11th International Advanced Computing Conference on 18th & 19th December, 2021 at University of Malta, Malta

Improving Recognition of Handwritten Kannada Characters using Mixup Regularization

Chandravva Hebbi, Anirudh Maiya and Mamatha H. R.

Department of Computer Science & Engineering, PES University, Bengaluru, India

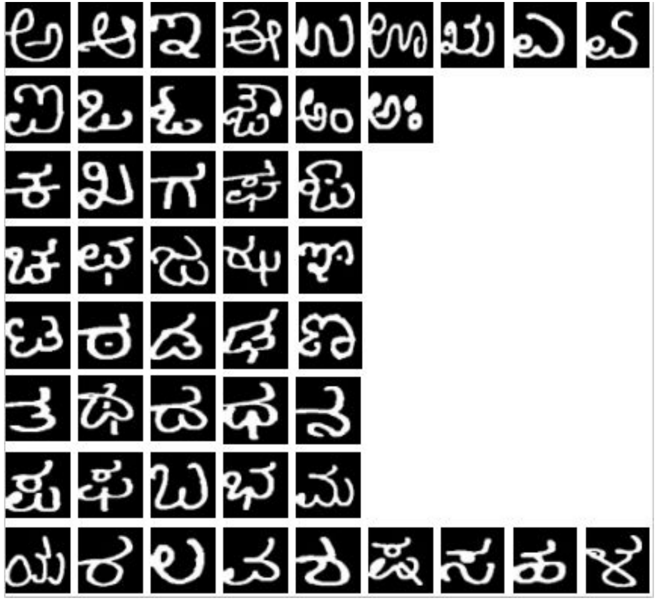
Presented by
Anirudh Maiya

1. ABSTRACT:

- Recognition of handwritten characters is essential for digitization of degraded/old documents
- A novel dataset called Kannada84 is created to tackle the problem of recognizing handwritten Kannada characters
- Dataset represents individuals from all walks of life
- State-of-the-art convolutional neural networks such as VGG Net, ResNet and Squeeze-and-Excitation Network are employed to solve the task of recognition
- Additionally, mixup regularization is used to reduce generalization error and boost performance
- Quantitative results such as top-1 and top-3 test accuracies are reported

2. INTRODUCTION:

- Handwritten characters are inconsistent unlike their printed/hand-drawn counterparts in terms of thickness, slope, size consistency, pen pressure etc.
- Hence recognition models must be capable to handle a diverse range of characters that are simple, complex and similar



Features of Kannada84:

1. 495 authors
2. Age: 9-60 years
3. Profession: Student, teacher, homemaker, clerk
4. Native Language: Both native and non-native speakers of Kannada
5. A total of 24265 letters of size 50x50

Hence Kannada84 represents characters that can be encountered in the real world

3. LITERATURE REVIEW:

Exploring deep learning techniques for Kannada handwritten character recognition: A boon for digitization

- Uses deep convolutional neural networks such as VGGNet, Inception Network
- Utilizes char74K dataset that consists of hand-drawn characters
- An accuracy of 86% is reported

Offline Character recognition on Segmented Handwritten Kannada Characters

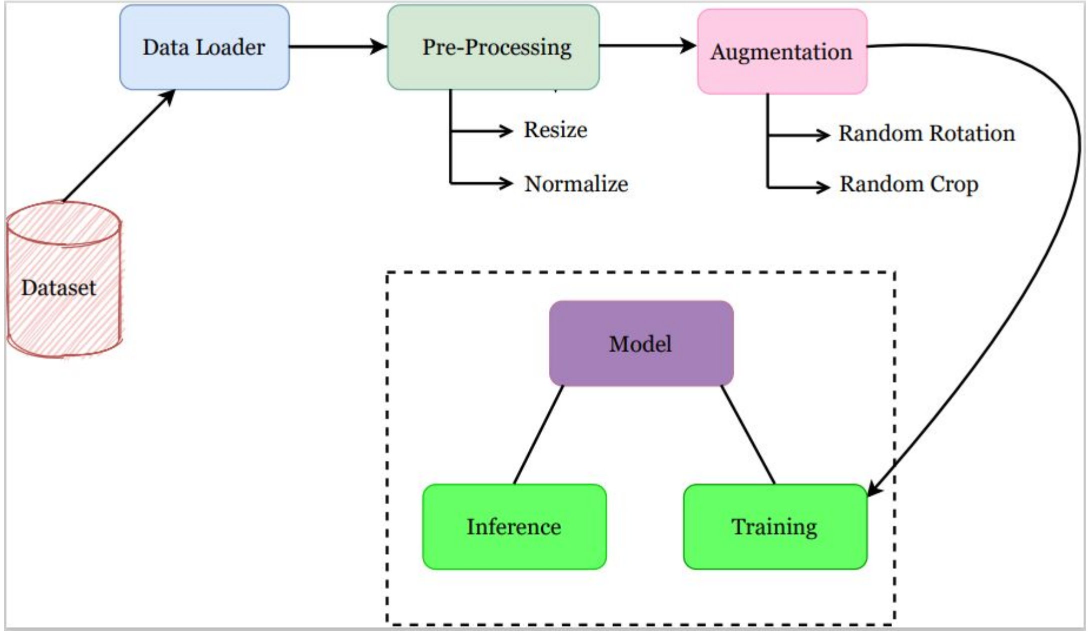
- Uses Multinomial Naive Bayes Classifier, Random Forest Classifier
- Utilizes char74K dataset that consists of hand-drawn characters

