AN INNOVATIVE METHOD TO ENHANCE THE ACCURACY IN CLASSIFICATION OF CAPTCHA RECOGNITION BY USING K-NEAREST NEIGHBOURS (KNN) ALGORITHM OVER SUPPORT VECTOR MACHINE

Gudaru Gopi Prasad[1], K. Malathi[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: Recognition of captcha to find the Best accuracy of the text based captcha by using the Algorithms in Machine Learning. The two algorithms are K-Nearest Neighbour(KNN) and the Novel Support Vector Machine(SVM).

Materials and Methods: The dataset is collected from www.kaggle.com. And the Two groups are K-Nearest Neighbour (N=10) and Novel Support Vector Machine(N=10) by using G-power and minimum power of the analysis is fixed as 80% and maximum accepted error is fixed as 0.5 with threshold value as 0.0805% and Confidence Interval is 95%.

Results: The Novel Support vector machine is used to recognize the captcha. The accuracy found for the improved captcha is 99.98% and for the K-Nearest Neighbour is 98.78%. The two algorithms are used to find the improved classification or complexity of the captcha. The significant value obtained is $p=0.002$ ($p<0.05$) i.e $\alpha=0.05$ and hence, there exists a statistically significant difference between the two groups with a confidence level of 95%.

Conclusion: Recognizing the Captcha Recognition significantly seems to be better in Novel Support vector machines than k-nearest Neighbour.

Keywords: K-Nearest Neighbour, Novel Support Vector Machine, Machine Learning, Text based, Captcha Recognition, Complexity.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai,Tamil Nadu, India, 602105.

[2*]Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India, 602105.

1. Introduction

It is one of the simplest classification algorithms available for supervised learning. Tha main idea is to be the closest match of the test data in the feature space of the digital image preprocessing. This algorithm will be increasing the complexity of the captcha Recognition (Z. Wang and Sun 2009). In today's world, this captcha has searched for closed matches of the test data in features space of the image. Manual Human test acknowledgement is a powerful method to keep up with network security and keep noiseus assaults from computer project.It has been generally utilised in major standard sites. Manual Human test (Captcha Recognition) is for the most past thought of to be converge turing test to order peoples and computers. This text based Captcha will be used or help us from the Hackers. By using this captcha ,the captcha has the main stream on the website to be accessed (Zbancioc and Feraru 2012). K-Nearest Neighbour(KNN) for Captcha Recognition(Kwon, Yoon, and Park 2020a). Also this article Captcha recognition and trends has become even more popular than before(Kwon, Yoon, and Park 2020a). The one more article is a survey on machine learning for Captcha recognition algorithms and techniques(Huber 2019; Sha 2011). Mainly the Novel Support vector Machine algorithm, these applications are to research time series recognition and Captcha recognition using Novel Support Vector Machine.

Captcha Recognition can be carried out from the researchers. There are 750 articles found on IEEE, and 536 articles were found in the Google Scholar. And the Capctha recognition using K-Nearest Neighbour can be found that of 98.78% with the accuracy (Singh et al. 2016) and then the captcha recognition by characters using Novel Support vector machine this algorithm brings the 99.98%. The captcha recognition also used the arabic languages like chinese language using digital image preprocessing with the accuracy of 97.85% (Kwon, Yoon, and Park 2020b) finally the key step of the captcha segmentation to extract individual letters has been reasonably successful with accuracy of 96.45% ("Video for A Low Complexity SVM Classifier for EEG Based Gesture Recognition Using Stochastic Computing," n.d.; Huber 2019).

