

# Thermal Neutron Classification in the Hohlräum Using Artificial Neural Networks

Jesus E. Rivero, Rosa M. Valdovinos, Edgar Herrera, Hector A. Montes-Venegas and Roberto Alejo

**Abstract**—One potential method for the treatment of brain tumors is the use of thermal neutrons flux. These type of neutrons are produced by the nuclear fission of uranium-235. When the nuclear fission occurs at the core of a reactor, part of the thermal neutrons flux go into the Hohlräum generating an isotopic expansion. To successfully carry out this radiation technique, it is necessary to evaluate the performance of the thermal neutrons flux at the Hohlräum during he isotropic expansion. In this paper, the classification of the resultant flux is achieved using a Multilayer Perceptron Artificial Neural Network (ANN). For the ANN topology configuration, a set of parameters specific to the structure of the horizontal thermal column of the Hohlräum reactor have been determined. Experimental results show the viability of the model and provide additional support to the reactor facilities remodeling tasks.

**Keywords:** Artificial Neural Networks, Multilayer perceptron, Back propagation algorithm, Thermal neutrons.

## I. INTRODUCTION

The National Institute for Nuclear Research (ININ, for its Spanish acronym) in Mexico boasts a TRIGA MARK III nuclear reactor, designed for nuclear research purposes, as well as for the production of radioisotopes and for operative staff training. The reactor routinely operates in steady state mode at thermal power levels up to 1000 KW and it is able to be repeatedly pulsed up to a peak power of approximately 2000 KW. Similarly, the ININ boasts research facilities to carry out gamma radiation and neutrons studies, as well as tests on the effects of large doses of radiation and sample activation [1]. The thermal column of the reactor is located at the north side of the pool structure, it goes from the outer structure of the core of the reactor to the external structure of the concrete shield construction (See figure 1).

Access to the thermal column is through a concrete gate similar to the one in the exposition room, only smaller. Inside the facilities there is an empty space or *Hohlräum* of size 91x91x100 cm, shrouded with borated polyethylene and graphite. Neutron shielding in this area is made using a large amount of graphite, the concrete structure, a 30 cm thick steel plate located at the top of the horizontal thermal column and next to the *liner*, and a wall of lead blocks right

This work has been partially supported by the UAEM grant 3834/2014/CIA, the SEP grant PROMEP/103.5/11/3796, the TESJO grant SDMAIA-010, and the ININ reactor operation UR-001.

Jesus E. Rivero (gregowrs@hotmail.com) is with the Centro Universitario UAEM Texcoco, Av. Jardín Zumpango s/n, Fracc. El Tejocote, Texcoco, Mexico. C.P. 56259

Rosa M. Valdovinos (rvaldovinosr@uaemex.mx) and Hector A. Montes-Venegas (hamontesv@uaemex.mx) are with the Universidad Autónoma del Estado de México, Facultad de Ingeniería, Cerro de Coatepec s/n, Ciudad Universitaria, Toluca, Mexico, C.P. 50130.

Roberto Alejo (ralejoll@hotmail.com) is with the Tecnológico de Estudios Superiores de Jocotitlán, Carretera Toluca-Atzacmulco KM. 44.8, Ejido de San Juan y San Agustín, 50700 Jocotitlán (Mexico)

Edgar Herrera (aha@nuclear.inin.mx) is with the Instituto Nacional de Investigación Nuclear ININ, Departamento del Reactor, Carretera México-Toluca s/n, 52750, La Marquesa, Ocoyoacac, Mexico

next to the reactor core used to reduce the levels of gamma radiation [2] (See Figure 2).

On one of the Hohlräum surfaces, there is a 15 cm diameter intake for the neutrons flux at various energy levels, e.g., cold, thermal, epithermal, slow, medium, fast and ultrafast. The neutrons flux interacts inside the reactor to generate the isotropic expansion and thus, irradiate the samples contained within.

Amongst the available neutrons, those of interest to our study are the thermal neutrons, as they are the ones with less energy suitable to be used in the the boron neutron capture therapy (BNCT) for the treatment of brain tumors. A small number of studies have been focused to the detection of thermal neutrons, due to that most of the elements in the Periodic Table yield nuclear reactions of type  $(n,g)$ , causing the formation of radioactive nucleus [3].

