Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 3 (2023)

A Novel Multi-Input Hybrid Deep Convolution Neural Network Methodology for Offline Writer Recognition

Naresh Purohit, Subhash Panwar
Department of Computer Science & Engineering, Engineering College Bikaner, Bikaner Technical University, Bikaner, India

ABSTRACT

Identifying a writer from a small sample of handwriting is a challenging task. In addition, it's a key domain of research for the field of forensic investigation of documents. A novel multi-input hybrid convolution neural network (CNN) is developed to address the offline writer recognition problem involving monolingual and bilingual handwritten scripts. Model generates global features for classification by combining local CNN features with traditional handcrafted features. HOG (Histogram of Oriented Gradients), a standard hand-crafted feature descriptor, is employed. Two distinct CNN paths are utilized to extract distinct feature maps. English script benchmark dataset CVL and an in-house bilingual dataset is utilized in experiments. The bilingual dataset consists mixture of handwritten English and Hindi text. Model achieved a 98.23% accuracy rate with CVL and a 96.93% accuracy rate with an in-house bilingual dataset.

Keywords: Convolution Neural Network, Global Features, Handcrafted Features, Histogram Oriented Gradients, Pre-Processing, Writer Recognition.

1. Introduction

Handwriting is still a key means of expressing an individual's uniqueness and plays a crucial part in bio-metric identification. The Writer recognition problem in document image analysis and recognition is a difficult task due to the wide variety of human writing styles. These systems aid in the authentication and validation processes, such as signature identification and user verification in banks, forensic document analysis, and criminal justice system personal identity [1]. It is the process of identifying the author of the requested image file. Based on the technique of composition, it is divided into two categories: offline and online. In online writer recognition, angle or pressure and writing speed are examined, whereas offline writer recognition focuses on words, characters, lines, and paragraphs [2].

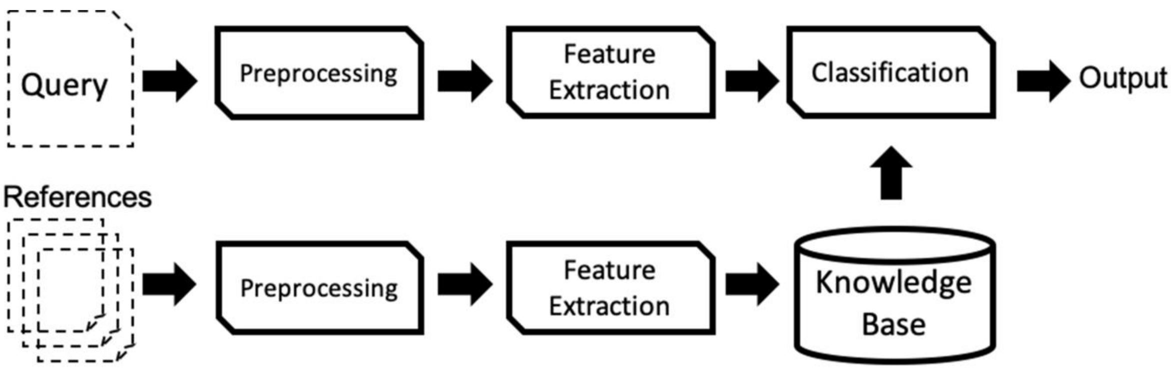


Figure 1: Basic Pipeline of Writer Recognition System

As depicted in figure-1, the methodology consisted of pre-processing, feature extraction, and recognition (classification) standard phases. The pre-processing module is used to clear up the noise from the handwritten input. The following operations may be included in the cleaning process: segmenting the handwritten picture into small image patches; normalizing the size of image patches; and performing the proper morphological operations for feature representation [3]. After that, the knowledge base stores the features that were derived from the references. The challenged document is subject to the same process. As shown in figure-1, trained classifiers assign the unknown query pattern to one of the known patterns throughout the classification process while taking the knowledge base into account [4].

This work proposes a novel multi-input and hybrid deep convolution neural network (CNN) model that uses local features to generate feature maps from original and skeleton [5] image patches. Also retrieved for the same image data was a manually created feature descriptor called a histogram of oriented gradients (HOG) [6]. Then local CNN features and HOG characteristics are integrated to create global features to classify the author of the handwritten image.

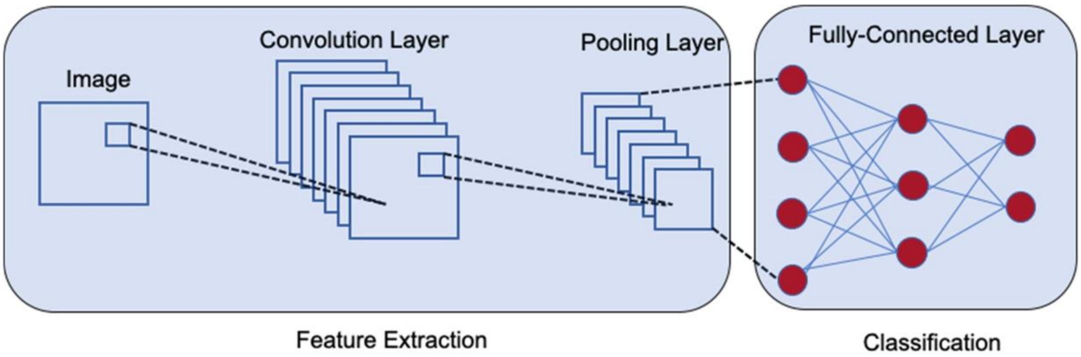


Figure 2: Basic Convolution Neural Network (CNN) Architecture

The foundation of the proposed model originates from the design and development of a Convolution Neural Network (CNN). Widely employed for feature recognition and classification in images, CNN is a deep learning approach in the field of computer vision. The underlying CNN architecture that was influenced by LeNet-5[7] consists of two main parts [8]. Figure-2 shows that the first block is for gathering features and the second block is for classifying. After getting an image as input, the Convolution Layer (CL) applies numerous filters to create different feature maps. The Pooling Layer (PL) down-samples the feature maps produced by the CL while keeping the most vital data. Repeating CL and PL will yield significant traits. All feature vectors are flattened after the final pooling layer in order to transmit 1D vector data to the fully connected layer for classification [9-10]. The last stage entails applying a probability loss function to forecast the writer class of a test query document. This suggested approach addresses the challenge of writer recognition using handwritten image datasets. The following are our contributions:

• An in-house bilingual dataset is prepared. The dataset consists of handwritten text with English and Hindi mixed script.

