Paper ID: NCGCE086

Vudoku

A Visual Sudoku Solver

*With **Digit Recognition***

TABLE OF CONTENTS

01

OBJECTIVES

The aim and purpose of this project

Objectives

- Why digit recognition
- Review current methodologies
- Sneak peek into the new techniques
- Solve sudoku from digit recognition results
- Study how changes affect the outcome
- Derive conclusions based on the study

02

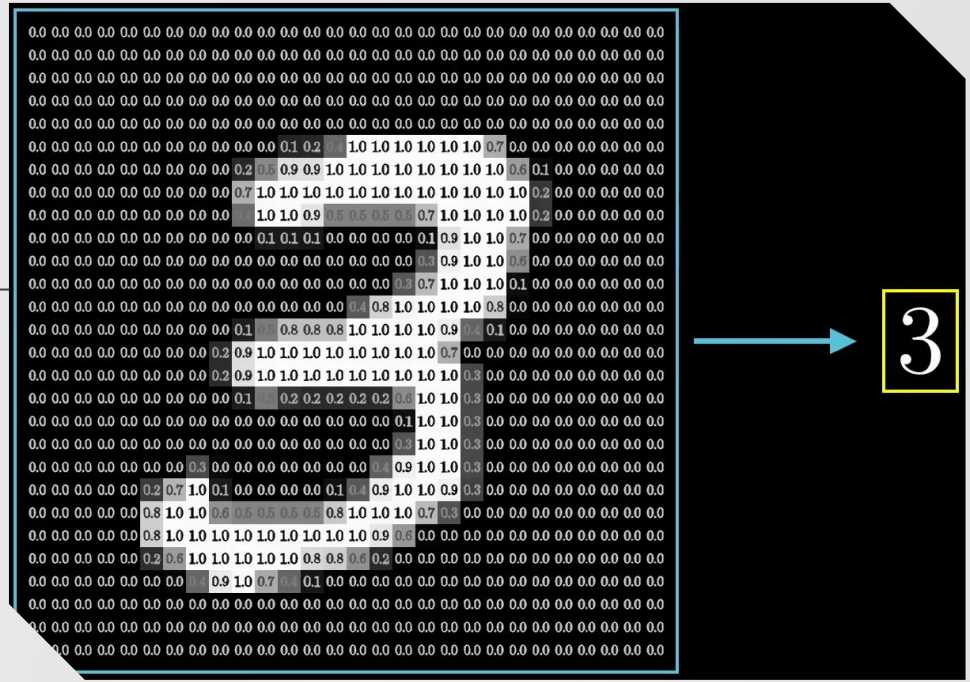
OVERVIEW

What will and won’t be covered in this review

Overview & Prerequisites

Why Digit Recognition?

- It's simple. (A classic example to begin with)
- A lot of work has been done in this field.
- Recommendation. Intuition. A gut feeling.



Overview & Prerequisites (Cont.)

Fourier Transform

A mathematical transform that decomposes a function (often a function of time, or a signal) into its constituent frequencies

It's first-order iterative optimization algorithm for finding a local minimum of a differentiable function

Gradient Descent

Activation Function

Defines the output of that node (usually a neuron) given an input or set of inputs. Sigmoid, ReLU etc.

Overview & Prerequisites (Cont.)

What part of the input should I focus on? Relevance of one entity with respect to others entities

Attention

Back Propagation

Is an algorithm, used to train feedforward neural networks for supervised learning

To find minimum cardinality subset of an arbitrary bipartite graph, with sets and the universe representing the inclusion of elements in sets

