NCE Journal of Science and Engineering (NJSE), Vol. 2, June 2025

Offline Rañjanā Lipi Handwritten Character Recognition Using CNN Model

Dinesh Maharjan[1] and Balkrishna Bal[2]
[1]Pokhara University, Kathmandu, Nepal
[2]Nepal College of Information Technology, Lalitpur, Nepal
*Correspondence: dinesh.maharjan@gmail.com and niranjan@ncit.edu.np
Manuscript received January 8, 2025; accepted April 20, 2025

***Abstract*—This research is focus on "Offline Rañjanā Lipi Handwritten Character Recognition (ORLHCR)" using a (CNN) Convolutional Neural Network and the development of Rañjanā Lipi Dataset (RLD). About 40 students voluntarily contribute to develop 6792 isolated character images of 62 classes. The developed dataset does not contain any data augmented data. Based on this RLD, the data's are divided into 75% Training, 25% Test set, and from Test set it is further split into 50-50 % i.e. 50% test and 50% validation set. Vgg19, InceptionV3, and ResNet50 CNN models are implemented to acquire the accuracy of ORLHCR. The test results delivers an accuracy rate of 96.82% by RestNet50 from RLD by tuning learning parameters only.**

***Keywords*—Convolutional Neural Network (CNN), handwritten character recognition, deep learning, Optical Character Recognition (OCR), hyper parameter - learning**

I. INTRODUCTION

HANDWRITTEN character recognition is one of the hot and interesting topic nowadays. Many researcher have worked on character recognition in Devanagari script but not in the field of Rañjanā Lipi/script. No digitization of this script found on the digital platform. The making of dataset of Rañjanā Lipi is a quite challenging job. The Rañjanā Lipi contain 16 vowels, 36 consonant and 0-9 digits. Compared to other scripts, the writing style of Rañjanā Lipi has a thick lining and beauty contains of this. The thick line of these scripts is to mark the paper for a longer duration. Rañjanā Lipi was used by inhabitant's people of Nepa Valley from Nepal to Tibet. The Lipi was developed during the 11th century which is Abugida's writing style. Many scripts were developed during middle ages from Nepal Bhasa such as Brahmi script, Prachalit script, Rañjanā Lipi/script, Bhujinmola script, Kunmol script, Kwenmol script, Golmol script, Pachumol script, Hinmol script, and Litumol script. The base of brahmi script was used to create other scripts.The Brahmi script was used to write Sanskrit or Pali.The inhabitants of Kathmandu valley are in contact with this script and they start to develop their own script known as Prachalit. As time flew away, the Newa people commence designing many beautiful calligraphies. Prachalit, Bhujinmola, and Rañjanā Lipi were broadly used in the middle ages.

Normally, handwritten character recognition has two approaches i.e. Offline Handwritten Recognition [3-4] and Online Handwritten recognition [1-2]. The information from the input scanned images are used in Offline Handwritten recognition. In contrast, online handwritten recognition utilizes the movement of the pen as its input image data. Every person has their own writing styles and even some people write so fast that makes the character so complicated and concise which makes it difficult for character recognition.

In this research, VGG19, ResNet50 and InceptionV3 are implemented to acquire the accuracy rate of ORLHCR with primary Rañjanā Lipi Dataset (RLD) which was collected with the help of Nepal Lipi Guthi teams and other calligraphy enthusiastic. About 6790 precise dataset of isolated character images of 62 classes were developed.

B. Problem statement

Many types of research have been done on recognizing Devanagari script but in the case of Rañjanā Lipi, only limited research found in the case of OCR or digitization of many religious documents, manuscripts, which are in fragile condition and are at risk of deterioration.

Till now, no any dataset were found in digital platform and if someone developed it already, which are not available publicly and handwritten character recognition is not done. So, this research paper address on the development of Rañjanā Lipi Dataset (RLD), and Offline Rañjanā Lipi Handwritten Character Recognition (ORLHCR) with precise result which can aid in digitization and preservation process.

C. Objectives of the study

The main objectives of this research paper is to develop a novel dataset for Rañjanā Lipi and to recognize Rañjanā Lipi Handwritten character using VGG19. ResNet50, and InceptionV3 and compare the accuracy results.

D. Significance of the study

The beneficial of this research paper is to aid for the recognition of Offline Rañjanā Lipi handwritten character with high ccuracy and also helps other to further research in compound character of Rañjanā Lipi.

