RV College of Engineering®

Go, change the world

Department of AIML

ANN & DL (AI253IA) LAB PROJECT

TITLE: OCR Sudoku Puzzle Solver

Presented by:

Srivanth Srinivasan - 1RV22AI058

Rajyalakshmi - 1RV22AI041

Faculty Mentors: Prof. Somesh Nandi and Prof. S. Anupama Kumar

Agenda

1. Introduction
2. Objectives
3. Literature Survey
4. Summary of Literature Survey
5. Requirement Analysis - Hardware & Software Specification
6. System Architecture - Architecture of CNN
7. Methodology
8. Module Specification
    a. Data collection & preprocessing
    b. Implementation of algorithm
    c. Testing & Validation
9. Conclusion

Introduction

- Sudoku puzzles are popular brain-teasers that enhance logical thinking and problem-solving skills.
- The OCR Sudoku Puzzle Solver project uses optical character recognition (OCR) and deep learning (DL) to automate solving Sudoku puzzles from captured or uploaded images.

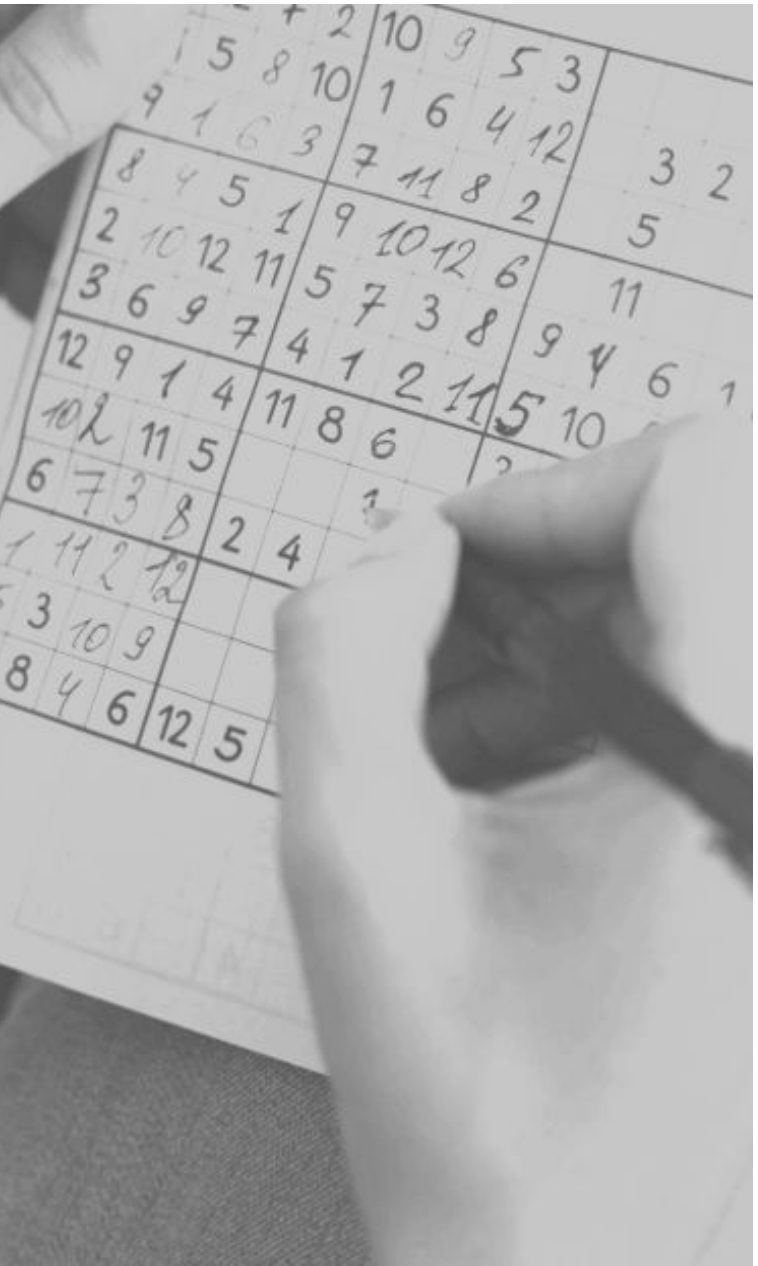
Domains:

- Computer Vision: Grid detection and segmentation
- Artificial Intelligence: OCR for digit recognition
- Game Theory: Algorithmic puzzle solving

Relevance:

- Assists Sudoku enthusiasts in solving or verifying solutions.
- Provides practical insights into OCR and DL applications.

Objectives



- To develop a system that can accurately solve Sudoku puzzles directly from real-world images.
- Use **optical character recognition (OCR)** to digitize Sudoku grids and recognize printed or handwritten digits.
- Apply a **Convolutional Neural Network (CNN)** to enhance digit recognition accuracy.
- Employ the **backtracking algorithm** to efficiently solve digitized Sudoku puzzles.
- Superimpose the solved puzzle onto the original image or display it as a clean digital grid.

Literature Survey

Sl No	Author and Paper title	Details of Publication	Summary of the Paper
1	A Dataset of Sudoku Puzzles With Difficulty Metrics Experienced by Human Players	Wang, Sheng-Wei. "A Dataset of Sudoku Puzzles with Difficulty Metrics Experienced by Human Players." IEEE Access (2024).	This paper addresses the problem of correct evaluation of difficulty level of Sudoku. Here, dataset with human player metrics are used collected via android app 'Cloud Sudoku'. Metrics collected are 1)Time spent 2)Success rate.
2	Augmented Reality-based Sudoku Solver with Training Module to Improve Cognitive Skills	Dugar, Abhinav, et al. "Augmented Reality-based Sudoku Solver with Training Module to Improve Cognitive Skills." 2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC). IEEE, 2024.	Abstract—This paper presents an interactive Augmented Reality (AR) Sudoku Scanner application that leverages Computer Vision and a dancing link algorithm to solve Sudoku puzzles in real-time. The proposed app allows users to view a leaderboard, improve cognitive skills, create a profile & view performance analytics.
3	Enhancing Classification Power: Tree Strength-Infused Enriched Random Forest	Jain, Vikas, Tej Bahadur Chandra, and Atul Kumar Srivastava. "Enhancing Classification Power: Tree Strength-Infused Enriched Random Forest." 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2024.	This paper proposes a new approach to building Random Forests called Enriched Random Forests (ERF). It uses a combination of Bag of Visual Words (BoVW) for feature extraction and Grey Wolf Optimization (GWO) for feature selection to improve the performance of the RF. The ERF model shows superior results compared to traditional RF methods.

