Indian Food Classification

MANOJ KUMAR
Computer Science Department
KLE Technological University
Hubbali,India
email address or ORCID

SRIVATHSA A
Computer Science Department
KLE Technological University
Hubbali,India
email address or ORCID

P SAI SUDHEER
Computer Science Department
KLE Technological University
Hubbali,India
email address or ORCID

PRANEET MANE
Computer Science Department
KLE Technological University
Hubbali,India
email address or ORCID

***Abstract*—In this study, we used convolutional neural networks to attempt to classify food-related photos. Convolutional neural networks are highly effective at solving the challenge of classifying pictures because they can extract spatial characteristics from images. People have recently started posting pictures of food on social media and reviewing feasts. Therefore, there are many food images on social media, but some of them might not be labelled. If restaurants can market their cuisine to consumers who are shopping for comparable dishes, it will be highly beneficial to them. Social media platforms may recognise food with the use of a food classification system. Food categorization is a particularly challenging undertaking since there is a lot of variation across photographs of the same food group. To categorise food photos in the food20 dataset, we created a convolutional neural network model. To classify photographs of food, we deployed a pre-trained InceptionV3 convolutional neural network model. In this design, the foundation is established by adding Squeeze-and-Excitation blocks, which facilitate feature recalibration by adaptively adjusting the relevance of distinct channels.**

***Index Terms*—Convolutional Neural Networks Food Image Classification Social Media Marketing InceptionV3 Model Food20 Dataset**

I. Introduction

The recognition of food photos is one of the significant applications used today. Because of this, smartphone users might be able to find the food's name. This is very important for tourists who go abroad. They are also useful when it comes to accurately ordering meals and knowing the specifics of the food, such its calorie content and potential allergies[9]. Nowadays, smartphone image recognition apps mostly require a computer because the identification procedure consumes a lot of resources in order to service the database. If the database is large, a smartphone's limited capabilities cannot handle the data. Consequently, the smartphone needs to transmit the data to the external computer for processing[22].in this case, the effectiveness of the recognition will depend on the computer's performance and the internet connection's speed.

Since many meals are similar in form, size, and color to one another and since regional variances in food preparation and cooking techniques occur, image-based food recognition is more difficult[2]. In compared to traditional image analysis approaches, deep learning boosted the accuracy of classification tasks. Machine Learning algorithms are unable to discern several complicated traits in food photographs owing to the low inter-class variation and large intra-class variance, whereas CNNs can do so with ease. As a result of its multiple practical benefits in fields such as nutrition calculation, allergy trace recommendation, foreign food prediction, and automated food monitoring and searching, researchers have recently started to pay increased attention to the categorization of food images[21].

Millions of food images are publicly available thanks to the growing use of cell phones and social media, and there is a great demand for automated categorization. The past few years have witnessed the compilation of numerous food photo databases. Since Deep Convolutional Neural Networks (DCNNs) may achieve improved accuracy by utilizing the availability of big labeled datasets with their millions of parameters, they have become the most common approach for image classification difficulties.

One of the successful ways for pattern and picture recognition is deep learning. Higher recognition accuracy is achieved utilizing deep learning learning algorithms compared to that from the previous techniques applied in the past[12]-16]. Deep learning algorithms are also well suited for learning from enormous amounts of data. This may be accomplished using an artificial neural network (ANN). Convolutional layer, pooling layer, fully connected layer, and other layers are merely a few examples of the layers to which ANN is applied.

A network known as a convolutional neural network (CNN) is generated when these layers are cascaded one after the other. One of the possible deep-learning algorithms is CNN[4]. It has been used to examine photos so that computers can understand them and group related images into the same class. It is advisable to employ supervised learning for this. There has been ongoing growth of CNN. The Residual network (ResNet) was proposed by Microsoft Research Asia in 2015[1]. (MSRA). ResNet achieved first place in both the Common Objects in Context (MSCOCO 2015) and the ImageNet Large Scale Visual Recognition Challenge (ILSVRC 2015) competitions in 2015. Identity mapping as a concept was put out.