Our team has extensive knowledge and research experience that has translated into high quality publications (K. Mohan et al. 2022; Vivek et al. 2022; Sathish et al. 2022; Kotteeswaran et al. 2022; Yaashikaa, Keerthana Devi, and Senthil Kumar 2022; Yaashikaa, Senthil Kumar, and Karishma 2022; Saravanan et al. 2022; Jayabal et al. 2022; Krishnan et al. 2022; Jayakodi et al. 2022; H. Mohan et al. 2022). The K-Nearest Neighbour has very little precision (accuracy), the exactness of the image is shown at an outstandingly low rate while examining the image and the manual data is ludicrous to hope to add to the dataset (Tamir 2019). The survey intends to chip away at the precision(accuracy) of the images, further fostering the exactness rate of the apparent images, and diminish the lack of data while planning and testing the dataset and also the training the dataset (Alaidi, Alsaidi, and Yahya 2017). The aim of this K-Nearest Neighbour(KNN) algorithm is implemented by using Novel Support Vector Machine to improve accuracy of Captcha Recognition.

2. Materials and Methods

The study setting of the proposed work is done in Saveetha School of Engineering, OOAD lab. The study setting of the proposed work is done in our university laboratory. The sample size was calculated by using clincalc.com by keeping G power (Mayr et al. 2007) and minimum power of the analysis is fixed as 0.8 and maximum accepted error is fixed as 0.5 with threshold value as 0.05% and Confidence Interval is 95%. Mean and standard deviation has been calculated based on the previous literature for size calculation. The two groups are used namely K-Nearest Neighbour (N=10) as an existing model as group 1 and Novel Support Vector Machine (N=10) as a Proposed model as Group 2.

Data Preparation

The Convolutional Neural Network is to find all the Captcha tests that are put away in the dataset, to prepare and test through The dataset comes from Wilhelmy, Rodrigo and Rosas, Horacio. (2013). manual human test dataset. The dataset remembers 10000 information for the type of pictures which are taken as an example from www.kaggle.com. with their separate penmanship styles. There are 1000 prepared pictures and 9000 tried pictures. The example pictures of manual human test present in the dataset has been displayed in Fig. 1 with the 36*36 pixel size.

K-Nearest Neighbour (KNN)

It is one of the simplest algorithms of classification of algorithms available for Supervised Learning. The idea is to search for the closest match of the test data in the featured space of the digital image preprocessing. The training and testing data set, all are hitting the accuracy nearly of 98.78% and also improving the complexity of the captcha

recognition. The Fig. 2 will be implemented as below

- Give the input as image with the 36*36 pixel size.
- Then it initialises the decoding of the image.
- After that it will be proposing the image and then it will be reading the image.
- It will be segmented and finally it will remain and return the image and letter or text based on the accuracy level of that image.

Support Vector Machine (SVM)

The utilising Novel Support vector machine calculation to do the Captcha acknowledgment. All might utilise that equivalent theta to move the obscure information to group it by the straight SVM model. This Novel Support vector machine model will work on the exactness of the other calculation or model.hen the default number of Novel Support vector machine components is shaky, it has prevailed with regards to accomplishing 92% free captchas, yet just 44% of the test components have been accomplished. With the boundary all have upgraded, the acknowledgment of the complexity captcha sound document has reached 89%. Free digital image preprocessing achievement has arrived at a generally excellent worth of 98%. The Fig. 3 will be implemented as below

- Downloading the dataset to stack.
- Introduce the factors to prepare and test the information.
- Characterise a model. fit () capacity to depict the parts which are to be gotten to for running the code to get precision.
- Characterise Categorical () capacity to arrange the information.
- Print the model. fit () work with the necessary ages and discover the exactness.

For comparing both the models, the dataset has been trained with ten different sample sizes. the accuracy values are recorded. The system configuration is used for the algorithm to run in a 64 - bit Operating System, 4GB RAM PC, and using Windows 10, Google Colab, and Microsoft Office for software specification. To assess the exhibition of the training model, the information has been parted for training and testing to approve the dataset. Then, at that point, stack and reshape the information clusters to arrange the numbers. Standardise the pixel upsides of grayscale pictures All the layers will be worked through the ReLU enactment capacity to the absolute cross_entrophy to discover the misfortune work. The model will be assessed with the fit() work which has the measurements capacity to approve the precision and loss of the information.

Statistical Analysis

The SPSS statistical software was used in the research for statistical analysis. Group statistics and independent sample t-tests were performed on the experimental results and the graph was built for two groups with two parameters under study. The independent variables are Datasets. The dependent variables are shape,size,Accuracy. The independent T test analysis was carried out to calculate the accuracy for both methods(Selamat, Hakeem Selamat, and Rais 2015).

