Development of Algorithms for Leaf Disease Diagnosis using Machine Learning Techniques

Member 1 (Roll No) Member 2 (Roll No)
Member 3 (Roll No) Member 4 (Roll No)

Guide Name
Project Guide

February 17, 2019

Contents

Project Introduction

The agricultural sector is going to face enormous challenges :

- In order to feed the 9.6 billion people that the FAO predicts are going to inhabit the planet by 2050; hence food production must increase by 70% by 2050.
- Due to diseases, the productivity of the crops is affected. Hence the efforts must be taken to diagnosis the disease at early stage.

Introduction

Limitations of existing systems :

- Farmers in rural India have minimal access to agricultural experts, who can inspect crop images and render advice.
- Delayed expert responses to queries often reach farmers too late.

Literature Survey on Project

Literature Survey on Project

Summary Literature Survey on Project

From the above literature survey it is found that the following methods are used by different researchers for plant disease detection and analysis:

- Feature Extraction: -
  - Combination of morphological features of leaves are used
  - Fuzzy surface selection technique for feature selection are used
  - Texture, color and shape features are used
  - New indices based on hyper spectral imagery are used
- Advanced Image Processing Techniques used
  - To detect severity of disease
  - To identify nutrition of plant

Continue Summary of Literature Survey

- Sensing techniques
  - Electronic nose/tounge are used
  - Devices such as Spectrometer,CCD used
  - Airborne hyperspectral imagery
  - Multispectral Imagery
- Classifiers
  - Neural network based classifier such as **Backpropagation**, Probability Based Neural Network, Self Organizing Maps, Deep Learning
  - Hybrid systems using Fuzzy logic, Genetic Algorithms and Artificial Neural Network
  - Support vector Machine

Formation of Problem Statement

Develop a leaf disease diagnosis system using new variants of the back propagation algorithm

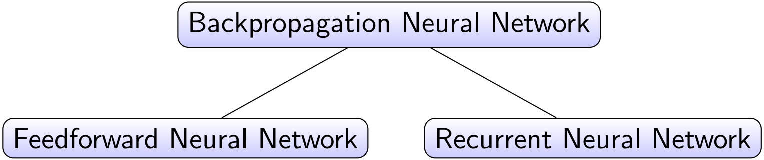
Plan of Research Conduction

- Thoroughly study and implement existing Backpropagation Neural Network and its variations
  - Identify major problems associated with them
  - Propose solution to solve those problem
- Analysis and study various methods to detect leaf diseases in early stages

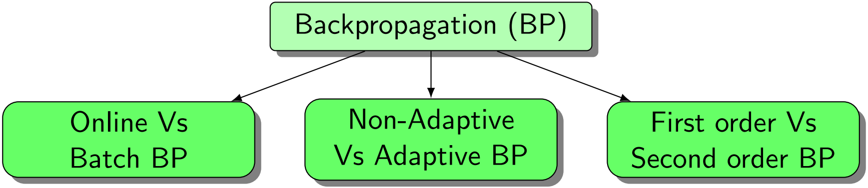
Objectives

1. Study Backpropagation and its existing variants.
2. Implement leaf disease diagnosis system using Backpropagation algorithm.
3. Propose new variants of the Backpropagation algorithm.
4. Implement Leaf disease diagnosis system using the proposed variants of the Backpropagation neural network.
5. Comparative analysis of the proposed algorithms for leaf disease diagnosis system.

Types of Backpropagation Neural Network



Types of Backpropagation Neural Network



Online Backpropagation Algorithm I

$(\mathbf{V}, \mathbf{v_b}) \leftarrow$ initialize with small random values in the range of -0.3 to $+0.3$.
$(\mathbf{W}, \mathbf{w_b}) \leftarrow$ initialize with small random values in the range of -0.3 to $+0.3$.
$\eta \leftarrow 0.4, \quad E_{max} \leftarrow 0.0001, \; SSE \leftarrow 0, \; MSE \leftarrow 0, \; epoch \leftarrow 0.$

---

while $epoch <= 30000$ or $MSE >= E_{max}$ **do**
    $SSE \leftarrow 0$
    for $p := 1 \rightarrow P$ **do**
        $\mathbf{z} \leftarrow p^{th}$ row of $\mathbf{X}$
        for $j := 1 \rightarrow J$ **do**