In the treatment of brain tumors, BNCT is based in the nuclear reaction that takes place when the stable nucleus of Boron-10 is irradiated with thermal neutrons (0.025 eV) to produce an alfa particle ( $2\text{He}4$ ) and a nucleus of Li-7. These particles have a motion length of approximately a cellular diameter ( $5 - 9 \mu\text{m}$ ), producing a high linear energy transference. Only the alfa particles at a narrow interval of 1.7 MeV produced during this nuclear reaction are enough to direct the cellular nucleus and destroy them. This destructive action occurs mainly in the cancer cells that have accumulated boron. Other cells with low concentrations of boron are not significantly damaged. As a result, the current tendency is to consider BNCT as a complementary therapy that could be more relevant for the treatment of tumors resistant to other therapies. Nevertheless, in order to make it viable at the ININ, it is required to evaluate the neutrons flux profile because when the TRIGA MARK III is fully operational, the behavior and the displacement of the produced thermal neutrons inside the Hohlräum cannot be verified, due to the lack of some measuring instrument or simulation that allow us to know the state of these particles.

In this study, an Artificial Neural Network (ANN) is used to model the thermal neutron flux in the Hohlräum to provide additional support for the reactor remodeling facilities.

An ANN is defined as a non-linear mapping system which structure is based on principles observed in the biological hu-

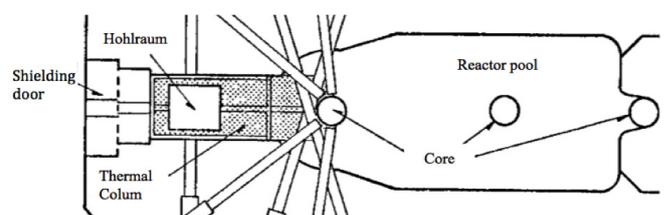


Fig. 1. Longitudinal view of the reactor.

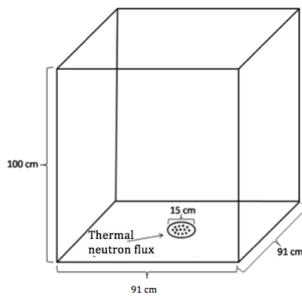


Fig. 2. Hohlraum.

man systems. It has a large number of simple neurons linked by a net of weighted connections. Each neuron takes input from other neurons and produces a single scalar output that depends on the available local information internally stored or arriving from the incoming connections as weights[4].

Some uses of ANNs are [5]: image classification, voice synthesis, sonar echoes classification, knowledge based systems, information coding, and many other classification and perception problems. They have also been used for different optimization problems, such as passenger transit time reduction in airports [6], artificial climate control and optimization in hotels [7], control and optimization of eolic turbines [8], among others.

In this paper, a study to estimate the neutron flux with respect to their energy of the thermal column in the Hohlraum of the TRIGA MARK III nuclear reactor using an ANN is presented. This study will be used for remodeling the facilities containing the reactor, as well as for evaluating the BNCT in patients suffering from brain tumors.

The overall organization of the paper discusses some basic terminology of thermal neutrons in Section II. Section III details the ANN model used, Section IV describes the methodology used in our study. Section V shows the experimental results of the work, and Section VI lists the paper conclusions and some possible future research work.

## II. THERMAL NEUTRONS

The neutrons produced by a nuclear reactor are unstable when they are outside the nucleus, and decay through the beta emission yielding a proton, an electron, and an antineutrino. The average decay time is 12.8 minutes. Depending on the energy groups or the given magnitude of energy, it is possible to find different types of neutrons as shown in Table I.

TABLE I  
NEUTRON ENERGY CLASSIFICATION

Type	Energy (eV)	Velocity (cm/s)
Cold	0.005	$9.66 \times 10^4$
Thermal	0.025	$2.2 \times 10^5$
Epithermal	1	$1.4 \times 10^6$
Slow	$10^2$	$1.4 \times 10^7$
Intermediate	$10^4$	$1.4 \times 10^8$
Fast	$10^6$	$1.4 \times 10^9$
Ultrafast	$10^8$	$1.4 \times 10^{10}$

- 1) Thermal Neutrons. Neutrons in thermal equilibrium with the medium they are interacting with in terms of their energy, which is much lower than those of the fast neutrons initially produced by fission.

- 2) Resonance or Intermediate Neutrons. They have average kinetic energy greater than the thermal ones, found in the 0.5 eV a 10 KeV interval.
- 3) Fast Neutrons. A neutron with a kinetic energy between 10 KeV and 10 MeV.
- 4) Relativistic Neutrons. The highest energy neutrons. They move at relativistic speeds (the speed of light) with energy greater than 10 MeV.

This study will be focused on the thermal neutrons. The speed distribution of thermal neutrons practically follows the Maxwell distribution, typical in the kinetic theory of gases [9]. The most probable speed has been shown to be:

$$v_o = \frac{2KT}{m_n}$$

where  $K$  is the Boltzmann constant,  $T$  is the absolute temperature, and  $m_n$  is the neutron mass. For  $t = 20 \text{ }^\circ\text{C}$ ,  $v_o$  yields a value of 2.200 m/s [9].

For the absorption case, it is defined as the lost of the neutron that becomes part of the new nucleus forming a composite nucleus. The absorption of a neutron could lead to a posterior fission process called radioactive capture forming radioisotopes or triggering the emission of loaded or neutral particles [10].

The neutrons are projectiles capable of inducing fission reactions in some elements that can emit or absorb the released energy in heavy nucleus. Such nuclear reactions are amongst the most important due to their various applications [3]. The most important properties of these processes are the following:

- 1) Fission is an atypical nuclear reaction in which a nucleus, usually heavy, is divided into two fragments, releasing 2 or 3 neutrons and a large amount of energy (about 200 MeV per fission).
- 2) Thermal neutrons are capable of inducing fission in heavy nucleoids of an odd mass number, such as  $^{235}\text{U}$  and  $^{239}\text{Pu}$ . Fast neutrons are capable of causing fission in some heavy nucleus, although with reduced probabilities or efficient sections.
- 3) The resulting fragments of the nucleus fission have many neutrons and are primary links of the beta negative disintegration chain.
- 4) The fact that on each fission process induced by the absorption, 2 a 3 neutrons are released, suggests the possibility that a chain reaction could be caused.

## III. ARTIFICIAL NEURAL NETWORK

One of the most popular models of an ANN is the Multilayer Perceptron due to its applicability on multi-class problems non-linearly separable using several neuron layers, i.e., the output of each neuron from one layer is an input to each neuron in the next layer, and the size of the input layer determines the size of the input pattern vector. The output layer will contain the number of neurons, i.e., classes, to classify[11].

The Back-propagation algorithm is commonly used to train the Multilayer ANNs [12]. The purpose of this algorithm is to adjust the weights in order to minimize the Mean Square Error (MSE) used to quantify the difference between the ANN output and the expected output given by

a set of training patterns. The whole process involves two stages, namely, *Forward propagation* and *Back propagation*.

#### A. Forward propagation

When a pattern  $x$  of  $n$  features is presented as input of an ANN, it is propagated through each weight  $w_{ji}$ , from the input layer to the hidden layer. In this way, the total input that a hidden neuron  $j$  takes is  $net_j = \sum_{i=1}^n w_{ij} \cdot x_i + \theta$ ; where  $x_1, x_2, \dots, x_n$  are the input signals;  $w_1, w_2, \dots, w_n$  are the synaptic of the neuron  $j$  and  $\theta$  is the *bias* (usually set to 1).

On the other hand, the output value of a hidden neuron  $j$ ,  $Z_j$ , is computed using an *activation function*  $f(\cdot)$  over its total input:  $Z_j = f(net_j)$ ; where  $f(net)$  is the sigmoid function:

$$f(net_j) = \frac{1}{1 + \exp[-(net + \theta)]} \quad (1)$$

Similarly, the total input that takes an output neuron  $t$ ,  $net_t$ , is defined as:

$$net_t = \sum_{j=1}^c w_{ij} \cdot Z_j + \theta_j \quad (2)$$

Lastly, the value of the output neuron  $t$ ,  $Z_t$ , is:

$$Z_t = f(net_t) \quad (3)$$

#### B. Back propagation

The minimization of the MSE is calculated using a gradient descent method that computes the gradient of the MSE of the link weights. The error function to be minimized for each pattern  $x$  is as follows:

$$E_x = \frac{1}{2} \sum_{t=1}^c (s_t - Z_t)^2 \quad (4)$$

Where  $Z_i$  is the desired output for the neuron  $t$  after observing pattern  $x$ , and yields 1 if  $x \in c$ , or 0 otherwise. From this expression, a general error measure can be computed using the following equation:

