International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 26 August 2023 **Revised:** 20 October 2023 **Accepted:** 02 November 2023*

Classification of Cotton Leaf Diseases using Whales Optimization Algorithm based Deep Neural Network

[1] T. Kalaiselvi and [2] V. Narmatha
[1] Research Scholar, [2] Assistant Professor
Department of Computer and Information Science, Faculty of Science, Annamalai University, Annamalai nagar, Tamilnadu, India-608002.
[1]Email: trkalaiselvi82@gmail.com [2]Email: balaji.narmatha8@gmail.com

Abstract: In the realm of agriculture, a current focal point of research revolves around the identification of plant diseases through leaf imagery. Employing image processing techniques for the recognition of agricultural plant diseases holds promise for reducing farmers' reliance on safeguarding their crops. This paper introduces a novel approach to classify cotton leaf diseases, utilizing a Deep Neural Network enhanced by the Whales Optimization Algorithm (WOA). The dataset comprises 10,000 cotton images sourced from Kaggle.com, directly captured from farm fields, covering healthy leaves, bacterial blight, Anthracnose, Cercospora leaf spot, and Alternaria diseases. Preprocessing involves the application of a median filter to eliminate image noise, and for segmenting diseased and normal regions, the Gustaffson-Kessel (G-K) fuzzy clustering method is employed. The WOA-augmented DNN demonstrates its effectiveness in classifying cotton images.

Keywords: Bacterial Blight, Anthracnose, Cercospora leaf spot, Alternaria,, Clustering, Neural Network.

1 Introduction

Agriculture stands as one of the oldest known professions, it continues to provide livelihoods for numerous families worldwide [1]. India, with its rich agricultural heritage, still relies on agriculture as the primary source of income for 58% of its population [2]. A report from the Food and Agricultural Organization (FAO) reveals that plant infections and pest attacks result in the loss of 20-40% of crops. Traditional methods for plant and leaf health assessment, such as flow cytometry, immunofluorescence, polymerase chain reaction, thermography, and chromatography [3], are often inconsistent, resource-intensive, and time-consuming, making them ill-suited for the current agricultural landscape. Given that a significant portion of agriculture occurs in rural areas where expert guidance is scarce, machine vision surpasses human observation [4]. The integration of automated technologies into agriculture holds the potential to yield remarkable results and usher in a new era of farming.

India ranks as the world's second-largest cotton producer, yet its substantial cotton production is consistently hampered by various diseases, including bacterial, viral, and parasitic types. Notable cotton diseases encompass bacterial blight, powdery mildew, leaf curl, boll rot, fusarium wilt, cercospora, anthracnose, alternaria, among others [5]. The application of image processing and deep learning techniques holds promise for the early detection of these ailments. Wang, Y., et al., in their study [6], introduce a novel approach to cotton disease recognition, leveraging color subtraction data from different cotton components. To further enhance accuracy in cotton identification, dynamic Freeman chain coding is employed for noise reduction.

Papageorgiou,E.I et.al., [7], introduced an innovative modeling and simulation approach that utilizes the Fuzzy Cognitive Maps (FCMs) soft computing technique to tackle the challenge of predicting crop yields. This new modeling methodology, employing FCM, was applied to the intricate process of site-specific management, offering an advanced knowledge representation and processing method capable of addressing the unique characteristics and management behaviors specific to cotton crop yields, thereby creating a model that is both interpretable and transparent.

Viraj A. Gulhane et.al., [8], proposed a color image segmentation technique to extract color features from cotton leaves. This technique offers a straightforward approach to extract various features from images of diseased cotton leaves. Subsequently, unsupervised Self-Organizing Feature Maps (SOFM) and backpropagation neural networks are employed to cluster the resulting color pixels and extract the color information of cotton leaves from the diseased portions of the images, respectively.