Hitting Set Problem

03

RESEARCH

Includes the research work for the topic

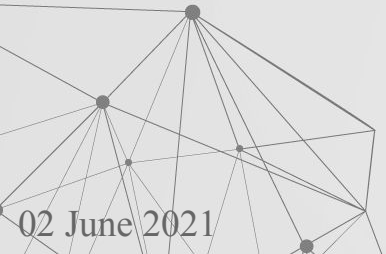
Evolution Of Neural Networks

1986 — Recurrent Neural Network

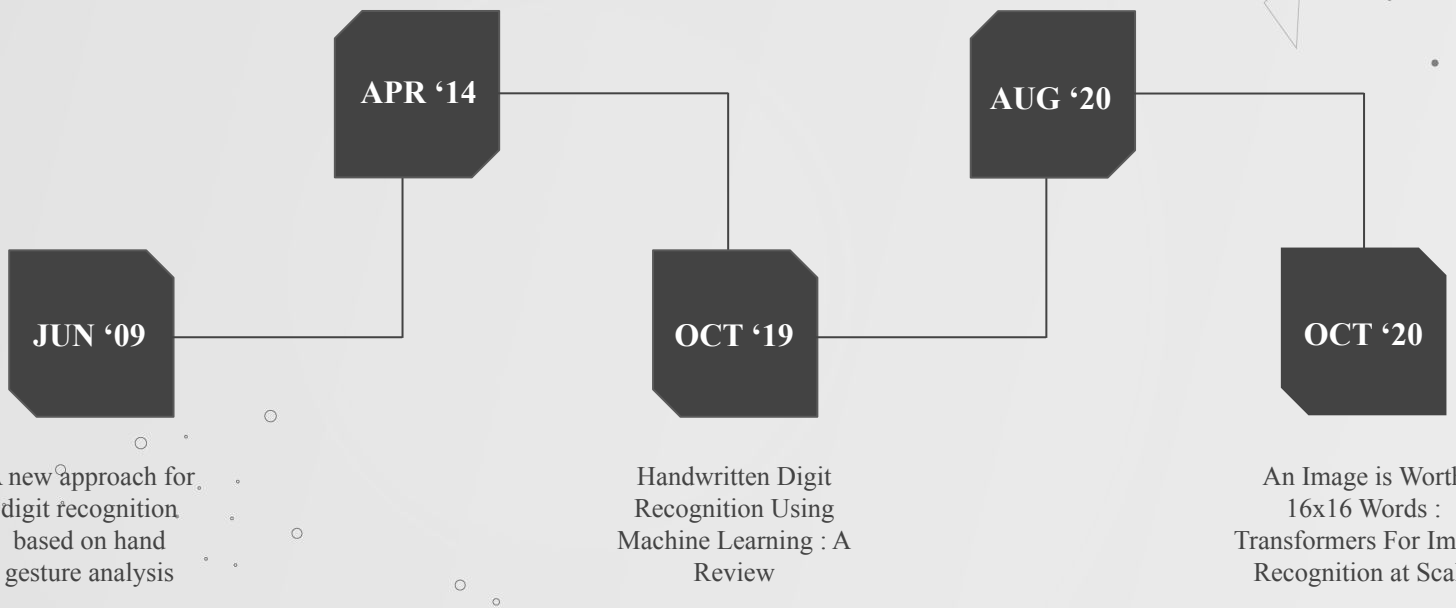
1997 — Long Short Term Memory

1998 — Convolution Neural Network

2017 — Transformer Network



Literature Review (Cont.)



Literature Review (Cont.)

A new approach for digit recognition based on hand gesture analysis

27 June 2009

Ahmed Ben Jmaa | Walid Mahdi | Yousra Ben Jemaa | Abdelmajid Ben Hmadou

- Analysis is based on extracting a set of features from a hand image and then combining them by using an induction graph.
- Apply a fuzzy classifier to identify the skin pixels.
- Remove all the hand components except the fingers.
- The final step consists on representing histogram of the detected fingers in order to extract features that will be used for digit recognition.

Literature Review (Cont.)

Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks

14 April 2014

Ahmed Ben Jmaa | Walid Mahdi | Yousra Ben Jemaa | Abdelmajid Ben Hmadou

- Traditional approaches is to separate out the localization, segmentation, and recognition steps.
- Unified approach that integrates these three steps via the use of a deep convolutional neural network that operates directly on the image pixels.
- Employs the DistBelief implementation of deep neural networks in order to train large, distributed neural networks on high quality images.

Literature Review (Cont.)

Handwritten Digit Recognition Using Machine Learning : A Review

18 October 2019

Anchit Shrivastava | Isha Jaggi | Sheifali Gupta | Deepali Gupta

- Claims to create a base for future researches in the area so that the researchers can overcome the existing problems.
- Focuses on digit recognition using multilayer neural network.
- It mainly comprises of three phases that is pre-processing phase, feature extraction and then the final phase of classification.

Literature Review (Cont.)

An ensemble of simple convolutional neural network models for MNIST digit recognition

12 Aug 2020

Sanghyeon An | Minjun Lee | Sanglee Park | Heerin Yang | Jungmin So

- They report a very high accuracy on the MNIST test set that can be achieved by using simple convolutional neural network (CNN) models.
- Uses three different models with different kernel size in the convolution layers, each followed by a single fully connected layer.
- A two-layer ensemble, a heterogeneous ensemble of three homogeneous ensemble networks, can achieve up to 99.91% test accuracy.

Literature Review (Cont.)

