International Journal of Emerging Technology and Innovative Engineering
Volume 6, Issue 04, April 2020 (ISSN: 2394 – 6598)

Food Detection and Recognition and using Deep learning Framework

M. Mohamed Foise[1], R. Karthik[2],

[1]Department of CSE, Sri Krishna College of Technology, Coimbatore.

[2]Assistant Professor, Sri Krishna College of Technology, Coimbatore.

ABSTRACT

Since health care on foods is drawing people's attention recently. Diet is very important in human life. Obtaining adequate nutrition from everyday meals is essential for our health. Such food recording is usually a manual exercise using textual description, but manual recording is tedious and time consuming. A problem that humans face on a daily basis is how to make a conscious decision regarding our daily food consumption that is nutritious and healthy. Deep learning has recently been used in image recognition. Deep learning is a collective term for algorithm shaving a deep architecture that solves complex problems. This works proposes inception v3 method using food images. In this project we used inception deep learning method to identify the food in effective manner.

Keywords: Deep Learning, Foods, Recognition, AI Techniques, Inception V3

1. INTRODUCTION

Food is an important part of everyday life. This clearly ripples through into digital life, as illustrated by the abundance of food photography in social networks, dedicated photo sharing sites and mobile applications. Automatic recognition of dishes would not only help users effortlessly organize their extensive photo collections but would also help online photo repositories make their content more accessible. Additionally, mobile food photography is now used to help patients estimate and track their daily calory intake, outside of any constraining clinical environment. However, current systems resort to nutrition experts or Amazon Mechanical Turk to label food items. Despite these numerous applications, the problem of recognizing dishes and the composition of their ingredients has not been fully addressed by the computer vision community. This is not due to the lack of challenges. classification or object detection, food typically does not exhibit any distinctive spatial layout: while we can decompose an outdoor scene with a ground plane, a horizon and a sky region, or a human as a trunk with a head and limbs, we cannot find similar patterns relating ingredients of a mixed salad. The point of view, the lighting conditions, but also (and not least) the very realization of a recipe are among the sources of high intra-class variations. On the bright side, the nature of dishes is often defined by the different colours and textures of its different local components, such that humans can identify them reasonably well from a single image, regardless of the above variations. Hence, food recognition is a specific classification problem calling for models that can exploit local information. Convolutional Neural Networks (CNN), a technique within the broader Deep Learning field, have been a revolutionary force in Computer Vision applications, especially in the past half-decade or so. One main use-case is that of image classification, e.g. determining whether a picture is that of a dog or cat.

2. Related Work

Image classification is a core problem for computer vision, with many recent advances coming from object recognition. Classical approaches exploit interest point descriptors, extracted locally or on a dense grid, then pooled into a vectoral representation to use SVM for classification. Recent advances highlight the importance of nonlinear feature encoding, e.g., Fisher Vectors and spatial pooling. A very recent and successful trend in classification is to try and identify discriminative object (or scene) parts (or patches) drawing on the success of deformable part-based models (DPM) for object detection. This can consist of (a) finding prototypes for regions of interest, (b) mining patches whose associated binary SVM obtains good classification accuracy on a validation set , (c) clustering patches with a multi-instance SVM (mi-SVM) on an external dataset, (d) optimizing part detectors in a latent SVM framework, (e) evaluating many exemplar-SVMS on sliding windows, exploiting discriminative decorrelation to speed-up the process, or (f) identifying discriminative modes in the hog feature space . While this work represents a variant of discriminative part mining, it differs in various ways from previous work. In contrast to all other discriminative part mining methods, we efficiently and simultaneously mine for discriminative parts for all the categories in our dataset thanks to the multi-class nature of Random Forests. Secondly, while all other methods employ a computationally expensive (often multi scale) sliding window detection approach to produce the part score maps for the final classification step, our approach employs a simple yet effective window selection by exploiting image super pixels. Concerning food recognition, most works follow a classical recognition pipeline, focusing on feature combination and on specialized datasets. uses a private dataset of Japanese food, later augmented with more features and classes. Similarly, jointly classifies and estimates quantity of 50 Chinese food categories using private data. uses DPM to locally pool features. Food images obtained in a controlled environment are also popular in the literature.

