Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 9
September 2024

Robust Devanagari Character Recognition Using Transfer Learning

Dr. Vina M Lomte[1] , Ms. Aakansha S. Mulik[2]
[1] Associate Professor, RMDSTIC, Pune, [2] Operation Analyst, Arcserve, Hyderabad, India

Abstract
Many applications, such as digital archiving, automatic transcription, and improving accessibility for native speakers, depended on Devanagari Character Recognition (DCR). This study used VGGNet and AlexNet, two pre-trained deep learning models, to examine the effectiveness of transfer learning in DCR. Utilizing the rich feature representations acquired from extensive image datasets, transfer learning significantly improved recognition performance even with a small amount of labeled data. Using a large dataset of handwritten and printed Devanagari letters, we optimized VGGNet and AlexNet to capture a variety of styles and variants. The outcomes of our investigation showed that both models were highly accurate at identifying letters, numbers, and conjunct characters in handwritten Devanagari text. Identifying characters, numbers, and conjunct characters from images required the detection, segmentation, and identification of Devanagari text. Once the text was discovered, it was converted into a machine- or digitally-encoded representation. VGGNet and AlexNet were used in this study as feature extractors inside a transfer learning framework. Our method reduced computing time and produced better results. We conducted experiments over 120 epochs for each model. The outcomes showed that VGGNet, with an accuracy 96.50 % in average period time of 121 seconds, obtained greater accuracy. In contrast, AlexNet had an average epoch time of 114 seconds and an accuracy of 95.97%. These results demonstrated the effectiveness of using VGGNet and AlexNet in conjunction with transfer learning to achieve reliable and effective Devanagari character recognition, implying notable gains in performance for real-world use.
Keywords: Vggnet, Alexnet, Deep Learning, CNN, Transfer Learning. ac tincidunt vitae semper.

1. Introduction

Text recognition in Devanagari is a field that is critical to the digitalization of textual information in many formats, such as texts, sheets, and notes. The process of digitalization makes it simpler to use and share content across geographical boundaries. With its extensive historical literature, Devanagari script stands to gain much from these technological developments, making it possible to preserve priceless manuscripts in digital format. Numerous real-world uses for handwriting character recognition exist, such as the identification of bank checks, postal codes, and bank signatures. Further expanding their accessibility, digitizing Devanagari writings also helps with their comprehension and translation into other languages. Text recognition is the act of detecting, separating, and locating text in photographs and then transforming that information into a machine- or digitally-readable format. A branch of this topic called character recognition deals with computers identifying and deciphering characters from text images and converting the resulting data into a form that can be understood by computational systems. [11] The Devanagari Handwritten Character Dataset (DHCD) is an example of how difficult it is to recognize Devanagari characters, which frequently consist of groups of symbols that resemble each other.Figures from the study—such as those between characters 1 and 2, 3 and 4, and 5 and 6—highlight the difficulties these similarities present. To enhance stroke recognition, researchers have suggested a number of manually created features,[14] However, the construction of a comprehensive feature set is complicated by the complex structure of Devanagari characters. The recognition of joint letters presents additional difficulties due to the presence of header lines and multiple modifiers. Such complexities make handwritten Devanagari characters particularly challenging to recognize, and errors often arise when joint letters are treated as single characters. This underscores the ongoing need for advanced research and development in Devanagari text recognition technology.

