Optical Character Recognition

Using Convolutional Neural Networks

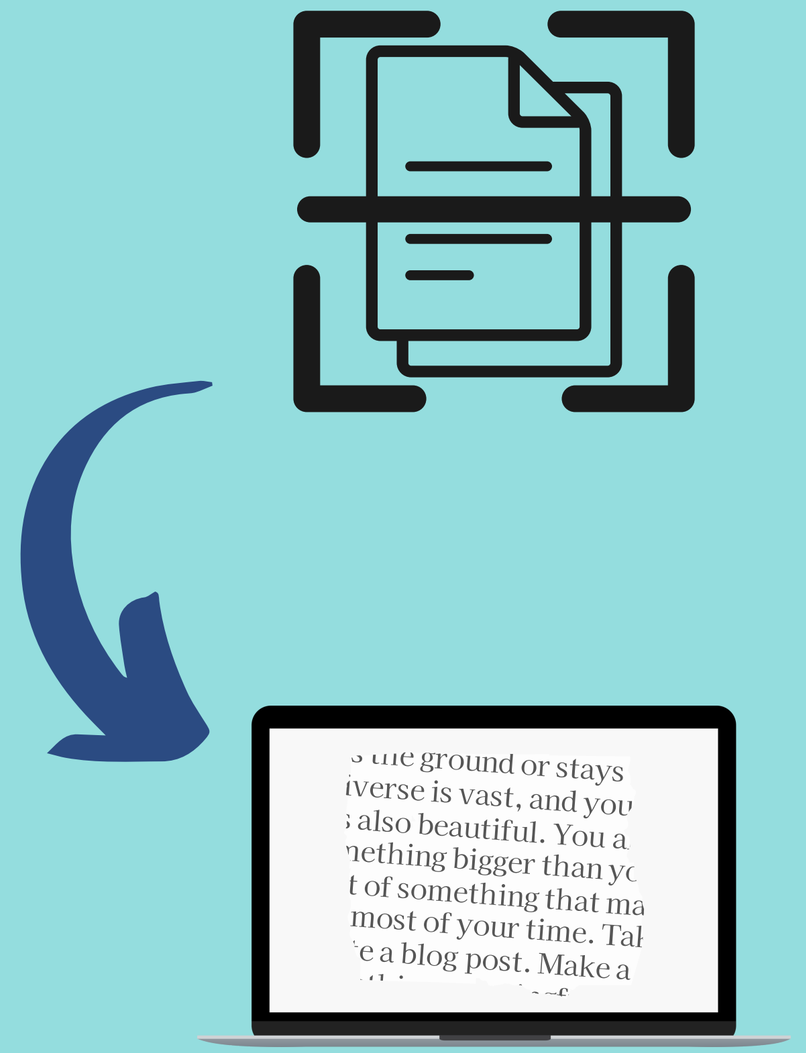
Agenda

KEY TOPICS DISCUSSED IN THIS PRESENTATION

- What is OCR ?
- Applications of OCR
- Goals of Implementation
- Objectives
- Dataset Used
- System Design
- Performance Analysis
- Future Work
- References

What is OCR ?

- OCR stands for Optical Character Recognition, which is a technology used to recognize and extract text from images, scanned documents, and other visual media.
- OCR technology has become increasingly important in today's digital world, as it allows for the digitization of printed materials, making them easier to search, edit, and share.



Applications of OCR

OCR: Because who has time to type out an entire document when you can just make your computer do it for you? #LazyLife

License plate recognition	Handwriting recognition
Receipt scanning to track expenses	Medical record digitization
Legal document processing	Educational material digitization

