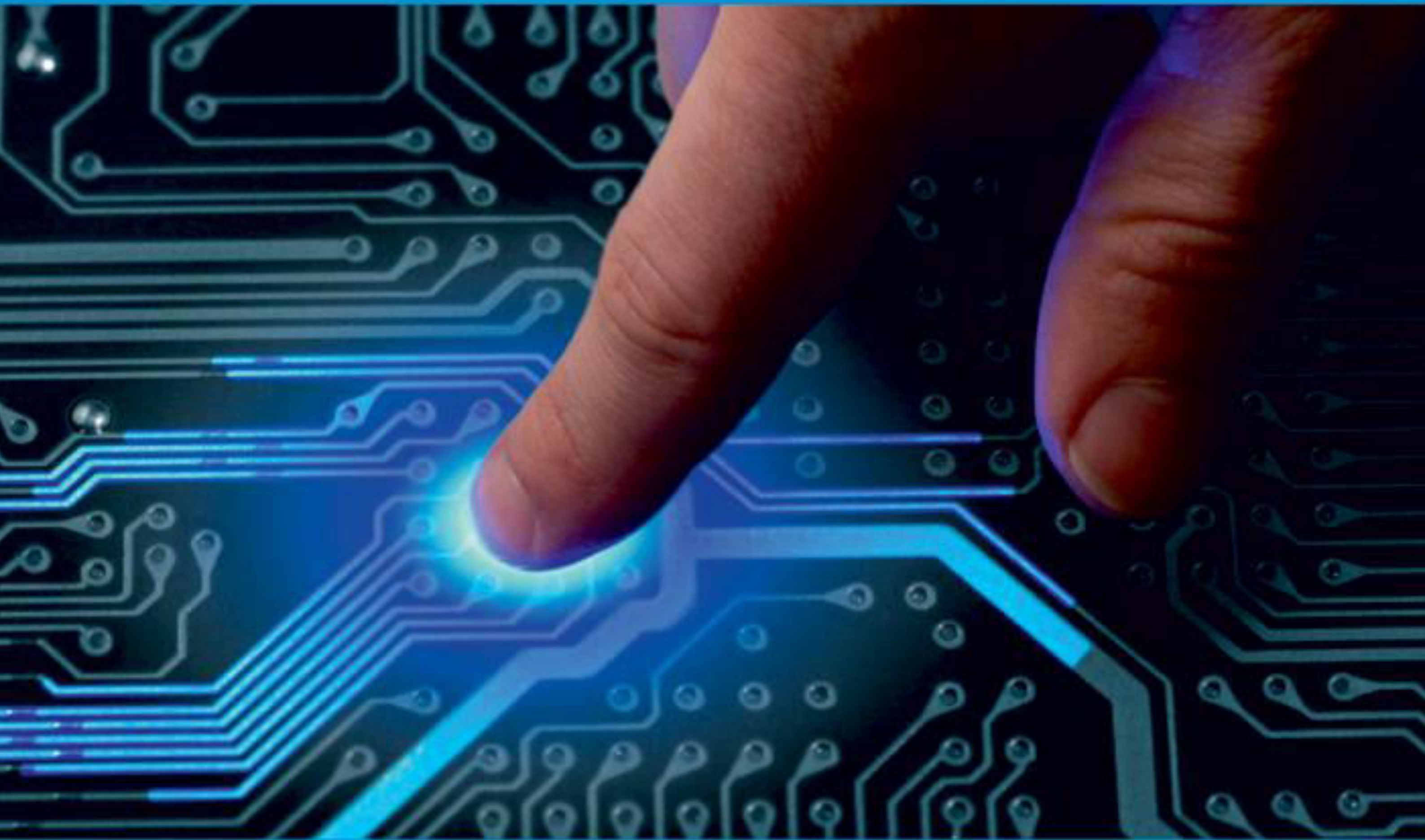
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205270 |

A CNN based Braille Conversion System

Dr. Archana P. Kale, Malan Shaikh, Pritee Shinde, Vaishnavi Korgaonkar, Akanksha Chopade

