Handwritten Numeral Recognition Using Transfer Learning Technique

Kazi Ziaul Hassan, 1607061

Md. Golam Kaochhar, 1607100

Supervised by:

Mohammad Insanur Rahman Shuvo

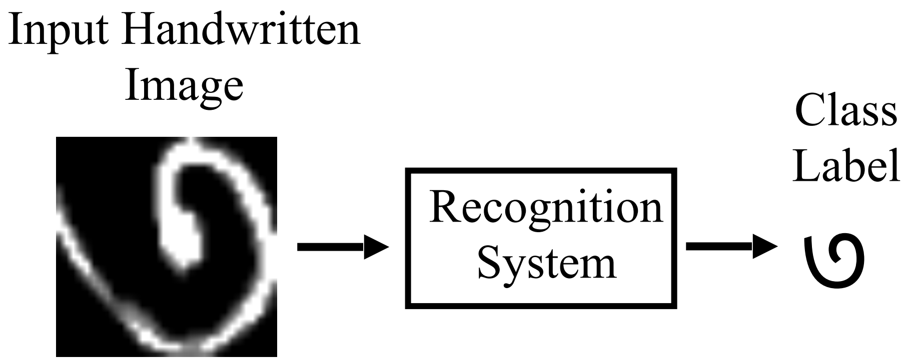
Assistant Professor, Dept. of CSE, KUET

Presentation Outline

- Introduction of Handwritten Numeral Recognition (HNR)
- Basic Steps of HNR
- Aim of this Study
- Related Works
- Proposed Methodology
- Experimental Studies
- Conclusions
- Future Works

Introduction of Handwritten Numeral Recognition (HNR)

Handwritten Numeral Recognition (HNR) and its importance



- HNR is a process which considers images of handwritten numerals as input and classifies the images into different numeral classes.

- Handwritten numerals are used in bank cheques, postal codes and in many other daily uses.

- In this ICT age, automatic recognition of handwritten numerals becomes an important research topic.

Handwritten Numeral Recognition (HNR) and its challenges

- Existence of similar shaped numerals make recognition task difficult.

This numeral is “୫” but it looks like “୬”

This numeral is “୬” but it looks like “୫”

- Same handwritten numeral may look very different in size, shape, and orientation due to different writing patterns of people.

Numeral “୨” looks different in handwritten format due to different writing patterns of people

