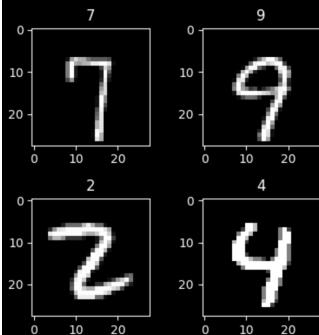


Image recognition using RNNs

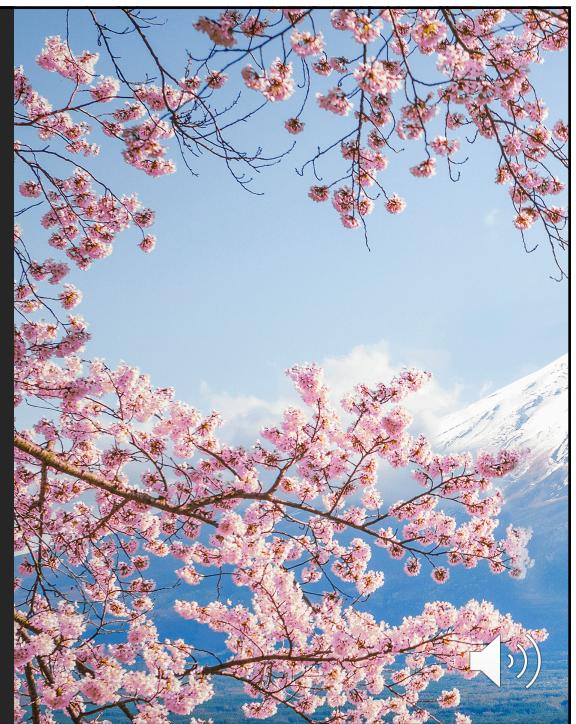
Yes, they can!



1

Introduction

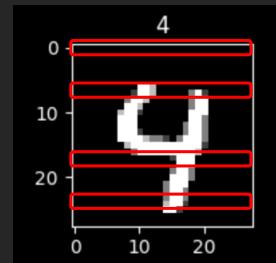
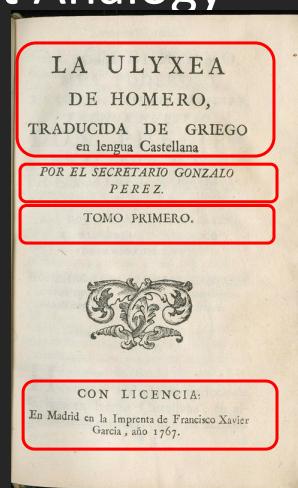
- Inspired by the paper from Sequencer: Deep LSTM for Image Classification by Yuki Tatsunami and Masato Taki. Rikkyo University



2

My hypothesis – Text Analogy

- LSTMs are well proven in NLP
- A text as a sequence of word sequences
- An image as a sequence of pixel sequences



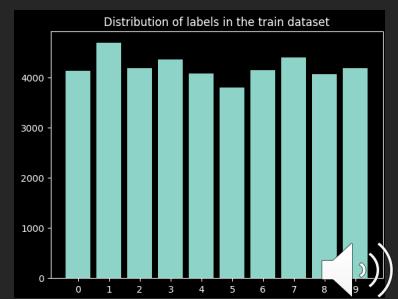
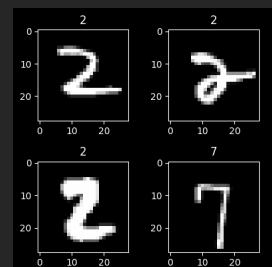
<https://apps.lib.umich.edu/online-exhibits/exhibits/show/translating-homer--from-papyri/homer-in-the-vernacular/spanish>



3

The dataset

- MNIST dataset
- Handwritten digits
- 40.000 images
- Each image is only 28x28 pixels
- It's a balanced dataset



4

2

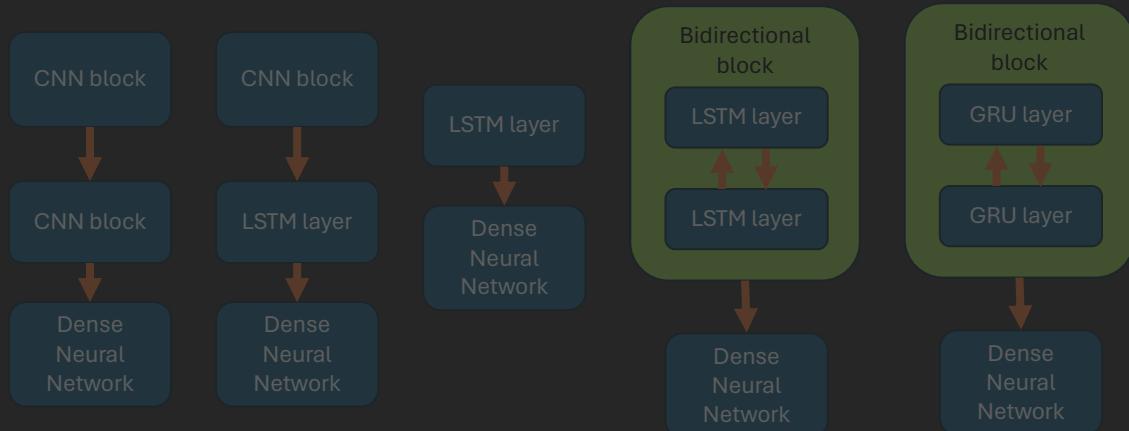
Some principles for training

- Keep complexity at the same level
- Train 5 epochs
- Use accuracy as metric
- Do hyperparameter tuning



5

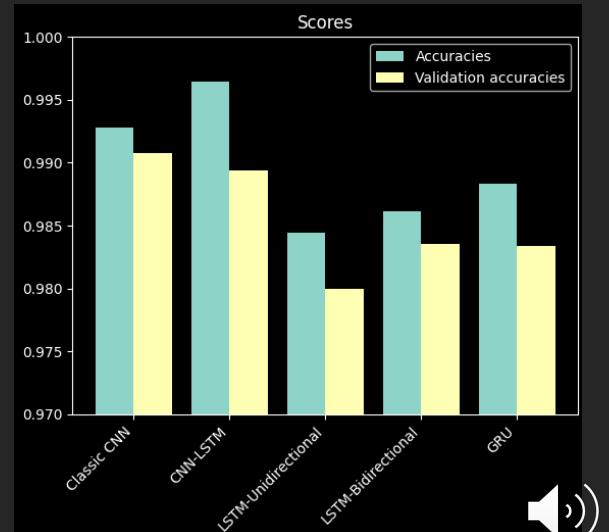
Models



6

The results (train and validation accuracy)

- CNN er very good
- RNNs can do the job
- The difficult ones: 5 – 3 and 4 - 9



7

Confusion matrices



5 vs 3

4 vs 9

9 vs 7

9 vs 3

1 vs 7

4 vs 9



8

Conclusions

- RNNs are good, but CNNs are better
- Misclassification is different
- RNNs were 40% to 55% faster
- “easy” images

