Fish Expert System-Diagnosis of Streptococcal Disease in Fish Using machine learning "Naive Bayes" Algorithm & PSO Algorithms

Ammar Adl
Faculty of Computers and Information
Beni-Suef University
ammar@fcis.bsu.edu

Sarah Abdelhakeem
Faculty of Computers and Information
Beni-Suef University
sarabahgt_sd@fcis.bsu.edu.eg

abstract:
This paper presents an application of image processing techniques and machine learning algorithm, to facilitate rapid diagnosis and identification of fish diseases and their causes, and therefore early identification of these diseases to reduce their spread and speed of treatment. In this paper we will recognize a disease called "streptococcal disease", this system will be diagnosed if the fish is infected with this disease or not. The methods that the user takes a picture of the fish and system show him if the fish is sick or not.This is done after preprocessing for input images then extract disease area to identify it by feature extraction technique and then calling Naive Bayes Algorithm & PSO(Particle Swarm Optimization) Algorithms to display If the fish is infected by streptococccal disease or not.

Keywords: Streptococcal Disease,Segmentation, Naive Bayes, ORB (Oriented Fast and rotated Brief),PSO(particle swarm optimization algorithm).

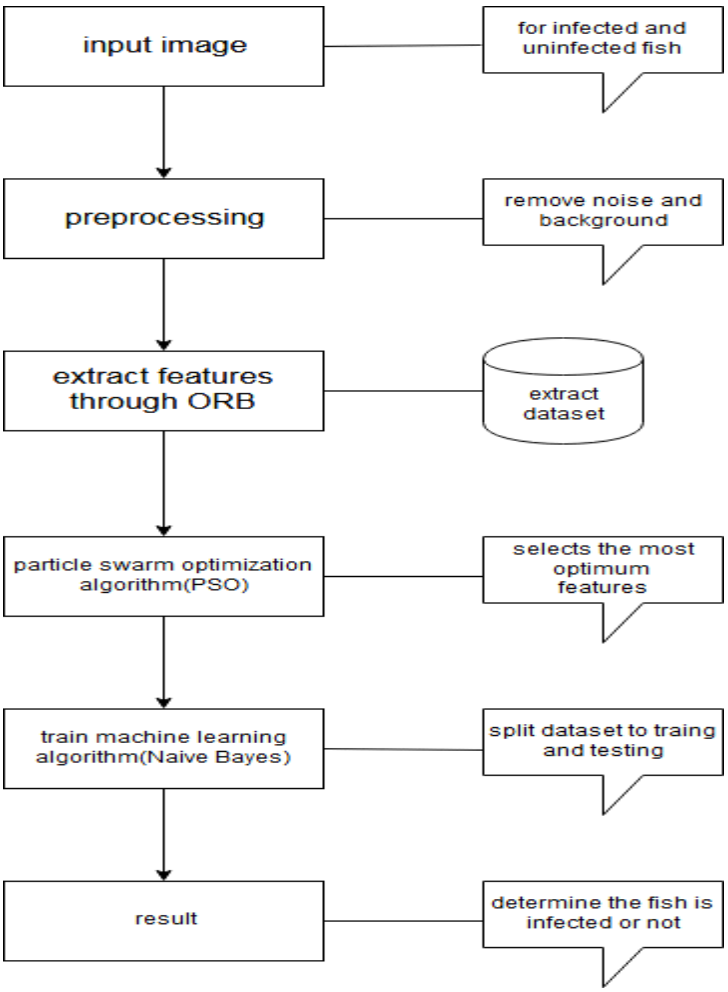
1-Introduction:

Fish is one of the most important vertebrates found throughout the world. Fish is one of the most important food sources in the world, and most countries rely on food as their main source.So these countries are interested in setting up fish farms to raise fish and with the spread of some bacteria that affect fish and cause some serious diseases that lead to the death of these fish as a disease called streptococcal caused by a type of bacteria called Gram-positive (they appear purple/blue when stained using a procedure called a Gram stain). By contrast, most of the common disease-causing bacteria of fish are Gram-negative (appear pink with a Gram stain). This is important when considering treatment options. This extremely serious disease causes the deaths of many fish and also infectious and also common highly in tilapia.When the fish is infected with the disease, it is separated from the rest of the fish and is done in traditional ways that take time and effort, especially in the large fishs farms.so, we offer this System(Fish Expert System) to diagnose fishs diseases and treated them as soon as possible.this system will diagnose streptococcal disease using machine learning algorithm, image processing techniques and feature extraction technique .So to build this system we will follow three steps:Preprocessing for images of diseases fishs using image processing techniques as(remove noise ,remove background) ,Extracting pathogen area from images of infected fish by applying ORB(Oriented Fast and rotated Brief) algorithm and do dataset then

using machine learning Naive Bayes classifier and PSO (particle swarm optimization algorithm) to be able to predict if any input image infected with streptococcal disease or not.This system is evaluated if it is good or not based on accuracy who gets from the classifier algorithm and also the speed of response to the correct diagnosis.this based on the dataset we are get for infected and non-infected fish to learn the algorithm well.

