A Review: Food Recognition for Dietary Assessment/Calorie Measurement using Machine Learning Techniques

Adhira Gupta*
School of Computer Science and Engineering
Shri Mata Vaishno Devi University (SMVDU),Katra, J&K

Sanjay Sharma
School of Computer Science and Engineering
Shri Mata Vaishno Devi University
(SMVDU), Katra, J&K

***Abstract*—Dieticians and healthcare convention sare concerned with the consumption of accurate quantity and right kind of food. There is no doubt that exercising alsoplays a vital role but what we are feeding our body plays a major role in obesity and many problems related to healthlikediabetes,stroke,andmanycardiovasculardiseases.Also,duetoadvancementintechnology,today'sgeneration can order food just with a click on their mobiledevices. Thus, acceleration in obesity is evident. For thepeople who are concerned with this problem, keeping therecordsoftheconsumptionofnutrientsmanuallyisdifficult. To combat this, a variety of health applicationsand Calorie measurement tools have emerged to reverse orshrink the effect of all the health-related troubles. Some oftheapplicationsalsoutilizestate-of-the-artMachineLearning algorithms. In this paper, we will take a look atsome of the methods used for food recognition and caloriemeasurements and also comparing their performance byputtingthemhead-to-headondifferentscales.**

Keywords—Convolutionalneuralnetwork(CNN)·Deeplearning·Food recognition·Machinelearning (ML)

I. INTRODUCTION

Recent studies by the WHO show that in 2016, more than 1.9billion adults aged 18 years and older were overweight. Ofthese over 650 million adults were obese. However, on theother side of the spectrum, a different study reveals that peopleare also leaning towards a healthy lifestyle more than ever intheview of adisease knownasobesity.

Collecting food recordings to keep an eye on the daily calorieintakeandmaintainingasetdietplanisnotasupernewconcept. It was done evenbefore the time when smartphonesandhigh-techspecializeddietarymeasurementtoolswereinvented and became popular. Unlike today, people used tophysically write their daily meals as well as diet plans on apieceofpaperoranotebook.Thisprocesswasclearlyinefficient,dull,andhadagreatamountof errorpossibility.

Modern technology entirely solvedthe issueand convertedthis tedious process into an exciting one by transforming thewholefoodrecordingprocessfromwritingeverythingdown to the matter of clicking just a single picture of the food itemon your smartphone or tablet and evaluating almost all of thepossible nutritional information .This all has become possiblewith the advancement inmachine learning and deep learningmodels. Now, takingphone out of the pocket and clicking apictureofthefoodtocalculatethenumberofcaloriesitcontains sounds so simple and magical but in reality, this jobrequires high skills and lots of complicated calculations. Thisall should work flawlessly behind the scenes so that the finaloutputhasbothhighaccuracywhilemaintaininggreatefficiency. Thatis why there is nosingle overall best methodtoperformfoodrecognition.Overtheyears,somanyresearchersallaroundtheworlddevelopednewandalsorefined existing methods and algorithms by using cutting edgetechniquestoachieve betterresultsthanbefore.

Moreover, one of the main hurdles in this path is to collectmassive datasets used in training the model. Not to mention,there are so many intraclass variations that could happen eveninonespecifictypeoffooditemwhichcouldalsoeasilybecome a major cause of worry down the road. In this paper,wearegoingtoexploreandanalyzedifferentmachinelearning approaches used by researchers for the recognition offood and nutrients estimation. Also we are going to do somecomparativeanalysisdependingondifferentparameterstofind the best approach which can be considered to improve thefoodrecognitionsysteminfuture.

II. RELATEDWORK

A.TraditionalApproaches

Fortreatingpeopleaffectedwithobesity,researchers[1]proposedasysteminwhichtheyidentifieddifferentfooditemsusingthe processofsegmentationbyapplyingtheGabor filter and hence classified them using SVM. Gabor Filter is afilter of a linear type specifically used for texture analysis,meaning that it scans for any specific frequency content in thepicturein certain directions inalocalized region throughoutthepoint.Thenutritionalvaluesofthefooditemswer ecalculatedonthebasisoftheportionoffoodmappedcorrespon ding to the nutrition tables. Also, for the estimationof the portion of food items, a thumb was placed with eachfood item while taking the picture so that it's trouble-free forthe algorithmto estimate the life-size portions of the fooditemswhich resultedin anaccuracy hike to about86%.

