International Journal of Pure and Applied Mathematics
Volume 116 No. 23 2017, 525-532
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

PALM-DORSAL VEIN PATTERN AUTHENTICATION USING CONVOLUTED NEURAL NETWORK (CNN)

[1]M. Rajalakshmi, [2]V. Ganapathy, [3]R. Rengaraj
[1]Asst.Prof, Dept of IT, Faculty of Engineering and Technology,
SRM University, Tamilnadu, India
[2]Professor, Dept of IT, Faculty of Engineering and Technology,
SRM University, Tamilnadu, India
[3]Assoc.Prof, Department of Electrical and Electronics Engineering,
SSN College of Engg, Tamilnadu, India
[1]srmraji@gmail.com, [2]ganapathy.v@ktr.srmuniv.ac.in, [3]rengaraj81@gmail.com

Abstract: with the advent of high power, cheap computing and greater complexity, biometric authentication has become possible at every scale because of its safer nature and also user-friendly. According to different researchers, vein biometric is a good biometric trait among others such as fingerprint, palm and finger veins, eyes, voice, signature, gait and DNA for authentication systems. In this paper, Palm-Dorsal (Back of hand) vein pattern authentication using a Convoluted Neural Network (CNN) is proposed. The main advantage of CNN over other traditional approaches is its ability to simultaneously extract features, reduce dimensionality of data, and classifying it using single network structure. Also, the method requires only minimal steps for image pre-processing since the CNN is robust to noise. The proposed method employs the existing database of Near Infra Red (NIR) images that contain palm dorsal hand vein structures of 50 subjects with 5 right hand samples for each subject. The proposed system includes: (1) Detection of Region of Interest (ROI) (2) Palm-Dorsal vein extraction (3) Training of Neural Networks using supervised learning and (4) Using trained CNN to perform biometric authentication. The ROI is extracted using edge detection and contour mapping. The extracted ROI is processed using threshold method followed by morphology operations to get the vein pattern and its features. In this paper, a reduced-complexity four-layer CNN with fused convoluted sub sampling architecture is proposed for Palm-Dorsal vein recognition. After the pre-processing stage, the Neural Network code is written in Python and is trained to classify input vein patterns using supervised learning. Once the training process is completed, optimum CNN weights are obtained. These weights are used in the evaluation stage in which an image is selected from the database and the identification process is invoked. Using trained CNN, an identification rate of 98% is achieved, with an 80/20 percent ratio for separation of training and test samples respectively.

Keywords: biometric authentication, Palm-Dorsal vein, Local dynamic thresholding, Sub sampling, Convoluted neural network

1. Introduction

Biometric technology is an efficient personal authentication and identification technique. Biometric is the term used in computer science to refer to the field of mathematical analysis of unique human features. Biometric solutions have witnessed an accelerated pace of growth in the global market of security over the past several decades, mainly by increasing requirements in public security against terrorist activities, sophisticated crimes, and electronic frauds. The following are some of the security parameters that are associated with biometrics.

- Biometrics is based on the uniqueness of human features
- Since biometrics is associated with individual, it cannot be shared with others.
- Biometrics properties cannot be lost, unless a serious accident happens.
- It cannot be copied.

As one of the main-stream branches, vein recognition has drawn much attention among researchers and diverse users. Anatomically, veins are blood carrying vessels interweaved with muscles and bones, and the key function of the vascular system is to supply oxygen to each part of the body. The spatial arrangement of vascular network in the human body is stable and

unique, and vein patterns of individuals are different, even between identical twins [1]. Biometrics is the science of identifying a person using its behavioural and physiological features [2]. Biometrics systems are classified in two categories that are physical and behavioural. Physical systems are related to the shape of the body such as fingerprints, face recognition, DNA, vascular patterns, iris of the eye etc. Behavioural biometrics system are related to behaviour of a person like voice, gait etc. In this paper, we focus on the vein or vascular pattern of the back of the hand (i.e., dorsal hand) because it is distinctly visible, easy to acquire, and efficient to process. As compared with other popular biometric traits, such as face or fingerprint, the hand vein has several distinguished merits. Comparatively, vein pattern based biometric systems have the following advantages [3].

