VISVESVARAYA TECHNOLOGICAL UNIVERSITY
BELGAUM, KARNATAKA

An Internship report
on
" FOOD CLASSIFIER AND NUTRITION INTERPRETER TOOL (FCNI)"

Submitted in partial fulfilment of the requirements for the award of the Degree

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

Prescribed by

Visvesvaraya Technological University

Submitted By

NIHAR.M 1BY15EC049

Under the Guidance of

Mrs. Anandi Giridharan
Principal Research Scientist
ECE dept, IISc

Mr.Suryakanth B
Assistant Professor, Dept. of ECE,
BMSIT & M

Department of Electronics and Communication Engineering B.M.S. Institute of Technology and Management
Yelahanka, Bengaluru – 560064

2019-20

B.M.S. INSTITUTE OF TECHNOLOGY AND MANAGEMENT

Yelahanka, Bengaluru-560064

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

Vision

Be a Pioneer in Providing Quality in Education in Electronics, Communication and Allied Engineering Fields to Serve as Valuable Resource for Industry and Society

Mission

1. **Impart Sound Theoretical Concepts & Practical Skills through innovative Pedagogy.**
2. **Promote Interdisciplinary Research.**
3. **Inculcate Professional Ethics.**

Program Outcomes (P.O.s)

1. **Engineering knowledge:** Apply the knowledge of Mathematics, Science, Engineering fundamentals and an engineering specialization to the solution of complex engineering problems
2. **Problem analysis:** Identity, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, Natural sciences and engineering sciences
3. **Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. **Conduct investigations of complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the Information to provide valid conclusions.
5. **Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern Engineering and I.T. tools, including prediction and modelling to complex engineering activities with an understanding of the limitations.

6. **The engineer and society:** Apply to reason informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for Sustainable development
8. **Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and teamwork:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings
10. **Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project management and finance:** Demonstrate knowledge and understanding of the Engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

12. **Life-long learning:** Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological.

Program Educational Objectives (P.E.O.s)

Our Graduates will be able to

P.E.O. – 1: Work as Professionals in the area of Electronics, Communication, and Allied Engineering fields.
P.E.O. – 2: Pursue Higher studies and involve in Interdisciplinary research work.
P.E.O. – 3: Exhibit Ethics, Professional skills and Leadership qualities in their profession.

Program Specific Outcomes (PSOs)

PSO – 1: Exhibit Competency in Embedded System and VLSI Design
PSO – 2: Capability to comprehend the technological advancements in Radio Frequency Communication and Digital Signal Processing.

BMS INSTITUTE OF TECHNOLOGY AND MANAGEMENT
Yelahanka, Bengaluru – 560064

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

CERTIFICATE

Certified that the Technical Seminar entitled “Food Classifier and Nutrition Interpreter Tool (FCNI)” presented by Mr.**NIHAR.M**, U.S.N. **1BY15EC049**, a bonafide student of **B.M.S Institute of Technology & Management** in partial fulfilment for the award of **Bachelor of Engineering** in **Electronics and Communication Engineering** of the Visvesvaraya Technological University, Belgavi during the year 2019-20. It is certified that all of the corrections/suggestions indicated for Internal Assessment have been incorporated in the report deposited in the departmental library. The technical seminar report has been approved as it satisfies the academic requirements in respect of Seminar work prescribed for the said Degree.

Signature of HOD	Signature of Guide	Signature of Principal
Dr. Ambika R	Prof.Suryakanth B	Dr. Mohan Babu G.N.
Dept. of ECE	Dept. of ECE	BMSIT&M

External Viva

Name of the Examiners
Signature with Date

1.
2.

INTERNSHIP CERTIFICATE

06.05.2020

DEPARTMENT OF

ELECTRICAL COMMUNICATION

ENGINEERING

INDIAN INSTITUTE OF SCIENCE BANGALORE-560012

TO WHOM IT MAY CONCERN

This is to certify that **Mr NIHAR.M**, fourth year, Eighth Semester student of Electronics and Communication Department of BMS Institute of Technology and Management has successfully completed the internship for the period of 2 months **01-March-2020 to 30April-2020**. He worked on "**Food Classifier and Nutrition Interpreter tool**" under my guidance at the Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore. He has successfully completed his internship. I wish him all the best for future endeavours.

Signature

Anandi Giridharan

Principal Research Scientist

ECE, Indian Institute of Science

06-05-2020

Acknowledgement

I express my sincere gratitude to Mrs. Anandi Giridharan, Principal Research Scientist, ECE Department, IISc, Bangalore, for the patient guidance, encouragement and advice she has provided throughout the project period. I am extremely lucky to have a mentor who cared so much about the work and responded to my questions & queries so promptly. Also, in spite of her busy schedule, she spent enough time in explaining the project requirements and made me to be on the right path to carry out the project successfully.

It is my privilege and pleasure to express our profound sense of respect, gratitude and indebtedness **to Dr Mohan Babu G N**, Principal, BMSIT&M, for guiding and providing facilities for the successful completion of my technical seminar.

I sincerely thank **Dr Ambika R**, Professor & H.O.D., Department of Electronics and Communication Engineering, BMSIT&M, for his valuable support and constant encouragement is given to me during this work.

I thank my guide, **Prof. Suryakanth B**, for his valuable time, patience, suggestion and periodic evaluation that was conducive for the project.

I also thank **Prof.Vinutha.B**, internship coordinator, for their contributions and technical support in conducting seminar and preparing this report.

I would also like to thank all the teaching and non-teaching staff of the Department of Electronics and Communication Engineering for their co-operation and motivation.

Finally, we express our cordial thanks to our parents and friends for their support and guidance throughout the project preparation.

NIHAR.M

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter-1

INTRODUCTION

The Indian Institute of Science (IISc) is a public, deemed, research university for higher education and research in science, engineering, design, and management. It is located in Bangalore city, in the Indian state of Karnataka. The institute was established in 1909 with active support from Jamsetji Tata and Krishna Raja Wadiyar IV and thus is also locally known as the "Tata Institute". It was granted the deemed to be university status in 1958 and the Institute of Eminence status in 2018.

The Institute places equal emphasis on student learning, with several postgraduate and PhD programmes, and a new dedicated four-year undergraduate programme, aimed at encouraging bright young students to pursue careers in the basic sciences.