Revathi, P., and Hemalatha, M., [9], outline a two-stage approach for identifying disease-affected areas. The initial phase involves edge-detection-based image segmentation, followed by image analysis and the classification of cotton diseases using the Homogeneous Pixel Counting Technique for Cotton Diseases Detection (HPCCDD) Algorithm. Azath M. et.al., [10] developed a model to detect cotton leaf

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 26 August 2023 **Revised:** 20 October 2023 **Accepted:** 02 November 2023*

diseases and pests utilizing deep learning techniques and Convolutional Neural Networks (CNN). The model encompasses common cotton leaf diseases and pests like bacterial blight, spider mite, and leaf miner, achieving an impressive 96.4% accuracy in identifying these categories of leaf diseases and pests in cotton plants. The research utilized 2,400 specimens for training purposes. Zekiwos et.al., [11] focused on enhancing a model through deep learning techniques for the detection of common cotton leaf diseases and pests, including bacterial blight, spider mite, and leaf miner. They employed a K-fold cross-validation strategy to split the dataset, enhancing the generalization of the CNN model. This research made use of approximately 2,400 specimens, with 600 images in each class, for training purposes.

Navina Pandhare [12] reported the achievement of an 89% accuracy using deep Convolutional Neural Network (CNN) models. She developed a web-based system tailored for farmers to identify diseases affecting their crops. The study involved a dataset of 520 cotton leaf images. The proposed system resized the input images, extracted relevant features, and employed them for disease classification. In a study by T. Shamyuktha Banu [13], ResNet and VGG16 CNN models achieved disease detection accuracies of 96.2% and 92.5%, respectively, when tested with 3,000 real-time sample images from cotton farms in Madurai.

Mirjalili et.al [14] introduced a novel metaheuristic algorithm inspired by the natural hunting behavior of humpback whales, as discussed in their work. This innovative approach, known as the Whale Optimization Algorithm (WOA), emulates the hunting techniques of humpback whales using bubble nets. In this work, a Gustaffson-Kessel (G-K) fuzzy clustering algorithm is employed for image segmentation, while a Whale Optimization Algorithm (WOA)-enhanced Deep Neural Network is utilized for disease classification. The diseases identified include Bacterial blight, Anthracnose, Cercospora leaf spot, and Alternaria in cotton plants, using a dataset comprising 10,000 images collected from direct capture in the field and Kaggle.com.

2. Proposed System

The system's workflow comprises data collection, preprocessing with median filtering for noise reduction and occlusion removal, segmentation via the G-K fuzzy clustering algorithm, and disease classification using the WOA-DNN classifier. Figure 1 illustrates the proposed process.

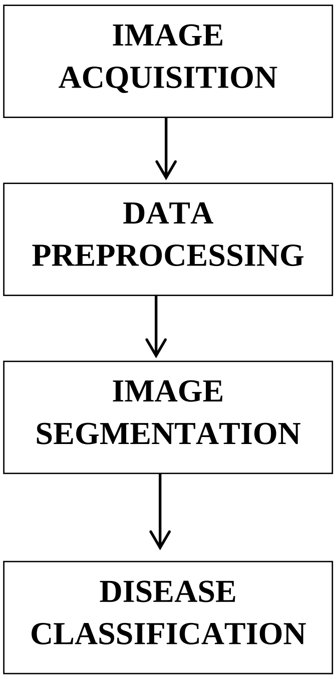

Fig1. Proposed system flow diagram.

The dataset used here is collected from Kaggle.com which contains Bacterial blight, Anthracnose, Cercospora leaf spot, Alternaria alternata and healthy images.

3. Gustaffson-Kessel (G-K) Fuzzy Clustering Algorithm

The G-K algorithm is same as that of FCM, only the computation of distance measure is calculated using co-variance matrices as follows:

Step(1) Initialization

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

Determine the number of clusters (c) you want to identify within the dataset and Initialize the cluster centers randomly or using another suitable method.

Step(2) Membership Function Calculation

Calculate the degree of membership for each data point to each cluster center. This involves assigning a membership value between 0 and 1 to each data point for each cluster based on its similarity to the cluster center.

Step(3) Updating Cluster Centers

Calculate the updated cluster centers by using the membership values. The new cluster centers are computed as weighted averages of the data points, where the weights are the membership values.

Step(4) Convergence Check

Check for convergence by evaluating whether the cluster centers have changed significantly between iterations. If they have, go back to step 2. If not, the algorithm has converged.

Step(5) Termination

End the algorithm when convergence is achieved. The cluster centers represent the final centroids of the clusters.