Professor, Department of Computer., MES Wadia College of Engineering Pune, India

UG Student, Department of Computer MES Wadia College of Engineering Pune, India

ABSTRACT: Deep learning models, especially Convolutional Neural Networks (CNN), can identify complex patterns in text and images, aiding in precise predictions and insights. Visually impaired individuals face challenges accessing literature and information. A Braille conversion system can significantly enhance their accessibility, promoting inclusion in education, employment, and daily life. Current Braille-to-text systems translate text into Braille but cannot extract text from images or large files. To address this, we propose the PCA-based Embedded Residual Network (PERN) Intelligent System, which uses ResNet for Braille character recognition. PERN aims to overcome the limitations of the Braille Enhance Network (BENet), offering improved accuracy and efficiency. It also supports image alignment and enhancement. Experimental results show PERN achieving a classification accuracy of 93.26% on the English Braille Dataset test set, outperforming BENet in computing efficiency and accuracy metrics. Statistical and hypothetical tests further validate its performance. PERN can handle larger datasets and more complex Braille languages, demonstrating its robustness and applicability.

KEYWORDS: Convolutional Neural Network (CNN), Dimensionality Reduction, Principal Component Analysis (PCA).

I. INTRODUCTION

A Deep Learning based recognition approach is proposed to convert Braille images into text. The purpose of the Braille Character Recognition System is to outline the functional and non-functional requirements for the Braille to text conversion using CNN software, detailing its features, constraints, and performance criteria. It also establishes a requirement baseline for development of the system [1]. The Braille Character Recognition System outlines both functional and non-functional requirements, detailing features, constraints, and performance criteria essential for the Braille-to-text conversion using Convolutional Neural Network (CNN) software. CNNs facilitate Braille image recognition by accurately identifying Braille characters within images. The system employs dimensionality reduction to simplify high-dimensional data, enhancing analysis efficiency [2]. The primary objective is to aid visually impaired users by transforming Braille images into readable text, thereby promoting independent access to information. This initiative is motivated by the challenges faced by visually impaired individuals, particularly in education and information access. The system utilizes a hybrid approach with CNN [3] for image recognition, ensuring precise Braille character identification. It encompasses a systematic process that includes image acquisition and preprocessing, Braille cell recognition, and Braille character mapping. The use of dimensionality reduction [4] techniques further improve the system's efficiency by focusing on key data components. This innovative Braille Character Recognition System signifies a significant advancement in assistive technologies, offering a reliable solution for Braille-to-text conversion. It aims for continuous improvement and future enhancements to better support the visually impaired community, contributing to the evolution of assistive technologies for enhanced accessibility and inclusivity [5].

II. RELATED WORK

The study involves collecting a dataset of 1560 images of English Braille Grade 1 character input by visually impaired individuals. Deep learning techniques are applied to assess the proposed method's performance. The results indicate that the GoogLeNet model, followed by the Sequential model, SVM, DT, KNN, and NB [6], achieved the highest accuracy in predicting user input. An automatic Braille image character recognition algorithm using CNN [7] is employed, incorporating Inverted Residual Block (IRB) modules for character recognition. The datasets used include English and Chinese Braille, achieving accuracies of 95.2% and 98.3% respectively. A system for recognizing Braille characters in six-dot patterns from scanned documents involves preprocessing, segmentation of Braille cells, centroid computation using Euclidean distance, and a generalized lookup table for text recognition [8]. This Braille Character Recognition (BCR) system demonstrates the feasibility of a cost-effective solution. Additionally, a parallel fully convolutional neural network (FCN) [9] for semantic segmentation integrates holistically-nested edge detection (HED) to capture image edge information, enhancing segmentation performance by combining coarse segmentation with edge

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205270 |

details, improving spatial consistency, and reducing information loss [10]. Experimental results on benchmark datasets such as PASCAL VOC 2012, PASCAL-Context, and Cityscapes [11] show this method's superiority over existing techniques with minimal additional parameters. An automatic image alignment technique using Principal Component Analysis (PCA) addresses the 180° rotation issue in existing methods. This algorithm efficiently aligns datasets like handwritten digits, rotated fingerprints, and brain MRI images [12], providing standard object orientations. It has practical applications in optical character recognition, fingerprint matching, and medical image registration, demonstrating robustness and effectiveness in real-world experiments. An improved principal component analysis (PCA) is also presented for face feature representation [13].