$$E_{avg} = \frac{1}{m} \sum_{x=1}^m E_x \quad (5)$$

To adjust the weights values, the back propagation algorithm uses the gradient descent [13]. The gradient moves in the direction of the fastest error increment, while the opposite direction (negative) determines the fastest error decrement. Taking into account that  $E_x$  is a function of all the weights of the ANN, the gradient of  $E_x$  is a vector equal to the partial derivative of  $E_x$  with respect to each weight. Thus, the error can be reduced adjusting each weight in the following direction:

$$\Delta w_{ij} = -\eta \frac{\delta E_x}{\delta w_{ij}} \quad (6)$$

where  $\eta$  is the learning rate. The value of  $\eta$  has a crucial role during the training process of the ANN because it controls the change size of the weights on each iteration. Therefore, the adjustment of the weights is computed by

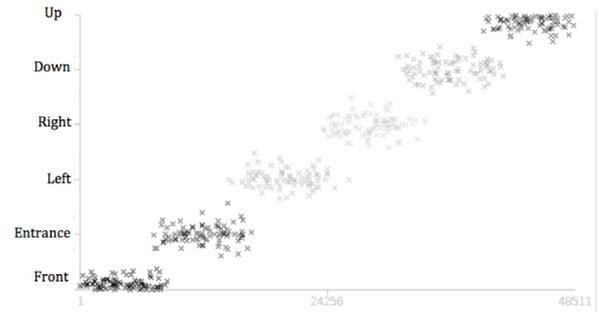


Fig. 3. Data distribution. The X axis corresponds to the pattern values and the Y axis corresponds to the classes under consideration.

$\Delta w(I + 1) = \eta \cdot \delta \cdot Z$ , where  $I$  is the iteration,  $Z$  the output of the neuron and  $\delta$  is the local gradient defined as:

$$\delta = \begin{cases} (s_t - Z_t) f'(net_t) & \text{output layer neurons} \\ f'(net_i) Z_i \sum_t \delta_t w_{t,i} & \text{hidden layer neurons} \end{cases} \quad (7)$$

It can be seen that the delta value associated to a neuron of the hidden layer  $j$  is determined by the sum of the errors that take place in the  $t$  output neurons that take as input the output of the hidden neuron  $j$ . Thus the algorithm's name *back propagation*.

## IV. METHODOLOGY

#### A. Data set

The data set used for training the ANN, consists of setting a grid on the 6 flat surfaces of the Hohlraum. Each surface was divided in 90 columns and 90 rows of 90x90 cm, which represent all the space under study of the Hohlraum. In this way, a balanced data set of 6 classes was obtained (front, intake, left, right, up, down), each with 90 patterns. Figure 3 shows the pattern distribution in the representation space where a disjoint distribution of the classes can be observed because each one corresponds to one of the Hohlraum sides.

For validation purposes, the *k-cross-fold-validation* [14] with 5 repetitions was used. We consider 80% of the patterns for training purposes of the ANN, and the remaining 20% for testing. With this model of the Hohlraum, the experiments followed the steps shown in Figure 2.

#### B. ANN configuration

The ANN configuration used in this study uses the Backpropagation training algorithm [12], considering 5000 iterations and a momentum factor of 0.2, i.e., the learning starts with an initial phase of forward propagation and a second phase of backward propagation of the error, which yields the synaptic weights and reduces the MSE. In the hidden layers the neurons use a transference sigmoid function (logsig) and a value of 0.05 as the learning rate. The number of hidden layers varied from 0 to 5.

## V. EXPERIMENTAL RESULTS

The experimental results shown in this section were obtained using the WEKA software (Waikato Environment for Knowledge Analysis) [15].

A confusion matrix was used to analyze the ANN results [16]. A confusion matrix is a table layout that shows the

classification performance of the ANN with respect to the test data. It is a 2D matrix, indexed in one dimension by the true class of the data and in the other by the response of the classifier under scrutiny [17]. A location  $m[i, j]$  of the confusion matrix indexes the data points from surface  $i$  of the Hohlraum that have been classified as hitting or striking surface  $j$ . Thus, all correct guesses are located in the diagonal of the table, and the classifier errors are represented by the values outside the diagonal.

A number of measures of classification performance are defined in terms of the confusion matrix values, namely, the Kappa Statistic, the General Precision, the Recall and the ROC curve.

