UNSUPERVISED PRE-TRAINING: SOLVING JIGSAW PUZZLES VIA NEURAL COMBINATORIAL OPTIMIZATION

Lucio Dery Robel Mengistu Oluwasanya Awe

BACKGROUND

There is a combinatorial explosion in the solution set of even a small Jigsaw puzzle ((3x3 = 9) ! = 362880). This makes solving Jigsaw puzzles an interesting vision task. The Jigsaw Puzzle Problem has potential applications in:

1. recovering shredded documents and
2. reconstructing archaeological artifacts.

Sholomon et. al. provide a Genetic Algorithm that achieve accuracies around 80-90%. This approach is slow at test time. **Noroozi and Favaro** also showed that a CNN trained to solve Jigsaw puzzles can be useful for unsupervised pre-training. However, the trained solver does not generalize to puzzles of higher dimensions due to factorial solution-set growth.

PROBLEM STATEMENT

Our Jigsaw Solver will take as input an arbitrary image split into $N = W \times H$ non-overlapping, equally-sized pieces in the same orientation as the original image. Specifically, we can define input,

$$I = (c_0, c_1, ..., c_N)$$

where I is the flattened representation of the N pieces. It will return an ordering $(x_1, x_2, ..., x_N) = \sigma(1, 2, ..., N)$ such that x_i is the index of the image that should be at position i in the flattened representation of the correct image. Using this ordering, we can reconstruct and visualize the original image.

