Dengue-Infected Mosquito Detection with Uncertainty Evaluation of the Wingbeats using Monte Carlo Dropout

Israel Torres, Mariko Nakano, *Member, IEEE,* Jorge Cime-Castillo, Enrique Escamilla-Hernandez, *Member, IEEE*, Osvaldo Lopez-Garcia, and Humberto Lanz-Mendoza

***Abstract*— Considering that Aedes mosquitoes are a principal vector of the Dengue virus, which can cause the death of infected people in the worst case, accurate detection of infected Aedes mosquitoes is very important to prevent the further spread of the virus. This paper proposes a detection algorithm for infected Aedes *aegypti* mosquitoes by Dengue Virus-2 using their wingbeat signals. The proposed algorithm uses Long Short-Term Memory (LSTM) as a classifier of the input wingbeat signal into healthy mosquitoes and infected mosquitoes. All living beings, even those of the same species, have different characteristics depending on the season in which they are born, temperature, humidity, food, etc. This individual difference perhaps influences the level of infection, although it is fed by the same blood infected. Considering these individual differences, an uncertainty measure based on Monte-Carlo dropout is introduced. The proposed algorithm detects approximately 5% of uncertainty data from all input wingbeat signals in the test set and provides a classification accuracy of 94.87% without any uncertainty.**

Link to graphical and video abstracts, and to code: https://latamt.ieeer9.org/index.php/transactions/article/view/9221

***Index Terms*—dengue fever, infected mosquitoes, Long Short-Term Memory, Monte Carlo dropout, uncertainty**

I. Introduction

CERTAIN mosquito species are responsible for the transmission of diseases such as dengue fever, Chikungunya fever, Zika, and malaria. Notably, *Aedes* aegypti and *Aedes* albopictus are the primary vectors of dengue fever, which can, in severe cases, result in mortality among infected individuals. According to a report from the Centers for Disease Control and Prevention (CDC) in the United States, approximately four billion people in the world are at risk of contracting this disease [1]. Many Latin American countries are located in tropical or subtropical regions, which coincide with the habitat of the *Aedes* mosquito. Therefore, these countries face the common problem of spreading the diseases transmitted by *Aedes* mosquitoes, especially dengue, which accounts for approximately 90% of all cases on the American continent [2].

A countermeasure to prevent further expansion of this disease is massive fumigation to exterminate these species of mosquitoes. However, massive fumigation causes non-invertible ecological damage in addition to economic burden. Therefore, in many countries, fumigation of a specific area is only permitted when human dengue cases are reported. This is because detecting infected mosquitoes is only possible through molecular biology in specialized laboratories, which is time-consuming and expensive [3].

Considering the above-mentioned situations, the final objective of this paper is to propose a technological solution to detect dengue-infected *Aedes* aegypti mosquitoes in the field. The detection must be rapid to avoid a possible virus spread between humans, and detection accuracy must be as high as possible, with the lowest uncertainty, to perform effective fumigation. Effective fumigation minimizes the impact on ecosystems and reduces the economic burden. This research could potentially contribute to two of the United Nations Sustainable Development Goals (SDGs): Goal 3, "Good health and well-being," and Goal 15, "Life on Land."

Some scientists have reported that the behavior of infected mosquitoes has changed compared to that of healthy mosquitoes [4-6]. For instance, the authors of [4] and [5] reported that the infected female *Aedes* mosquitoes bite more frequently to obtain the same quantity of blood required to lay her eggs. The authors of [6] also observed that infected mosquitoes exhibited increased locomotor activity compared to healthy mosquitoes. These observations related to the change in behavior of infected mosquitoes suggest a possible hypothesis about the change in flight form of infected mosquitoes, and the possibility that infected mosquitoes can be detected by analyzing their wingbeat signals, which directly reflect the flight form of mosquitoes.

All living organisms, even those of the same species, have individual differences. The differences between individual mosquitoes lie in their birth season, feeding habits during their immature stages, and environmental factors such as temperature and humidity. These individual differences are

The associate editor coordinating the review of this manuscript and approving it for publication was Suelia Fleury (*Corresponding author: Mariko Nakano*).

I. Torres, M. Nakano, E. Escamilla-Hernandez, and O. Lopez-Garcia are with Instituto Politecnico Nacional, Culhuacan, Coyoacan, México (e-mails: itorres@ipn.mx, mnakano@ipn.mx, and eescamillah@ipn.mx, olopezg@ipn.mx).

J. Cime-Castillo, and H. Lanz-Mendoza are with Instituto Nacional de Salud Público, Cuernavaca, Morelos, México (e-mails: jorge.cime@insp.mx, and humberto@insp.mx).

thought to affect the degree of infection with the dengue virus and the degree of behavioral changes in mosquitoes that become infected. Considering these individual differences, the classification system should introduce a measure of uncertainty to ensure a high degree of certainty in the final decision. The number of factors influencing variation in individual mosquitoes is almost infinite, which is the principal restriction on obtaining and managing all possible variations.

To achieve the goal, this paper proposes an algorithm to detect dengue-infected *Aedes* aegypti mosquitoes using their wingbeat signals as an input signal. First, the wingbeat signal is transformed into its time-frequency representation, called a spectrogram, using the Short Time Fourier Transform (STFT). The Mel-scale in the frequency domain is used to generate the Mel-spectrogram, which is the input data for the proposed algorithm. A recurrent neural network called Long Short-Term Memory (LSTM), proposed by [7], is used as a classifier to discriminate infected mosquitoes from healthy ones.