3. Results

K-Nearest Neighbour calculation frames the layers with every one of the pictures of each number, at whatever point it runs at various times because of the introduction of test size (N=10).The layers are shaped because of the cycles, the precision(accuracy) esteem changes with the span of running time and delivers the exactness and misfortune concerning the period which is displayed in Table 1. K-Nearest has preferable exactness and less misfortune over the support vector machine. Because of the enhancement capacities and measurements, the K-Nearest neighbour calculation has not visible with the progressed actuation capacities which are simply limited to the adam, adaleta, and adagrad which takes additional time and the capacities are not taking the entire information to break down the comparable digits in the dataset through the Support vector machine takes the information and structures layers with every captcha separately lastly gives the outcome. Concerning the enhancement capacities, the Accuracy and misfortune have changed and have demonstrated that the support vector machine is superior to k-nearest neighbour. Table1 represents the data collection from the N=10 samples of the dataset for k-nearest neighbour with the size of 36*36 pixels to gain accuracy (%) and support vector machines to gain accuracy (%). Table2 represents the data collection from the N=10 samples of the dataset for K-Nearest Neighbour with the size of 28*28 pixels to gain accuracy (%) and support vector machines to gain accuracy (%) .

$$\textbf{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{--(1)}$$

Where, in equation (1)
TP = True Positive
TN = True Negative
FP = False Positive
FN = False Negative

The IBM SPSS version 21 statistical software is used for our study. The independent variables are shape and size and the dependent variable is accuracy (%) for our study Captcha Recognition.
In SPSS, the datasets are prepared using N=10 as the sample size for K-Nearest Neighbour and Support Vector Machine. GroupID is given as a grouping variable and accuracy is given as the testing variable. GroupID is given as 1 for support vector machines and group 2 for K-Nearest Neighbours. Group Statistics is applied for the Statistical Package for the Social Sciences (SPSS) dataset and shown in Table 2. By performing the statistical analysis group statistics represents the comparison of the accuracy and Loss of Captcha recognition of k-nearest neighbour and support vector machine. The support vector machine algorithm had the highest accuracy (99.98%). K-Nearest Neighbour had the lowest accuracy (98.78%) in Table 2. Table 3 represents the Independent Sample T-Test that is applied for the sample collections by fixing the level of significance as 0.005 with a confidence interval of 95 %. After applying the SPSS calculation, the support vector machine has accepted a statistically significant value($p<0.05$). From Fig. 4 it was represented by a simple bar Mean of Accuracy K-Nearest Neighbour error range (0.99 - 0.98) and support vector machine error range (0.99 - 0.98) .

4. Discussion

General outcomes show that there are a few varieties saw in the precision and misfortune esteems due to the headways of the enhancement capacities which demonstrated that the Support vector Machine with a precision of 99.98% is superior to the K-Nearest Neighbour with an exactness of 98.78% in perceiving the captcha. There is a genuinely huge distinction in creative Captcha Recognition exactness of two calculations having the critical precision worth of 0.001 ($p<0.005$ Independent Sample t-Test). The institution is passionate about high-quality text based research and has excelled in various fields (Brodić and Amelio 2019). All hope this study adds to this rich legacy. In the future, all can improve this classification for further development in K-Nearest Neighbour Architecture and the applications of some big complexity noisy data. All can improve this system to recognize characters in different languages. Systems are to be developed to analyse over Retinal Real-time Captcha Image ("Captcha Recognition Using GAN" 2020a). Previously our team has a rich experience in working on various research projects across multiple disciplines ((Kwon, Yoon, and Park 2020b),(Jia et al. 2018),(Shu and Xu 2019a; Panwar et al. 2018a),(Shu and Xu 2019b; Panwar et al. 2018b),(Shu and Xu 2019b),(Shu and Xu 2019b; J. Wang et al. 2019),(Shu and Xu 2019b; J. Wang et al. 2019; "Captcha Recognition Using GAN" 2020b),(Shu and Xu 2019b; J. Wang et al. 2019; "Captcha Recognition Using GAN" 2020b; Sachdev 2020),(Singh et al. 2016),(T., R., and J. 2016),(Alsuhibany and Parvez 2016),(Alsuhibany and Parvez 2016; Panwar et al. 2018a),(Cao 2021),(Selamat, Hakeem Selamat, and Rais 2015)). Now the growing trend in this area motivated us to pursue this project.(Kwon, Yoon, and Park 2020a).