1.1 History:

It was an accidental meeting between Jamsetji Tata and Swami Vivekananda, on a ship in 1893 where they discussed Tata's plan of bringing the steel industry to India, Tata wrote to Vivekananda five years later: "I trust, you remember me as a fellow-traveller on your voyage from Japan to Chicago. I very much recall at this moment your views on the growth of the ascetic spirit in India... I recall these ideas in connection with my scheme of Research Institute of Science for India, of which you have doubtless heard or read."

Impressed by Vivekananda's views on science and leadership abilities, Tata wanted him to guide his campaign. Vivekananda endorsed the project with enthusiasm, and Tata, with the aim of advancing the scientific capabilities of the country, constituted a Provisional Committee to prepare a plan for setting up of an Institute of research and higher education. The committee presented a draft proposal to Lord Curzon on 31 December 1898. Subsequently, Sir William Ramsay, a Nobel Laureate, was called on to propose a suitable place for such an institution who suggested Bangalore as the best location. Mir Osman Ali Khan made the most significant contribution in terms of money which amounted to 3 lakh Rupees over a period of 31 years.

The land and other facilities for the institution were donated by on behalf of the state of Mysore by Krishnaraja Wodeyar IV, and Tata himself. The Karnataka donated about 371 acres (1.50 km^2) of land. (now valued about Rs 200 billion). Tata gave several buildings towards the creation of IISc. State of Karnataka also contributed Rs

500000 (now valued 12.5 million) towards capital expenditure and Rs 50000 (now valued 1.25 million) for annual expense.

The constitution of the Institute was approved by the Viceroy, Lord Minto, and the necessary Vesting Order to enable it to function was signed on 27 May 1909. Early in 1911, the Maharaja of Mysore laid the foundation stone of the Institute, and on 24 July, the first batch of students were admitted to the Departments of General and Applied Chemistry under Norman Rudolf and Electro-Technology under Alferd Hay. Within two months, the Department of Organic Chemistry was opened. In 1958 the institute was granted the deemed university status by the UGC.

At the time of the inception of IISc in 1909, Morris Travers, Sir William Ramsay's co-worker in the discovery of the noble gases, became its first Director. For Travers, this was a natural continuation of his work on the Institute, since he had played a role in its founding. The first Indian Director was the Nobel Laureate Sir C.V.Raman. Raman was the Indian Science-based Nobel Laureate. The current Director is Anurag Kumar.

The Institute was the first to introduce master's programs in Engineering. It has also started integrated doctoral programs in Biological, Chemical, Physical and Mathematical Sciences for natural science graduates.

1.2 Vision :

IISc aims to be among the world's foremost academic institutions through the pursuit of excellence in research and promotion of innovation by offering world-class education to train future leaders in science and technology and by applying science and technology breakthroughs for India's wealth creation and social welfare.

1.3 Mission :

I. Imparting world-class higher education in an environment of fundamental and applied research in science and engineering.

II. Conducting high-impact research, generating new knowledge, and disseminating this knowledge through publications in top journals and conferences.

III. Applying faculty expertise towards the success of national science and technology initiatives.

IV. Applying in-depth knowledge in various areas to create knowhow and developing such knowhow for utilization by industry and society.

1.4 Objective/Purpose :

I. To provide for advanced instruction and to conduct investigations in all branches of of knowledge and, in particular such branches of knowledge as are likely to promote the material and industrial welfare of India

II. To establish and maintain chairs and lectureships in science, Arts and technology,

III. To provide suitable libraries laboratories and equipment's,

IV. To cooperate as far as possible with such recognized institutions as exist or are found in future for cognate objects in India.

V. To do such things as are incidental, necessary or conducive to attainment of all or any of the objects of the Institute.

1.5 Rankings:

QS World University rankings ranked IISc second in the world in terms of citations per faculty, IISc was ranked 251–300 in the world by the Times Higher Education World University Rankings of 2018, the top institute in India, as well as 21 in Asian the 2018 ranking and 14 among BRICS & Emerging Economies University Rankings in 2017. The QS World University Rankings of 2019 ranked IISc 170 in the world, as well as 51 in Asia and 10 among BRICS nations. In 2019, it ranked 353rd among the universities across the globe by SC Imago Institutions Rankings. The Academic Ranking of World Universities ranked it 301–400 in the world in 2017, the only institute in India to be listed by this ranking. It was ranked top university (1st place) in India by the National Institutional Ranking Framework university ranking for 2018 and 2019, and overall ranking for 2018.

1.6 Institute Services And Functions :

The main emphasis of the Institute is on research, and it expects its faculty members to initiate and carry out research programs of a quality that are on par with international standards. The Institute continuously strives to provide the necessary support and the entire setup (including administration) is geared to facilitate the pursuit of excellence. The Institute offers complete freedom to the faculty to pursue their research interests. The Institute offers both graduate and undergraduate programs. It started a unique four-year Bachelor of Science (Research) programme in 2011. The programme is designed as a blend of core science and interdisciplinary topics that will serve as a launching pad for attractive career opportunities in academia and industry.

The Institute faculty receive research funding from various government agencies such as the Aeronautical Research Development Board (ARDB), Council for Scientific and Industrial Research (CSIR), Department of Atomic Energy (DAE), Department of Science and Technology (DST), Department of Biotechnology (DBT), Defence Research and Development Organization (DRDO), Ministry of Human Resources Development (MHRD), Ministry of Information Technology (MIT), Ministry of Non-Conventional Energy Source (MNES), Aeronautical Development Agency (ADA), and Indian Space Research Organization(ISRO). Research collaboration and funding are also obtained from industry. A representative list of companies with which IISc faculty have research interaction include Boeing, CadilaPharmceuticals, Daimler Chrysler, General Electric, General Motors, Hindustan Lever Limited, Indian Immunologies, Intel, IBM, Microsoft, Nokia, Nortel, Pratt & Whitney, Sir Dorabji Tata Centre, Texas Instruments etc. The research projects are handled through the Centre for Sponsored Schemes and Projects (CSSP) or the Society for Innovation and Development (SID). The Institute provides a handsome seed grant to all the new faculty members to assist in setting up laboratories etc.