Basic Steps of HNR

Basic Steps of Traditional HNR

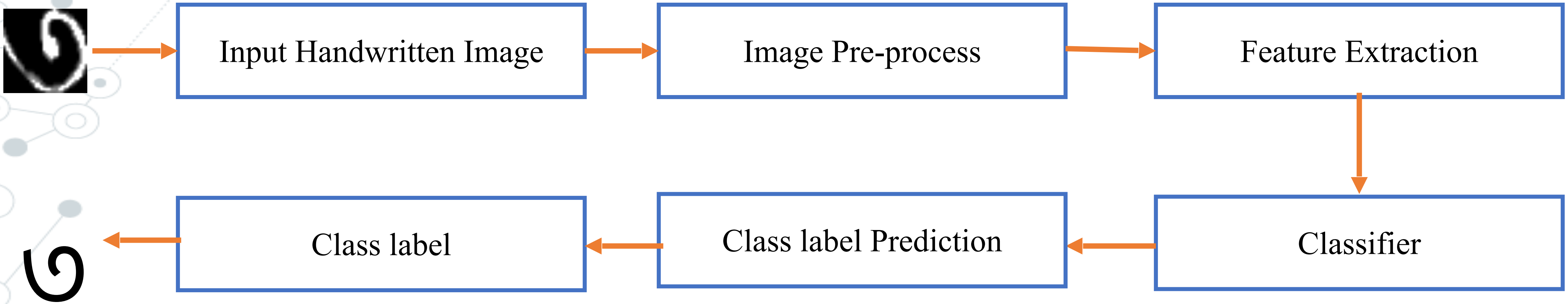


Figure 01: Basic Steps of HNR using Traditional HNR

Basic Steps of HNR using Deep Learning

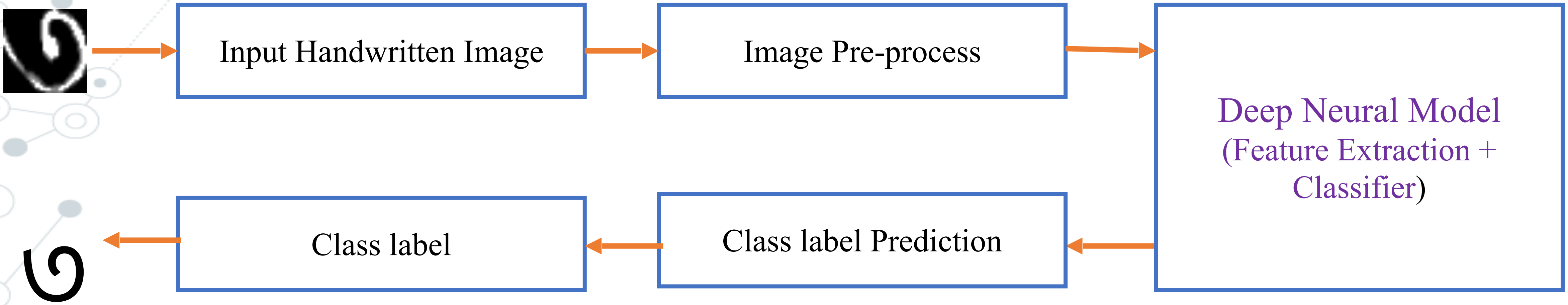


Figure 02: Basic Steps of HNR using Deep Learning

Basic Steps of HNR using Transfer Learning

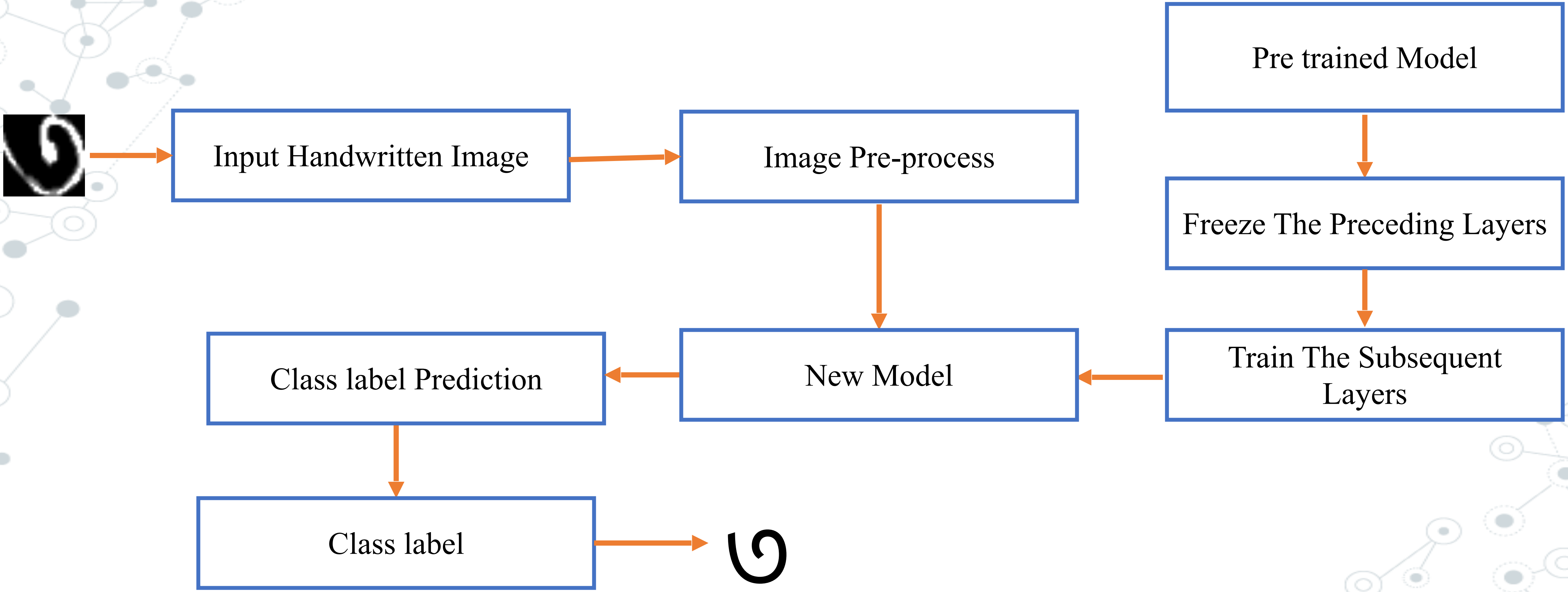


Figure 03: Basic Steps of HNR using Transfer Learning

Basic Steps of HNR with Fine Tuning

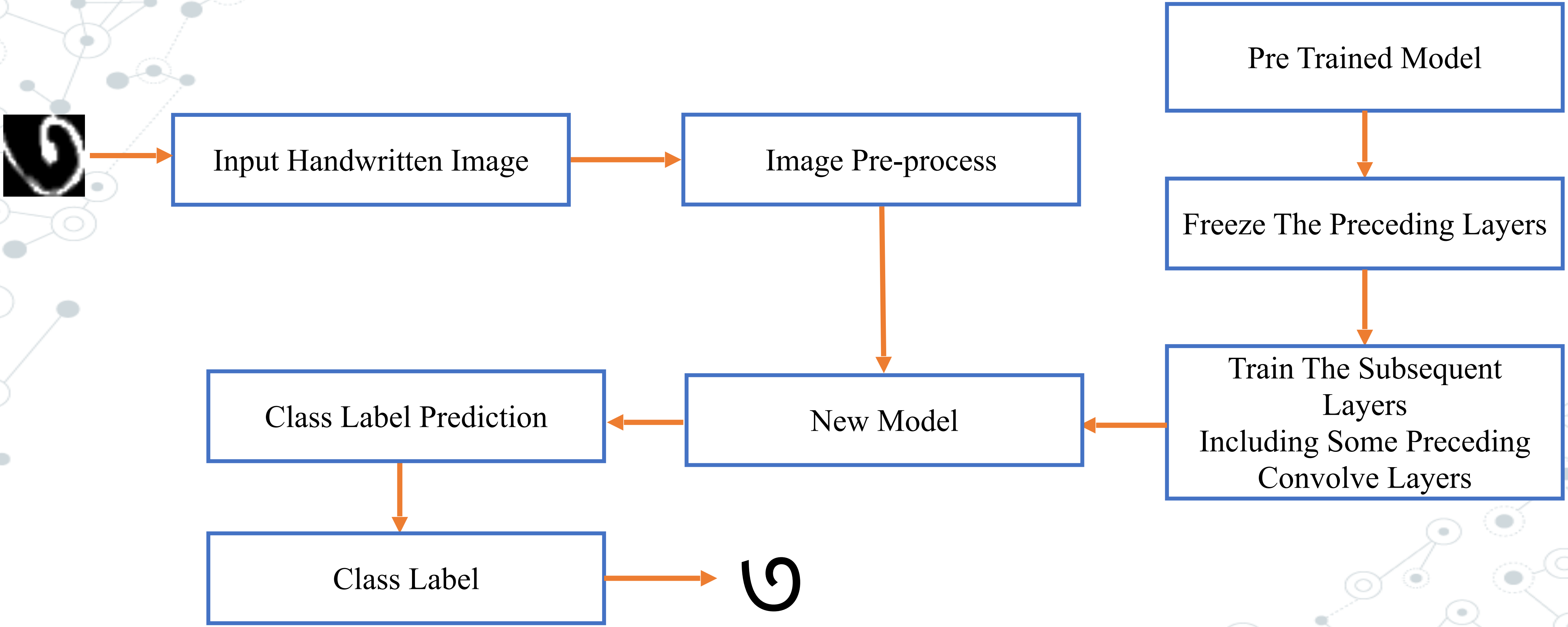


Figure 04: Basic Steps of HNR with fine tuning

Aim of this Study

- Study of existing classification techniques for HNR of different languages.
- Propose a HNR system using transfer learning technique.
- Performance comparison of the proposed system with other prominent works on HNR.

Related Works

Work Ref. [Author, Year]	Method
Basu et al., 2005	MLP with Dempster-Shafter techniques
Pal et al., 2006	Binary decision tree
Wen et al., 2007	Support Vector Machine (SVM)
Akhand et al., 2016	Convolutional neural network (CNN)
Shopon et al., 2016	Auto-encoder (AE) with CNN
Guha et al., 2019	Multilayer perceptron (MLP)
Mahtab et al., 2016	Stacked Auto-encoder (SAE)
Akhand et al., 2018	Convolutional neural network (CNN)
Mahtab et al., 2019	Deep Long Short Term Memory (DLSTM)
Shuvo et al., 2019	Convolutional Auto-Encoder(CAE) with CNN

In Dept. of CSE, KUET

Contribution of the Study

- Build a potential HNR system based on the transfer learning technique.

- Investigate fine tuning (FT), without fine tuning, normal as well as rotational augmentation ($\pm 10^0$, $\pm 20^0$) on the dataset and batch size (BS) to discover the most accurate HNR system.

- Investigate the Bengali and Devanagari language to compare the models to see how well the pattern recognition and feature extraction work.

Proposed Methodology

- Basic Structure of the Proposed HNR System
- Development of Classifier
- Development of Pre-Trained Model
- Development of Transfer Learning Technique

Proposed Methodology: Basic Structure of the Proposed HNR System

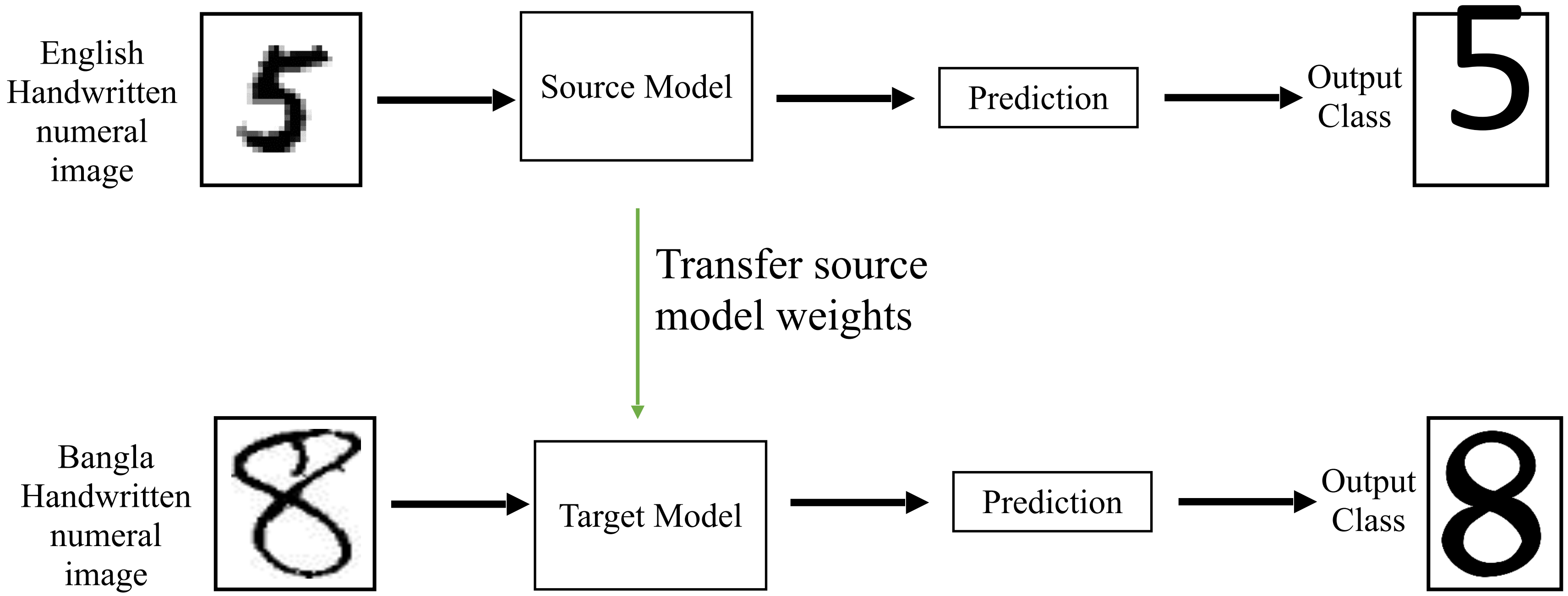


Figure 05 : Structure of the Handwritten numeral recognition using transfer learning technique