III. PROPOSED ALGORITHM

Character recognition in Braille involves four main stages. Initially, text input is converted to Braille and Braille images are transformed into text. The second step employs image preparation techniques, including alignment and augmentation. The third step uses a Convolutional Neural Network (CNN) for character identification. Figure 1 illustrates the Braille image processing workflow, which enhances recognition accuracy and visual quality through several interconnected stages.

First, Principal Component Analysis (PCA) aligns images for accurate orientation. Next, image enhancement techniques improve contrast, brightness, and sharpness. Then, a CNN recognizes Braille patterns, moving to the final step if successful. The framework comprises four components:
1. Input (Dataset)
2. Image Alignment
3. Image Enhancement
4. Convolutional Neural Network.

A. *Input (Dataset):*
The Braille character dataset consists of digital images representing raised-dot symbols used by visually impaired individuals. These images enable the CNN to learn Braille characters, facilitating model development to assist blind individuals in reading and writing. The dataset includes various Braille characters and their corresponding digital representations for effective CNN training. It contains 1560 grayscale images of English Braille characters, each 28 x 28 pixels.

B. Pre-processing:-
1. Image Alignment
PCA is employed to correct geometric distortions in Braille images captured by scanners or cameras, which may arise from improper positioning or rotation during image acquisition. These misalignments can significantly impair character recognition accuracy. PCA, a dimensionality reduction technique, identifies the primary directions of variance in image data. By computing the mean and covariance matrix of pixel coordinates, PCA determines the optimal rotation angle for correct image alignment. Once misaligned images are identified, PCA facilitates their rotation using bilinear interpolation, ensuring their orientation matches the standard layout. This process enhances overall consistency and accuracy in Braille character recognition by addressing geometric irregularities in the input images.

2. Image Enhancement and Noise Removal
Enhancing the quality of Braille images is crucial for optimizing CNN model recognition performance. The preprocessing pipeline employs several techniques to reduce noise and improve visual clarity. Wiener filtering is initially applied to reduce noise and enhance dot patterns while preserving image sharpness. Hysteresis thresholding then segments the image into binary regions based on predefined intensity thresholds. Morphological operations, including erosion and dilation, refine the binary image by removing extraneous elements and emphasizing core features. These preprocessing steps collectively improve image quality, facilitating robust and accurate character recognition in subsequent CNN processing stages.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205270 |

