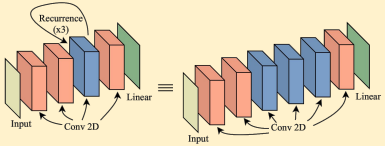
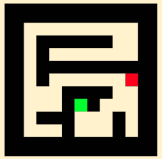
Thinking Deeply With Recurrence: Generalizing From Easy to Hard Sequential Reasoning Problems

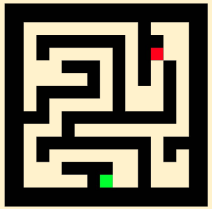
Avi Schwarzschild, Arjun Gupta, Micah Goldblum, Tom Goldstein

Can neural networks solve problems harder than those used for training?

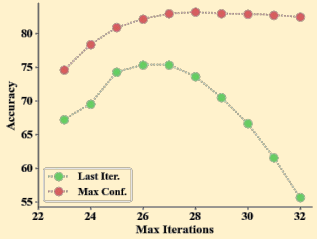


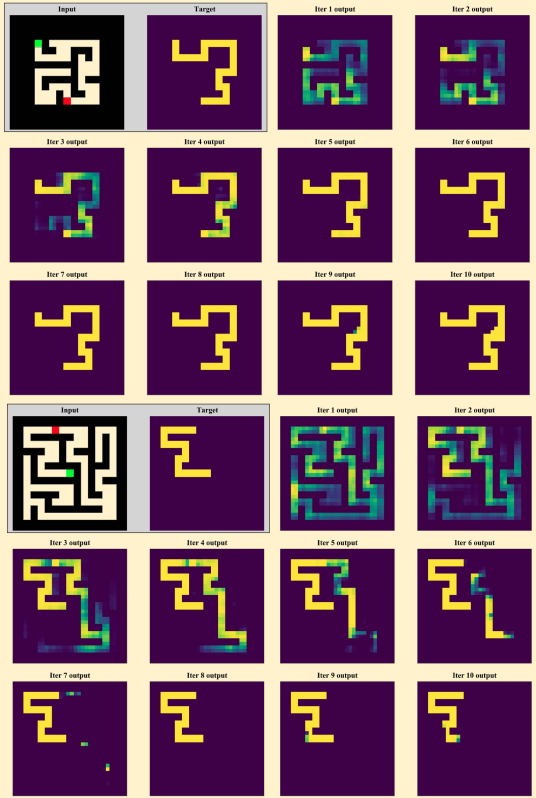
- Recurrent models can, simply by increasing their test time iteration budget (i.e., thinking for longer than they did at train time).
- We find that the performance of recurrent models improves as they recur for more iterations, even without adding parameters or re-training in the new, more challenging problem domain.





- We train on small mazes and test on large mazes. While the best feed forward models achieve only 22%, the recurrent models -- by thinking deeper -- solve over 80% of large mazes.
- We compare the accuracy on *large mazes* when using the output at the last iteration to taking the output from the iteration with the highest confidence (this model was trained to solve *small mazes* in 20 iterations.





- The output after each iteration of a model trained to solve *small mazes* in 6 iterations.
- Shown here is the confidence that each pixel is on the optimal path (purple is 0, yellow is 1).
- Top: Note that after 6 iterations the model has solved the maze
- Bottom: The model can't solve the large maze until 8 iterations.

Yes! These recurrent networks can solve harder problems -- by thinking for longer.

Read the whole paper: