



A Food Recognition System for Diabetic Patients Based on an Optimized Bag-of-Features Model

Pranathi D K¹, Roshini Sri S², Rakshitha L³, Shalini A⁴, Sanjay Dhanasekaran⁵,
Suguna A⁶

^{1, 2, 3, 4, 5, 6}Computer and Science Department, Sri Sairam College of Engineering Leo Muthu
Nagar, Anekal, Bangalore-562106

¹sc e23cs087@sairamtap.edu.in , ²sce23cs005@sairamtap.edu.in,

³sce23cs070@sairamtap.edu.in , ⁴sce23cs020@sairamtap.edu.in ,

⁵sce23cs016@sairamtap.edu.in , ⁶sugunaa.cse@sairamce.edu.in

ABSTRACT

Healthcare delivery might be completely changed by computer vision technologies. It might be used, in instance, to assess the carbohydrate content of meals for diabetes patients. Based on the bag-of-features (BoF) paradigm, this paper suggests a way for automated food recognition. A thorough technical analysis was carried out to identify and optimise the components that make up the BoF architecture in order to guarantee the best outcomes. Also estimated were the associated parameters. A visual dataset with about 5000 food photos that was divided into 11 groups was developed in order to construct and test the prototype system. Using the scale-invariant feature transform on the HSV colour space, the optimised method calculates dense local features. After that, it utilises a visual dictionary with 10,000 visual words to create a linear support vector machine classifier is then used to categorise the food photos. The system demonstrated the viability of the suggested strategy in a very difficult picture dataset by achieving a classification accuracy of about 88%. By giving diabetes patients a more precise and effective means to control their carbohydrate consumption, this cutting-edge technology might have a profound effect on their quality of life.

Keywords: Based on the bag-of-features (BoF)

1 Introduction

Patients with Type 1 diabetes (T1D) get daily injections of exogenous insulin as part of their therapy. In order to counteract the impact of a meal, a prandial insulin dosage is administered. On the other hand, determining the prandial dosage is a difficult and time-consuming operation that depends on a variety of variables, with carbohydrate (CHO) counts being a crucial component. Clinical trials have demonstrated that postprandial management in children and adolescents receiving intense insulin treatment is not compromised by a CHO counting error of less than 10 g. However, a difference of less than 30 g has a big effect on postprandial glycaemia. It can be challenging for T1D patients, even those who are trained, to estimate their CHO intake precisely, which might result in less than ideal glycemic control. Various technologies have been developed, such as continuous glucose monitoring, to increase the



precision of prandial insulin administration. Pumps and continuous glucose monitoring. Through the provision of real-time glucose measurements and the delivery of insulin dosages based on those readings, these devices can assist patients in better managing their diabetes. For optimum glycemic control in T1D patients, precise prandial insulin dose is essential. Despite being a crucial step in this procedure, CHO counting can be difficult for even experienced individuals. The management of T1D can be made more accurate and straightforward with the use of cutting-edge technology like CGM and insulin pumps. Systems that can help T1D patients correctly measure their diet are now more important than ever due to the surge in diabetes individuals throughout the world. Unfortunately, a lot of diabetes people find it challenging to precisely measure their carbohydrate intake. The research has suggested a wide range of mobile phone applications to solve this problem. These programmes include anything from interactive diaries to food tracking using on-body sensors. Image and video analysis-based applications for diet control have been released thanks to improvements in computer vision technology and the processing capability of mobile devices. The user often uses their phone's camera to take a photo of their future meal. After that, the picture is analysed locally or remotely to extract a number of characteristics that detail its aesthetic attributes. The classifier uses these attributes to identify the various food categories in the image. The estimated carbohydrate content of the meal is then determined using this information. Overall, these mobile phone applications might be a huge help to T1D sufferers in controlling their food and keeping up their health. Patients who carefully monitor their carbohydrate consumption are better able to control their blood sugar levels and prevent diabetes-related problems. Fast food photographs may be divided into four different groups using a food identification tool created by Shroff et al. [9]. Each segmented food item is represented by a vector of colour, size, texture, form, and context-based attributes in the application. An excellent identification accuracy of 96%, 85%, 89%, and 92% for hamburgers, fries, chicken nuggets, and apple pies, respectively, is achieved by feeding these attributes into a feed-forward artificial neural network (ANN). The topic of food recognition has also been advanced by Zhu who identified 22 different food classes by combining a support vector machine (SVM) classifier with a collection of colour and texture data. Their methodology produced an identification rate of 84% for fake food and 59% for actual food items. Using scale-invariant feature transform (SIFT) data grouped into visual words, Kong and Tan [11] have suggested a novel approach for food recognition. Then, using a straightforward Bayesian probabilistic classifier, these attributes are used to match the food items to a library of photos of fast food, homecooked meals, and fruits. Given that there are more than 50 references per food class in the database and fewer than six food items to be recognised, their method has produced a recognition performance of 95%. Overall, these developments in food identification technology have the potential to completely transform the food sector by making precise and effective tracking and analysis of food possible.