Features of char74K:

1. Comprises of characters that are hand-drawn
2. A total of 1225 characters of size 1200x900

4. PROPOSED METHODOLOGY:

a) Empirical Risk Minimization



Models used:

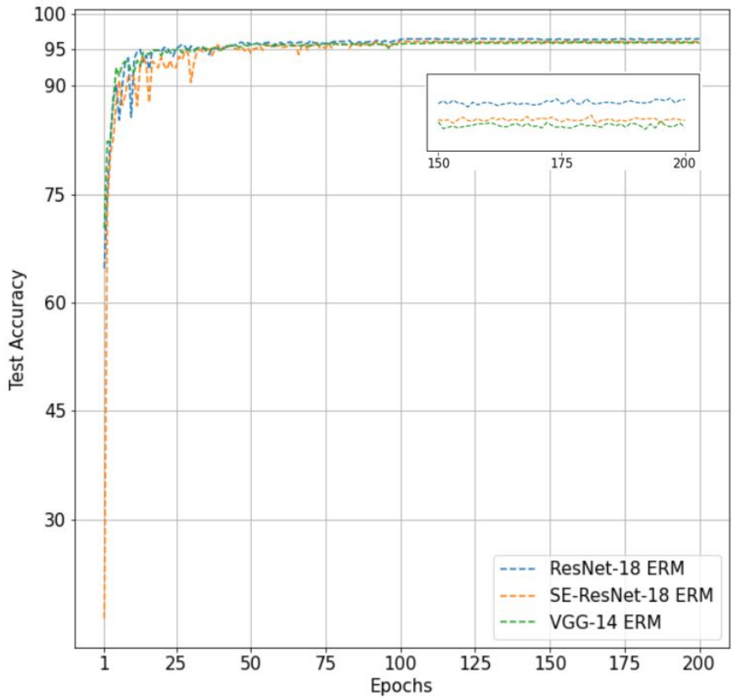
1. VGG Network
2. Residual Neural Network
3. Squeeze and Excitation Network

Training:

1. Empirical Risk Minimization
2. 70-30 split
3. SGD-M, 200 epochs, with step size 0.01 decayed by a factor of 10 (100, 150)

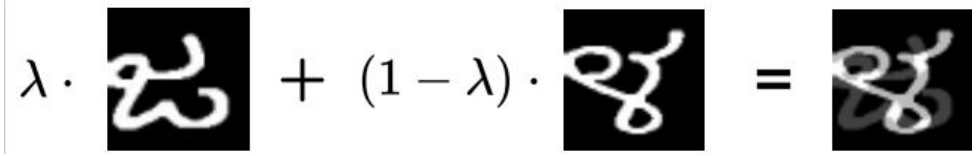
5. RESULTS:

Model	Method	Kannada84	
		Top-1 Test Accuracy	Top-3 Test Accuracy
VGG-14	ERM	96.03	99.28
ResNet-18	ERM	**96.49**	99.47
SE-ResNet-18	ERM	96.16	**99.56**



4. PROPOSED METHODOLOGY:

How can mixup regularization help?



- Introduces linear behaviour between 2 training instances
- Virtual training samples on the fly
- Minimal training overhead
- Reduces generalization error