The *Kappa Statistic* is a metric that compares the *observed accuracy* with an *expected accuracy*, and also compares the agreement among different ANN models with the accuracy that could be observed by random chance. When all test cases are correctly classified the maximum value agreement of the index is achieved, i.e.,  $Kappa = 1$ . When the values of the coefficient is 0, the observed agreement is considered the same as the observed by random chance.

On the other hand, the *General Precision* is the number of correctly classified test cases for each of the classes considered. In its formulation, two cases are involved: the *true positives* (TP) and the *false positives* (FP). The TP are actual neutron surface strikes that were correctly classified, whereas the FP are strikes that were incorrectly labeled as correct surface hits. Thus,

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

The *Recall* is the ratio of test cases that are correctly classified,

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

Lastly, the f-measure represents the test's accuracy and represents the weighted average of the precision and recall. The closest a f-measure is to 1, the higher is the accuracy of the classifier.

$$f - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (10)$$

Table II summarizes the average results of different ANNs models. Table II shows that a bigger number of hidden neurons yields a better classification of the neutron surface strike. Similarly, as the General Precision increases, the ROC curve achieves a higher agreement when using 5 hidden neuron in the ANN.

A closer look to the FP with respect to the precision levels reveals something interesting. On one hand, the general precision increases as more hidden layers are added to the ANN, however, FP remains stable, mainly when adding 0 to 2 hidden layers. This suggests that despite the FP, the ANN achieves an increasing classification rate. This behavior is confirmed by analyzing the MSE among the different ANN models (Figure 4).

Lastly, the ROC curve analysis yields information to select a possible optimal model and, thus, avoid sub-optimal models regardless of the cost of the distribution of the objects been classified. The ROC curve is also independent of the class

TABLE II  
CLASSIFICATION RESULTS

ANN hidden layers	TP	FP	Area under the ROC curve	General Precision	f-measure	Kappa statistic
0	0.674	0.065	0.938	67.40	0.653	0.6089
1	0.665	0.067	0.945	66.48	0.646	0.5978
2	0.967	0.060	0.979	84.81	0.844	0.8178
3	0.848	0.03	0.977	87.04	0.866	0.8444
4	0.944	0.011	0.990	94.44	0.943	0.9330
5	0.983	0.003	0.999	98.33	0.983	0.9800

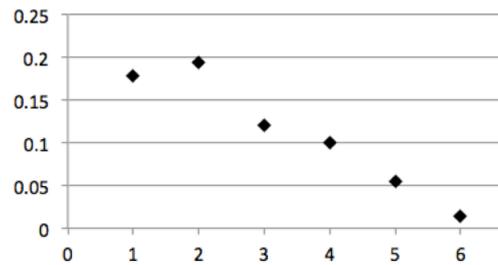


Fig. 4. Observed error by the ANNs models.

distribution. Figure 5 shows the ROC curve that correctly classified neutrons hitting the left wall using a ANN model with 5 hidden layers.

## VI. CONCLUSIONS AND FUTURE WORK

Despite the high potential that the thermal neutrons produced with the TRIGA MARK III reactor have for brain tumor treatment, it is necessary to carry out a study that shows their behavior when going into the Hohlraum, i.e., the space destined for the irradiation. In this sense, the profile identification of the thermal neutrons flux is the starting point to validate its use at the ININ facilities. This paper has proposed the use of an ANN to identify this profile. In addition, a 2D model of the Hohlraum has been designed to validate the results. With the Hohlraum model, the necessary representation and data set for training were presented to the ANN. Once the abstract classification of the thermal neutron flux has been made, an acceptable recognition rate was achieved, and it was observed that a larger the number of hidden neurons causes a more accurate class prediction.

It will be interesting to compare these results with other classification techniques, as well as designing a 3D model of the Hohlraum to simulate the displacement of the thermal neutrons flux.