2-Material and Methods:

In this section we will talk about the steps that system is going through to diagnose the streptococcal disease from input images to the output. The first step preprocessing for input images for infected and uninfected fish,then extracting feature from input images to extract dataset and then applying machine learning algorithm by spliting dataset to training and testing data to learn the algorithm we evaluate the algorithm if it is good or not according to accuracy and prcentage of error of it.



Fg.1: the workflow of the total process of the system

2.1-Input Image

Collecting images of infected fish and healthy fish to do classifier by using machine learning algorithm

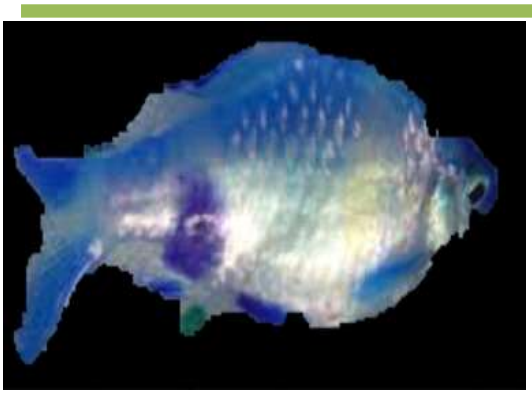



Fig2.1:fish infected by streptococcal disease Fig2.2:healthy fish

2.2:Pre-processing

Any input image do for it preprocssing using image processing technique removing noise ,segmentation(remove background) to detect object(fish)only.




Fig3.1:image before preprocessing Fig3.2:image after preprocessing

2.3:Feature Extraction

Using feature ettraction technique ORB(Oriented Fast and rotated Brief) algorithm.as it ORB technique is better than SIFT, Fast, and HOG feature extraction as the prevalent SURF and SIFT feature descriptors are computationally complex and demand very high level hardware requirements for processing. Off late, ORB, a cost effective feature descriptor, has been effective for such systems.Using feature extraction technique(ORB) to extract feature from object(fish) and identify the affected area for infected fish.first step in feature convert image from color level to gray level and then extract four feature from image (Mean,Median,Standard Deviation, and Variance).and put it in excel file.csv to get dataset.

