Machine Learning Approach of Hybrid KSVN Algorithm to Detect DDoS Attack in VANET

Ms.Nivedita Kadam[1]
PhD Research Scholar
Department of Computer Science and Engineering
Koneru Lakhmaiah Education Foundation
Vaddeswaram, A.P., India

Dr. Krovi Raja Sekhar[2]
Professor
Department of Computer Science and Engineering
Koneru Lakhmaiah Education Foundation
Vaddeswaram, A.P., India

***Abstract*—Most of the self-driving vehicles are suspect of the of the different types attacks due to its communication pattern and changing network topology characteristics, these types of vehicles are dependent on external communication sources of VANET, which is a vehicular network, It has attracted great interest of industry and academia, but it is having a number of issues like security, traffic congestion, road safety which are not addressed properly in recent years. To address these issues it's required to build secure framework for the communication system in VANET and to detect different types of attack are the most important needs of the network security, which has been studied adequately by many researchers. However to improve the performance and to adapt the scenario of VANET, here in this paper we proposed a novel Hybrid KSVM scheme which is based on KNN and SVM algorithm to build a secure framework to detect Distributed Daniel of Service attack *which* is the part of Machine Learning approach. The experimental results shows that this approach gives the better results as compared to different Machine Learning based Algorithms to detect Distributed Daniel of Service attack.**

***Keywords*—K-Nearest neighbor (KNN); support vector machine (SVM); DDoS (distributed denial of service attack)**

I. Introduction

Vehicular Ad Hoc Networks (VANETs) are a sub type of mobile ad hoc network used for communication from vehicles to vehicle and vehicle to roadside units. VANETs is a new technology for the wide range of applications, including traffic congestion monitoring and controlling traffic, and is associated with the passenger and the driver safety program For example, a vehicle on the road where the accident occurred, is able to warn each other in order to find an alternate route to avoid the traffic congestion that occurs after an accident. The special characteristics of VANET, such as high mobility, dynamic network topology, etc., that compromises the quality of service of the application. [1] As vehicles are moving from normal to fully autonomous generation and its era is changed in to the driverless cars, they are providing the lots of information to their surroundings like Vehicle to vehicle, Vehicle to infrastructure and RSU's, as they are interconnected with each other. WAVE protocol is used for communication in VANET. The rising transport network has been thought to be a vital element of the event of the intelligent transportation system (ITS) and smart cities. It's expected to modify an entire new set of applications, starting from road safety improvement to traffic potency optimization. This new generation of networks can ultimately have a profound impact on society and also the daily lives of a lot of folks around the world. A Machine -learning approach is roughly divided into 3 major categories: supervised, un-supervised and reinforcement learning. Classification algorithms assign a categorical label to every sample which is in the system. In wireless networks, the classification detect whether or not the networks are misbehaving node or elements of the system are not functioning properly. The algorithms such as Bayesian classifiers, k-nearest neighbors, decision trees, support vector machines, and neural networks are considered.[2] In machine learning knowledge driven higher cognitive process approach is employed to overcome the new issues coming in transport networks, it is necessary to rethink ancient approaches to wireless network style, particularly given the made sources of knowledge from numerous on board sensors, margin observation facilities, historical transmissions, and then forth. Indeed, it's extremely fascinating to plot economical ways to interpret and mine the huge amounts of knowledge and facilitate a lot of data-driven higher cognitive process to enhance transport network performance. Machine learning represents smart tool to serve such functions with evidenced good performance in an exceedingly wide range of applications.[2].

DDoS attack uses the client/server technology with multiple computers as an attack platform to perform attacks on one or more targets to increase the power of the attack is becoming most of the intense major now a days however, because of the diversity of DDoS attack modes and the variable size of attack traffic, there is not any full proof model which gives the accurate detection method of attack presently.[3-4]. Confidentiality, integrity and availability are the main properties of the any communication in VANET [5] for security, network nodes are vulnerable to capture, hack, and communicate with attackers, and node messages may be eavesdropped and fake messages introduced or replayed into the network. Malicious nodes interfere purposefully with the regular actions of the network, aiming to disrupt the ordinary functionalities of the network. These issues would have a wide-scale effect on Vehicular network implementation, due to this the defence at vehicular network has more complex task, and such issues are the absence of data connectivity due to the low performance of wireless networks between different nodes and the lack of a source-destination connection. Therefore, it is extremely necessary for vehicular network to be highly flexible

to respond to any situation. So to overcome these issues a framework is required which detects the intrusion and store the information in communication by preserving the data integrity. By applying ML algorithms the privacy and data integrity issues can be overcome. There are many unaddressed issues of the security in previous work. This paper is divided in to the following sections: Section 2, Related work; Section 3, Methodology and workflow; Section 4, Experimental results and Discussion; Section 5, Conclusion.

