**Abstract**

- **Context:** Chess players take photographs of board positions for future analysis.
- **Purpose:** Our model turns photo into text. Player can paste text into a chess engine.

**Model:**
- Crop the image into squares of one piece
- Apply transfer learning from ResNet50
- Train a neural network on 15,000+ images
- Achieved average $F_1 = 0.94$

**Task 1: Cropping**
- Extract board from input image using Canny Edge detection, Hough line detection, corner extraction, then geometric transformation
- Evenly crop transformed board into 64 images of individual squares

**Task 2: Piece Classification**
- Substituted last FC layer with 13 unit layer, softmax activation
- Adam optimization, Loss: sparse categorical cross-entropy
- Each square is classified separately and then the board is reconstructed

**Data**
- 10,000 labelled chess-piece images compiled by Daylen Yang [1].
  - Mostly pawns, empty
  - We hand-labelled additional 10,000 images produced from the cropping algorithm in Model Task 1, then augmented as well

**Data Augmentation**
- rotate
- flip

**Results/Discussion**
- Best model achieves $F_1=0.94$ with few mistakes on a real board transcription
- Other models overfit -> perform poorly on real boards
- Model inevitably overfits to limited data
- Best classification of empty squares and worst of black bishops and kings

**Future Work**
- Increase size of database (currently prohibitively small)
- Generalize model to more piece/board styles
- Build end-to-end object detector and classifier similar to YOLO

**Acknowledgements**
We would like to thank our adviser Sarah Najm for constant feedback and support. Also thank you to Professors Ng and Stanford for teaching us.