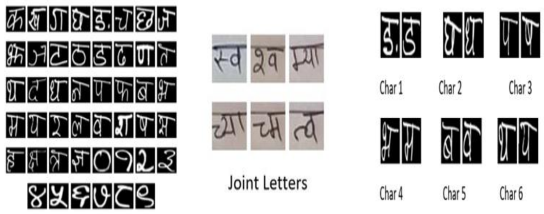


Fig. 01: Devnagari Character, number,joint & Confusing Characters

2. Literature Survey

Researchers have been interested in Abu Sayeed's handwritten Bengali persona for many years. Bengali has no longer produced many notable contributions in the field of handwritten person reputation, despite the fact that other well-known languages, such as English, Chinese, Hindi, Spanish, etc., have made numerous contributions in this area. This is because Bengali characters have similar writing models and complex curvatures. Shikhar Prateek Pandey et.al [2], Character reputation, a sub area of sample reputation, has been a topic of studies for the reason that early twentieth century. Hindi is the not unusual place and maximum famous language within side the international locations which includes India, Nepal etc. People use this language now no longer most effective for communiqué however additionally of their automobiles license plates, documents, signal boards, handwritten notes etc. In latest years, many processes were proposed for Hindi person reputation and numerous packages which include textual content to speech translator, automated registration code reputation etc. are proposed for these. Some computationally highly-priced processes have done acceptable accuracy however for mild computing devices; reputation of handwritten characters continues to be hard task. This paper proposes an method for reputation of handwritten Devanagari person reputation. Shalaka prasad deore et.al [3], A brand-new, publicly accessible dataset of handwritten Devanagari characters is created. 5880 remoted images of certain man or woman classes (12 vowels, 36 consonants, and 10 numbers) are included in the general datasets. Furthermore, a -degree VGG16 deep learning version is conducted in addition to this database in order to comprehend the characters through the application of advanced adaptive gradient techniques. To enhance the standard implementation of the suggested Devanagari Handwritten Character Recognition System (DHCRS), a -degree deep learning approach is employed. On a fresh dataset, the first version achieves 94.84% testing accuracy with a schooling deficiency of 0.18. Furthermore, to get modern overall performance on a very small dataset, the second refined version requires far less training time and fewer trainable parameters. It attains 96.55% testing accuracy with a 0.12 education gap. Kavitha B.R et.al [4] In recent times, Convolutional Neural Networks (CNN) have gained significant prominence in all aspects of laptop vision applications. In order to identify handwritten Tamil characters in offline mode, we have utilized the nation of the artwork CNN in this article. When it comes to automatically extracting capabilities, CNNs differ from the traditional method of Handwritten Tamil Character Recognition (HTCR). We used a handwritten, remotely accessible Tamil man or woman dataset that was developed by HP Labs India. We have developed a CNN version from the ground up by training the offline version with Tamil characters and included accurate reputation effects on all training and testing datasets. This painting is an attempt to establish a standard for offline HTCR deep learning methods. These paintings have produced a schooling accuracy of 95.16% that is some distance higher as compared to the conventional approaches P. Giri Kishore1 et.al [5] this author’s paper highlights how EMNIST database may be positioned to use, to create smooth and artificial snap shots of texts in special handwriting styles. This authors paper offers an in depth evaluate of Intelligent Character Recognition, the techniques the usage of which they could classify man or woman via way of means of detecting and extracting the placement of a man or woman from artificial snap shots. With the assist of the photo processing, the uncooked photo is cleaned, earlier than its miles dispatched for classification that allows you to supply a better chance of a hit popularity of characters. Nagender Aneja et.al [6] a review of amateur models for Deep Convolution Neural Network (DCNN) transfers learning in order to understand handwritten Devanagari alphabets. AlexNet, DenseNet, Vgg, and

Inception ConvNet are used in this study as a robust and efficient feature extractor. For AlexNet, DenseNet 121, DenseNet 201, Vgg 11, Vgg 16, Vgg 19, and Inception V3, we used 15 epochs. The results indicate that while AlexNet works fastest with 2.2 minutes in line with epoch and engages in 98% accuracy, Inception V3 performs better in terms of accuracy, engaging in 99% accuracy with not exceptional place epoch time of 16.three minutes on the same time. According to Asfi Fardous et al. [7], the Bangla alphabet contains anglo composite characters. Bangla OCR is insufficient without improving the recognizer for distant compound characters. While most research has been done on Bangla numerals and easy characters, not much has been done on handwritten Bangla compound characters. For this reason, a convolutional neural network (CNN) based full model has proven to be more effective in understanding handwritten distant Bangla compound letters. By training the suggested model using CMATERdb 3.1.3.3 and comparing the final result over the test dataset with novel, modern methodologies for handwriting Bangla compound person popularity, the average performance of the model has been examined. The network's ultimate end result demonstrated 95.5% accuracy on the test dataset, which is superior to several existing methods. Rajarajan Lalitha and Sridevi T.N. [8] Three ranges are included in the structure: ReLu in the primary ranges, Max pooling layers, and ReLu alone in the 1/3 stage. Their output enters an absolutely linked layer that plays a certain kind of character.