- Direct Liveness Detection: Vein pattern can be obtained from a person who is alive.
- Internal: As the veins are hidden inside the body it is not damageable and non-wearable.
- Highly Secure: It is extremely difficult to counterfeit or duplicate.
- Universality: It can be obtained from anyone.
- Consistency: Vein patterns do not change with time.
- Uniqueness: Vein patterns are different for different individuals. Even identical twins have different pattern.

It is also a best variant to biometric systems that require physical contact because it extracts the vein pattern, with the hand not in contact with the device instead hand is just easily stretched and the capturing of vein pattern is easily carried out. Since the system is based on three features like live body, internal veins and non-contact type, there is no possibility for forgery, and no misuse by evildoers, thus it can be used at places requiring high level of security.

Despite the merits mentioned above, there are some challenges still needed to be sorted out in order to achieve greater performance required in real-world deployments of Palm-Dorsal vein biometric identification systems. First, poor lighting at the capturing device may cause the image to appear extremely dark or bright [4]. Secondly, in a Palm-Dorsal vein acquisition process, the position of the back of the hand and the camera is very close, which causes optical blurring on the captured image [5] and if the hand is not guided with a proper hand-docking frame, then a slight misalignment will be there, which may decrease the recognition rate [6] eventually causing the matching process to be inaccurate. Third fact is, each individual has different pressure levels in keeping the hand in a fisted position, resulting in varied size of palm dorsal surface. The resulting noise has to be removed as much as possible and traditional palm-vein recognition methods employ some complex image processing algorithms to overcome the issues mentioned.

Several works in the literature have been focused on applying Computer Intelligence (CI) methods in biometric identification especially using CNN. Among the early works, face recognition research using CNN was reported by Lawrence et al. in 1997 [7]. In their work, input samples were first processed by a Kohonen self-organizing map neural network to reduce dimensionality, and then followed by the operation on a LeNet-5 CNN. Learning was performed using a standard back propagation algorithm. Their method had high complexity since two different neural networks were combined to perform the recognition tasks.

From the recent researches, it has been inferred that the above-said problems can be solved by adapting to CI methods, which include neural networks [8], fuzzy logic, and evolutionary computing. A neural network is a cognitive model trying to simulate biological neurons inside human brain, typically known as a multilayer perceptron (MLP), which will work as a powerful classifier. The Neural Network (NN) is formed by several processing elements that provide the ability to do parallel processing, to handle noisy images and learn from experience.

1.1 Convolution Neural Network:

Convolution Neural Network (CNN) was proposed by LeCun in 1989 through LeNet-5 architecture, which was first applied in a handwriting recognition problem. It is a NN Variant that has been widely used for several applications such as face detection [8][9], gender, object, character and texture recognitions. CNN is a type of feed-forward artificial neural network in which the connectivity patterns amongst its neurons are inspired by the organization of the animal visual cortex. LeNet-5 contains several important features; namely, convolution layers, sub sampling layers and two or three fully connected layers. Convolution and sub sampling layers are interleaved twice among each other and become the feature extraction layer arriving at a feature map, while the two or three fully connected layers become the MLP for the classification process. CNN combines segmentation, feature extraction and classification in a single trainable module with minimal pre-processing steps. CNN is normally trained with standard back propagation algorithm. The position and values of filters are automatically decided by the CNN during training process. Classification is performed during training, and at the end of training, the final weights obtained behave as a feature extractor to classify the query input sample.

To the best of our knowledge, this is the first attempt to apply a CNN for Palm-Dorsal vein pattern authentication. The paper is organised as follows: Section 1 gives the introduction about the proposed work; Section 2 describes the theoretical background of the current research; Section 3 discusses in detail the proposed approach, followed by experimental work and results. The final section concludes with the findings of the paper.

