

Stopping Wildlife Poaching Using Face Recognition

Rajyashree, Monojit Debnath, Sanyam Malik, Shashank Mishra

Abstract: Poaching is illegal hunting, killing of wild animals also referred to the illegal harvesting of wild plant species. It's considered as an ecological wrong doing against the regular assets, unlawful catching of natural life for creature extravagances, for example, ivory, horn, teeth, skin and bone. India is home to some of the most beautiful animals on the planet such as Tiger, Elephant, Rhino, Leopard, Lizard and many types of snakes. These glorious creatures were utilized for games of chasing, presently they are ensured under the wildlife protection act of India. Sadly enough, the population of these beautiful wild animals of India are going down because of Poaching for ivory, horn, teeth and skin. In this project, we proposed to stop the poaching by using face recognition technology for the detection unauthorized persons responsible for poaching. Government has installed several cameras for the purpose of monitoring the wildlife in the forest. We are planning to connect the face recognition to that and with we will be able to detect the person, his location in the forest and an alarm will be sent if the person is not on the database of the forest department. Here we are also planning to update the architecture of the system so the detection will be much better from the previously designed systems. In this paper an algorithm is provided for face detection in noisy background with additional spoof face detection. The implemented algorithms are CNN algorithm, SVM classifier, Local Binary Pattern (LBP), Micro Texture Analysis. For fast face detection the LBP is used. The error face detection rate is decreased using eye detection algorithm. To increase the contrast and orientation the detected facial image is processed with maintaining high face recognition accuracy. Here large dataset of facial and fake images is used to train the dataset. True positive result of this algorithm is 98.8% and the correct facial recognition is 99.2%. Here we have done work on spoof face detection with Micro Texture Analysis with multi scaling LBP. Two databases are used one is Yale database and other is publicly available database.

Index: Machine Learning, Image Detection, Face detection, LBP, SVM classifier, Micro-Texture Analysis, Back-Propagation.

I. INTRODUCTION

In India every day the crime rate in the field of wildlife poaching is increasing. By our survey we have seen that the wild life sanctuary has cameras to watch over the animals and to watch over their breeding sites. Also, all the outside

Revised Manuscript Received on April 05, 2019.

Rajyashree(Guide) , Computer Science, Assistant Professor, SRM Institute Of Science And Technology, Chennai, India

Monojit Debnath, , Computer Science, SRM Institute Of Science And Technology, Chennai, India

Sanyam Malik, Computer Science, SRM Institute Of Science And Technology, Chennai, India

Shashank Mishra, Computer Science, SRM Institute Of Science And Technology, Chennai, India.

premises have its own surveillance cameras but that's no use in night or if someone pass by and the guard is not watching. Because the sites that are shown to the visitors they always have guards , but the restricted sites are the places where the main crime is happening because that are watched by normal surveillance lance systems , so if any imposter wear a guards uniform he can easily get through , so here in this project we proposed a method to utilize the Surveillance cameras with help of the face recognition and spoof face detection , here face detection will work because the sites we have to protect is restricted to common people so only limited amount of authority is there so if any unknown face goes there our method will alert the guards, also the could try to use authorized people fake face to get there with that in mind we have decided to implement the spoof face detection also . In this project we have proposed enhanced version of both face detection and spoof face attack detection. Below section (II) discuss about Logistic Regression and Convolutional Neural Network, in section (III) we have discussed the face detection proposed method and also in (A) we have discussed our proposed spoof face detection method. In (B) we have discussed how we are implementing the multi scalar LBP and store it in Histogram. In (C) we have discussed about the SVM classifier that we have used in this project. In section (IV) we have discussed the experimental analysis in (F) we have discussed setup, in (G) Experimental results. In section (V) we have discussed the conclusion.

II. LOGISTIC REGRESSION AND CONVOLUTION NEURAL NETWORK

In this project we briefly describe the methods of convolution and logistic regression. The intention is to detect their strength and weaknesses. After this we will combine both methods strengths.

A. Convolution Network

Convolution Network is a network which is able to extract the topological properties from the input images. Features from raw images is extracted by CNN and then classified by a classifier. CNN is not affected by distortion and simple geometric variations. Local receptive fields, shared weights, spatial or temporal sub-sampling is utilized so the network stays unaffected by distortion. For training purpose of the network the back propagation is used .