3. Proposed System

Our proposed system leverages the power of deep learning algorithms, specifically VGGNet and AlexNet, to achieve high accuracy in Devanagari character recognition. Through the use of VGGNet's deep architecture, the system is able to better recognize characters by capturing their complex patterns and properties. AlexNet's effectiveness and capacity for handling big datasets improve the system's speed and scalability at the same time. [15] A large variety of Devanagari characters can be recognized with resilience and versatility because to the merging of these two models. We also integrate pre-processing in our method. Pre-processing, division, extraction of features, and categorization. Pre-processing entails a number of essential procedures to improve the consistency and quality of the input data. To start, greyscale images are created in order to simplify computation. Artifacts are eliminated by applying Gaussian blur noise reduction techniques. Normalization then guarantees consistency in stroke thickness and character size.

Fig 02: Architecture of proposed system

known as histogram equalization is applied in computer image processing. The segmentation stage is crucial to Devanagari text recognition since it extracts individual letters or words from the input text. The text graphic is initially binarized in order to extract the text from the background. Line segmentation is used to split text lines, and word segmentation is used to identify words inside each line. Character segmentation then divides each character into discrete parts for identification, ensuring precision in the following stages. [17] Complex script structures and overlapping characters are handled with advanced approaches such as connected component analysis. To aid in precise recognition, unique traits of Devanagari characters are located and extracted during the feature extraction stage. Key features like edges, forms, and textures are captured using techniques like the histogram. This procedure is further improved by deep learning models, particularly VGGNet and AlexNet, which automatically extract hierarchical features from the data. The robust representation of each character that is produced using these extracted features allows for accurate discrimination throughout the classification stage.

Dataset Description: Two dataset have been used for experimentation. Normal Characters-This dataset, which has a total size of 92000, includes mixed categories for consonants (36 classes) and Devanagari numerals (10 classes). 78,200 samples altogether in the train set, with 1700 samples each class for a total of 46 Segmentation prior to processing Sorting by Gray Scale Histogram Sharpening Edge Detection. VGG Alexne t Post-Processing Recognition Database 6 6 10 classes. There are 13,800 samples in the test set overall, 300 samples in each of the 46 classes. Test set contains total 13,800 samples with 300 samples per class for total 46 classes. The Joint Characters dataset –This is created by 52 writers which contains 6 classes with total 42 images. Same 80/20 train-test split ratio is used.

VGGNet Architecture: The VGG architecture, which stands for Visual Geometry Group, is a multilayer deep Convolutional Neural Network (CNN) architecture. [19]The terms "deep" and "vgg-19" denote the number of convolutional layers in VGG-16 and VGG-19, respectively. Thirteen convolutional layers, five max-pooling layers, and three thick layers comprise VGG16. As seen in Figure 7, Conv1 has 64 3X3 kernel filters, Conv2 has 128 3X3 kernel filters, Conv3 has 256 3X3 kernel filters, and Conv4 and Conv5 have 512 3X3 kernel filters. 2X2 is the kernel of every max-pooling layer. To stop negative values from being sent to the following layers, the ReLu activation function is added to the two dense layers and each convolution layer. The last dense output layer uses the Softmax activation function for prediction.