Various LSTM configurations are evaluated to determine the best one for the objective. Furthermore, to more accurately determine whether a mosquito is infected, an uncertainty measure based on Monte Carlo theory is introduced into the proposed algorithm using Monte Carlo dropout. In the inference phase, the proposed algorithm determines that approximately 5% of the input data contains some grade of uncertainty, and the rest, 95%, does not contain uncertainty. Using input data without uncertainty, the proposed algorithm provides an accuracy of 94.87%, which is better than the performance provided by previous work [8].

The rest of the paper is organized as follows: Section II provides a brief description of concepts used in the proposed algorithm, along with related works. The proposed algorithm is described in detail in Section III, and the experimental results of the proposed algorithm are presented in Section IV. Finally, in Section V, the conclusions are provided.

II. Background and Related Works

A. LSTM Recurrent Neural Networks

LSTM proposed by [7] is one of the most utilized recurrent neural networks for classification or estimation of the time-serial data, such as voice, video data and biomedical signals [9, 10]. The LSTM contains several cells formed sequentially through time, as shown in Fig. 1(a). The LSTM layer is composed of *T* number of LSTM cells, receiving *T* input vectors $X = [x_1, x_2, \cdots, x_T]$, calculating internally T hidden vectors $H = [h_0, h_1, \cdots, h_{T-1}]$ and *T* state vectors $C = [c_0, c_1, \cdots, c_{T-1}]$ to obtain T output vectors $Y = [y_1, y_2, \cdots, y_T]$. Each LSTM cell is composed of three gates: forget gate "***f***", input gate "***i***" and output gate "***o***". Using these three gates, the LSTM networks determine the importance grade of relationship between data of different time-steps. The LSTM cell of time-step "t" receives input vector x_t, hidden vector h_{t-1} and cell state vector c_{t-1} of the previous time-step to calculate the output vector y_t as shown in Fig. 1(b).

The outputs of these three gates, and outputs of hidden and cell state vectors are given by (1)-(5).

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big) \tag{1}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\begin{aligned} \tilde{C} &= \tanh\,(W_c \cdot [h_{t-1}, x_t] + b_C) \\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \end{aligned} \tag{3}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{4}$$

$$h_t = o_t * \tanh\,(C_t) \tag{5}$$

where $\sigma(\cdot)$ and $\tanh(\cdot)$ are sigmoidal and hyper-tangent activation functions, W_f, W_i and W_o are weight vectors of three gates, and W_c is weight vector for state cell. The operator "$\cdot$" is inner product, while "$*$" is element-wise multiplication. $[h_{t-1}, x_t]$ means concatenation of two vectors h_{t-1} and x_t.

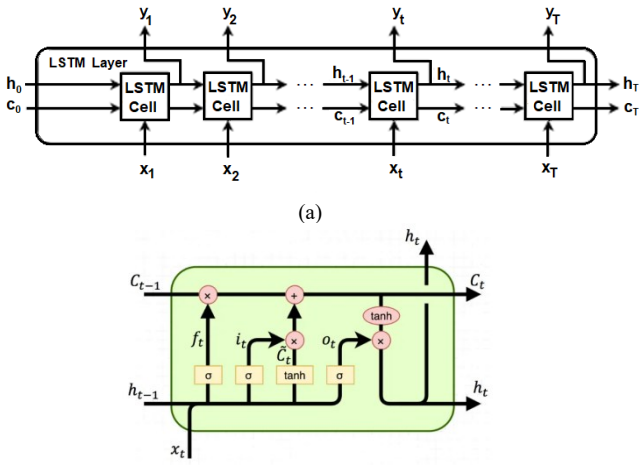


Fig. 1. (a) LSTM layer and (b) LSTM cell [7].

Bidirectional LSTM (Bi-LSTM) is a variant of LSTM, in which two LSTM layers are combined, performing forward and backward data flows. Bi-LSTM allows capturing both past and future contexts of the input sequence, therefore in several applications, such as speech recognition and natural language processing, Bi-LSTM performs better than ordinal LSTM.

B. Monte Carlo Dropout

When the input data contains uncertainty for a classification or regression system, Bayesian inference can be used to estimate the uncertainty of these data and improve system reliability. For example, a regression system has been trained using N input data $X = \{x_1, \cdots, x_N\}$ and its corresponding output data $Y = \{y_1, \cdots, y_N\}$. When this regression system receives new data $\tilde{x}$, the distribution of output data $\tilde{y}$ can be estimated using Bayesian inference, as given by (6).

$$p(\tilde{y}|\tilde{x}, X, Y) = \int p(\tilde{y}|w, \tilde{x})\, p(w|X, Y) dw \tag{6}$$

where w is the weight of the regression system and $p(w|X, Y)$ is the posterior distribution of w given X and Y.

Bayesian inference can be applied to conventional machine learning systems with reasonable computational complexity. However, because classification or regression system based on

deep learning contains highly non-linearity, the application of Bayesian inference is not feasible [11].

Monte Carlo dropout was proposed to evaluate uncertainty of the data for deep learning-based classification or regression systems [11]. Generally, the dropout operation is used to mitigate overfitting during the training phase by randomly removing a portion of neurons in the hidden layers. Random removing weight connection avoids memorized relationships between inputs and outputs, improving the generalization capability of the system. However, the authors of [11] proposed application of the dropout operation at inference phase to detect uncertainties in input data.

During the inference phase, the trained deep learning system produces the same output if the same input data is introduced. However, the trained system with Monte Carlo dropout can produce different outputs if input data contains some uncertainty due to a dropout operation during the inference phase.