Stopping Wildlife Poaching Using Face Recognition

Feature map $C_{k,l}^i$ in CNN is shown below by (1),

$$C_{k,l}^i = g(I_{k,l} \circ W_{k,l} + B_{k,l}) \quad (1)$$

$S_{k,l}$ feature in Non overlapping sub sampling is shown by (2)

$$S_{k,l}^i = g(I_{k,l} \downarrow w_{k,l} + E_{k,l}) \quad (2)$$

here $g(x)=\tan H(x)$ sigmoid activation function, biases are B and b , weights are W and w , i th input is denoted by $I_{k,l}^i$, \downarrow is used for down-sampling symbol, matrix whose all elements are one is denoted by E , is used for 2 dimensional convolution. Here the upper class letter represents the matrices and lower class letter represents the vectors [Lyons M et., al].

local receptive fields are extracted from preceding layers by a convolutional layer. For extracting the local features feature maps is used for detecting the features of face. In a network of 5*5 convolution kernel every unit has 25 inputs which is connected to 5*5 area in previous layer also known as local receptive field. Each connection has a trainable weight but every unit of same layer share the same weight, also known as weight sharing technique.

32*32 pixel of image is taken by LetNet(5). LetNet(5) is consist of three convolution layers (C1,C2,C3), two sub sampling layers (S2 AND S4), a fully connected layer (F6) and the output layer. Euclidean RBF layer is connected to 10 layer (for 10 classes) which is known as output layer, shown in fig :1 ,

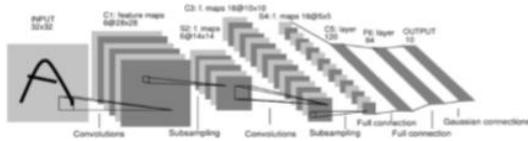


Figure 1: LetNet(5) Architecture

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X		X	X
1	X	X				X	X	X			X	X	X		X	X
2	X	X	X			X	X	X			X	X	X		X	X
3		X	X	X			X	X	X			X	X	X		X
4			X	X	X			X	X	X			X	X	X	
5				X	X	X			X	X	X			X	X	X

Table 1 : The interconnection among C3 and S2 layer
The table map indicates that every feature map of S2 is not related to C3.

B. Logistic Regression and Classifier

Logistic Regression is an equation of form

f: $X \rightarrow Y$ or $p(Y|X)$,

here the Y is discrete value and $X = \langle X_1, \dots, X_n \rangle$ is any vector whose containing any discrete or continuous variably = $\langle Y_1, \dots, Y_n \rangle$ which is able to take any discrete values [Mitchell T. M. (2010)et.,al]. Logistic Regression directly assume the parameters from the training set by assuming a parametric form for the distribution $P(X|Y)$. When Y is Boolean value the parametric model is:

$$P(Y=1|X) = \frac{1}{1 + \exp(w_0 + \sum_{i=1}^n W(i) X(i))} \quad (3)$$

and

$$P(Y=0|X) = \frac{\exp(w_0 + \sum_{i=1}^n w_i x_i)}{1 + \exp(w_0 + w_i x_i)} \quad (4)$$

The sum of two profanities must be equal that's why the (4) follows directly from (3)[Mitchell T. M. (2010) et.,al].

We assignee the label $Y=0$ if the below condition is there [Mitchell T. M. (2010) et.,al]:

$$1 < \frac{P(Y=0|X)}{P(Y=1|X)}$$

By substituting the (3) and (4) we get :

$$1 < \exp(w_0 + \sum_{i=1}^n w_i X_i)$$

The form of the Logistic function is shown in figure 2:

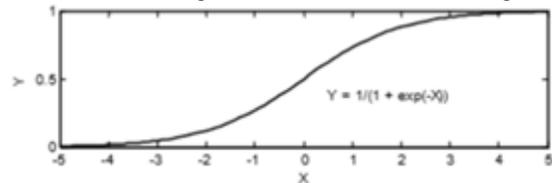


Figure 2: The form of $P(X|Y)$

Now we have linear classification rule by taking log on both sides, it assigns $Y=0$ if X satisfies

$$0 < w_0 + \sum_{i=1}^n w_i X_i \quad (5)$$

and assigns $Y=1$ otherwise [Mitchell T. M. (2010)et.,al]

