1. Introduction/Background/Motivation

This paper investigated the use of deep learning models for recipe generation. Specifically, it looked at how a deep learning model can take a food image, parse the underlying components, and produce a set of outputs (ingredients, recipes) useful to an end user (im2recipe); a process that can also be done in reverse (recipe2im). This is an interesting problem as it shows applications of deep learning models for practical use; the recreation of a food one saw online, a calorie counter for one on a diet, or understanding of food ingredients for allergy considerations.

Specifically, we’re interested in learning how to build these models, understanding and potentially improving on the used architectures. Our dataset [1] contains cross-modal data (both image and textual info). As such, our approach enables us to investigate both CNN and LSTM architectures - an exciting prospect as it melds multiple aspects of the course together.

1. Approach
2. S
3. Work done
   1. Justin
      1. developed the main.py function to run the code for necessary use in this project
      2. modified data\_loader to load appropriate data for both image and text analysis
      3. generated and tuned the CNN architecture for image processing in accordance with <cite paper>
   2. Geoff
      1. Set up and populated initial github (for code base), box (for data store), and jupyterlab repository code & interfaces to enable collaboration and to utilize cloud services
      2. Wrote grid search training script and plotting methods to help tune parameters and display results
      3. generated and tuned the LSTM architecture for ingredient and image processing in accordance with <cite paper>
4. Database

Recipe 1M+ [1]: <http://im2recipe.csail.mit.edu/>

1. References

[1] J. Marín et al., "Recipe1M+: A Dataset for Learning Cross-Modal Embeddings for Cooking Recipes and Food Images," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 1, pp. 187-203, 1 Jan. 2021, doi: 10.1109/TPAMI.2019.2927476.

[2] J.-J. Chen, C.-W. Ngo, F.-L. Feng, and T.-S. Chua, “Deep understanding of cooking procedure for cross-modal recipe retrieval,” in Proc. 26th ACM Int. Conf. Multimedia, 2018, pp. 1020–1028. [Online]. Available: <http://doi.acm.org/10.1145/3240508.3240627>

[3] M. Carvalho, R. Cadene, D. Picard, L. Soulier, N. Thome, and M. Cord, “Cross-modal retrieval in the cooking context: Learning semantic text-image embeddings,” in Proc. 41st Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2018, pp. 35–44.

[4] M. Kumari and T. Singh, "Food Image to Cooking Instructions Conversion Through Compressed Embeddings Using Deep Learning," 2019 IEEE 35th International Conference on Data Engineering Workshops (ICDEW), 2019, pp. 81-84, doi: 10.1109/ICDEW.2019.00-31.

[5] H. Rawlani, J. Saita, V. Zambre and R. L. Priya, "Deep Learning based approach to suggest recipes," 2018 International Conference on Smart City and Emerging Technology (ICSCET), 2018, pp. 1-4, doi: 10.1109/ICSCET.2018.8537350.