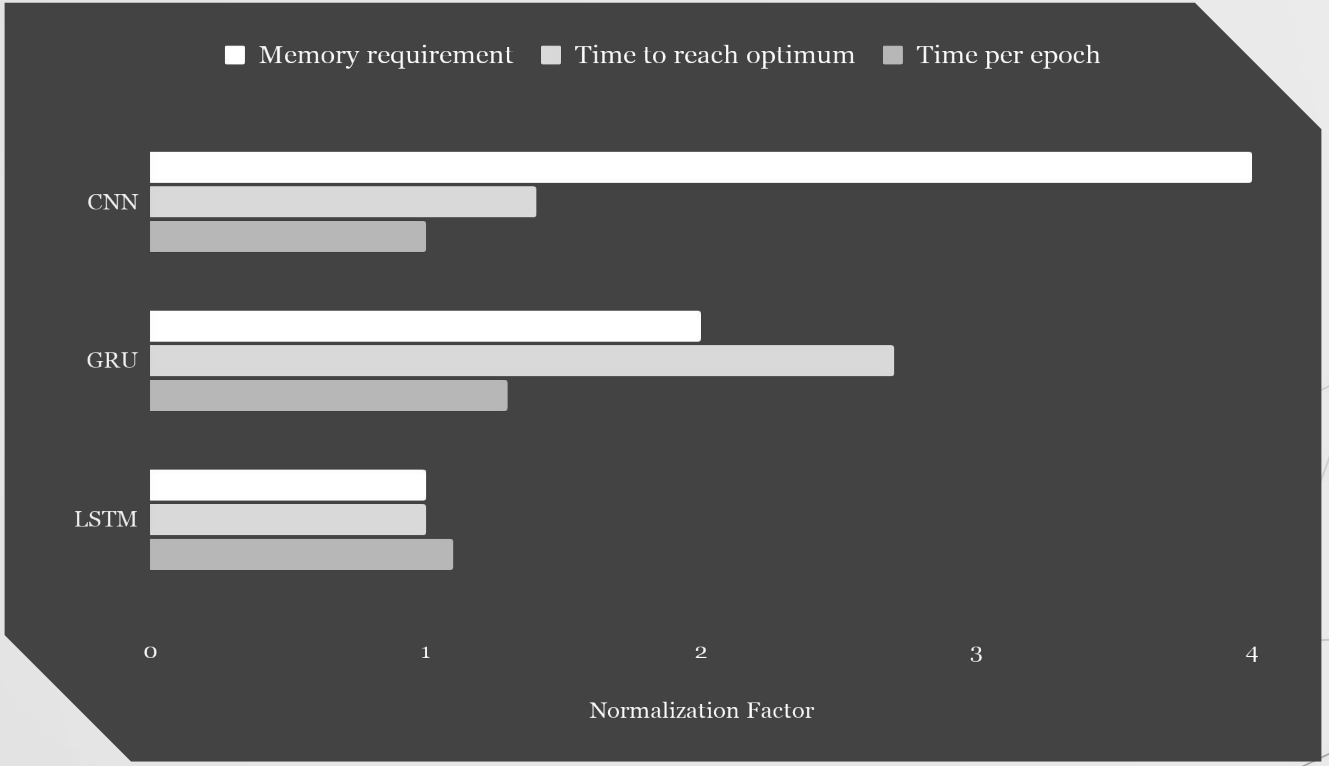
An Image is Worth 16x16 Words : Transformers For Image Recognition at Scale

22 Oct 2020

Alexey Dosovitskiy | Lucas Beyer | Alexander Kolesnikov | Dirk Weissenborn | Xiaohua Zhai

- Claim reliance on CNN's is not necessary and a pure transformer applied directly to sequences of image patches can perform.
- Vision Transformer (ViT) attains excellent results compared to state-of-the-art convolutional networks.
- Requires substantially fewer computational resources to train.