3. Dataset: Food-101

As noted above, to date, only the paid dataset is publicly available. However, it contains only standardized fast food images taken under laboratory conditions. Therefore, we have collected a novel real-world food dataset by downloading images from foodspotting.com. The site allows users to take images of what they are eating, annotate place and type of food and upload this information online. We chose the top 101 most popular and consistently named dishes and randomly sampled 750 training images. Additionally, 250 test images were collected for each class, and were manually cleaned.

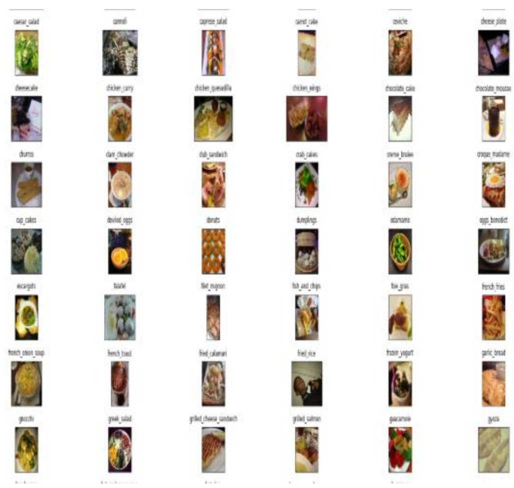
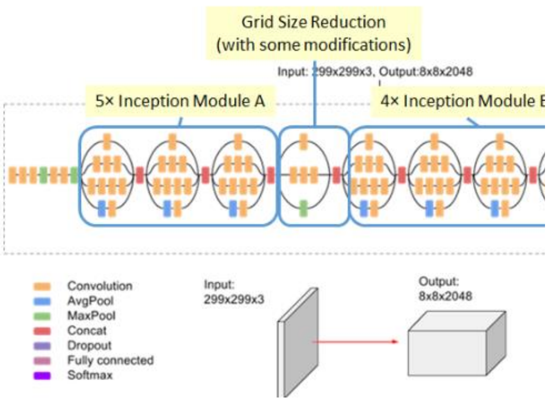


FIG. 1.1. VISUALIZE RANDOM IMAGE FROM EACH OF THE 101 CLASSES

On purpose, the training images were not cleaned, and thus still contain some amount of noise. This comes mostly in the form of intense colours and sometimes wrong labels. We believe that real-world computer vision algorithms should be able to cope with such weakly labelled data if they are meant to scale well with the number of classes to recognise. All images were rescaled to have a maximum side length of 512 pixels and smaller ones were excluded from the whole process. This leaves us with a dataset of 101’000 real world images in total, including very diverse but also visually and semantically similar food classes such as Apple pie, Waffles, Escargots, Sashimi, Onion rings, Mussels, Edamame, Paella, Risotto, Omelette, Bibimbap, Lobster bisque, Eggs benedict, Macarons to name a few. The dataset is available for download at http://www.vision.ee.ethz.ch/datasets/food-101/.

4. Deep learning



Inception v3 is a widely-used image recognition model that has been shown to attain greater than 78.1% accuracy on the ImageNet dataset. The model is the culmination of many ideas developed by multiple researchers over the years. It is based on the original paper: "Rethinking the Inception Architecture for Computer Vision" by Szegedy, et. al. The model itself is made up of symmetric and asymmetric building blocks, including convolutions, average pooling, max pooling, concats, dropouts, and fully connected layers. Batch norm is used extensively throughout the model and applied to activation inputs. Loss is computed via Softmax. The TPU version of Inception v3 is written using TPU Estimator, an API designed to facilitate development, so that you can focus on the models themselves rather than on the details of the underlying hardware. The API does most of the low-level grunge work necessary for running models on TPUs behind the scenes, while automating common functions, such as saving and restoring checkpoints. The Estimator API enforces separation of model and input portions of the code. You have to define model_fn and input_fn functions, corresponding to model definition and input pipeline / pre-processing stages of the TensorFlow graph, respectively. Each Cloud TPU device has 8 cores and is connected to a host (CPU). Larger slices have multiple hosts. Other larger configurations interact with multiple hosts. For instance, v2-256 communicates with 16 hosts.