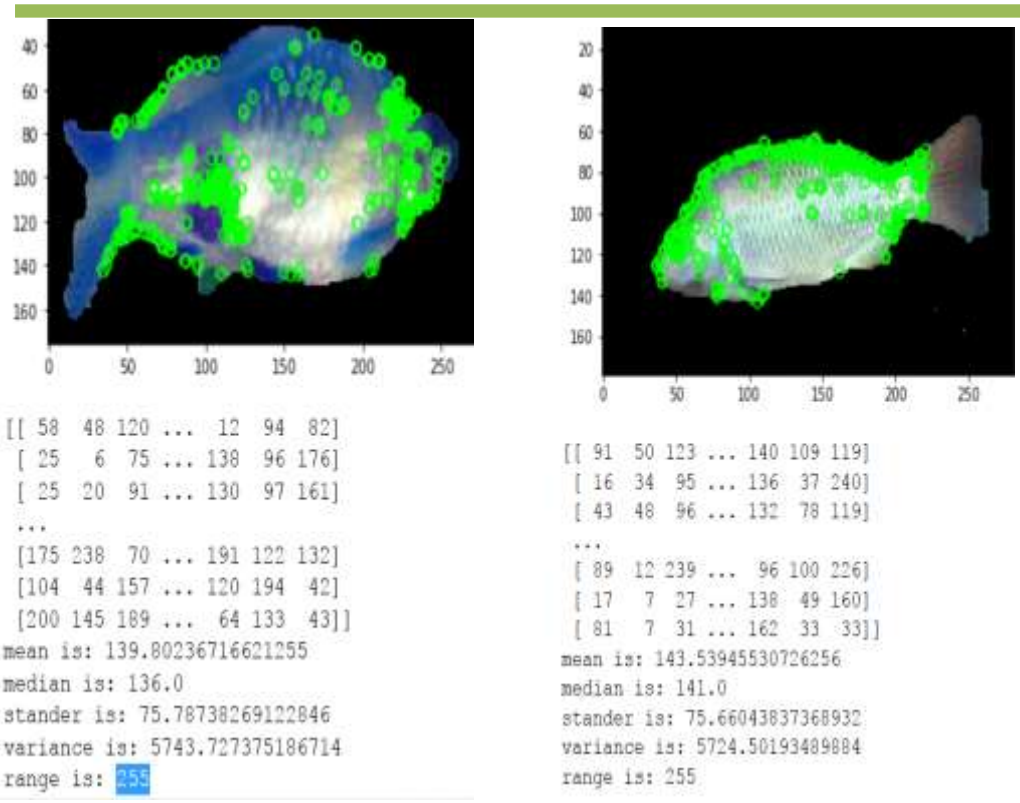


Fig4:applying feature(ORB) to extract the value from image

2.4:Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is an efficient evolutionary computation technique which selects the most optimum features which contribute more to the result which reduces the computation time and increases the accuracy. Experimental result shows that the proposed model with PSO as feature selection increases the predictive accuracy of the Naive Bayes to classify fish disease.

2.5:Machine Learning Naive Bayes Classifier

After extrction five feature from infected and healthy fish images to make classification between the image if infected by strptococcal disease or not and do dataset based on this two class(infected by strptococcal or not)and then applying machine learning Naive Bayes classifier on dataset extracted to train on this dataset and produce the accuracy for the system after then predict if fish is infected by strptococcal disease or not.

3:Dataset

collected dataset from different site for fish images and from YouTube videos and these aren't copy righted for infected by strptococcal disease and healthy fish in this paper .he following table is sample of database that contains extracted features that I extracted from the result of preprocessing step which is infected and healthy samples:

range,variance,stander,mean,median,class
255,5712.026858,75.57795219,135.0802596,128,0
255,5273.166978,72.61657509,164.43822,176,0
255,5737.111493,75.74372247,130.4437023,128,0
255,5477.357396,74.00917103,138.8391029,137,0
255,5522.206815,74.31155237,140.7420233,144,0
255,6184.910817,78.64420396,134.0147377,134,0
255,5879.140597,76.67555411,128.3479056,126,0
255,5602.659118,74.85091261,126.4616678,128,0
255,5614.398145,74.92928763,146.1917873,149,0
255,5799.173121,76.15230214,135.5543515,135,0
255,5671.585897,75.30993226,124.1598125,124,0
255,5745.883939,75.80160908,126.572875,125,0
255,5838.353599,76.40911987,129.08675,128,1
255,5798.47207,76.14769905,136.6621875,136,1
255,5623.394757,74.98929761,121.3338125,119,1
255,5930.945325,77.01263094,133.64975,130,1
255,5522.973693,74.31671207,142.6833813,143,1
255,5743.659029,75.78693178,129.7895625,128,1
255,6057.638641,77.83083348,142.3225625,142,1
255,5602.331266,74.84872254,144.7573492,148,1
255,5821.191075,76.29673043,155.28775,166,1
255,5736.595206,75.74031427,143.5876567,141,1

Fig5: this fig represents samples of extracted features where class(1) for healthy fish and class(0)for infected fish

4-Result:

We evaluate the speed of system in the correct diagnosis of the fish disease based on the accuracy. the accuracy of this system after applying Naive Bayes algorithm is 85%. and we applying this dataset on anthor machine learning(Random ForestClassifer,DecisionTreeClasifier algorithm,SuportVectorMachineClassifier).and we geet the accuracy for all we show it in this Fig.