5. Conclusion

In this examination, the inventive Captcha Recognition framework utilises Wilhelmy, Rodrigo & Rosas, Horacio (2013). The captcha dataset is by all accounts better precision (99.98%) utilising Support Vector Machine than K-Nearest Neighbour (98.78%). The lucidity of Captcha is found with acceptable precision and less misfortune is accomplished and the complexity of the manual human test (Captcha recognition) is improved.

Declarations

Conflicts of Interest

No conflicts of interest in this manuscript.

Authors Contribution

Author GGP was involved in data collection, data analysis, and manuscript writing. Author KM was involved in conceptualization, guidance and critical review of manuscript.

Acknowledgments

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

We thank the following organisations for providing financial support that enabled us to complete the study.

1. SM DSQUARE interiors and solutions Pvt.Ltd, Bangalore.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

6. References

Alaidi, Abdul Hadi M., Saif Ali Abd Alsaidi, and Omar Hashim Yahya. 2017. "Plate Detection and Recognition of Iraqi License Plate Using

KNN Algorithm." Journal of Education College Wasit University. https://doi.org/10.31185/eduj.vol1.iss26.102.

Alsuhibany, Suliman A., and Mohammad Tanvir Parvez. 2016. "Secure Arabic Handwritten CAPTCHA Generation Using OCR Operations." 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR). https://doi.org/10.1109/icfhr.2016.0035.

Brodić, Darko, and Alessia Amelio. 2019. The CAPTCHA: Perspectives and Challenges: Perspectives and Challenges in Artificial Intelligence. Springer Nature.

Cao, Yu. 2021. "Digital Character CAPTCHA Recognition Using Convolution Network." 2021 2nd International Conference on Computing and Data Science (CDS). https://doi.org/10.1109/cds52072.2021.00029.

"Captcha Recognition Using GAN." 2020a. Journal of Xidian University. https://doi.org/10.37896/jxu14.6/325.

2020b. Journal of Xidian University. https://doi.org/10.37896/jxu14.6/325.

Huber, Manuel. 2019. CAPTCHA Recognition with Adaptable Neural Networks.

Jayabal, Ravikumar, Sekar Subramani, Damodharan Dillikannan, Yuvarajan Devarajan, Lakshmanan Thangavelu, Mukilarasan Nedunchezhiyan, Gopal Kaliyaperumal, and Melvin Victor De Poures. 2022. "Multi-Objective Optimization of Performance and Emission Characteristics of a CRDI Diesel Engine Fueled with Sapota Methyl Ester/diesel Blends." Energy. https://doi.org/10.1016/j.energy.2022.123709.

Jayakodi, Santhoshkumar, Rajeshkumar Shanmugam, Bader O. Almutairi, Mikhlid H. Almutairi, Shahid Mahboob, M. R. Kavipriya, Ramesh Gandusekar, Marcello Nicoletti, and Marimuthu Govindarajan. 2022. "Azadirachta Indica-Wrapped Copper Oxide Nanoparticles as a Novel Functional Material in Cardiomyocyte Cells: An Ecotoxicity Assessment on the Embryonic Development of Danio Rerio." Environmental Research 212 (Pt A): 113153.

Jia, Yang, Wang Fan, Chen Zhao, and Jungang Han. 2018. "An Approach for Chinese Character Captcha Recognition Using CNN." Journal of Physics: Conference Series. https://doi.org/10.1088/1742-6596/1087/2/022015.