Online Backpropagation Algorithm II

$$net_j = (\sum_{i=1}^{I} v_{j,i} * z_i) + v_{j,I+1}$$
$$y_j = \frac{1}{1+exp^{-\lambda * net_j}}$$

end for

for $k := 1 \rightarrow K$ **do**

$$net_k = (\sum_{j=1}^{J} w_{k,j} * y_j) + w_{k,J+1}$$
$$o_k = \frac{1}{1+exp^{-\lambda * net_k}}$$
$$e_k = d_k - o_k$$
$$SSE = SSE + \frac{1}{2} * \sum_{k=1}^{K} {e_k}^2$$
$$\delta_{ok} = e_k * f'(net_k) = e_k * o_k * (1 - o_k)$$

end for

for $j := 1 \rightarrow J$ **do**

Online Backpropagation Algorithm III

$\delta_{y_j} = (y_j * (1 - y_j)) * (\sum_{k=1}^{K} \delta_{o_k} * w_{k,j});$

end for

for $k := 1 \rightarrow K$ **do**

$\Delta w_{k,J+1} = \eta * \delta_{o_k}$

for $j := 1 \rightarrow J$ **do**

$\Delta w_{k,j} = \eta * \delta_{o_k} * y_j$

end for

end for

for $j := 1 \rightarrow J$ **do**

$\Delta v_{j,I+1} = \eta * \delta_{y_j}$

for $i := 1 \rightarrow I$ **do**

Online Backpropagation Algorithm IV

$\Delta v_{j,i} = \eta * \delta_{y_j} * z_i$

end for

end for

for $k := 1 \rightarrow K$ **do**

$w_{k,J+1} = w_{k,J+1} + \Delta w_{k,J+1}$

for $j := 1 \rightarrow J$ **do**

$w_{k,j} = w_{k,j} + \Delta w_{k,j}$

end for

end for

for $j := 1 \rightarrow J$ **do**

$v_{j,I+1} = v_{j,I+1} + \Delta v_{j,I+1}$

Online Backpropagation Algorithm V

for $i := 1 \to \ I$ **do**

$v_{j,i} = v_{j,i} + \Delta v_{j,i}$

end for

end for

end for

$MSE = \frac{SSE}{P}$;

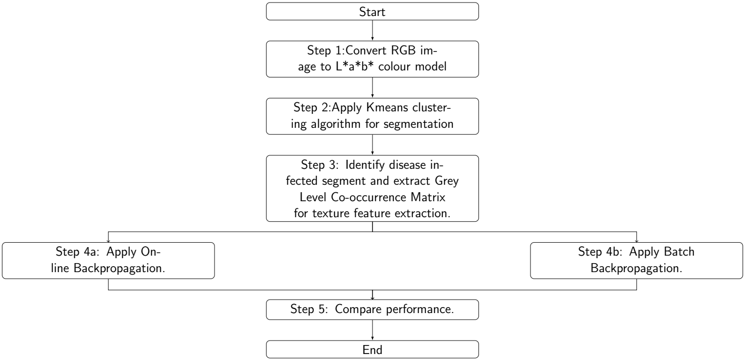
$epoch = epoch + 1$;

end while

Problems of Backpropagation Neural Network

- Slow Convergence Rate
  - Local Minima problem
  - Flatspot problem

Approach 1: Leaf Disease Diagnosis using Backpropagation Neural Network



Performance Analysis of the Leaf Disease Diagnosis system (LDDS) using online and Batch BP

Figure 1: Maximum accuracy obtained in 10 fold validation in 10 runs

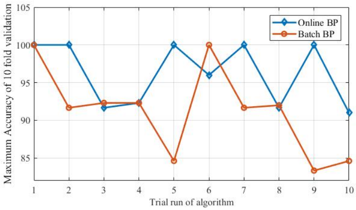
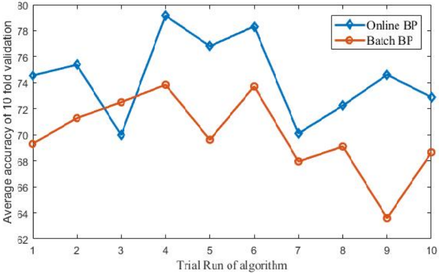


Figure 2: Average accuracy obtained in 10 fold validation in 10 runs



Summary of Approach 1

- The Backpropagation algorithm is well proven for leaf disease classification purpose.
- Here two versions of the Backpropagation i.e. online and batch BP are applied for leaf disease diagnosis system. It is observed that the performance of the online BP is better as compared to the batch BP.
- Here the dataset used is small in size. Further the performance can be verified on large amount of dataset available on plantvillage website.

Publication

Published paper on *Leaf Disease diagnosis using online and batch Backpropagation neural network* in International Journal of Computer Sciences and Engineering, Vol 6, Issue 6.

Approach 2: Modified BP by adding White Gaussian Noise in the Weighted Sum (NNBP)

- In Backpropagation the weighted sum is calculated using equation 1.
- In NNBP the *net* is modified by adding white Gaussian noise using equation 2. Here awgn() function of matlab is used.

$$
\begin{aligned}
net_j &= (\sum_{i=1}^{I} v_{j,i} * z_i) + v_{j,I+1}, \quad where \quad j = 1 \quad to \quad J, \\
net_k &= (\sum_{j=1}^{J} w_{k,j} * y_j) + w_{k,J+1} \quad where \quad k = 1 \quad to \quad K.
\end{aligned}
\tag{1}
$$

The awgn() function is described below.

- awgn(*net*,snr) – Add white Gaussian noise to signal *net*.
- It is a Matlab function which adds white Gaussian noise in *net* with signal to noise ratio (SNR) given as second argument. In equation 2 50 value is used as SNR.

$$\begin{aligned} net_j^{new} &= awgn(net_j, 50) \quad where \quad j = 1 \quad to \quad J, \\ net_k^{new} &= awgn(net_k, 50), \quad where \quad k = 1 \quad to \quad K. \end{aligned} \tag{2}$$

Literature Survey on Noise Injection Techniques in BP

- Noise Benefits in Backpropagation and Deep Bidirectional Training by Karthik Audhkashi, IJCNN 2013
- Kernel Regression and Backpropagation training with noise by Koistinen, Petri, and Lasse Holmström, Advances in Neural Information Processing Systems. 1992.
- Training neural network with additive noise in the desired signal by Chuan wang in IEEE Transaction on NN, 1999
- Convergence Analysis on online weight noise injection based training for MLP by John sum, IEEE Transactions on NN, 2012

Summary of Literature Survey on Noise Injection Techniques in BP

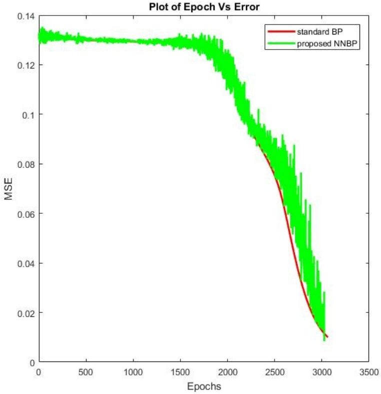
Noise is added in three entity in Backpropagation

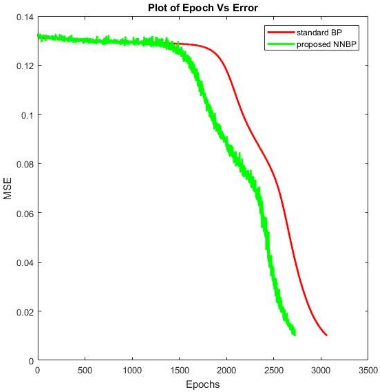
- Input
- Weights
- Desired Output

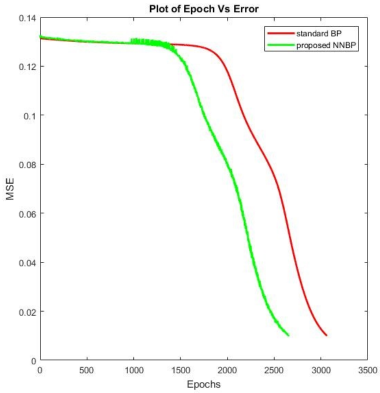
Here we added noise in the weighted sum entity of the Backpropagation.

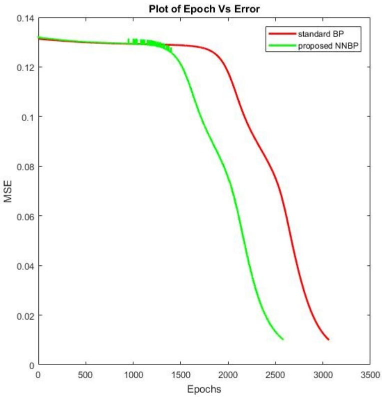
Experimentation on different SNR values

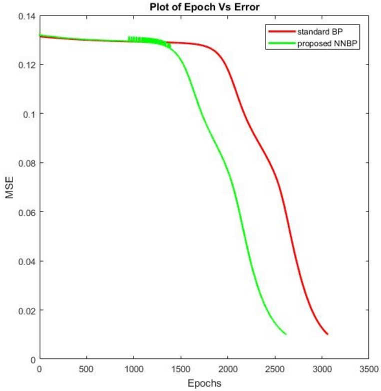
- The NNBP is applied to the parity problem and the IRIS dataset.
- The experimentation is conducted using different SNR values in the range of 10 to 80.
- It is observed that NNBP works better with SNR value in the range of 40 to 70.

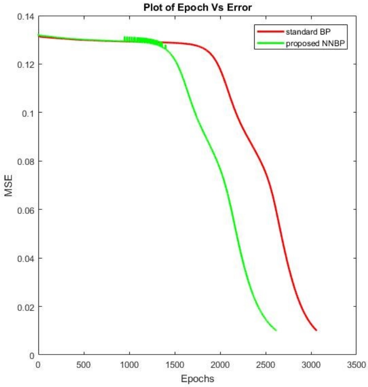

Plot of Epoch Vs Error
standard BP
proposed NNBP
MSE
Epochs
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0
500
1000
1500
2000
2500
3000
3500

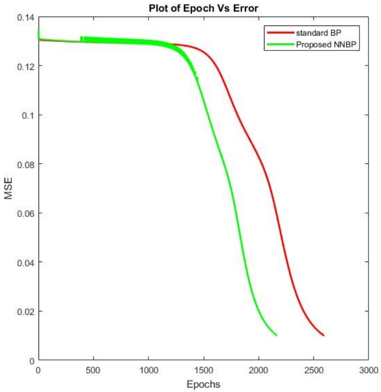

Plot of Epoch Vs Error
standard BP
proposed NNBP
MSE
Epochs
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0
500
1000
1500
2000
2500
3000
3500

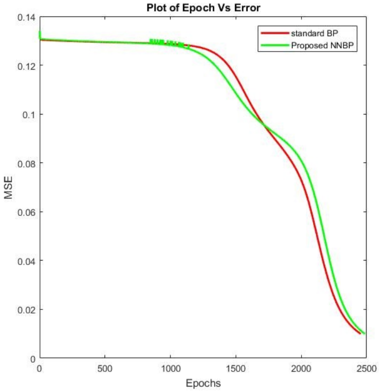

Plot of Epoch Vs Error
standard BP
proposed NNBP
MSE
Epochs
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
500
1000
1500
2000
2500
3000
3500


Plot of Epoch Vs Error
standard BP
proposed NNBP
MSE
Epochs
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
500
1000
1500
2000
2500
3000
3500


Plot of Epoch Vs Error
standard BP
proposed NNBP
MSE
Epochs
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
500
1000
1500
2000
2500
3000
3500


Plot of Epoch Vs Error
standard BP
proposed NNBP
MSE
Epochs
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
500
1000
1500
2000
2500
3000
3500


Plot of Epoch Vs Error
standard BP
Proposed NNBP
MSE
Epochs
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0
500
1000
1500
2000
2500
3000


Plot of Epoch Vs Error
standard BP
Proposed NNBP
MSE
Epochs
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0
500
1000
1500
2000
2500

NNBP applied on IRIS dataset

NNBP is applied on IRIS dataset. The results are given below

Table 1 IRIS dataset results

Algorithm	Average Epochs	Validation Accuracy(%)	Testing Accuracy(%)
BP	6050.0	96.63	95.87
Proposed NNBP	605.2	96.39	95.28

NNBP applied on Leaf disease detection system using Texture Features

NNBP is applied on leaf disease diagnosis system. The results are given below

Table 2 Leaf disease detection system results

Algorithm	Average Epochs	Validation Accuracy(%)
BP	>5000	73.26
Proposed NNBP	>5000	43.50

Summary of Approach 2

- NNBP converges faster as compared to the standard BP when applied on parity, encoder and IRIS dataset.
- In case of IRIS dataset, the accuracy of the BP is better than NNBP.
- BP performance is better as compared to the NNBP, when applied on leaf disease detection system with Texture features.

Publication

Paper Published "Modified Backpropagation with added white Gaussian noise in weighted sum for convergence improvement 8^{th} International Conference On Advances In Computing & Communications, Kochi (ICACC -2018) on 15^{th} September(All published paper in conference are going to be published in Elsevier Procedia Computer Science Journal).