Cost Analysis

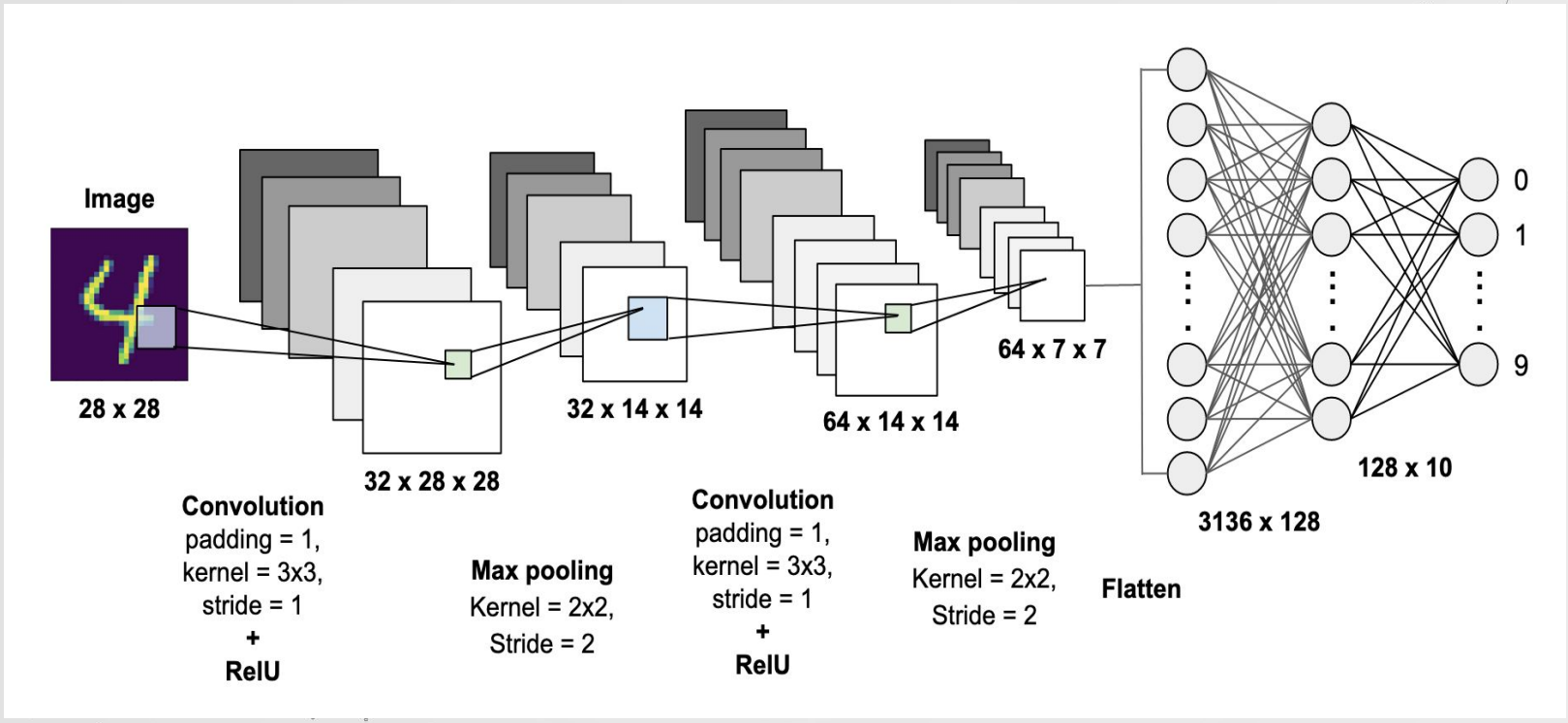


04

WHAT'S NEW?