4. Whale Optimization Algorithm (WOA)

The Whale Optimization Algorithm (WOA) emulates the hunting behavior of humpback whales. These magnificent creatures are renowned as the largest mammals globally, with adult whales reaching lengths of up to 30 meters and weighing as much as 180 tons. There are seven main species of whales, including killer whales, Minke whales, Sei whales, humpback whales, right whales, finback whales, and blue whales. Whales are primarily known as predators and have a distinctive behavior of never fully sleeping, as they need to surface for breathing. Remarkably, only half of their brain sleeps at a time. Whales are intriguing due to their high intelligence and emotional capacities, as noted by Hof and Van Der Gucht [16].

Whales exhibit both solitary and group living patterns, with group behavior being more commonly observed. Some species, such as killer whales, maintain family units throughout their entire lifespans. Among the largest baleen whales, humpback whales stand out. What makes humpback whales particularly intriguing is their distinctive hunting method known as "bubble-net feeding." This foraging technique involves the creation of distinct bubbles in a circular or '9'-shaped pattern, primarily used to hunt schools of krill or small fish near the water's surface. It's important to note that this unique bubble-net feeding behavior is exclusive to humpback whales. In this study, the mathematical modeling of the spiral bubble-net feeding maneuver is employed for optimization purposes.

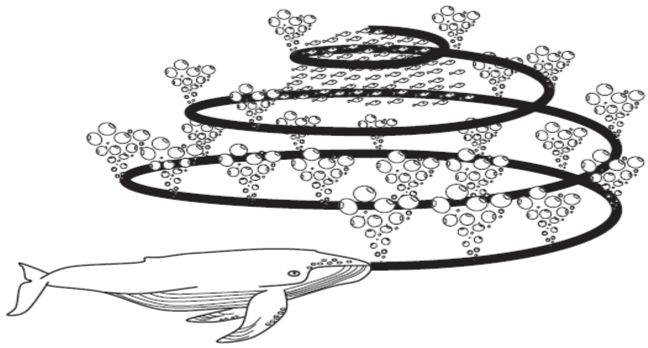

Fig. 2. Bubble-net feeding behavior of humpback whales.

The steps involved in WOA are as follows,

Step 1: Commence by initializing the optimization process with random weight values for hidden neurons in the hidden layers.

Step 2: Calculate the fitness function, which aims to minimize the Mean Square Error (MSE) using the following equation.

$$MSE=\sum_{i=1}^{n}\frac{(x_i-y_i)^2}{n} \tag{5}$$

Step 3: Calculate the location for updating, as determined by the following equations, with x_i representing the input data, y_i denoting the output data, and n representing the number of iterations.

$$\vec{U}=\left|\vec{\gamma}.\ \vec{p}_{best}(t_{it})-\vec{p}_{best}(t_{it})\right| \tag{6}$$

$$\vec{p}(t_{it}+1)=\vec{p}_{best}(t_{it})-\vec{\omega}.\ \vec{U} \tag{7}$$

$$\vec{\omega}=2\vec{\mu}.\vec{r}_i-\vec{\mu} \tag{8}$$

$$\vec{\gamma}=2.\vec{r} \tag{9}$$

Here, $\vec{\mu}$ is minimized from 2 to 0 and $\vec{r}_i$ indicates the unsystematic vector in (0,1).

$$\vec{p}(t+1)=\vec{U}.\ e^{\eta\varphi}.\cos(2\pi\varphi)+\vec{p}_{best}(t) \tag{10}$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 26 August 2023 **Revised:** 20 October 2023 **Accepted:** 02 November 2023*

$$\vec{p}(t+1)=\begin{cases} \vec{p}_{best}(i)-\vec{\omega}.\,\vec{U} & \text{if } R<0.5 \\ \vec{U}.\,e^{\eta\varphi}.\;\cos(2\pi k)+\vec{p}_{best}(t) & \text{if } R\geq 0.5 \end{cases}$$

$$\vec{U}=\left|\vec{\gamma}.\vec{p}_{rand}-\vec{p}\right| \quad (11)$$

$$\vec{p}(t+1)=\vec{p}_{rand}-\vec{\omega}.\vec{U} \quad (12)$$

5. Classification using WOA-DNN

Cotton disease classification using the Whale Optimization Algorithm (WOA) combined with a Deep Neural Network (DNN) involves using the WOA to optimize the DNN's parameters for accurate classification of cotton diseases. This technique leverages WOA's optimization capabilities to enhance the performance of the DNN in distinguishing between different cotton diseases based on input data, such as features. The WOA-DNN approach aims to achieve higher disease classification accuracy and can be a valuable tool in agricultural and crop management systems. In the hidden layer the total number of nodes is evaluated by using Eqn. (13).