Kotteeswaran, C., Indrajit Patra, Regonda Nagaraju, D. Sungeetha, Bapayya Naidu Kommula, Yousef Methkal Abd Algani, S. Murugavalli, and B. Kiran Bala. 2022. "Autonomous Detection of Malevolent Nodes Using Secure Heterogeneous Cluster Protocol." Computers and Electrical Engineering. https://doi.org/10.1016/j.compeleceng.2022.107902.

Krishnan, Anbarasu, Duraisami Dhamodharan, Thanigaivel Sundaram, Vickram Sundaram, and Hun-Soo Byun. 2022. "Computational Discovery of Novel Human LMTK3 Inhibitors by High Throughput Virtual Screening Using NCI Database." Korean Journal of Chemical Engineering. https://doi.org/10.1007/s11814-022-1120-5.

Kwon, Hyun, Hyunsoo Yoon, and Ki-Woong Park. 2020a. "CAPTCHA Image Generation: Two-Step Style-Transfer Learning in Deep Neural Networks." Sensors 20 (5). https://doi.org/10.3390/s20051495.

2020b. "CAPTCHA Image Generation: Two-Step Style-Transfer Learning in Deep Neural Networks." Sensors 20 (5). https://doi.org/10.3390/s20051495.

Mayr, Susanne, Edgar Erdfelder, Axel Buchner, and Franz Faul. 2007. "A Short Tutorial of GPower." Tutorials in Quantitative Methods for Psychology. https://doi.org/10.20982/tqmp.03.2.p051.

Mohan, Harshavardhan, Sethumathavan Vadivel, Se-Won Lee, Jeong-Muk Lim, Nanh Lovanh, Yool-Jin Park, Taeho Shin, Kamala-Kannan Seralathan, and Byung-Taek Oh. 2022. "Improved Visible-Light-Driven Photocatalytic Removal of Bisphenol A Using V2O5/WO3 Decorated over Zeolite: Degradation Mechanism and Toxicity." Environmental Research. https://doi.org/10.1016/j.envres.2022.113136.

Mohan, Kannan, Abirami Ramu Ganesan, P. N. Ezhilarasi, Kiran Kumar Kondamareddy, Durairaj Karthick Rajan, Palanivel Sathishkumar, Jayakumar Rajarajeswaran, and Lorenza Conterno. 2022. "Green and Eco-Friendly Approaches for the Extraction of Chitin and Chitosan: A Review." Carbohydrate Polymers 287 (July): 119349.

Panwar, Pooja, Monika, Parveen Kumar, and Ambalika Sharma. 2018a. "CHGR: Captcha Generation Using Hand Gesture Recognition." 2018 Conference on Information and Communication Technology (CICT). https://doi.org/10.1109/infocomtech.2018.8722409.

2018b. "CHGR: Captcha Generation Using Hand Gesture Recognition." 2018 Conference on Information and Communication Technology (CICT). https://doi.org/10.1109/infocomtech.2018.8722409.

Sachdev, Sumeet. 2020. "Breaking CAPTCHA Characters Using Multi-Task Learning CNN and SVM." 2020 4th International Conference on Computational Intelligence and Networks (CINE). https://doi.org/10.1109/cine48825.2020.234400.

Saravanan, A., P. Senthil Kumar, B. Ramesh, and S. Srinivasan. 2022. "Removal of Toxic Heavy Metals Using Genetically Engineered Microbes: Molecular Tools, Risk Assessment and Management Strategies." Chemosphere 298 (July): 134341.

Sathish, T., R. Saravanan, V. Vijayan, and S. Dinesh Kumar. 2022. "Investigations on Influences of MWCNT Composite Membranes in Oil Refineries Waste Water Treatment with Taguchi Route." Chemosphere 298 (July): 134265.

Selamat, M. Hakeem, M. Hakeem Selamat, and Helmi Md Rais. 2015. "Image Face Recognition Using Hybrid Multiclass SVM (HM-SVM)." 2015 International Conference on Computer, Control, Informatics and Its Applications (IC3INA). https://doi.org/10.1109/ic3ina.2015.7377765.

Sha, Nong. 2011. Breaking Image Recognition CAPTCHA: An Approach Using Web-Trained Classifiers.