II. Related Work

Literature study has done by considering various papers which are based on intrusion detection systems in vanet through machine learning approach and various algorithms which are used for security in vanet. Elvin Eziama proposed the trust based model with classification approach using Bayesian NN(BNN). This model is generic in nature with classification approach. Author says that when trust mechanism is applied to this framework it gives better results with respect to performance prediction, classification accuracy and low detection latency. And when it is compared with Neural Network, with the uncertainty in the information it leads to result in over-fitting of the data which are collected from the nodes during the training phase. Timing attack, sybil attack and false positioning attack are considered in this work.[6] Carlos H. O. O. Quevedo, Ana M. B. C. Quevedo and Ahmed Serhrouchni Proposed a SyDVELM system for detection of Sybil attack which fast fast, scalable with low complexity Vehicle mobility pattern is considered as vector matrix, for displacement movements of the sybil attack. For some features of Vehicular nodes it applies the Extreme Learning machine. SyDVELM's proposed approach elaborates that from the urban scenarios the mobility pattern of the vehicle nodes are compared with real vehicle reliability in terms of inaccuracies in the relocation of Sybil nodes. The advantage of this mechanism is that it confirms a versatile detection process for sybil attacks with high detection rate, very low error rates and improves the scalability. But it this mechanism is not suitable for the low density scenarios, for more than 300 epochs its detect attacks quickly but when epochs are small, it limits the result. Other ML approaches can be applied with ELM for better results for detection. The authors in [7] Fabio Goncalves and Bruno Ribeiro explains in their work about different types of attacks are identified including the intrusion detection systems with anomaly detection by Machine learning algorithms. Like for malicious packet attack hierarchical IDS used to detect anomaly by applying learning automata algorithm. Based on their survey it is observed that no specific information is provided regarding datasets and its operations which are performed on it like training, validating of data and detection of attacks and beacon messages in communication. The authors in [8] Mohammad Asif Hossain and Wahidah Md Shah proposed they have considers the award and penalty scheme which is based on the case of student and teacher learning process. With this scheme it can reduce the network overhead and delay. The authors in [9] Stefan Mihai, Nedzhmi Dokuz and Meer Saqib Ali in this work explains on technical advances of leading to a wired, mobile, cooperative transport system. And the most important safety consequences of VANETs which provide a complete analysis of existing possible methods for keeping vehicle network communications private, stable and confidential. Different types of security issues and attacks they have elaborated in their work. The issues they have raised about the Mass acceptance depend on closing the gap in terms of both secure automobile accessibility and road networks. Flexibility and reliability while upholding appropriate levels of security and privacy is also the main concern [10]. Steven So, Prinkle Sharma and Jonathan Petit, proposed the misbehaviour detection scheme for VANET, by applying the plausibility checks feature of vector for machine learning models under the SVM algorithm for classification with this KNN- 1-NN algorithm is applied. Data set is KDD for flooding attack detection is possible in WSN VeReMi dataset is used for Detecting and classifying location spoofing misbehavior, this dataset is balanced by normal to attack ratio(70:30) is trained and Check is done by precision-recall curve. Drawback is that it considered only next neighbor for this work. The authors in [10] Francisco Sales de Lima Filho and 1 Frederico A. F. Silveira proposed the method for identify the DoS/DDoS attack by applying the Random Forest Tree algorithm, which segregates the samples collected from the flow protocol directly from network devices. It gives the better results with respect to better data rate, false alarm rate and precision. With this it is possible to to separate the attacks such as TCP, UDP and HTTP flood, HTTP slow. The authors in [11] Uzma Khana, Shikha Agrawala and Sanjay Silakaria proposed that Node centric detection scheme, this algorithm provides the better throughput, high packet delivery ratio and less end to end delay considers.