Basic Structure of the Overall Systems

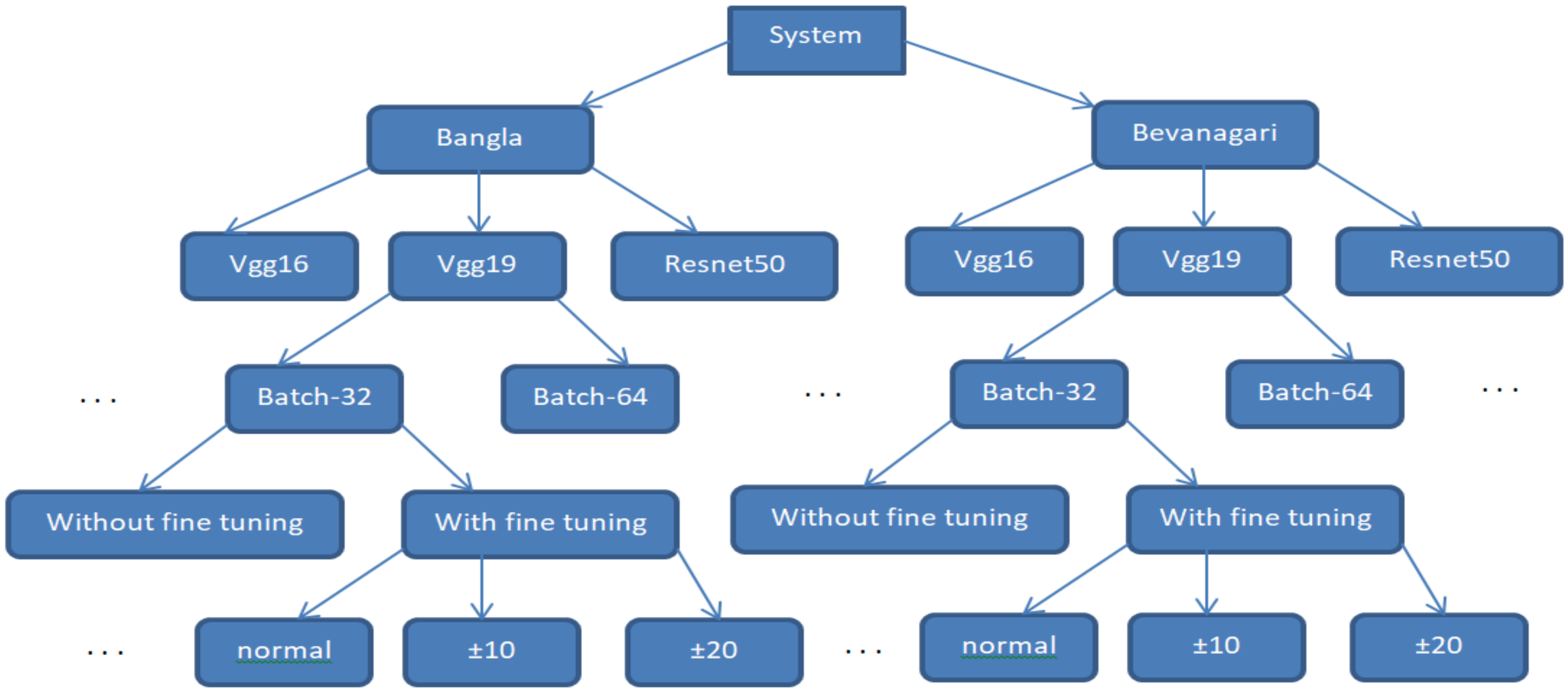


Figure 06 : An illustration of how all the systems have been categorized

Proposed Methodology: Development of Classifier

Convolutional Neural Network (CNN) as Classifier

- The basic CNN structure contains an input layer, subsequent convolution-pooling layer(s), one or more dense layer(s) and an output layer.

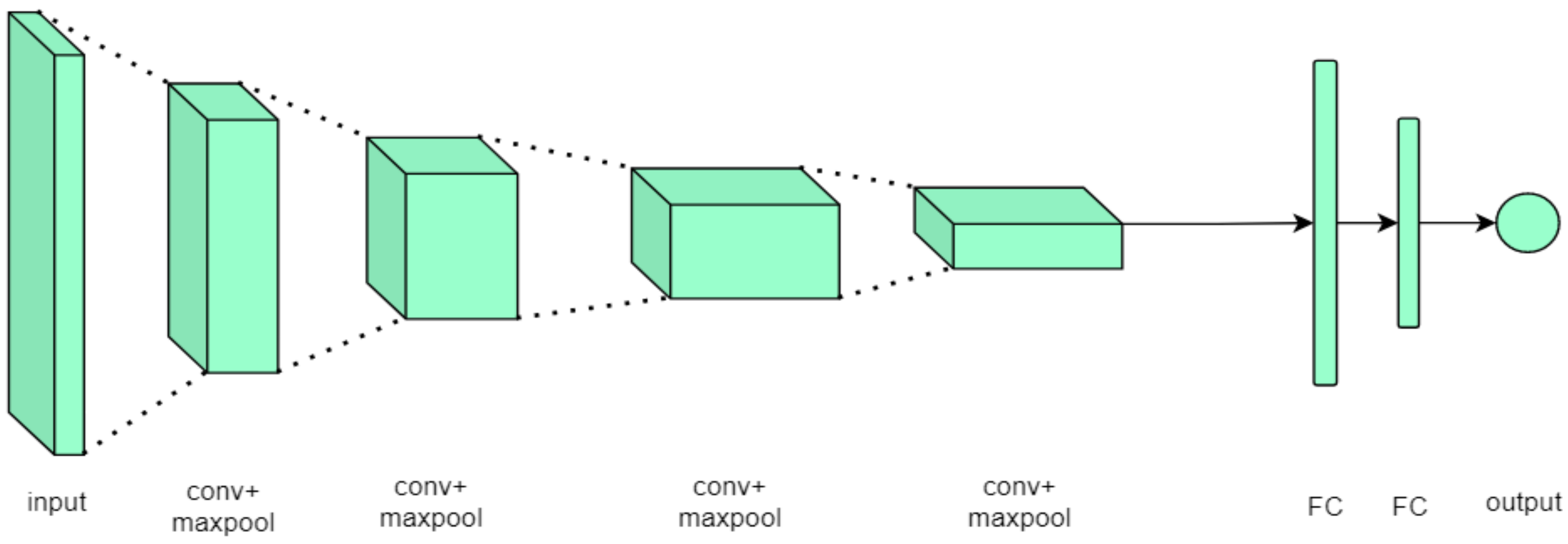


Figure 07 : Basic structure of Convolutional Neural Network (CNN).

Proposed Methodology: Development of Pre Trained Model

Development of Pre Trained Model

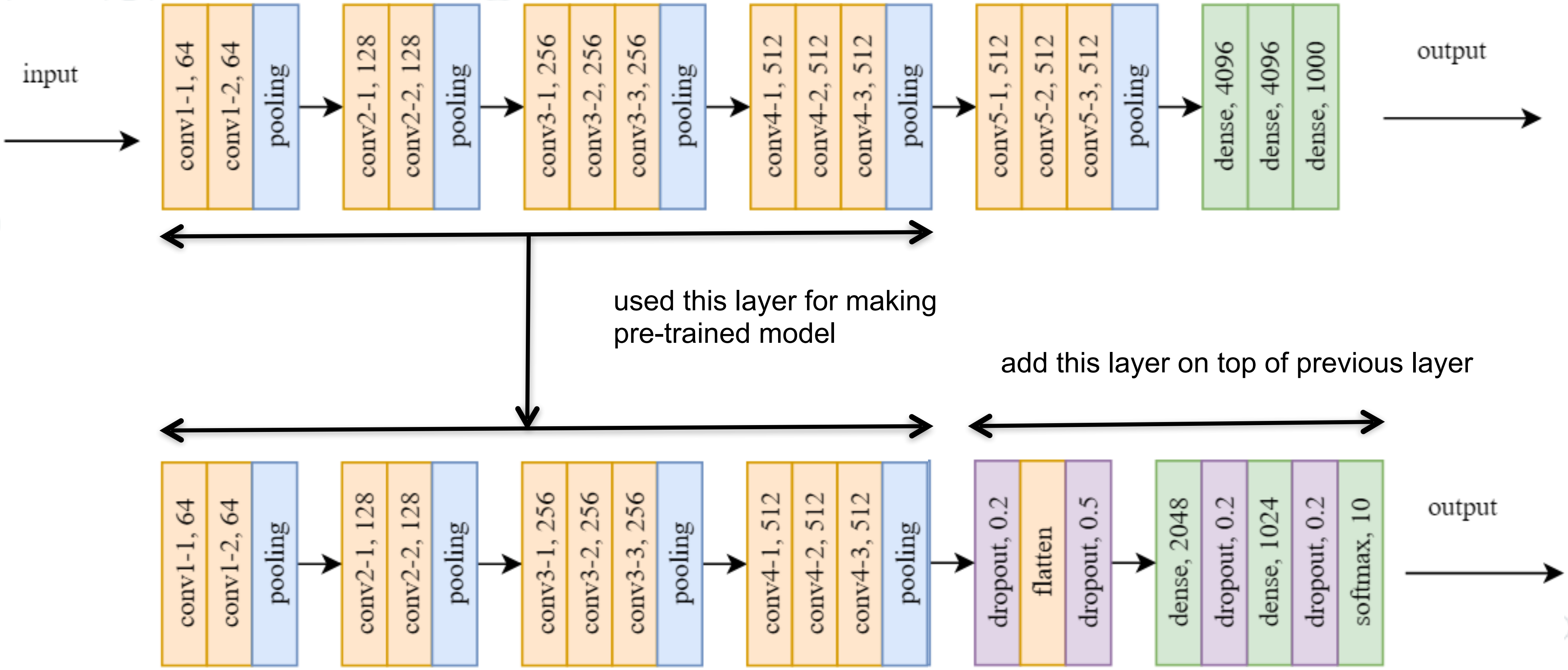


Figure 08: Pre-trained model structure with vgg16

Development of Pre Trained Model

Table 01: Partial summary of pre-trained model vgg16

Layer(Type)	Output shape
input_1 (Input Layer)	(32, 32, 1)
zero_padding2d(Zero Padding 2D)	(38, 38, 1)
block1_conv1 (Conv2D)	(38, 38, 64)
block1_conv2 (Conv2D	(38, 38, 64)
block1_pool (MaxPooling2D)	(19, 19, 64)
block2_conv1 (Conv2D)	(19, 19, 128)
block2_conv2 (Conv2D)	(19, 19, 128)
block2_pool (MaxPooling2D)	(9, 9, 128)
...	…
block4_conv1 (Conv2D)	(4, 4, 512)
block4_conv2 (Conv2D)	(4, 4, 512)
block4_conv3 (Conv2D)	(4, 4, 512)
block4_pool (MaxPooling2D)	(2, 2, 512)
flatten (Flatten)	(2048)
fc1 (Dense)	(2048)
fc2 (Dense)	(1024)
predictions (Dense)	(10)