2 Existing System

A substantial addition to the field of food recognition has been made by the current work. We have developed a visual dataset with approximately 5000 photographs of home-cooked meals that properly depicts dietary practises in central Europe. Eleven classes with strong intravariability have been created from the photos. Six classifiers, fourteen local image descriptors, two clustering algorithms for building the visual dictionary, and three key point extraction techniques were all examined and compared to make sure our conclusions were accurate. Additionally, we experimented with feature selection and descriptor fusion. Our rigorous testing allowed us to demonstrate the effects of many factors, including the quantity of important points collected, the size(s) of the descriptors, and the quantity of visual words. These discoveries will be extremely helpful to researchers and experts in the subject of food identification. We are optimistic that our findings will have a big influence on the profession since this study offers a thorough and complete investigation of food identification.

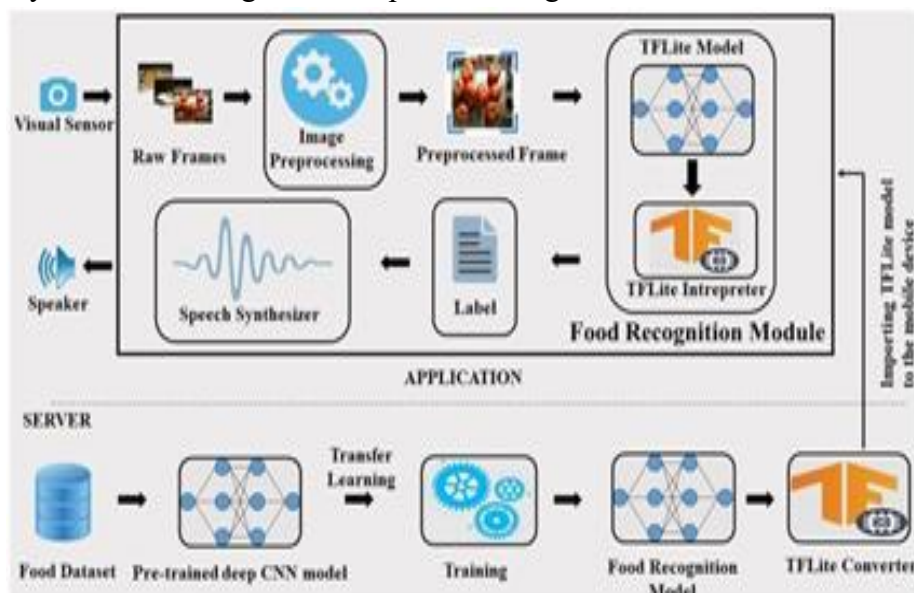


Figure 1: Existing System

Disadvantages

The system learns the lexicon offline while it is in the training phase. In light of the system's performance in terms of recognition, its efficiency is therefore viewed as being of small consequence. It is important to note that improving the training vectors has little effect on the effectiveness of classification as a whole. Additionally, it's critical to remember that the branch factor's powers determine how many words may be generated using hk-means.