When considered the load, distance and distrust value, based on these parametersThey suggested that PSO can be selected to enhance the optimization. The author in [12] Abdulaziz Alshammari, Mohamed A. Zohdy, Debatosh Debnath and George Corser proposed the classification approach for the IDS in Vehicular networks, Machine learning techniques to cluster and classify the intrusions in VANET by applying KNN and SVM algorithms. It uses the CAN protocol by considering its request and response time for evaluating the time interval and offset ratio. Focus of this is to compare and detect the Dos attack and Fuzzy. Here KNN gave better results as compared to SVM on predefined data sets. CAN is more prone to attacks as it is a bus communication protocol. The author in [13] Khaoula Jeffane and Khalil Ibrahimi proposed the method DoS attacks detection with packet delivery ratio metric on the physical and MAC layers. Here black list is encrypted which is to be sent to the Road side unit by distributing to the users of the network to thwart packets sent by attackers. Also, they have suggested that DOS problem can be studied based on the stochastic learning game. The author in [14] Wenjia Li proposed the An SVM based Security Framework for VANET, SVM is used as a classifier to classify and train the patterns of the misbehaving nodes It gives better results with higher precision and recall values with less communication overhead. The author in [15] Kajal Rai et.al worked on the IDS based on the Decision Tree algorithm, which gives near about 80% uniform weightage to all attribute values. Drawback of this system is unknown attacks detection fails and no data Pre-processing is done. The author in [16] Mabayoje also proposed the Gain ratio and decision tree classifier for intrusion detection, in which 97 % accuracy for

DoS attack and drawback is prone to unknown attacks and no pre-processing is done on the data. K. M. A. Alheeti, A. Gruebler, and K. D. McDonald proposed a method to identify the Black Hole attack by observing the anomalies and misuse of the network, and Neural Network ML approach is used. It is observed in this paper accuracy and efficiency is less for detection of Black Hole attack because they used less data for training [17].

III. Methodology and Workflow

In ML approach, a huge variety of methods are developed for the classification approach such as K-NN, SVM, ANN, Deep Learning/CNN, out of that SVM is a most demanding and robust classifier. SVM uses the kernel functions to map the data of non –linearly to high dimensional space.[18] In proposed work, as indicated in Fig. 1 firstly we collect the data from different sources here in this paper we have collected data from kaggle for DDoS attack, this is the input to the system. Then we are training this data by considering the 70-30 percent ratio. 70 % data is trained properly. After that we are applying the SVM algorithm for classification purpose to detect the normal and malicious activity in database we have used the DDoD attack field. Then we are applying the KNN algorithm for finding the nearest neighbours which are trained and tested properly then simulation is done on dataset and it gives the results like if malicious node or attack is detected it predicts and provide security otherwise it send the normal report [19]. For this work input is taken as a text file of collected different IP addresses.

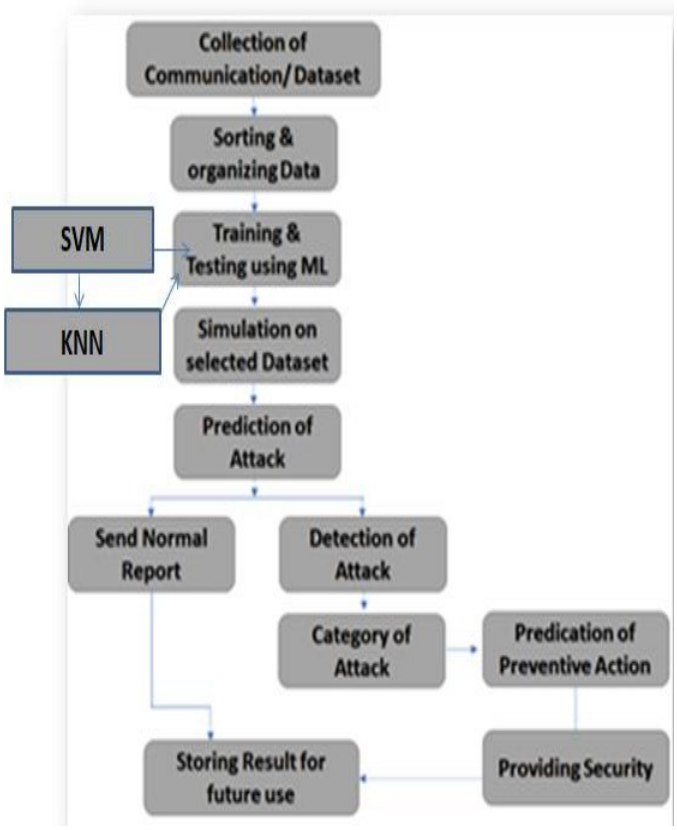

Fig. 1. Proposed System Workflow.