Literature Survey

Sl No	Author and Paper title	Details of Publication	Summary of the Paper
4	MNSIT Handwritten Digit Recognition using CNN	Ahmareen, Shafaque, Alreem Khalid Khamies Alabdouli, and Sirisha Polturi. "MNSIT Handwritten Digit Recognition using CNN." 2024 5th International Conference on Image Processing and Capsule Networks (ICIPCN). IEEE, 2024.	This research aims to use Convolutional Neural Networks (CNN) to recognize handwritten digits from the MNIST dataset. The study involves data preprocessing, model architecture design, hyperparameter tuning, and performance evaluation. By experimenting with different CNN configurations
5	Handwritten Image Detection using DCGAN with SIFT and ORB Optical Features	Dubey, Rohan, and Ipshita Das. "Handwritten Image Detection using DCGAN with SIFT and ORB Optical Features." 2023 6th International Conference on Information Systems and Computer Networks (ISCON). IEEE, 2023.	This paper proposes a novel approach for handwritten image detection that combines deep convolutional GANs with optical feature extraction techniques like ORB and SIFT. It incorporates optical features into the GAN model which improves accuracy of digit recognition on MNIST, EMNIST, and MADBase datasets. This outperforms traditional GAN-based methods.
6	MNSIT Handwritten Digit Recognition Using Machine Learning Classification Algorithms	Anoop, R., et al. "MNSIT Handwritten Digit Recognition Using Machine Learning Classification Algorithms." 2023 World Conference on Communication & Computing (WCONF). IEEE, 2023.	This paper explores the use of machine learning for recognizing handwritten digits and text. Algorithms like SVM, KNN,RFC with varying hidden layers.The study highlights the challenges of handwriting recognition due to variations in writing styles and emphasizes the importance of accurate recognition in various applications

Literature Survey

Sl No	Author and Paper title	Details of Publication	Summary of the Paper
7	Research on English Automatic Recognition System Based on OCR Technology	Li, Jieying. "Research on English Automatic Recognition System Based on OCR Technology." 2023 7th Asian Conference on Artificial Intelligence Technology (ACAIT). IEEE, 2023.	This paper proposes a system for recognizing handwritten English text using OCR technology. The system combines Tesseract OCR with a CNN+LSTM neural network to improve accuracy. The experimental results show that the system is effective in recognizing text with high integrity, especially when compared to using CNN alone.
8.	Image based Sudoku Solver using Applied Recursive Backtracking	Anasune, Aditya, and Sakshi Bhavsar. "Image based Sudoku Solver using Applied Recursive Backtracking." 2023 2nd International Conference on Futuristic Technologies (INCOFT). IEEE, 2023.	The paper describes how recursive backtracking algorithm is utilized to make a sudoku solver. This involves detection of sudoku grid (shows 80% accuracy) followed by solving it (shows 100% accuracy). This was possible due to use of CNN & image pre-processing techniques like gaussian blur.
9	The Algorithm Selection Problem for Solving Sudoku with Metaheuristics	Notice, Danielle, Ahmed Kheiri, and Nicos G. Pavlidis. "The Algorithm Selection Problem for Solving Sudoku with Metaheuristics." 2023 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2023.	In this paper we study the algorithm selection problem and instance space analysis for solving Sudoku puzzles with metaheuristic algorithms. Sudoku is treated as a com-binatorial optimisation problem and we implement four local-search metaheuristics to solve the problem instances.

Literature Survey

Sl No	Author and Paper title	Details of Publication	Summary of the Paper
10	Implementation of a Sudoku Puzzle Solver on a FPGA	Ciantar, Keith George, and Owen Casha. "Implementation of a Sudoku Puzzle Solver on a FPGA." 2023 9th International Conference on Control, Decision and Information Technologies (CoDIT). IEEE, 2023.	In this paper, the design, implementation and evaluation of a hybrid Sudoku puzzle solver on a Field- Programmable Gate Array (FPGA) is presented. It uses pen & paper technique to fill initial numbers, then uses DFS to backtrack through the puzzle till a solution is reached.Implementation done on Xilinix Spartan-6 XC6SLX45 FPGA.
11	Design and Analysis of a Modified 3D Sudoku Solver	Jana, Sunanda, et al. "Design and Analysis of a Modified 3D Sudoku Solver." IEEE Access 11 (2023): 27352-27368.	The paper presents the design and analysis of two modified algorithms for solving 3D Sudoku puzzles, which are more complex than traditional 2D Sudoku due to their additional dimension and constraints
12	An Improved Sudoku Solver	Wadud, Adiba Afif Suha Binta, and Mohammad Abdullah-Al-Wadud. 2022 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2022.	In this paper, we have improved the existing HBPnP method by creating additional data structures and modified algorithms.It focuses on Sudoku-solving methods and backtracking algorithms.

Literature Survey

Sl No	Author and Paper title	Details of Publication	Summary of the Paper
13	Realization of Computer-Assisted Language Reading Training Platform based on Non-Destructive Scanning Technology and OCR Recognition Optimization	He, Xiaofang. "Realization of Computer-Assisted Language Reading Training Platform based on Non-Destructive Scanning Technology and OCR Recognition Optimization." 2022 International Conference on Inventive Computation Technologies (ICICT). IEEE, 2022.	In this paper, the system uses OCR to scan and recognize student ID information. This information is then used to create a personalized reading training platform based on ASP.Net. It uses tesseract OCR to recognize characters.
14	Two Decades of Bengali Handwritten Digit Recognition: A Survey	Rahman, ABM Ashikur, et al. "Two decades of bengali handwritten digit recognition: A survey." IEEE Access 10 (2022): 92597-92632.	Handwritten Digit Recognition (HDR) is a challenging task due to variations in writing styles and the complexity of certain languages like Bengali. This paper reviews the state-of-the-art techniques for Bengali HDR, including traditional machine learning and deep learning approaches.
15	Design and Simulation of Optical Neural Network	Al-Hasnawi, Zainab Faris, and Ibrahim A. Murdas. "Design and Simulation of Optical Neural Network." 2022 2nd International Conference on Advances in Engineering Science and Technology	This paper explores the use of optical neural networks for image classification tasks. By leveraging the benefits of light, such as low power consumption and high speed,they propose an optical neural network. It is trained and tested on the MNIST and EMNIST datasets, achieving an accuracy of 99.58% on MNIST. The use of backpropagation allows for efficient training of the network.