E. Limitation

This research paper is only limited to recognize only isolated characters and does not address to compound letter recognition, consonant with identifier, and dynamic segmentation of data.

II. Review of Related Literature

In research, entitled Assamese Character Recognition using (CNN) Convolutional Neural Network, two datasets have been used, which contain a total of 12,863 images. The images consist of 52 Assamese characters, which were collected from a group of more than 200 individuals including students, faulty, and staff from IIT Guwahati in India. The datasets were split into training (70%), validation (20%), and testing (10%) sets. Several architectures such as DenseNet201, ResNet50, LeNet 5, and InceptionV3 were trained and compared. The highest accuracy of 94% was achieved using DenseNet201, in compared to other CNN algorithms [11].

Similarly, in research, entitled "Vietnamese Handwritten Character Recognition using Convolution Neural Network", different classification and algorithms are in use. SVM classification with 39,800 samples achieved an accuracy of 91.3%, while a CNN with 3 convolutional layers and 2 fully connected layers had a 97.2% accuracy rate [12].

In research, named "Manipuri Handwritten Character Recognition by Convolutional Neural Network", 90 different people of age groups and education were involved in the creation of a handwritten dataset. About 4900 sample image datasets are used. Various recognition methods are in use and their accuracy rate varies such as fuzzy features and probabilistic with Artificial Neural Network (ANN) shows 90.3% accuracy rate, binary pattern as vector and NN with back-propagation shows 80%, Binary pattern of pixel density using NN shows 85%, Gabor filter using with SVM show 89.48% accuracy rate, Lenet-5 shows 96.02% accuracy rate, and CNN model shows 98.86% accuracy rate. Among them, the CNN model with 98.86% highest accuracy rate is shown [13].

In a Journal named "Handwritten Devanagari (Marathi) Compound Character Recognition using Seventh Central Moment", the samples with 'Shirorekha', without 'Shirorekha', and the combination of different characters' data is used and the average accuracy rate is 93.87% of some compound characters. The recognition of composite characters was performed using a combination of 7 Invariant Moments and 7th order Central Moments, and a classifier based on SVM was employed. The outcome accuracy rate is 93.87% [16].

In the research paper of "Rañjanā Script Handwritten Character Recognition using CNN", about 17,360 data has been collected from 150 people and later 173,600 data's are generated with data augmentation. The datasets are divided into Training, Testing, and Validation set of 60:20:20 ration and tested using LeNet, AlexNET, ZFNET and a proposed CNN model. The test accuracy rate of 99.73% result achieved by proposed model with 64x64 pixels resolution [25].

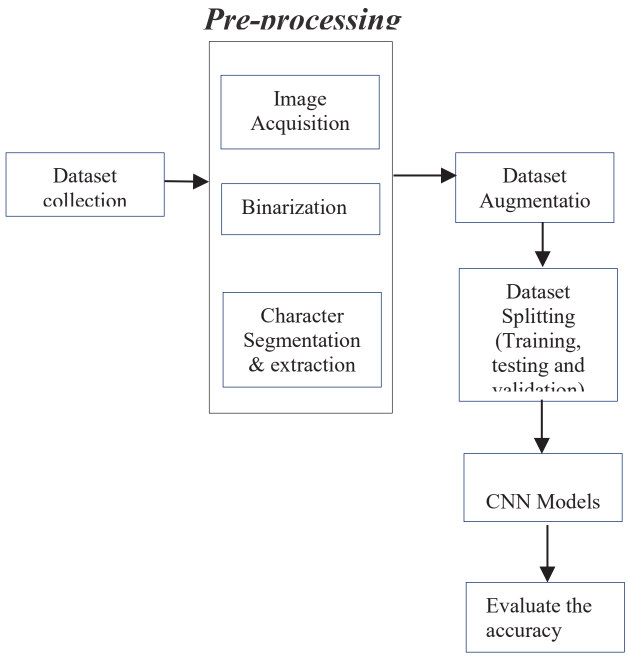

Fig. 1. Methodology

II. Methodology

Rañjanā Lipi dataset needs to develop from the scratch. The dataset is collected with the help of Nepal Lipi Guthi team, calligraphy enthusiasts, learners, and by capturing the image from different places where Rañjanā Lipi is used. This research methodology consists of six parts which are dataset collection or development, pre-processing, dataset augmentation, dataset splitting, fit to CNN models, and evaluate the accuracy rate.