2. Theoretical Background

A well-known CNN architecture called LeNet-5 which was first applied in a handwriting recognition problem, consisted of seven layers performing convolution alternated with sub sampling operations. The first convolution layer, which convolves the input with a convolution kernel, essentially acts an edge detector that extracts salient features of the input samples. The kernel, which is of size 5 x 5 in this case, consists of weighting coefficients that create blurring (low-pass filter), sharpening (high-pass filter), or edge enhancement effects. The convolution process is performed by moving a flipped kernel through the images and the resulting output is placed as a new pixel of a feature map at the succeeding layer. The second layer performs sub sampling, that is, a local averaging on a non overlapping small window size of 2 x 2 in this case. This operation reduces the resolution of the feature maps from the previous layer, essentially adding robustness against small distortions in translation, rotation, and scaling. The final two layers of this architecture are MLPs that act as classifiers.

The CNN architecture proposed in this paper has convolution and sub sampling layers which were fused into one single layer. The idea of reducing the number of layers, known as fusion convolution or subsampling concept, was inspired by Simard et al. in 2002 [10]. This fusion significantly reduces the total number of layers in the CNN from seven to four, where seven layers are used in LeNet-5 architecture. The fused convolution and subsampling layer can be described by Eq. (1):

$$y_{pj}^{(l)} = f\left(\sum_{i \in M_j^{(l-1)}} \sum_{(u,v) \in K^{(l)}} w_{ji(u,v)}^{(l)} x_{pi}^{(l-1)}(c+u, r+v) + b_j^{(l)} \right) \quad (1)$$

$$c \bmod S_x = 0,\ r \bmod S_y = 0$$

where $K = \{(u,v) \in N^2 \mid 0 \le u < k_x; 0 \le v < k_y\}$, k_x and k_y are the width and the height of the convolution kernels $w_{ji}^{(l)}$ of layer (l) and $b_j^{(l)}$ is the bias of feature map n in layer (l), c and r respectively refer to the column and row of a current pixel, and p refers to the particular training sample. The set $M_j^{(l-1)}$ contains the feature maps in the preceding layer $(l-1)$ that are connected to feature map n in layer (l). The notation f is the activation function of layer (l). The variables u and v describe the horizontal and vertical step sizes of the convolution Kernel in layer (l).

3. The Proposed Approach

The development of the proposed Palm-Dorsal vein biometric identification system is discussed in two parts: the pre-processing stage and CNN design. Typical image processing designs require image pre-processing and segmentation to be performed before any further image processing and feature extraction process can be conducted. Pre-processing involves algorithms that include region of interest (ROI) extraction [11][12], image enhancement (denoising), and image normalization (resizing). The segmentation process essentially transforms the original image into a meaningful representation making it easier to analyze.

3.1 Pre-processing Stage

The original image captured is of size 320 x 240 pixels. Instead of going for any fixed size, a cropping is done in a dynamic manner in this proposed work. This size covers almost the entire Palm-Dorsal surface. This image cropping is carried out in order to reduce the time consumed by the process of local dynamic thresholding, which will be subsequently applied in the segmentation phase. The image is then resized to 45 x 60 pixels, reducing some unnecessary portions of vein image to ease subsequent NN training. After resizing, the binarized image is obtained by applying some local dynamic thresholding method, which is an algorithm for converting the gray scale image to a bi-level representation of black and white pixels. The result obtained is black and white image, what is referred as binary image here may contain some noisy data, which is removed by an image processing technique called erosion. Thus, with these minimal pre-processing steps, the vein map is extracted.