Literature Survey

Sl No	Author and Paper title	Details of Publication	Summary of the Paper
16	A Classical Approach to Handcrafted Feature Extraction Techniques for Bangla Handwritten Digit Recognition	Ferdous Wahid, Md, Fahim Shahriar, and Shohanur Islam Sobuj. "A Classical Approach to Handcrafted Feature Extraction Techniques for Bangla Handwritten Digit Recognition." arXiv e-prints (2022): arXiv-2201.	This paper focuses on recognizing handwritten Bengali digits, a challenging task due to their complex structure. The researchers experimented with various machine learning classifiers (KNN, SVM, RF, GBDT) and feature extraction techniques (HOG, LBP, Gabor) on four different datasets.
17	Searching for Explainable Solutions in Sudoku	Björnsson, Yngvi, Sigurður Helgason, and Aðalsteinn Pálsson. "Searching for explainable solutions in Sudoku." 2021 IEEE Conference on Games (CoG). IEEE, 2021.	This paper explores how to make the decision-making process of AI systems more human-understandable, focusing on heuristic search. It uses Sudoku puzzles as a testbed to develop techniques that generate solutions that are easy for humans to comprehend.
18	Strengthening Probabilistic Graphical Models: The Purge-and-Merge Algorithm	Streicher, Simon, and Johan A. Du Preez. "Strengthening probabilistic graphical models: The purge-and-merge algorithm." IEEE Access 9 (2021): 149423-149432.	This paper proposes a new algorithm called "purge-and-merge" to solve constraint satisfaction problems (CSPs) using probabilistic graphical models (PGMs).

Literature Survey

Sl No	Author and Paper title	Details of Publication	Summary of the Paper
19	Key-Value Pair Searching System via Tesseract OCR and Post Processing	Kaló, Áron Zoltán, and Miklós László Sipos. "Key-Value Pair Searching System via Tesseract OCR and Post Processing." 2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI). IEEE, 2021.	This paper proposes a system that uses Tesseract OCR to extract text from images and then performs a key-value pair search on the extracted text. To improve accuracy, image preprocessing is used to minimize noise, and post-processing techniques like regular expressions are applied to the OCR output.
20	Solving Sudoku With Ant Colony Optimization	Lloyd, Huw, and Martyn Amos. "Solving Sudoku with ant colony optimization." IEEE Transactions on Games 12.3 (2019): 302-311.	This paper introduces ant-colony-optimization-based method significantly outperforms the state-of-the-art algorithm on the hardest, large instances of Sudoku. It performs better tha backtracking algorithms.
21	Image retrieval based on a hybrid model of deep convolutional encoder	Qin, Jingkun, et al. "Image retrieval based on a hybrid model of deep convolutional encoder." 2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE). IEEE, 2018.	This paper proposes a hybrid model combining CNN and autoencoder to improve image search accuracy. The CNN extracts high-level semantic features, while the autoencoder reduces dimensionality.The proposed method outperforms state-of-the-art techniques on CIFAR-10 and MNIST datasets, achieving high accuracy and recall rates

Summary of Literature Survey

- Sudoku-solving methods include recursive backtracking, FPGA-based solvers, metaheuristics (e.g., ant colony optimization), and human-comprehensible AI.
- Advanced techniques like 3D Sudoku algorithms address more complex constraints.
- Handwritten recognition is improved using CNN, GAN, OCR (e.g., Tesseract), and hybrid CNN+LSTM on datasets like MNIST and Bengali scripts.
- AR-based Sudoku scanners and OCR systems integrate image preprocessing for real-world applications.
- Enriched Random Forests (ERF) and CNN-autoencoder hybrids enhance feature extraction and dimensionality reduction.
- Novel algorithms like "purge-and-merge" solve constraint satisfaction problems efficiently.
- Optical neural networks leverage light-based hardware for faster, energy-efficient computation.
- AI-driven platforms for student IDs and personalized training demonstrate practical applications.
- Hybrid AI approaches improve decision-making transparency and usability.
- Interdisciplinary solutions highlight advancements in AI, hardware integration, and real-world deployment.

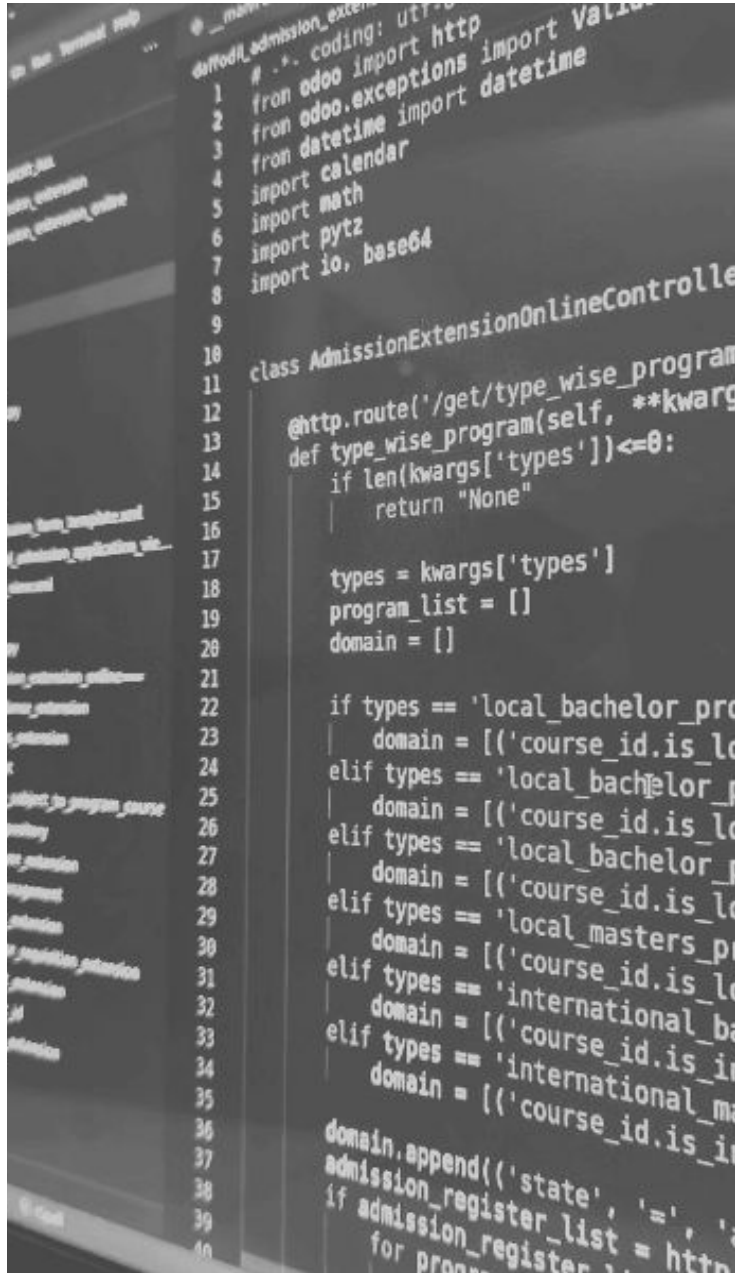
Requirement Analysis

Hardware Requirements

- **Processor:** Intel Core i5 (8th gen) or higher
- **RAM:** 8 GB or more
- **GPU:** NVIDIA GTX 1050 or better (for DL tasks)
- **Storage:** 256GB SSD