Proposed hybrid algorithm of KSVM for training and validating the data:

- Training set Ta={x1,x2,x3..xn}
- Testing set Tb={x1,x2,…xn}
- T=Ta U Tb

1. begin
   int i,j,k; // test samples i=1,2,3..n , j=1,2,3,..m.
2. SVM:T- > SV_{ij};
3. input X_k;Calculate nearest neighbour dist. of X_k and SV_{ij};
4. Put in to the nearest neighbour
5. k=k+1;// add all samples

   Repeat step 3 to 5 until all samples values are calculated end.

Proposed Algorithm for data prediction and Accuracy is given below:

Algorithm (Predict, Accur) = Hybrid KSVM (W+O) (Train, Div, Test, Test-Final, θ)

Require: Train, Div, Test, Test-Final, Datasets;

θ - Termination condition

Ensure: Predict→ Predicted sentiment output;

Accur→ Accuracy

1: Net → Create_Network()
2: Network_initialize(Net)
3: for error >= θ do
4: error Netwrok_Train(Net, Train, Div)
5: end for
6: /* Training completed */
7: Featureopt → CSVM (Train, Div)
8: HTrain → GetTop_HiddenLayer (Net, Train)
9: Train_combined ← HTrain + Featureopt
10: ModelSVM ← SVMLinear (Traincombined)
11:HTest←GetTop_HiddenLayer(Net, Test)
12: Testcombined← HTest + Featureopt
13: Predict ←SVMLineart(ModelSVM , Testcombined)
14: Accur← Evaluation (Test-Final, Predict)
15: return (Predict, Accur)

By using above algorithm, we are trying to find the misbehaving activity in network with secure communication between the Autonomous Vehicles.

IV. Experimental Results and Discussion

The aim of this proposed work is to detect DDoS attack in a vehicular network where the network is choked. This would prevent those cars from having enough access to systems. This leads to the field of prediction using the machine learning method, to incorporate more accurate and stable security results. Client-server model is created for the experiment in python.

The DDoS attack is a method to spread an internet service or web-site by overburdening it with massive traffic floods caused by several outlets. In contrast to a service-default attack, that involves one device and an Internet link, a DDoS attack uses several machines and several Internet access. Data is collected from Kaggle dataset. Name of dataset is DDos data. The attributes of dataset are like protocol which is used for communication, source and destination IP address, source and destination port no. etc., predefined IP address is considered for this work to detect the malicious activity, if any message encountered in the system with different IP address which not from the predefined set then this is detected as malicious node. The evaluation criteria are based on the prediction and accuracy algorithm which performs the training and testing of data from the data set. Implementation is done using the python language. Different ML algorithms are considered which are shown below in the simulation results as indicated that our KSVM hybrid algorithm is giving the better results with different parameters like in the Fig. 2, the accuracy of data is more when we compare other ML algorithms. In Fig. 3, sensitivity results are better by comparing other ML algorithms. In Fig. 4, precision and in Fig. 5, recall parameters are giving the better results. And in Fig. 6, error percentage is less as compared to other algorithms of machine learning algorithms.

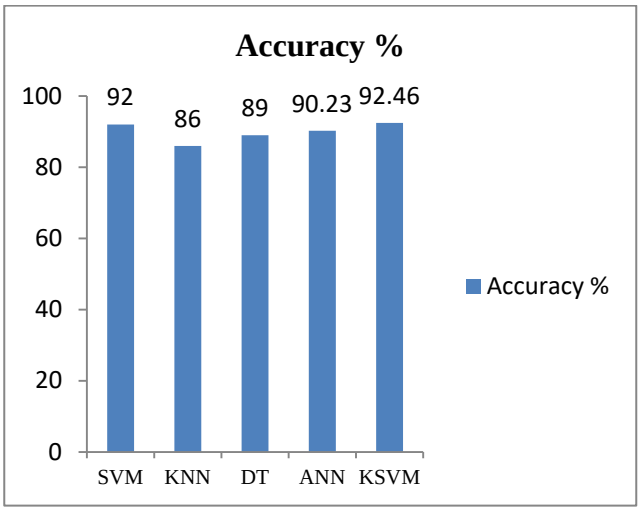

Fig. 2. Accuracy % as Compared to other ML Algorithms.