In this work, the black and white pixels are set to (0 and 1) respectively. Local dynamic thresholding allows adapting the threshold to the pixel within the window size. Let g(x, y) ϵ [0, 1] be the intensity of a pixel at location (x, y) in a gray-scale image. Threshold t(x, y) is calculated by finding the median pixel value

within a window size shown by Eq. (2). For every pixel at location (x, y), t(x, y) is its local threshold and m(x, y) window centered on the pixel (x, y)as shown in Eq. (3) [13].

$$t(x,y)=m(x,y) \quad (2)$$

$$out(x,y)=\begin{cases}0 & if\ g(x,y)<t(x,y)\\ 1 & otherwise\end{cases} \quad (3)$$

After thresholding, each pixel at location (x,y) is replaced with $out(x,y)$. This method of segmentation is simpler than NiBlack segmentation that requires the mean and variance to be computed. Thus, the pre-processing applied in this work has lower computation complexity than the traditional approach.

is the median of the pixel intensities in a *w* x *w* sized

4. Proposed CNN Design

As mentioned earlier, the proposed CNN architecture is based on a design that consists of fused convolutional and sub sampling layers. The CNN consists of four layers (namely C1, C2, C3, and C4 convolutional layers), and this does not include the input layer. We refer to this architecture as the 6-32-50 model, implying that there are 6, 32, and 50 feature maps in layers C1, C2, and C3, respectively. Layer 4 (i.e. the output layer) is fixed at 50 neurons since the target number of categories to classify is 50 subjects; hence, this information is implicit and therefore unnecessary to include in the model referencing name.

In our CNN design, the input image size is set to 45 x 60 pixels and a 7 x 7 convolution kernel size is used. Each convolution and sub sampling operation uses the same kernel size for convolution. For Subsampling operations, the kernel size is 2 x 2.The selection of

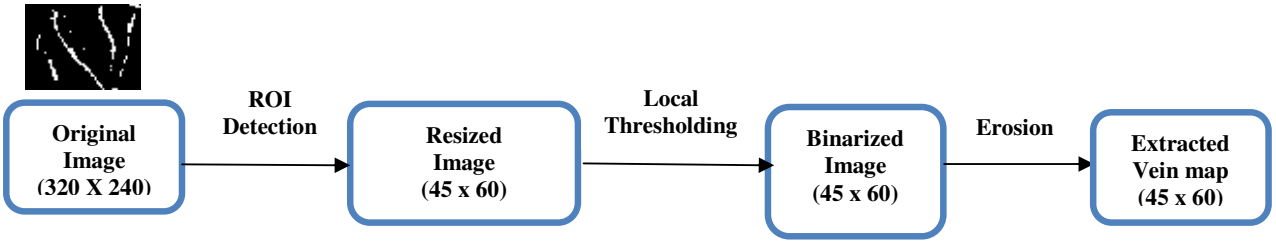


Figure 1. Stages in pre-processing

convolution and sub sampling sizes determines the sizes of the feature maps at layer C3. We are targeting for 1x1 neurons at the final two layers of the MLP classifier. Each layer performs feature extraction tasks and dimensionality reduction processes at the same time.

The size of feature maps at each layer follows the following equation [13]:

$$M_x^{(l)}=\frac{M_x^{(l-1)}-\left(K_x^{(l)}-S_x^{(l)}\right)}{S_x^{(l)}};M_y^{(l)}=\frac{M_y^{(l-1)}-\left(K_y^{(l)}-S_y^{(l)}\right)}{S_y^{(l)}} \quad (4)$$

Where *Mx* and *My* give the sizes of the output map in *x* and *y* directions, *Kx* and *Ky* indicate the convolution kernel size, S_x and S_y determine the subsampling kernel size, and (*l*) and (*l*-1) refer to the current and previous layers. The hyperbolic tangent is used as the activation function in this work, as given in Eq. (5). The notation x is the net input of a neuron. The activation function scales the net input into a specified range. The active range is [0, 1] while the constant value is the asymptote of a sigmoid.