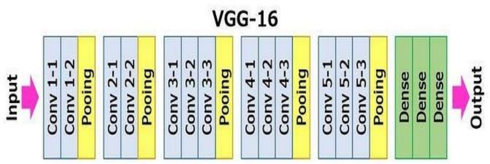


Fig. 03 Vggnet Architecture[12]

AlexNet Architecture: An eight-layer convolutional neural network is called AlexNet.Each of the eight levels has a unique set of parameters that can be acquired. With the exception of the output layer, which combines max pooling and three fully connected layers, the model's five layers all employ Relu activation. Five convolutional layers make up the design; the first, second, and fifth have Max-Pooling layers for appropriate feature extraction. The Max-Pooling layers have overlapping strides of two and a filter size of three by three. Conv1 has 96 3X3 kernel filters, Conv2 has 256 3X3 kernel filters, Conv3 and Conv4 have 384 3X3 kernel filters, and Conv5 has 256 3X3 kernel filters, as shown in Figure.

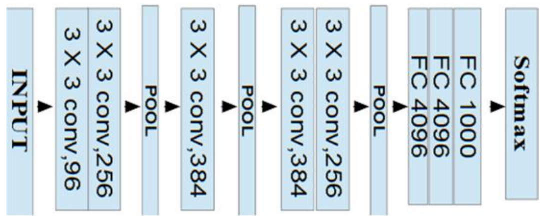


Fig. 04 Alexnet Architecture [13]

Here Input X is a devanagari character and $C1, C2, \ldots Cn$ are the classes.

$x \in R^{m \times n \times c}$ ....(1)

m & n are height and width. C is the channel of the input image shown in equation (1)

(Xi, Yi) is the pair Of the dataset, where the Xi is ith input image $Yi \in \{1,2,..k\}$ is the corresponding class labels.

$h = F_{pre-int}(X;\ \theta_{pre-int})$..........(2)

Equation (2) represents pre-trained model used for feature representation, where $\theta_{pre-int}$ are the pre-trained parameters and $F_{pre-int}$ denotes pre-trained model up to a certain layer.

$z = wh + b$ .... (3)

After feature extraction extracted features needs to be feed to the classifier to find the closed match. Here fully connected layer ρ at the top of the feature extractor, where all neurons are connected with each other

The w and b are the weight & bias added in the new layer for correct recognition shown in equation (3). The logits z will be applied to softmax classifier to obtain class probabilities using equation (4).

$$\breve{y}_{i=softmax(z)}$$

$$softmaz(z) = \frac{e^{2j}}{\sum_{l=1}^{k} e^{2l}} \text{ ....(4)}$$

At this point, the predicted error. is measured using the cross entropy loss to calculate loss.

$L(Xi, Yi; \theta) = -\log \hat{Y}\imath, Yi$ ....(5)

$$\theta \text{ includes both } \theta_{pre-int} \text{ and parameters of } \rho(w \text{ and } b)$$

Equation(5) is used to calculate the loss.

during classification process. To minimize the loss gradient decent Adam optimization gives fine outcomes using equation (6).

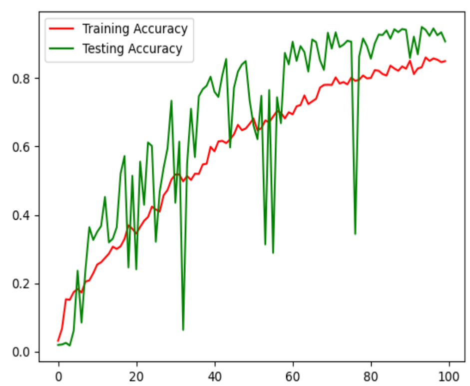
4. Results and Discussion

Proposed system is aimed at recognizing the handwritten Devanagari characters. We have developed system for following two aspects of handwritten Devanagari character recognition using algorithms VggNet and Alexnet. For Vggnet, we obtained testing accuracy of 99.8% and training accuracy of 99.2%; meanwhile, for Alexnet, we obtained testing accuracy of 97.1% and training accuracy of 96.1%.