Other methods exist to quantify uncertainty besides the Monte Carlo dropout method. The ensemble-based method and the Softmax-based method are the most used [12]. The ensemble-based method needs several pre-trained classifiers with different initial values or configurations. The uncertainty is quantified using discrepancies among the classifiers' outputs. This method requires large memory space to store all classifiers during the inference phase [12]. The Softmax-based method evaluates the output values of the Softmax activation function [13]. In this model, uncertainty is calculated by the difference between the probabilities of the top-1 and top-2 classes. A small difference provides large uncertainty, and vice versa. This method is meaningful only for multiclass classification.

In [14], several uncertainty quantification methods are compared using a common deep neural network under the same criteria. The comparison result indicates that the Monte Carlo dropout method is the most accurate in different datasets. Considering the excellent classification accuracy, the compactness of the system, and the purpose of use being binary classification, the Monte Carlo dropout method seems to be the most suitable for this proposal.

C. Related Works

To date, several scientific studies have been proposed to classify mosquitoes by species, such as *Aedes sp*, *Anopheles sp* and *Culex sp*. The main objective of these works is vector surveillance in the entomological field and future vector control planning, calculating the distribution of different mosquitos' species in determined regions. Some works are focused on the immature stage of mosquitoes [15], such as eggs and larvae, and other works focus on the adult stage [16-18]. Within these proposals that focus on the adult stage of mosquitoes for species classification, [16] uses images data taken by camera to classify males and females of three species: *Aedes aegypti*, *Aedes albopictus* and *Culex quinquefasciatus*. The works [17, 18] use wingbeat signals to classify species of mosquitoes. In [17], wingbeat signals were captured by optical sensor, and the classification of six species of mosquitoes was performed by several off-the-shelf CNNs. The authors of [18] proposed a classification algorithm for 20 species of mosquitoes using wingbeat signals captured by smartphones. The principal objective of this work is to show that relatively cheap smartphones are useful for vector surveillance.

The objective of all previous works is a classification of mosquito species, which is different from the objective of this paper. Our previous proposal [8] has the same objective as this paper's. In [8], 15 spectral features based on spectral analysis are extracted from the wingbeat signal of dengue virus-infected *Aedes aegypti* mosquitoes and healthy mosquitoes of the same species. The extracted 15 spectral features are Spectral Rolloff, Spectral Centroid, Spectral Spread, Spectral Contrast, Spectral Flux, Spectral Flatness, Spectral Slop, and so on. Two common classifiers, Support Vector Machine (SVM) and K Nearest Neighbors (KNN) are used. Additionally, the LSTM recurrent network with two layers of 64 hidden neurons is employed for the same purpose, showing better performance (89.35%) compared with the performance obtained by SVM and KNN.

III. Proposed System

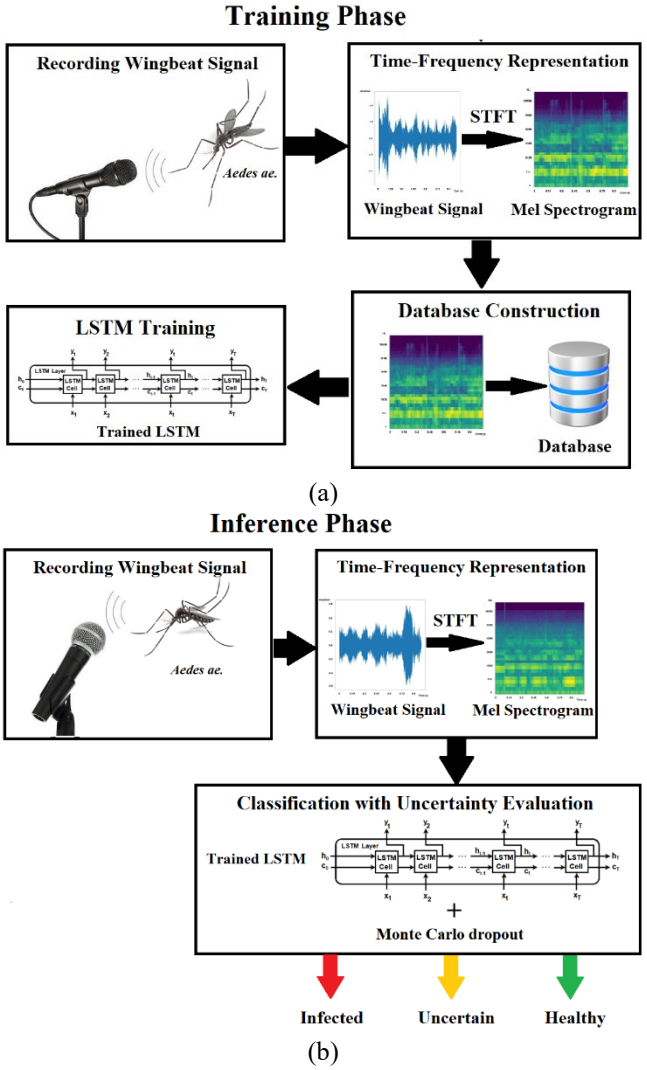


Fig. 2. Global diagram of proposed system. (a) Training phase, (b) Inference phase.

The proposed system consists of two phases: the training phase and the inference phase. Both phases are very similar and consist of two steps: the preprocessing step, in which the input wingbeat signal is transformed into time-frequency data, and training or testing phase based on the LSTM recurrent neural network. In the training phase, the LSTM network is trained using a time-frequency representation of the wingbeat signal in the training set. Meanwhile, in the inference phase, the trained LSTM with Monte Carlo dropout uses the time-frequency representation of the testing data to classify the input data into infected mosquitoes and healthy mosquitoes with high certainty or determine that the input data is uncertain. The block diagram of the training phase of the proposed algorithm is depicted in Fig. 2 (a), while the inference phase is depicted in Fig. 2(b).