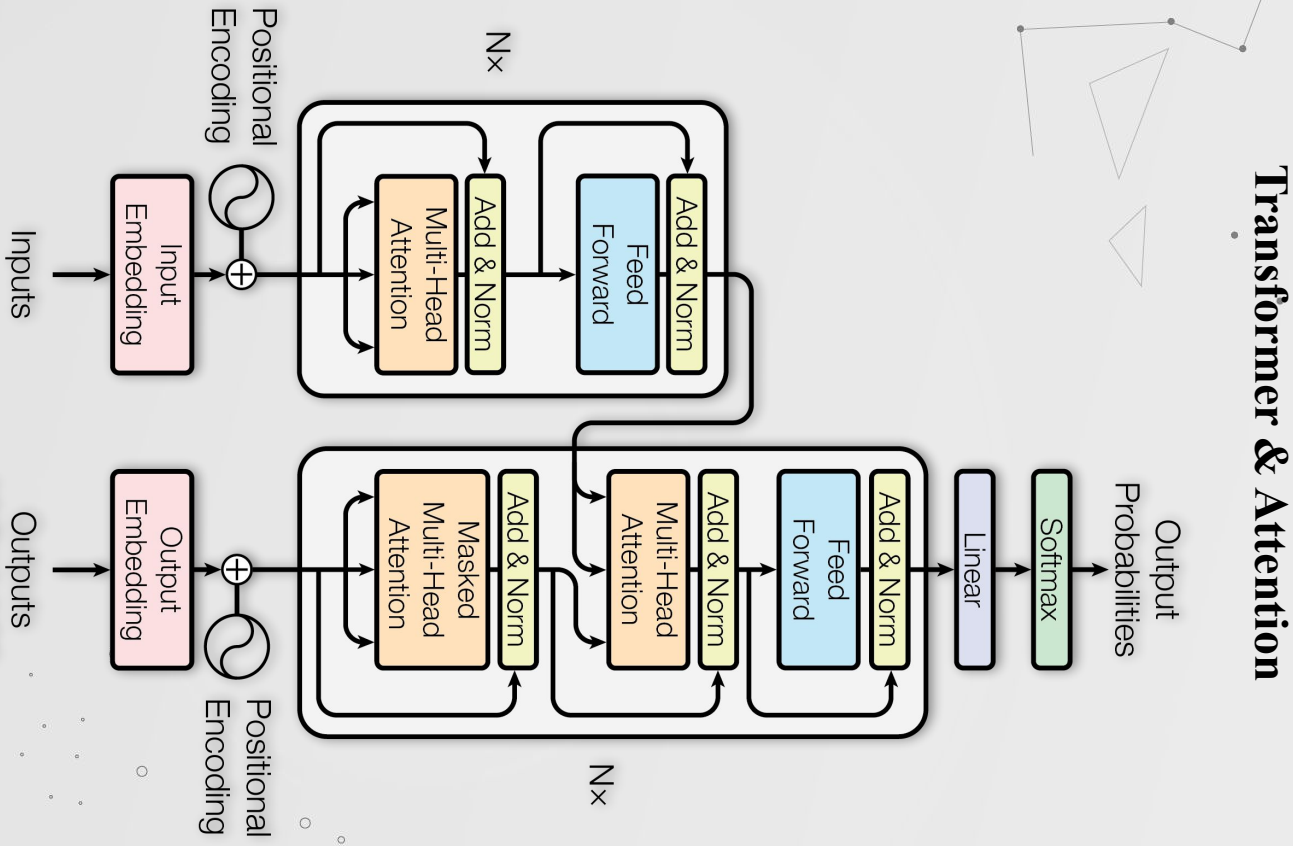
Introducing novelty in this project

Architecture - Part 1 : CNN

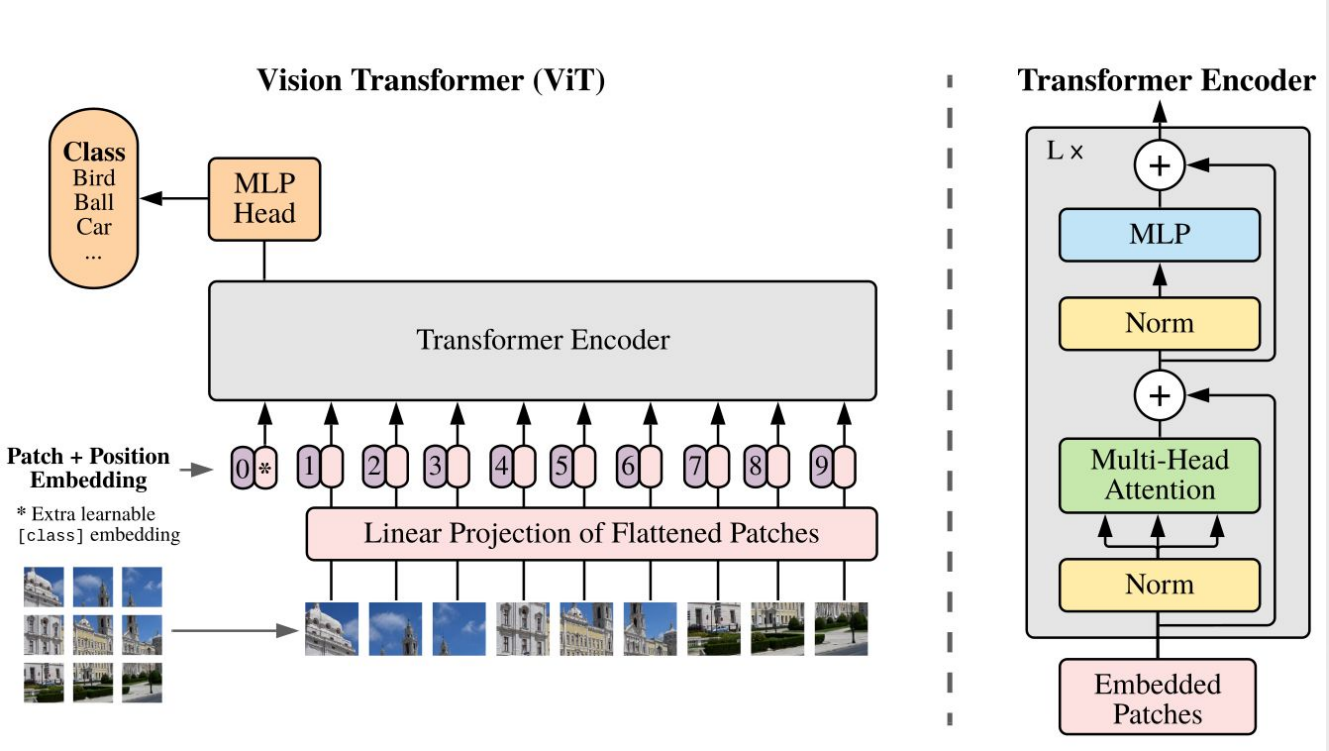


Convolution Neural Networks