3 Proposed System

Food Image Description and Food Image Classification are the first two steps of the suggested food identification system. A collection of traits that reflect the image's visual content are retrieved and quantified in the first step. In the second stage, a classifier uses this collection of



features as input and assigns the image to one of the established food groups. Training and testing are two processes that both stages' design and development go through. The algorithm picks up new information during the training phase while recognising food categories from fresh, untried photos during the testing phase. By offering a quicker and more precise method of classifying foods, this food identification technology has the potential to revolutionise the food sector. It a useful tool for food identification. Overall, the suggested technique for identifying foods is a potential advancement in the realm of food technology. Its capacity to enhance food categorization and identification may result in greater food safety, inventory control, and general consumer pleasure. The food image description the BoF model has been used to successfully describe the visual qualities of numerous food groups. There is no conventional spatial layout within each class, and this approach has been demonstrated to accommodate a broad variety of visual variability. This method enables us to succinctly and clearly describe how various dishes seem while maintaining accuracy. The classification of food images is a vital step in both the computer vision training and testing processes. We carried out multiple trials with three supervised classification techniques: Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Random Forests (RF), in order to determine the best classifier for a given task. The most popular classifier among Bag of Features (BoF) techniques is SVM with linear or nonlinear kernels. Another well-liked machine learning technique for tackling challenging computer vision issues is the artificial neural network (ANN). In this work, we used two alternative feedforward ANN models: a linear model without hidden layers (ANNnh) and a nonlinear model with one hidden layer (ANNwh). While ANNwh trained with the scaled conjugate gradient back-propagation approach, ANNnh employed the standard gradient-descent algorithm. The goal of our tests was to find the best classifier for identifying food images. We were able to identify the most effective strategy for our unique challenge by contrasting the performance of SVM, ANNnh, and ANNwh. Our findings will aid in the creation of methods for classifying food images that are both accurate and effective.

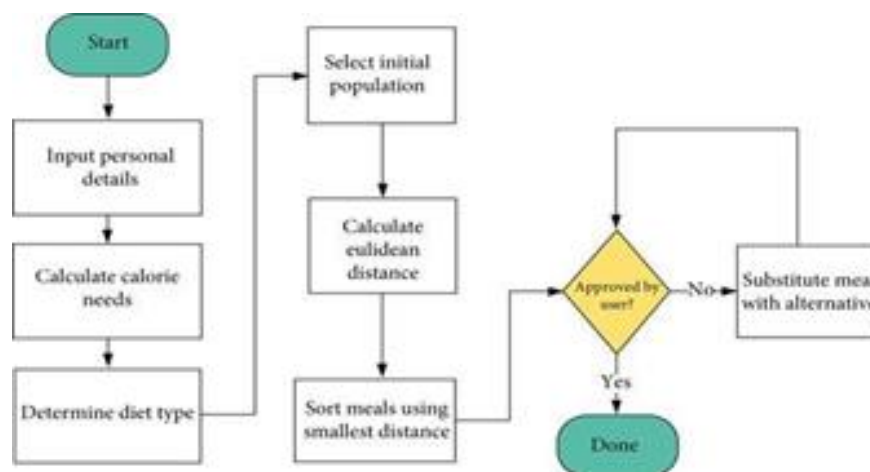


Figure 2: Block Diagram of Proposed System



Practical Applications

We effectively optimised the BoF-based system after rigorous testing with different elements and factors related to the proposed food identification architecture. With our method, we compute a set of hsvSIFT descriptors for every picture in densely extracted rectangular patches with widths of 6, 14, and 42 and spacings of 9, 13, and 15, respectively. Over 1 million randomly chosen descriptors were clustered into 20,000 clusters using the hkmeans method, which also employed a branch factor of 200 and two hierarchical levels. In order to normalise the resultant histograms of visual words, the L2-norm of each histogram was set to 1, the smallest value was taken out, and the highest value was divided. We used SVMlinear and empirically established the best classification algorithm for each image. In addition, we supplied the SVM with class weights, which are the inverse of the occurrence frequency for each class in the dataset. Our improved method has shown to be quite proficient at correctly identifying a variety of foods.