II. Related Work

A research study named ”Deep Indian Delicacy: Classification of Indian Food Images using Convolutional Neural Networks” explores the use of CNN for the discovery and labeling of photos of Indian food. It shows how difficult it is to spot different types of food because of their varied arrangement methods. The model gets a classification accuracy of 0.879 for correct sorting of all studied pictures in a single session. The study uses a small collection of 60,000 grayscale pictures relating to 10 types of Indian food. In order to show how dataset size and epochs affect classification accuracy, the model's performance when training on CPU and GPU is also compared in this study. The study offers views into the difficulties faced in the process of building the model, including the effects of using grayscale pictures and computing limits. It also makes suggestions for possible changes, such as increasing the amount of training pictures and exploring the use of food image recognition in areas like robots and smartphone technology.

The first paper on picture recognition of Thai food was presented in 2017. The first and biggest file of shots of Thai food in existence was presented by THFOOD50. They suggested NUInNet1.1, a new network that uses the Inception module with the suitable depth, to keep the model's accuracy while lowering processing time and model size. In comparison to Google Net, the suggested model beats it, getting 69.8 accuracy and cutting the forecast time to 18.16 ms[8]. An change to NU-InNet1.1 was suggested. With an accuracy of 80.34 on the THFOOD-50 dataset, four stacks were shown to work best in a study on adding multiple modified inception modules.The NU-ResNet1.1 Depth 4 was finally created by combining the idea of leftover layers from ResNet with the NU-InNet1.1 Depth 4. Skip layers were used to split the start and end of each origin module. For the THFOOD-50 dataset, it was the most accurate classification model, with an accuracy of 83.07 and an inference time of 44.60 ms[11].Provided was an extra collection of 3,961 images of Thai food, divided into 11 groups. Attaining an accuracy of 88.33, they improved a GoogLeNet that had been taught using data from ImageNet beforehand.

Photographs of Thai food were divided into 13 groups using a food recognition program. They were able to get an accuracy of 82.00 by using a VGG19 network using transfer learning on their own collection of 7,632 pictures of Thai food. Due to the size of the VGG19 network (144 106 parameters), the forecasts were handled on the server and the results were sent back to the clients' devices. From these past efforts, we may draw some important lessons for this specific case. First, they are all depending on teaching neural networks to provide state-of-the-art results. Second, greatly improving classifier performance on a small sample via the use of transfer learning[3]. Nevertheless, the use of Transfer Learning for the largest dataset on Thai food—THFOOD-50—has never been thoroughly examined. Comparing the benefits of Transfer Learning with different networks, training methods, and augmentations is the goal of this study.

III. Background

A family of deep learning models called Convolutional Neural Networks (CNNs) is especially made for handling structured grid data, like photographs. When it comes to image processing , CNNs have been very effective. They have shown higher performance in tasks like object identification, face recognition, and picture categorization. The network is

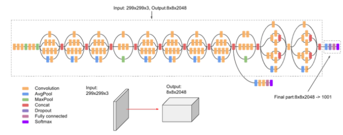

Fig. 1. Inception V3 architecture

able to recognize patterns because they use convolutional layers to automatically learn hierarchical characteristics from input photos. The design generally includes of convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification. In addition to being used in computer vision, CNNs have revolutionized picture identification and are now widely used in fields including natural language processing and medical image analysis.

Conversely, Google has built a more complex and powerful convolutional neural network design called InceptionV3. It belongs to the Inception model family and is renowned for its effectiveness in striking a balance between model accuracy and computing complexity. InceptionV3 makes use of a module known as the Inception module, also known as GoogLeNet, which captures characteristics at various scales by using several convolutional filters of varying sizes inside a single layer. As a result, the model performs better and can extract a wider range of properties. Many image recognition tasks have made extensive use of InceptionV3, which has shown state-of-the-art performance on benchmark datasets.