The research work is primarily carried out by the students and all the students are paid scholarships by the Institute. In addition, the faculty can recruit project assistants/post-doctoral fellows through individual projects. The Institute also permits faculty members to undertake consultancy work, which is handled either by the Centre for Scientific and Industrial Consultancy (CSIC) or the Society for Innovation and Development (SID). The Institute encourages collaborative programs with various international universities and research organizations. Such collaborations allow the exchange of visitors for enhanced cooperation. The Institute has set up the Office of International Relations(OIR) to handle these matters. Our academic calendar consists of two semesters, August-December and January-April. The period, May-July, is considered a vacation period during which the faculty can undertake short/long term visits within India or abroad.

1.7 Academic Division, Department and Centres :

Division	Departments, Centres, and Units
Biological Sciences	• Department of Biochemistry • Central Animal Facility • Centre for Ecological Sciences • Centre for Infectious Disease Research • Centre for Neuroscience • Department of Microbiology and Cell Biology • Molecular Biophysics Unit • Department of Molecular Reproduction, Development and Genetics
Chemical Sciences	• Department of Inorganic and Physical Chemistry • Materials Research Centre • Nuclear Magnetic Resonance Research Centre • Department of Organic Chemistry • Solid State and Structural Chemistry Unit
Physical and Mathematical Sciences	• Astronomy and Astrophysics Programme • Centre for Cryogenic Technology • Centre for High Energy Physics • Department of Instrumentation and Applied Physics • Department of Mathematics • Department of Physics
Electrical, Electronic, and Computer Sciences	• Department of Computer Science and Automation • Department of Electrical Communication Engineering • Department of Electrical Engineering • Department of Electronic Systems Engineering
Mechanical Sciences	• Department of Aerospace Engineering • Centre for Atmospheric and Oceanic Sciences • Centre for Earth Sciences • Centre for Product Design and Manufacturing • Centre for Sustainable Technologies • Department of Chemical Engineering • Department of Civil Engineering • Divecha Centre for Climate Change • Department of Materials Engineering • Department of Mechanical Engineering
Interdisciplinary Research	• Centre for Infrastructure, Sustainable Transportation and Urban Planning • Centre for Biosystems Science and Engineering • Centre for Contemporary Studies • Centre for Nano Science and Engineering • Department of Computational and Data Sciences • Department of Management Studies • Interdisciplinary Centre for Energy Research • Interdisciplinary Centre for Water Research • Interdisciplinary Mathematical Science • Robert Bosch Centre for Cyber Physical Systems • Supercomputer Education and Research Centre

Chapter-2

ELECTRICAL COMMUNICATION ENGINEERING DEPARTMENT

The ECE Department is recognized by UGC as a Center for Advanced Studies and has a rich heritage and a strong reputation for R&D activities of internationally acclaimed standards, predominantly in the areas of Communications, Signal Processing, Microelectronics and RF/Photonics.

Presently, with 128 Ph. D students, over 84 ME students, 11 M.Sc. students and 31 faculty members, the Department does cutting edge research in areas such as sensor networks, wireless networks, QoS Architectures, Wireless Communications, MIMO Technologies, Coding Theory, Information Theory, Cognitive Radio, Speech/Audio, Video, Biomedical, SP for Communication, Compressive Sensing, Low Power Circuits, CMOS for RF, Nano Devices, MEMS and Sensors, RF MEMS, Antennas, Imaging, Optical communication, Biosensors.

Students, while working on algorithmic development, also have state-of-the-art experimental facilities to support the theoretical research, apart from research-based degree programs such as Ph. D, direct PhD and MSc (Engg), the Department offers ME programs in Telecommunication Engineering, Microelectronic Systems (jointly with CEDT) and Signal Processing (together with EE).

Thus, in addition to research activities, the Department adds to the quality of education by teaching several subjects covering fundamentals and advanced topics in the fields mentioned above.

The department faculty have several patents, publications and sponsored industry projects to their credit. Over the years, the graduating students have gone on to occupy key positions in education, science, industry and government administration in India and abroad. Over a span of sixty-seven years, the Department has been successful in attracting some of the best minds in the country, both in the student and teaching community.

In the future, the Department hopes to extend its research into newer areas like Energy Efficient Networks, Body Area Networks, High Precision Location and Navigation, Green Communication and Computing, Organic Electronics, Cyber-Physical Interfaces, Healthcare Informatics, Millimeter Wave Systems Imaging.

Chapter-3

TASK PERFORMED

3.1 Description About the Topics :

There were a number of topics which were discussed and explained during the period of work. The topics which were stressed upon were based on food segmentation, classification and Recognition, which dealt with convolutional neural networks, machine learning, API and GUI. The adopted following methodology:

I. Food Image Dataset

II. Food segmentation and classification

3.1.1 Food Image Dataset:

Fig.3.1 Food image dataset

The Food-101N dataset is introduced in a CVPR 2018 paper Clearnet: Transfer Learning for scalable Image Classifier Training from Microsoft AI & Research. The dataset is designed for learning to address label noise with minimum human supervision.Food-101N is an image dataset containing about 310,009 images of food recipes classified in 101 classes (categories). Food-101N and the Food 101 dataset share the same 101 classes, whereas Food-101N has much more images and is noisier, with 101'000 images. For each category, 250 manually reviewed test images are provided as well as 750 training images. 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. All images were rescaled to have a maximum side length of 512 pixels.

In this dataset, we define two types of labels for images:

Class labels: a class label describes the class of an image. Class labels are noisy, which means they could be incorrect. Each image in this dataset has a class label. The estimated noise rate is ~20%.

Verification labels: a verification label marks whether the class label is correct for an image.

Training a food image classifier relies on a comprehensive collection of food images. An assembled image dataset can be used subsequently to benchmark the recognition performance of other approaches. Several food image datasets have been created for this purpose. It has been a common practice to verify new classifier performance in contrast with the previous methods by training it with an abundant food image dataset such as Food101, PFID, UEC Food100, and UEC Food 256. Existing food image datasets have diverse characteristics, such as food categories, cuisine type, and the total images in the dataset/per food class.