• The techniques for data preparation which includes image filtration, word segmentation and image skeleton are applied to the dataset.

• Multi-Input convolution neural network (CNN) is designed to process both original and skeleton image patches.

• To create a hybrid model, HOG (Histogram of Oriented Gradients), a classic handcrafted feature, are calculated and combined with the local features of CNN model.

Consequently, the remainder of the paper is structured as follows: Section two discusses related work. The approach is then described in section three, which includes the proposed methodology of data pre-processing and a multi-input hybrid deep convolution neural network (CNN) model. The outcomes of the investigations are analyzed in part four, and the conclusion and next measures are summarized in section five.

2. RELATED WORK

Several common offline writer recognition approaches are reviewed in this section. Prior to the advent of deep learning-based neural networks, classic handcrafted feature-based techniques were employed. Typically, texture and shape properties of picture data are taken into account in handcrafted feature-based approaches.

Six distinct probability distribution functions (PDFs) were utilized by Bulacu et al. [11] to identify writers in the investigated manuscripts. These PDFs were taken from handwritten images with texture and allograph level. The high precision was generated by combining two texture (Direction Co-Occurrence PDFs & Run-Length PDFs) and one allograph (Grapheme Emission PDF) features. Following this, Rajiv Jain et al. [12] presented a new technique in 2011 to enhance performance by utilizing K-adjacent segment (KAS) characteristics. They also employed Euclidean distance for writer categorization and applied exemplar clustering to the KAS (K-adjacent segment) characteristics, and on the IAM[13] dataset, they achieved an accuracy rate of 89.6%. Additionally, the accuracy was increased from 89.6% to 94.1% in 2014 by the same author by combining the KAS features with SURF (Speeded Up Robust Features) and contour gradient descriptors [14]. Xiangqian et al. [15] employed a LoG (Laplacian of Gaussian filter) with Scale-Invariant Feature Transform(SIFT) descriptors to extract words regions from documents using the HIT-MW Chinese dataset, and they were able to do so with an accuracy rate of 95.4%.

Deep learning-based algorithms typically outperform manually created feature-based approaches in terms of recognition performance. Large amounts of data with several features are the cause. A deep learning method called Convolution Neural Network (CNN) can recognize and categories characteristics in visual data. Figure-2 depicts the basic architecture of CNN. Recently, various CNN deep learning models paired with VLAD (Vector of Locally Aggregated Descriptor) encodings, SVM (Support Vector Machine), and other procedures applied to extracted features achieved good writer recognition accuracy rates. Xing et al. [16] used a LeNet-5 like CNN architecture with five convolution layers (CL) and three max pooling layers (MP) to extract feature maps from the IAM dataset. A further approach suggested by Christlein et al. [17] is the clustering of SIFT (Scale-invariant feature transform) key-points extracted from image data, which is subsequently used as the image target in their CNN Model. One illustration of a deep learning model that can work with unlabeled data comes from Shiming Chen et al. [18], who suggested a model using a weighted uniform label distribution method called WLSR(Weighted label smoothing regularization) for unlabeled data. They used ResNet-50[19], a CNN variation that yields 99.2% accuracy, on the CVL [20] dataset.

The method for recognizing hollow Hindi characters that is invariant to scale, rotation, and distortion was proposed by Kumar et al. in [21]. With vertical and horizontal projection, the feature extraction methods SURF (Speeded-Up Robust Features), SIFT (Scale-invariant feature transform), and ORB (Oriented FAST and Rotated BRIEF) are applied. The classifiers used are k-NN (k-Nearest Neighbor), SVM (Support Vector Machine), and Random Forest. They used the Random Forest classifier and got an accuracy of 91.10%. A Boundary line detection method with RLSA (Run Length Smearing Algorithm) was proposed by Jindal et al. in [22] for the segmentation of text and graphics from newspapers in Gurmukhi with a 96.15% accuracy rate.

Kaur et al. [23] presented a study evaluating the performance of various features and classifiers for Gurumukhi newspaper text recognition. The advantages and disadvantages of multiple feature extraction techniques based on a number of classifications are discussed. In conclusion, multi-layer perceptron produced the highest recognition accuracy, i.e. 96.5%, using diagonal features and a partitioning strategy of 70%-30%.

Evident from the preceding discussion are the numerous methods proposed in the literature to solve the problem of offline writer recognition. Table-1 gives overview of comparison of other methodologies with proposed methodology. The majority of methods can only classify the author of monolingual handwritten documents. In this paper, we propose a novel hybrid deep learning approach for handling both monolingual (English) and bilingual (English + Hindi) handwritten documents.