Another food recognition and calorie measurement techniquethat was proposed by Turmchokksam et al. [4] uses a uniquecombination of nutritional data of food in addition to foodtemperature and brightness levels information captured by thethermalaswellasaCCD(chargecoupleddevice)camera.Thi sunionofhardwareworkingwiththesoftwaremanaged to give them more accurate outcomes than the othertraditionalmethodsof foodrecognition.

A system was developed by He et al. [5] named Dietcam toovercomethechallengesthatarisewiththeintraclassvariatio ns while doing food detection. It primarily has twomain bits: ingredient detection and food classification. First,the program scans for all the ingredients in the food items bytakingadvantageoftextureverificationandapart-basedmodel. Second, it categorizes the food items with the help of amulti-viewmulti-kernelSupportVectorMachineorSVM.They used DietCam on 15262 images of around 55 differentfood categories and obtained great accuracy on food items thatmadeupof several elements.

M.A.Subhietal.[9]didasurveyandexploredseveraltraditional practices plus some neural networks for the purposeoffood recognition andnutrients estimationbut concludedthatestimatingthevolumeofthefoodisstillthemostc hallenging process. SVM and MLP have been brought intoimplementationusing MATLABto getdesirableresults[17].

.

B.Deeplearning Approaches

Deep learning has been very widely used especially in projectswithhugedatasets because ofits powerfullearningabilityalong with the luxury of automatic extraction of new featuresfrom raw data.[2] Additionally, the Deep Learning algorithmCNNsprovedtobeverysuccessfulinpatternrecognit ion,imageprocessing,andin reducingthe number of parametersby using spatial relationships without compromising on themodel quality.

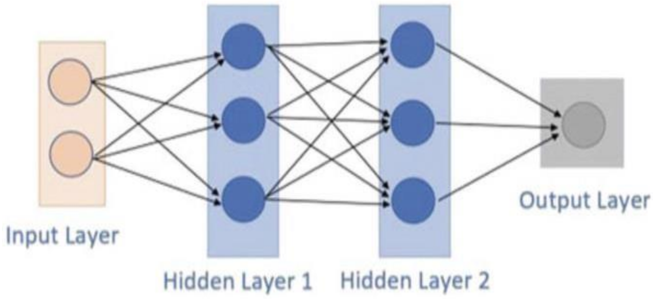

Fig.1Structureofneuralnetwork

Cioccaetal.[3]cameupwithanewfooddatasetconsistingof1,0 27canteentraysfilledwithfooditemsdividedinto73food classes. They used CNN for image recognition andattained an accuracy rate of 79%. Along with that, theysuccessfully build a pipeline that takes an image containingnumerous food items presented in a tray as input, finds theregionofinterestwhichthenfinallyoutputsalistofidentified fooditems.

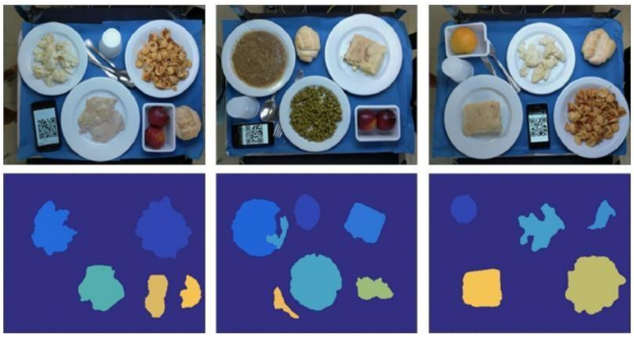

Fig.2 Examples of Segmentation results of some canteen trayimages

Authors [7] focussed on creating a new dataset (Fruits-360)consistingof90483high-qualityimagesoffruits.Theydetecteddifferentfruitclassesusin gneuralnetworksandobtainedanastonishingaccuracyof95%. In[8],afoodrecognitionsystemwasdesignedwhereedgecomp utingserviceswereemployedandnewalgorithmswereevolved basedondeeplearningandimagepre-processing.Segmentationalgorithmswerealsodevelopedfore nsuringthat the quality of the food images is adequate. The systemconsistedofthreemaincomponents:FrontendCompon ent(wherewatershedsegmentationalgorithmwasused),Backe ndComponent(whereCNNbasedalgorithmswereused),andth eCommunicationComponent(CC).Theexperimentwascond uctedonUEC-256, UEC-100, andFood-

101datasetswhichobtainedanaccuracyof(63-87)%, (76-94)%,and(77-94)%respectively.