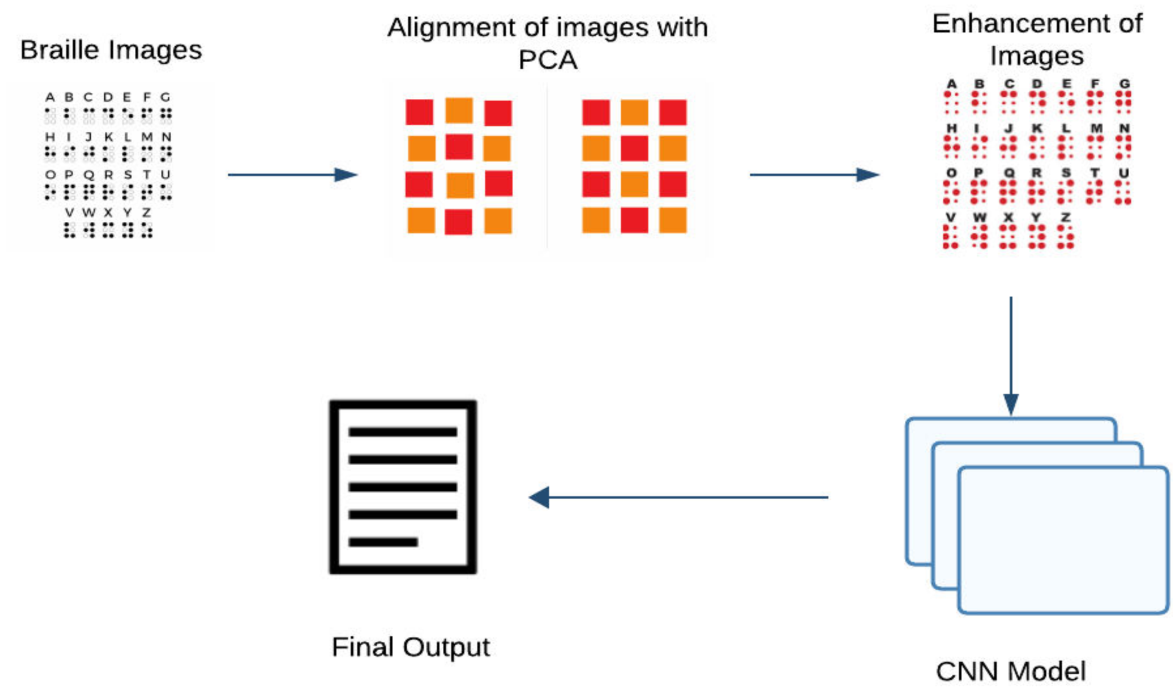


Figure 1. Proposed Architecture

C. Convolutional Neural Network: -
The CNN architecture processes 28 x 28 grayscale Braille images, categorizing them into 26 classes. It includes several key layer:
1. Convolutional Layer
This layer extracts features from input images using filters called kernels. These 3x3 filters slide over the input image, computing dot products to produce feature maps. With 64 filters applied, the resulting volume is 28 x 28 x 64, transforming images into numerical representations for pattern analysis.

2. Activation Function
Activation functions like Rectified Linear Unit (ReLU) introduce non-linearities essential for discovering complex patterns within data. ReLU outputs the input value for positive inputs and zero for negatives, fostering efficient learning by focusing on relevant features. The SoftMax activation function, used in the output layer, converts network outputs into probability distributions over 26 classes, facilitating multiclass classification.

3. Pooling Layer
Pooling layers, typically following convolutional layers, reduce computational complexity and overfitting while emphasizing prominent features. They aid in providing translation invariance and improving model performance by down sampling feature maps. For instance, a 28 x 28 x 64 volume can be reduced to 14 x 14 x 64 dimensions.

The CNN also includes dense layers for learning intricate feature dependencies and an output layer for classification. This meticulously designed architecture excels at extracting meaningful features from input images, enabling effective interpretation of complex visual information.

IV. SIMULATION RESULTS

The image preprocessing algorithm has been implemented using Jupyter Notebook, while the CNN is developed with TensorFlow and Keras [14] on the same platform. During the model evaluation phase in machine learning, it is essential to utilize various metrics to assess the model's performance, especially in classification problems. Key evaluation metrics include accuracy, precision, recall, and F1-score [15]. Accuracy, the most basic evaluation metric in

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205270 |

classification, is defined as the proportion of correctly predicted results out of the total sample. Precision indicates the percentage of correctly classified positive instances within the test set. Recall [16] measures the proportion of actual positive cases that are correctly identified as positive. The F1-score, a balanced measure of a model's overall performance, is calculated as the weighted harmonic mean of recall and precision. The Geometric Mean (G-Mean) [17] evaluates the balance between classification performance on both the majority and minority classes. A low G-Mean indicates poor performance in classifying positive cases, even if negative cases are correctly classified. This metric is crucial for avoiding overfitting to the negative class and underfitting [18] to the positive class.

The following formulas are used to calculate these metrics:

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1\ score = 2 \times \frac{(Precision \times Recall)}{(Precision+Recall)}$$

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FN+FP)}$$

Here, TN, TP, FN, and FP denote the counts of true negatives, true positives, false negatives, and false positives respectively. This framework provides a comprehensive evaluation of the classification model's performance. The TABLE I summarizes the outcomes of a series of experiments designed to evaluate the effectiveness of a proposed method within the realm of Convolutional Neural Network (CNN).