Regarding your project on "Indian food classification," InceptionV3 may be a useful tool. It is appropriate for differentiating between many Indian cuisine varieties because to its capacity to catch minute nuances and a variety of attributes. Accurate categorization of Indian food is challenging due to its richness and diversity; nevertheless, the multi-scale feature extraction capabilities of InceptionV3 may help to alleviate this difficulty. By using InceptionV3, your project may have access to a formidable pre-trained model that has proven successful in picture classification tasks, which might increase the precision and resilience of your system for classifying Indian cuisine.

IV. Methodology

A. Dataset Description

With 4018 training shots and 1250 validation images spanning 20 important areas of Indian food, the Food20 dataset provides researchers with a comprehensive resource.

studying the topic of photo classification as it applies to the kitchen. Datasets are methodically sorted into train, val, and test folders to create a user-friendly and structured environment for model construction and assessment. The unique emphasis on the complexity of Indian cuisine, which are impacted by several geographical, climatic, and historical elements, making it a one-of-a-kind task. Multiple dishes on one plate and identical sides for each item highlight the dataset's practical usage. As a consequence, models need to

Fig. 2. Sample images of Food20 dataset

be able to make sophisticated selections when presented with diverse sorts of food. Researchers also have a fantastic base to work upon because a pre-trained Inception v3 model was supplied in the Model directory as an optional extra. This model was constructed via transfer learning. Not only may this pretrained model be utilized as a baseline for performance, but researchers are also invited to refine its architecture for even higher accuracy and precision. By capturing the enormous range of Indian food, the Food20 dataset basically allows researchers to dig into the difficult picture classification issues encountered by this domain.

B. Preprocessing and Augmentation

In the preprocessing and augmentation phase, different strategies were utilized to increase the quality and variety of the dataset, supporting the training of a robust convolutional neural network (CNN). The issues addressed included cleansing the dataset owing to misspelled labels and an excessive amount of raw photos not connected to any class. To acquire insights into the data and undertake early analysis, a picture was imported, and pixel intensity distributions were shown in the RGB color space. Additionally, clustering methods, such as K-means clustering, were applied to find patterns in the picture data.

Exploratory Data Analysis (EDA) was undertaken to show the distribution of training instances for each food item. It indicated class inequalities, with two classes, Pani puri and Kulfi, having less training examples. To alleviate this problem and increase model generalization, data augmentation approaches were employed employing offline augmentation. Randomly picked photographs from each class were displayed to offer an overview of the dataset.

C. Model training

First, the InceptionV3 model, pretrained on the ImageNet dataset, was selected as the fundamental model for transfer learning. This conclusion was informed by the model's demonstrated ability to catch precise features and patterns from various photos. To tailor the architecture to the unique task of identifying Indian food, additional layers were added, including global average pooling for spatial information aggregation, dense layers for feature learning, and dropout for regularization to minimize overfitting. During training, a huge dataset was built, employing data augmentation methodologies to address class imbalances and increase the model's potential to generalize to new food items. The ImageDataGenerator from TensorFlow's Keras API supported real-time data augmentation, enabling transformations such as shearing, zooming, and horizontal flipping. This extended dataset was then utilized to train the model using the Stochastic Gradient Descent (SGD) optimizer, employing a learning rate of 0.0001 and a momentum of 0.9. The use of categorical cross-entropy as the loss function aims to assess the dissimilarity between predicted and true class labels. The training process was monitored using checkpoints and a CSV logger, keeping the best-performing model and documenting the training history for subsequent research. After 20 epochs, the model demonstrated convergence, capturing the underlying patterns within the dataset. The testing stage involves assessing the trained model on a separate test set, grading its performance on previously unknown data. This quantitative test was necessary to confirm the model's generalization capabilities and prove its accuracy in identifying various Indian cuisine products. The model obtained an excellent test accuracy f 0.897, encouraging confidence in its potential to provide reliable predictions. Prediction, a basic element of the model's practical value, was established by the usage of both individual picture prediction and batch prediction functions. The algorithm was capable of providing accurate predictions on single food photos, indicating its capabilities in detecting many culinary items. Furthermore, batch predictions on full directories allowed for a thorough analysis of the model's performance across multiple food classes