Development of Pre Trained Model

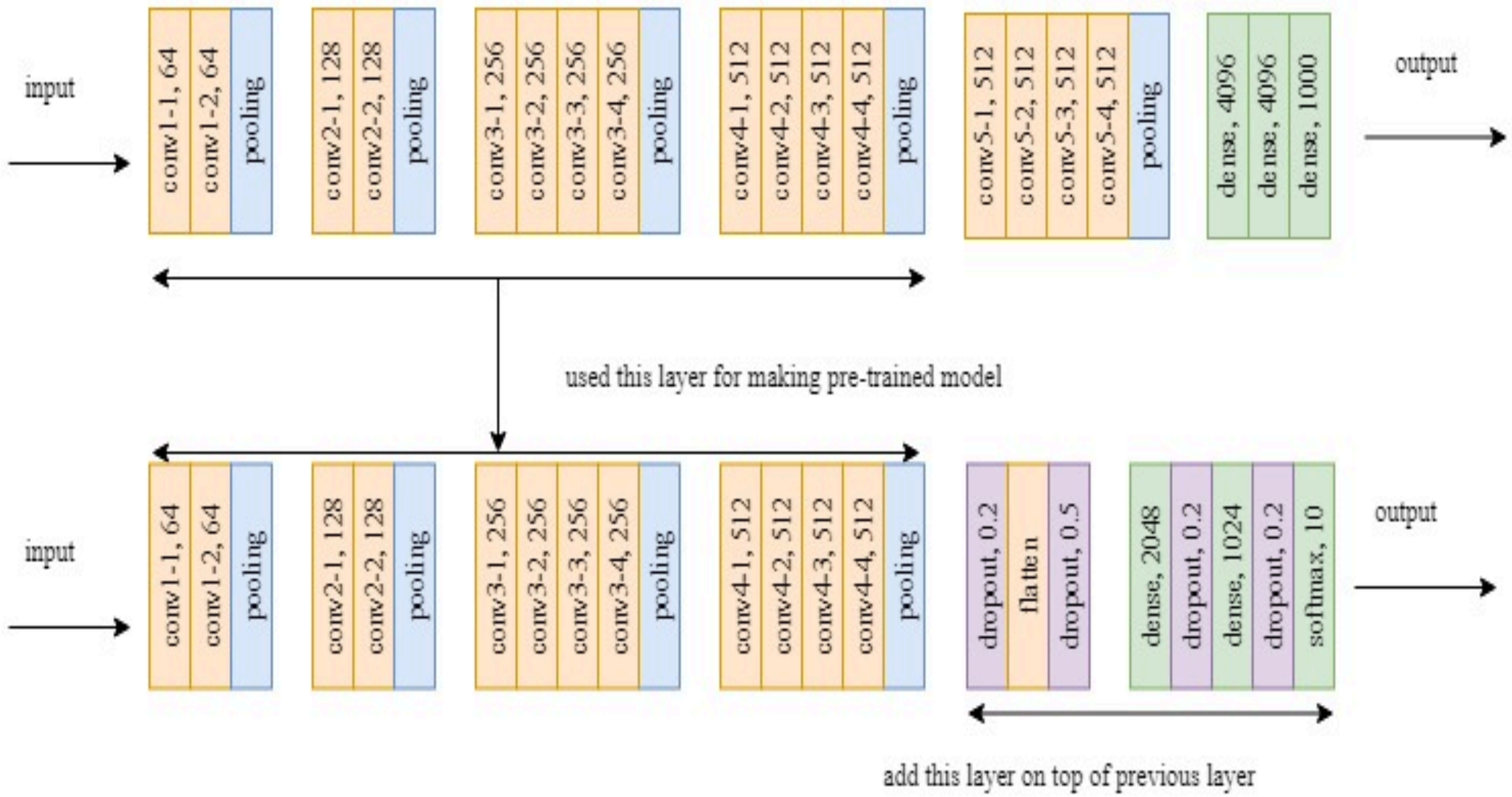


Figure 09: Pre-trained model structure with vgg19

Development of Pre Trained Model

Table 02: Partial summary of pre-trained model vgg19

Layer(Type)	Output shape
input_1 (Input Layer)	(32, 32, 1)
zero_padding2d (Zero Padding 2D)	(38, 38, 1)
block1_conv1 (Conv2D)	(38, 38, 64)
block1_conv2 (Conv2D)	(38, 38, 64)
block1_pool (MaxPooling2D)	(19, 19, 64)
…	…
block3_conv3 (Conv2D)	(9, 9, 256)
block3_pool (MaxPooling2D)	(4, 4, 256)
block4_conv1 (Conv2D)	(4, 4, 512)
block4_conv2 (Conv2D)	(4, 4, 512)
block4_conv3 (Conv2D)	(4, 4, 512)
block4_conv4 (Conv2D)	(4, 4, 512)
block4_pool (MaxPooling2D)	(2, 2, 512)
dropout (Dropout)	(2, 2, 512)
flatten (Flatten)	(2048)
dropout_1 (Dropout)	(2048)
fc1 (Dense)	(2048
fc2 (Dense)	(1024)
predictions (Dense)	(10)

Development of Pre Trained Model

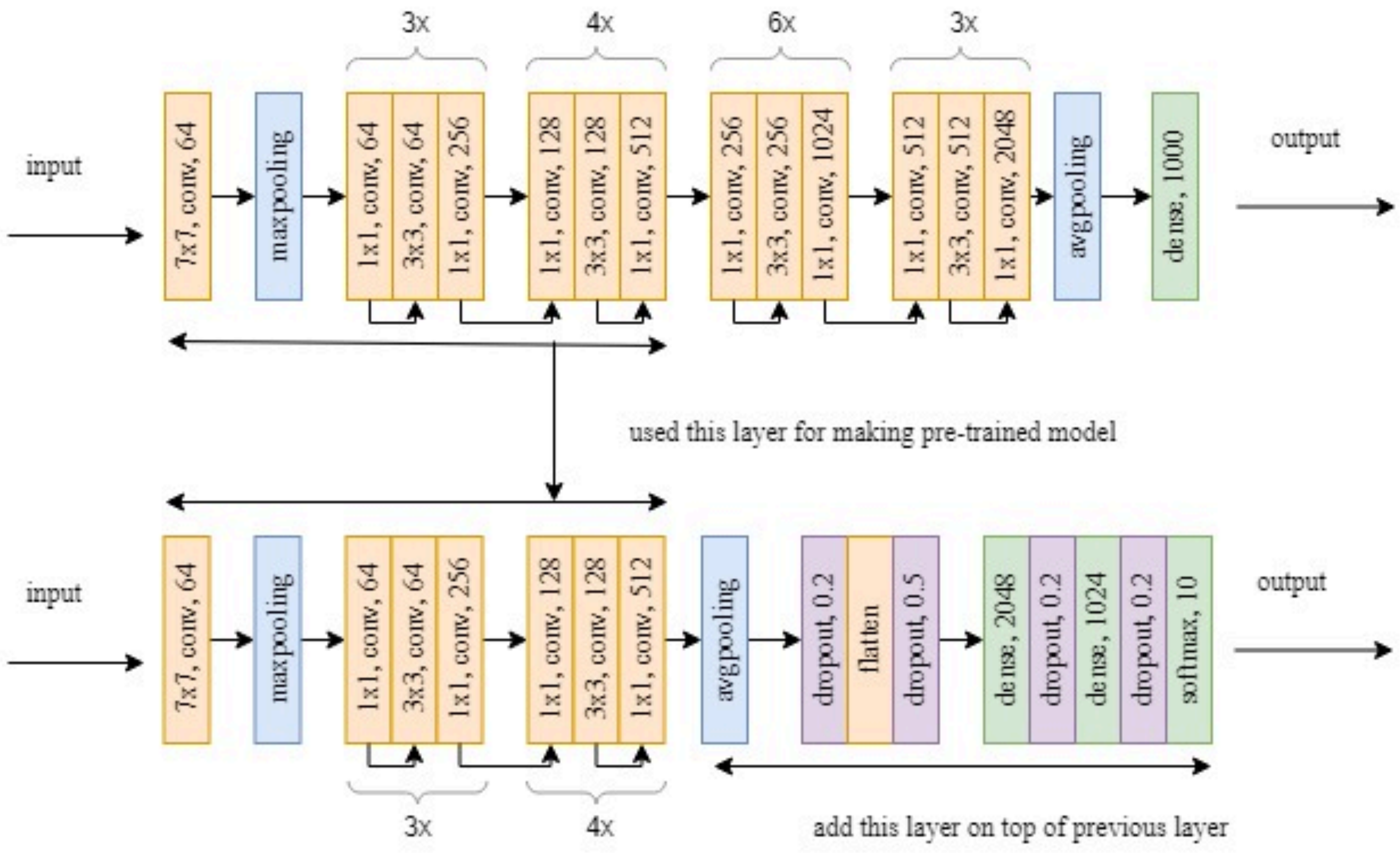


Figure 10: Pre-trained model structure with resnet50

Development of Pre Trained Model

Table 03: Standard structure of resnet50 model

layer name	output size	50-layer	No. of layer
conv1	112x112	7x7, 64, stride 2 3x3 maxpool, stride 2	1
conv2_x	56x56	1x1, 64 3x3, 64 1x1, 256	3x3 = 9
conv3_x	28x28	1x1, 128 3x3, 128 1x1, 512	4x3 = 12
conv4_x	14x14	1x1, 256 3x3, 256 1x1, 1024	6x3 = 18
conv5_x	7x7	1x1, 512 3x3, 512 1x1, 2048	3x3 = 9
-	1x1	average pool, 1000-d, fc, softmax	1