$$N=\sqrt{a+b}+c \quad (13)$$

The softmax as an activation function is given as,

$$S(y)_i=\frac{\exp(y_i)}{\sum_{j=1}^{n}\exp(y_i)} \quad (14)$$

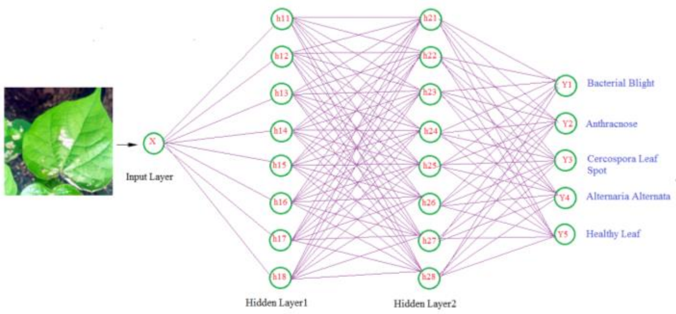

Fig.3 Architecture of DNN with two hidden layers.

The WOA algorithm is applied to determine the optimal weight selection for the Deep Neural Network (DNN). This algorithm works by continuously updating the values, gradually shifting the fitness value towards the best possible solution. Subsequently, a comparison is made between the new and old solutions, with only the best solutions retained for the next iteration. The adjusted solution is then compared to the previous one, and if an improvement is observed, it replaces the previous solution; otherwise, the previous solution is retained. This iterative process continues until the termination criteria are met. The flowchart illustrating the optimal weight selection process is depicted in Figure 4.

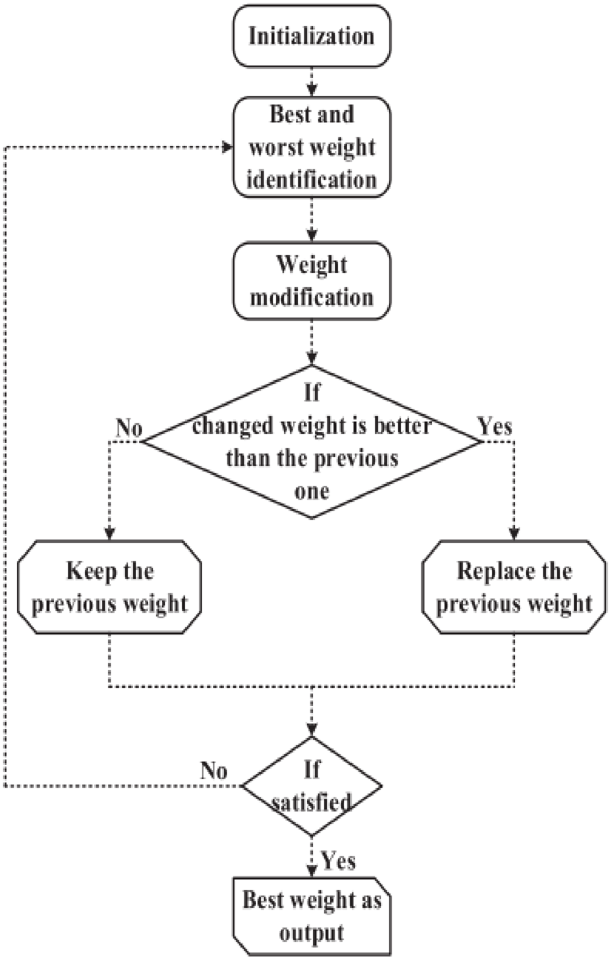

Fig.4. Flow chart of Weight updation using WOA.

6. Experimental Results and Discussion

The output images presented have undergone preprocessing, segmentation, and classification utilizing the previously mentioned proposed algorithms.