The activation function is given by

$$f(x)=\frac{e^{y_i}}{\sum_{j=1}^{J}e^{y_i}} \quad (5)$$

In this CNN design, classification is based on the winner-takes-all rule [6]. The neuron with the value of +1 corresponds to the winning neuron (that corresponds to the correct class) while 0 corresponds to the incorrectly classified neuron. In the winner-takes-all approach, each output neuron represents a single predefined class of the pattern set. The output neurons compete among each other, in which only the neuron with maximum output value is determined as the winner. This indicates the class of the particular input pattern. This approach is simpler and computationally effective, since only the operation of

finding the maximum output value among these output neurons is required in order to determine the class of the recognition approach that is commonly used in biometric matching processes. A true match is assigned during training and the optimal weights obtained enable direct identification of an unknown subject in the test phase. Cross Entropy Error is the error function used in this work, which is given by

$$C = -\frac{1}{n}\sum_{x} y \ln a + (1-y)\ln(1-a) \qquad (6)$$

where n is the total number of items of training data, the sum is overall training inputs, x and y are the corresponding desired outputs.

4. Experimental Results and Analysis

This paper presents a new method for personal identification of Palm-Dorsal vein patterns using a CNN. The database used in this research was provided by Prof. Ahmed A. Badawi , Professor and Head of Biomedical Engineering, Cairo University, Egypt. It consists of 50 subjects with 5 right hand samples for each subject. In the mentioned database, the infrared image for each subject was captured at different poses. Therefore, there are a total of 250 samples, among which 200 sample images are used for training purpose and 50 sample images are used for testing. The algorithms are implemented in Python and are utilized for pre-processing.

The performance measure of the system is based on the percentage of correct classification of the test samples. We measure the performance of the network by the accuracy of the classification process. In this work, instead of measuring the output error, we train all the samples and identify the number of misclassified samples. Misclassified samples (or samples being unrecognized) occur due to insufficient learning of the features contained in that particular sample by the NN. Correct classification is obtained by comparing the maximum value of the output vector and its corresponding class (target). If the position of the maximum value in the output vector is not equal to the target value, the number of misclassified samples is increased by 1. Then the accuracy is obtained by finding the percentage of correctly classified samples as given in the equation (7) [13]:

$$Accuracy = \left(1 - \frac{No.ofcorrectly\ classified\ test\ samples}{Total\ no.of\ test\ samples}\right) \times 100\% \qquad (7)$$

Table 1 summarizes the performance characteristics in terms of accuracy. A high accuracy of 98% is achieved by the 6-32-50 model.

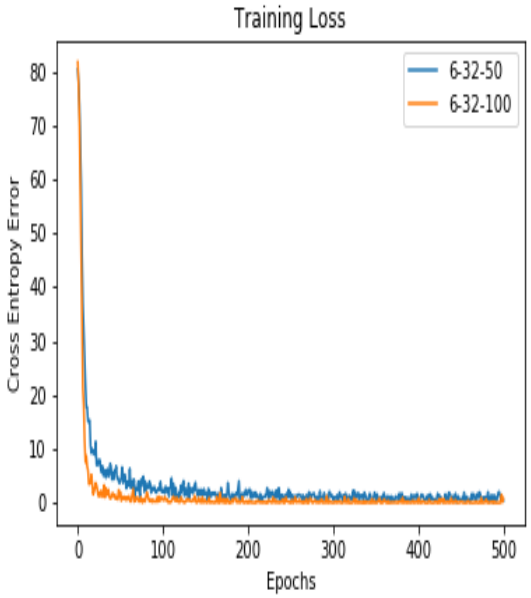
pattern. This approach replaces the similarity metric

Table 1. Performance Accuracy of Each CNN Model.