Shu, Yujin, and Yongjin Xu. 2019a. "End-to-End Captcha Recognition Using Deep CNN-RNN Network." 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). https://doi.org/10.1109/imcec46724.2019.8983895.

2019b. "End-to-End Captcha Recognition Using Deep CNN-RNN Network." 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). https://doi.org/10.1109/imcec46724.2019.8983895.

Singh, Bhargavi, M. Tech Scholar, G.I.E.T, and Sonipat. 2016. "Iris Recognition Using Curve Let Transformation Based on Gabor Filter& SVM." International Journal Of Engineering And Computer Science. https://doi.org/10.18535/ijecs/v5i8.32.

Tamir, Kassahun. 2019. "Handwritten Amharic Characters Recognition Using CNN." 2019 IEEE AFRICON. https://doi.org/10.1109/africon46755.2019.9133925.

T., Manoj, Basavaraj R., and Jyoti J. 2016. "Security Enhancement in Captcha Recognition Using Animated GIF Images." International Journal of Computer Applications. https://doi.org/10.5120/ijca2016912473.

"Video for A Low Complexity SVM Classifier for EEG Based Gesture Recognition Using Stochastic Computing." n.d. https://doi.org/10.1109/iscas45731.2020.9180431/video.

Vivek, J., T. Maridurai, K. Anton Savio Lewise, R. Pandiyarajan, and K. Chandrasekaran. 2022. "Recast Layer Thickness and Residual Stress Analysis for EDD AA8011/h-BN/B4C Composites Using Cryogenically Treated SiC and CFRP Powder-Added Kerosene." Arabian Journal for Science and Engineering. https://doi.org/10.1007/s13369-022-06636-5.

Wang, Jing, Jiao Hua Qin, Xu Yu Xiang, Yun Tan, and Nan Pan. 2019. "CAPTCHA Recognition Based on Deep Convolutional Neural Network." Mathematical Biosciences and Engineering: MBE 16 (5): 5851–61.

Wang, Ziqiang, and Xia Sun. 2009. "Image Retrieval Using Discriminant Embedding and LS-SVM." 2009 Chinese Conference on Pattern Recognition. https://doi.org/10.1109/ccpr.2009.5344079.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Algal Biofuels: Technological Perspective on Cultivation, Fuel Extraction and Engineering Genetic Pathway for Enhancing Productivity." Fuel. https://doi.org/10.1016/j.fuel.2022.123814.

Yaashikaa, P. R., P. Senthil Kumar, and S. Karishma. 2022. "Review on Biopolymers and Composites – Evolving Material as Adsorbents in Removal of Environmental Pollutants." Environmental Research. https://doi.org/10.1016/j.envres.2022.113114.

Zbancioc, Marius, and Silvia Monica Feraru. 2012. "Emotion Recognition of the SROL Romanian Database Using Fuzzy KNN Algorithm." 2012 10th International Symposium on Electronics and Telecommunications. https://doi.org/10.1109/isetc.2012.6408133.

Tables and Figures

Table 1. Data collection from the N=10 samples of the dataset for K-Nearest Neighbour with the size of 36*36 pixels to gain accuracy (%) and Support vector machine to gain accuracy (%).

Samples (N)	K-Nearest Neighbour (KNN)	Support Vector Machine (SVM)
	Accuracy (%)	Accuracy (%)
1	98.78	99.98
2	98.56	99.56
3	97.48	98.48
4	96.87	97.48
5	96.14	97.89
6	95.79	96.87
7	94.58	96.12
8	95.06	95.79
9	94.18	94.89
10	93.14	93.58

Table 2. Comparison of the accuracy of Captcha Recognition of K-Nearest Neighbour and support vector machine. The support vector machine algorithm had the highest accuracy (99.98%) and K-Nearest Neighbour had the lowest accuracy (98.78%).