## REFERENCES

- [1] R. Histrica, "Instituto nacional de investigaciones nucleares," *Contacto Nuclear, ININ*, vol. 3, no. 4, pp. 22–35, 2011.
- [2] Reactor, "Instalaciones del reactor triga mark iii," *Contacto nuclear, ININ*, vol. 1, no. 3, pp. 22–31, 2009.
- [3] G. F. Knoll, *Radiation, detection and measurement*, J. W. . Sons, Ed. USA: John Wiley & Sons, 2011.
- [4] C. P. Ponce, *Inteligencia Artificial con aplicaciones a la ingeniera*, Alfaomega, Ed. Mxico: Alfaomega, 2010.
- [5] O. X. Basogain, *Redes neuronales artificiales y sus aplicaciones*, E. Escuela superior de ingeniera de Bilbao, Ed. Espaa: Escuela superior de ingeniera de Bilbao, EHU, 2006.

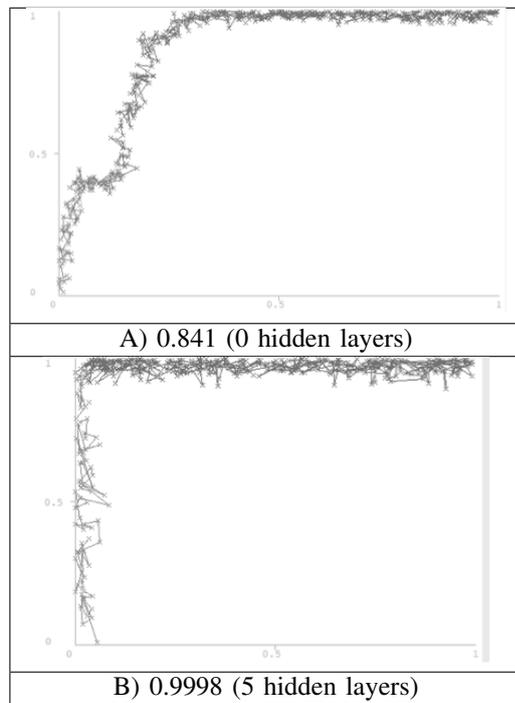


Fig. 5. ROC curve of the left surface. A) area under the curve using an ANN with 0 hidden layers. B) area under the curve using an ANN with 5 hidden layers.

- [6] A. Aragn, S. Casado, and J. Pacheco, "Uso de redes neuronales para optimizar simulaciones en un modelo de flujo de pasajeros en aeropuerto," *Actas de Congreso ASEPUMA*, vol. 10, no. 7, pp. 0–0, 2002.
- [7] H. S. Montelier, N. A. Borroto, T. M. de Armas, S. J. Gmez, T. C. Prez, and L. O. Goza, "Estimacin de cargas trmicas de climatizacin de hoteles mediante simulacin y redes neuronales artificiales," *Ingeniera energtica*, vol. XXXI, no. 3, pp. 13–18, 2010.
- [8] O. Lpez, D. Rojas, M. Vilaragut, and A. Costa, "Simulacin de aerogeneradores de velocidad y paso variable utilizando redes neuronales artificiales," *Ingeniera energtica*, vol. XXXI, no. 1, pp. 35–42, 2010.
- [9] J. N. Dodd, *Atoms and Light: Interactions*, P. Press, Ed. New York: Plenum Press, 1991.
- [10] A. Haroldde, *The elements of neutron interaction theory*, T. M. Press, Ed. USA: The Mit Press, 1971.
- [11] M. G. Pajares and P. M. Santos, *Inteligencia artificial e ingeniera del conocimiento*, Alfaomega, Ed. Mxico: Alfaomega, 2006.
- [12] J. L. McClelland, *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs and Exercises*, P. Books, Ed. Cambridge, MA, EUA: Press/Bradford Books, 2001, vol. 1.
- [13] D. E. Rumelhart, J. L. McClelland, and P. R. Group, *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*, M. Press, Ed. Cambridge, MA: MIT Press, 1986.
- [14] B. Parmanto, P. W. Munro, and H. R. Doyle, "Improving committee diagnosis with resampling techniques," *Advances in Neural Information Processing Systems*, vol. 1, no. 2, pp. 882–888, 1996.
- [15] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques, 2nd Edition.*, M. Kaufmann, Ed. San Francisco, USA: Morgan Kaufmann, 2005.
- [16] R. Congalton and K. Green, *Assesing the Accuracy of Remotely Sensed Data: Principles and Practices*, C. P. Tylor and Francis, Eds. USA: CRC Press/Tylor and Francis, 2009.
- [17] K. Ting, "Confusion matrix," in *Encyclopedia of Machine Learning*, C. Sammut and G. Webb, Eds. Springer US, 2010, pp. 209–209.