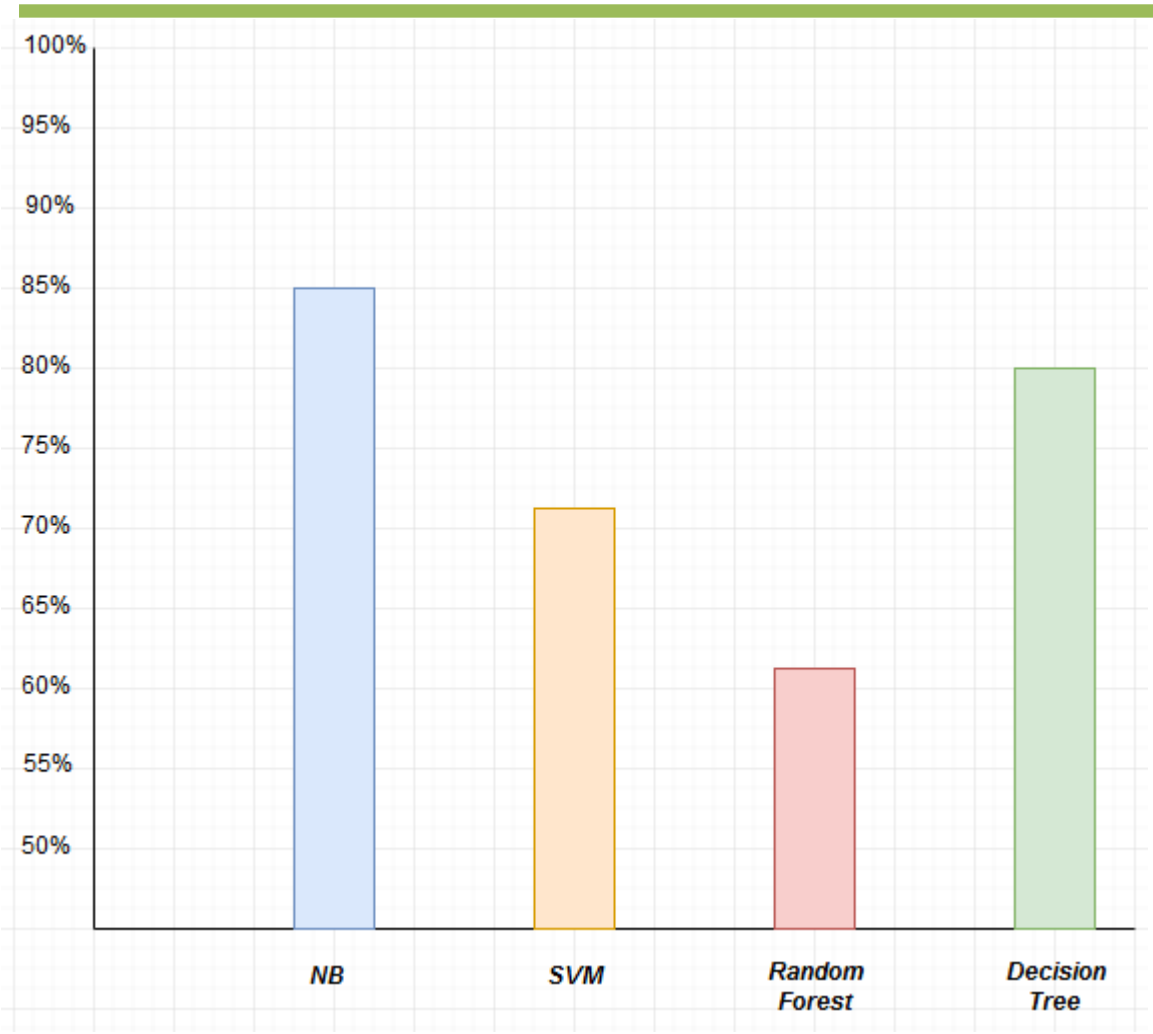


Fig6:the result of (NB, SVM, Decision Tree, Random Forest) Algorithms in evaluation process

this fig show that the Naive Bayes Algorithm is best in classifiction and this will show in confusion matrix.

***Confusion Matrix**

A confusion matrix is a table that is often used to describe the performance of a classification model(Naive Bayes) thet shows 85% accuracy through the Naive Bayes clasifier

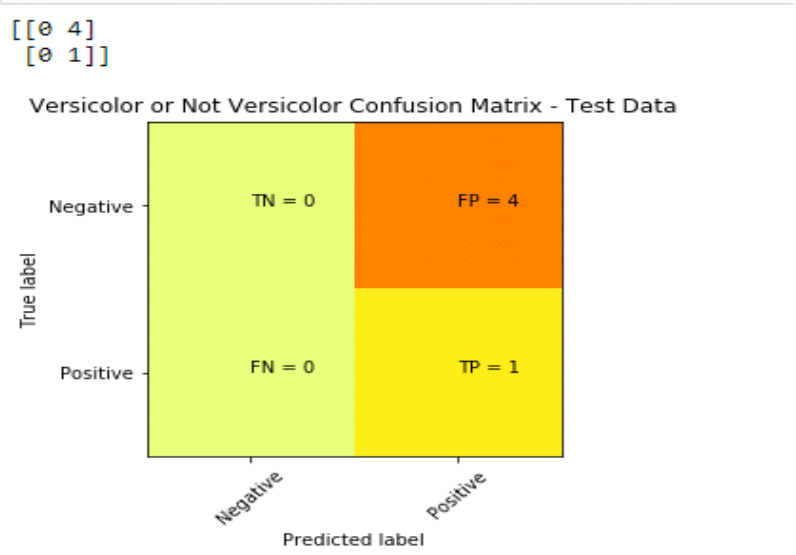


Fig7:confusion matrix of Naive Bayes classifier

*__Classification Report__ of Naive Bayes

Class	Precision	Recall	F1-Score	Support
Healthy	0.80	1.00	0.89	4
Infected	0.00	0.00	0.00	1
Avg/Total	0.64	0.80	0.71	5

Fig8:Classification Report of confusion matrix

5-Conclusion and Discussion:

In this paper we displays To diagnose fish diseases(streptococcal disease)with accuracy 85% by using machine learning algorithm(Naive Bayes)and image processing techinque(segmentation).This system helps us to identify fish diseases(strptococcal disease)quickly,less cost and and effortlessly from traditional methods.For quick treatment and separate them from healthy fish in Fish Farms.