By inspecting food image datasets, it is clear that most of the existing datasets are designated to a specific type of Food. Thus, there is a need for a generic and comprehensive food image dataset that can be used for benchmarking and general classification purposes. A dataset of food images is required to evaluate the performance of the different feature extraction and classification algorithms proposed. Also, the images can have multiple Food depicted. This makes the dataset more challenging since it requires the segmentation of each Food in the image. Random forest is used to mine discriminant super pixel-grouped parts in the food images. These parts are then classified with SVM achieving an average accuracy of 50.76% on the 101classes—a dataset used for testing a system that recognizes foods and estimates the classification of Nutrition. The system is designed to help everyone in controlling their vitamins, carbohydrates, amount of calories, sodium, potassium and different fats daily consumption. Various visual features and classification strategies are tested, and the best combination shows a classification accuracy of 96.81%.

Name	Year	#Images	#Classes	Type	Acquisition	Task	Annotation	Availability	Reference
Food50	2009	5000	50	Single	Wild	Food Recognition	Label	Proprietary	[30]
PFID	2009	1098[a]	61[a]	Single	Wild and Lab	Food Recognition	Label	Public	[63]
TADA	2009	50/256	–	Single and Multi	Lab	Food Recognition	–	Proprietary	[22]
Food85[b]	2010	8500	85	Single	Wild	Food Recognition	Label	Proprietary	[31]
Chen	2012	5000	50	Single	Wild	Food Recognition	Label	Public	[34]
UEC FOOD-100	2012	9060	100	Single and Multi	Wild	Food Recognition	BBox	Public	[32], [64]
Food-101	2014	101 000	101	Single	Wild	Food Recognition	Label	Public	[35]
UEC FOOD-256[c]	2014	31 397	256	Single and Multi	Wild	Food Recognition	BBox	Public	[33], [65]
UNICT-FD889	2014	3583	889	Single	Wild	Near Duplicate Food Retrieval	Label	Public	[66]
Diabetes	2014	4868	11	Single	Wild	Food Recognition	Label	Public	[5]
UNIMIB2015	2015	1000 × 2	15	Multi	Wild/Canteen	Food Recognition and Leftover Estimation	Poly	Public[d]	[10]
UNIMIB2016	2016[d]	1027	73	Multi	Wild/Canteen	Food Recognition	Poly	Public[d]	–

Table 3.1 Different datasets used for Food Recognition

The datasets have been categorized according to the type of images considered (i.e., images containing a single food or a set of foods), the acquisition procedure (e.g., in-the-wild for unconstrained acquisitions, or in the lab for constrained assets), the task for which it is used or created, the annotation type (label only, bounding boxes, or polygonal areas), and the availability (i.e., either public or proprietary).

3.1.1.1 Brief Overview ：

The increase in awareness of people towards their nutritional habits has drawn considerable attention to the field of automatic food analysis and Semantic Food Detection, which integrates with the same framework food localization, Recognition and segmentation. We demonstrate that our method improves the state-of-the-art food detection by a considerable margin on achieving about 96.81% through visualizations and audio processing which makes it so easy to identify and know the facts of what kind of Food would be suitable for the customer by checking on different nutrient levels with the help of Deep Learning. Deep learning has been proved to be an advanced technology for big data analysis with a large number of successful cases in image processing, speech recognition, object detection, and so on. Recently, it has also been introduced in food science and engineering as the data analysis tool to solve the problems and challenges in the food domain, including food recognition, calories estimation, quality detection of fruits, vegetables, meat and aquatic products, food supply chain, and food contamination.

Convolutional Neural Network (CNN) is one of the most prominent techniques in deep learning.CNN's is widely preferred in computer vision applications owing to its exceptional ability to learn operations on visual data and obtain high accuracies in challenging tasks with large-scale image data. CNN, in contrast to other traditional methods, outperforms by a large margin.

YOLOv2 is currently one of the best object detection approaches in state of the art. It allows predicting the bounding boxes and class probability of any object with a single convolutional neural network in real-time. As for the model, the authors propose a new FCN called Darket-19, composed of 19 convolutional layers and five max-pooling layers to tackle the recognition task.

3.1.2 Food segmentation and classification

I. Food recognition systems.
II. A traditional machine learning approach.
III. Deep learning approach.

3.1.2.1 Food recognition systems :

Research works have often focused on different aspects of food recognition problems. Many works address the challenges in Recognition of Food by developing recognition strategies that differ in terms of features and classification methodologies. The context of where the pictures are taken is also exploited along with the visual elements. With respect to the classification strategies, the most widely used are k-nearest neighbour (k-NN) classifiers, and support vector machines (SVMs), Bag of Features (BoF), Multiple Kernel Learning (MKL), and Random Forests (RF). An evaluation of different classification methodologies is reported where SVM, artificial neural networks, and random forest classification methods are analyzed. Recently, the convolutional neural network (CNN) is used in the context of food recognition.

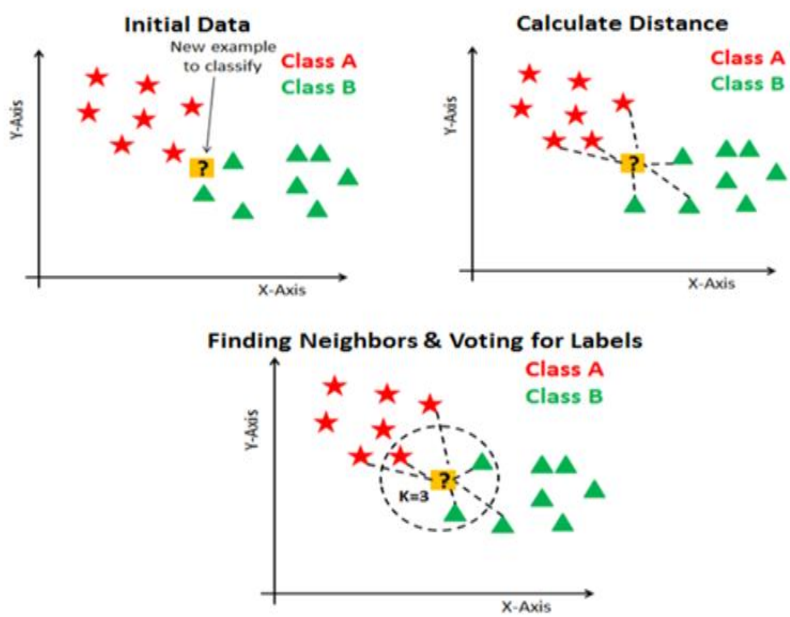

Fig. 3.2 k-nearest neighbour (k-NN)