III. PROPOSED SYSTEM

Here the Fig:3 is describing the feature extraction of CNN. One input layer, two convolutional layer, one subsampling layer has been used here in CNN. To make input images suitable with the designed structure the images of database were converted to 64*64. Here we have used first convolutional layer with 6 feature maps, in which each have a resolution of 54*54, with field of 7*7. With a receptive field of 2*2 the second subsampling layer consist of 16 feature maps with a size of 34*34. The second convolutional layer with resolution of 22*22, has 16 feature maps with a respective field of 8*8. The output field with a field of 11*11 and has 15 feature maps with the size of 1*1. Table:2 s showing the interconnection between first and second layer.

	0	1	2	3	4	5	6	7	8	9
0	X				X	X	X			
1	X	X				X	X	X		
2	X	X	X				X	X	X	
3		X	X	X				X	X	X
4			X	X	X				X	X
5				X	X	X				X

Table 2: interconnection between first subsampling layer with second convolutional layer

	Accuracy (%)	Time (s)
NaiveBayes	75.15	0.05
NaiveBayesSimple	75.15	0.01
NaiveBayesUpdatable	75.15	0
Logistic	65.45	1.23
MultiLayerPerceptron	81.21	2.88
RBFNetwork	72.72	23
SimpleLogistic	86.06	1.22
SupportVectorMachine	83.03	9.79
IB1	76.36	0
IBk	83.63	0
LWL	78.78	0
ClassificationViaRegression	81.21	0.28
BFTree	75.15	0.21
FT	71.51	0.64
LMT	86.06	1.64
SimpleCart	75.75	0.22

Figure 10: Comparison of different algorithms

As shown above the Simple Logistic algorithm can detect the unseen faces with most accuracy of 86.06% and with a time of 1.22 sec, The LMT algorithm has the same accuracy but the time taken by it is 1.64 sec, the simulation was done by Weka software .

The Classification accuracy and time taken by different algorithm is shown below in figure 11,

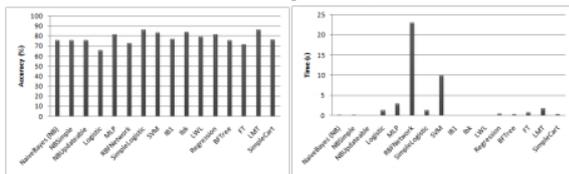


Figure 11: classification accuracy and the time test on the test set

So in this project we have proposed a two step learning process method: view the first N-1 layer as a feature extractor by training a CNN. Then we have applied a simple logistic regression method on the CNN features, Backpropagation gradient descent algorithm is used to train the network. Here we have shown that the simple logistic regression method classifies the features which were extracted with the highest accuracy and with the lowest time.

E. Spoofing Detection Using Micro Texture Analysis

Face images captured by picture and the live faces at first look very similar but the picture is a 2d rigid object where the face is a 3d non-rigid structure that's why they both reflect light in a very different way. By using micro texture analysis, we can differentiate that. Here to describe the micro texture and the special information we have used LBP (Linear Binary Patterns). To determine whether the image is live or a picture we feed the vectors in feature space to SVM classifier. In Fig 6, to images is shown in original space and LBP images of them using LBP as basic feature. At first both live face and the picture look same but in the micro analysis texture detection both have some differences. The enhancement made in this project is shown in next stanza.



Figure 6: Example of two images (a live face and face print) in the pivotal space and LPB images using LBP as a crude architecture

F. Discriminative Feature Space Using LBP

To create labels for image pixels the original LBP operator consider the result as binary value and threshold the value 3*3 neighbourhood of each pixel with center value. Fig:7 shows a histogram of which means 256 different labels.