A. Data collection

It's the initial step with lots of hard work that need to be done in the case of Rañjanā Lipi. The isolated handwritten Rañjanā Lipi character will be collected by capturing the images where this script is used. Similarly, can collaborate with the team of Nepal Lipi Guthi, teachers, students, artists, designers, and interested people. Every character data should be of 52-60 different types which have different unique styles. The total number of isolated characters is 62 classes which contains vowels, consonants, and digits. So, about 40 volunteers involve to create a 6792 dataset of 67 x78 pixels.

The filled dataset collected from the volunteers are:
The 62 isolated characters were collected from filled A4 sheet paper.

B. Data pre-processing

Data pre-processing is a significant steps. This steps involve noise filtering techniques. The pixels in the image are classified as either background (colour white) or character (colour black). However, some writers may have mistakenly added extra, unwanted characters or dots in the image, which need to be

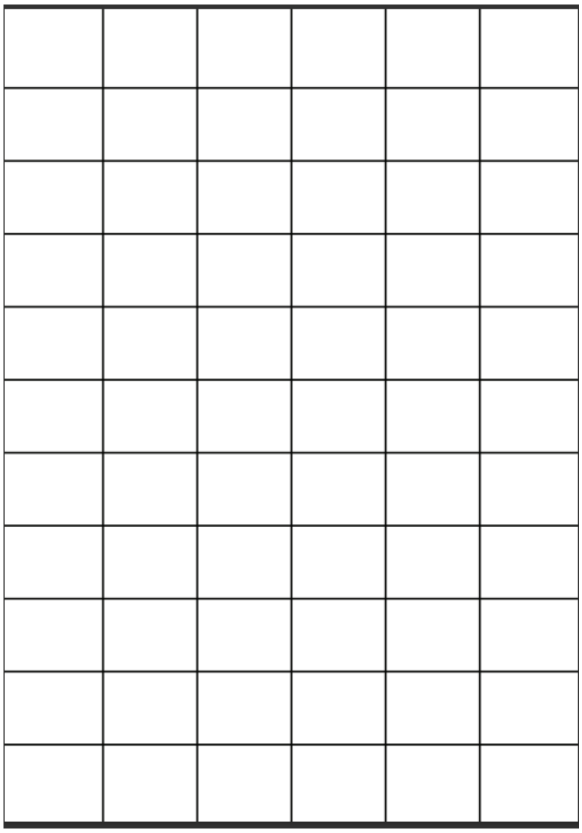

Fig. 2. Sample form to collect the dataset from an A4 sheet.

Fig. 3. Voluntarily collected dataset.

removed through the use of outlier elimination. After this step, image smoothing is performed using the average filter method [11].

In the case of Rañjanā Lipi, the input image is transformed into grayscale and added the Gaussian blur to the converted gray scaled image. A certain threshold is chosen to choose the portion of an image that are relevant to it while ignoring the rest. The threshold value and techniques like THRESH_BINARY_INV AND THRESH_OTSU were used. The morphological transformations required the initial image and the structuring element which decides what kind of action to be performed. The kernel size for MORPH_RECT is 44 x 44 and performs the Dilation, erosion, and open morphological operations. At last, the red rectangle data is segmented or extracted and saved to the segmented datasets directory.

Note: The above explain process is implemented for few dataset to extract the dataset automatically but other dataset are segmented with the help of CamScanner to scan images and convert image to B&W and Figma is used to segment the dataset.

C. Data augmentation

Data augmentation involves using various transformation methods to produce additional data and reduce high variability, leading to the development of improved models [18]. Deep

learning algorithms typically perform best when trained on a large amount of data, but if the available data is limited, the recognition results can be poor. To address this issue, data augmentation techniques are employed to increase the amount of training data and improve recognition accuracy.

Similarly, data augmentation is also used to avoid data over fitting. To generate artificial data into our dataset, rotate a certain degree an image, de-centre it or zoom in or zoom out tiny, and different techniques are used likes vertical/horizontal flips, rescaling, random crops, adjust brightness, rotations, translations, width shift, height shift and many more.

D. Data splitting

Data division is very significant steps and complex part. Destitute preparing and testing of data sets can lead to eccentric impacts on the output of the model which may lead to over-fitting or under-fitting of the data and our model may conclusion up giving one-sided comes about. Normally data are split into 3 sets i.e. train, test, and validation set.