Table 1: Comparison of other methodologies with proposed methodology

Author	Feature Extraction Methodology	Classifier
Bulacu et al. [11]	Probability Distribution Function: Texture level and the Character-shape (allograph) level features	Nearest Neighbor
Jain et al. [12]	K-adjacent segment KAS: K neighboring line segments features	Proposed Distance Function
Jain et al. [14]	KAS with Scale Invariant Feature Transform (SIFT) and Contour Gradient Descriptors	Fisher Vector + Cosine Distance
Xing et al. [16]	Alex-Net based CNN (Convolution Neural Network)	SoftMax
Christlein et al. [17]	CNN with Scale Invariant Feature Transform (SIFT) features	VLAD (Vector of Locally Aggregated Descriptors) Encoding, E-SVM (Support Vector Machine)
Chen et al. [18]	ResNet-50 : 50-layer Convolutional Neural Network(CNN) with stacking residual blocks	VLAD (Vector of Locally Aggregated Descriptors) Encoding + Nearest Neighbor
B. Kumar et al. [24]	Alex-Net based CNN (Convolution Neural Network)	SoftMax
Adak et al. [25]	Squeeze-Net CNN (Convolution Neural Network): CNN model consists two module squeeze and expand layer	SoftMax
Sheng et al. [26]	Deep Adaptive CNN: Model consists two paths for auxiliary and main task with local feature sharing	SoftMax
Sheng et al. [27]	Frag-Net CNN: Model consists two paths, first feature pyramid which accepts the whole word image as the input and second Fragment pathway (green color) which accepts the fragment as the input	SoftMax
Proposed	**Multi-Path Hybrid Deep CNN: Model consists two paths for local features extraction then summed with HOG (Histogram Oriented Gradients) features to construct global features**	**SoftMax, SVM (Support Vector Machine)**

3. Proposed Methodology

The suggested methodology for off-line writer recognition utilizing a novel multi-path hybrid CNN (Convolution Neural Network) deep learning model is discussed in this section. Data pre-processing is the preliminary stage of this model pipeline. To acquire the necessary data, some morphological procedures are applied to datasets during the model learning phase; the model evaluation is carried out to examine the accuracy rate. The model’s pipeline is covered in detail in later subsections.

3.1 Data Pre-Processing

Data pre-processing introduced in order to increase image quality by removing unnecessary data and enhancing information readability for creating appropriate images. This study uses CVL and a bilingual dataset. CVL is an open-source benchmark dataset, which includes scanned handwritten images of English corpus. The proposed model is then used to extract the features from the image data. The model's learning phase is carried out using the features that were extracted. CVL dataset contains scanned image set from 310 writers of English and German corpus. The sample handwritten paragraph image from CVL dataset shown in figure-3(a).

The bilingual dataset is an in-house handwritten dataset with Hindi and English mixed script as in figure-3(b). There are average six handwritten text paragraph images for every writer. It is essential to have a large amount of training data in order to construct an effective deep learning model. In the preprocessing phase, a data augmentation algorithm as shown in algorithm-1 is used to generate additional data from an existing dataset. The proposed model is trained with augmented paragraph image patches, which are generated using word images. Word images are already available for the CVL dataset. For in-house dataset, first we have word extraction algorithm to segment the words from given paragraph images. Word segmentation procedure is given in figure-4. Sample handwritten text line image is shown in figure-4(a) from bi-lingual database for word segmentation. Text line image converted into binary form using Otsu[28] method as in figure-4(b). Then Gaussian filter[29] is applied on binary text line image to find connected components as a word, shown in figure-4(c). Again binarization is applied on filtered blurred image to get region of the bounding box of a word as in figure-4(d) & figure-4(e).

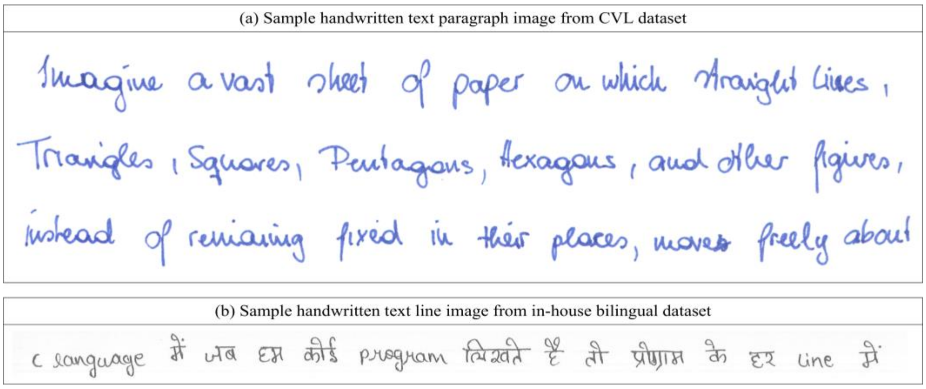


Figure 3: Sample images from CVL and in-house datase

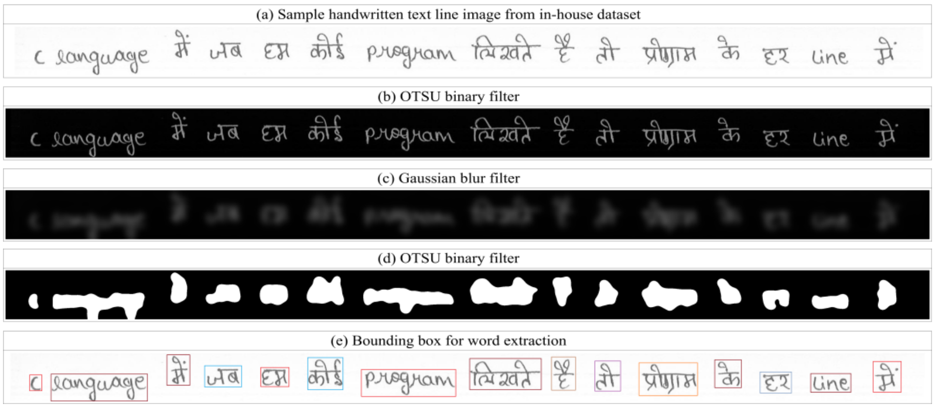


Figure 4: Word Segmentation Procedure

Figure-5 illustrates the complete data augmentation workflow. Consequently, during the dataset preprocessing step, binary and skeleton image patches are constructed for the proposed CNN model. Binarization is a method that is frequently used to reduce background noise in images. It also improves the sharpness and clarity of the contours of various items in the image. This feature extraction helps model learning. Since there are so many writers, the dataset contains a wide range of diverse handwriting styles. Consequently, the potential for varying pen stroke widths also impacts the model's performance during training. Binary images are skeletonized so that the stroke width is uniform throughout all of the data. While most of the original foreground pixels are discarded, the region's original size and connectivity are preserved when this method is used to decrease foreground regions in a binary image.