6-Future work:

We will also increase the number of dataset to increase the accuracy more than 85%.We plan to diagnose all fish diseases found in all fish species and also offer solutions for treatment of diseases.and also we will work by deep learning for the system for become a Sensor to be able to take images for the fish under water and identify the patient and the proper and determine the type of disease.

7-References:

1-https://thefishsite.com/articles/streptococcus-in-tilapia
2-Jaglarz, A., Gurgul, A., Leigh, W. J., Costa, J. Z., & Thompson, K. D. (2018). Complete Genome Sequences of Three Streptococcus agalactiae Serotype Ia Isolates Obtained from Disease Outbreaks in Nile Tilapia (Oreochromis niloticus). Genome announcements, 6(1), e01432-17.

3-Jaglarz, A., Gurgul, A., Leigh, W. J., Costa, J. Z., & Thompson, K. D. (2018). Complete Genome Sequences of Three Streptococcus agalactiae Serotype Ia Isolates Obtained from Disease Outbreaks in Nile Tilapia (Oreochromis niloticus). Genome announcements, 6(1), e01432-17.
4-Liu, H., Zhang, S., Shen, Z., Ren, G., Liu, L., Ma, Y., ... & Wang, W. (2016). Development of a vaccine against Streptococcus agalactiae in fish based on truncated cell wall surface anchor proteins. The Veterinary record, 179(14), 359-359.
5-Pasaribu, W., Sukenda, S., & Nuryati, S. (2018). The Efficacy of Nile Tilapia (Oreochromis niloticus) Broodstock and Larval Immunization against Streptococcus agalactiae and Aeromonas hydrophila. Fishes, 3(1), 16.
6-HARDI, E. H., NUGROHO, R. A., SAPTIANI, G., SARINAH, R., AGRIANDINI, M., & MAWARDI, M. (2018). Identification of potentially pathogenic bacteria from tilapia (Oreochromis niloticus) and channel catfish (Clarias batrachus) culture in Samarinda, East Kalimantan, Indonesia. Biodiversitas Journal of Biological Diversity, 19(2), 480-488.
7-Zhu, B., Macleod, L. C., Kitten, T., & Xu, P. (2018). Streptococcus sanguinis biofilm formation & interaction with oral pathogens. Future microbiology, (0).
8-Sirimanapong, W., Thompson, K. D., Shinn, A. P., Adams, A., & Withyachumnarnkul, B. (2018). Streptococcus agalactiae infection kills red tilapia with chronic Francisella noatunensis infection more rapidly than the fish without the infection. Fish & shellfish immunology, 81, 221-232.
9-Farhi, E., & Neven, H. (2018). Classification with quantum neural networks on near term processors. arXiv preprint arXiv:1802.06002.
10-http://bigdata-madesimple.com/10-machine-learning-algorithms-know-2018/
11-https://blog.easysol.net/machine-learning-algorithms-4/
12-https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/
13-Rusland, N. F., Wahid, N., Kasim, S., & Hafit, H. (2017, August). Analysis of Naïve Bayes Algorithm for Email Spam Filtering across Multiple Datasets. In IOP Conference Series: Materials Science and Engineering (Vol. 226, No. 1, p. 012091). IOP Publishing.
14-Mansour, A. M. (2018). Texture Classification using Naïve Bayes Classifier. INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND NETWORK SECURITY, 18(1), 112-120.
15-Liu, J., Yu, J., Wang, C., & Zhang, X. (2018, February). An ORB Feature Matching Algorithm for Mobile Devices. In Proceedings of the 2018 10th International Conference on Computer and Automation Engineering (pp. 116-119). ACM.
16-Karami, E., Prasad, S., & Shehata, M. (2017). Image matching using SIFT,SURF, BRIEF and ORB: performance comparison for distorted images. arXivpreprint arXiv:1710.02726.
17-
]https://docs.opencv.org/3.0beta/doc/py_tutorials/py_feature2d/py_orb/py_orb.html