Proposed Methodology: Development of Transfer Learning Technique

Development of Transfer Learning Technique

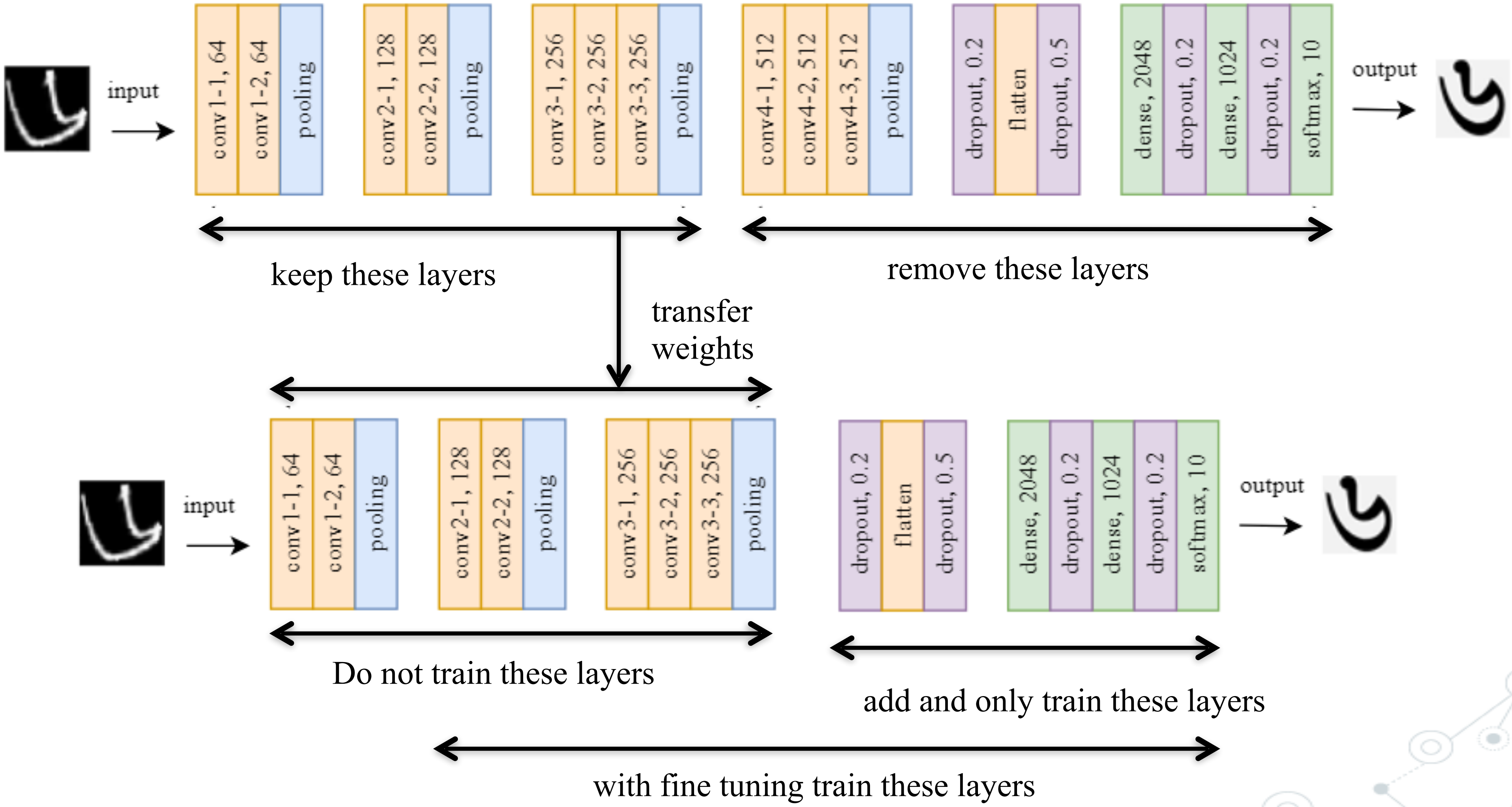


Figure 11: Proposed method structure for vgg16, with and without fine tuning

Development of Transfer Learning Technique

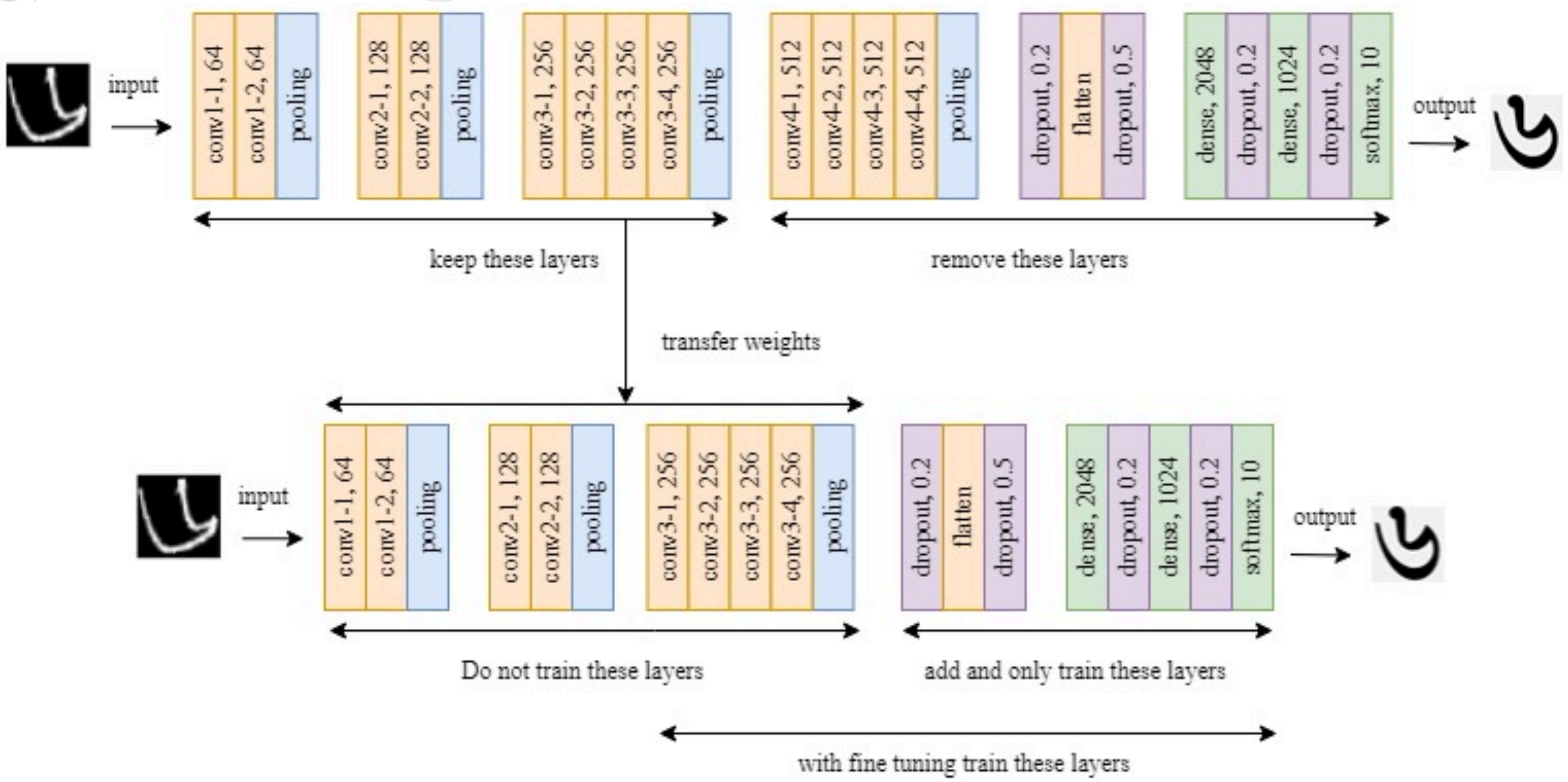


Figure 12: Proposed method structure vgg19, with and without fine tuning

Development of Transfer Learning Technique

Figure 13: Proposed method structure resnet50, with and without fine tuning

Development of Transfer Learning Technique

Table 04: Training Parameter with Transfer Learning Approaches for vgg16

Type of Parameter	Without fine tuning	With fine tuning
Total	13,949,642	13,949,642
Trainable	6,304,778	13,832,586
Non-trainable	7,644,864	117,056

Table 05: Training parameter with transfer learning approaches for vgg19

Type of Parameter	Without fine tuning	With fine tuning
Total	12,829,130	16,902,602
Trainable	10,499,082	16,635,914
Non-trainable	2,330,048	266,688

Experimental Studies

- Experimental Pre-trained Dataset
- Benchmark Dataset
- Experimental Setup
- Experimental Results and Analysis

Experimental Pre-trained Dataset

- English handwritten numeral images MNIST dataset a subset of a larger set available from NIST.