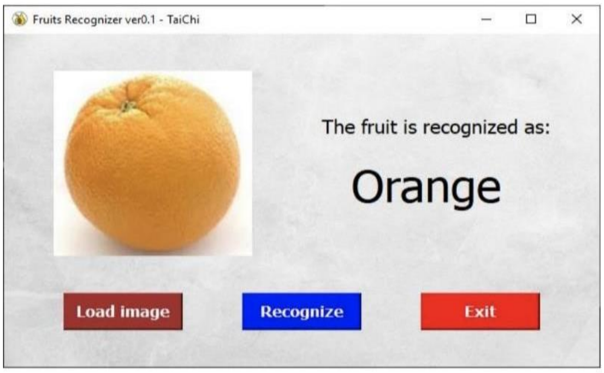

Fig3 :Mainwindowoftheprogram[7]

Moreover, volume estimation can be performed only for thesolidfood itemsand thattooisanerror-proneprocess.L.Jiang et al [10] took two datasets: UEC-FOOD100 and UEC-FOOD256 for training and testing their deep learning modelforrecognizingfooditemsandimplementingnutritional analysis using the Region Proposal Network (RPN) which is apart of the Faster R-CNN model. The project executes in athree-stepprocesswhichincorporatesdetectingregionsofinterest, applying feature maps over them and then identifyingthecomponentsofeachpicture.Intheend,adietaryassessment report is generated based on the existing data andtheresearcherssawincredibleresultsfromthefoodrecognitiondeeplearningmodel.

Nearly all of the food recognition techniques that exist useCNNs and almost every one of them struggles to get excellentresultsasitbecomestoughforthemodeltohuntforthecommonsemanticpatternsinfooditemsbecauseoftheircountless appearances and intraclass variations. However, ascheme called (MSMVFA) or Multi-Scale Multi-View FeatureAggregation can be utilized to eliminate all of those issues. Itcanbe usedtoaggregate various features withina unifiedrepresentation which then resolves the problem of commonpattern recognition in food images. Various tests have beendone on some of the most popular food datasets using thisscheme which has shown exceptional results.[11]. In additiontothis,researchersinitiatedexploringandcomparingdifferent attribute to find the best way to recognize food andthem applying the best method to estimate the calories presentin food [12][13] .In [14] , the latest vision-based methods havebeen more explorative to outline the present approaches andmethodswhichareusedforautomaticdietaryassessment,feasibility,and challengeswhich are unaddressed .

TogetridoftheoverfittingprobleminDNN,differentclassifiers were combined to get better performance [15] andthedatasetsusedwereFood-11andFood-101.Whencomparedwithstateoftheartmethods,transferlearningdeliversmoreoptimizedresults[16].Alsosixvotingcombinationruleswereappliedon UEC-FOOD256,Food-101andUEC-FOOD100whichshowedanaccuracyof77.20%, 84.28%and 84.52%respectively.Anewdatabasewascreatedwhichconsistedof9classesofhealthdrinkpowders[19],D-CNNwasusedforpredictingtheproteincontent and image attributes with linear regression were alsoused.Anerrorofapprox±2.71wasfoundinpredictingprotein content and deep learning improved predicting errorby±1.96.In[18],authorspreparedtheirowndatasetconsistingof360categoriesoffoodwhereeachclassconsisted of at least 500 images . Theymade use of D-CNNfor identifying cooking methods,food ingredients and dishtypesandreceived agood accuracyof81.55%.

In [20], a system called goFOOD was introduced for calorieestimation. They evaluated their system on MADiMa databasewhich contains 319 fine grained categories of food and theyalso prepared their own Fast food database which consisted of20 meals which included 14 different categories of food. Theyobtained lowest top-3 accuracy of 71.8%for fine grainedcategories.