Software Requirements

- **OS:** Windows/Linux/MacOS
- **Development Tools:** IDE - PyCharm or Visual Studio Code
- **Additional Software:** Tesseract Installation (and added to the system path)



Requirement Analysis

Required libraries and tools:

LIBRARY	VERSION	PURPOSE
OpenCV	4.5.5	Image processing
Tesseract OCR	4.1.1	OCR for digit extraction
pytesseract	0.3.8	Python wrapper for Tesseract OCR
TensorFlow/Keras	2.8.0	DL & training
Numpy	1.22.0	Numerical computations
Matplotlib	3.5.1	Visualization
scikit-learn	1.0.2	Model evaluation & metrics
Pillow	9.0.1	Image manipulation & handling
imutils	0.5.4	Image processing utilities for OpenCV

Requirement Analysis

USER REQUIREMENTS

Hardware Requirements

- **Processor:** Intel Core i3 (7th gen) or higher
- **RAM:** 4 GB or more
- **GPU:** Integrated GPU
- **Storage:** 128GB SSD

Software Requirements

- **OS**: Windows/Linux/MacOS
- **Application interface:** Web
- **Dependencies:** OpenCV, Tesseract, TensorFlow etc.

System Architecture

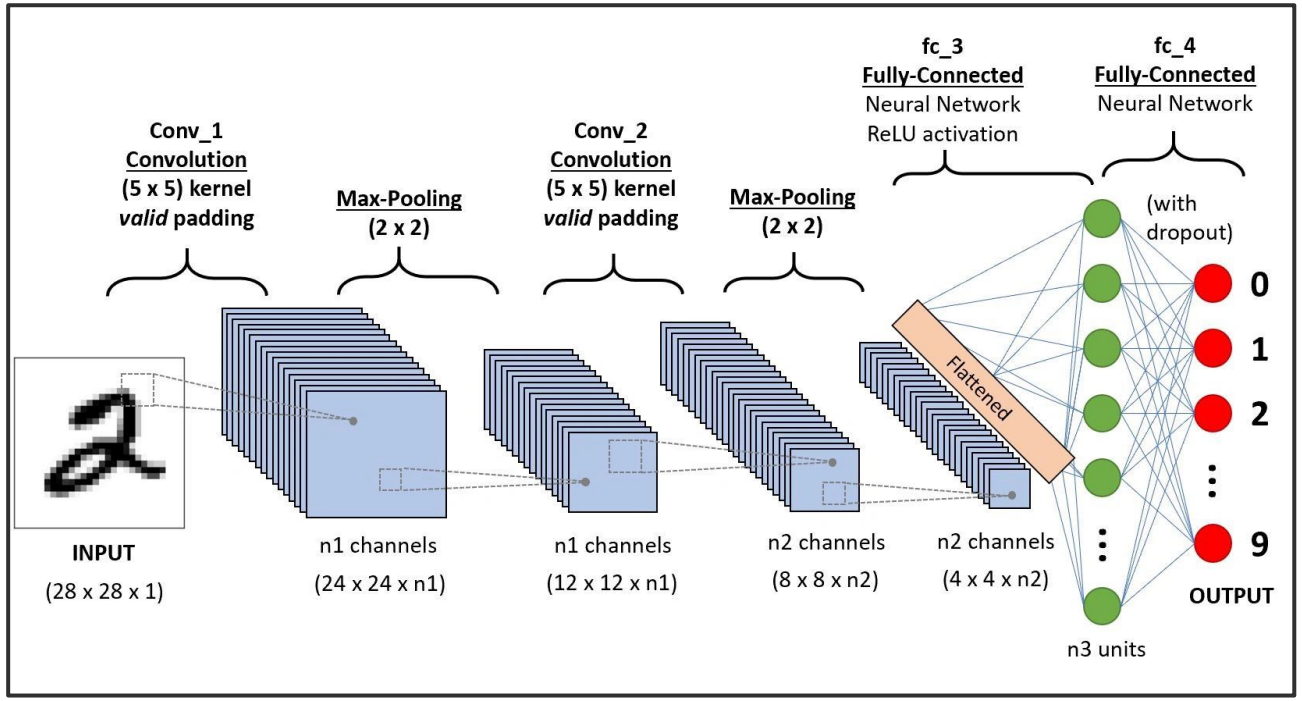
ARCHITECTURE OF CNN

1) Input Layer

- Accepts 28x28 grayscale images of individual Sudoku cells.
- Pixel values are normalized between 0 and 1.

2) Convolution Layer

- Extract low-level features such as edges and textures.
- Example: First layer with 32 filters of size 3x3, followed by ReLU activation.
- Each convolutional layer learns progressively complex features.



System Architecture

3) Pooling Layers:

- Downsample feature maps to reduce computational complexity.
- Example: **Max Pooling** with a **2x2 kernel** reduces spatial dimensions.

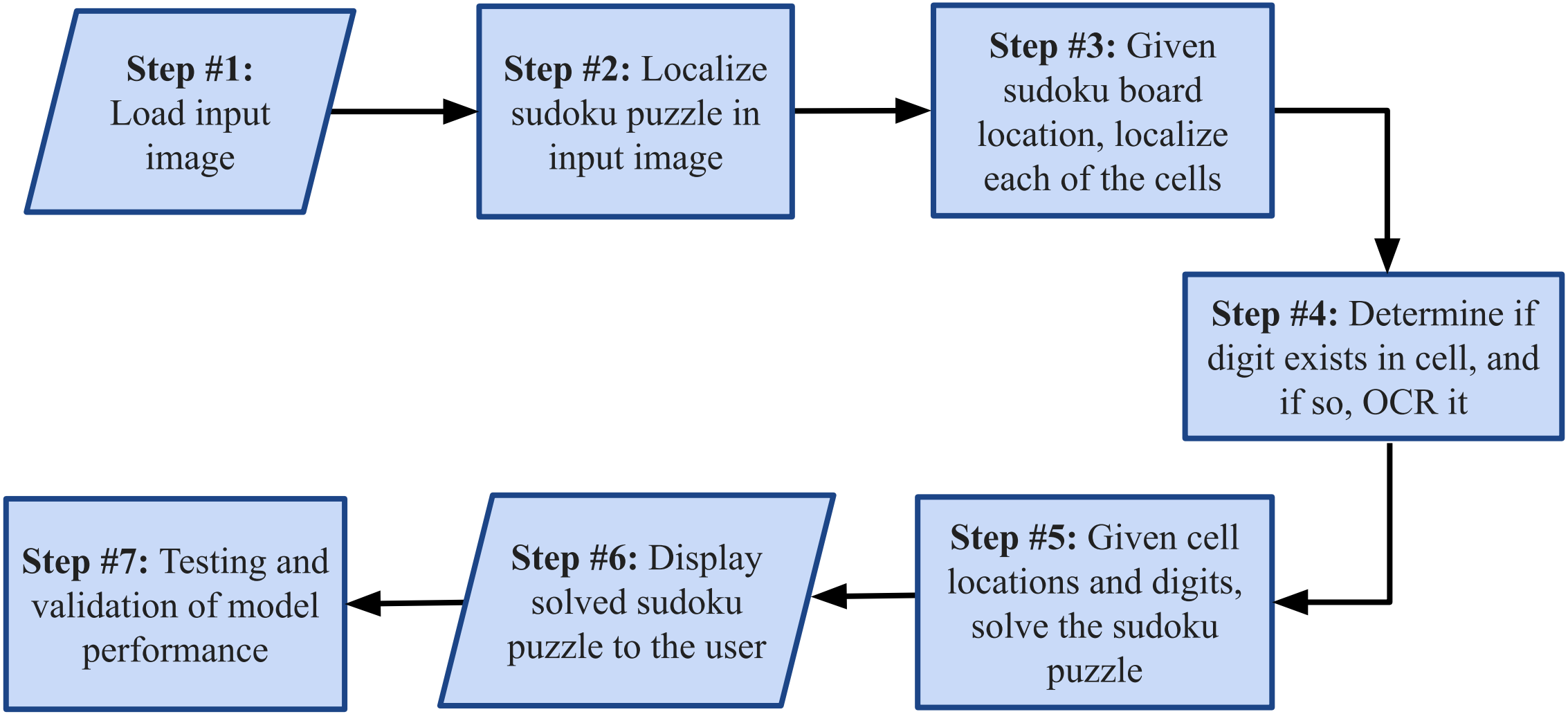
4) Fully Connected Layers:

- Flatten the feature maps into a 1D vector.
- Pass through one or more dense layers for final classification.

5) Output Layer:

- A softmax layer outputs probabilities for each digit class (0–9).
- The highest probability determines the recognized digit.

Methodology



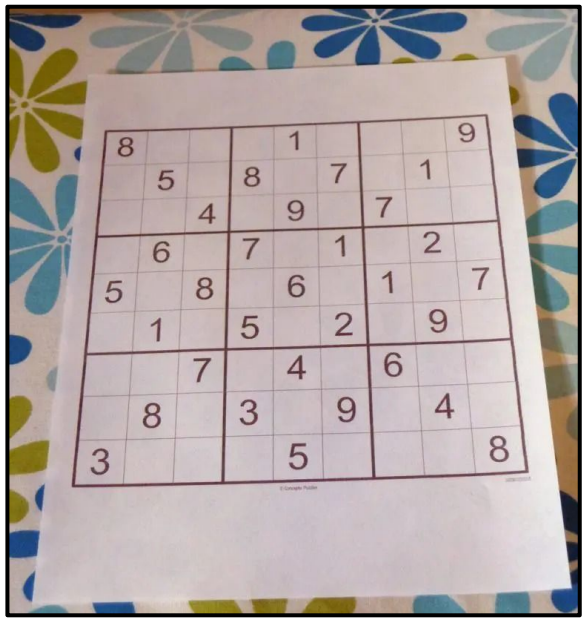
Methodology

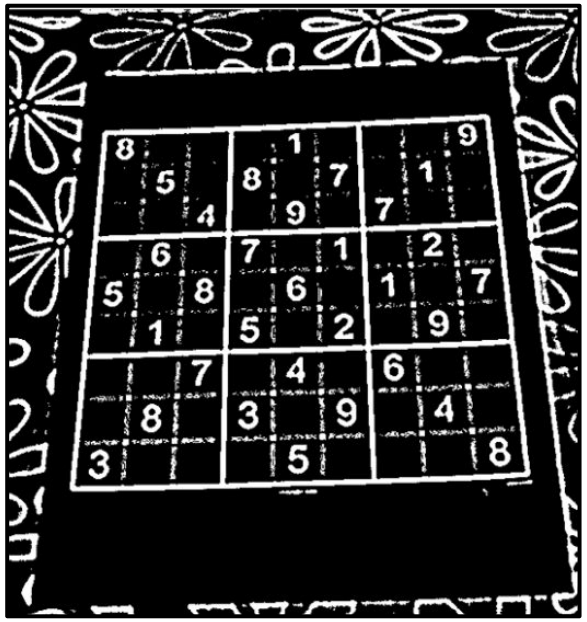
1) Load input image

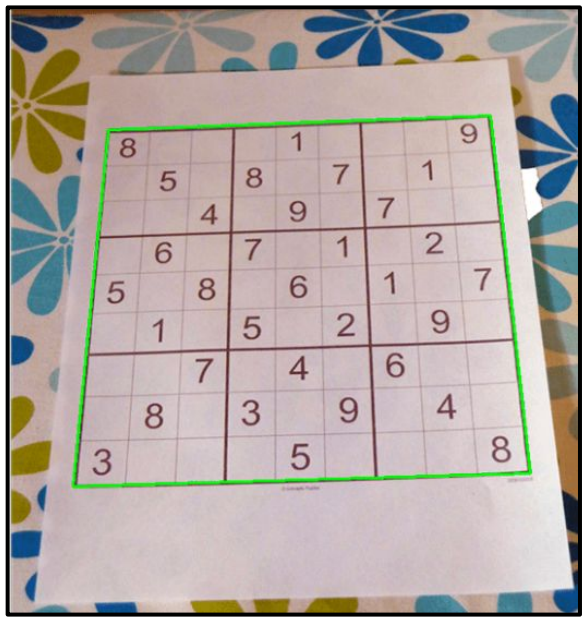
Capture & upload the image of sudoku puzzle.