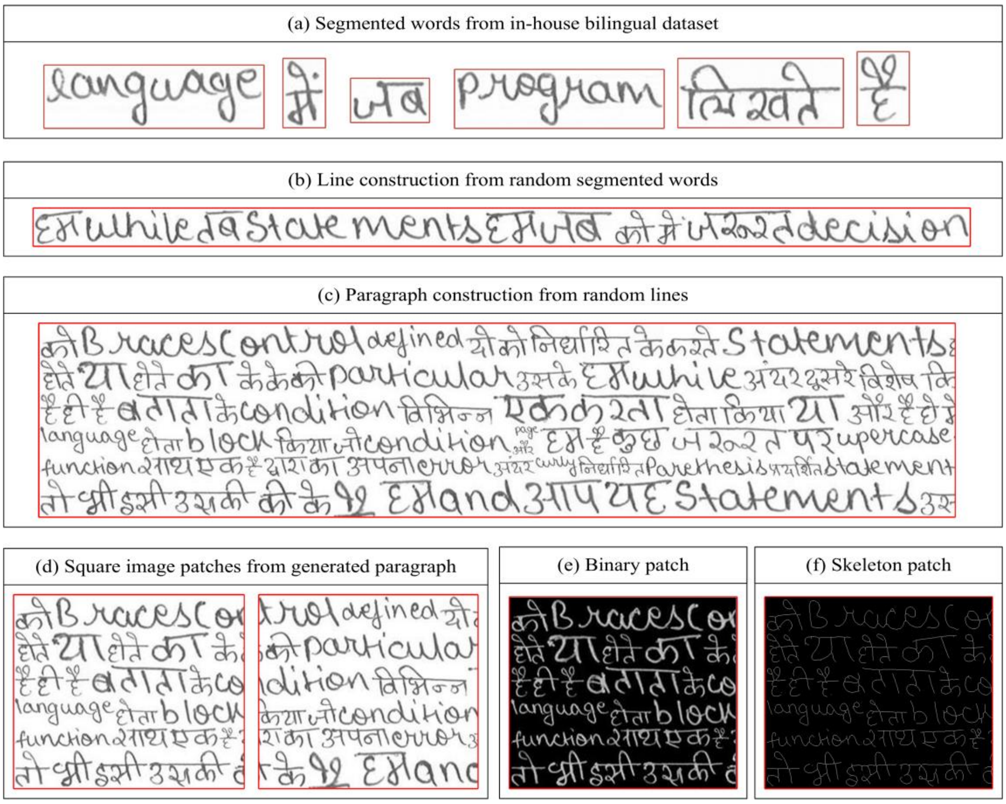


Figure 5: Data augmentation procedure (a) The input word images (b) Constructed line from word images (c) Constructed paragraph from line (d) Three square image patches from augmented paragraph (e) Binary paragraph image patch (f) Skeleton paragraph image

Algorithm 1: Data Augmentation

1: Input ←segmented words as shown in figure-5(a).

2: Permutation of 35 random words is utilized to construct a line as shown in figure-5(b). Total 120 lines constructed.

3: Permutation of 6 random lines is utilized to construct a paragraph as shown in figure-5(c). Total 500 paragraphs constructed for each class.

4: Due to the CNN model's constant requirement for squared shape image input for classification tasks. Therefore, a patch cropping approach is employed to deal with non-squared paragraph images. All paragraph images should first be resized to a fixed width and height of 224 pixels while keeping aspect ratio. Then, from the resized paragraph image data, crop the 224x224 image patches as shown in figure-

5(d).

5: Using Otsu's approach, all paragraph image patches are turned into a single channel, known as binarization as shown in figure-5(e).

6: The binary image patches are used to generate the skeleton image. Skeleton is referred to as a morphological operation since it gives useful information about the image's shapes. Sample skeleton image patch shown in figure-5(f)

3.2 Multi-Input Hybrid Deep CNN

The AlexNet[30] design serves as inspiration for the suggested CNN model. It performed exceptionally well in the 2012 ImageNet[31] challenge. Given that it is a multi-input CNN, the suggested architecture comprises of two pathways (Path-I and Path-II) as shown in figure-6. Both path-I and path-II accept input in the form of skeletonized binary squared image patches and binarized squared image patches, respectively. Both routes are a part of the model's feature extraction stage. After each convolution layer in the feature extraction phase, the model outputs local feature maps, which are subsequently down-sampled by a pooling layer. As it comprises of three convolution layers and three max-pooling layers to increase the depth of the network, and two completely connected (dense) layers. Both paths have the same arrangement of convolution and pooling layers. The first convolution layer's kernel/filter size is 5x5, while the next convolution layers have a size of 3x3. For all convolution layers, the stride value is fix s=1, and for all pooling levels, it is 2.

After each convolution and dense layer, the ReLU (Rectified Linear Unit) [32] function is employed to keep the neurons in the network active. If the input values of the neuron fall below the threshold, the neuron dies or deactivates. ReLU returns 0 for negative input values and returns the input value for positive input values. Thus, it can be expressed as:

$$\mathrm{ReLU}(x) = \max(0, x) \tag{1}$$

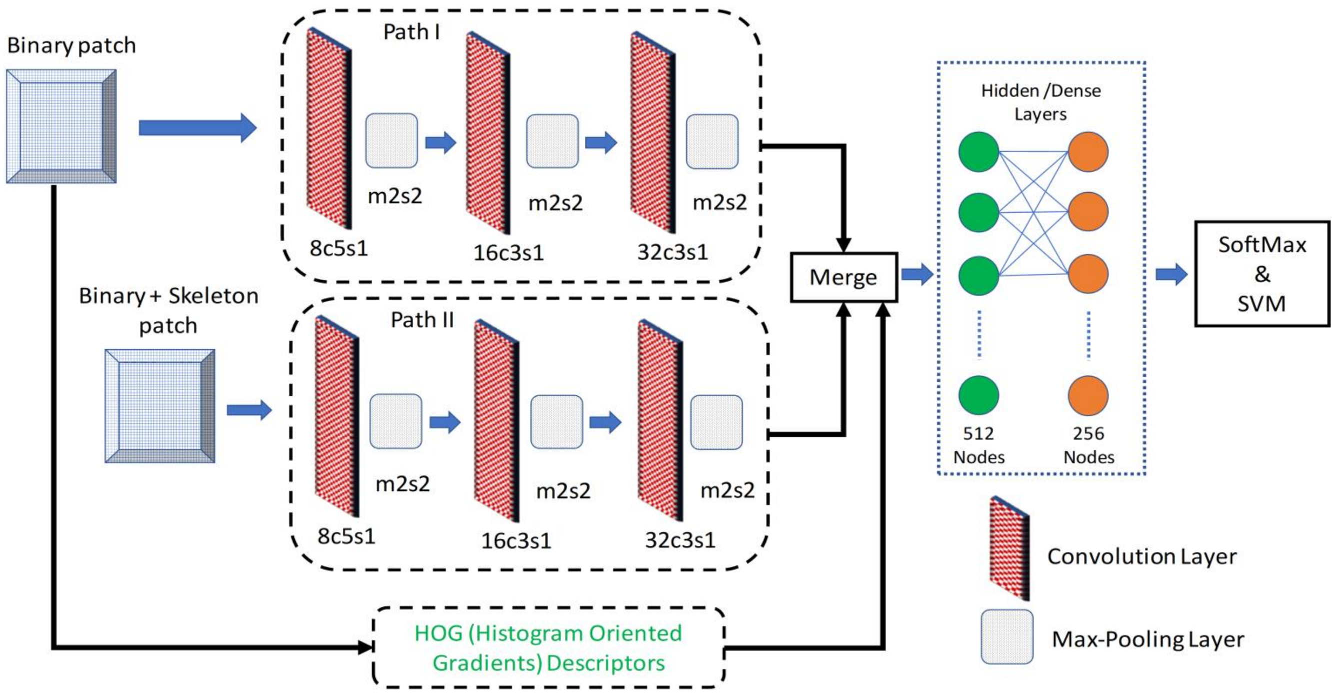


Figure 6: Multi-Input Hybrid Deep CNN. 8c5s1 notation means that convolution layer with total 8 filters, 5x5 filter size and stride is 1. m2s2 means max pooling layer with filter 2x2 and stride=2. D512 denotes that dense layer with 512 nodes.

A conventional manual feature extraction technique known as HOG(Histogram Oriented Gradients) is combined with CNN local features to create the hybrid model. HOG uses the orientation or angle information to extract the edge information of objects in an image. The gradient or magnitude of the pixels from the image patch is used to determine the angle values. The figure-7 presents a visual depiction of the bilingual in-house dataset enhanced paragraph patch's HOG feature descriptors.

HOG are calculated on a binary paragraph image patch. First binary patch is separated into a predetermined number of pixel blocks. The gradient or magnitude of the pixels in the x-y direction is then determined for each block using the following equation:

$$D_x = X2 - X1 \tag{2}$$

$$D_y = Y2 - Y1 \tag{3}$$

Where D_x is X-gradient & D_y is Y-gradient. X1 and X2 are neighboring pixels in X direction. Y1 and Y2 are neighboring pixels in Y direction. In addition, the following equation is applied to the X-Y gradients to determine the orientation:

$$\theta = \tan^{-1}(D_x/D_y) \tag{4}$$

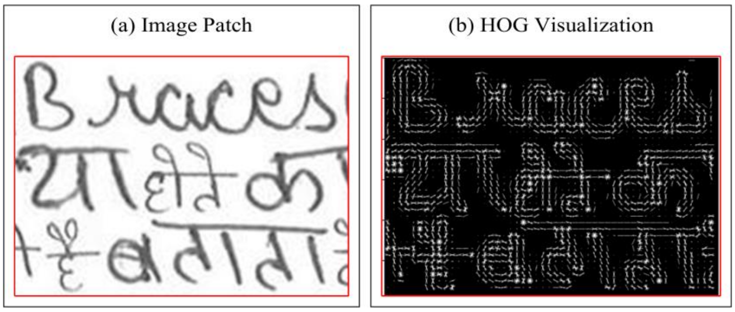


Figure 7: HOG Visualization

In model architecture, the retrieved local features of CNN are formed as the output of paths I and II. As a result, all of the retrieved features, such as those from CNN and HOG, are combined to generate global hybrid features. These hybrid attributes are then used in the model's training phase. The following dense layers receive these combined features as input in the form of 1D vectors. The model has two hidden layers that are fully connected layers with 512 and 128 nodes, as seen in figure-6. The multi-layer perceptron network for model training is created by these two dense/hidden layers. Following model training, the writer class of the queried image document is predicted using two classifiers separately for multi-class classification: Softmax[33] and Support Vector Machine(SVM)[34].

4. Experimental Results

Various types of handwritten document images from the English corpus are used in the experiments. It utilizes the benchmark dataset CVL. A total of 310 CVL writers contributed samples of their handwriting. A bilingual in-house dataset is also prepared utilized. Total 50 writers with 6 paragraphs for each contributed. The augmented paragraph generated during the preprocessing phase. The preprocessed dataset is divided into training and testing part in 70:30 ratios.

The feature extraction phases of proposed model have three convolution layer and max-pooling layers. Convolution layer generate several distinct feature maps. Further the feature maps are down-sampled to reduce it's dimensionality. The convolution and pooling feature maps shown in figure-8. By equation-2, 3, 4 gradients and orientations are calculated. Pixel blocks of 16x16 are applied on image patch for gradient orientation. These orientation values are used as handcrafted features and further merged into CNN local features. The histogram of calculated gradient orientations is shown in figure-9.