Fig.5 (a). Preprocessed Healthy leaf **Fig.5 (b) Segmented image using G-K clustering**

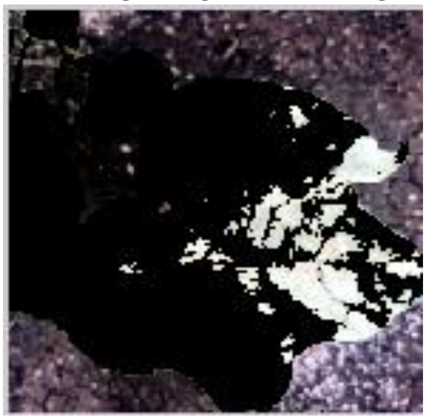

Fig.6(a) Preprocessed diseased leaf **Fig.6 (b) Segmented image using GK clustering**

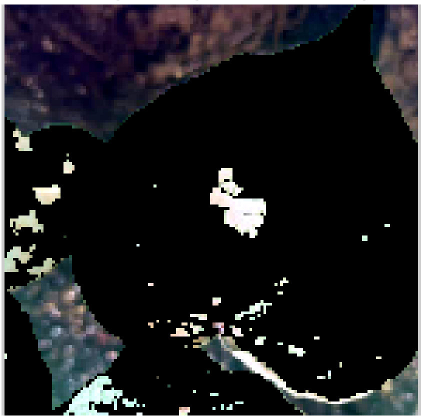

Fig.7(a) Preprocessed diseased leaf **Fig.7 (b) Segmented image using GK clustering**

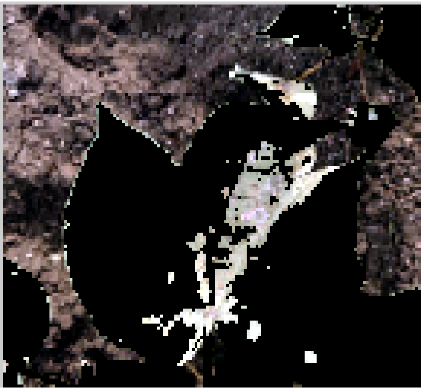

Fig.8(a) Preprocessed diseased leaf **Fig.8 (b) Segmented image using GK clustering**

Fig.9(a) Preprocessed diseased leaf

Fig.9 (b) Segmented image using GK clustering

Fig.10(a) Preprocessed Cercospora Leaf Spot Leaf

Fig.10 (b) Segmented image using GK clustering

Fig.11 (a) Preprocessed Alternaria infected image

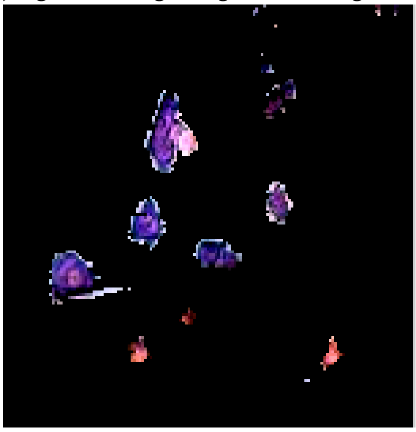

Fig.11 (b) Segmented image using GK clustering

Fig.12(a)Preprocessed Anthracnose infected Leaf

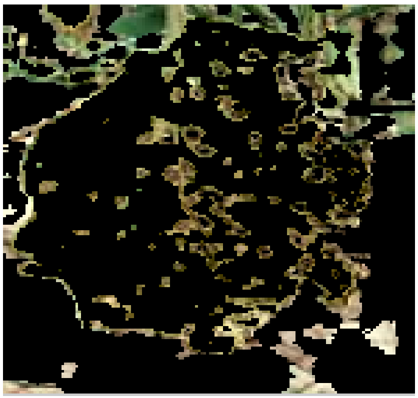

Fig.12(b)Segmented image using GK clustering

4.Classification Performance Metrics

A total of 10,000 images were employed to classify four distinct cotton diseases and healthy leaves. Table 1 presents the classification accuracy. Classification performance metrics are presented in Table 2.