Sr No	Technique	Precision	Sensitivity	F-score	Accuracy
1	CNN	87.96	86.85	86.85	86.85
2	CNN+X	91.32	90.38	90.47	90.38
3	CNN+Y	91.69	90.70	90.38	90.70
4	CNN+X+Y	92.14	91.35	91.23	91.35

Table I: Analysis Of Impact In Performance Of Image Preprocessing (X: Alignment, Y: Enhancement), Results Are Evaluated On English Braille Dataset

In the initial experiment, the proposed method was applied to original images using a base CNN architecture, resulting in a precision of 87.96%, sensitivity of 86.85%, and an F-score of 86.85%. Subsequent experiments introduced enhancements to investigate their impact on performance metrics. In the second experiment, image alignment (X) was incorporated, leading to significant improvements with precision, sensitivity, and F-score reaching 91.32%, 90.38%, and 90.47%, respectively. The third experiment added an enhancement component (Y), further boosting performance, achieving a precision of 91.69%, sensitivity of 90.70%, and an F-score of 90.38%. In the fourth experiment, the combined impact of additional CNN types, represented by variables D, R, and W, was explored, resulting in a precision of 90.16%, sensitivity of 88.14%, and an F-score of 88.27%. The fifth experiment combined multiple enhancements, including alignment (X) and enhancement (Y), demonstrating the potential for optimizing CNN applications in image processing through alignment, enhancement, and the integration of various CNN types. These findings provide valuable insights for improving CNN performance in image processing tasks.

V. CONCLUSION AND FUTURE WORK

The project represents a pioneering effort to empower visually impaired individuals by providing an advanced and accurate system for Braille character recognition. Through meticulous image preprocessing and alignment, along with feature extraction using CNN, the project tackles the complexities of diverse Braille dot patterns and image qualities. Its success is not only evident in its commendable accuracy but also in its dedication to user-centric design. However, the project envisions future enhancements such as multilingual support, real-time mobile recognition, and continuous learning mechanisms to further enhance the system's capabilities. It concludes with a commitment to ongoing development and collaboration, poised to make a lasting impact on the accessibility and usability of Braille content for the visually impaired community. The proposed algorithm achieves an accuracy of 92.26%. Future research may focus on investigating the influence of Braille

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205270 |

image quality on system performance, considering variations from diverse image sources. Additionally, exploring the feasibility of adapting the system to recognize Grade Braille, a more complex and challenging language format, could be beneficial. This paper aims to contribute to the advancement of Braille character recognition technology, striving for higher accuracy and broader applicability.