V. RESULTS

Three pre-trained models were evaluated in our study on Indian food classification: Xception, InceptionResNetV2, and InceptionV3. The InceptionV3 model stood out for its exceptional accuracy, reaching an impressive 0.87. This result demonstrates how well InceptionV3 captures complex patterns and traits that are important to the wide variety of Indian cuisines. When compared to this, the Xception model's 0.834 accuracy was rather good, while InceptionResNetV2's 0.805 accuracy was just below average. The findings show that these pre-trained models aren't exactly the same when it comes to fine-grained food categorization .

Model	Accuracy (%)
InceptionV3	87.0
Xception	83.4
InceptionResNetV2	80.5

TABLE I
COMPARISON OF ACCURACY FOR DIFFERENT PRE-TRAINED MODELS

The model's practical value could be proved via prediction, which was tested using both individual picture prediction and batch prediction methods. The model's precision in identifying different food items was shown by its ability to provide correct predictions from individual food photos. Furthermore, batch predictions on whole directories allowed for a complete examination of the model's performance across numerous food groups.

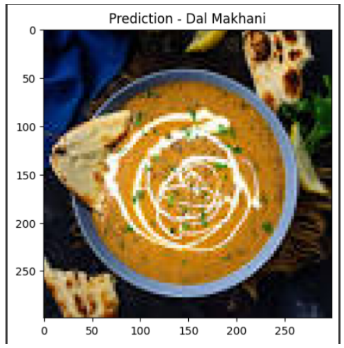

Fig. 3. Lung Image with Masked Tumour Area

.

VI. CONCLUSION

In conclusion, this research study on Indian Food categorization effectively exhibits the usefulness of utilizing the InceptionV3 deep learning model, attaining a remarkable accuracy of 0.87 in automating the identification and categorization of varied Indian food specialties. Through rigorous data preparation, data augmentation, and model training, the study overcomes difficulties such as class imbalances and changes in picture properties. Comparative comparison with various pretrained models, including Xception and InceptionResNetV2, reveals the higher performance of InceptionV3. The conclusions of this study have major implications for computer vision applications in the food sector, highlighting the model's practical value for tasks such as nutritional evaluation and menu item identification. The effort adds to increasing the knowledge of deep learning applications in picture classification and emphasizes the relevance of model selection and data pretreatment for obtaining high accuracy in complicated tasks like Indian cuisine categorization. Future initiatives may comprise more fine-tuning, research of ensemble techniques, and dataset enlargement for larger generalization capabilities.

.

VII. REFERENCES

[1] Lu, Yuzhen. "Food image recognition by using convolutional neural networks (cnns)." arXiv preprint arXiv:1612.00983 (2016).

[2] Kaur, P., Sikka, K., Wang, W., Belongie, S., Divakaran, A. (2019). Foodx-251: a dataset for fine-grained food classification. arXiv preprint arXiv:1907.06167.

[3] Termritthikun, Chakkrit, Paisarn Muneesawang, and Surachet Kanprachar. "NU-InNet: Thai food image recognition using convolutional neural networks on smartphone." Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 9.2-6 (2017): 63-67.

[4] ttisoponpisan, Sirawit, et al. "Pushing the Accuracy of Thai Food Image Classification with Transfer Learning." Engineering Journal 26.10 (2022): 57-71.

[5] Hnoohom, Narit, and Sumeth Yuenyong. "Thai fast food image classification using deep learning." 2018 International ECTI northern section conference on electrical, electronics, computer and telecommunications engineering (ECTI-NCON). IEEE, 2018.

