. 158-162, doi: 10.1109/WiSEE.2018.8637354.
[9] Yael Edan, Shufeng Han, and Naoshi Kondo, "Automation in Agriculture," pages 1095–1128, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
[10] Yaqi Cui and You He and Tiantian Tang and Yu Liu, "A new target tracking filter based on deep learning," pages 11-24, vol. 35, Chinese Journal of Aeronautics, 2022.
[11] Byunggu Yu, and Junwan Kim, "Using a Neural Network to Detect Anomalies given an N-gram Profile," 5th International Symposium on Cyber Security Cryptology and Machine Learning (CSCML 2021), pp 451–466, July, 2021
[12] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour and E. -H. M. Aggoune, "Internet-of-Things (IoT)-Based Smart Agriculture: Toward Making the Fields Talk," in IEEE Access, vol. 7, pp. 129551-129583, 2019, doi: 10.1109/ACCESS.2019.2932609.