Table 01: Model Performance

MODEL ACCURACY(%)	VGGNET	ALEXNET
TRAINING ACCURACY	99.20	96.10
TESTING ACCURACY	99.80	97.10
TRAINING LOSS	0.36	0.46
TESTING LOSS	0.39	0.15

Graph 01: Model Performance



from the Table 01. The testing loss and training loss for vggnet and alexnet, respectively, are displayed as 0.39 and 0.46 ms. Utilizing VGGNet architecture, our system achieves an impressive accuracy rate of 98% in recognizing joint letters in Devanagari script. The deep layers of VGGNet effectively capture intricate patterns and structures within the joint letters,[19] ensuring precise recognition. Extensive training with diverse datasets enhances the model's ability to accurately identify and classify joint letter combinations. Through rigorous testing and validation, our system demonstrates exceptional performance, making it a reliable tool for Devanagari text recognition applications.

Table 03 Alexnet joint letters Recognition & Avg. Time

Char	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	Avg. Time (ms)	Accuracy (%)
च्म	0.064	0.008	0.00	0.008	0.00	0.008	0.008	0.00	0.008	0.007	0.011	100
च्या	0.007	0.0017	0.064	0.0017	0.008	0.007	0.064	0.078	0.00	0.00	0.038	100
म्या	0.078	0.007	0.078	0.064	0.008	0.078	0.00	0.078	0.007	0.078	0.047	90
श्व	0.008	0.007	0.008	0.064	0.008	0.00	0.006	0.064	0.007	0.008	0.018	100
स्व	0.007	0.078	0.008	0.078	0.078	0.008	0.078	0.015	0.007	0.008	0.036	100
त्व	0.015	0.008	0.015	0.008	0.008	0.064	0.015	0.008	0.064	0.008	0.021	100
Average Value											**0.028**	**98.33**

Here, X stands for a picture that is not recognized. Using the AlexNet architecture, our system recognizes joint letter combinations in Devanagari letters with an impress

Char	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7	Image 8	Image 9	Image 10	Avg. Time (ms)	Accuracy (%)
च्म	0.015	0.088	X	0.015	0.089	0.096	0.104	X	0.015	0.089	0.063	80
च्या	0.088	0.089	0.097	0.089	X	0.097	0007	0.015	X	0.007	0.061	80
म्या	0.089	0.096	0.095	0.088	0.097	0.140	X	0.096	0.088	0.095	0.094	90
श्व	0.096	0.064	0.007	0.096	X	0.088	0.015	0.064	X	0.075	0.063	80
स्व	0.104	0.0188	0.015	0.097	0.096	0.015	X	0.095	0.104	X	0.076	80
त्व	0.088	0.015	0.097	0.096	0.089	0.015	X	0.015	0.097	0.088	0.061	90
Average Value											0.071	83

ive 83% accuracy rate. Because of its strong architecture and deep learning capabilities, AlexNet can effectively extract features that are essential for joint letter recognition. After a rigorous training and optimization process, our model exhibits notable competence in identifying and categorizing intricate joint letter arrangements. This precision highlights the dependability and efficiency of our method in Devanagari text recognition assignments. We investigate how well VGGNet and AlexNet architectures perform in the crucial task of Devanagari character recognition, which is relevant to a wide range of language remarkable outcomes. Comparing Accuracy: VGGNet recognized Devanagari characters, processing applications. By means of meticulous testing and analysis, we were able to attain including intricate joint letters, with an astounding 98% accuracy rate, while AlexNet only managed 83%.

VGGNet Performance: The deep convolutional layers of the VGGNet architecture shown exceptional efficacy in capturing the complex characteristics and patterns seen in Devanagari scripts.[20] Better recognition performance was the outcome, particularly when it came to picking up on minute details and variances. AlexNet Performance: When it came to identifying Devanagari characters, AlexNet performed admirably even though it had less layers than VGGNet. Its capacity to manage intricate data structures helped it to get an acceptable 83% accuracy rate.