III. CHALLENGES

A. *Identificationoffood isNotEasy*

The procedureof analysis of foodis quite reliable andshowsresultsaccuratelywiththecomputervisionmethodology .But still it is very challenging process to recognize differentkinds of food in the image correctly. There are a large numberof image recognition tools available currently, the methodsincluding identification of food are still dependent on self-reporteddietaryintakes.Itisduetothereasonthatdeformability can be seen in food items when compared toother things present in the real world. It is difficult to define afooditem'sstructure,andahighintra-class(similarfoodslookvery different) and low inter-class (different foods look verysimilar)variancealsoexist.

B. *Timeconsuming process*

Time is also a major factor in this process as it takes quite longto train the model. However, the training time hugely dependson the computation power of the machine.Training time alsoincreasesastheimageandclassesinthedatasetgetsmassive

C. *Overfitting*

Overfitting could become a serious issue so it should be takencare of while tuning the model parameters. It occurs when themodel picks up the noise and random details present in thedata.Many researches on Food recognition show that it is acommon issue which drops the accuracy to a significant bit.longtotrainthemodel.dependsonthecomputationpower

D. Volumeestimation

Volume estimation still remains the most challenging area aspredicting the portion size of the food by looking at the 2Dimages is far from acceptable range. Also ,volume estimationmethodsarenotapplicabletoliquiditems.Theycanbeappliedonlyonsolidandclearlydistinguishableitemslikefruit.

E. Nutrientandcalorieestimation

This stage remains the most error prone stage as it depends onthe food recognition and volume estimation of the food. Ifthere is any error in the above stages ,it will automaticallygeneratewrongresults.

Table1:Tabularrepresentationofthesurvey

R.No./Year	Approach	DatasetUsed	Advantages	Drawbacks	Accuracy
[1](2014)	Gabor filter andSVM	-	- Uses a devicebuilt-incameratocaptureimages. - Also measuresthe volume of theportions of thefood items.	- Mixed or Liquidfood items are notsupported. - Couldnotachieve accuracyas high as othermethods studiedin thispaper.	86%
[3](2016)	CNNsandJPEGalgorithm	UNIMIB2016	- Largefoodcategoriessupport. - Custom madedataset which hasfooditemsplacedin trays for betteroverall results inrecognition.	-Segmentationprocess is notautomatic andthereforeconsumes moretime. - Accuracy is notvery high	79%
[4](2018)	Fuzzy C-means,weighted FCMand SVM	-	-Uniquetechnique forfood recognitionwhich producesacceptableresults	-Cost of thesystem is veryexpensive as itneeds thermalcameras forfunctioning.	2.28%& 2.21% errorinhardwareandsoftwarerespectively.
[5](2015)	Multi-viewrecognition andMulti KernelSVM	PFID	-Great andreliableresultsinfood items madeupofcomplicatedelementscompared tocommonly usedmethods.	-Real-timeperformance isnotexceptional. -Database islimited and somefood categoriesarenotcovered.	90% (forgeneralfood items)followed by85% fordifficultcategories(DCs)
[7](2019)	EfficientNetArchitecture andCNNs	(Fruits-360)	-Accuracyinthissystem isimpressive.	-Model onlyrecognizes fooditemsatthispoint.	95%