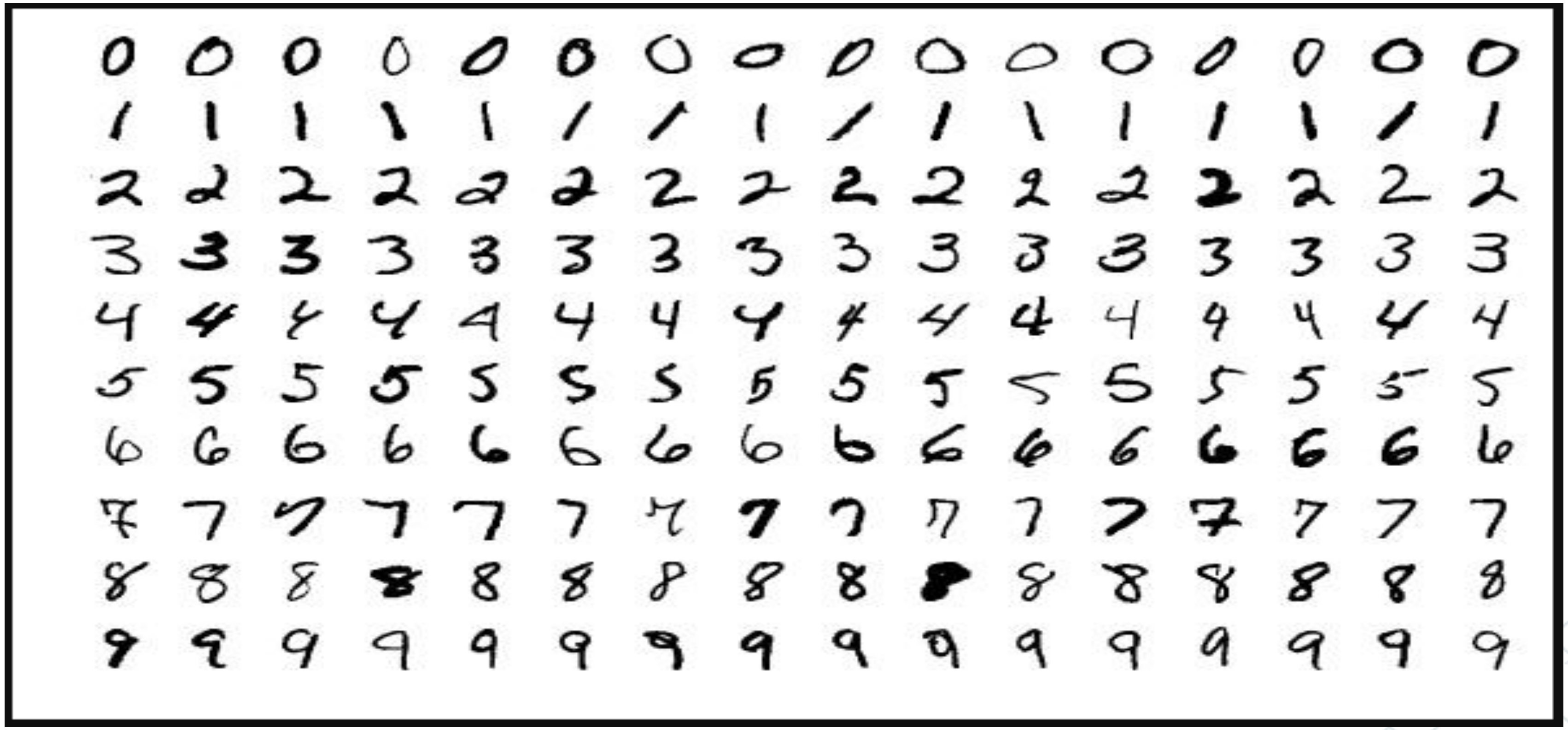


Figure 14: Samples of English handwritten numerals from MNIST dataset [8].

Experimental Pre-trained Dataset Info

- English handwritten numeral images MNIST dataset a subset of a larger set available from NIST.

- MNIST Stands for the Modified National Institute of Standards and Technology

- Having set of 60,000 training examples and a test set of about 10,000 examples

- The digits have been size-normalized and centered in a fixed-size image having square shape 28x28 pixel grayscale images of handwritten single digits between 0 and 9.

Experimental Studies: Benchmark Dataset

- Bengali handwritten numeral images dataset preserved by Computer Vision and Pattern Recognition (CVPR) Unit of Indian Statistical Institute (ISI).

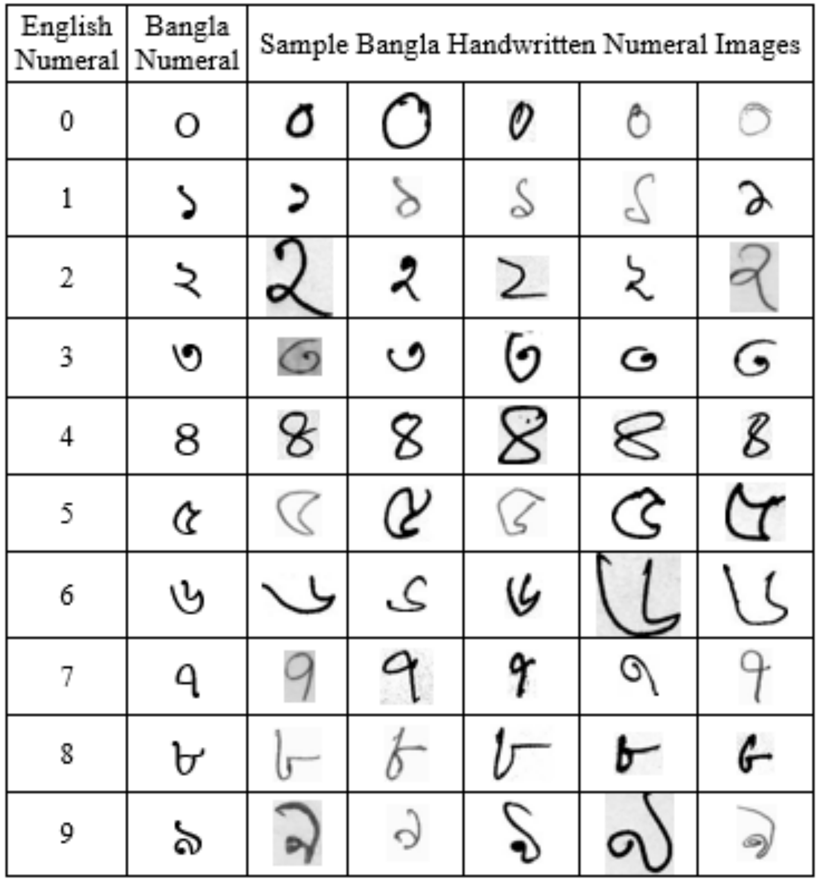


Figure 15: Samples of Bengali handwritten numerals from the CVRP, ISI dataset. [9]

Experimental Studies: Experimental Setup

- The size of the input handwritten numeral images is fixed which is 32x32.
- The considered rotation angles for training data augmentation are ±10° and ± 20°.
- Batch Size (BS) is 32 and 64.
- No of training images: 18000 (=1800×10) and testing images: 4000 (=400×10).

Experimental Studies: Benchmark Dataset

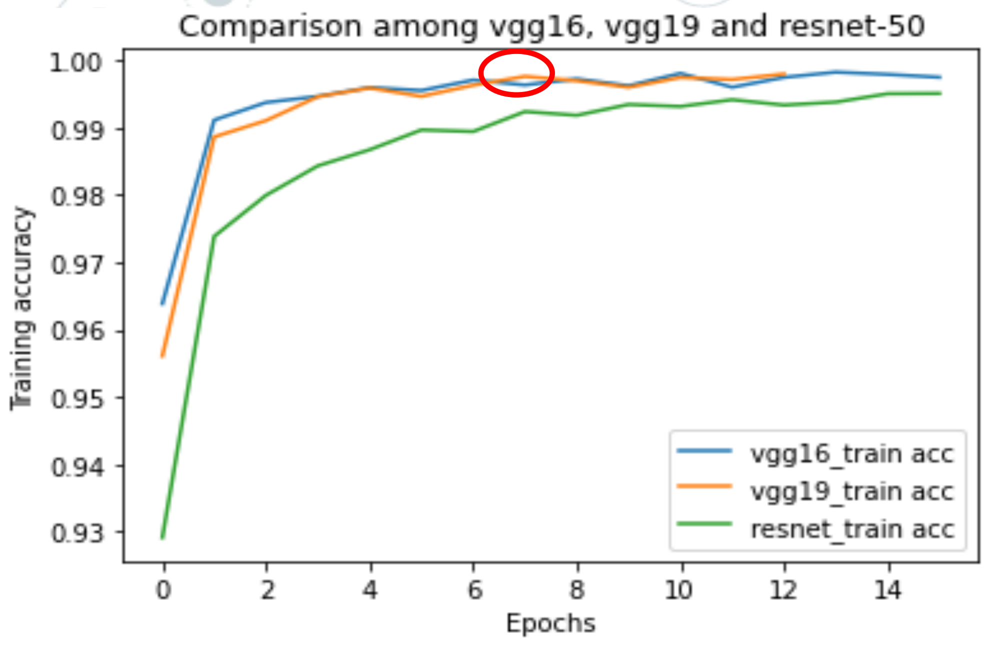
- Devanagari handwritten numeral images dataset preserved by Computer Vision and Pattern Recognition (CVPR) Unit of Indian Statistical Institute (ISI).