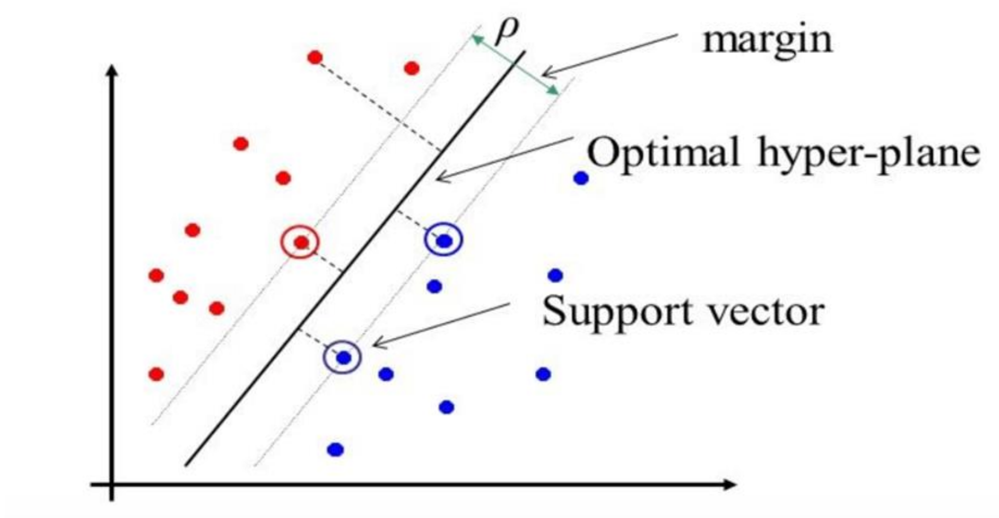

Fig.3.3 support vector machines (SVMs)

Often these systems exploit mobile application for food recognition, assessment, and logging. Food quantity estimation is fundamental in the context of dietary monitoring applications since it depends on the evaluation of the food intakes. Very few works individually consider the problem of food recognition and Nutrition estimation.

3.1.2.2 Traditional Machine Learning Approach :

- The process of feature extraction in this category is implemented manually by inspecting the visual features found in the food images, such as colour, shape, and texture. These features are then used to train a prediction model based on existing algorithms such as support vector machines (SVM), K-Nearest Neighbors (KNN), Bag of Features (BoF), Multiple Kernel Learning (MKL), and Random Forests (RF). The traditional classification methods basically execute three following tasks: segmentation, feature extraction, and classification. Segmentation is an essential step in identifying different regions of an image and then extracting the locations of the objects.
- Segmentation, when appropriately implemented, improves the classification accuracy, especially when multiple food items have to be identified within a single image or volume and nutrient contents have to be extracted. Food segmentation is yet a challenging task, as some food images may not present features such as shape contours and food edges. The segmentation could be more challenging when food items are minced, mixed in the food preparation process, and occluded food items laying on top of each other and hiding other parts of the Food.
- In another approach, several segmentation methods, including image colour, saturation, JSEG segmentation, and noise removal was combined to address the issue of multiple food identification. In this work, 73 food classes, found in a real food tray served in a canteen, were considered. The results showed that the classification accuracy was significantly improved. However, the tray images were manually segmented by drawing polygonal boundaries.

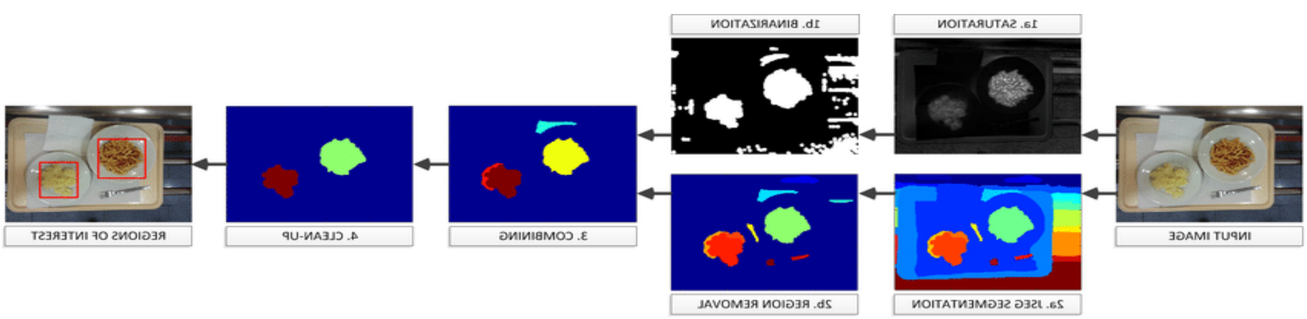

Fig.3.4 Segmentation Methord

- Another study attempted an ingredient-based segmentation based on the spatial relationships between the objects in the image by applying a Semantic Texton Forest (STF) algorithm. The overall classification accuracy was improved when compared with the traditional methods. However, this method relies on the composition of visually distinctive ingredients organized in predictable spatial settings.

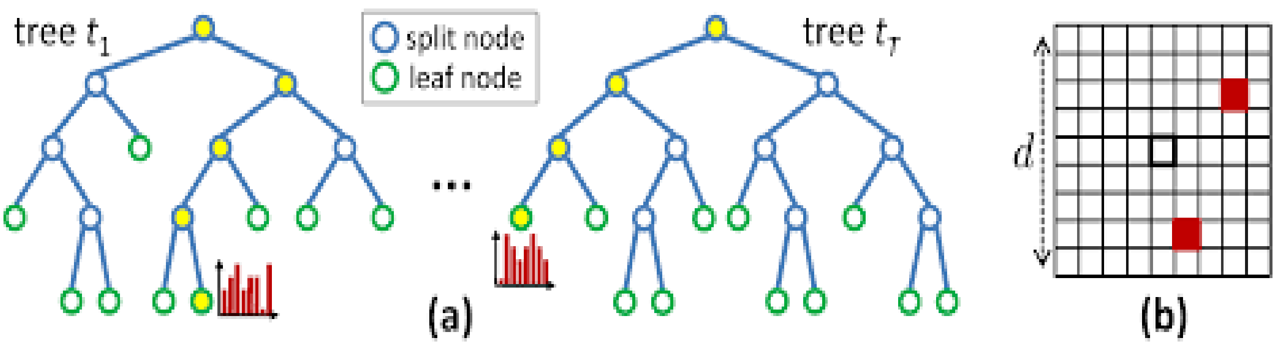

Fig.3.5 Semantic Texton Forest

- Zhu et al. implemented multiple segmentation hypotheses by assigning a class label to each pixel in an image. By using the classifier results as feedback to the segmentation, the number of segments in the image was estimated, considering the confidence scores assigned to each section. This approach outperformed the normalized cut method as in (1), where (assoc) computes the total edge associations from nodes in A or B to all nodes in the graph (V).