[8](2018)	Watershedsegmentatio nAlgorithm andCNNs	UEC-256,UEC-100 andFood-101	-Response timeand Energyconsumption ofthe system isclose to minimalof existingtechniques. -Performance isbetter thanexistingprinciples.	-Response time isfastbutitisstill5percent slowerthan the bestsystem availablein the market.	(63-87)%, (76-94)%, and(77-94)% respectively.
[11](2020)	CNN's,Multi-Scale Multi-View FeatureAggregation	ETHFood-101, VireoFood-172andChineseFoo-dNet	-Model achievesstate-of-the-artrecognitionperforman ce onsome of the bestdatasets.	-This methoddoes not provideperfect accuracyfor some foodcategories.	90.59%, 98.31 %and96.94 % respectively.
[12](2015)	CNN: GoogLeNet,NegativeCl assifier	Food-101	-This Model isable to get highperformancenumber s andconverge faster toadapt the newfood categoriesbecauseof the "negativeclassifier".	-Accuracy scoreis notmagnificent.	78.11%
[13](2015)	CNN,GMM	Fish and UECFOOD-100	-Betterperformance thansimple linearmodel	-	-
[16](2020)	Voting-basedfine-tunedCNNs, ResNet,GoogleNet,VG GNet, andInceptionV3	Food-101,UEC-FOOD100,and UEC-FOOD256	Promising resultscompared to thefinest methods onFood-101, UEC-FOOD100, andUEC-FOOD256 imagedatasets.	-Slightly lesseraccuracy andspeed comparedto the use ofsingle CNNarchitectures.	84.28%, 84.52 %and77.20 % respectively.
[17](2021)	MLP, SVM,MATLAB, MultilayerPerceptionm odel	-	Project showedacceptable resultsusing both SVMand improvedMLPmodels.	Complex fooditems are notconsidered.	96.5% (forSVM) 97.2%(forMLP)
[18](2016)	CNNs, Multi-tasklearning	ILSVRC	-Higher accuracyscore thantraditionalmethods. -Includes	Results are notalwaysoptimalasexpe cted.	81.55%
			cooking methodrecognizer andingredientdetector forreferenceinformation incase, the dishesare not present inthe trainingdataset.		
[19](2019)	Deep Learning,LinearRegress ion usingSVM	IDPE	-Predictionaccuracy isacceptableconsidering thetypeofproject.	-Extra equipmentis required tominimize theerrorfigures.	-
[20](2020)	CNN's, Segmentation	MADiMadatabase	Supports a widerange of foodcategories.	Dietaryassessment resultswere inferiorcompared toprofessionals insomecases.	71.8%

IV. Conclusion

In this paper, we have surveyed numerousmethods of foodrecognition, some are traditional which uses SVM and othersare more advanced and hence quicker and more precise, mostoftheseapplythepowerofDeeplearningandNeuralNetworks.Someauthorsalsodevelopedconvenientandeasyto use softwares for the user to simply take pictures with theirdeviceinrealtimeandobtainnutritionalinformationandcalorie estimation of the food. Additionally, the accuracy ofthesesystemswillcontinuetoimprovefromhereincludingthe more accurate volume estimation and better coverage of avariety offoodcategories.

REFERENCES

[1] P.Pouladzadeh,S.ShirmohammadiandR.Al-Maghrabi,"MeasuringCalorieandNutritionFromFoodImage,"inIEEETransactionsonInstrumentation and Measurement, vol. 63, no. 8, pp.1947-1956,Aug. 2014

[2] P. Kuang, W. Cao and Q. Wu, "Preview on structures and algorithms of deep learning," 2014 11th International Computer Conference on Wavelet Actiev Media Technology and Information Processing(ICCWAMTIP), 2014, pp. 176-179, doi: 10.1109/ICCWAMTIP.2014.7073385.

[3] G. Ciocca, P. Napoletano and R. Schettini, "Food Recognition: A New Dataset, Experiments, and Results," in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, pp. 588-598, May 2017, doi: 10.1109/JBHI.2016.2636441.

[4] S. Turmchokkasam and K. Chamnongthai, "The Design and Implementation of an Ingredient-Based Food Calorie Estimation System Using Nutrition Knowledge and Fusion of Brightness and Heat Information," in IEEE Access, vol. 6, pp. 46863-46876, 2018, doi: 10.1109/ACCESS.2018.2837046.

[5] H. He, F. Kong and J. Tan, "DietCam: Multiview Food Recognition Using a Multikernel SVM," in IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 3, pp. 848-855, May 2016, doi: 10.1109/JBHI.2015.2419251.

[6] .https://en.wikipedia.org/wiki/Machine_learning

[7] ChungDTP,TaiDV(2019)Afruitsrecognitionsystem based on a modern deep learning technique.In:IOPConferenceSeries:JournalofPhysics

[8] C. Liu et al., "A New Deep Learning-Based Food Recognition System for Dietary Assessment on An Edge Computing Service Infrastructure," in IEEE Transactions on Services Computing, vol. 11, no. 2, pp. 249-261, 1 March-April 2018, doi: 10.1109/TSC.2017.2662008.

[9] M. A. Subhi, S. H. Ali and M. A. Mohammed, "Vision-Based Approaches for Automatic Food Recognition and Dietary Assessment: A Survey," in IEEE Access, vol. 7, pp. 35370-35381, 2019, doi: 10.1109/ACCESS.2019.2904519.