A. Time-Frequency Transformation

TABLE I
PARAMETERS USED FOR MEL SPECTROGRAM

Parameters	Values
Sampling rate	48kHz
Length of the FFT window	1024
Hop length (length of overlapping)	256
Window length	1024
Window type	Hanning
Number of Mel bands	32

First, the input mosquito's wingbeat signals are transformed into a time-frequency representation using the Short-Time Fourier Transform (STFT), and the Mel-scale is applied in the frequency domain to obtain the Mel spectrogram. The parameters shown in Table I are used to obtain the Mel spectrograms. Fig. 3 shows examples of input wingbeat signals of infected and healthy mosquitoes and their Mel spectrogram, respectively.

All mosquitoes (infected and healthy) are female *Aedes ae.* mosquitoes. The fundamental frequency, as the first frequency zone with a high-power element (yellow part) in the Mel spectrogram, is approximately 500Hz, which coincides with the fundamental frequency of female *Aedes ae.* mosquitoes reported by [19].

B. Training Phase and Inference Phases of LSTM

Several architectures of LSTM-based networks are evaluated to select the best ones, varying the number of LSTM layers, the number of hidden neurons in the ordinal LSTM or Bi-LSTM, and the dropout rate. The selection of the best networks is based on classification accuracy and the number of trainable parameters. This evaluation process of several LSTM architectures is provided in Section IV in detail. The hyperparameters used to train the selected LSTM networks are given in Table II.

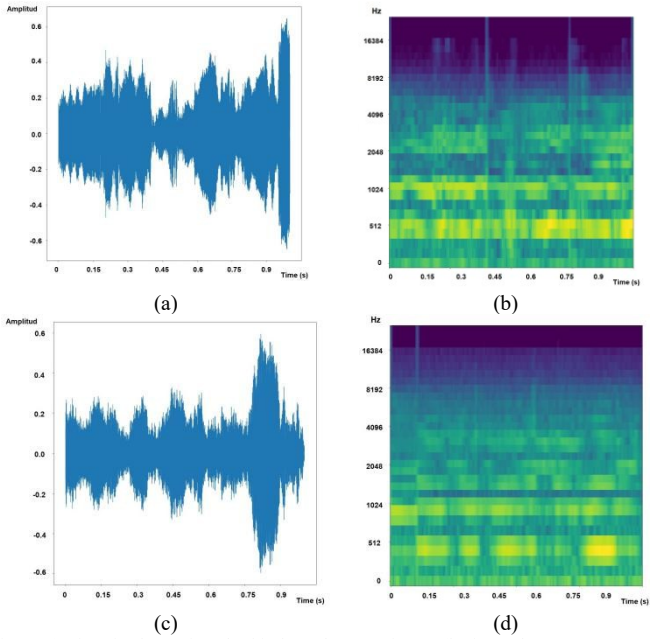


Fig. 3. (a) is wingbeat signal of infected mosquitoes, (b) is Mel spectrogram of (a), and (c) is wingbeat signal of healthy mosquitoes, and (d) is the corresponding Mel spectrogram.

TABLE II
HYPERPARAMETERS USED IN TRAINING

Hyperparameters	Values
Loss function	Binary Cross-Entropy Function
Optimizer	Adam, lr =0.000005, $\beta_1 = 0.9, \beta_2 = 0.999$
Number of Epochs	300
Batch size	16

Once the best architectures based on LSTM and Bi-LSTM were selected and trained using the hyperparameters given in Table II, in the inference phase, each test data was introduced N times to evaluate its uncertainty using the Monte Carlo dropout technique. If the degree of uncertainty of input data is greater than a predetermined threshold, it is considered uncertain, and the classification result is not taken into account. In the inference phase, the proposed algorithm provides one of three possible outputs, which are "infected", "healthy" or "uncertain input data" as shown in Fig.2 (b).

IV. EXPERIMENTAL RESULTS

A. Data Collection and Dataset Construction

Due to the nonexistence of a public domain dataset with desired characteristics whose wingbeat signal is labeled as infected or healthy, our own dataset is constructed, recording the wingbeat signals of female *Aedes ae.* mosquitoes infected with dengue type 2 and healthy female *Aedes ae.* mosquitoes. An acoustic chamber, shown in Fig. 4, is constructed to record the wingbeat signal. The chamber uses sound-blocking material to reduce environmental noise. The upper part of the acoustic chamber has a hole for introducing an omnidirectional condenser microphone to capture sound.

Fig. 4. Acoustic chamber used for recording of wingbeat signals.

Using this acoustic chamber, approximately 150 wingbeat signals from infected mosquitoes and another 150 wingbeat signals from healthy mosquitoes are recorded. Infected mosquitoes are fed on blood infected with Dengue type 2 four days before recording. The sound capture was performed with a sampling rate of 48 kHz, and the duration of all signals was more than 2 seconds. From each recorded signal, a one-second signal is extracted, removing the initial and the final part because these parts do not have any wingbeat signals. In total 300 wingbeat signals (150 infected and 150 healthy) are divided equitably into 80% for training set and 20% for test set. The temperature and humidity during recording are 19.75±1ºC and 42.28±3 %rh, respectively.