$$Ncut\,(A,B) = \frac{cut\,(A,B)}{assoc\,(A,V)} + \frac{cut\,(A,B)}{assoc\,(B,V)} \qquad \textbf{Eq 3.1}$$

- Another study proposed a JSEG segmentation approach linked with several object detectors, including circle detector, whole image, and Deformable Part Model (DPM) combination. It was shown that overall classification accuracy could improve in relation to using the DPM model alone.

- He et al. implemented a local variation segmentation algorithm, applied along with a segmentation refinement as feedback to increase the score of the assorted items. The overall classification was improved when compared with the normalized cuts approach.
- In traditional machine learning, a proper selection of these features significantly improves the classification accuracy and vice versa. The term handcrafted features come from the researcher's ability to identify the relevant features of the desired objects in the image. In the case of food classification, food items vary in shape, colour, and texture. The selection of associated features must relate to these three aspects. To date, the challenge remains when prepared Food is to be identified. Different methods of food preparation may result in different distinguishing features s. In order to find an optimal feature extraction process, informative visual data must be extracted from food images. These data can be found in general information descriptors, which are a set of visual descriptors that collect information about different basic features, including colour, texture, shape, and others. The descriptors, including Local Binary Patterns (LBP), Gabor filter, colour information, and Scale Invariant Feature Transform (SIFT) can be applied individually to extract image features. However, multiple descriptors can be implemented simultaneously to improve the overall classification accuracy. For example, a study implemented LBP and SIFT features individually on a food image dataset [20], the results showed that the accuracy of using SIFT features only is 53% while using the LBP features only resulted in 46% accuracy. Combining both features, along with additional Gabor filter and colour features, improved the efficiency to 68%. In another study, the same dataset was used and SIFT, LBP and colour features were extracted in addition to other features such as Histogram of Oriented Gradients (HOG) and MR8 filter. A combination of these handcrafted features obtained an accuracy of 77.4%. The study revealed that different parameters of the same extracted features might add up to the overall classification accuracy.
- There are several classification approaches with a variety of manually extracted features. Support Vector Machines (SVM) and K-Nearest Neighbor (KNN) have been the chosen traditional methods in several investigations in the field of food image recognition, mostly due to their substantial performance compared with other methods.
- A recent study applied colour, texture, and SIFT features to train a KNN classifier for food recognition. In contrast with an SVM classifier, KNN achieved a better classification accuracy of 70% while SVM classification achieved only 57%.

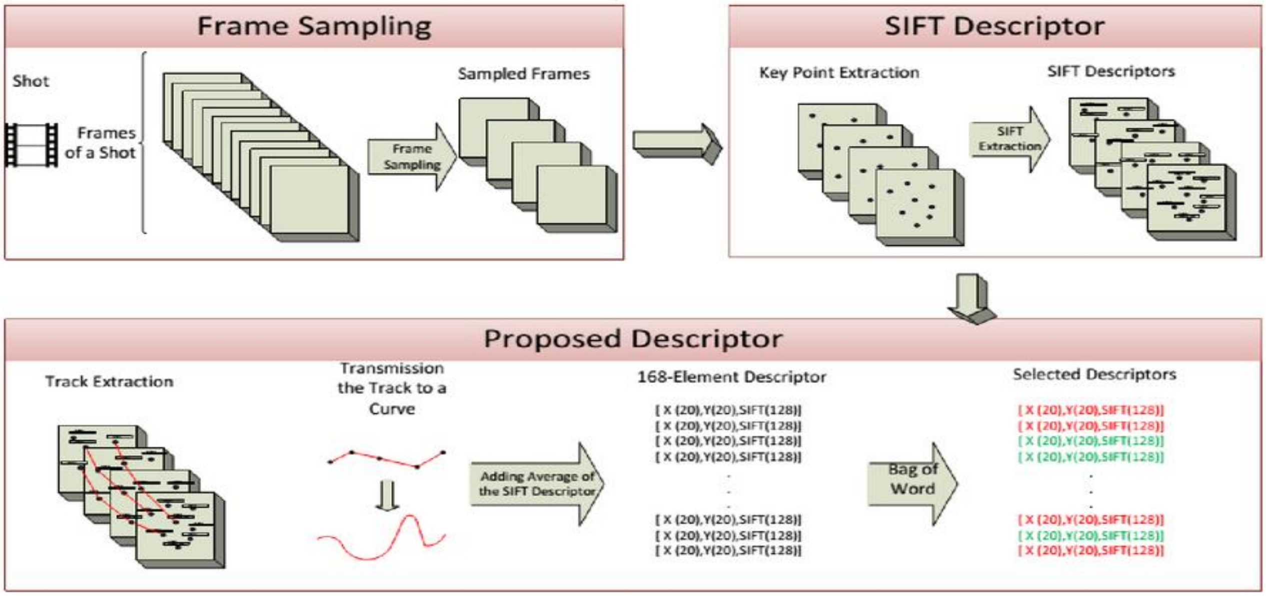

Fig.3.6 SIFT Extracting feautures

- Anthimopoulos et al. implemented a bag-of-features (BoF) model with SIFT extracted features. The authors trained an SVM linear image classifier to identify 11 classes of Food and obtained an accuracy of 78%.
- Chen et al. Implemented a multi-class SVM classifier to identify 50 classes of Chinese Food with 100 images in each category. Further, the authors added a multi-class Adaboost algorithm and improved the classification accuracy to 68.3%, followed by 62.7%, when SVM was implemented separately.
- Moreover, Beijbom et al. applied SIFT, LBP, colour, HOG and MR8 features and developed an SVM image classifier. An evaluation of their work was applied to two food image datasets and achieved a 77.4% accuracy in the dataset presented earlier [20], while they obtained only 51.2% precision using their menu-match dataset.
- The traditional food classification methods, summarized, highlighted the type of the implemented classifiers, the selected visual features, and the overall performance. Thus far, the process of features selection remains a challenging task regarding food image classification. Food items, such as fruits and vegetables, come in distinctive shapes, colours, and textures that are easily separable. Traditional classification approaches. Could be identified. However, the resemblance in the colour and texture of mixed and prepared Food renders the conventional classification methods ineffective. Alternatively, with

the development of deep learning algorithms, the need for manual feature selection as well as any user intervention has been eradicated or reduced. Hence, it may form a strong foundation for a prospective, fully automatic food identification system.