Hosts retrieve data from the file system or local memory, do whatever data pre-processing is required, and then transfer pre-processed data to the TPU cores. We consider these three phases of data handling done by the host individually and refer to the phases as: 1) *Storage*, 2) *Pre-processing*, 3) *Transfer*. To yield good performance, the system should be balanced. Whatever amount of time a host CPU spends retrieving images, decoding them, and doing relevant pre-processing, should ideally be slightly less or about the same as that spent by the TPU doing computations. If the host CPU takes longer than the TPU to complete the three data handling phases, then execution will be host bound. The current implementation of Inception v3 lives right at the edge of being input-bound. Images have to be retrieved from the file system, decoded, and then pre-processed. Different types of pre-processing stages are available, ranging from moderate to complex. If we use the most complex of pre-processing stages, the large number of expensive operations executed by the pre-processing stage will push the system over the edge and the training pipeline will be pre-processing bound.

5. Results of proposed work

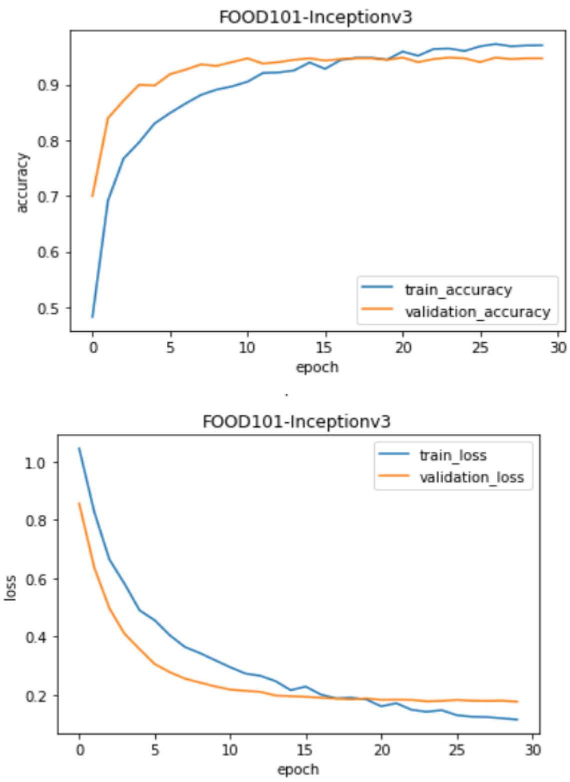


Fig.2. Inception V3 results

Table 1. SUMMARY RESULTS OF PROPOSED WORK

CLASS	METHOD	TRAIN ACCURACY	TEST ACCURACY
3 class	Inception V3	97%	95%
11 class	Inception V3	97%	95%

cup_cakes

spring_rolls

garlic_bread

6. CONCLUSION:

We evaluated the effectiveness in classifying food images of a deep-learning approach based on the specifications of Google's image recognition architecture Inception. To the best of our knowledge, these results significantly improve the best published results obtained on the same datasets, while requiring less computation power, since the number of parameters and the computational complexity are much smaller than the competitors. For future work, we will implement the proposed framework on mobile devices. This works proposes inception v3 using food images. In this project we used inception deep learning method to identify the food in effective manner. Regarding Twitter food photo mining, we plan to extend the framework to analysis food distribution and preference from the geo-spatial and temporal aspects.

REFERENCES

[1] N. Krishnaraj, Mohamed Elhoseny, M. Thenmozhi, Mahmoud M. Selim, K. Shankar, “Deep learning model for real-time image compression in Internet of Underwater Things (IoUT)”, Journal of Real-Time Image Processing, May 2019. In Press. https://doi.org/10.1007/s11554-019-00879-6

[2] Elhoseny, M., & Shankar, K. (2019). Reliable data transmission model for mobile ad hoc network using signcryption technique. IEEE Transactions on Reliability.