2) Localise sudoku puzzle in input image *(Tool - OpenCV)*

- Convert image to grayscale, apply Gaussian blur and binary adaptive thresholding.
- Detect edges with Canny edge detection.
- The border of the Sudoku puzzle board is found by means of determining the largest contour with four points using OpenCV's contour operations.
- Apply perspective transformation to largest contour.







Methodology

3) Localise each cell *(Tool - OpenCV)*

Divide grid into 9x9 matrix (81 cells) & extract each cell for individual analysis.

4) Determine if a Digit Exists in Each Cell and Perform OCR

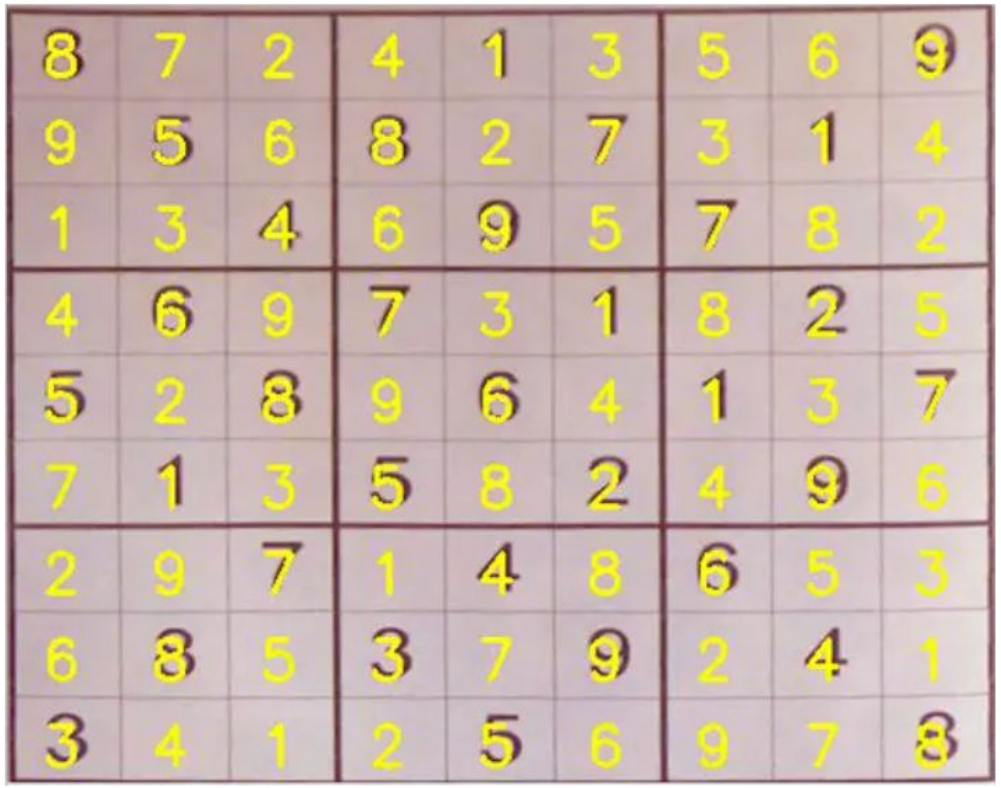
(Tool - Tesseract/TensorFlow)

Use Tesseract OCR or custom CNN to detect, recognize and store digits into the matrix.

8				1				9
	5		8		7		1	
		4		9		7		
	6		7		1		2	
5		8		6		1		7
	1		5		2		9	
		7		4		6		
	8		3		9		4	
3				5				8

A sample of the MNIST dataset on which the CNN will be trained

Methodology



5) **Solve sudoku puzzle** *(Tool - Python)*
Apply backtracking algorithm in python. It identifies empty cells, tries each number and checks validity, similarly continues recursion and backtracking.

6) **Display solved sudoku to user**
(Tool - Matplotlib)
Map solved digits back onto the grid and overlay it over the original image via OpenCV.

7) **Testing & Validation**
Test performance on multiple printed and handwritten sudokus, evaluate performance with scikit-learn and perform optimization.

Module Specification

Module 1: Data Collection and Preprocessing

Input:

Sudoku image (captured or uploaded)

Process:

- Image preprocessing using OpenCV (grayscale conversion, Gaussian blur, thresholding).
- Edge detection and cell extraction.

Output:

Straightened and segmented grid with individual cells

Module 2: Implementation of ANN/DL Algorithm

Input:

81 individual cell images (some may be blank)

Process:

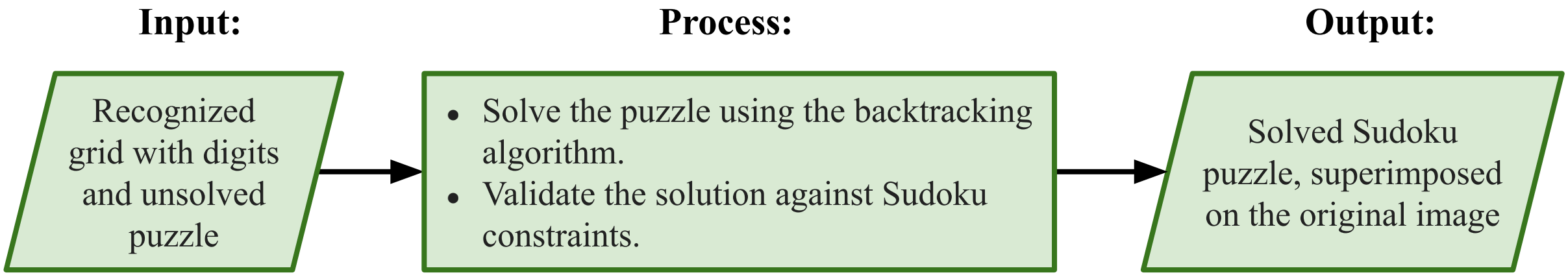
- Perform OCR using Tesseract or CNN for digit recognition.
- Train and validate CNN on MNIST dataset for high accuracy.