3.1.2.3 Deep Learning Approach :

Deep learning, a subset of machine learning, is a new approach to learn and train a more effective neural network. The built-in mechanism of deep learning algorithms adopts the features extraction automatically through a series of connected layers followed by a fully connected layer which is responsible for the final classification. It has recently become popular owing to its marginally exceptional performance with enhanced processing abilities, large datasets, and outstanding classification ability compared to other traditional methods.

Deep learning has the following deep neural networks which are capable of giving the best accuracy rates with the high amount of data been given to the network they can be classified as :

I. Convolutional Neural Network (CNN)
II. You Only Look Once (Yolo)

3.1.2.3.1 Convolutional Neural Network :

Convolutional Neural Network (CNN) is one of the most prominent techniques in deep learning. It was introduced by LeCun et al. for the classification of handwritten digits. CNN's is widely preferred in computer vision applications owing to its exceptional ability to learn operations on visual data and obtain high accuracies in challenging tasks with large-scale image data. CNN, in contrast to other traditional methods, outperforms by a large margin. A convolutional neural network is one of the most popular algorithms for deep learning, a type of machine learning in which a model learns to perform classification tasks directly from images, video, text, or sound.

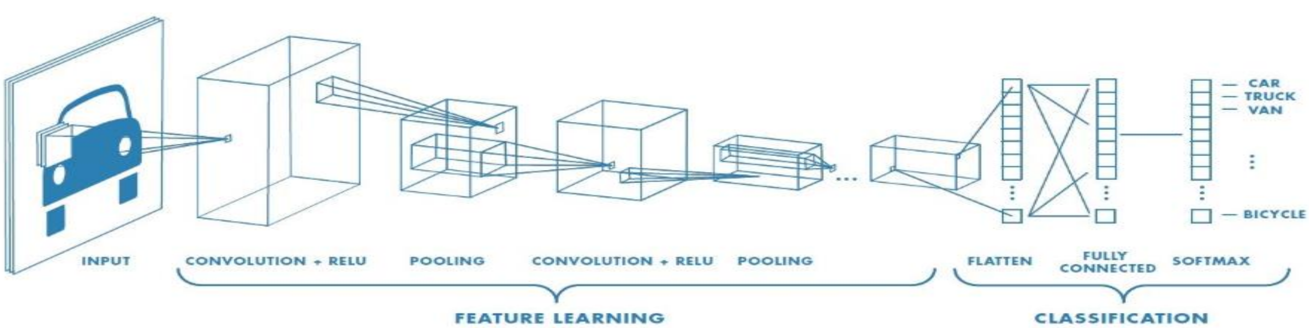

Fig.3.7Convolutional Neural Network (CNN)

CNN's are particularly useful for finding patterns in images to recognize objects, faces, and scenes. They learn directly from image data using models to classify images and eliminating the need for manual feature extraction.

Three of the most common layers are convolution, activation or ReLU, and pooling.

- Convolution puts the input images through a set of convolutional filters, each of which activates certain features from the images.
- Rectified linear unit (ReLU) allows for faster and more effective training by mapping negative values to zero and maintaining positive values. This is sometimes referred to as activation because only the activated features are carried forward into the next layer.
- Pooling simplifies the output by performing nonlinear downsampling, reducing the number of parameters that the network needs to learn.
- After learning features in many layers, the architecture of CNN shifts to classification.
- The next-to-last layer is a fully connected layer that outputs a vector of K dimensions, where K is the number of classes that the network will be able to predict. This vector contains the probabilities for each level of any image being classified.
- The final layer of the CNN architecture uses a classification layer such as SoftMax to provide the classification output
- The fully connected network (FNN) block, composed by fully connected neural units, is usually placed at the end as the classifier or used to generate numerical output for regression problems exploiting the learned feature map.

3.1.2.3.1.1 Applications of CNN’s :

I. Decoding Facial Recognition
II. Analyzing Documents
III. Historic and Environmental Collections
IV. Food Recognition with the nutrient analyzer
V. Understanding Climate
VI. Legal, Banking, Insurance, Document digitization - Optical Character Recognition
VII. Medical Image Computing - Healthcare Data Science / Predictive Analytics
VIII. Scene Labelling
IX. Natural Language Processing

3.1.2.3.1.2 CNNs Enable Advances in Object Detection and Object Recognition :

CNN's provide an optimal architecture for image recognition and pattern detection. Combined with advances in GPUs and parallel computing, CNNs are a vital technology underlying new developments in automated driving and facial Recognition.

For example, deep learning applications use CNNs to examine thousands of pathology reports to detect cancer cells visually. CNN's also enable self-driving cars to detect objects and learn to tell the difference between a street sign and a pedestrian.

3.1.2.3.2 You Only Look Once (YOLO) :

YOLOv2 is currently one of the best object detection approaches in state of the art. It allows predicting the bounding boxes and class probability of any object with a single convolutional neural network in real-time. As for the model, the authors propose a new FCN called Darket-19, composed of 19 convolutional layers and five max-pooling layers to tackle the recognition task. This network is modified for object detection by 1) removing the last convolutional layer and adding four convolutional layers for producing 13x13 feature maps, and 2) providing a region selection that enables to predict B bounding boxes at each cell on the output feature maps. The network predicts five coordinates for each bounding box, among them is the confidence score to, which represents both the confidence that the box contains an object and the accuracy at which the object is believed to be predicted; and c = 1, . . . , C conditional class probabilities, P r(Classc|Object). Predictions are obtained from the last convolutional layer having a size equal to 1 × one and F filters, where the number of filters is calculated as F = (B × (5 + C)).