Approach 3:Leaf Disease Diagnosis system using Method 3

- Explain method 3 similar to approach 1 and 2.

$$
\begin{aligned}
net_j &= (\sum_{i=1}^{I} v_{j,i} * z_i) + v_{j,I+1}, \quad where \quad j = 1 \quad to \quad J, \\
net_k &= (\sum_{j=1}^{J} w_{k,j} * y_j) + w_{k,J+1} \quad where \quad k = 1 \quad to \quad K. \\
net_j^{new} &= (max - min) * net_j + min, \quad where \quad j = 1 \quad to \quad J, \\
net_k^{new} &= (max - min) * net_k + min, \quad where \quad k = 1 \quad to \quad K.
\end{aligned} \tag{3}
$$

Approach 4: Leaf disease diagnosis system using Deep learning feature

- In this approach, instead of using texture features calculated using Gray Level Covariance Matrix, features of the leaf images are extracted using pretrained deep learning neural network mode 'Alexnet'.
- The comparative analysis of these two feature extraction technique is done here.
- The Backpropagation Neural Network is used as a classifier.

Table 3 Alexnet Layers Layout-I

Sr No	Label	Layer	Description
1	' data'	Image Input	227x227x3 images with'zerocenter normalization
2	'conv1'	Convolution	96 11x11x3 convolutions with stride [4 4] and padding [0 0]
3	'relu1'	ReLU	ReLU
4	'norm1'	Cross Channel Normalization	cross channel normalization with 5 channels per element
5	'pool1'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0]
6	'conv2'	Convolution	256 5x5x48 convolutions with stride [1 1] and padding [2 2]
7	'relu2'	ReLU	ReLU
8	'norm2'	Cross Channel Normalization	cross channel normalization with 5 channels per element
9	'pool2'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0]

Table 4 Alexnet Layers Layout-II

Sr No	Label	Layer	Description
10	'conv3'	Convolution	384 3x3x256 convolutions with stride [1 1] and padding [1 1]
11	'relu3'	ReLU	ReLU
12	'conv4'	Convolution	384 3x3x192 convolutions with stride [1 1] and padding [1 1]
13	'relu4'	ReLU	ReLU
14	'conv5'	Convolution	256 3x3x192 convolutions with stride [1 1] and padding [1 1]
15	'relu5'	ReLU	ReLU
16	'pool5'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0]
17	'fc6'	Fully Connected	4096 fully connected layer

Table 5 Alexnet Layers Layout-III

Sr No	Label	Layer	Description
18	'relu6'	ReLU	ReLU
19	'drop6'	Dropout	50% dropout
20	'fc7'	Fully Connected	4096 fully connected layer
21	'relu7'	ReLU	ReLU
22	'drop7'	Dropout	50% dropout
23	'fc8'	Fully Connected	1000 fully connected layer
24	'prob'	Softmax	softmax
25	'output'	Classification Output	crossentropyex with 'tench', 'goldfish', and 998 other classes

Procedure of texture feature and deep learning feature extraction techniques

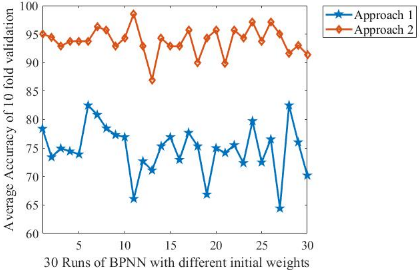
Procedure 1:Texture Feature Extraction Technique

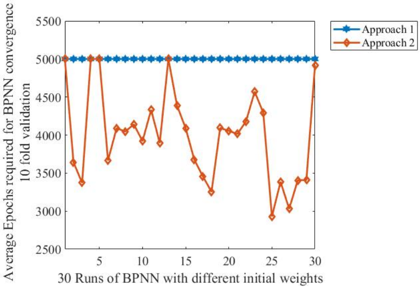
1: **for each** $p := 1 \rightarrow P$ **do**
2: Read p^{th} RGB leaf Image.
3: Resize it to 256 X 256 size & convert it into gray scale..
4: Apply K-Means on it. Choose the infected segment.
5: Calculate 12 features using GLCM & store it row wise.
6: **end for**
7: Input feature vector to BPNN and calculate its performance.
8: End