Output:

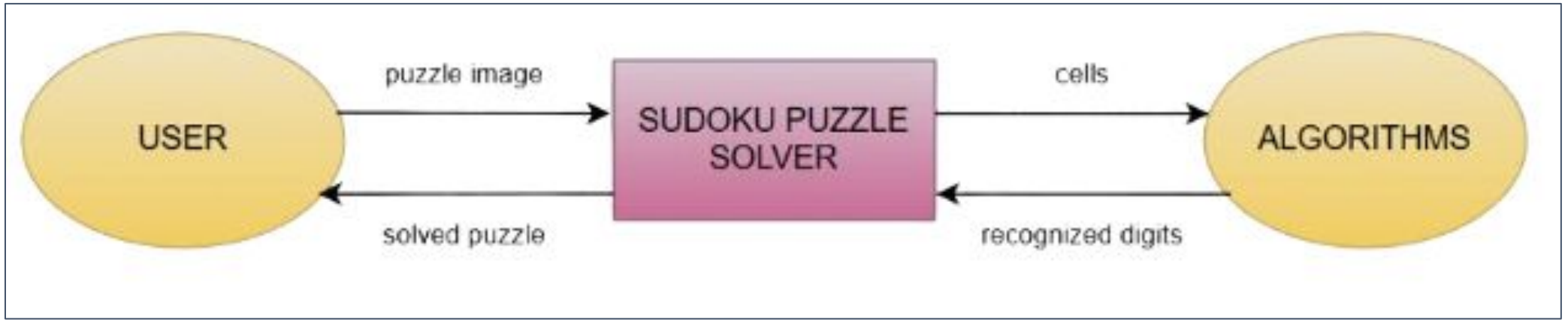
Recognized digits mapped to their respective grid positions

Module Specification

Module 3: Testing and Validation



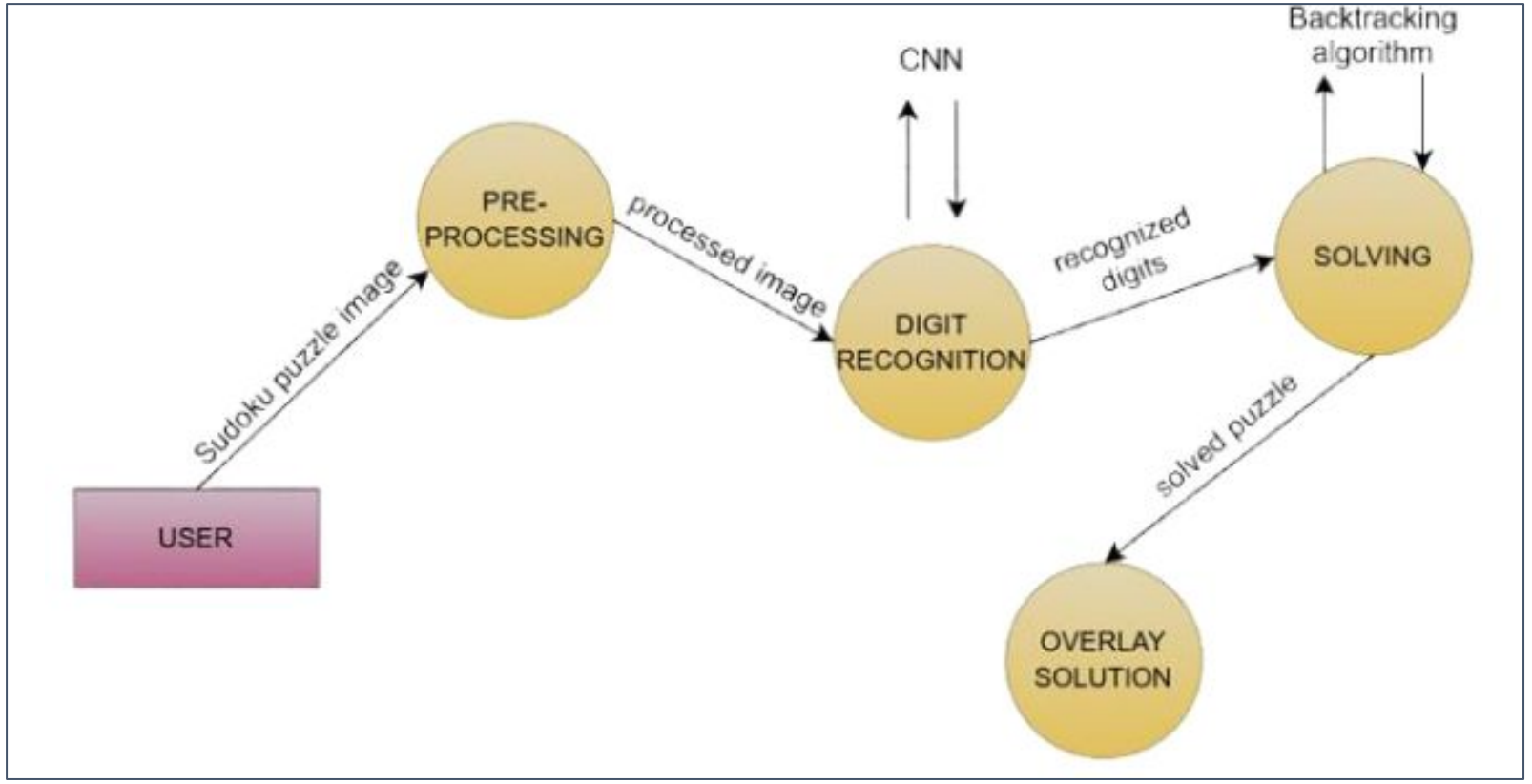
DFD Level 0



It is the highest-level Data Flow Diagram (DFD), which provides an overview of the entire system without providing any details about the internal workings of these processes.