Same as, in this research, data sets are split into 75% train set, 25% test set and from 25% test set data are further divided equally into validation set and test set from test set with random state. To partition data sets into train and test, sklearn model is used. The total number of dataset used in this research is 6,792.

E. Classification and Fit model

In this research, 62 classes (10 digits from 0-9, 16 vowels, and 36 consonants) are classified and the class labels are categorical. Classification is achieved by employing a fully connected layer, where probabilities for each class are obtained from the input data. The input is then assigned to the appropriate class by selecting the class with the highest probability. To predict the model VGG19, ResNet50, and InceptionV3 model is used. Every sample from the training, testing, and validation set is processed to produce a good fit model. The result is the accuracy rate from VGG19, InceptionV3, and ResNet50.

F. Model Evaluation and Verification

In order to determine if a model is producing a good results, it is significant to assess its performance using metrics designed to evaluate the effectiveness of the model. In this research, the

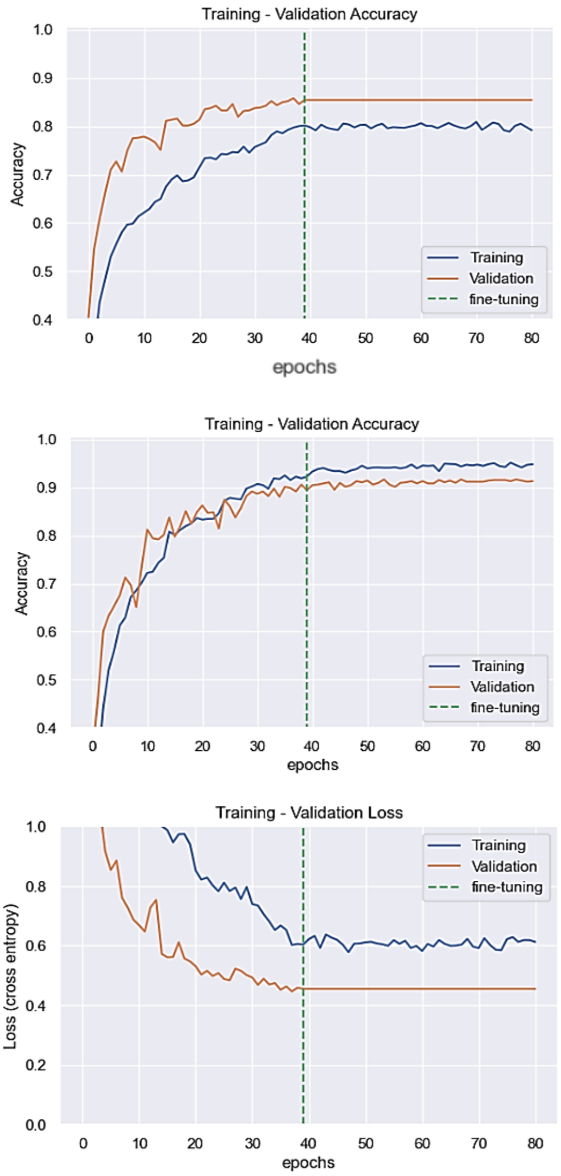

Fig. 4. ResNet50 validation accuracy and loss without tuning learning parameter.

Hold-out method is used where datasets are divided into 75/25 splitting of training, testing, and validation sets and to appraise the model results average precision, recall and F1 score across multiple classes.

III. Results and Discussion

To train the model initially Kaggle is used but due to the limitation of jupyter notebook, later the model was trained in personal PC.

The data are loaded from the local dataset directory with shuffle. The loaded data files are segregate into Target label with their classes. The class are labelled as numerically as [49 42 49 ... 13 37 11] and class label as:

['a', 'aa', 'ah', 'ai', 'am', 'au', 'ba', 'bha', 'ca', 'cha', 'da', 'dda', 'ddha',

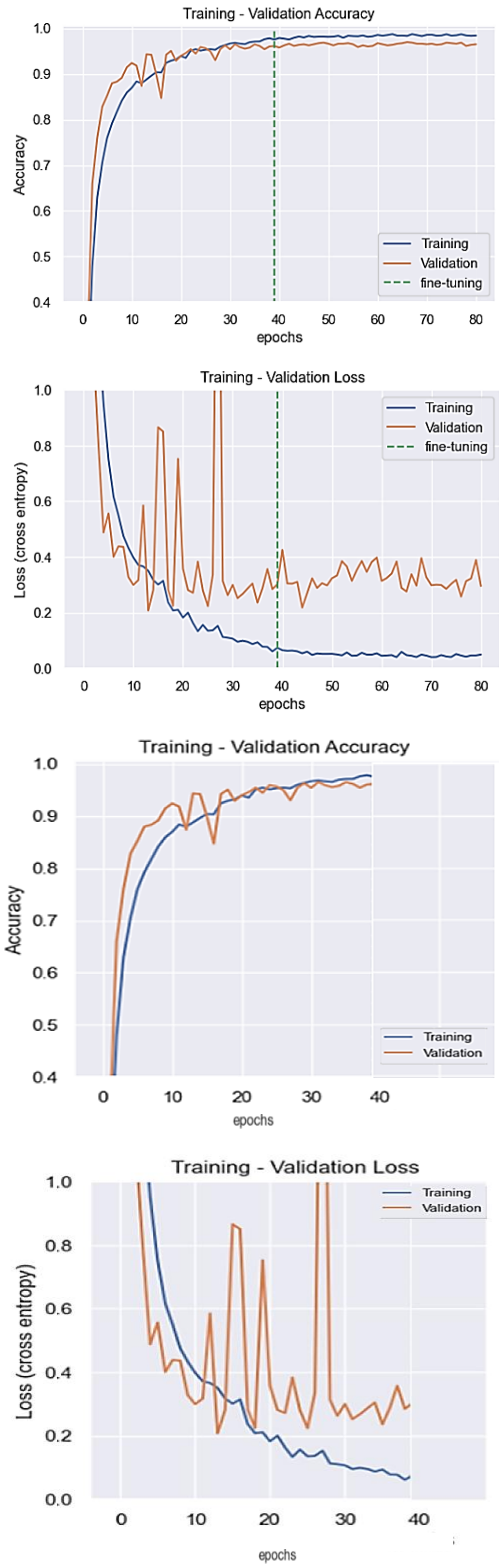

Fig. 5. ResNet50 validation accuracy and loss without tuning learning parameters.

'dha', 'e', 'eight', 'five', 'four', 'ga', 'gha', 'gyan', 'ha', 'i', 'ii', 'ja', 'jha', 'ka', 'kha', 'ksa', 'la', 'lu', 'luu', 'ma', 'na', 'nine,' 'nna', 'nnna', 'nya', 'o',

