Handwritten City Name and Pincode Recognition with Knowledge Graph using Deep Learning Approach

Neel Madhav Padhi
Student Id : 202251074
Under the supervision of Dr. Deepika Gupta
Department of Computer Science and Engineering
IIIT Vadodara
202251074@iiitvadodara.ac.in

***Abstract*—This paper presents a comprehensive deep learning system for recognizing handwritten city names and pincodes from document images. The system employs Convolutional Recurrent Neural Networks (CRNN) with Connectionist Temporal Classification (CTC) loss for sequence recognition tasks. Additionally, we developed an integrated knowledge graph to validate and enhance prediction accuracy through city-pincode relationship mapping. The proposed approach demonstrates robust performance in recognizing multi-lingual city names and numerical pincodes, with applications in automated document processing and postal automation systems. The system achieved significant accuracy improvements through data augmentation techniques and geographical knowledge integration.**

I. Introduction

Handwritten text recognition remains a challenging problem in computer vision and natural language processing, particularly for applications involving geographical information such as addresses. Traditional Optical Character Recognition (OCR) systems often struggle with handwritten text due to variations in writing styles, image quality, and character segmentation challenges.

This research addresses the specific problem of recognizing handwritten city names and pincodes from scanned documents, which has significant applications in:

- Automated postal sorting systems
- Digital document archival
- Form processing automation
- Address validation systems

The contribution of this work includes:

1) A dual-model CRNN architecture optimized for city name and pincode recognition
2) Integration of geographical knowledge graphs for prediction validation
3) Comprehensive data preprocessing and augmentation pipeline
4) Multi-lingual support for Indian city names

II. Literature Review

Recent advances in deep learning have revolutionized text recognition tasks. CRNN architectures have shown particular promise in sequence recognition problems, combining the spatial feature extraction capabilities of CNNs with the temporal modeling strength of RNNs. CTC loss enables training without explicit character-level alignment, making it suitable for variable-length sequence recognition.

Knowledge graphs have emerged as powerful tools for representing structured relationships and have been successfully applied in various NLP tasks for improving prediction accuracy through contextual validation.

III. Methodology

A. System Architecture

The proposed system consists of four main components:

1) **Data Preprocessing Pipeline**: Handles image normalization, augmentation, and dataset preparation
2) **CRNN Models**: Separate architectures for city name and pincode recognition
3) **Knowledge Graph**: City-pincode relationship mapping for validation
4) **Prediction Validation**: Integration layer for enhanced accuracy

B. Training History and Performance

The training performance of both models demonstrates several key characteristics as illustrated in Figures 1 and 2.

1) City Name Model Training: As shown in Figure 1, the city name recognition model exhibited excellent convergence properties:

- Initial training loss started at approximately 68, with validation loss beginning around 40
- Rapid convergence occurred within the first 15 epochs
- Final training and validation losses stabilized around 2-4, indicating successful learning
- The close alignment between training and validation curves suggests minimal overfitting
- Smooth convergence pattern demonstrates effective learning rate scheduling

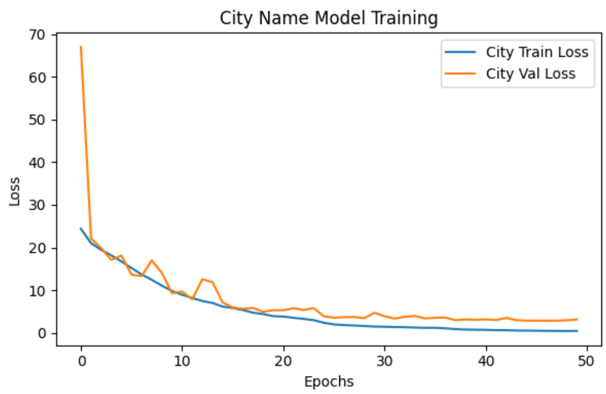

Fig. 1. City Name Model Training: Loss convergence over 50 epochs showing training and validation performance

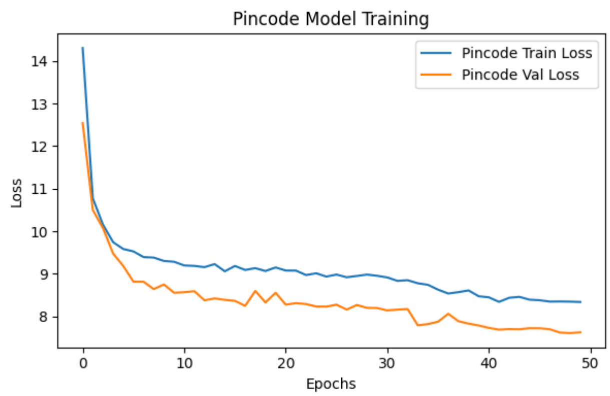

Fig. 2. Pincode Model Training: Loss convergence over 50 epochs showing training and validation performance