REFERENCES

1. T. Kausar, S. Manzoor, A. Kausar, Y. Lu, M. Wasif, and M. A. Ashraf, “Deep learning strategy for braille character recognition,” IEEE Access, vol. 9, pp. 169357–169371, 2021.
2. S. Shokat, R. Riaz, S. S. Rizvi, A. M. Abbasi, A. A. Abbasi, and S. J. Kwon, “Deep learning scheme for character prediction with position-free touch screen-based braille input method,” Human-centric Computing and Information Sciences, vol. 10, no. 1, pp. 1–24, 2020.
3. K. Sugirtha and M. Dhanalakshmi, “Transliteration of braille code into text in english language,” in Computational Signal Processing and Analysis: Select Proceedings of ICNETS2, Volume I, pp. 373–381, Springer, 2018.
4. J. Ji, X. Lu, M. Luo, M. Yin, Q. Miao, and X. Liu, “Parallel fully convolutional network for semantic segmentation,” IEEE Access, vol. 9, pp. 673–682, 2020.
5. H. Z. U. Rehman and S. Lee, “Automatic image alignment using principal component analysis,” IEEE Access, vol. 6, pp. 72063–72072, 2018.
6. Y. Zhu, C. Zhu, and X. Li, “Improved principal component analysis and linear regression classification for face recognition,” Signal Processing, vol. 145, pp. 175–182, 2018.
7. H. M. PANDEY, C. Kang, X. Yu, S.-H. Wang, D. S. Guttery, Y. Tian, and Y.-D. Zhang, “A heuristic neural network structure relying on fuzzy logic for images scoring,” IEEE Transactions on Fuzzy Systems, vol. 29, no. 1, pp. 34–45, 2021.
8. B. Kanawade, R. Wahul, A. Kale, J. Pansare, P. Patil, M. Tungar, N. Verma, and A. Tarte, “Scaled densenet: An efficient deep learning architecture for skin lesion identification,” Revue d’Intelligence Artificielle, vol. 37, pp. 977–983, 08 2023.
9. Y. I. Daradkeh, I. Tvoroshenko, V. Gorokhovatskyi, L. A. Latiff, and N. Ahmad, “Development of effective methods for structural image recognition using the principles of data granulation and apparatus of fuzzy logic,” IEEE Access, vol. 9, pp. 13417–13428, 2021.
10. W. Ren, O. Ma, H. Ji, and X. Liu, “Human posture recognition using a hybrid of fuzzy logic and machine le Technologies, vol. 13, no. 1, pp. 14-25, 1 Jan.-March 2020, doi: 10.1109/TLT.2018.2889100.
11. J. Serrano-Guerrero, F. P. Romero, and J. A. Olivas, “Fuzzy logic applied to opinion mining: a review,” Knowledge-Based Systems, vol. 222, p. 107018, 2021.
12. S. Dai, Y. Wang, D. Mei, K. Tao, and S. Wang, “A portable one hand braille input system enabled by dual-interlocked tactile sensors for communication,” IEEE Transactions on Electron Devices, 2023.
13. A. P. Kale, R. M. Wahul, A. D. Patange, R. Soman, and W. Ostachowicz, “Development of deep belief network for tool faults recognition,” Sensors, vol. 23, no. 4, 2023.
14. M. M. H. Saikot and K. R. I. Sanim, “Refreshable braille display with adjustable cell size for learners with different tactile sensitivity,” IEEE Transactions on Haptics, vol. 15, no. 3, pp. 582–591, 2022.
15. Y. Yuan, C. Wang, and Z. Jiang, “Proxy-based deep learning framework for spectral–spatial hyperspectral image classification: Efficient and robust,” IEEE transactions on geoscience and remote sensing, vol. 60, pp. 1–15, 2021.
16. P. Wang, E. Fan, and P. Wang, “Comparative analysis of image classification algorithms based on traditional machine learning and deep learning,” Pattern Recognition Letters, vol. 141, pp. 61–67, 2021.
17. F. Jing, H. Shaohai, and M. Xiaole, “Sar image de-noising via grouping based pca and guided filter,” Journal of Systems Engineering and Electronics, vol. 32, no. 1, pp. 81–91, 2021.
18. Z. Zhou, Y. Yang, and H. Liu, “A braille reading system based on electrotactile display with flexible electrode array,” IEEE/CAA Journal of Automatica Sinica, vol. 9, no. 4, pp. 735–737, 2022.
19. C. Park, S. W. Cho, N. R. Baek, J. Choi, and K. R. Park, “Deep feature based three-stage detection of banknotes and coins for assisting visually impaired people,” IEEE Access, vol. 8, pp. 184598–184613, 2020.
20. T.-Z. Xiang, G.-S. Xia, X. Bai, and L. Zhang, “Image stitching by line-guided local warping with global similarity constraint,” Pattern recognition, vol. 83, pp. 481–497, 2018.
21. G. Zhang, L. Pan, J. Chen, Y. Gong, and F. Liu, “Joint alignment of image faces,” IEEE Access, vol. 8, pp. 114884–114891, 2020.
22. Y. Xu, Y. Bin, J. Wei, Y. Yang, G. Wang, and H. T. Shen, “Multimodal transformer with global-local alignment for composed query image retrieval,” IEEE Transactions on Multimedia, vol. 25, pp. 8346– 8357, 2023.
23. Q. Zheng, Y. Wang, and P. A. Heng, “Online subspace learning from gradient orientations for robust image alignment,” IEEE Transactions on Image Processing, vol. 28, no. 7, pp. 3383–3394, 2019.
24. S. Sharif, R. A. Naqvi, M. Biswas, and W.-K. Loh, “Deep perceptual enhancement for medical image analysis,” IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 10, pp. 4826–4836, 2022.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1205270 |

25. T. Matsui and M. Ikehara, “Low-light image enhancement using a simple network structure,” IEEE Access, vol. 11, pp. 65507–65516, 2023.Learning Technologies, vol. 13, no. 1, pp. 14-25, 1 Jan.-March 2020, doi: 10.1109/TLT.2018.2889100.