MODELS

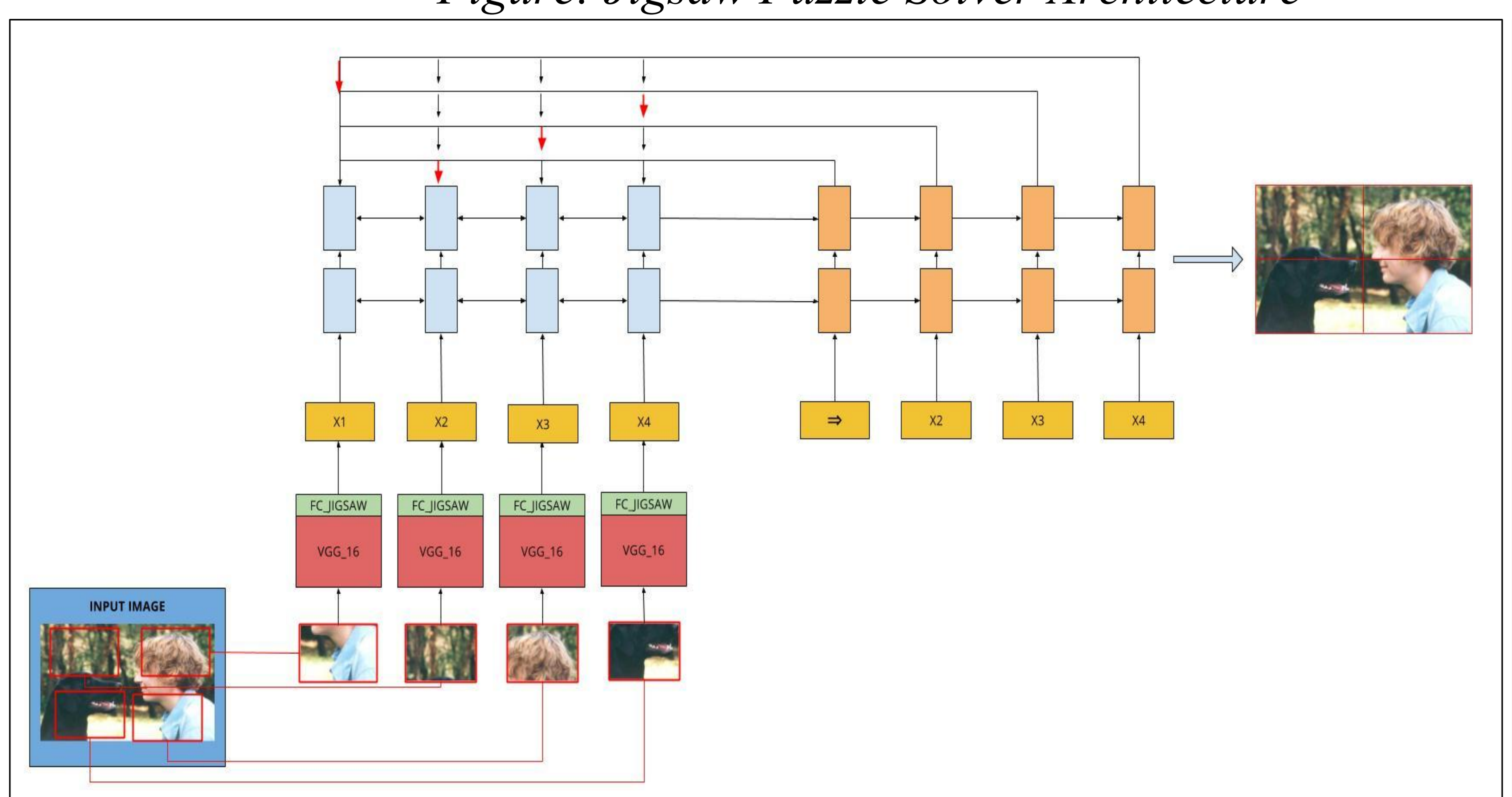
Preprocessing

1. Resize to [H x (64 + ϵ)] x [W x (64 + ϵ)] x D
2. Divide into H x W cells
3. Sample random 64 x 64 patch

Post Processing

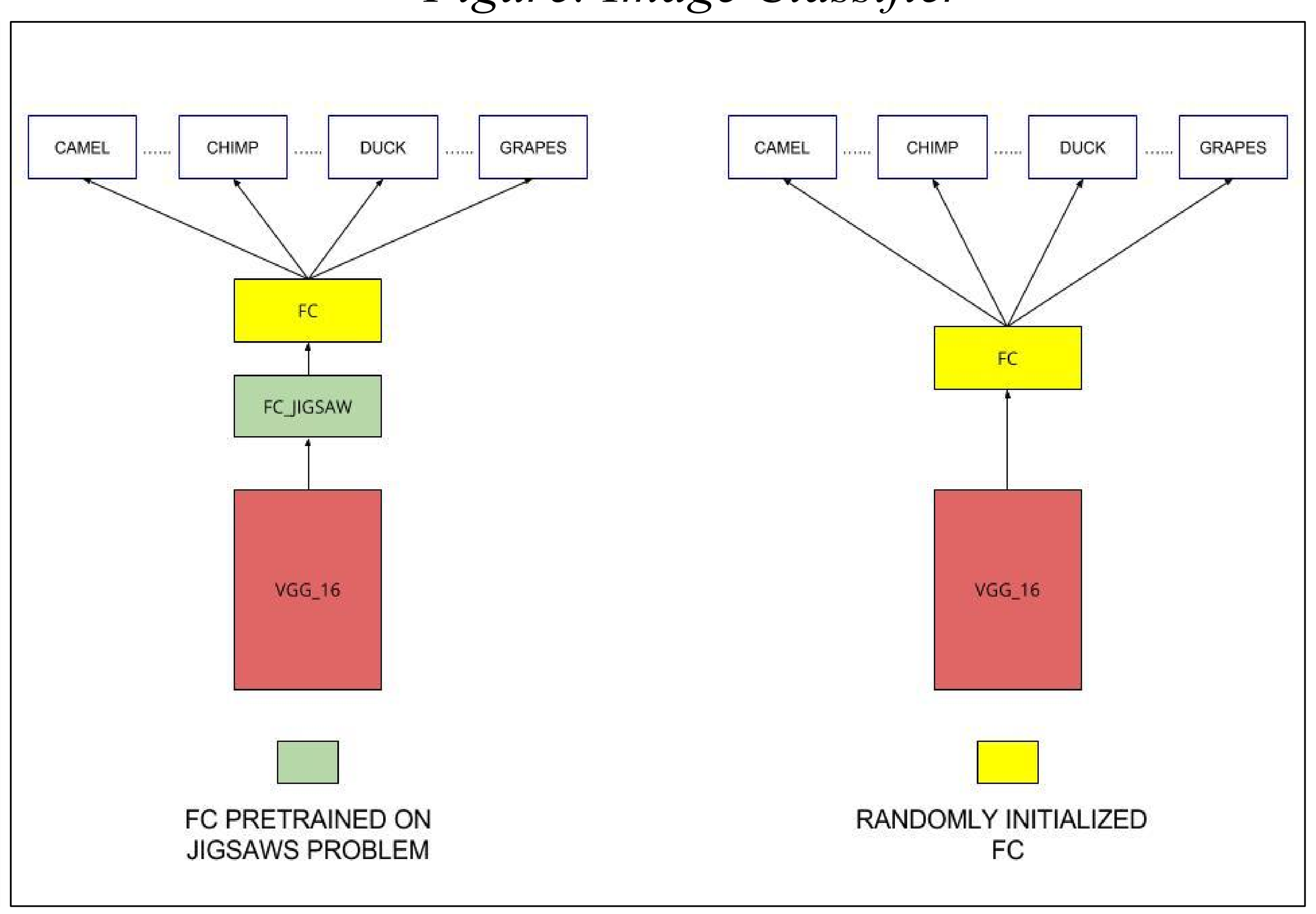
We force network to make unique predictions at every timestep by setting logits of already predicted indices to -∞.