- ➢ Entities: Users, Algorithms
- ➢ Process: Sudoku puzzle solver

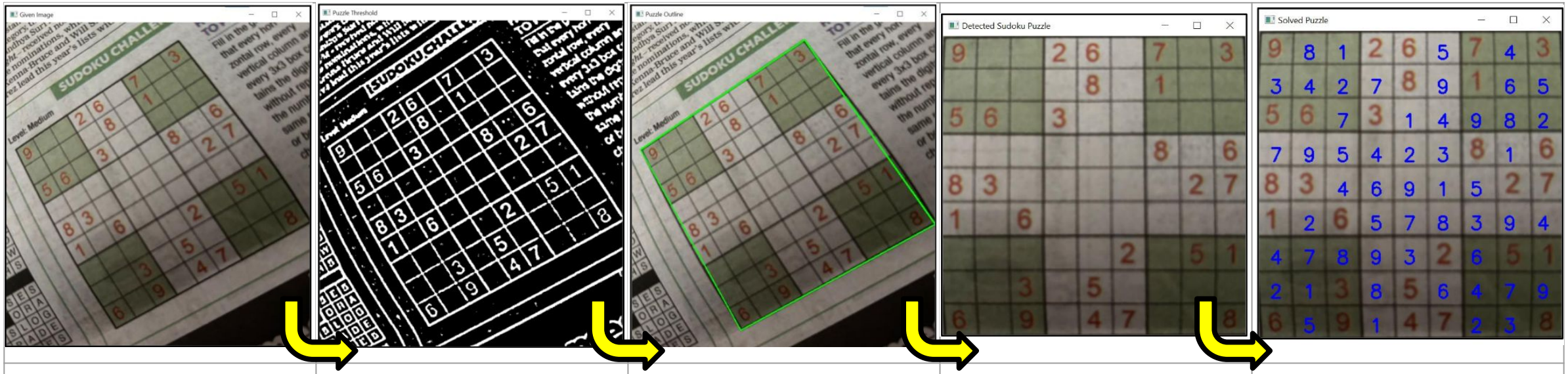
DFD Level 1



It provides a more detailed view of the system by breaking down the major processes identified in the level 0 Data Flow Diagram (DFD) into sub-processes. Each of these being represented separately.

- ➢ Entities: User
- ➢ Processes: Pre-processing, Digit recognition, Solving, Overlay solution

Implementation

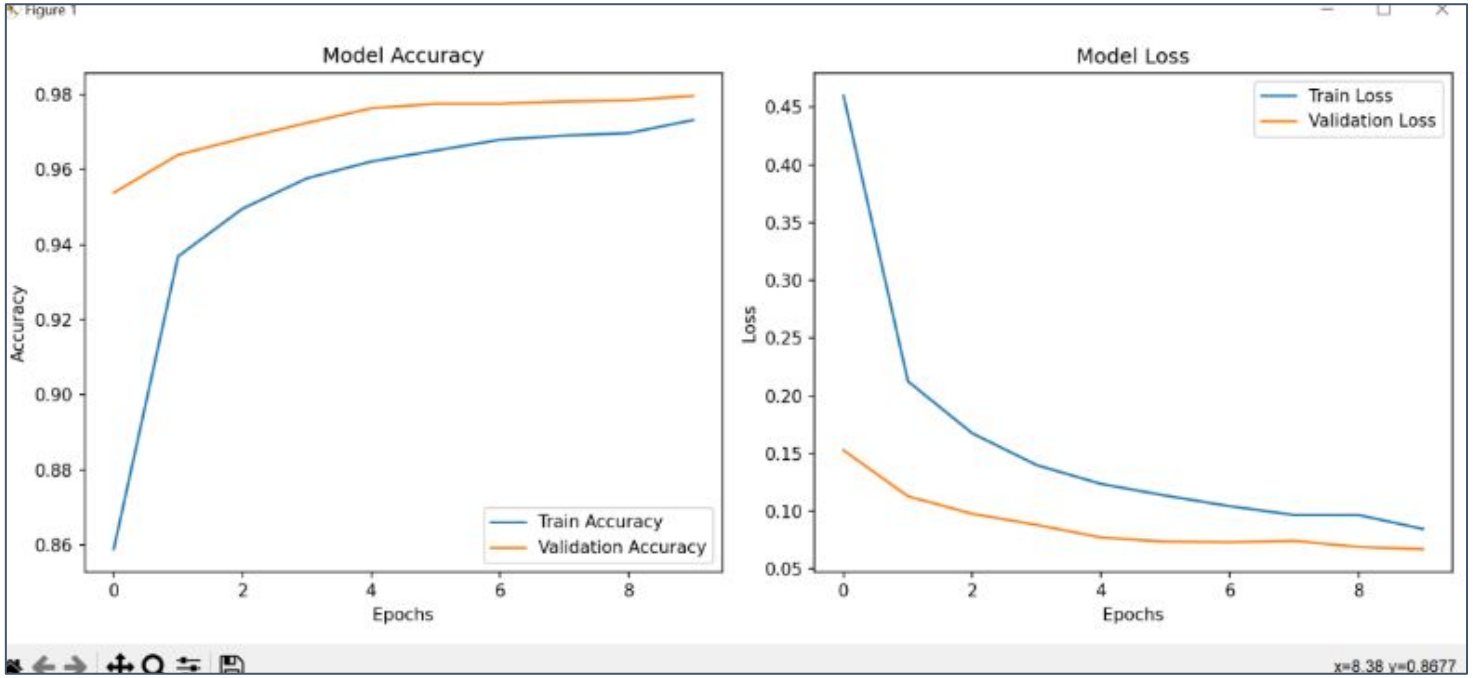


Step 1	Step 2	Step 3	Step 4	Step 5
Input image of sudoku puzzle	Image after applying gaussian blur, grayscale & adaptive thresholding to reduce noise & increase contrast	Identifying largest contour (puzzle outline)	Image after applying four point perspective transformation to obtain a top-down view. This helps extract digits	Solved sudoku puzzle with digits overlaid in correct positions on original image

Model Accuracy

- The trained digit classifier, which is based on the MNIST dataset, achieved an **accuracy of approximately 98%** on the test set. This high accuracy allowed for reliable digit recognition from the cells of the Sudoku puzzle, which was crucial for board reconstruction.

- The model correctly identified digits even in noisy or distorted images, showcasing the robustness of the CNN architecture used in the project.



Plot Showing Model Accuracy & Loss for Each Epoch During Testing

Future Scope

The proposed Sudoku solver has significant potential for further enhancement and real-world applications. Future developments may include:

- **Real-Time AR Integration**: Augmented reality (AR) overlays to display solved puzzles directly on physical Sudoku grids.
- **Mobile App Development**: Deploying the system as a smartphone app for on-the-go puzzle solving.
- **Enhanced AI Models**: Using more advanced deep learning architectures for improved digit recognition accuracy.
- **Multiple Puzzle Types**: Extending the system to recognize and solve different puzzle types beyond Sudoku.
- **Self-Learning System**: Implementing reinforcement learning techniques to improve performance over time

Conclusion

- The OCR Sudoku Puzzle Solver successfully integrates image processing, deep learning and algorithmic problem-solving to automate Sudoku solving from real-world images.
- By employing a robust CNN architecture and advanced backtracking techniques, the system achieves high accuracy in digit recognition and puzzle-solving.
- This project demonstrates the potential of combining computer vision and AI to solve real-world challenges, paving the way for future applications in similar domains.

THANK YOU