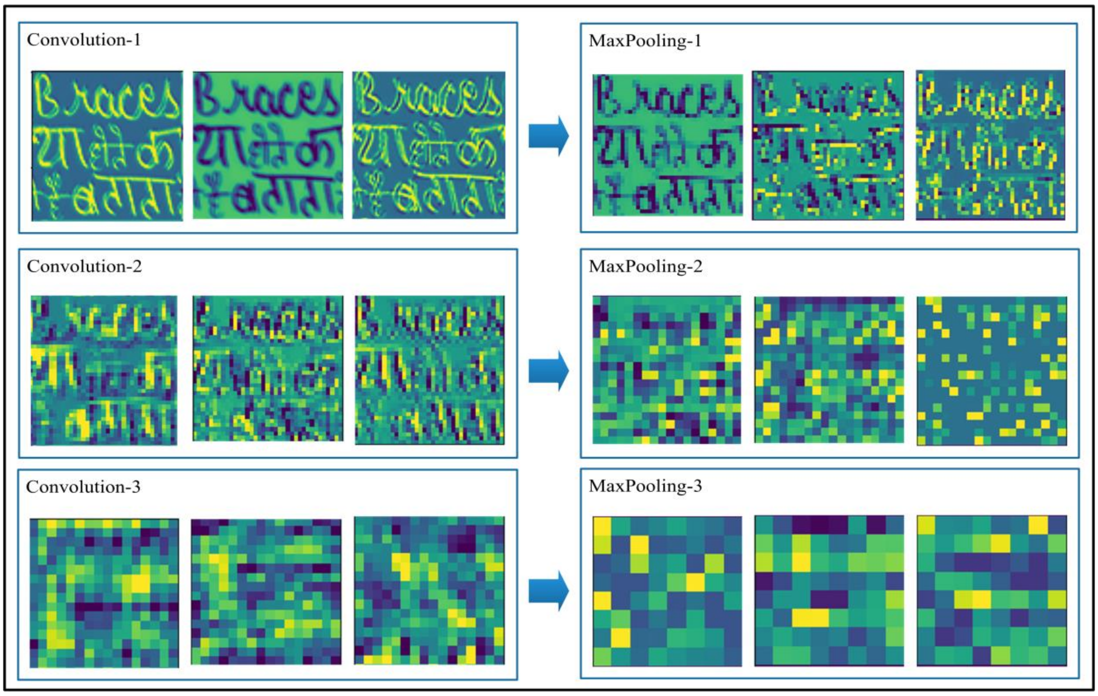


Figure 8: Feature maps generated by CNN

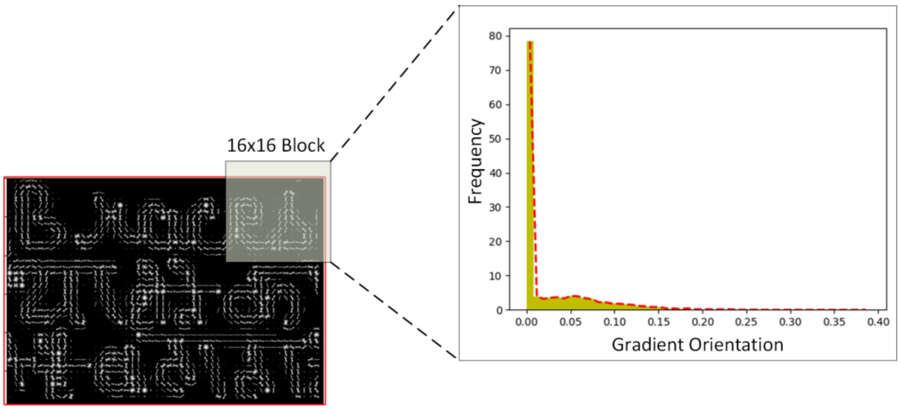


Figure 9: Histogram of gradient orientations

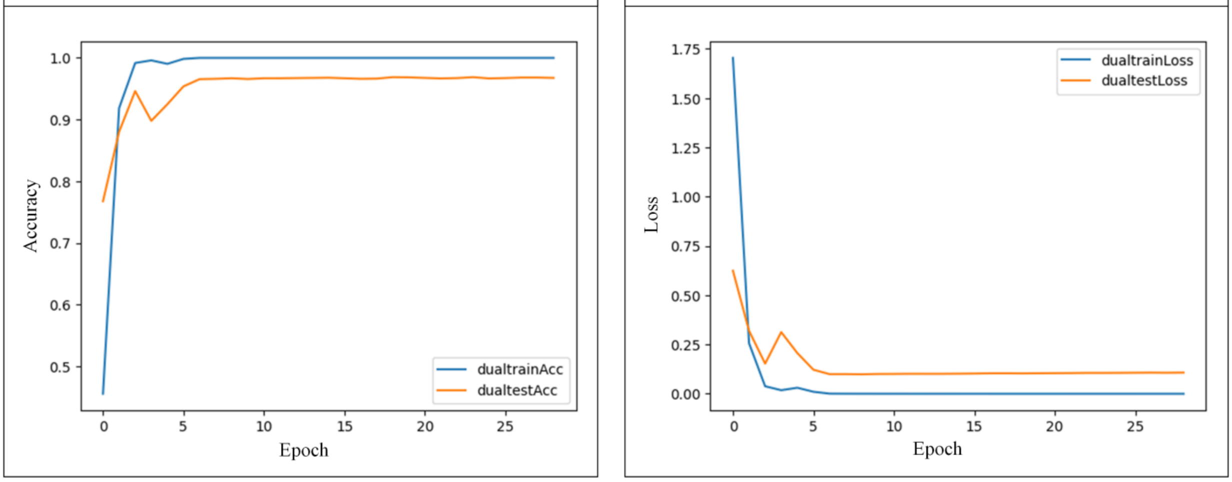
Figure-10 depicts the display of accuracy and loss curves for training and validation phases of the proposed model. CVL and in-house bilingual dataset are used to train and validate the model separately. Table-2 shows the accuracy results of model with both datasets. Model performed better with SVM classifier by utilizing CNN local features. Table-3 displays the performance comparison between the proposed model and the other models. Multi-Input Hybrid Deep CNN has improved accuracy rates than the other methods stated.