4 Methodology

The two steps of the suggested food identification system are picture classification and food image description. To represent the content of the image, a collection of visual traits is first extracted and measured. The second stage, in which a classifier assigns the image to one of many predetermined food classifications, begins with this collection of features as input. Training and testing are the two processes that make up the design and development of both stages. The system picks up new information during the training phase, while during the testing phase it recognises food categories from fresh, untested photographs. By offering a quicker, more precise method of identifying and categorising food products, this food recognition technology has the potential to completely transform the food sector. It can be employed in a variety of locations, including eateries, grocers, and kitchens, to assure quality control and raise client happiness. The system is useful for the food business because of its capacity to learn and adjust to new information.

5 Results and Discussion

This study introduces a unique BoF-based method for food picture categorization, which represents a fundamental first step towards the development of a portable application that provides nutritional guidance to diabetes patients via automatic CHO counting. To do this, we chose and optimised the system's parts and settings in a series of five significant studies. 4868 photos of foods from 11 different food classes from central Europe were used in our studies, which were performed using a recently created food image collection. The Bag-of-Features (BoF) methodology, a popular picture categorization method, is what our suggested system makes use of. Each picture may be visualised as a histogram of visual words by extracting and quantizing local image elements. Following that, we employ a Support Vector Machine (SVM) classifier to Place the photographs in the appropriate food categories. In our tests, we determined which feature extraction, quantization, and SVM kernel techniques were the most



suitable. For the highest level of classification accuracy, we additionally optimised the settings of each component. According to our findings, the suggested approach achieves a remarkable classification accuracy of 82.5%. In conclusion, our BoF-based system for food picture recognition is an encouraging step towards the creation of a portable application that may give nutritional recommendations to diabetes patients through automatic CHO counting. Our research shows that our method works as intended, and we think that diabetes patients' life might be considerably improved by it.

6 Conclusion

The first experiment proved the advantages of dense key point extraction, resulting in a high number of patches with little overlap. We investigated the effect of descriptor size on performance in the second experiment. Descriptors of sizes 16, 24, and 32 were combined to provide the best results. By accounting for differences in food scale between photos, this method gave the BoF system multiresolution capabilities, which increased performance. Swift was chosen as the best efficient colour and texture descriptor out of 14 alternatives. With the use of this differential descriptor, which describes local texture in all three colour channels of the HSV colour space, colour information may be included while still preserving consistency in intensity and colour variations. The hswSIFT's capacity for colour capture was verified by the experiment with combining the best colour descriptors with other descriptors, which failed to improve performance. We contrasted k-means with hk-means in terms of the learning of the visual dictionary. The results from the latter were almost equal to those from the k-means, with 20,000 visual words being the ideal quantity. While fewer words produced worse results, additional words had no discernible impact on output. Two linear and four nonlinear machine learning algorithms were used in the final classification, with the linear approach producing the best results, especially for features containing a lot of data. The high dimension of the feature space, which to some part makes the issue linearly separable, is probably to blame for this. The final system, which was optimised, had an ORA of about 78%, proving the viability of a BoF-based system for identifying foods. In order to further increase accuracy, we intend to investigate a hierarchical categorization strategy that combines visually comparable elements classes for the initial levels of the hierarchical model, from which it is possible to discriminate at a later level by utilising the proper discriminative characteristics.

References

1. C. E. Smart, K. Ross, J. A. Edge, B. R. King, P. McElduff, and C. E. Collins, Year: 2019, "Can children with type 1 diabetes and their carers estimate the carbohydrate content of meals and snacks?" *Diabetes Medicine*, Vol: 27, pp. 348-353.
2. O. Amft and G. Troster, Year: 2017, "Recognition of dietary activity events using on-body sensors, *Artif. Intell. Med*, Vol: 42, no: 2, pp. 121-136.



3. F. Mindru, T. Tuytelaars, L. Van Gool, and T. Year: 2020, "Moons. Moment invariants for recognition under shifting perspective and light, Comput. Vis. Image Understanding, Vol: 94, pp. 3-27.
4. K. G. M. M. Alberti and P. Z. Zimmet. Year: 2017, Diabetes mellitus: Definition, Diagnosis, and Classification, Part 1: Classification and diagnosis of diabetes mellitus. Provisional report from a WHO Consultation, Diabetes Medicine, Vol: 15, no: 7, pp. 539-553.
5. Standards of medical treatment in diabetes, Year: 2021, American Diabetes Association Diabetes Care, Vol: 29, pp. S4–S42.