B. LSTM Architecture

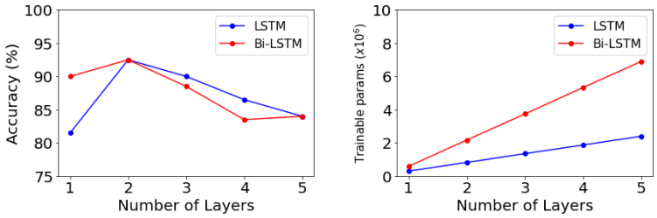
As mentioned before, the best architecture that provides the highest classification accuracy using a reasonable number of trainable parameters is determined. To this end, several parameters related to LSTM architecture are considered, which are given in Table III.

TABLE III
PARAMETERS USED FOR LSTM ARCHITECTURE

Configuration	Values
Type	{LSTM, Bi-LSTM}
Number of LSTM layers	[1,2,3,4,5]
Number of hidden neurons	[32, 64, 128, 256, 512]
Dropout rate	[0.1,0.2,0.3,0.4,0.5]

As mentioned in Section II.A, Bi-LSTM can leverage the bi-directional relationship in the input sequence, increasing classification accuracy in some applications. On the other hand, the selection of the number of LSTM layers is essential because a deeper network tends to overfit, losing generalization ability; however, a few layers can cause underfitting in some tasks. The best number of LSTM layers depends on the applications and the amount of available training data. The selection of the number of hidden neurons in each LSTM layer is also a crucial issue in any deep learning system [20]. The main advantage of a wider network is its generalization ability, especially when the input data contains noise elements, or the diversity of the input data is relatively large. However, a wider network structure leads to an exponential increase in trainable parameters, so-called as “the curse of dimensionality”.

First, to define the number of layers in the LSTM and Bi-LSTM architectures, both architectures with five different numbers of layers are trained using the training set and evaluated using the test set. Fig. 5 shows classification accuracy and the number of trainable parameters, varying the number of layers in LSTM and Bi-LSTM architectures. Fig. 5(a) shows the relationship between the number of layers and classification accuracy. Fig. 5(b) shows the relationship between the number of layers and the required number of trainable parameters. In all cases, the number of hidden neurons in each layer is 256, although the tendency is the same regardless of the number of hidden neurons. From this figure, the best number of layers in both architectures is determined as 2.



(a) (b)

Fig. 5. (a) Relationship between the number of layers and classification accuracy. (b) Relationship between the number of layers and number of trainable parameters.

To define the best dropout rate, both types of architecture are evaluated with five different dropout rates, as shown in Table III. The relationship between dropout rates and accuracy is shown in Fig. 6. From this figure, the best dropout rate is equal to 0.2.

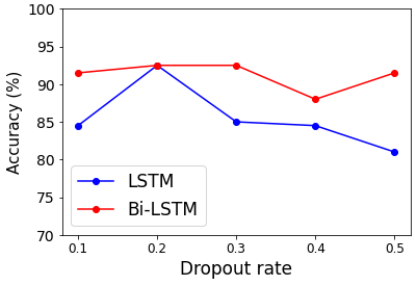
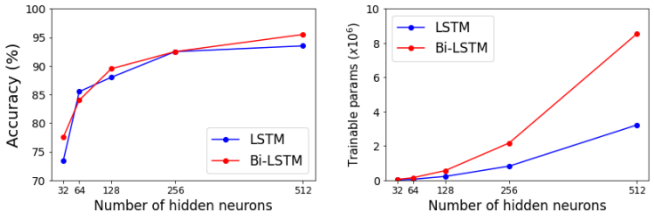


Fig. 6. Relationship between dropout rate and classification accuracy.

For two-layer LSTM and two-layer Bi-LSTM with a dropout rate of 0.2, the best number of hidden neurons is determined, varying this as shown in Table III. Fig. 7(a) shows the relationship between the number of hidden neurons and the classification accuracy. Fig. 7 (b) shows the relationship between the number of hidden neurons and the number of required trainable parameters. Considering the results shown in Fig.7, two architectures to evaluate the proposed algorithm are selected, which are two-layer LSTM and two-layer Bi-LSTM with 256 hidden neurons. In both architectures, the dropout rate

is set to 0.2. It is worth noting that in all evaluations, the accuracies are average values of 10 trials, randomly introducing test data on both architectures.



(a) (b)

Fig. 7. (a) Relationship between the number of hidden neurons and classification accuracy for two-layer LSTM and two-layer Bi-LSTM (b) Relationship between the number of hidden neurons and number of trainable parameters for two-layer LSTM and two-layer Bi-LSTM.

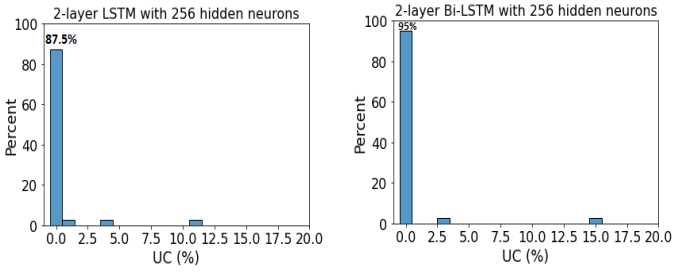
The trainable parameters of the selected architectures are 0.8 million for two-layer LSTM and 2.1 million for two-layer Bi-LSTM, which are smaller than the small version of MobileNetV3 with 2.5 million parameters [21]. This ensures future deployment of the proposed algorithm on a mobile device and performs real-time detection.