[3] Elhoseny, M., & Shankar, K. (2019). Optimal bilateral filter and convolutional neural network based denoising method of medical image measurements. Measurement, 143, 125-135.

[4] Dutta, A. K., Elhoseny, M., Dahiya, V., & Shankar, K. (2019). An efficient hierarchical clustering protocol for multihop Internet of vehicles communication. Transactions on Emerging Telecommunications Technologies, e3690.

[5] Elhoseny, M., & Shankar, K. (2020). Energy Efficient Optimal Routing for Communication in VANETs via Clustering Model. In Emerging Technologies for Connected Internet of Vehicles and Intelligent Transportation System Networks (pp. 1-14). Springer, Cham.

[6] RL Babu, S Vijayan,(2016) “Wrapper based feature selection in semantic medical information retrieval”, Journal of Medical Imaging and Health Informatics, 6 (3), pp.802-805.

[7] RA Suji, DB Anand, RL Babu,(2019) “An Efficient Image Enhancement Technique using Stationary Wavelet Transform Based Image Fusion”, International Journal of Recent Technology and Engineering 8 (1), pp.2277-3878.

[8] RA Suji, DB Anand, RL Babu,(2019) “Polar ice image segmentation using improved estimation and normalization of

illumination", Cluster Computing 22 (1), pp.S3811–S3819.

[9] Bright Anand, D, Saravanan, R & Adaline Suji, R 2017, 'Adaptive Batch Mode Active Learning Technique Usingwith an Improved Time Adaptive Support Vector Machine for Classification of Remote Sensing Applications', Journal of Computational and Theoretical Nanoscience, vol. 14, no. 2, pp. 1108-1113.

[10] Wang G-G, Tan Y (2019) Improving metaheuristic algorithms with information feedback models. IEEE Trans Cybern 49(2): 542-555. doi: 10.1109/TCYB.2017.2780274

[11] Wang G-G, Gandomi AH, Alavi AH (2014) An effective krill herd algorithm with migration operator in biogeography-based optimization. Appl Math Model 38(9-10): 2454-2462. doi: 10.1016/j.apm.2013.10.052

[12] Wang G-G, Guo L, Gandomi AH, Hao G-S, Wang H (2014) Chaotic krill herd algorithm. Inf Sci 274: 17-34. doi: 10.1016/j.ins.2014.02.123

[13] Wang G-G (2018) Moth search algorithm: a bio-inspired metaheuristic algorithm for global optimization problems. Memetic Computing 10(2): 151-164. doi: 10.1007/s12293-016-0212-3

[14] Wang G-G, Deb S, Cui Z (2019) Monarch butterfly optimization. Neural Comput Appl 31(7): 1995-2014. doi: 10.1007/s00521-015-1923-y

[15] Uma Maheswari, P., et al. "Bat optimization algorithm with fuzzy based PIT sharing (BF-PIT) algorithm for Named Data Networking (NDN)." Journal of Intelligent & Fuzzy Systems 37.1 (2019): 293-300.

[16] Manickam, P., et al. "Performance comparisons of routing protocols in mobile ad hoc networks." arXiv preprint arXiv:1103.0658 (2011).

[17] Manickam, P., and D. Manimegalai. "A HIGHLY ADAPTIVE FAULT TOLERANT ROUTING PROTOCOL FOR ENERGY CONSTRAINED MOBILE AD HOC NETWORKS." Journal of Theoretical & Applied Information Technology 57.3 (2013).

[18] Manickam, P., et al. "Secure data transmission through reliable vehicles in VANET using optimal lightweight cryptography." Cybersecurity and secure information systems. Springer, Cham, 2019. 193-204.

[19] Muslihudin, Muhamad, et al. "Expert system in determining baby disease using web mobile-based weighted product method." International Journal of Recent Technology and Engineering 8.1 (2019): 3299-3308.

