CSE472

SOLVING VISUAL JIGSAW PUZZLE USING DEEP LEARNING

Presented By :

1805072 1805079

9 January, 2024

BASE PAPER

- **Title:** Deepzzle: Solving Visual Jigsaw Puzzles with Deep Learning and Shortest Path Optimization
- **Publication:** IEEE Transactions on Image Processing (2020)
- **Date of Publication:** 26 May 2020
- **Link:** https://arxiv.org/abs/2005.12548

CODE & DATA AVAILABILITY

1 **Code:**

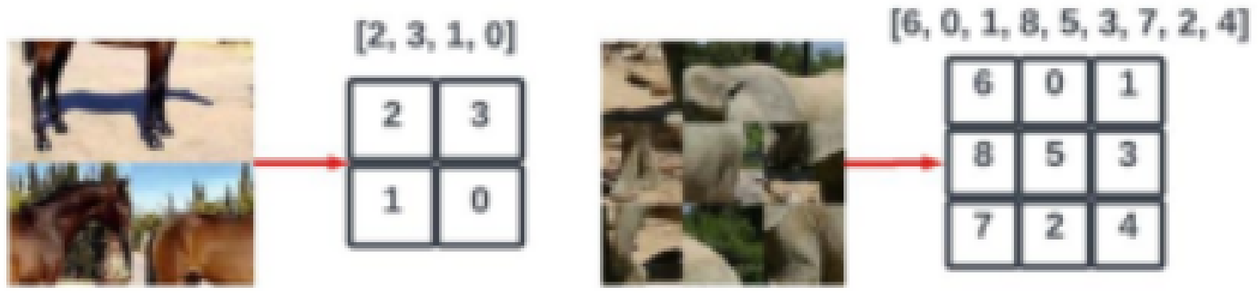
At present, we are not aware of any specific code repositories associated with this project.

2 **Dataset:**

https://www.kaggle.com/datasets/shivajbd/jigsawpuzzle

DATASET

- 100K images in the form of a Jigsaw puzzle.
- There are two types of puzzle data, 2x2 and 3x3.
- Each image has 4/9 pieces which are randomly positioned.
- The label is an array of 4/9 integers, which represents the correct position of the piece in the puzzle.



BASE PAPER

Problem Definition

The task is to find out the correct positioning of all pieces using neural networks. The base paper tried to automatically reassemble jigsaw puzzles with deep learning even when the fragments have wide spaces between them, are eroded, or come from different sources.

Key Idea

Use a two-step method:

1. A neural network predicts the relative positions of the fragments
2. A graph is used to find the best reassembly that maximizes the joint probability of the predictions.

BASE PAPER

Solution Overview

Deep Learning: Siamese network architecture, which is a type of neural network that consists of two identical subnetworks is used for the task of solving jigsaw puzzles.

Graph Theory: The graph is constructed based on the predicted probabilities. The graph is then searched for the path that maximizes the joint probability of the predictions.

Key Results

Solving jigsaw puzzles that can outperform other methods in terms of accuracy and speed.

PROPOSED EXPERIMENT

Models To Be Evaluated:

- Fully Connected Neural Network with Softmax Loss Function
- Convolution Neural Network (VGGNet Architecture)
- Convolution Neural Network (ResNet Architecture)

Evaluation Metrics

- Accuracy
- Precision
- FI Score
- Recall

METHODOLOGY

1. Training

- CNN models are trained on a dataset of jigsaw puzzles.
- The model learns to predict the correct placement of each piece based on its visual features.

2. Prediction:

The model outputs a probability distribution over the possible placements of the piece.

3. Refinement

The initial solution may contain errors, especially for challenging puzzles. By iteratively re-predicting the placements of pieces that are inconsistent with their neighbors, an optimal solution can be found .