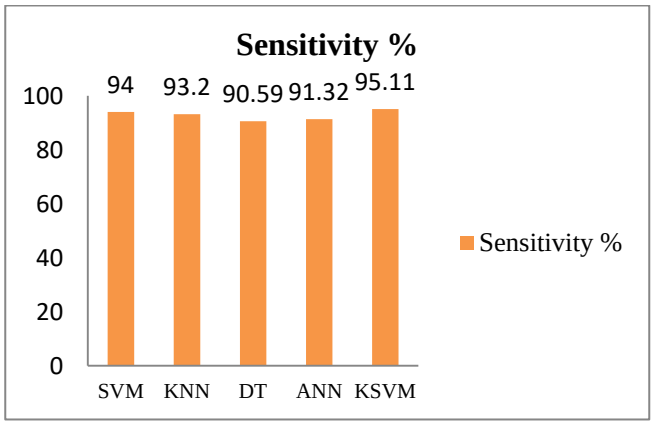

Fig. 3. Sensitivity % as Compared to other ML Algorithms.

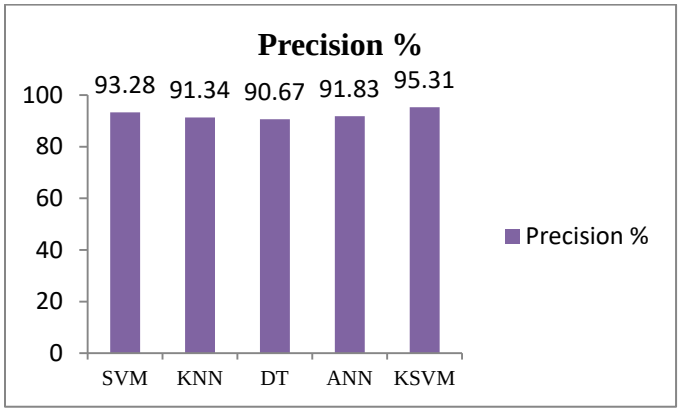

Fig. 4. Precision % as Compared to other ML Algorithms.

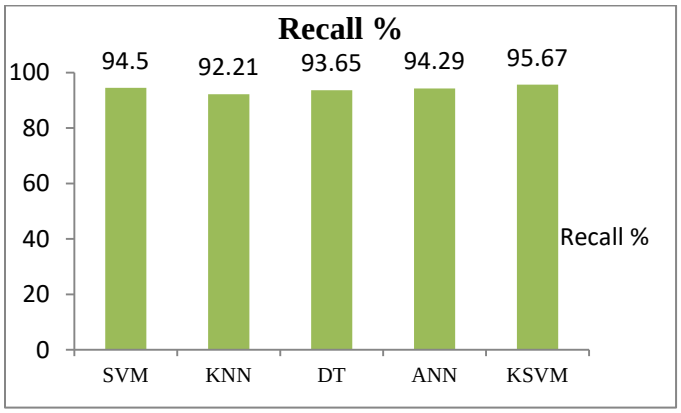

Fig. 5. Recall % as Compared to other ML Algorithms.

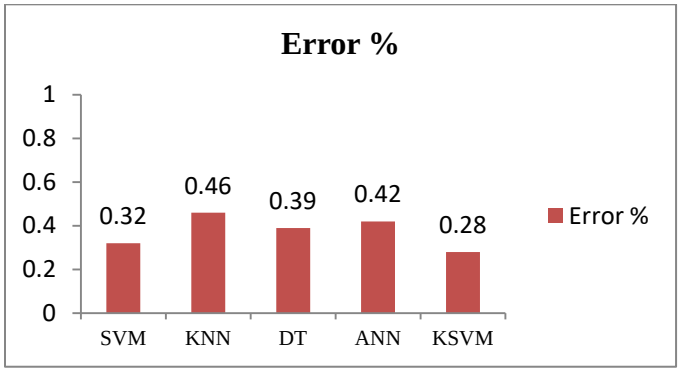

Fig. 6. Error % as compared to other ML algorithms

V. Conclusion and Future Scope

Here in this work, different ML approaches and proposed algorithm Hybrid KSVM for DDoS attack to detect malicious activity is considered with respect to a. accuracy b. sensitivity c. precision d. recall and e. error are shown in above figures and as compare to other ML algorithms our Hybrid KSVM is giving better results. Future scope of this work is further we can apply the same method to detect different attacks like Dos, Sybil. The more secure framework can be built for communication which preserves the privacy and integrity of the message in transit by considering different network security algorithms.