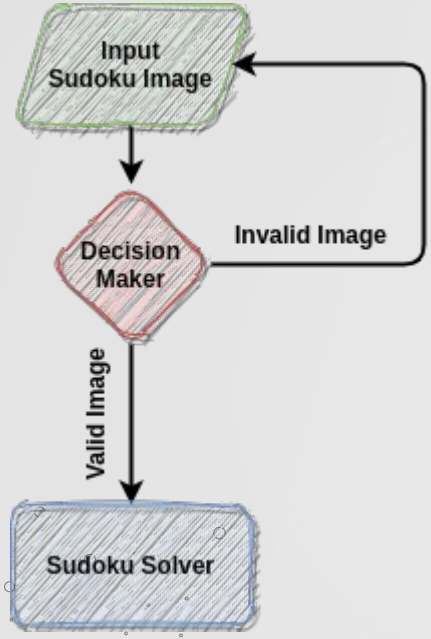
Architecture - Part 2 : Components

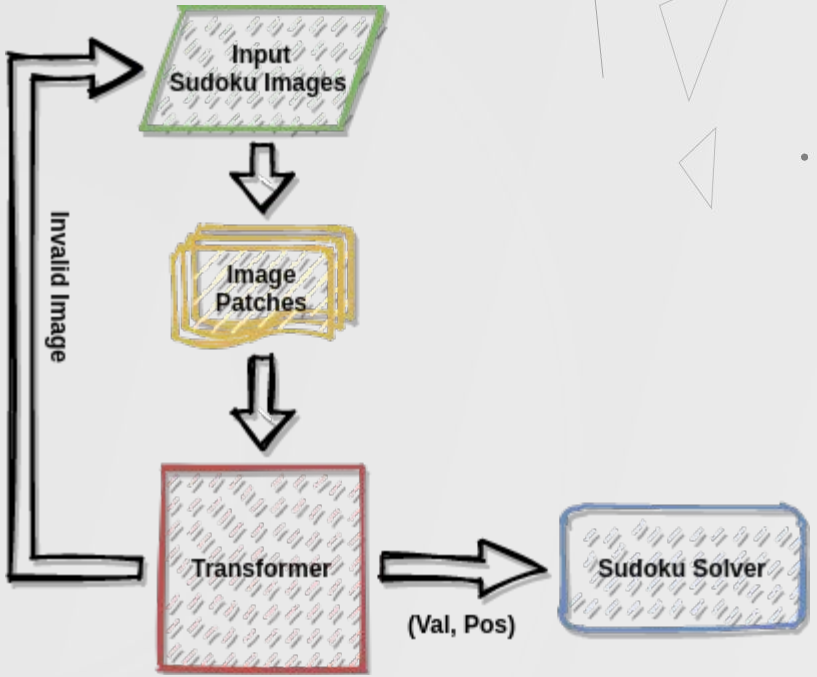


Architecture - Part 2 : Components (Cont.)