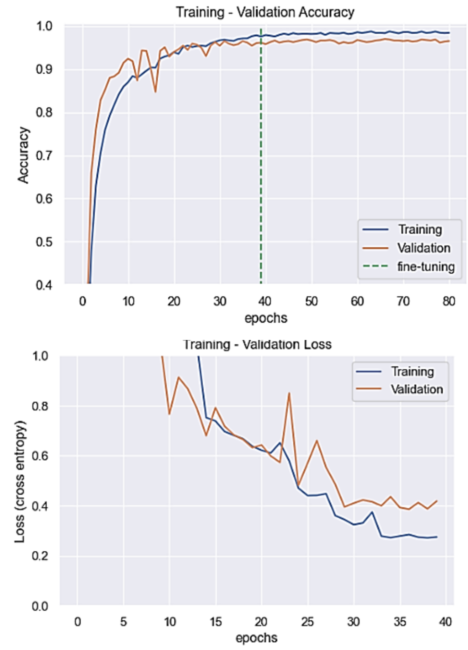

Fig. 6. ResNet50 validation accuracy and loss without tuning learning parameters

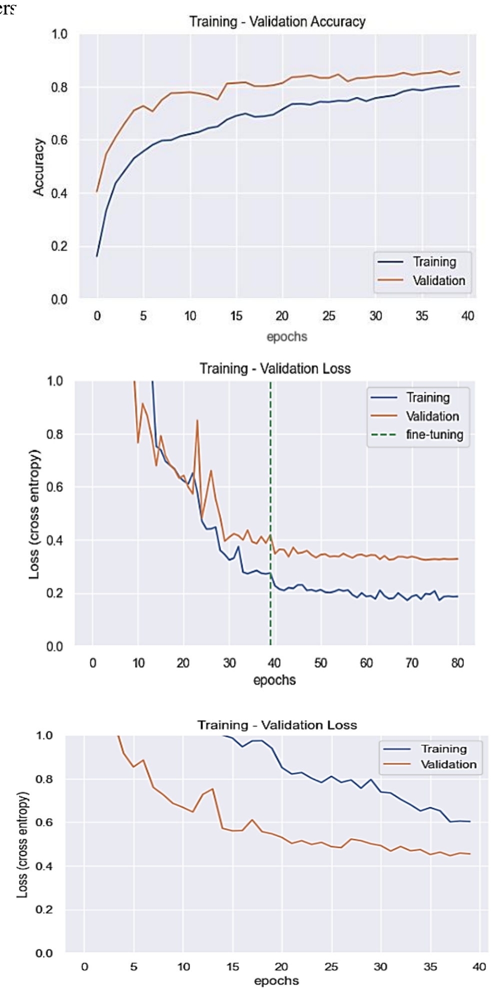

Fig. 7. InceptionV3 training-validation accuracy and loss with and without tuning learning parameters.

TABLE I: VALIDATION ACCURACY AND VALIDATION LOSS OF DIFFERENT ARCHITECTURES

Architecture	Batch Size	Epochs	Accuracy (%)	Loss
ResNet50	32	80	96.82	0.16
VGG19	32	80	92.58	0.29
Inception v3	32	80	85.39	0.43

TABLE II: RESNET50 METRICS WITH ACCURACY, MACRO AVERAGE AND WEIGHTED AVERAGE

Metrics	Precision	Recall	F1 - Score	Support
accuracy			0.97	849
macro avg	0.97	0.96	0.96	849
weighted avg	0.97	0.97	0.97	849

TABLE III: VGG19 METRICS WITH ACCURACY, MACRO AVERAGE AND WEIGHTED AVERAGE

Metrics	Precision	Recall	F1 - Score	Support
accuracy			0.93	849
macro avg	0.91	0.90	0.90	849
weighted avg	0.94	0.93	0.93	849

'one', 'pa', 'pha', 'ra', 'ri', 'rii', 'sa', 'saa', 'seven', 'sha', 'six', 'ta', 'tha', 'three', 'tra', 'tta', 'ttha', 'two', 'u', 'uu', 'wo', 'ya', 'zero']

As we have 62 class which includes vowel, consonants and digits. Every class have their class or labels which is associate with unique value which is also known as character number. The total number of loaded data are 6792. Later the loaded data are further converted into numpy arrays and also convert the image size into 224 x224 images sizes. As if the loaded images contains any varies size then it is convert the images into specific sizes of numpy array which helps to normalize the variability.

The experimental result of Vgg19, ResNet50, and InceptionV3 with 40 epochs without hyper parameter tuning and with tuning learning parameter 80 epochs with accuracy and loss of Training-Validation are calculated. The below graph shows the relationship between the training and validation with Accuracy and Loss.

From figure 4, 5, 6, and 7, by tuning the hyper parameter learning rate from 0.001 to 0.001, the performance of the model is improved in performance on ResNet50, VGG19, and InceptinV3. The tune epochs is the total epochs from the summation of previous epochs and tune hyper parameter learning epochs. The vertical dotted green line represents with and without tuning learning hyper parameter.

By fine-tuning the learning parameters, the deep learning models, ResNet50, VGG19, and Inception v3 trained on a specific Rañjanā Lipi Dataset with 80 epochs and a batch size of 32. An epoch is one complete iteration through the dataset, and the batch size refers to the number of samples processed and pass through the network. When the data is tested from Rañjanā Lipi Dataset, 96.82% accuracy rate is achieved by using ResNet50. Similarly, 92.58 % by using VGG19, and 85.39% by using Inception v3. These accuracy rates are specific to the Rañjanā Lipi Dataset, training conditions, and evaluation criteria and may note generalize to other datasets or use cases.

Fig. 8. Prediction of Rañjanā Lipi image by using ResNet50

TABLE IV: INCEPTIONV3 METRICS WITH ACCURACY, MACRO AVERAGE AND WEIGHTED AVERAGE

Metrics	Precision	Recall	F1 - Score	Support
accuracy			0.85	849
macro avg	0.84	0.84	0.83	849
weighted avg	0.87	0.85	0.85	849

The Rañjanā Lipi dataset of every class does not contains equal set of data. So, the weighted average metrics is used. Table II-IV show the difference in architecture performance metrics. The combined accuracy is calculated from the custom Rañjanā Lipi Dataset. The results with accuracy and loss with tuning the learning parameters are shown in table below:

Tables II, III, and IV show the performance metrics commonly used to evaluate the classification tasks of ResetNet50, VGG19, and InceptionV3. The accuracy metric is used to measure the percentage of correct predictions made by a model. Similarly, Macro average is a metric that is used to calculate the average of metric scores for each class, treating each class as equally important, and weighted average is a metric

In Table II, the accuracy rate for ResNet50 show the 0.97 which means that the model has a 97% accuracy in its positive class predictions, recall of 0.96 means that the model is able to detect 96% of the positive class samples, and F1 score of 0.96 means that the model has a good balance between precision and recall, with a score close to 1.