[20] Sudhakar, Sengan, and S. Chenthur Pandian. "Hybrid cluster-based geographical routing protocol to mitigate malicious nodes in mobile ad hoc network." International Journal of Ad Hoc and Ubiquitous Computing 21.4 (2016): 224-236.

[21] Sengan, Sudhakar, and S. Chethur Pandian. "An efficient agent-based intrusion detection system for detecting malicious nodes in MANET routing,(2012)." International Review on Computers and Software (IRECOS) 7.6: 3037-3041.

[22] Sudhakar Sengan, Dr S. "A trust and co-operative nodes with affects of malicious attacks and measure the performance degradation on geographic aided routing in mobile ad hoc network." Life Science Journal 10.4s (2013).

[23] Sudhakar Sengan, S. Chenthur P. "Secure packet encryption and key exchange system in mobile ad hoc nerwork." J. Comput. Sci. 2012.

[24] Sengan, Sudhakar, and Pandian S. Chenthur. "Investigation of attribute aided data aggregation over dynamic routing in wireless sensor networks." Journal of Engineering Science and Technology 10.11 (2015): 1465-1476.

[25] Lakshmanaprabu, S. K., et al. "Effective features to classify big data using social internet of things." IEEE access 6 (2018): 24196-24204.

[26] Gupta, Deepak, et al. "Optimized cuttlefish algorithm for diagnosis of Parkinson's disease." Cognitive systems research 52 (2018): 36-48.

[27] Gupta, Deepak, Anil K. Ahlawat, and Kalpna Sagar. "Usability prediction & ranking of SDLC models using fuzzy hierarchical usability model." Open Engineering 7.1 (2017): 161-168.

[28] Gupta, Deepak, and Anil K. Ahlawat. "Taxonomy of GUM and usability prediction using GUM multistage fuzzy expert system." Int. Arab J. Inf. Technol. 16.3 (2019): 357-363.

[29] Gupta, Deepak, et al. "Improved diagnosis of Parkinson's disease using optimized crow search algorithm." Computers & Electrical Engineering 68 (2018): 412-424.
[30] Raja Senthil Ragavan, V. K., Mohamed Elhoseny, and K. Shankar. "An enhanced whale optimization algorithm for vehicular communication networks." International Journal of Communication Systems (2019).
[31] Uthayakumar, J., et al. "Intelligent hybrid model for financial crisis prediction using machine learning techniques." Information Systems and e-Business Management (2018): 1-29.
[32] Muslihudin, Muhammad, et al. "Prediction of layer chicken disease using fuzzy analytical hierarcy process." Int. J. Eng. Technol 7.2.26 (2018): 90-94.
[33] Lydia, E. L., Shankar, K., Pamina, J., & Raja, J. B. (2018). Correlating NoSQL Databases With a Relational Database: Performance and Space. International Journal of Pure and Applied Mathematics, 118(7), 235-244.
[34] Gupta, Deepak, et al. "Efficient artificial fish swarm based clustering approach on mobility aware energy-efficient for MANET." Transactions on Emerging Telecommunications Technologies 30.9 (2019): e3524.
[35] Murugan, B. S., et al. "Region-based scalable smart system for anomaly detection in pedestrian walkways." Computers & Electrical Engineering 75 (2019): 146-160., J. Beschi, et al. "Exploring Effective Feature Selection Methods for Telecom Churn Prediction."
[36] Pamina, J., et al. "Inferring Machine Learning Based Parameter Estimation for Telecom Churn Prediction." International Conference On Computational Vision and Bio Inspired Computing. Springer, Cham, 2019.
[37] Raja, J. Beschi, S. Chenthur Pandian, and J. Pamina. "Certificate revocation mechanism in mobile ADHOC grid architecture." Int. J. Comput. Sci. Trends Technol 5 (2017): 125-130.
[38] Aishwarya, R., et al. "A Novel Adaptable Approach for Sentiment Analysis." (2019).
[39] Lydia, E. Laxmi, et al. "Correlating NoSQL Databases With a Relational Database: Performance and Space." International Journal of Pure and Applied Mathematics 118.7 (2018): 235-244.