Table 4: System Throughput

Category	VGGNET Recognition Accuracy (%)	ALEXNET Recognition Accuracy (%)
Character	98.99	98
Number	100	99
Joint Character	98	83
Confusing Character	91.33	97.85

Table 05 : Evaluation Parameters

Evaluation Parameter	**Model**	
	VGGNET	**ALEXNET**
Precision	0.96	0.94
Recall	0.95	0.93
F1-Score	0.95	0.94
Sensitivity	0.95	0.92
Accuracy	96.50	95.97
Epoch Time(s)	121	114

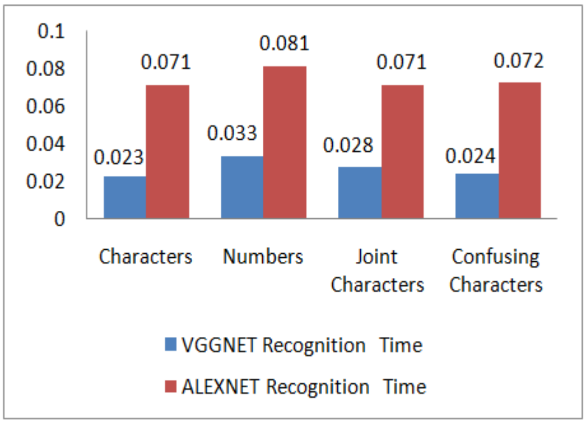


Fig. 05: Recognition Time in ms

Accurate recognition of Devanagari characters is a big difficulty due to their complicated nature, particularly combined letters. This problem was successfully solved by both VGGNet and AlexNet, albeit VGGNet performed somewhat better. Both models' recognition abilities were greatly improved by thorough optimization and extensive training on huge datasets. Although VGGNet performed better in terms of accuracy, its deeper design resulted in greater processing costs. Nevertheless, AlexNet provided a compromise between computing efficiency and accuracy thanks to its somewhat shallower structure.

Table. 06: Confusion Matrix

Predicted /Actual	*C1*	*C2*	*C3*	*C4*	*C5*
C1	*180*	*10*	*5*	*0*	*5*
C2	*15*	*150*	*10*	*5*	*10*
C3	*5*	*20*	*160*	*10*	*5*
C4	*0*	*5*	*15*	*170*	*10*
C5	*10*	*15*	*5*	*15*	*145*

5. Conclusion and Future Directions

Our research shows that deep learning models—VGGNet and AlexNet in particular—are useful at recognizing Devanagari characters. While VGGNet is the most accurate, AlexNet provides a more computationally efficient option without sacrificing performance noticeably. These results open up new avenues for research in this area and further the development of Devanagari text recognition technology. To attain even greater accuracy rates, future research could concentrate on ensemble methods that incorporate the advantages of both VGGNet and AlexNet. Furthermore, looking into the application of more sophisticated architectures like attention-based models or ResNet may produce better outcomes. The Devanagari character recognition capabilities of both VGGNet and AlexNet are strong, opening up a range of practical uses, such as document digitization, Natural language processing (NLP) and optical character recognition (OCR) are available for languages written in the Devanagari script.

References

[1] Abu Sayeed 1, Jungpil Shin 2, Md. Al Mehedi Hasan 2, Azmain Yakin Srizon 1 and Md. Mehedi Hasan, "BengaliNet: A Low-Cost Novel Convolutional Neural Network for Bengali Handwritten Characters Recognition," IEEE 2021.