Fig 16: Samples of Devanagari handwritten numerals from the CVRP, ISI dataset [9].

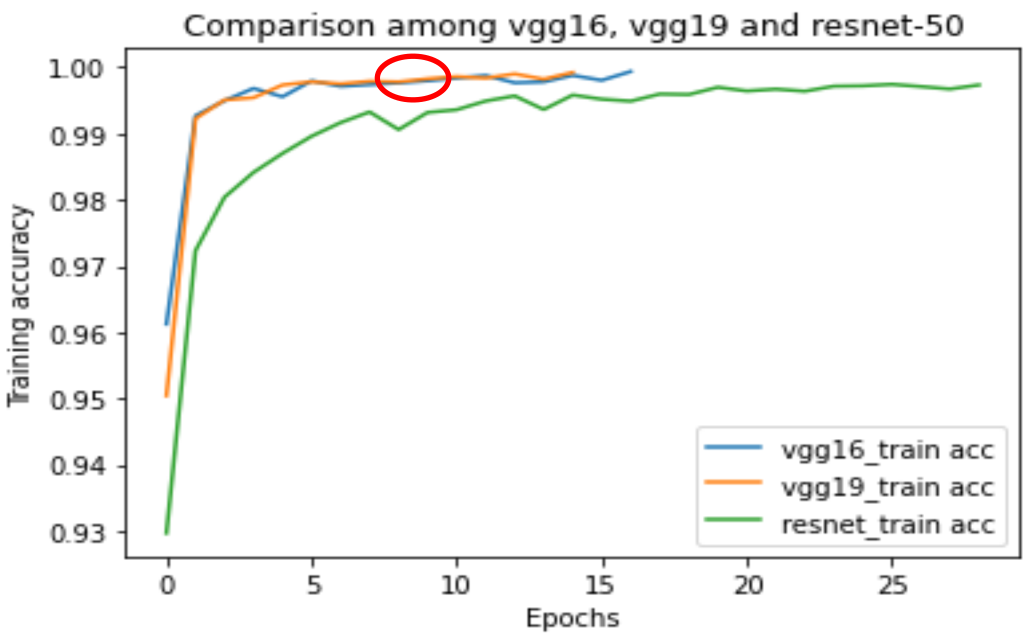
Experimental Studies: Experimental Setup

- The size of the input handwritten numeral images is fixed which is 32x32.
- The considered rotation angles for training data augmentation are ±10° and ± 20°.
- Batch Size (BS) is 32 and 64.
- No of training images: 18000 (=1800×10) and testing images: 3750(=375×10).

Experimental Studies: Experimental Results and Analysis



(a) BS=32

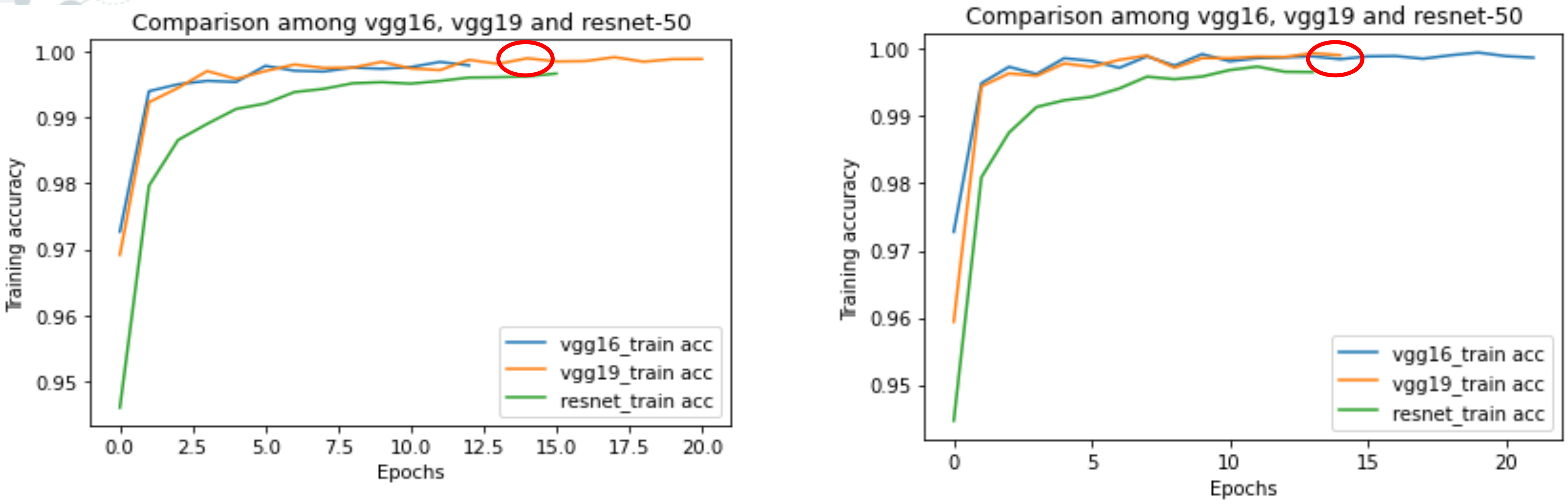


(b) BS=64

Fig 17: Performance of HNRUTL on Bengali CVPR, ISI dataset.

- Pre Trained model weights helps reduce training time as well as achieve high accuracy.

Experimental Results and Analysis (contd.)



(a) BS=32 (b) BS=64

Fig 18: Performance of HNRUTL on Devanagari CVPR, ISI dataset.

- Performance with augmentation for rotation angle of ±10° or ±20° is better than without augmentation.

Experimental Results and Analysis (contd.)

Table 06: Testing accuracies for CVPR Bengali dataset with batch size 64

Model Used	Normal(%)	Normal FT(%)	±10(%)	± 10 FT(%)	± 20(%)	±20 FT(%)
VGG16	98.48	99.18	99.15	99.58	98.78	99.58
VGG19	98.88	99.28	98.90	99.68	98.75	99.58
RESNET50	97.50	98.05	98.08	98.93	98.23	98.55

±20 FT(%)

Table 07: Testing accuracies for CVPR Devanagari dataset with batch size 64

Model Used	Normal(%)	Normal FT(%)	±10(%)	±10 FT(%)	±20(%)	±20 FT(%)
VGG16	99.25	99.49	99.44	99.60	99.23	98.80
VGG19	98.75	99.28	99.09	99.63	99.04	99.52
RESNET50	98.72	98.91	99.07	99.31	98.85	99.31

Experimental Results and Analysis (contd.)

Table 08: Testing accuracies for CVPR Bengali dataset with batch size 32

Model Used	Normal(%)	Normal FT(%)	±10(%)	± 10 FT(%)	± 20(%)	±20 FT(%)
VGG16	98.38	99.18	98.98	99.73	98.88	99.43
VGG19	98.35	99.48	99.08	99.40	98.53	99.48
RESNET50	97.05	98.25	97.78	98.80	97.88	98.70

Table 09: Testing accuracies for CVPR Devanagari dataset with batch size 32

Model Used	Normal(%)	Normal FT(%)	±10(%)	±10 FT(%)	±20(%)	±20 FT(%)
VGG16	99.04	98.77	99.25	99.52	99.15	99.55
VGG19	99.07	99.31	99.33	99.55	99.25	99.60
RESNET50	98.64	99.01	99.01	99.49	98.43	99.31

Experimental Results and Analysis (contd.)

Table 10: Confusion matrix for test samples of handwritten Bengali numeral dataset.

	Number of samples classified as									
Bengali Numeral	০	১	২	৩	৪	৫	৬	৭	৮	৯
০	399	0	0	1	0	0	0	0	0	0
১	0	399	0	0	0	0	0	0	0	1
২	0	0	400	0	0	0	0	0	0	0
৩	0	0	0	400	0	0	0	0	0	0
৪	0	0	0	0	400	0	0	0	0	0
৫	1	1	0	0	2	396	0	0	0	0
৬	0	0	0	0	0	0	400	0	0	0
৭	0	0	0	0	1	0	0	399	0	0
৮	0	0	0	0	0	0	0	0	400	0
৯	0	3	0	1	0	0	0	0	0	396

- 11 misclassified samples.

Experimental Results and Analysis (contd.)

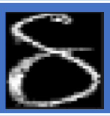
Table 11: Confusion matrix for test samples of handwritten Devanagari numeral dataset.