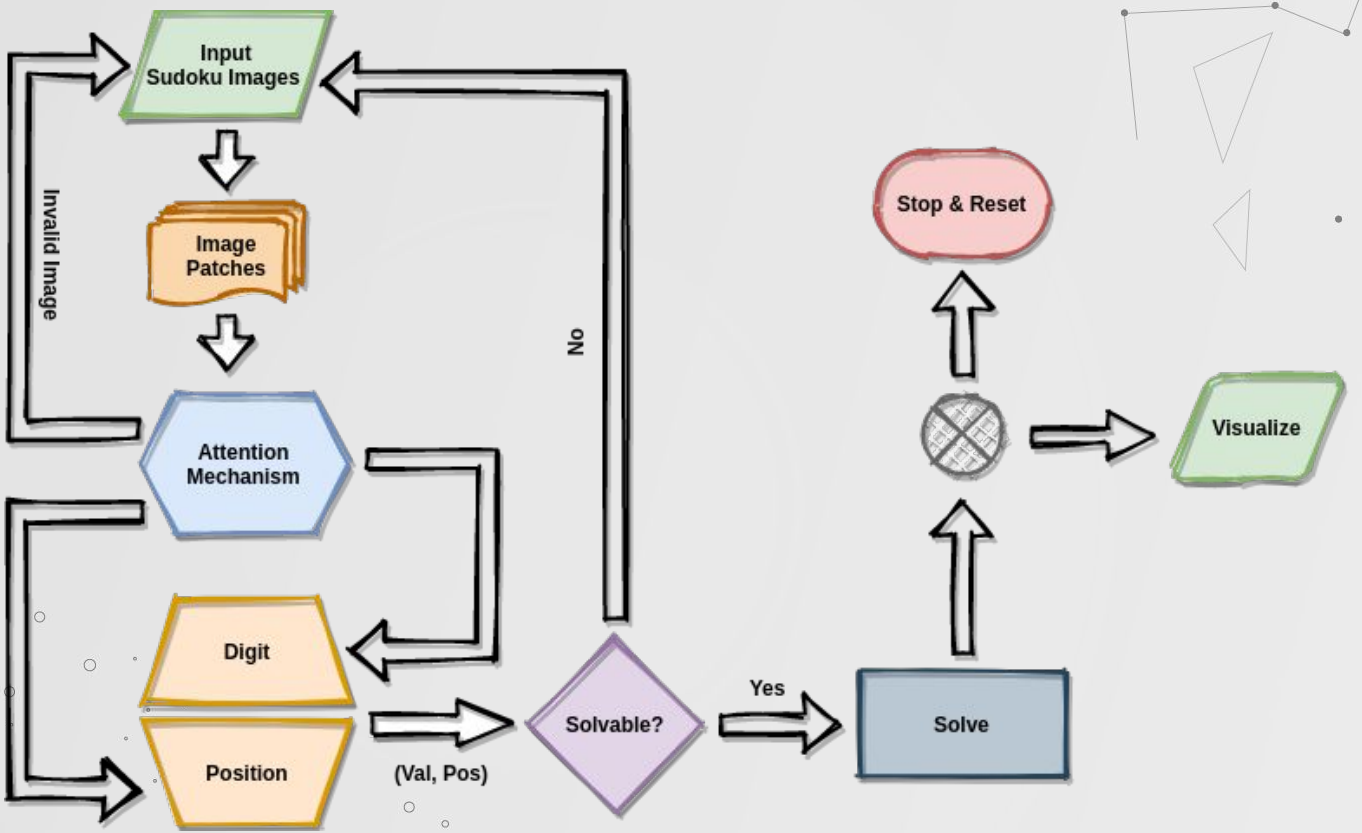


Architecture - Part 3 : Pass 1 & 2





Architecture - Part 3 : Final Pass



Project Modules

- Convoluting
  - Edge Detection
  - Digit Recognition
- Image Patching
- Encoder & Decoder
- Attention Mechanism
  - Edge / Box Detection
  - Digit Recognition

Artificial Intelligence

- Algorithms
  - Backtracking (or)
  - Stochastic Search (or)
  - Exact Cover
- Visualization

Solving Sudoku

Project Path



Input Image **Preprocessing** **Attention**



Solving Sudoku **Visualizing** **Product**

05

REQUIREMENTS

Necessary software and hardware entities

Dependencies

Hardware

- **CPU:** 2+ Cores, 2.46+ GHz
- **RAM:** 6 - 8 GB
- **Disk:** 5+ GB free space
- **Camera / Scanner**

If the system does not meet these requirements, then any system which can run a modern browser smoothly is recommended - to work on Colabs.

Software

- **Code Editor:** VSCode/ium
- **Language(s):** Python/Nim
- **Libraries (primary):**
  - OpenCV
  - Tensorflow / Keras
  - Numpy
  - Pillow
- **Dataset:** MNIST