Goals of Implementation

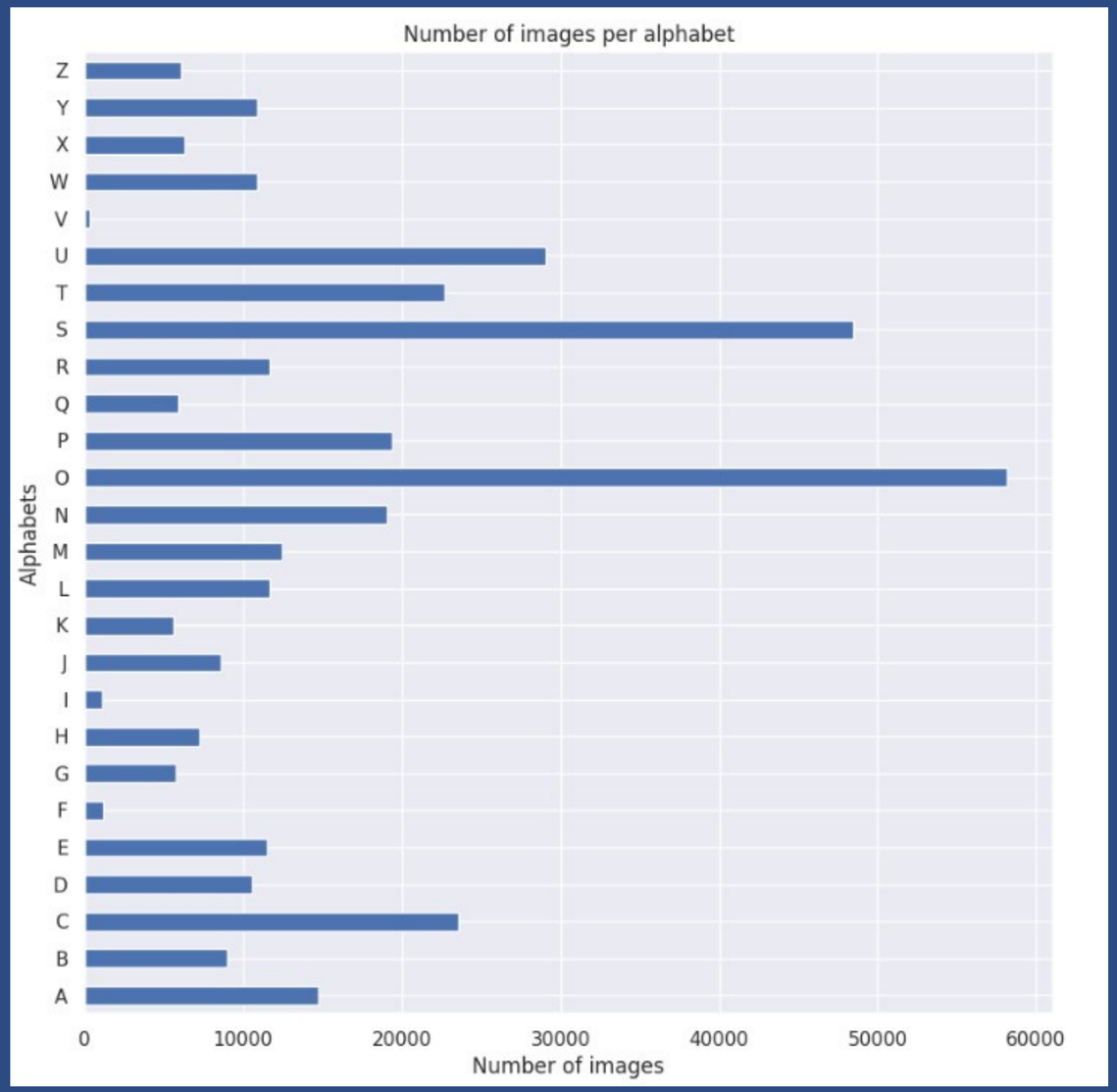
TO DEVELOP AN OCR MODEL THAT IS ACCURATE, EFFICIENT, AND SCALABLE

Objectives

- To evaluate and select the most appropriate OCR model for our needs.
- To train our OCR model on a large and diverse dataset of character images.
- To test the accuracy of our OCR model.
- To document our findings and share our results with the broader community.

Dataset Used

- Consists of 28x28 pixel images of handwritten alphabets.
- Total of 372,037 rows of image representing an alphabet.
- The images are in a pixel format.
- It has 785 columns.
- 1st column representing the labelling of the alphabet from 0-25 as A-Z.