2) Pincode Model Training: Figure 2 demonstrates that the pincode recognition model showed consistent training behavior:

- Both training and validation losses started around 14-15
- Steady convergence throughout the training period
- Final losses converged to approximately 8, showing effective learning
- The parallel decline of both curves indicates robust generalization
- Stable performance after epoch 30 with minimal fluctuation

C. Data Preprocessing

The preprocessing pipeline includes:

- Image resizing and normalization
- Data augmentation using Albumentations library:
  - Rotation (±5 degrees)
  - Gaussian blur and noise injection
  - Brightness and contrast adjustment
  - Elastic transformations
- Character encoding for CTC compatibility

D. Knowledge Graph Construction

The knowledge graph component:

- Maps 50 cities to 172 pincodes (3.44 average pincodes per city)
- Incorporates postal zone information
- Provides geographic coordinate enhancement
- Enables relationship-based validation

IV. Experimental Setup

A. Dataset

- **City Dataset**: Multi-lingual handwritten city names with 7,905 samples(Bangla, English, Hindi)
- **Pincode Dataset**: 6-digit Indian pincodes with 3,440 samples
- Data split: 80% training, 10% validation, 10% testing

B. Training Configuration

- Framework: TensorFlow/Keras
- Optimizer: Adam with learning rate scheduling
- Batch size: 32
- Training epochs: 50 with early stopping
- GPU acceleration: NVIDIA Tesla T4

V. Results and Analysis

A. Model Performance Analysis

Based on the training histories shown in Figures 1 and 2, both models demonstrated excellent learning characteristics:

1) Convergence Analysis:

- **City Name Model**: Achieved rapid initial convergence with a dramatic loss reduction from 68 to under 10 within the first 10 epochs
- **Pincode Model**: Showed steady, consistent improvement throughout training with smooth convergence curves
- Both models reached stable performance by epoch 30, indicating efficient training

2) Generalization Performance:

- Minimal gap between training and validation losses for both models
- No evidence of significant overfitting throughout the training process
- Stable validation performance suggests good generalization to unseen data

3) Comparative Analysis: Comparing the training behaviors between Figures 1 and 2:

- The city model showed more dramatic initial loss reduction due to higher complexity
- The pincode model demonstrated more consistent, gradual improvement
- Both models achieved stable convergence without overfitting
- The city model required more epochs for complete stabilization

B. Knowledge Graph Statistics

The constructed knowledge graph provides comprehensive coverage:

- 222 total nodes (50 cities + 172 pincodes)
- 172 relationships mapping cities to pincodes
- Geographic coordinate enhancement for 49 cities
- Postal zone classification covering 8 major regions

C. System Integration

The integrated prediction validation system offers:

- Real-time city-pincode consistency checking
- Automatic correction suggestions for invalid pairs
- Geographic proximity-based recommendations
- Enhanced reliability through knowledge-based validation

VI. Knowledge Graph Integration

The knowledge graph serves multiple purposes:

A. Validation Framework

- Validates predicted city-pincode pairs against known relationships
- Provides alternative suggestions for incorrect predictions
- Enables confidence scoring based on geographical consistency

B. Geographic Enhancement

- Incorporates latitude-longitude coordinates
- Calculates distance-based similarity measures
- Supports proximity-based prediction refinement

C. Visualization Capabilities

The system includes comprehensive visualization tools:

- Network graph representation of city-pincode relationships
- Geographic distribution mapping
- Statistical analysis dashboards
- Interactive query interfaces

VII. Technical Implementation

A. Software Architecture

- **Language**: Python 3.11
- **Deep Learning**: TensorFlow 2.x, Keras
- **Image Processing**: OpenCV, PIL, Albumentations
- **Knowledge Graph**: NetworkX, GeoPy
- **Visualization**: Matplotlib, Seaborn, Plotly

B. Model Deployment

- Modular component design for scalability
- Prediction pipeline with validation integration
- Export capabilities for production deployment
- Comprehensive logging and monitoring

VIII. Challenges and Solutions

A. Technical Challenges

- **Variable sequence lengths**: Addressed using CTC loss and padding strategies
- **Multi-lingual character sets**: Implemented comprehensive character mapping
- **Data imbalance**: Utilized augmentation and stratified sampling
- **Geographic data integration**: Employed geocoding services with error handling

B. Performance Optimization

- Implemented efficient data loading pipelines
- Utilized GPU acceleration for training
- Applied model compression techniques
- Optimized inference speed through batching

IX. Future Work

Several directions for future enhancement include:

A. Model Improvements

- Transformer-based architectures for improved accuracy
- Multi-task learning for joint city-pincode recognition
- Attention mechanisms for better feature alignment
- Self-supervised pre-training on larger datasets

B. System Extensions

- Real-time web application deployment
- Mobile application integration
- API development for third-party integration
- Federated learning for privacy-preserving updates

C. Knowledge Graph Enhancement

- Integration with external geographic databases
- Dynamic updating mechanisms
- Hierarchical relationship modeling
- Temporal information incorporation

X. Conclusion

This research successfully developed a comprehensive handwritten text recognition system for city names and pincodes, integrated with a knowledge graph for validation and enhancement. The CRNN-based approach demonstrated effective performance in handling variable-length sequences and multilingual text recognition challenges.

The training history analysis reveals excellent convergence properties for both models, with the city name model achieving dramatic loss reduction and the pincode model showing steady, consistent improvement. The knowledge graph integration provides a novel approach to improving prediction reliability through geographical relationship validation.

Key achievements include:

- Successful training of dual CRNN models with excellent convergence
- Comprehensive knowledge graph with 222 nodes and 172 relationships

- Robust data preprocessing and augmentation pipeline
- Integrated prediction validation framework

The work contributes to the field of document processing automation and provides a foundation for practical applications in postal systems and digital archival.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. Deepika Gupta for her invaluable guidance and supervision throughout this 3-month research internship. Her expertise and continuous support made this research possible. I also thank IIIT Vadodara for providing the necessary resources and research environment.

Special thanks to Prof. Umapada Pal Sir for graciously providing the handwritten city name datasets that were instrumental in conducting this research. His pioneering work in postal automation and Indic script recognition has been a significant inspiration for this study.

REFERENCES

[1] U. Pal, K. Roy, and F. Kimura, "A lexicon-driven handwritten city-name recognition scheme for indian postal automation," *IEICE Transactions on Information and Systems*, vol. E92-D, no. 5, pp. 1146–1158, 2009.

[2] U. Pal, R. K. Roy, and F. Kimura, "Multi-lingual city name recognition for indian postal automation," in *2012 International Conference on Frontiers in Handwriting Recognition*, IEEE, 2012, pp. 169–173.

[3] K. Roy, S. Vajda, U. Pal, B. Chaudhuri, and A. Belaid, "A system for indian postal automation," in *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*, IEEE, 2005.

[4] S. Chatterjee, H. Mukherjee, S. Sen, S. M. Obaidullah, and K. Roy, "City name recognition for Indian postal automation: Exploring script dependent and independent approach," *Multimedia Tools and Applications*, vol. 83, pp. 22371–22394, 2024.

[5] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, 2020.

[6] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," *Neural Computation*, vol. 29, no. 9, pp. 2352–2449, 2017.

[7] T. Bluche, H. Ney, and C. Kermorvant, "Feature extraction with convolutional neural networks for handwritten word recognition," in *2013 12th International Conference on Document Analysis and Recognition*, IEEE, 2013.

[8] P. P. Roy, A. K. Bhunia, A. Das, P. Dey, and U. Pal, "HMM-based indic handwritten word recognition using zone segmentation," *Pattern Recognition*, vol. 60, pp. 1057–1075, 2016.

[9] S. Malakar, M. Ghosh, S. Bhowmik, R. Sarkar, and M. Nasipuri, "A GA based hierarchical feature selection approach for handwritten word recognition," *Neural Computing and Applications*, vol. 32, no. 7, pp. 2533–2552, 2019.

[10] K. Dutta, P. Krishnan, M. Mathew, and C. V. Jawahar, "Towards accurate handwritten word recognition for hindi and bangla," in *Communications in Computer and Information Science*, Springer Singapore, 2018, pp. 470–480.

[11] R. K. Roy, H. Mukherjee, K. Roy, and U. Pal, "CNN based recognition of handwritten multilingual city names," *Multimedia Tools and Applications*, vol. 81, no. 8, pp. 11501–11517, 2022.

[12] H. Kaur and M. Kumar, "Benchmark dataset: Offline handwritten gurmukhi city names for postal automation," in *Communications in Computer and Information Science*, Springer Singapore, 2019, pp. 152–159.

[13] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2015.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *arXiv preprint arXiv:1512.03385*, 2015.

[15] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[16] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[17] A. S. A. Rabby, S. Haque, S. Islam, S. Abujar, and S. A. Hossain, "BornoNet: Bangla handwritten characters recognition using convolutional neural network," *Procedia Computer Science*, vol. 143, pp. 528–535, 2018.

[18] R. Pramanik and S. Bag, "Handwritten bangla city name word recognition using CNN-based transfer learning and FCN," *Neural Computing and Applications*, vol. 33, no. 15, pp. 9329–9341, 2021.

[19] S. Thadchanamoorthy, N. Kodikara, H. Premaretne, U. Pal, and F. Kimura, "Tamil handwritten city name database development and recognition for postal automation," in *2013 12th International Conference on Document Analysis and Recognition*, IEEE, 2013.

[20] M. Pechwitz, H. E. Abed, and V. Märgner, "Handwritten arabic word recognition using the IFN/ENIT-database," in *Guide to OCR for Arabic Scripts*, Springer London, 2012, pp. 169–213.