Procedure 2:Deep Learning Feature Extraction Technique

1: **for each** $p := 1 \rightarrow P$ **do**
2: Read p^{th} RGB leaf Image.
3: Resize image to 227X 227.
4: Load the pretrained 'Alexnet' model
5: Extract the features using the 'fc8' layer & store it row wise.
6: **end for**
7: Input feature vector to BPNN & calculate its performance.
8: End

Comparative analysis of Texture features & Deep learning features, when applied on Backpropagation Neural Network





Summary of Approach 4

- In this model, the pretrained deep learning model 'Alexnet' is used to extract features from the infected leaf to detect disease accurately.
- The accuracy of the model implemented with deep learning features is quite high as compared to the model implemented with texture features,.
- The model with deep learning features gives average accuracy 93.5% when applied to Backpropagation neural network algorithm.

Publication on Approach 4

Paper on *Comparative analysis of the leaf disease diagnosis system using texture features and deep learning features* in International Journal of Applied Engineering Research (IJAER), Research India Publication.

Conclusion I

- The Backpropagation algorithm suffers with the problem of slow convergence rate. Here two approaches are proposed to increase the convergence speed of the BP.
- In approach 1, The leaf disease detection system is implemented using texture features applied on online and batch Backpropagation. The online BP gives better results as compared to the batch BP.

Conclusion II

- In approach 2, the weighted sum is modified by adding white Guassian noise in it. This approach is named as NNBP (Noisy Net Backpropagation) which outperforms as compared to the standard BP when applied on parity problem and Iris dataset. But its performance is average when applied on leaf disease diagnosis system using texture features.
- In approach 3, explain method 3.
- In approach 4, instead of using texture features, automatic feature extraction technique is used by applying pretrained deep learning model 'Alexnet'.

References I

[1] D. E. Rumelhart, G. E. Hinton and R. J. Williams, Learning Internal Representations by Error Propagation, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition 8*, (MIT Press,1986).

[2] S. E. Fahlman, Faster-learning variations on Backpropagation: An empirical study, *Proceedings of the 1988 Connectionist Models Summer School*, (1988) pp. 38-51.

[3] C. Cheung, S. Ng, A. K. Lui and S. S. Xu, Solving the Local Minimum and Flat-Spot Problem by Modifying Wrong Outputs for Feed-Forward Neural Networks, *Proceedings of IJCNN* (Dallas, Texas, USA, 2013) pp. 1463–1469.

References II

[4] B. Yang, Y. Wang, X. Sui and L. Wang, Solving Flat-Spot Problem in Backpropagation Learning Algorithm Based on Magnified Error, *Proceedings of the Third International Conference on Machine Learning and Cybernetics* (Shanghai, 2004) pp. 1784-1788.

[5] S. C. Ng, S. H. Leung and A. Luk, An Integrated Algorithm of Magnified Gradient Function and Weight Evolution for Solving Local Minima Problem, *Proceedings of International Conference on Neural Networks 1*, (2002) pp. 586-591.

References III

[6] P. Moallem and S. A. Ayoughi, Removing potential flat spots on error surface of multilayer perceptron (MLP) neural networks, *International Journal of Computer Mathematics Taylor & Francis,* **88***(1)*, (Bristol, PA, USA, 2011) pp. 21-36.

[7] Wang Chuan, and Jose C. Principe. Training neural networks with additive noise in the desired signal, *IEEE Transactions on Neural Networks 10(6)* (1999) pp. 1511-1517.

[8] Ho Kevin, Chi-Sing Leung, and John Sum. Objective functions of online weight noise injection training algorithms for MLPs, *IEEE transactions on neural networks 22(2)* (2011) pp. 317-323.

References IV

[9] Kiyotoshi Matsuokax. Noise Injection into Inputs in Back-Propagation Learning, *IEEE Transactions on System, Man and Cybernetics 21(6)* (2011) pp. 436-440.

[10] Amara, Jihen, Bassem Bouaziz, and Alsayed Algergawy. A Deep Learning-based Approach for Banana Leaf Diseases Classification,*In BTW (Workshops)* (2017) pp. 79-88.

[11] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. Using deep learning for image-based plant disease detection, *Frontiers in plant science 7* (2016) pp. 14-19.

References V

[12] Zhang, Xihai, Yue Qiao, Fanfeng Meng, Chengguo Fan, and Mingming Zhang. Identification of Maize Leaf Diseases Using Improved Deep Convolutional Neural Networks. *IEEE Access* (2018).

Thank You !
&
Questions ??