C. Uncertainty Evaluation

As mentioned in Section III.B, each test data is evaluated for its uncertainty by feeding N times the same data into the trained network. If all N results are equal, the certainty of the input data is 100%. The uncertainty degree (UC) of the input data is calculated in the following manner:

$$UC\ (\%) = \begin{cases} \dfrac{N-T}{N} \times 100 & if\ T > \dfrac{N}{2} \\ \dfrac{T}{N} \times 100 & otherwise \end{cases} \tag{6}$$

where UC is the degree of uncertainty, N is the number of times to introduce the same input data, and T is the sum of outcomes in N trials. For example, if N is equal to 100 and T is equal to 99, the UC is 1%, while if N is equal to 100 and T is equal to 45, the UC is 45%.



(a) (b)

Fig. 8 (a) Histogram of uncertainty data detected using two-layer LSTM, (b) Histogram of uncertainty data detected using two-layer Bi-LSTM.

For the two selected networks, two-layer LSTM and two-layer Bi-LSTM, with 256 hidden neurons and a dropout rate of 0.2, the UC (%) on the test data is calculated. Where N set to 100. Fig. 8 shows a normalized histogram of UC (%), in which it can be observed that very few data have uncertainty, with more than 85% of test data having 0% uncertainty (100% reliability). In the case of two-layer LSTM, the network classifies 87.5% of the test data as 100% reliability. And the two-layer Bi-LSTM considers that 95% of the test data are classified as 100% reliability.

D. Classification Accuracy

In the inference phase, using the UC(%) value and T value calculated in uncertainty evaluation, as shown by (6), the network's output is obtained by following the process given by (7).

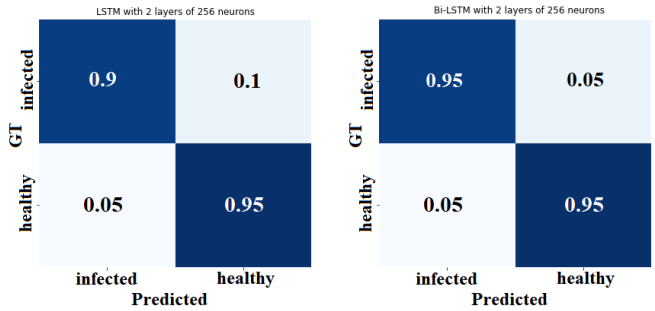
$$\begin{aligned} &if\ UC \le th\ \ then \\ &\quad if\ T \le \frac{N}{2}\ \ then \\ &\qquad output \leftarrow \text{"infected mosquito"} \\ &\quad otherwise\ \ output \leftarrow \text{"healthy mosquito"} \\ &otherwise\ \ output \leftarrow \text{"}uncertain\ input\text{"} \end{aligned} \tag{7}$$

where th is a threshold value, which is set at 0%. N is the number of times to introduce the same input data to the trained networks. And T is the sum of outcomes in N trials. The threshold value is 0% means that only 100% reliable data provides classification results, such as "infected mosquito" or "healthy mosquito". The selection of the threshold th may depend on the amplitude of environmental noise. In a noisy environment, this value must increase because a large part of the wingbeat signal would be determined as uncertain data.

Table IV shows the performances of both architectures, and Fig. 9 shows the confusion matrices obtained by both architectures. From the confusion matrices provided in Fig. 9, it can be concluded that in the proposed algorithm, classification confusion is minimal. Considering that all test data are reliable with high certainty (100%).

TABLE IV
PERFORMANCE OF THE PROPOSED ALGORITHM

Architectures	Accuracy	Precision	Recall	F1-score
2-layer LSTM	91.67%	92.06%	91.89%	91.88%
2-layer Bi-LSTM	94.87%	94.87%	94.87%	94.87%



(a) (b)

Fig. 9 Confusion matrices of both best architectures. (a) 2-layer LSTM with 256 neurons. (b) 2-layer Bi-LSTM with 256 neurons.

TABLE V
PERFORMANCE COMPARISON

Methods	Accuracy (%)
KNN with K=3 [8]	84.32
KNN with K=5 [8]	80.39
SVM with linear kernel [8]	80.39
SVM with Gaussian kernel [8]	78.84
2-layer LSTM with 64 HNs without uncertainty evaluation [8]	89.35
2-layer LSTM with 256 HNs with uncertainty evaluation (Proposed)	**91.67**
2-layer Bi-LSTM with 256 HNs with uncertainty evaluation (Proposed)	**94.87**

The performance of the proposed algorithm is compared with our previous work [8], in which the same dataset was used for the classification of infected mosquitoes and healthy mosquitoes. This previous work used two common machine learning techniques: SVM and KNN classification with two different hyperparameters and a two-layer LSTM recurrent network with 64 neurons without uncertainty evaluation. Table V shows performance comparisons between five previous works and two proposed algorithms.

Table V shows that the proposed algorithm provides much better classification accuracy, increasing to 10% accuracy with respect to the common classifiers, such as SVM and KNN. Additionally, the proposed algorithm shows approximately 5% better performance with respect to the previously proposed LSTM architecture. It is worth noting that the proposed algorithm provides accuracy with high reliability using only high certainty test data. These certainty data are detected by uncertainty evaluation based on Monte Carlo theory applied to deep learning, denominated as Monte Carlo dropout. Providing highly reliable results is very important for decision making in several fields, such as the medical and public health field.

V. Conclusions

This paper proposes LSTM-based classification algorithms to distinguish dengue type 2 infected mosquitoes from healthy ones, using their wingbeat signals as input data. The proposed algorithm is based on several observations about the change of behavior of dengue-infected mosquitoes reported in scientific papers [4-6] and on the hypothesis that the flight morphology of infected mosquitoes can be modified.