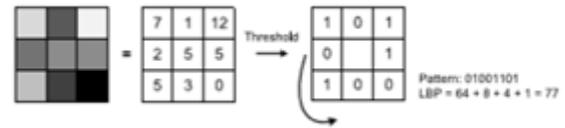


Figure 7: The basic Operator of LBP

We had used a circular neighborhood and bilinearly interpolating values at non-integer pixel coordinates for allowing any number of radius and umber of pixes in the neighbourhood. For denoting P sampling point on the circle of R radius we normally use (P, R).Pixel (xc,yc)'s LBP code is given by;

$$LBP_{p,r} = \sum_{P=0}^{P-1} s(g_p - g_c) 2^P$$

g_c is the gray value of the center pixel (xc,yc), the gray value of P, that is equally spaced pixel in the R radius of circle is denoted by g_p , and the thresholding function is defined by s as follows:

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0; \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

The uniform pattern is another extension of the original operator. The uniform pattern was brought in field by the fact that there are some common binary patterns that made occurrence more than other binary patterns in texture images.

If a binary pattern consists of 2 bitwise transitions traversed circularly from 0 to 1 or vice versa that is called uniform binary pattern. Uniform pattern used in LBP so that every uniform pattern has its own separate label and oth ($LBP_{P,R}^{u2}$) er non-uniformed patterns are labeled with a single label.u2 is used to notify all the uniform pattern and for labelling all the remaining pattern in a single label [T. Ojala,et...al].

Every LBP label can be treated as micro-texture. Local features is defined first by this labels include different type of curved edges, flat spaces, spots, etc. Normally the Histogram collects the LBP code occurs in the image. Then the computation of Histogram similarities is done to perform the classification. To extract the LBP Histogram and to concatenate them into one enhanced feature Histogram the facial images are first divided into several local regions. For face recognition such representation is very adequate [T. Ahonen,et...al]. But we have seen that the details to discriminating a real human face form a fake human face for micro-texture can be best detected using different combination of LBP operators. For better partition of real and fake human face we have derived a facial representation using multi-scale LBP operators. The proposal is shown below:

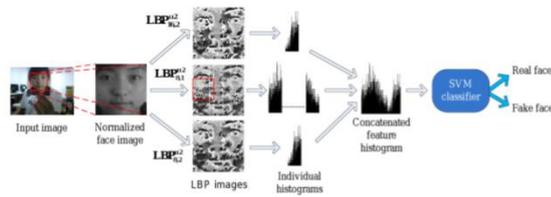


Figure 12: The detected face is normalized in 58*58 pixel image. After we applied LBP8,1u2 on normalized face and divided that 3*3 overlapping area. The local 59-bin Histograms from each region are calculated and stored in 531-bin Histogram. Then we calculated two other Histograms from the whole face using LBP8,2u2, LBP16,2u2 creating 59-bin and 243-bin Histograms that are added to 531-bin Histogram. Then we used Non-linear SVM classifier to decide the face is real or fake.

In above fig:12, we have proposed a method to capture the special information enhance the holistic description by including global LBP histograms calculated over the whole face image this is done by calculating LBP features of the 3*3 overlapping region. This previously mentioned work is done by following order: first the face is detected with the help of enhanced CNN technique we proposed then the face is normalized to 58*58 pixel image. Then an application is made of LBP_{8,1}^{u2} on the normalized facial images and divide the out coming value of LBP face image in 3*3 overlapping area (the size of overlapping is 14 pixel). The local 59-bin Histogram from each region is collected by 531-bin histogram after computing. After this we used LBP_{16,2}^{u2} and LBP_{8,2}^{u2} operators to compute two other whole images, to create 59-bin and 243-bin histograms that are added to 531-bin histogram previously computed. With that in mind the final histogram's length is 833 (i.e.. 531+59+243).

G. Classification

After computing the enhanced histogram, we used a non-linear SVM classifier with radial basis function kernel [V. N. Vapnik, et., al] to detect whether the input image is live or fake. The classifier (SVM) is first trained with a set of positive samples and then with a set of negative samples.

IV. EXPERIMENTAL ANALYSIS

To test this system, we have used the publicly available NUAA photograph Imposter database [X. Tan, et., al]. Both the live and face images of human were collected in three sessions about interval of two weeks. In addition, during every session the lighting and environmental conditions were changing.