06

EFFECTUATION

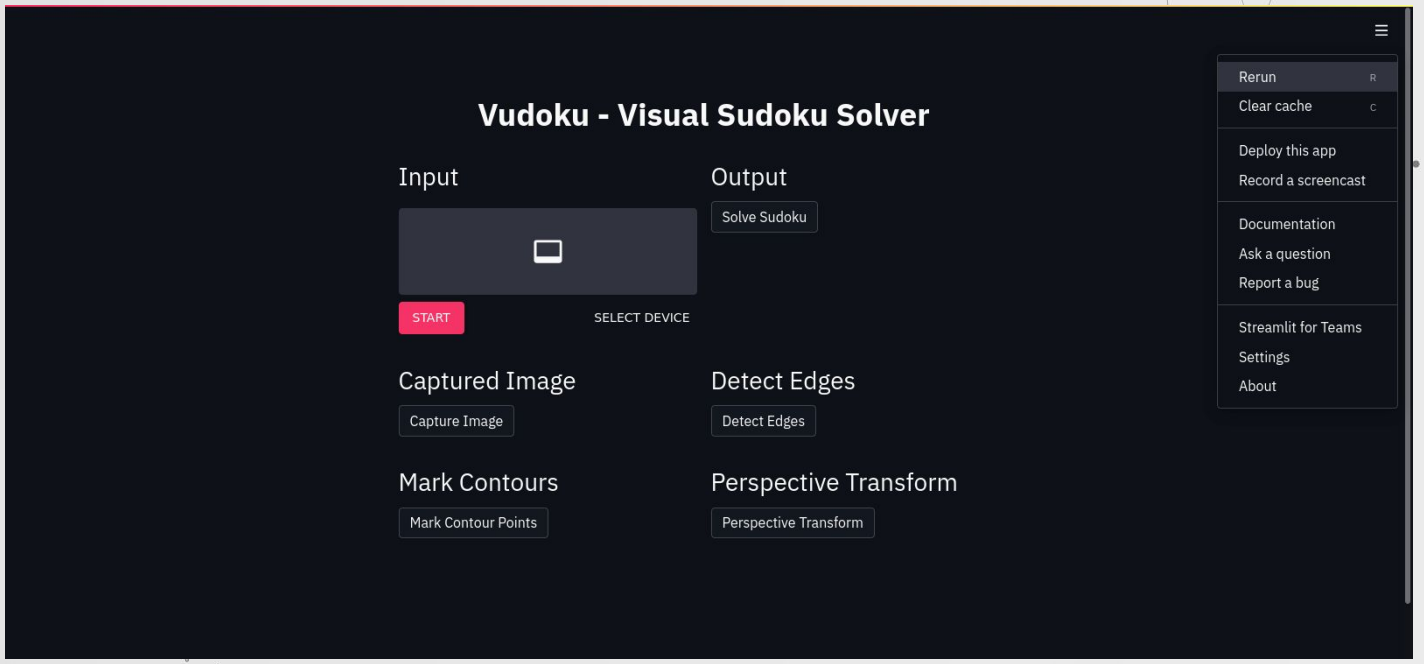
Putting ideas and things into effect

Source Modules

2D sequential image recognition model in Keras	Grabs the sudoku image and performs preprocessing	Digit extraction from grid via contour transform
classifier.py	**scavenger.py**	**extractor.py**

Solves 81-bit long encoded string using backtracking	Builds back an image of solved sudoku grid	Controls the overall process with streamlit
solver.py	**builder.py**	**main.py**

Interface Screenshot



07

CONCLUSION

Concluding remarks about the research

Closing Thoughts

When the Transformer was first released back in 2017 by the Google-Brain team, it got a lot of traction from the research community. The OpenAI created three successive versions of GPT, which took the media by the wind, was based on the same Transformer. Not a month ago Google-Research released their Vision Transformer (ViT), which is at the heart of Vudoku. Visualizing sudoku will give an insight on a lot many aspects, including - accuracy, performance, precision, recall etc.

Digit recognition is an interesting classical problem to embark the journey in the field of artificial intelligence. Sudoku, a mathematical puzzle is captivating beauty of the mystery that lies in various patterns and combinations in itself. This project could be an inspiring stepping stone for futures enthusiasts to work upon. On a personal note, I believe this project has been a crescendo of learning adventure.

References

Papers

1. Ben Jmaa, A., Mahdi, W., Ben Jemaa, Y., and Ben Hamadou, A., “A new approach for digit recognition based on hand gesture analysis”, *arXiv e-prints*, 2009.
2. Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., and Shet, V., “Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks”, *arXiv e-prints*, 2013.
3. Shrivastava, Anchit & Jaggi, Isha & Gupta, Sheifali & Deepali, Gupta. (2019). Handwritten Digit Recognition Using Machine Learning: A Review. 322-326. 10.1109/PEEIC47157.2019.8976601. [base]
4. An, S., Lee, M., Park, S., Yang, H., and So, J., “An Ensemble of Simple Convolutional Neural Network Models for MNIST Digit Recognition”, *arXiv e-prints*, 2020. [base]
5. Dosovitskiy, A., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale”, *arXiv e-prints*, 2020. [base]

References (Cont.)

Internet

1. Artificial neural network
2. The Illustrated Transformer – Jay Alammar – Visualizing machine learning one concept at a time
3. Meta learning
4. Visualizing A Neural Machine Translation Model (Mechanics of Seq2seq Models With Attention)
5. Transformer Neural Networks - EXPLAINED! (Attention is all you need)
6. How to code The Transformer in PyTorch
7. Attention Mechanism
8. Conv Nets: A Modular Perspective
9. Understanding Convolutions - colah's blog
10. AliShazly/sudoku-py: An augmented reality sudoku solver using OpenCV
11. google-research/vision_transformer

Project Influencers

Grant Sanderson

Software Engineer | Mathophila | TEDx Speaker | 3Blue1Brown

Parul Pandey

Data Scientist / Evangelist at H2O.ai | Kaggle GrandMaster (Notebooks)

Harrison Kinsley

Programmer | Race car driver | Entrepreneur | Teacher | Sentdex

Michael Nielsen

Quantum Physicist | Science Writer | Computer Programming Researcher

Thank You

Please feel free to ask questions

Ebin P M (Guide)
https://itsmeebin.wordpress.com/
pm.ebin74@gmail.com

Jovial Joe Jayarson
https://joe733.github.io/profile
jovial7joe33@gmail.com