System Design

1 — 2 — 3

STEP

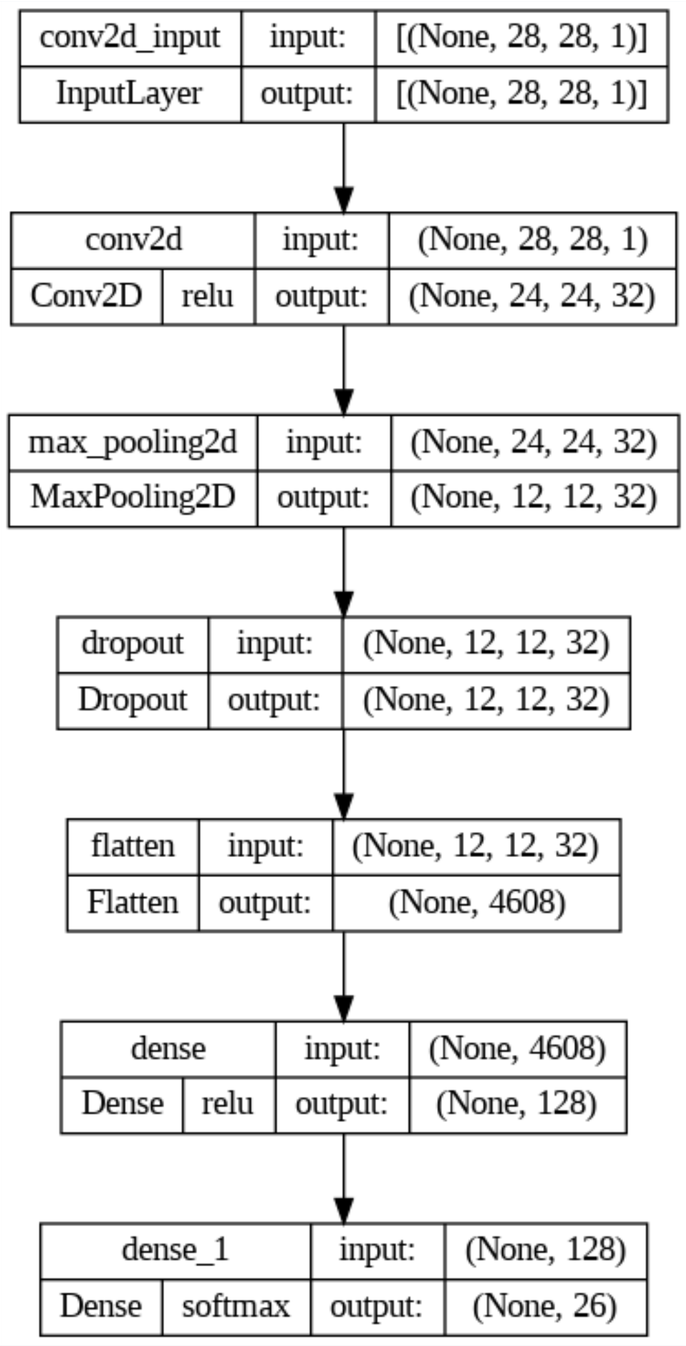
Data Preprocessing
- Scaling the pixel values

STEP

Model architecture - Convolutional Neural Network (CNN).

STEP

Model's performance - Accuracy



Model Architecture

1. "Conv2D(32, (5, 5), input_shape=input_shape, activation='relu')": Imagine a bunch of 32 detective filters roaming around, investigating a 5x5 grid of pixels with the hope of finding something exciting. When they do, they get really excited and activate their "relu" mode!
2. "MaxPooling2D(pool_size=(2, 2))": Now that the detectives have collected some valuable clues, they decide to take a break and dive into a mini swimming pool with a size of 2x2. It's their way of cooling down and shrinking their findings to a more manageable size.
3. "Dropout(0.3)": Just when you thought the detectives were focused, they randomly decide to drop 30% of their collected clues. It's like they have a mischievous side and enjoy keeping things unpredictable!
4. "Flatten()": After their swimming pool adventure, the detectives decide to flatten themselves like pancakes. It's quite a sight to see these once three-dimensional investigators turning into a single line of data.
5. "Dense(128, activation='relu')": Now the detectives enter a room with 128 energetic colleagues, all fired up in their "relu" mode. They exchange information and amplify their excitement as they dive deeper into the investigation.
6. "Dense(num_classes, activation='softmax')": Finally, the detectives gather in a meeting room where they represent the different classes of suspects. They vote and determine the probabilities of each suspect being guilty, using a "softmax" voting system.

Our CNN model’s overall accuracy on the test set is 98.16%.

- Accuracy on the training set - 98.12%.
- Accuracy on the validation set - 98.22%.

What's Next ?

- Using Image augmentation, which can create synthetic training samples by applying a variety of image transformations such as rotation, flipping, and zooming.
- Investigating the impact of hyperparameters tuning can help to further optimize the model performance.
- Extending the model to recognize characters in other languages for creating a universal OCR model.

OCR may soon become advanced enough to recognize the doctor's handwriting. A true miracle in the making!

What we Learnt ?

- Use of Convolutional Neural Networks (CNNs) in image recognition tasks.
- Significance of hyperparameter tuning for optimizing model performance.
- The practical applications of OCR technology.
- The impact of OCR technology in various fields.

References

I didn't come up with this all by myself, you know? Here are the sources that made me look smart:

T. Jamtsho, K. Powdyel, R. K. Powrel, R. Bhujel and K. Muramatsu, "OCR and Speech Recognition System Using Machine Learning," 2021 Innovations in Power and Advanced Computing Technologies (i-PACT), Kuala Lumpur, Malaysia, 2021, pp. 1-5, doi: 10.1109/i-PACT52855.2021.9697030.
April 15 2023

https://www.kaggle.com/datasets/sachinpatel21/az-handwritten-alphabets-in-csv-format
April 15 2023

https://www.tensorflow.org/api_docs/python/tf/keras
April 15 2023

https://docs.opencv.org/4.x/
April 15 2023

https://medium.com/analytics-vidhya/optical-character-recognition-using-tensorflow-533061285dd3
April 15 2023

THANK YOU