Figure: Jigsaw Puzzle Solver Architecture



(HxW)x64x64 ⇒ VGG Features
⇒ Jigsaw Embedder ⇒ Pointer Network Encoder ⇒ Pointer Network Decoder
⇒ Predicted Order

Figure: Image Classifier

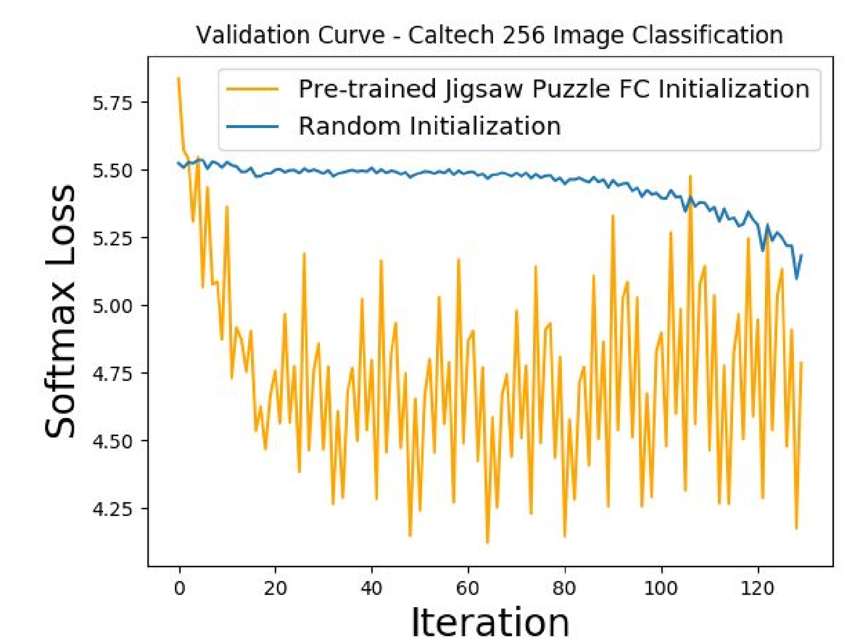


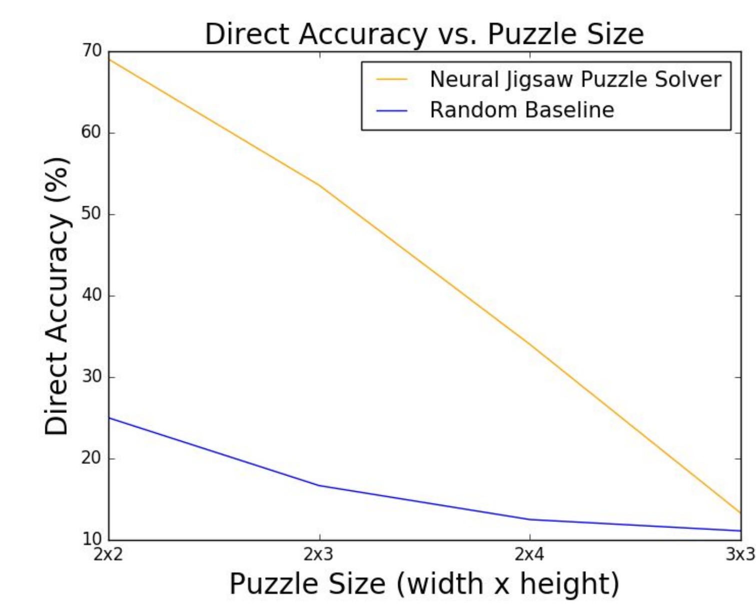
DATASET

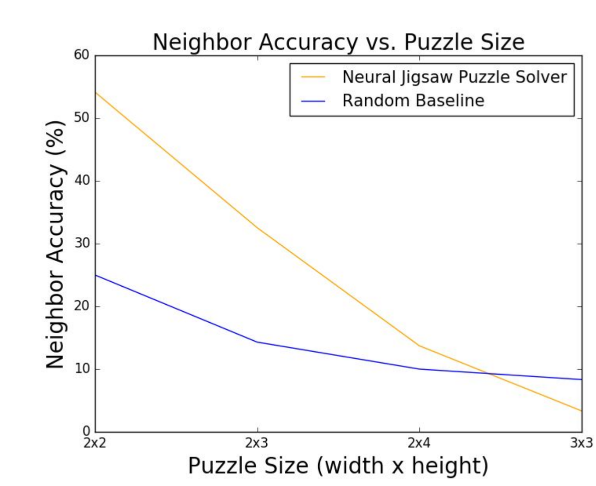
Caltech256 Image Dataset (2006)
92,652 Images from Google and Picsearch
256 Object Categories
At least 80 images per category
Train = 10240, Val = 1024, Test = 512

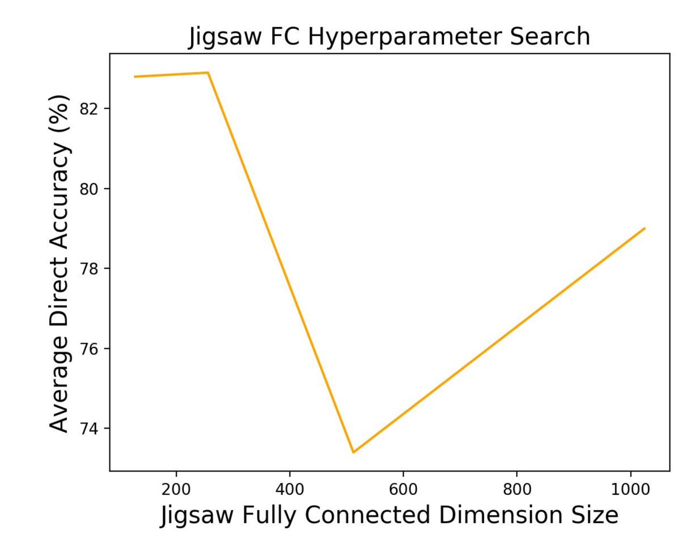
EVALUATION AND FINDINGS

Performance on Different Puzzle Size









FUTURE DIRECTIONS

1. **Reinforcement Learning**
Using edge compatibility as a reward function and training the network via policy gradient, effectively removing the need for labels.
2. **Set2Seq**
Allows us to disregard the order of our input. And thus the probabilistic effect of input order on our output.

REFERENCES

D. Sholomon, O. David, and N. S. Netanyahu. A genetic algorithm-based solver for very large jigsaw puzzles. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2013.

M. Noroozi and P. Favaro. Unsupervised learning of visual representations by solving jigsaw puzzles. In European Conference on Computer Vision, pages 69–84. Springer, 2016.