Model	Accuracy (%)
6-32-50	98.8
6-32-100	97.6

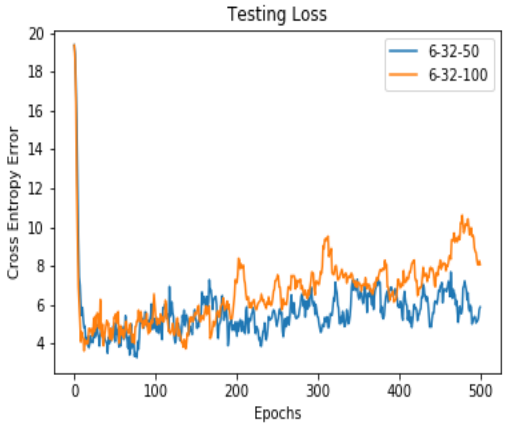
4.1 Error rates in both phases

The number of epochs used in this proposed work is 500. The results of training and testing error are compared as depicted in the Figures 8 (i) and (ii). It is clear from the graphs that the training loss obtained by plotting the cross entropy error against the number of epochs is slightly higher for 6-32-50 model, when compared to the 6-32-100 model, whereas the testing loss is slightly lesser for 6-32-50 model, when compared to the 6-32-100 model. Thus the 6-32-50 model is selected for both training and testing purposes.



Figures 2. (i) Error rates in training mode

The effect of reducing samples in increasing the accuracy can be seen in Table 2. The network has classified the images moderately when trained with a larger dataset.



Figures 2. (ii) Error rates in testing mode

Table 2. The Accuracies obtained with 50 and 40 subjects

Number of Subjects	Number of Training samples	Number of Testing samples	Accuracy (%)
50	200	50	98%
40	160	40	97%

5. Conclusion and Future work

In this paper, vein pattern authentication using a Convoluted Neural Network (CNN) is proposed. The main advantage of the CNN over other traditional approaches is its ability to simultaneously extract features, reduce dimensionality of data, and classifying it using single network structure. Also, the method requires only minimal steps for image pre-processing since the CNN is robust to noise. The proposed method employed the existing database of Near Infra Red (NIR) images that contained palm dorsal hand vein surface of 50 subjects with 5 right hand samples for each subject. The Region of Interest (ROI) is extracted using edge detection and contour mapping. The extracted ROI is processed using threshold method followed by morphology operations to get the vein pattern and its features. In this paper, a reduced-complexity four-layer CNN with fused convoluted sub sampling architecture has been proposed for Palm-Dorsal vein recognition. After the pre-processing stage, the samples are trained by the CNN and accuracy is displayed. The Neural Network code is written in Python and is trained to classify input vein patterns using supervised learning. Using the trained CNN, an identification rate of 98% is achieved, with an 80/20 percent ratio for separation of training and test samples respectively.

Future work could include but is not limited to development of vein image capture tool for generation of a larger dataset and training of other machine learning algorithms for comparison and ensemble predictions. One of the largest problems facing data acquisition for vein pattern analysis is the high cost of infrared imaging tools. Development of low cost methods for imaging would enable building of larger databases. Implementation of the above in a standalone hardware solution would provide a gateway to popularising newer methods of biometric analysis for a new generation of privacy and security solutions.

6. Acknowledgement

The Authors would like to thank Prof.Ahmed M. Badawi, Professor and Head of Biomedical Engineering, Cairo University, Egypt for providing us with the palm- Dorsal hand vein database which consists of 50 subjects (250 images in total).

References

[1] Mohamed Shahin, Biometric Authentication Using Fast Correlation of Near Infrared Hand Vein Patterns, World Academy of Science, Engineering and Technology, Vol.2, pp.141-148, 2008.

[2] A K Jain, A Ross and S Prabhakar, An Introduction to Biometric Recognition, Circuits and Systems for Video Technology, IEEE Trans, ISSN: 1051-8215, Vol. 14, pp. 4-20, 2004.

[3] Toshiyuki Tanaka, Naohiko Kubo, Biometric Authentication by Hand Vein Patterns, SICE Annual Conference, Sapporo, ISBN: 4-907764-22-7, Vol.1, pp. 249- 253, Aug 2004.

[4] Yanagawa T, Aoki S, Ohyama T. Human finger vein images are diverse and its patterns are useful for personal identification in 21st Century COE Program, Development of Dynamic Mathematics with High Functionality; 2007. pp. 1-8.