REFERENCES

[1] Kadam N.N., Sekhar K.R. (2019), "Secure and congestion free routing techniques in vehicular Ad-Hoc Network (VANET)", International Journal of Recent Technology and Engineering, 8(0), PP.915-922.

[2] Hao ye, le liang, geoffrey ye li, joonBeom Kim, lu lu, and may Wu," Machine Learning For Vehicular Networks",2018 IEEE

[3] Zargar S T, Joshi J, Tipper D.,"A survey of defense mechanisms against distributed denial ofservice(DDoS)flooding attacks". IEEE Corn— munications Surveys & Tutorials.2013, 15(4) : 2046—2069.

[4] Jiangtao Pei1, Yunli Chen1*, Wei Ji1*,"A DDoS Attack Detection Method Based on Machine Learning",ICSP 2019 conference.

[5] Stefan Mihai, Nedzhmi Dokuz, Meer Saqib Ali, Purav Shah, and Ramona Trestian" Security Aspects of Communications in VANETs",. @2020IEEE.

[6] Elvin Eziama, Kemal Tepe, Ali Balador, Kenneth Sorle Nwizege and Luz M. S. Jaimes ," Malicious Node Detection in Vehicular Ad-Hoc Network Using Machine Learning and Deep Learning,2018 IEEE.

[7] Carlos H. O. O. Quevedo, Ana M. B. C. Quevedo, Ahmed Serhrouchni , "An Intelligent Mechanism for Sybil Attacks Detection in VANETs", 978-1-7281-5089-5/20/$31.00 ©2020 IEEE.

[8] Fabio Gonc¸alves, Bruno Ribeiro, Oscar Gama,"A Systematic Review on Intelligent Intrusion Detection Systems for VANETs", 978-1-7281-5764-1/19/$31.00 ©2019 IEEE.

[9] Mohammad Asif Hossain,Wahidah Md Shah, "Faster Convergence of Q-Learning in Cognitive Radio-VANET Scenario" , Springer 2019.

[10] Steven So , Prinkle Sharma, Jonathan Petit ," Integrating Plausibility Checks and Machine Learning for Misbehavior Detection in VANET",2018IEEE.

[11] Francisco Sales de Lima Filho ,1 Frederico A. F. Silveira," Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using MachineLearning", HindawiSecurity and Communication Networks Volume 2019, Article ID 1574749.

[12] Uzma Khana, Shikha Agrawala, Sanjay Silakaria, "Detection of Malicious Nodes (DMN) in Vehicular Ad-Hoc Networks", ELSEVIER, 2015.

[13] Abdulaziz Alshammari1, Mohamed A. Zohdy1, Debatosh Debnath1, George Corser2, "Classification Approach for Intrusion Detection in Vehicle Systems" , Wireless Engineering and Technology SciRP 2018.

[14] Khaoula Jeffane, Khalil Ibrahimi, "Detection and identification of attacks in Vehicular Ad-Hoc Network" , IEEE 2016.

[15] WANG TONG, AZHAR HUSSAIN , WANG XI BO , AND SABITA MAHARJAN ,"Artificial Intelligence for Vehicle-to-Everything: a Survey", 2169-3536 (c) 2019 IEEE.

[16] Kajal Rai et.al," Decision tree based algorithm for intrusion detection", International Journal Advanced Networking and Applications,2015.

[17] Mabayoje, "Gain ratio and decision tree classifier for intrusion detection", Int. Journal of Computer applications.2015.

[18] K. M. A. Alheeti, A. Gruebler, and K. D. McDonald- Maier, "An Intrusion Detection System against Black Hole Attacks on the Communication Network of Self-Driving Cars," 2015 Sixth International Conference on Emerging Security Technologies (EST), pp. 86–91, 2015.

[19] Gulnaz Alimjan, Tieli Sun, Yi Liang, Hurxida Jumahun and Yu Guan,."A New Technique for Remote Sensing ImageClassi¯cation Based on Combinatorial Algorithm of SVM and KNN", International Journal of Pattern Recognitionand Artificial Intelligence, Vol. 32, No. 7 2018.