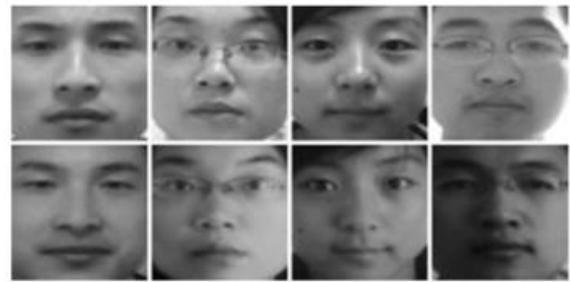


Figure 13: Example of the real face in upper row and fake face in lower row

Above picture fig:13, is showing the images of the database. A 20fps camera is used for the client accesses and spoofing attacks were recorded. 500 images recorded for each subject. In the time of capturing the data, the main idea was to static live objects (live objects look like static object by little or no movements or eye-blinking), furthermore using 2-D facial prints with varying motion 5 different attacks were simulated. The images were captured using conventional web cams with a resolution of 640*480 pixels.

H. Setup

The 15 subject's images in database are divided into two setups one for training and the other one is for test purposes. The first two sessions were in training set, and the third set was in test set. The training set contains of total number of 1743 images out of which 9 is real clients (first batch - 889 and second batch - 854) and 1748 fake images with same 9 clients (855-first batch and 893-second batch). With 3362 client samples and 5761 fake images taken during the third session the test set were created. In addition, 6 more clients were added in the test set to increase the difficulty. Then the face images were normalized using the previous formula to 58*58 pixels.

V. EXPERIMENTAL RESULTS

Here we have evaluated the performance of three powerful texture analysis operators LBP, Local Phase Quantization (LPQ) [V. Ojansivu, et., al], and Gabor wavelets [B. S. Manjunath, et., al] for differentiating real and fake faces of the image. Then the SVM were feed the computed features. To do fair comparison, for texture description optimal SVM parameters were determined which is used for texture description. To apply SVM implementation we have used LibSVM Library in whole project [C.-C. Chang, et., al]. In terms of ROC (Receiver Operating Characteristic) the performance of three texture operator is shown in below picture,

Stopping Wildlife Poaching Using Face Recognition

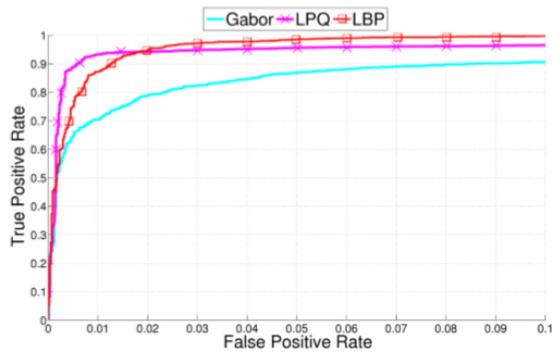


Figure14:performance of three texture operator

Table 5:Error equal rate of three texture operator

Descriptor	LBP	LPQ	Gabor
Error Equal Rate (EER)	2.9 %	4.6 %	9.5 %

From the graph we can say that three operators performed well. In the above table we can see the EER (Equal Error Rate) of LBP(2.9%),LPQ(4.6%),Gabor(9.5%).LPQ is more successful because its blur tolerant and most of the print images were having some amount of blur [V. Ojansivu,et..al.]. One big differences in fake and live face is the 2-D images don't have depth reflection but 3-D images have that like the reflection of the check .Some Differences between real and fake faces is shown in below picture fig:13,



Figure13:Typical characteristics of real client image

This details could be lost if only a single LBP Histogram has been used .

Method	Tan et al.	Our approach
AUC	0.94	0.99

Table 5: Accuracy difference between our approach and Tan et al.

Above table:4, represents the difference between our results and the result of Tan et al [X. Tan, et.,al]. To be fairer we have done the comparison under the terms of Area Under Curve (AUC)as Tan et al. In that also we have achieved the superiority (0.99 vs 0.94). Our method has accuracy of 98.0% and the false rate is 0.6% and the false rejection rate is 4.4%.