Similarly, in Table III, the accuracy rate for VGG19 show the 0.93 which means that the model has a 93% accuracy in its positive class predictions, recall of 0.90 means that the model is able to detect 90% of the positive class samples, and F1 score of 0.83 means that the model has a good balance between precision and recall, with a score close to 1, and in Table IV, the accuracy rate for InceptionV3 show the 0.85 which means that the model has a 85% accuracy in its positive class predictions, recall of 0.84 means that the model is able to detect 84% of the positive class samples, and F1 score of 0.83 means that the has not a good balance between precision and recall for and its score is not above 0.9 and close to 1. At last, the model has successfully recognized the text from the images. The sample recognized data from the sample data and mobile application are shown in Fig. 9.

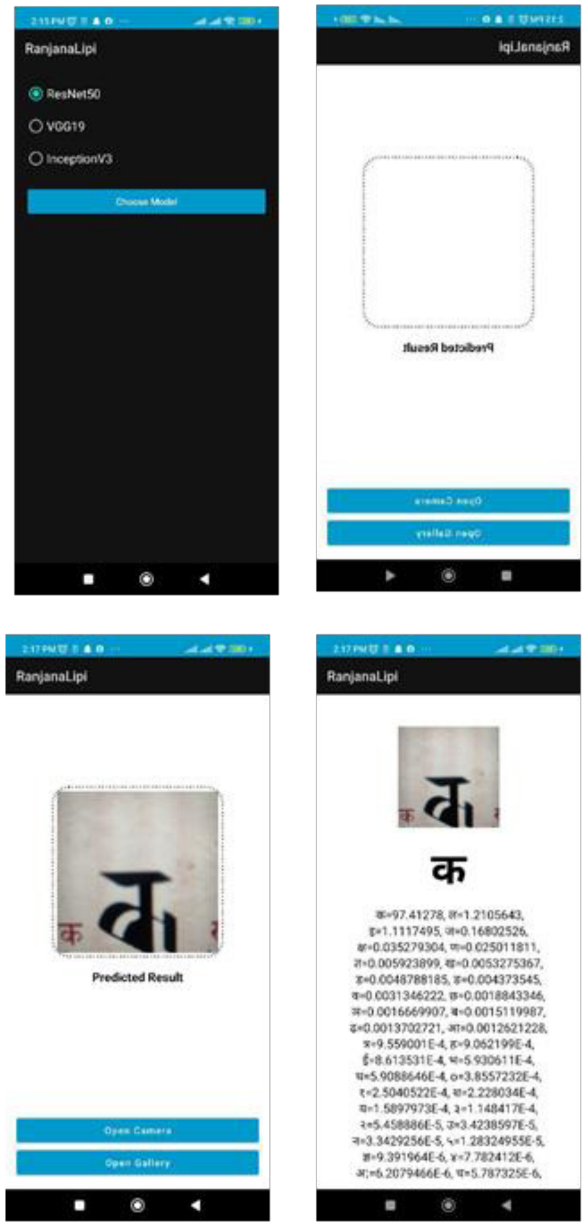

Fig. 9. Prediction with the ResNet50 model using a custom-made Android application.

The trained model is further converted into.tflite to integrate the model in android application and to implement OCR of Rañjanā Lipi Character offline.

V. CONCLUSION

In this research, the best performing model is ResNet 50 which achieved the accuracy rate of 96.82% by improving the learning parameters. Comparing Precision, recall, and F1 score of ResNet50, VGG19, and InceptionV3, and find that ResNet50 has scores closer to 1 than VGG19 and InceptionV3, this suggest that ResNet50 is a more accurate model than VGG19 and InceptionV3. One important note that the performance of a model can vary depending on the specific dataset and task being evaluated.

REFERENCES

[1] S. España-Boquera, M.J. Castro-Bleda, J. Gorbe-Moya, F. Zamora-Martinez, "Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 4, pp. 767 - 779, 2011.