Table 2: Proposed model results

Classifier	**Accuracy (%)**	
	CVL	**Bilingual In-house**
Soft-max	97.92	96.84
SVM	98.23	96.93

Table 3: Performance analysis of other methodologies with proposed methodology

Author	Dataset	Script Language	Accuracy (%)
Bulacu et al. [11]	Firemaker	Dutch	86.0
	IAM	English	89.0
Jain et al. [12]	IAM	English	89.6
Xing et al. [16]	IAM	English	89.0
Christlein et al. [17]	ICDAR 2017	Historical Documents (English, Dutch)	88.9
Adak et al. [25]	In-house Bengali	Bengali	90.43
Sheng et al. [26]	CVL	English	94.3
	IAM	English	85.2
Proposed	**CVL**	**English**	**98.23**
	Bi-Lingual	**English + Hindi**	**96.93**



(a) Accuracy Curve of Training and Testing phase of Proposed Model with Bilingual In-house Dataset

(b) Loss Curve of Training and Testing phase of Proposed Model with Bilingual In-house Dataset

Figure 10: Model training-testing performance with bilingual in-house dataset (a) Accuracy Curve (b) Loss Curve

5. Conclusion

This study presents a novel hybrid feature-based approach for author recognition in handwritten monolingual and bilingual corpora. CVL, a English script benchmark dataset is utilized, which were contributed by multiple authors with different handwriting styles. A bilingual dataset with English and Hindi is created inhouse. The advantage of the proposed multi-input hybrid CNN is that it can learn patterns using global features that combine deep learning attributes with handcrafted attributes. A data augmentation strategy is also employed to produce additional data from an existing dataset, as a substantial amount of training data is required to construct a reliable CNN model. The system demonstrated highest accuracy of 98.23% with CVL and 96.93% with in-house data based on the results of experiments. The proposed model is capable of handling handwritten scripts in single and mixed languages.

References

[1] Dargan S., Kumar M. , "Writer Identification System for Indic and Non-Indic Scripts: State-of-the-Art Survey", Arch Computat Methods Eng 26, 12831311 (2019). doi: 10.1007/s11831-018-9278-z

[2] Xiong, Yu-Jie, et al. "Off-line Text-Independent Writer Recognition: A Survey.", International Journal of Pattern Recognition and Artificial Intelligence, vol. 31, no. 05, World Scientific Pub Co Pte Lt, Feb. 2017, p. 1756008. Crossref, doi: 10.1142/s0218001417560080.

[3] Halder, C., Obaidullah, S.M., Roy, K. (2016). "Offline Writer Identification and VerificationA State-of-the-Art", In: Satapathy, S., Mandal, J., Udgata, S., Bhateja, V. (eds) Information Systems Design and Intelligent Applications. Advances in Intelligent Systems and Computing, vol 435. Springer, New Delhi. doi: 10.1007/978-81-322-2757-1 17

A. Kumar and K. Bhatia, "A survey on offline handwritten signature verification system using writer dependent and independent approaches", 2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Fall), Bareilly, India, 2016, pp. 1-6, doi: 10.1109/ICACCAF.2016.7748998.

[4] Waleed Abu-Ain, Siti Norul Huda Sheikh Abdullah, Bilal Bataineh, Tarik Abu-Ain, Khairuddin Omar, "Skeletonization Algorithm for Binary Images", Procedia Technology, Volume 11, 2013, Pages 704-709, ISSN 2212-0173. doi:10.1016/j.protcy.2013.12.248.

[5] Reza Ebrahimzadeh and Mahdi Jampour. "Efficient Handwritten Digit Recognition based on Histogram of Oriented Gradients and SVM", International Journal of Computer Applications 104(9):10-13, October 2014. doi:10.5120/18229-9167

[6] E. Chen, X. Wu, C. Wang and Y. Du, "Application of Improved Convolutional Neural Network in Image Classification", 2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 2019, pp. 109- 113, doi: 10.1109/MLBDBI48998.2019.00027.

[7] Z. Li, F. Liu, W. Yang, S. Peng and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects", in IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 12, pp. 6999-7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.

[8] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng and M. Chen, "Medical image classification with convolutional neural network", 2014 13th International Conference on Control Automation Robotics & Vision (ICARCV), Singapore, 2014, pp. 844-848, doi: 10.1109/ICARCV.2014.7064414.

[9] H. Lee and H. Kwon,"Going Deeper With Contextual CNN for Hyperspectral Image Classification", in IEEE Transactions on Image Processing, vol. 26, no. 10, pp. 4843-4855, Oct. 2017, doi: 10.1109/TIP.2017.2725580.

[10] M. Bulacu and L. Schomaker, "Text-Independent Writer Identification and Verification Using Textural and Allographic Features", in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 4, pp. 701-717, April 2007, doi: 10.1109/TPAMI.2007.1009

[11] R. Jain and D. Doermann, "Offline Writer Identification Using K-Adjacent Segments", 2011 International Conference on Document Analysis and Recognition, Beijing, China, 2011, pp. 769-773, doi: 10.1109/ICDAR.2011.159.

[12] Marti, UV., Bunke, H. "The IAM-database: an English sentence database for offline handwriting recognition", IJDAR 5, 3946 (2002). doi: 10.1007/s100320200071

[13] R. Jain and D. Doermann, "Combining Local Features for Offline Writer Identification", 2014 14th International Conference on Frontiers in Handwriting Recognition, Hersonissos, Greece, 2014, pp. 583-588, doi: 10.1109/ICFHR.2014.103.