Group Statistics

GROUP		N	Mean	Std Deviation	Std Error Mean
ACCURACY	SVM	10	99.9830	2.74173	.86701
	KNN	10	98.7800	.78249	.24744

Table 3. Independent Sample T-Test is applied for the sample collections by fixing the level of significance as 0.05 with confidence interval as 95%. After applying the SPSS calculation, the Support vector machine has accepted a statistically significant value ($p<0.05$).

	Levene's Test for Equality of Variance		T-test for Equality of Means						
	f	Sig	t	df	Sig.(2-tailed)	Mean Difference	Std.Error Difference	95% Confidence of the Differences	
								Lower	Upper

Accuracy	**Equal variances assumed**	12.421	.002	4.965	18	.000	4.47700	.90163	2.47987	6.37125
	Equal variances not assumed			4.965	10.456	.000	4.47700	.90163	2.47987	6.37125

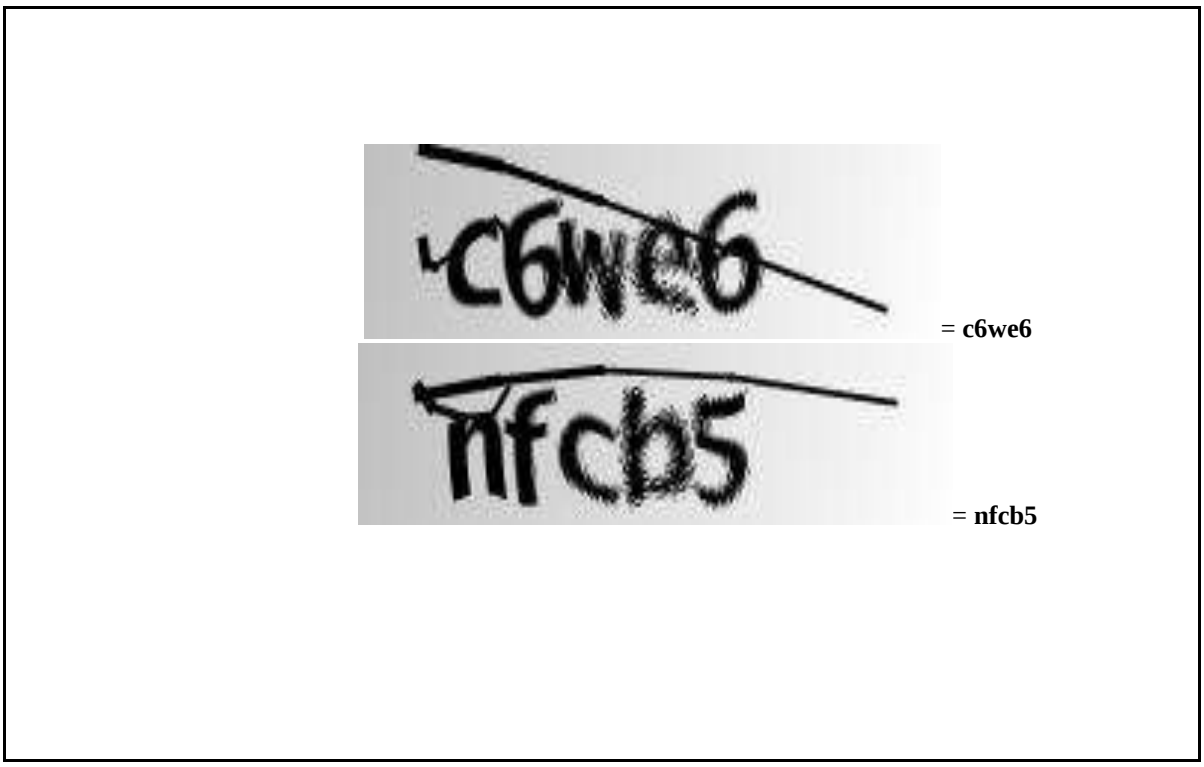

Fig. 1. Captchas from the dataset

Fig. 2. Flowchart of K-Nearest Neighbour Algorithm

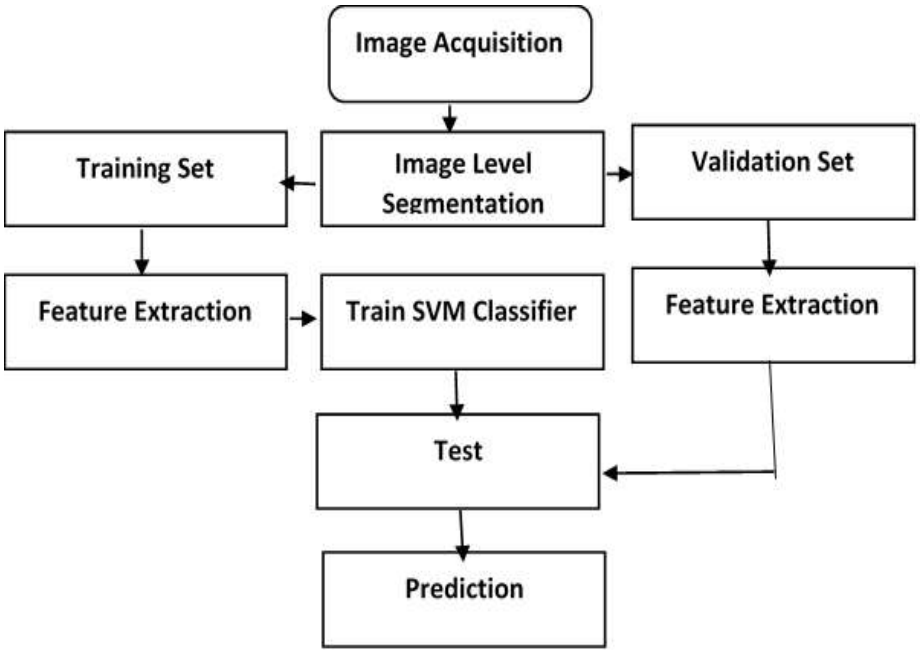

Fig. 3. Flowchart of Support Vector Machine

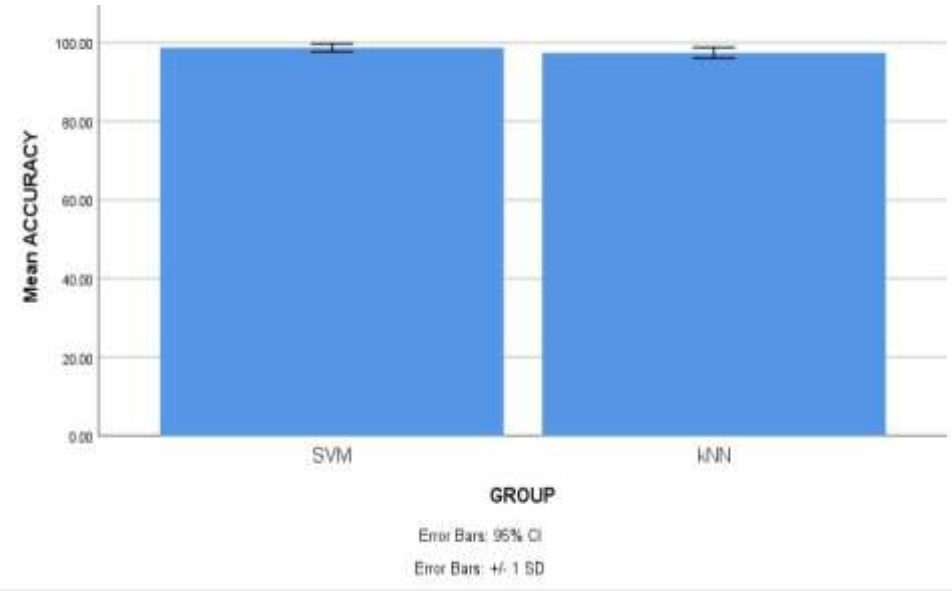

Fig. 4. Simple Bar Mean of Accuracy k-nearest neighbour error range (0.99 - 0.98) and support vector machine error range (0.98- 0.99) and with Mean accuracy of detection ± 1 SD.X Axis: k-nearest vs support vector machine Y-Axis: Mean accuracy ± 1 SD.