4. PROPOSED METHODOLOGY:

Algorithm 1: Mixup

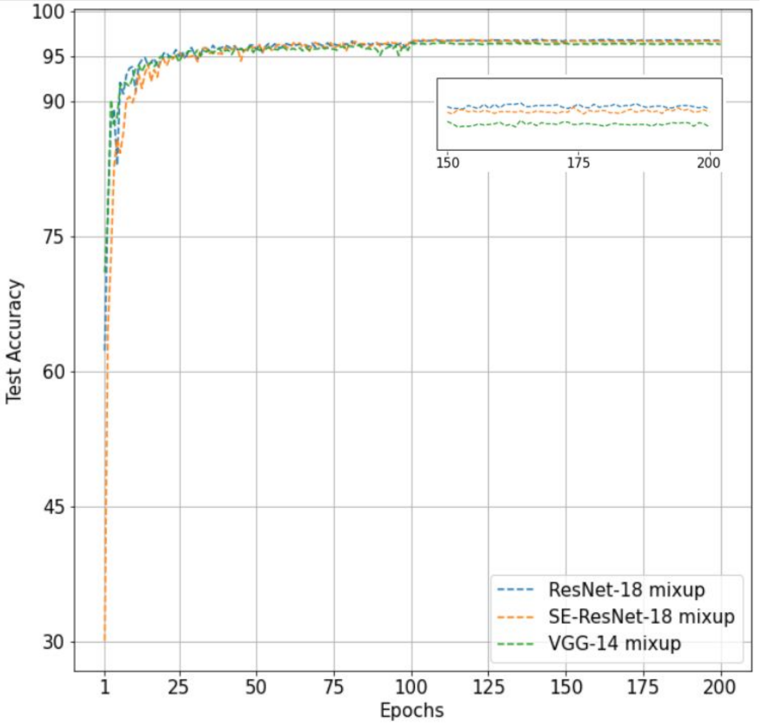
Require: Dataset $\mathcal{D}$ consisting of $1,2,\ldots,n$ I.I.D samples
Require: $f(\cdot\,;\theta)$ is a ConvNet with parameters θ
Require: $\mathcal{L}(\cdot)$ is the objective function that must be minimized
Require: η is the step size
Require: α, β are the parameters for Beta distribution
Result: Parameters θ^* after the model has converged
while θ not converged do
    Sample minibatch of m samples $\mathcal{A} = \{(X_1, y_1), \ldots, (X_m, y_m)\}$ from $\mathcal{D}$;
    Create a shuffled minibatch of m samples $\mathcal{B}$ from $\mathcal{A}$ such that $\mathcal{A}_i \neq \mathcal{B}_i\ \forall i \in \{1,2,\ldots,m\}$;
    $\mathcal{A}^{\mathcal{X}}, \mathcal{A}^{\mathcal{Y}} = \{X_1, \ldots, X_m\},\ \{y_1, \ldots, y_m\}$ such that $X_i \in \mathcal{A}$ and $y_i \in \mathcal{A}$;
    $\mathcal{B}^{\mathcal{X}}, \mathcal{B}^{\mathcal{Y}} = \{X_1, \ldots, X_m\},\ \{y_1, \ldots, y_m\}$ such that $X_i \in \mathcal{B}$ and $y_i \in \mathcal{B}$;
    $\lambda \sim Beta(\alpha, \beta)$;
    $X' = \lambda \cdot \mathcal{A}^{\mathcal{X}} + (1-\lambda) \cdot \mathcal{B}^{\mathcal{X}}$;
    $y' = \lambda \cdot \mathcal{A}^{\mathcal{Y}} + (1-\lambda) \cdot \mathcal{B}^{\mathcal{Y}}$;
    $\hat{y} = f(X'; \theta)$; ▷ Forward Propagation
    $G = \nabla_\theta \mathcal{L}(\hat{y}, y')$; ▷ Compute derivative of the objective function w.r.t parameters θ
    $\theta = \theta - \eta \cdot G$; ▷ Perform a step with gradient descent

Training:

1. Mixup
2. 70-30 split
3. SGD-M, 200 epochs, with step size 0.01 decayed by a factor of 10 (100, 150)
4. $\alpha = \beta = 1$

5. RESULTS:

Model	Method	Kannada84	
		Top-1 Test Accuracy	Top-3 Test Accuracy
VGG-14	ERM	96.03	99.28
	mixup	96.56	99.43
ResNet-18	ERM	96.49	99.47
	mixup	**96.92**	99.56
SE-ResNet-18	ERM	96.16	99.49
	mixup	96.90	**99.57**



6. ADDITIONAL RESULTS:

Model	Method	Kannada84		Char74K	
		Top-1 Test Accuracy	Top-3 Test Accuracy	Top-1 Test Accuracy	Top-3 Test Accuracy
VGG-14	ERM mixup	96.03 96.56	99.28 99.43	86.96 92.11	95.65 97.55
ResNet-18	ERM mixup	96.49 **96.92**	99.47 99.56	93.75 **94.84**	98.10 97.28
SE-ResNet-18	ERM mixup	96.16 96.90	99.49 **99.57**	92.66 94.30	98.13 **98.37**

6. CONCLUSION:

- A new dataset, Kannada84 is developed
- Kannada84 is large enough for deep learning architectures to take advantage of
- State-of-the-art deep convolutional neural networks are employed
- Performance is enhanced with mixup regularization/augmentation on the fly
- Top-1 and top-3 test accuracies are reported

7. FUTURE WORK:

- Recognize modifiers and vattakshara's present in Kannada language
- Employ the network in an end-to-end OCR engine

8. REFERENCES:

1. Rao, Abhishek and Arpitha, Anusha and Nayak, Chandana and Meghana, Sneha and Nayak, Sneha and S., Sandhya: Exploring deep learning techniques for kannada handwritten character recognition: A boon for digitization 29, 11078–11093 (07 2020)
2. Joe, K.G., Savit, M., Chandrasekaran, K.: Offline Character recognition on Segmented Handwritten Kannada Characters. In: 2019 Global Conference for Advancement in Technology (GCAT). pp. 1–5 (2019)
3. Zhang, H., Cissé, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: Beyond Empirical Risk Minimization. In: 6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings. OpenReview.net (2018)

Q & A