

Visual Sudoku Puzzle Classification: A Suite of Collective Neuro-Symbolic Tasks



Eriq Augustine¹, Connor Pryor¹, Charles Dickens¹, Jay Pujara², William Yang Wang³, Lise Getoor¹

¹ University of California, Santa Cruz

² University of Southern California

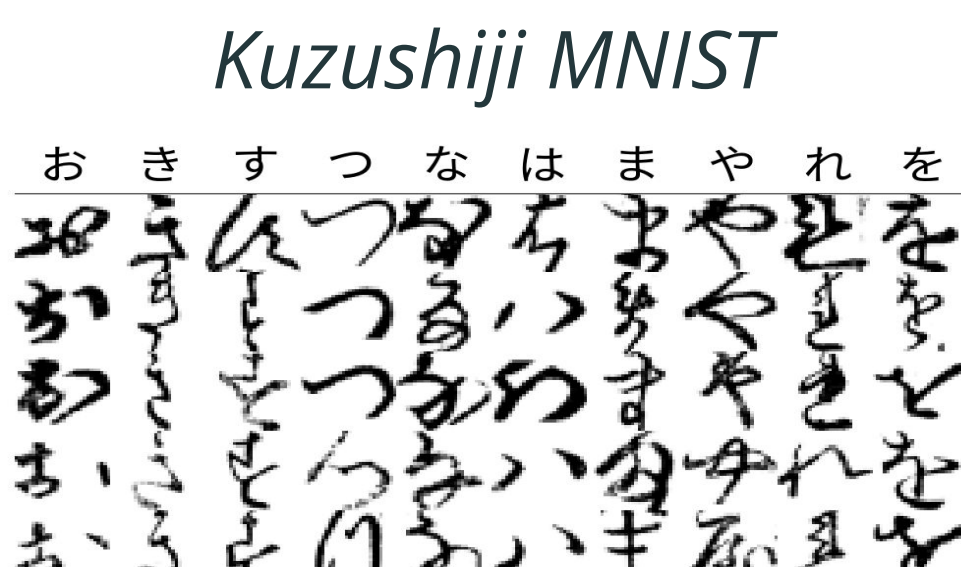
³ University of California, Santa Barbara

Goal

Create a suite of tasks that:

- require neural perception and symbolic reasoning
- provide multiple levels of challenge
- can be used in low and noisy data settings
- are easily understood both inside and outside of the NeSy community
- provide utility to the NeSy community