[2] Shikhar Prateek Pandey, "HANDWRITTEN DEVANAGARI DIGITS RECOGNITION USING DEEP LEARNING," International Journal of Education, Modern Management, Applied Science & Social Science

[3] SHALAKA PRASAD DEORE1,2, * and ALBERT PRAVIN, "Devanagari Handwritten Character Recognition using fine-tuned Deep Convolutional Neural Network on trivial dataset," Sådhanå (2020)45:243 © Indian Academy of Sciences

[4] B.R. Kavitha, C. Srimathi, Benchmarking on offline Handwritten Tamil Character Recognition using convolutional neural networks, Journal of King Saud University - Computer and Information Sciences, 2019,

[5] P. Giri Kishore1, Athira Ajayakumar2 , N Nitin3 and Arun Natarajan "Intelligent Character Recognition- Character detection using Neural Networks" International Journal of Engineering Trends and Technology (IJETT-Scopus Indexed) – Special Issues - ICT 2020

[6] Nagender Aneja and Sandhya Aneja, Universiti Brunei Darussalam, "Transfer Learning using CNN for Handwritten Devanagari Character Recognition," Accepted for publication in IEEE International Conference on Advances in Information Technology (ICAIT), ICAIT - 2019, Please refer IEEE Explore for final version

[7] Asfi Fardous,Shyla Afroge "Handwritten Isolated Bangla Compound Character Recognition," 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February 2019

[8] Sridevi T.N and Lalitha Rangarajan "DEEP CONVOLUTIONAL NEURAL NETWORKS FOR DEGRADED PRINTED KANNADA CHARACTER RECOGNITION" DOI : 10.21817/indjcse/2021/v12i3/211203187 Vol. 12 No. 3 May-Jun 2021

[9] Mayur Bhargab Bora, Dinthisrang Daimary, Khwairakpam Amitab¬, Debdatta Kandar "Handwritten Character Recognition from Images using CNN-ECOC" International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

[10] JAMSHED MEMON 1 , MAIRA SAMI 2 , RIZWAN AHMED KHAN 3 , AND MUEEN UDDIN 4 "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)".

[11] To Design and Develop Efficient Method for Devanagari Text Recognition using GoogLeNet Inception. Siddhi N. Pawar1 Prof. Vina M. Lomte2 Ranjit T. Makawne3 Riddhi N. Pawar4 Ramprabhu S. Khakare

[12] VGG16 - Convolutional Network for Classification and Detection neurohive.io

[13] Layout diagram of Alexnet Architecture | Download Scientific Diagram (researchgate.net)

[14] Dr. Dharmpal D. Doye, etal, JOURNAL OF ALGEBRAIC STATISTICS Volume 13, No. 3, 2022, p.2190- 2198, ISSN: 1309-3452

[15]V M Lomte, etal , Computer Integrated Manufacturing Systems, Vol.29, ISSN-1006-5911, DOI: 10.24297/j.cims.2023.1.7

[16] Mahesh S Shewale, etal , Novel machine health monitoring system, Proceedings of the 2nd International Conference on Data Engineering and Communication Technology: ICDECT 2017, Pg-461-468.

[17] Ronit Keshave, etal Advancing Document Accessibility for the Visually Impaired: Comparative Analysis of Integrated OCR, Translation, and TTS Technologies, Journal of

Technology, ISSN: 10123407, 5, 2024, PAGE NO:429-435,

[18] Ishita Satija1 ,etal, Text-To-Speech Synthesis Using Transfer Learning, International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) Volume 4, Issue 1, April 2021, DOI: 10.48175/IJARSCT-956

[19] Sneha Punde, etal, SURVEY ON OBJECT DETECTION TECHNIQUES BASED ON GPU FOR MARATHI CHARACTER RECOGNITION, International Journal of Advances in Engineering Research, Vol. No. 21, Issue No. VI, June e-ISSN: 2231-5152, p-ISSN: 2454-1796.

[20] Harsh Shah, REVIEW ON SEGMENTATION AND RECOGNITION METHODOLOGIES FOR OCR SYSTEM OF DEVANAGARI SCRIPT, 2021, Vol. No. 21, Issue No. III, March, e-ISSN: 2231-5152, p-ISSN: 2454-1796.