[6] Salim, Nareen OM, et al. "Study for Food Recognition System Using Deep Learning." Journal of Physics: Conference Series. Vol. 1963. No. 1. IOP Publishing, 2021.

[7] J. Panprasittikit and K. Amnatchotiphan, "Thai Tangerine Size Classification via Computer Vision," 2021 IEEE 8th International Conference on Industrial Engineering and Applications (ICIEA), 2021, pp. 571-574, doi: 10.1109/ICIEA52957.2021.9436701.

[8] T. Van Phat, D. X. Tien, Q. Pham, N. Pham and B. T. Nguyen, "Vietnamese food recognition using convolutional neural networks," 2017 9th International Conference on Knowledge and Systems Engineering (KSE), 2017, pp. 124-129, doi: 10.1109/KSE.2017.8119446.

[9] M. A. Subhi and S. Md. Ali, "A Deep Convolutional Neural Network for Food Detection and Recognition," 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), 2018, pp. 284-287, doi: 10.1109/IECBES.2018.8626720.

[10] A. M. Uddin, A. Al Miraj, M. Sen Sarma, A. Das and M. M. Gani, "Traditional Bengali Food Classification Using Convolutional Neural Network," 2021 IEEE Region 10 Symposium (TENSYMP), 2021, pp. 1-8, doi: 10.1109/TENSYMP52854.2021.9550874.

[11] S. Yadav, Alpana and S. Chand, "Automated Food image Classification using Deep Learning approach," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 542-545, doi: 10.1109/ICACCS51430.2021.9441889.

[12] A. Samraj, S. D., D. K.A. and O. R., "Food Genre Classification from Food Images by Deep Neural Network with Tensorflow and Keras," 2020 Seventh International Conference on Information Technology Trends (ITT), 2020, pp. 228-231, doi: 10.1109/ITT51279.2020.9320870.

[13] J. R. Rajayogi, G. Manjunath and G. Shobha, "Indian Food Image Classification with Transfer Learning," 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), 2019, pp. 1-4, doi: 10.1109/CSITSS47250.2019.9031051.

[14] P. C. Patil and V. C. Burkapalli, "Food Cuisine Classification by Convolutional Neural Network based Transfer Learning Approach," 2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC), 2021, pp. 1-5, doi: 10.1109/ICMNWC52512.2021.9688333.

[15] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[16] K. Srigurulekha and V. Ramachandran, ”Food image recognition using CNN,” 2020 International Conference on Computer Communication and Informatics (ICCCI), 2020, pp. 1-7, doi: 10.1109/ICCCI48352.2020.9104078.

[17] Z. Abidin, R. I. Borman, F. B. Ananda, P. Prasetyawan, F. Rossi and Y. Jusman, ”Classification of Indonesian Traditional Snacks Based on Image Using Convolutional Neural Network (CNN) Algorithm,” 2021 1st International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), 2021, pp. 18-23, doi: 10.1109/ICE3IS54102.2021.9649707.

[18] A. Deshmukh, G. Punjabi and S. Joshi, ”Enhancing the Food Image Classification Accuracy Using Ensemble of CNNs and HelperNet,” 2021 7th International Conference on Signal Processing and Communication (ICSC), 2021, pp. 166-170, doi: 10.1109/ICSC53193.2021.9673186.

[19] D. J. Attokaren, I. G. Fernandes, A. Sriram, Y. V. S. Murthy and S. G. Koolagudi, ”Food classification from images using convolutional neural networks,” TENCON 2017 - 2017 IEEE Region 10 Conference, 2017, pp. 2801-2806, doi: 10.1109/TENCON.2017.8228338.

[20] K. Yanai and Y. Kawano, ”Food image recognition using deep convolutional network with pre-training and fine-tuning,” 2015 IEEE International Conference on Multimedia Expo Workshops (ICMEW), 2015, pp. 1-6, doi: 10.1109/ICMEW.2015.7169816