[10] L. Jiang, B. Qiu, X. Liu, C. Huang and K. Lin, "DeepFood: Food Image Analysis and Dietary Assessment via Deep Model," in IEEE Access, vol. 8, pp. 47477-47489, 2020, doi: 10.1109/ACCESS.2020.2973625.

[11] S. Jiang, W. Min, L. Liu and Z. Luo, "Multi-Scale Multi-View Deep Feature Aggregation for Food Recognition," in IEEE Transactions on Image Processing, vol. 29, pp. 265-276, 2020, doi: 10.1109/TIP.2019.2929447.

[12] S. Ao and C. X. Ling, "Adapting New Categories for Food Recognition with Deep Representation," 2015 IEEE International Conference on Data Mining Workshop (ICDMW), 2015, pp. 1196-1203, doi: 10.1109/ICDMW.2015.203.

[13] Z. Ge, C. McCool, C. Sanderson and P. Corke, "Modelling local deep convolutional neural network features to improve fine-grained image classification," 2015 IEEE International Conference on Image Processing (ICIP), 2015, pp. 4112-4116, doi: 10.1109/ICIP.2015.7351579.

[14] K, Kagaya H, Ogawa M (2014) Food detection andrecognition using convolutional neural network. In:'14Proceedingsofthe22ndACMinternationalconferenceonmultimediadoi.org:10.1145/2647868.2654970

[15] Aguilar E., Bolaños M., Radeva P. (2017) Food Recognition Using Fusion of Classifiers Based on CNNs. In: Battiato S., Gallo G., Schettini R., Stanco F. (eds) Image Analysis and Processing - ICIAP 2017. ICIAP 2017. Lecture Notes in Computer Science, vol 10485. Springer, Cham. doi.org :0.1007/978-3-319-68548-9_20

[16] Tasci, E. Voting combinations-based ensemble of fine-tuned convolutional neural networks for food image recognition. Multimed Tools Appl 79, 30397–30418 (2020). doi: 10.1007/s11042-020-09486-1

[17] Kumar, R.D., Julie, E.G., Robinson, Y.H. et al. Recognition of food type and calorie estimation using neural network. J Supercomput (2021). https://doi.org/10.1007/s11227-021-03622-w

[18] Zhang XJ, Lu YF, Zhang SH. Multi-task learning for food identification and analysis with deep convolutional neural networks.JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 31(3): 489–500 May 2016. DOI 10.1007/s11390-016-1642-6

[19] Shermila, P.J., Milton, A. Estimation of protein from the images of health drink powders. J Food Sci Technol 57, 1887–1895 (2020). https://doi.org/10.1007/s13197-019-04224-4

[20] Lu, Y.; Stathopoulou, T.; Vasiloglou, M.F.; Pinault, L.F.; Kiley, C.; Spanakis, E.K.; Mougiakakou, S. goFOODTM: An Artificial Intelligence System for Dietary Assessment. Sensors 2020, 20, 4283. https://doi.org/10.3390/s20154283

[21] Fatehah AA, Poh BK, Shanita SN, Wong JE. Feasibility of Reviewing Digital Food Images for Dietary Assessment among Nutrition Professionals. Nutrients. 2018;10(8):984. Published 2018 Jul 27. doi:10.3390/nu10080984

[22] Singla, L. Yuan, and T. Ebrahimi, ''Food/non-foodimageclassificationandfoodcategorizationusingpre-trainedgooglenetmodel,''inProc.2ndInt.WorkshopMultimediaAssist.DietaryManage.,2016,pp.3–11.

[23] C.Termritthikun,P.Muneesawang,andS.Kanprachar,''NU-InNet:Thaifoodimagerecognition using convolutional neural networks onsmartphone,''J.Telecommun.,Electron.Comput.Eng.,vol. 9, nos.2–6,pp.63–67, 201724.

[24] Y. He, C. Xu, N. Khanna, C. J. Boushey, and E. J.Delp,''Analysisoffoodimages:Featuresandclassification,''inProc.IEEEInt.Conf.ImageProcess.(ICIP), Oct. 2014, pp.2744–2748doi: 10.1109/ICIP.2014.7025555