[5] Lee EC, Park KR., Image restoration of skin scattering and optical blurring for finger vein recognition, Opt Lasers Eng 2011; 49: 816-828.

[6] Mulyono D, Jinn HS,A study of finger vein biometric for personal identification, International Symposium on Biometrics and Security Technologies; 2008.

[7] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back. 1997. Face Recognition: A Convolutional Neural Network Approach. IEEE Transactions on Neural Networks. 8: 98-113.

[8] Jürgen Schmidhu, Deep learning in neural networks: An overview, Elsevier, 2014.
[9] Syafeeza, A Convolutional Neural Network for Face Recognition with Pose and Illumination Variation, International Journal of Engineering & Technology, 6(1): 0975-4024, 2014.
[10] P. Y.Simard, D. Steinkraus and J. C.Platt. 2003. Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis. In: Seventh International Conference on Document Analysis and Recognition. pp. 958-963.
[11] Yang J, Shi Y, Finger-vein ROI localization and vein ridge enhancement, Pattern Recognition Letter 2012; 33: 1569-1579.
[12] Podgantwar UD, Raut, U. Extraction of finger-vein patterns using Gabor filter in finger vein image profiles, International Journal of Engineering Research and Technology 2013; 2: 60861.
[13] Syafeeza ahmad radzi, mohamed khalil-hani, rabia bakhteri, Finger-vein biometric identification using convolutional neural network, Turkish Journal of Electrical Engineering & Computer Sciences, (2016) 24: 1863-1878.
[14] http://opencv-python tutorials.readthedocs.io/en/ latest/py_tutorials/py_tutorials. Html Date accessed: 10/08/2016.
[15] G. Bradski. The OpenCV library, Dr. Dobb's J. Software Tools, 2000.
[16] T. Padmapriya and V. Saminadan, “Improving Throughput for Downlink Multi user MIMO-LTE Advanced Networks using SINR approximation and Hierarchical CSI feedback”, International Journal of Mobile Design Network and Innovation- Inderscience Publisher, ISSN : 1744-2850 vol. 6, no.1, pp. 14-23, May 2015.
[17] S.V.Manikanthan and K.srividhya "An Android based secure access control using ARM and cloud computing", Published in: Electronics and Communication Systems (ICECS), 2015 2nd International Conference on 26-27 Feb. 2015,Publisher: IEEE,DOI: 10.1109/ECS.2015.7124833.
[18] T. Padmapriya, V.Saminadan, “Performance Improvement in long term Evolution-advanced network using multiple imput multiple output technique”, Journal of Advanced Research in Dynamical and Control Systems, Vol. 9, Sp-6, pp: 990-1010, 2017.
[19] S.V.Manikanthan, Padmapriya.T, “RECENT TRENDS IN M2M COMMUNICATIONS IN 4G NETWORKS AND EVOLUTION TOWARDS 5G”, International Journal of Pure and Applied Mathematics, Vol. 115, No. 8, pp: 623-630, 2017.
[20] SHUBHANSHU GUPTA, S. KOLANGIAMMAL, T.PADMAPRIYA, “Smart Curtain Using Internet Of Things” International Innovative Research Journal of Engineering and Technology, Vol. 2, Special Issue, pp. 4-8.
[21] Rajesh, M., and J. M. Gnanasekar. "Path observation-based physical routing protocol forwireless ad hoc networks." International Journal of Wireless and Mobile Computing 11.3(2016): 244-257.
[22] S.V.Manikanthan and V.Rama“Optimal Performance Of Key Predistribution Protocol In Wireless Sensor Networks” International Innovative Research Journal of Engineering and Technology ,ISSN NO: 2456-1983,Vol-2,Issue –Special –March 2017.
[23] Rajesh, M., and J. M. Gnanasekar. "Congestion control in heterogeneous wireless ad hoc network using FRCC." Australian Journal of Basic and Applied Sciences 9.7 (2015): 698-702.