The main contributions of this work are as follows:

1. The best architectures based on LSTM and Bi-LSTN recurrent neural networks are selected from several experiments
2. Incorporate an uncertainty measure based on the Monte Carlo theory into the proposed architectures, providing highly reliable results.
3. The proposed algorithm shows 10% higher accuracy than the standard hand-crafted classifiers and 5% higher than the LSTM-based classifier.
4. The proposed algorithms can be implemented in a mobile device because they require a small number of parameters to perform. This can offer possible implementation in the field.
5. Our own dataset was constructed, which contained 150 wingbeat signals of dengue-infected mosquitoes and 150 wingbeat signals of healthy mosquitoes. As far as our knowledge, it is the first dataset involved in the wingbeat signal of infected mosquitoes.

This model can facilitate an alternative method for diagnosing insect disease vectors, wherein the distinction between the flight frequency of an infected organism and an uninfected one is isolated. This approach is applicable to sand flies, vectors of Leishmania, or other mosquito species, such as those of the genus Anopheles, vectors of the malaria-causing parasite. This methodology has implications for utilizing flight signal measurements as a diagnostic tool for infected mosquitoes, which may prove valuable when implemented in areas endemic to dengue mosquito vectors as a strategy to mitigate disease transmission.

The main limitation of this study is that the dataset is small and has only slight variations since it was collected during the same season. The temperature and humidity were 19.75±1ºC and 42.28±3%rh, respectively. According to [19], the fundamental frequency of wingbeat signal varies depending on environmental and physiological factors, such as flight time (day or night), the age of the mosquitoes, and whether they are gravid. Because the proposed method is based on the transition of frequency components of the wingbeat signal during mosquitoes' flight, the variation of fundamental frequency may influence the experimental results. Therefore, it is essential to collect more data under different conditions to further support and substantiate the hypothesis that dengue-infected mosquitoes change their flight morphology. The collected data must include several variations, such as the age of mosquitoes, the number of days since the last feeding, and environmental variations, such as temperature, humidity, and recording time.

The issue in actual field deployment of the proposed system depends on how environmental noise can be controlled. An adequate noise-canceling algorithm must be introduced to the proposed system. After noise cancelation, a certain amplitude of residual noise may cause higher uncertainty of recorded wingbeat signals. In this case, to detect possible infected mosquitoes, the threshold for discarding uncertain signals must be increased. Increasing the threshold relaxes the uncertainty criterion and may result in an increase in false positive or false negative errors. To accurately identify infected mosquitoes in the field, the threshold value must be adjusted according to the residual noise level. This is considered a major challenge.

References

[1] CDC Homepage about Dengue. [Online]. Available: https://cdc.gov/dengue/about/index.html.

[2] Pan American Health Organization, Epidemiological update for dengue, chikungunya and zika in 2022. [Online]. Available: https://ais.paho.org/PAHOArboBulletin2022.pdf

[3] E. Harris *et al.*, "Typing of Dengue Viruses in Clinical Specimens and Mosquitoes by Single-Tube Multiplex Reverse Transcriptase PCR", AMS Journals, Journal of Clinical Microbiology, vol. 36, no. 9, pp2634-2639, 1998

[4] B. Xiang *et al*., "Dengue virus infection modifies mosquito blood-feeding behavior to increase transmission to the host", *PNAS*, vol. 119, no. 3, e2117589199, 2022. doi: 10.1073/pnas.2117589119.

[5] P. Mendez-Luz *et al*., "Potential impact of a presumed increase in the biting activity of dengue-virus-infected *Aedes aegypti* (Diptera: Culicidae) female on virus transmission dynamics", *Mem. Inst.*

Oswaldo Cruz, vol. 106, no. 6, pp. 755-758, 2011. doi: 10.1590/s0074-02762011000600017

[6] T. Lima-Camara *et al.*, "Dengue infection increases the locomotor activity of *Aedes aegypti* females", *PLOS One*, vol. 6, no. 3, e17890, 2011. doi: 10.1371/journal.pone.0017690

[7] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", *Neural Computation*. vol. 9, no. 8, pp. 1735-1780, 1997 doi: 10.1162/neco.1997.9.8.1735

[8] M. Haro, M. Nakano, I. Torres, M. Gonzalez and J. Cime, "LSTM-based Infected Mosquitos Detection using Wingbeat Sound", in *Proc. the 22nd Mexican Int. Conf. on Artificial Intelligence*, Merida, Mexico, 2023.pp. 157-164. doi: 10.1007/978-3-031-47640-1_13

[9] H. Zehir, T. Hafs and S. Daas, "Empirical mode decomposition-based biometric identification using GRU and LSTM deep neural netwoks on ECG signals", *Evolving Systems*, vol. 15, pp. 2193-2209, 2024, doi: 10.1007/s12530-024-09611-7.

[10] Z. He, J. Yang, R. Alroobaea and L.Y. Por, "SeizureLSTM: An optimal attention-based trans-LSTM network for epileptic seizure detection using optimal weighted feature integration", *Biomedical signal Processing and Control,* vol. 96, Part B, 106603, 2024, doi: 10.1016/j.bspc.2024.106603.

[11] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: representing model uncertainty in deep learning". in *Proc. International Conference on Machine Learning.* New York, USA, 2016. pp. 1050–1059. doi: 10.1145/1553374.1553375

[12] B. Lambert, F. Forbes, S. Doyle, H. Dehaene and M. Dojat, "Trustworthy clinical AI solutions: A unified review of uncertainty quantification in deep learning models for medical image analysis", *Artificial Intelligence In Medicine*, vol. 150, 102830, 2024, doi:20.2016/j.artmed.2024.102830.