Data Sources



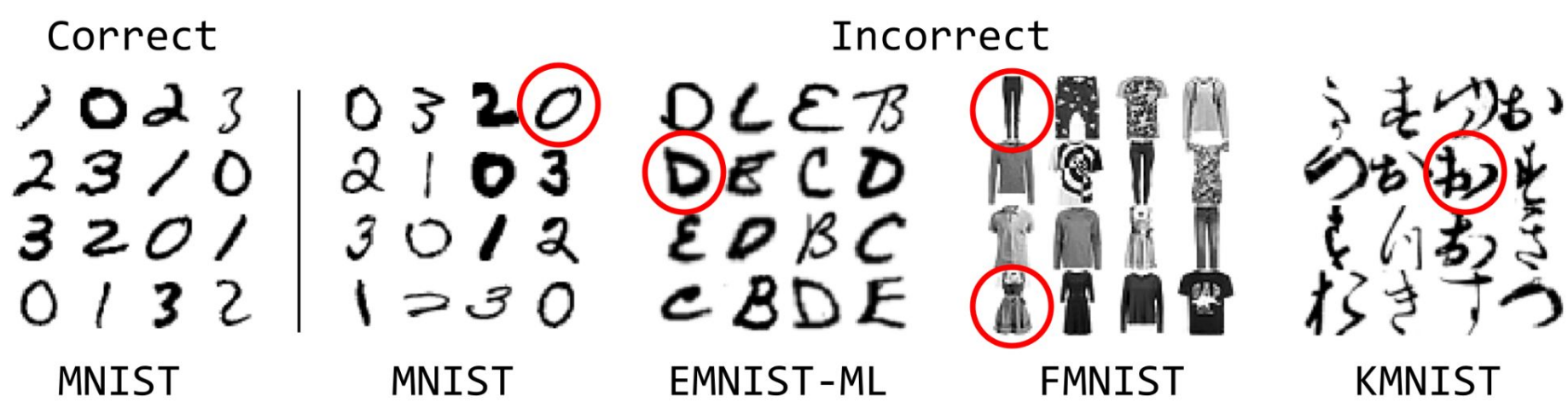
Visual Sudoku Puzzle Classification



Tasks

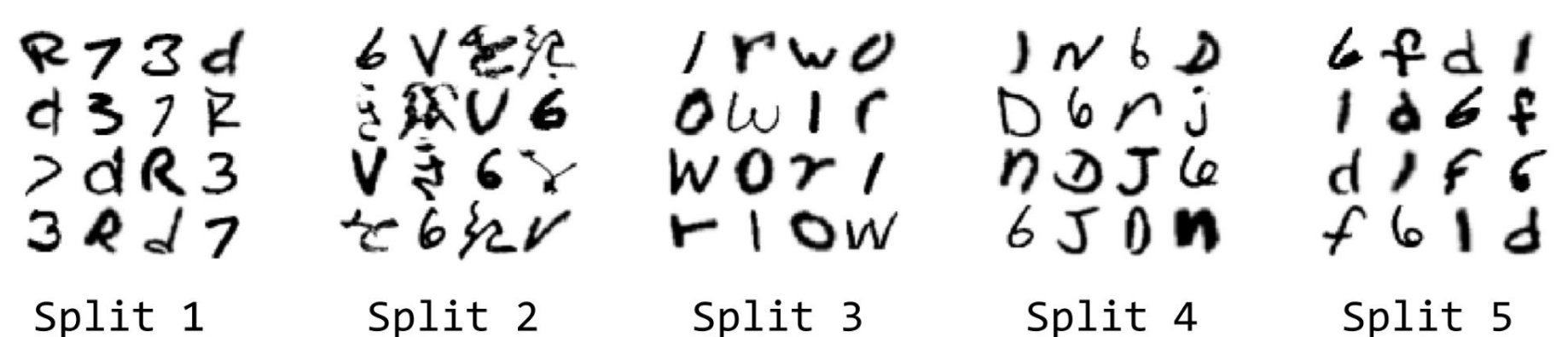
Basic

Labels chosen from a single data source, puzzles corrupted for negative examples (red).



Random Label Per Split

Allow labels from multiple data sources, labels randomly chosen for each split.



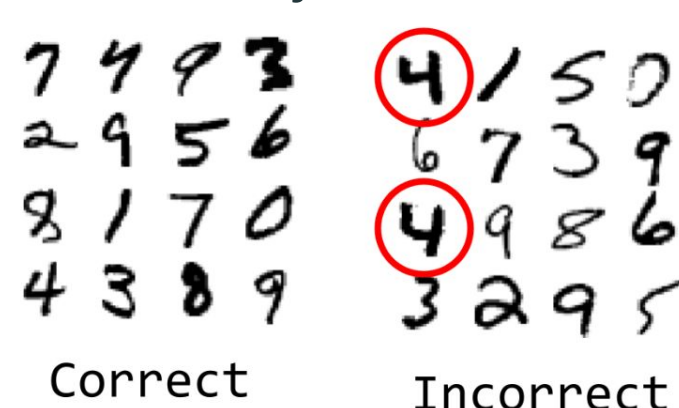
Random Label Per Puzzle

Allow labels from multiple data sources, labels randomly chosen for each puzzle.



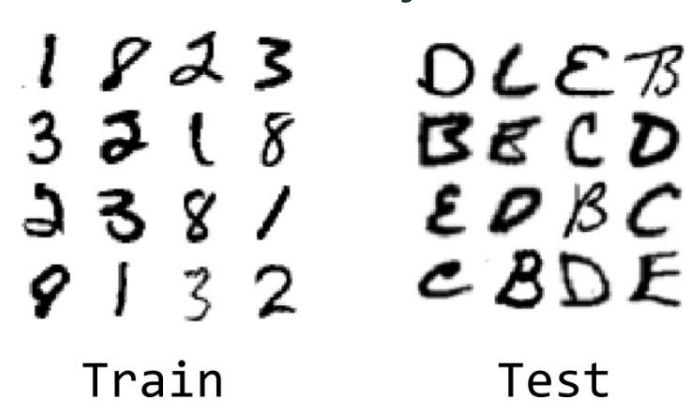
Random Label Per Cell

Allow labels from multiple data sources, labels randomly chosen for each cell.



Transfer

Use labels from a single data source, train and test receive disjoint label sets.



Results

SOTA Theoretical IID Best

Data Source	Model	Image Accuracy	Expected 4x4 Accuracy	Expected 9x9 Accuracy
MNIST	CNN Ensemble [9]	99.91%	98.57%	98.57%
EMNIST-Merged ⁴	WaveMix [18]	91.59%	24.52%	00.08%
FMNIST	DARTS [19]	96.91%	60.52%	07.87%
KMNIST	SpinalNet [20]	99.15%	87.23%	50.09%

Basic, 4x4, Full Data, No Digit Labels

Data Source	Baseline-Digit	Baseline-Visual	NeuPSL [1]
MNIST	0.71 ± 0.04	0.51 ± 0.02	0.88 ± 0.02
EMNIST-ML	0.69 ± 0.05	0.50 ± 0.01	0.79 ± 0.09
FMNIST	0.70 ± 0.04	0.50 ± 0.03	0.74 ± 0.04
KMNIST	0.71 ± 0.03	0.50 ± 0.02	0.65 ± 0.12

[1] C. Pryor*, C. Dickens*, E. Augustine, A. Albalak, W.Y. Wang, and L. Getoor. NeuPSL: Neural Probabilistic Soft Logic. arXiv. 2022.

Random Label Per Split, 4x4, No Digit Labels