In this project we proposed multiscale LBP and have computed features from 3*3 overlapping regions for capturing special information and the holistic description with inclusion of computation of global LBP over the whole face, we have performed a set of tests to measure the impact of the choices we have made, that shown in the below graph,

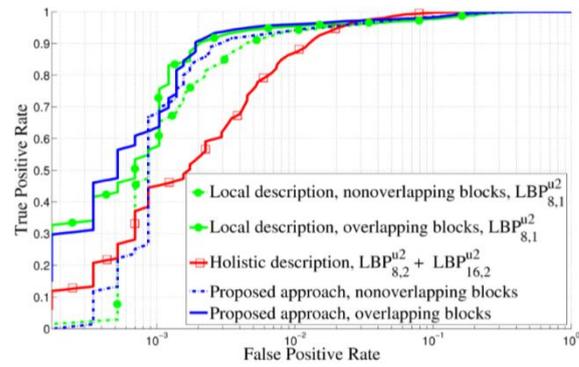


Figure15:Performance of multiscale LBP,overlapping blocks,feature computation from whole image

We can see above at fig:15, that block processing methodology significantly improves performance with lower false acceptance rate (Using the fusion of overlapping regions from 54.0% to 89.0%). Furthermore, Global and local representation combination achieves 91.2% at 2FAR by the uses of overlapping and the multi-scaling LBP at whole face area. We also noticed that bigger block sizes lead to better results.

We performed various 2D face spoofing attacks using print images, the system takes 13.5 MS in average to process an image on 2.4 GHz Intel i5 5th generation CPU with 8Gb of ram.

V. CONCLUSION

Present day face detection technology is very vulnerable in most of the cases the attacker will use a Face print which is 2-D object. Here in this project we have proposed a method to discriminate the fake face from live faces on the basis of Micro-texture analysis pattern. The face prints contain some defects so that can be easily detected using micro-texture analysis, furthermore the face print which is a 2-D object and live face which is a 3D object with a non-rigid structure will reflect the light differently. To encode the micro texture patterns in enhanced Histograms we have used multi scale Local Binary Patterns (LBP), then the results were given to SVM classifier which decides that whether the face is real imposter or not. Hitherto, in relation to the paradigm between this and other experiments, the proposed idea is quite promiscuous and does no longer require the cooperation the consumer.For the future work the Back propagation should be more optimized and the LBP should be more optimized .

REFERENCES

1. Lyons M., Akamatsu S., Kamachi M., Gyoba J. (1998), 'Coding Facial Expressions with Gabor.
2. 2.Wavelets'. In Third IEEE International Conference on Automatic Face and Gesture Recognition, 200-205
3. 3.Mitchell T. M. (2010) 'Generative and Discriminative Classifiers: Naive Bayes and Logistic Regression', Machine Learning
4. 4.Lawrence S., Giles C. L., Tsoi A. C., Back A.D. (1997). 'Face Recognition: A Convolutional Neural Network Approach', IEEE Transactions on Neural Networks, 98-113. 8
5. 5.T. Ojala, M. Pietikainen, and T. Maenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans.



- Pattern Anal. Mach. Intell., 24:971–987, July 2002. 2, 3, 4
6. T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell., 28:2037–2041, December 2006. 2, 4
 7. V. N. Vapnik. Statistical Learning Theory. WileyInterscience, 1998. 2, 4
 8. X. Tan, Y. Li, J. Liu, and L. Jiang. Face liveness detection from a single image with sparse low rank bilinear discriminative model. In Proceedings of the 11th European conference on Computer vision: Part VI, ECCV'10, pages 504– 517, Berlin, Heidelberg, 2010. Springer-Verlag. 2, 4, 5, 6
 9. V. Ojansivu and J. Heikkilä. Blur insensitive texture classification using local phase quantization. In Proceedings of the 3rd international conference on Image and Signal Processing, ICISP '08, pages 236–243, Berlin, Heidelberg, 2008. Springer-Verlag. 5
 10. C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. 5.

AUTHORS PROFILE



Rajyashree (Guide), SRM University, assistant professor, M.E Computer science engineering. Research interests includes network security, ad hoc networks



Monojit Debnath: B.Tech third Year student, SRM University, Research work: prediction of heart attack using AI, RFID based medical Information system.



Sanyam Malik: B.Tech third Year student, SRM University, Ramapuram, Chennai, India



Shashank Mishra: B.Tech third Year student, SRM University, Ramapuram, Chennai, India