	Number of samples classified as									
Devanagari Numeral	०	१	२	३	४	५	६	७	८	९
०	375	0	0	0	0	0	0	0	0	0
१	0	370	1	0	0	0	0	2	2	0
२	0	0	375	0	0	0	0	0	0	0
३	0	0	1	373	0	0	0	1	0	0
४	1	0	0	0	374	0	0	0	0	0
५	0	0	0	0	0	375	0	0	0	0
६	0	1	0	0	0	0	372	0	0	2
७	0	0	0	0	0	0	0	375	0	0
८	0	0	1	0	0	0	0	0	374	0
९	0	0	0	0	0	0	1	0	1	373

- 14 misclassified samples.

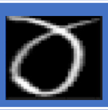
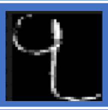
Experimental Results and Analysis (contd.)

Table 12: Misclassified handwritten Bengali numerals by HNRUTL.

Handwritten Bengali numeral image	Predicted value	Actual value
	৩	০
	৪	০
	৯	১
	৬	৩
	৫	৬
	৬	৫

Experimental Results and Analysis (contd.)

Table 13: Misclassified handwritten Devanagari numerals by HNRUTL.

Handwritten Devanagari numeral image	Predicted value	Actual value
	७	९
	०	४
	७	३
	७	९
	८	९
	२	३

Experimental Studies: Comparison of Results

Table 14: Comparison between traditional models and proposed model for the Bengali language

Work Reference and Year	Feature Selection	Method	Dataset; Training and Test Samples	Recognition Accuracy
Pal et al. [6], 2006	Feature based on reservoir and water overflow	Binary decision tree	Self-prepared, 12000	92.80%
Wen et al. [5], 2007	Principal component analysis (PCA) and Kernel PCA	SVM	Bangladesh postal system; 6000 and 10000	95.05%
Wen and He [4], 2012	Eigenvalues and eigenvectors	Kernel and Bayesian discriminant	Bangladesh postal system; 30000 and 15000	96.91%
Nasir and Uddin [17], 2013	Bayes' theorem, K-means clustering and Maximum Posteriori	SVM	Self-prepared, 300	96.80%
Akhand et al. [7], 2016	No	CNN	CVPR, ISI [9]; 18000 and 4000	98.45%
Mahtab et al. [18], 2016	No	SAE	CVPR, ISI [9]; 18000 and 4000	96.30%
Akhand et al. [19], 2018	No	CNN	CVPR, ISI [9]; 18000 and 4000	98.98%
Shuvo et al. [3], 2019	No	CAE with CNN	CVPR, ISI [9]; 18000 and 4000	99.68%
Proposed Method	No	TL with VGG-16 using FT, BS 32	10 CVPR, ISI [9]; 54000 and 4000	99.73%

Experimental Studies: Comparison of Results

Table 15: Comparison between traditional models and proposed model for the Devanagari language

Work Reference and Year	Method	Dataset, Training and Test Samples	Recognition Accuracy
Arya et al. [16], 2015	SVM and KNN	CVPR, ISI [9]; 19798 and 3763	98.06 %
Singh et al. [15], 2015	Ensemble of NNs	CVPR, ISI [9]; 16794 and 3762	99.37 %
Singh et al. [14], 2016	MLP	CMATERdb 3.2.1 [16]; 2000 and 1000	98.92 %
Singh et al. [10], 2017	MLP	CMATERdb 3.2.1 [16]; 2000 and 1000	97.50 %
Akhand et al. [7], 2018	CNN with data augmentation	CVPR, ISI [9]; 18000 and 3763	98.96 %
Trivedi et al. [13], 2018	CNN with GA	CVPR, ISI [9]; 18784 and 3762	96.41%
Jiang et al. [12], 2020	Edge-TripleNet	CMATERdb [16]; 2400 and 600	97.33%
Gupta et al. [11],2021	Transfer Learning (MNIST Pre-trained)	CVPR, ISI [9]; 15789 and 2256	99.04%
Proposed Method	TL with VGG-19 using FT, BS 64	10 CVPR, ISI [9]; 54000 and 3750	99.63 %

Conclusions

- Challenges tell us that normal dataset is not enough to handle such a convoluted task.

- Rotational augmentation is introduced to enrich the dataset with some rotated images as if it had some distorted images.

- The model that used fine tuning and $\pm 10°$ rotational augmentation gives better result than others.

- Since we have pre trained models, we do not need to train all the layers. That is why transfer learning is introduced to reduce the computation time, training time.

Future Works

- Proposed HNR can be ameliorated to an extent for other languages also.
- Other CNN architectures or adjusting the working CNN architectures may work fine.
- Any image size other than 32x32 may improve the accuracy.
- Necessary modifications out of all possible combinations may bring a satisfactory outcome.

Important References

1. A. K. Tushar, A. Ashiquzzaman, A. Afrin and M. R. Islam, “A novel transfer learning approach upon hindi, arabic and bangla numerals using convolutional neural networks,” in Computational Vision an Bio Inspired Computing. Springer, 2018, pp. 9972-981

2. Sakib Reza, Ohinda Binte Amin and M. M. A. Hashem “Basic to Compound: A Novel Transfer Learning Approach for Bengali Handwritten Character Recognition” in 2nd International Conference on Bangla Speech and Language Processing (ICBSLP), September 2019

3. M. I. R. Shuvo, M. A. H. Akhand, and N. Siddique, “Handwritten Numeral Superposition to Printed Form Using Convolutional Auto-encoder and Recognition Using Convolutional Neural Network”, International Joint Conference on Computational Intelligence (IJCCI), Dhaka, Bangladesh, 2019.

4. Y. Wen and L. He, “A classifier for Bangla handwritten numeral recognition,” Expert Syst. Appl., vol. 39, no. 1, pp. 948–953, Jan. 2012.

5. Y. Wen, Y. Lu, and P. Shi, “Handwritten Bangla numeral recognition system and its application to postal automation,” Pattern Recognit., vol. 40, no. 1, pp. 99–107, Jan. 2007.
6. U. Pal, B. B. Chaudhuri, and A. Belaid, “A Complete System for Bangla Handwritten Numeral Recognition,” IETE J. Res., vol. 52, no. 1, pp. 27–34, Jan. 2006.
7. M. M. H. Akhand, M. A. H.; Ahmed, Mahtab; Rahman, “Convolutional Neural Network based Handwritten Bengali and Bengali-English Mixed Numeral Recognition,” Int. J. Image, Graph. Signal Process., vol. 8, no. 9, pp. 40–50, 2016.

8. Y. C. C. C. J. C. B. LeCun, “THE MNIST DATABASE of handwritten digits”

9. B. . Bhattacharya, U.; Chaudhuri, “Offline Handwritten Bangla and Devanagari Numeral Databases. Kolkata: Computer Vision and Pattern Recognition Unit, Indian Statistical Institute.” 2009

10. M. A. H. Akhand, Deep Learning Fundamentals - A Practical Approach to Understanding Deep Learning Methods. Dhaka: University Grants Commission of Bangladesh, 2021.

Important References

11. D. Gupta and S. Bag, "CNN-based multilingual handwritten numeral recognition: A fusion-free approach," Expert Syst. 26 Appl., vol. 165, p. 113784, Mar. 2021, doi: 10.1016/j.eswa.2020.113784.

12. W. JIANG and L. ZHANG, "Edge-SiamNet and Edge-TripleNet: New Deep Learning Models for Handwritten Numeral Recognition," IEICE Trans. Inf. Syst., vol. E103.D, no. 3, pp. 720–723, Mar. 2020, doi: 10.1587/transinf.2019EDL8199

13. A. Trivedi, S. Srivastava, A. Mishra, A. Shukla, and R. Tiwari, "Hybrid evolutionary approach for Devanagari handwritten numeral recognition using Convolutional Neural Network," Procedia Comput. Sci., vol. 125, pp. 525–532, 2018, doi: 10.1016/j.procs.2017.12.068.

14. P. K. Singh, R. Sarkar, and M. Nasipuri, "A Study of Moment Based Features on Handwritten Digit Recognition," Appl. Comput. Intell. Soft Comput., vol. 2016, pp. 1–17, 2016, doi: 10.1155/2016/2796863.

15. P. SINGH, A. VERMA, and N. S. CHAUDHARI, "Feature selection based classifier combination approach for handwritten Devanagari numeral recognition," Sadhana, vol. 40, no. 6, pp. 1701–1714, Sep. 2015, doi: 10.1007/s12046-015-0419-x.

16. S. Arya, I. Chhabra, and G. S. Lehal, "Recognition of Devnagari Numerals using Gabor Filter," Indian J. Sci. Technol., vol. 8, no. 27, pp. 1–6, Oct. 2015, doi: 10.17485/ijst/2015/v8i27/81856.

17. M. K. Nasir, "Hand Written Bangla Numerals Recognition for Automated Postal System," IOSR J. Comput. Eng., vol. 8, no. 6, pp. 43–48, 2013.

18. M. Ahmed, A. K. Paul, and M. A. H. Akhand, "Stacked auto encoder training incorporating printed text data for handwritten bangla numeral recognition," in 2016 19th International Conference on Computer and Information Technology (ICCIT), 016, pp. 437–442

19. M. A. H. Akhand, M. Ahmed, M. M. H. Rahman, and M. M. Islam, "Convolutional Neural Network Training incorporating Rotation-Based Generated Patterns and Handwritten Numeral Recognition of Major Indian Scripts," IETE J. Res., vol. 64, no. 2, pp. 176–194, Mar. 2018

Thank You