[13] Y. Zheng et al. "Elongated Physiological Structure Segmentation via Spatial and Scale Uncertainty-Aware Network",in Proc. *MICCAI 2023*, Vancouver, Canada, 2023. pp. 323-332, doi: 10.1007/978-3-031-43901-8_31

[14] M. Sensoy, L. Kaplan, and M. Kandemir, "Evidential deep learning to quantify classification uncertainty", in *Proc. 32nd Conf. on Neural Information Processing Systems,* Montreal, Canada, 2018, pp.3179-3189,

[15] A. Arista-Jalife, *et al*., "*Aedes* mosquito detection in its larval stage using deep neural networks", *Knowledge-based systems*, vol. 189, 104841, 2020. doi: 10.1016/j.knosys.2019.07.012

[16] D. Motta *et al*., "Application of convolutional neural networks for classification of adult mosquitoes in the field", *PLOS One*, vol. 14, no.1, e0210829. 2019. doi: 10.1371/journal.pone.0210829

[17] E. Fanioudakis, M. Geismar and I. Potamitis. "Mosquito wingbeat análisis and classification using deep learning", in *Proc. 26th EUSIPCO*, Rome, Italy, 2018, pp. 2410-2413. doi: 10.23919/eusipco.2018.8553542

[18] H. Mukundarajan, F. Hol, E. Castillo, C. Newby and M. Prakash, "Using mobile phones as acoustic sensors for high-throughput mosquito surveillance", *e-life*, vol.6, e27854, 2017. doi: 10.7554/elife.27854

[19] B. Arthur, K. Emr, R. Wyttenbach and R. Hoy, "Mosquito (*Aedes aegypti*) flight tones: Frequency, harmonicity, spherical spreading, and phase relationships", *J. Acoust. Soc. Am*. vol. 135, no. 2, pp. 941, 2014. doi: 10.1121/1.4861233

[20] L. J. Ba and R. Caruana, "Do deep nets really need to be deep?", *Advances in Neural Information Processing systems*, vol. 3, no. 1, pp. 2654-2662, 2014. doi: 10.4135/9781071878965.n2

[21] A. Howard *et al*., "Searching for mobileNetV3", in *Proc .ICCV*, Seoul, Korea, 2019, pp. 1314-1324. doi: 10.1109/iccv.2019.00140

[22] D. Kim, T.J. DeBriere and N. D. Burkett-Cadena, "Effect of physiological and environmental factors on mosquito wingbeat frequency", *J. of Vector Ecology*, vol. 42, no. 2, 2024, pp. R70-R77. doi: 10.52707/1081-1710-49.2.R70.

Israel Torres was born in México. He received the B.S. degree in communications and electronic engineering with acoustic specialty in 2022 and Currently he courses M.S. degree in microelectronics Engineering both from Instituto Politécnico Nacional (IPN) His principal research interest are: sound and image processing, AI, ML, RL, positional encoding, deep dream, LLM & MML.

Mariko Nakano received the bachelor's and master's degrees from The University of Electro Communication, Tokyo, in 1983 and 1985, respectively, and the Ph.D. degree in science from Universidad Autónoma Metropolitana, Mexico, in 1999. She is currently a professor in the Graduate and Research Section, Instituto Politécnico Nacional, Mexico. Her research interests include image processing, information security, and machine learning. To date, she has supervised 20 doctoral theses and 72 master's theses.

Jorge Armando Cime Castillo earned his biologist degree in 2013, and his M. S. in 2005 from the National Autonomous University of Mexico, his PhD in Science from the Institute of Biomedical at UNAM 2014, he was a fellow of the National Council of Science and Technology at the National Institute of Public Health's Center (INSP). He has also completed research stays at the Institute of Science and Technology in Lille, France, and at the Southwest Foundation for Biomedical Research Institute in San Antonio, Texas in 2015. Currently he has joined the National Institute of Public Health, and is a researcher in Medical Sciences and level I of the National System of Researchers Mexico.

Enrique Escamilla Hernández received his degree in Electronics Engineering from the UAM in 1997. In 2002 and 2006, he received his Master of Science degree in Microelectronics Engineering and his Ph.D. degree in Communications and Electronics, respectively, from the ESIME Culhuacan at IPN Mexico. In 2008, he joined the Graduate Studies and Research Section of the ESIME, Culhuacan, where he is currently a full-time professor. His research interests are in the fields of pattern recognition, digital signal processing, and electronic design.

Osvaldo López García received his degree in communications and electronic engineering from Instituto Politecnico Nacional (IPN) in 2000. Beginning as professor of communication in Escuela Superior de Ingenieria Mecanica y Electrica (ESIME) Culhuacan in 2001, In 2005, head of dept. of material resources and in 2012 head of Graduate department. From 2019 until now he is head of the Graduate Section of ESIME Culhuacan.

Humberto Lanz Mendoza obtained a degree in biology from the National Autonomous University of Mexico in 1985, a Master of Science in Immunology from the National School of Sciences of the IPN in 1989, and a Doctor of Science in Immunology in 1992 from the same institution. He was a fellow of the National Council of Sciences and Technology. He completed postdoctoral studies at the University of Stockholm, at the Polytechnic School of Zurich and at the Insect Research Center of the University of Arizona. He joined the INSP in 1998, where he is currently Director of the Area of Infection and Immunity, Level “F” Researcher, and Level III in the National System of Researchers.