[2] Yanwei Wang; Xiaoqing Ding; Changsong Liu, "Topic Language Model Adaption for Recognition of Homologous Offline Handwritten Chinese Text Image", IEEE Signal Processing Letters, vol. 21, no.5, pp. 550 - 553, 2014.

[3] Daniel Keysers, Thomas Deselaers, Henry A. Rowley, Li-Lun Wang, Victor Carbune, "Multi-Language Online Handwriting Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1180 - 1194, 2017.

[4] Zecheng Xie, Zenghui Sun, Lianwen Jin, Hao Ni, Terry Lyons, "Learning Spatial-Semantic Context with Fully Convolutional Recurrent Network for Online Handwritten Chinese Text Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 8, pp. 1903 - 1917, 2018.

[5] Cheng-Lin Liu, Fei Yin, Qiu-Feng Wang, Da-Han Wang, "ICDAR 2011 Chinese Handwriting Recognition Competition", 2011 International Conference on Document Analysis and Recognition, 2011.

[6] Michael Murdock, Jack Reese, Shawn Reid, "ICFHR2016 Competition on Local Attribute Detection for Handwriting Recognition", 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2016

[7] Jangid Mahesh and Sumit Srivastava, "Handwritten Devanagari Character Recognition Using Layer-Wise Training of Deep Convolutional Neural Network," Journal of Imaging, February 2018.

[8] H. Singh, R.K. Sharma and V.P Singh," Online handwriting recognition systems for Indic and non-Indic scripts: a review" Artif Intell Rev 54, 1525–1579, 2021.

[9] Manoj Sonkusare, Gupta Roopam, and Asmita Moghe, "Palm-Leaf Manuscript Character Recognition and Classification Using Convolutional Neural Networks," January 2019.

[10] Adith Narayan, Raja Muthalagu, "Image Character Recognition using Convolutional Neural Networks", June 2021

[11] Mihir Yadav, Divyansh Mangal, Srinivasan Natesan, Marcin Paprzycki, and maria Ganzha, "Assamese Character Recognition using Convolutional Neural Networks", 2021

[12] Truong Quang Vinh, Le Hoai Duy, Nguyen Thanh Nhan, "Vietnamese handwritten character recognition using convolutional neural network", April 2020

[13] Sanasam Inunganbi, Prakash Choudhary, and Khumanthem Manglem, "Manipuri Handwritten Character Recognition by Convolutional Neural Network", pp, 307-318, 2020

[14] Saniya Ansari, Bhavani S, and Udaysingh Sutar, "A Novel approach towards Online Devanagari Handwrtten word recognition based on robust feature extraction method and FFNN classifier," International Journal of Advanced Research(IJAR), vol. 5, no. 8, pp. 764-776, 2017.

[15] Mohammed N AlJarrah, Mo'ath M Zyout, Rehab Duwairi, "Arabic Handwritten Characters Recognition Using Convolutional Neural Network", 2021

[16] P. E. Ajmire, R. V. Dharaskar, and V. M. Thakare, "Handwritten Devanagari (Marathi) Compound Character Recognition using Seventh Central Moment," International Journal of Innovative Research in Computer and Communication Engineering, vol. 3, no. 6, June 2015.

[17] Kaggle, Competition Notebook, Kaggle, https://www.kaggle.com/cdeotte/25-million-images-0-99757-mnist

[18] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," J. Big Data, vol. 6, no. 1, pp. 1– 48, Dec. 2019, doi: 10.1186/s40537-019-0197-0

[19] Mohammed N AJarrah, Mo'ath M Zyout, Rehab Duwairi, "Arabic Handwritten characters recognition Using Convolutional Neural Network", 978-1-6654-3351-8/21/$31.00 ©2021 IEEE

[20] Kavitha B.R, Srimathi C, "Benchmarking on offline Handwritten Tamil Character Recognition using convolutional neural networks", 7 June 2019

[21] Hassan El Bahi, Zouhir Mahani and Abdelkarim Zatni, "An offline handwritten character recognition system for image obtained by camera phone", 15 October 2015

[22] I Khandokar, Md M Hasan, F Ernawan, Md S Islam, M N Kabir, "Handwritten character recognition using convolutional neural network", 2021

[23] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J and Chen T 2018 Patt. Recognition 77 354

[24] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998; DOI: 10.1109/5.726791,

[25] Jen Bati, "Rañjanā Script Handwritten Character Recognition Using CNN",2022