[14] X. Wu, Y. Tang and W. Bu, "Offline Text-Independent Writer Identification Based on Scale Invariant Feature Transform", in IEEE Transactions on Information Forensics and Security, vol. 9, no. 3, pp. 526-536, March 2014, doi: 10.1109/TIFS.2014.2301274.

[15] L. Xing and Y. Qiao, "DeepWriter: A Multi-stream Deep CNN for Text-Independent Writer Identification", 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, 2016, pp. 584-589, doi: 10.1109/ICFHR.2016.0112.

[16] V. Christlein, M. Gropp, S. Fiel and A. Maier, "Unsupervised Feature Learning for Writer Identification and Writer Retrieval", 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017, pp. 991-997, doi: 10.1109/ICDAR.2017.165.
[17] Shiming Chen, Yisong Wang, Chin-Teng Lin, Weiping Ding, Zehong Cao, "Semi-supervised feature learning for improving writer identification", Information Sciences, Volume 482, 2019, Pages 156-170, ISSN 0020-0255, doi: 10.1016/j.ins.2019.01.024.
[18] Bin Li, Dimas Lima, "Facial expression recognition via ResNet-50", International Journal of Cognitive Computing in Engineering, Volume 2, 2021, Pages 57-64, ISSN 2666-3074. doi: 10.1016/j.ijcce.2021.02.002.
[19] F. Kleber, S. Fiel, M. Diem and R. Sablatnig, "CVL-DataBase: An Off-Line Database for Writer Retrieval, Writer Identification and Word Spotting", 2013 12th International Conference on Document Analysis and Recognition, 2013, pp. 560-564, doi: 10.1109/ICDAR.2013.117.
[20] Kumar, M., Jindal, M.K. & Kumar, M. , "Distortion, rotation and scale invariant recognition of hollow Hindi characters", Sdhan 47, 92 (2022). doi: 10.1007/s12046-022-01847-w.
[21] Kaur, R.P., Jindal, M.K. & Kumar, M. , "Text and graphics segmentation of newspapers printed in Gurmukhi script: a hybrid approach", Vis Comput 37, 16371659 (2021). doi: 10.1007/s00371-020-01927-0.
[22] Kaur, R.P., Kumar, M. & Jindal, M.K. , "Performance evaluation of different features and classifiers for Gurumukhi newspaper text recognition", J Ambient Intell Human Comput (2022). doi: 10.1007/s12652-021-03687-8.
[23] B. Kumar, P. Kumar and A. Sharma, "RWIL: Robust Writer Identification for Indic Language", 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2018, pp. 695-700, doi: 10.1109/ICCONS.2018.8662997.
[24] C. Adak, B. B. Chaudhuri and M. Blumenstein, "A Study on Idiosyncratic Handwriting with Impact on Writer Identification", 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, 2018, pp. 193-198, doi: 10.1109/ICFHR-2018.2018.00042.
[25] Sheng He, Lambert Schomaker, "Deep adaptive learning for writer identification based on single handwritten word images", Pattern Recognition, Volume 88, 2019, Pages 64-74, ISSN 0031-3203, doi: 10.1016/j.patcog.2018.11.003.
[26] S. He and L. Schomaker, "FragNet: Writer Identification Using Deep Fragment Networks", in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3013-3022, 2020, doi: 10.1109/TIFS.2020.2981236.
[27] Xiangyang Xu, Shengzhou Xu, Lianghai Jin, Enmin Song, "Characteristic analysis of Otsu threshold and its applications", Pattern Recognition Letters, Volume 32, Issue 7, 2011, Pages 956-961, ISSN 0167-8655, doi: 10.1016/j.patrec.2011.01.021.
[28] M. Wang, S. Zheng, X. Li and X. Qin, "A new image denoising method based on Gaussian filter", 2014 International Conference on Information Science, Electronics and Electrical Engineering, Sapporo, Japan, 2014, pp. 163-167, doi: 10.1109/InfoSEEE.2014.6948089.
[29] Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Van Esesn, B.C., Awwal, A.A.S. and Asari, V.K., "The history began from alexnet: A comprehensive survey on deep learning approaches", arXiv preprint , doi.: 10.48550/arXiv.1803.01164.
[30] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017, "ImageNet classification with deep convolutional neural network", Commun. ACM 60, 6 (June 2017), 8490, doi: 10.1145/3065386.
[31] Agarap, Abien Fred. , "Deep Learning Using Rectified Linear Units (ReLU)", arXiv, 7 Feb. 2019. arXiv.org, doi:10.48550/arXiv.1803.08375.
[32] Y. Gao, W. Liu and F. Lombardi, "Design and Implementation of an Approximate Softmax Layer for Deep Neural Networks", 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain, 2020, pp. 1-5, doi: 10.1109/ISCAS45731.2020.9180870.

[33] X. Li, L. Wang and E. Sung, "Multilabel SVM active learning for image classification", 2004 International Conference on Image Processing, 2004. ICIP '04., Singapore, 2004, pp. 2207-2210 Vol. 4, doi: 10.1109/ICIP.2004.1421535

Naresh Purohit received his BTech degree in Information Technology from Rajasthan Technical University, Kota, India in 2010, MTech degree in computer engineering from Jodhpur National University, Jodhpur, India, in 2016, and pursuing PhD degree from Bikaner Technical University, Bikaner, India. He is currently working at Jodhpur Institute of Engineering & Technology, Rajasthan, India as an assistant professor. His current research interests include computer vision, pattern recognition using deep learning algorithms. He can be contacted at email: iteng.naresh@gmail.com

Subhash Panwar is an assistant professor in the Department of Computer Science & Engineering, Government Engineering College, Bikaner. He received his MTech degree in computer engineering from Motilal Nehru National Institute of Technology, Allahabad, India and PhD degree from the Malviya National Institute of Technology, Jaipur, India. He has more than two decades of teaching experience. His research area is image fusion, computer vision, computational intelligence and pattern recognition. He has published number of articles in international journal and conferences. He can be contacted at email: panwar.subhash@gmail.com