Table 1. Classification Accuracy

S.No	Disease	Samples taken	WOA-DNN Rightly detected	WOA-DNN Accuracy
1.	Bacterial blight	2000	1873	93.65
2.	Anthracnose	2000	1881	94.05
3.	Cercospora Leaf Spot	2000	1875	93.75
4.	Alternaria Alternata	2000	1887	94.35
5.	Healthy Leaf	2000	1912	95.6
	Total	10000	9088	94.28

Table 2. Classification Performance Metrics by implementing WOA-DNN.

S. No	Name of the Disease	Samples	Precision	Recall	F-Score	Specificity
1	Bacterial blight	2000	0.89	0.91	0.90	0.91
2	Anthracnose	2000	0.89	0.89	0.89	0.90
3	Cercospora Leaf Spot	2000	0.90	0.89	0.88	0.92
4	Alternaria Alternata	2000	0.90	0.88	0.91	0.90
5	Healthy Leaves	2000	0.90	0.88	0.89	0.90

5. Conclusion

The proposed system employs G-K clustering for segmentation and WOA-DNN for classification. Experimental results demonstrate the remarkable efficiency of the WOA-DNN classifier, with accuracy reaching 94.28%. Currently, the system processes single-digitized color leaf images as input. Future enhancements may extend its capabilities to handle batches of images encompassing all parts of a plant, thereby improving output predictions.

Reference

1. Bellwood, P. (2023). *First farmers: the origins of agricultural societies*. John Wiley & Sons.
2. Kadiyala, S., Harris, J., Headey, D., Yosef, S., & Gillespie, S. (2014). Agriculture and nutrition in India: mapping evidence to pathways. *Annals of the New York academy of sciences*, *1331*(1), 43-56.
3. Fang, Y., & Ramasamy, R. P. (2015). Current and prospective methods for plant disease detection. *Biosensors*, *5*(3), 537-561.
4. Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2020). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, *9*, 4843-4873.
5. Sharma, O. P., & Bambawale, O. M. (2008). Integrated management of key diseases of cotton and rice. *Integrated management of diseases caused by fungi, phytoplasma and bacteria*, 271-302.
6. Wang, Y., Peng, S., Huang, J., Zhang, Y., Feng, L., Zhao, W., ... & Deng, N. (2022). Prospects for cotton self-sufficiency in China by closing yield gaps. *European Journal of Agronomy*, *133*, 126437.
7. Papageorgiou, E. I., Markinos, A. T., & Gemtos, T. A. (2011). Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application. *Applied Soft Computing*, *11*(4), 3643-3657.
8. Gulhane, V. A., & Gurjar, A. A. (2011). Detection of diseases on cotton leaves and its possible diagnosis. *International Journal of Image Processing (IJIP)*, *5*(5), 590-598.
9. Revathi, P., & Hemalatha, M. (2012, December). Classification of cotton leaf spot diseases using image processing edge detection techniques. In *2012 International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET)* (pp. 169-173). IEEE.
10. Azath, M., Zekiwos, M., & Bruck, A. (2021). Research Article Deep Learning-Based Image Processing for Cotton Leaf Disease and Pest Diagnosis.

11. Zekiwos, M., & Bruck, A. (2021). Deep learning-based image processing for cotton leaf disease and pest diagnosis. *Journal of Electrical and Computer Engineering*, *2021*, 1-10.
12. Pandhare, N., Panchal, V., Mishra, S. S., & Tambe, M. D. Cotton plant disease detection using deep learning.
13. Banu, T. S., & Mani, D. (2020). Cotton crop monitoring system using CNN. *Xi'an Jianzhu Keji Daxue Xuebao/Journal Xi'an Univ Archit Technol*, *12*, 2.
14. Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in engineering software*, *95*, 51-67.
15. Klawonn, F., & Keller, A. (1999). Fuzzy clustering based on modified distance measures. In *Advances in Intelligent Data Analysis: Third International Symposium, IDA-99 Amsterdam, The Netherlands, August 9–11, 1999 Proceedings 3* (pp. 291-301). Springer Berlin Heidelberg.
16. Hof, P. R., & Van der Gucht, E. (2007). Structure of the cerebral cortex of the humpback whale, Megaptera novaeangliae (Cetacea, Mysticeti, Balaenopteridae). *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology: Advances in Integrative Anatomy and Evolutionary Biology*, *290*(1), 1-31.