From this, it is possible to determine the class-specific confidence score, CSc for each bounding box as follows:

$$CS_c = Pr(Class_c|Object) * \sigma(t_o)$$

Eq 3.2

where σ(.) stands for a logistic activation to constrain the predictions to fall in the range between 0 and 1.

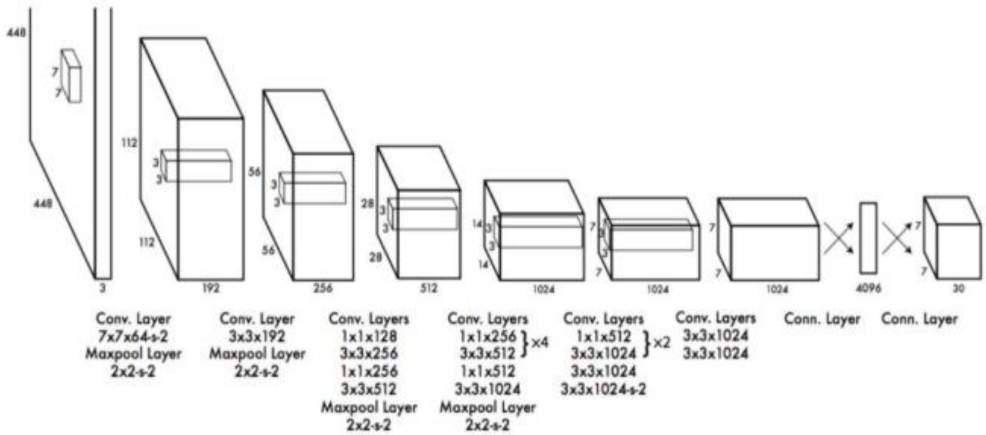

Fig 3.8 YOLO v2 Architecture

As for the experimental setup, YOLOv2 was pre-trained on the ILSVRC dataset. Following, we adapted it by changing the output of the model to 65 classes and applied a fine-tuning using UNIMIB2016 images. For training, the model used the framework Darknet [48]. The models trained during 4000 iterations with a batch size of 32, and a learning rate of 1e−3. Also, then applied a decay of 0.9 to the iterations 3000 and 3500. Once YOLOv2 training completed, the next step is to determine the confidence threshold used during localization and Recognition of the Food. A low confidence threshold implies a greater number of detections, which maximizes the likelihood that all the foods present in the image will be detected. At the same time, it also increases the chances of obtaining false detection. Results obtained by YOLOv2 and the proposed approach in the training set using different confidence thresholds.

In the training set using different confidence thresholds. The tested thresholds range from the minimum and maximum values mentioned above. As we can observe, when the threshold increases, the precision also increases considerably, but the rest of the indicators are affected. When comparing the results obtained between YOLOv2 and the proposed method, for the minimum threshold, it can be observed that a significant improvement in precision is obtained (≈40%) with only a slight decrease in the other indicators (0.1%-0.2%). Another interesting aspect to highlight is when comparing the results using the maximum threshold, in which case the results are practically identical for both methods. Means that, for a limit of 1, 2, there are almost no false detections that reduced with our procedure. For the remaining experiments, the minimum threshold chosen for two main reasons: 1) it obtains the best results for the Recall, MAA and T A indicators; and 2) it allows us to discard the false positives that appear when combining the results with the food segmentation procedure.

Chapter-4

Methodology Of Recognition

The Recognition of Food with different significant aspects of classification helps the segmentation process much faster due to the properly trained model and its labelled data with having the exact cluster of predicted classes understands the nutrients of the specific Food given as data. There are two stages of the process to track the Food and estimate the nutrients present in the Food:

4.1 Nanonets:

Nanonets is a framework which allows to train models of different weighted data and helps to train, test and validate the results on according to the given input of the data into the training model. The model built on the nanonets helps to create an API which used in the code which acts as a trained model with accuracy, which helps to create a machine learning models or deep learning models based on the data given to training the model. Nanonets have language bindings in Shell, Ruby, Golang, Java, C# and Python which indeed makes it feasible to code in any language. Their stack in the backend uses Golang, Python, Cassandra, Mysql, Tensorflow, Caffe, Torch, Keras, Theano, Elasticsearch and Docker, all of which hosted on AWS.

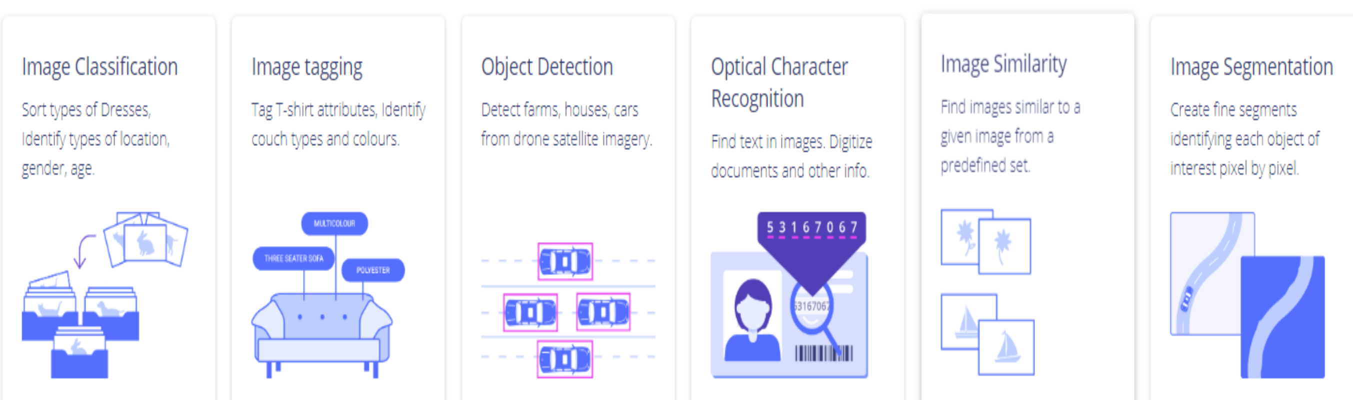

Fig.4.1 Nanonet API's

Nanonets helps to classify images based on classification, image tagging, object Detection, Optical Character Recognition, Image similarity and Image Segmentation. The categories for choosing image classification should be done before the uploading of data and once the names are specified, then upload the data. The data uploaded into the nanonets categories for the training of the data for further classification.

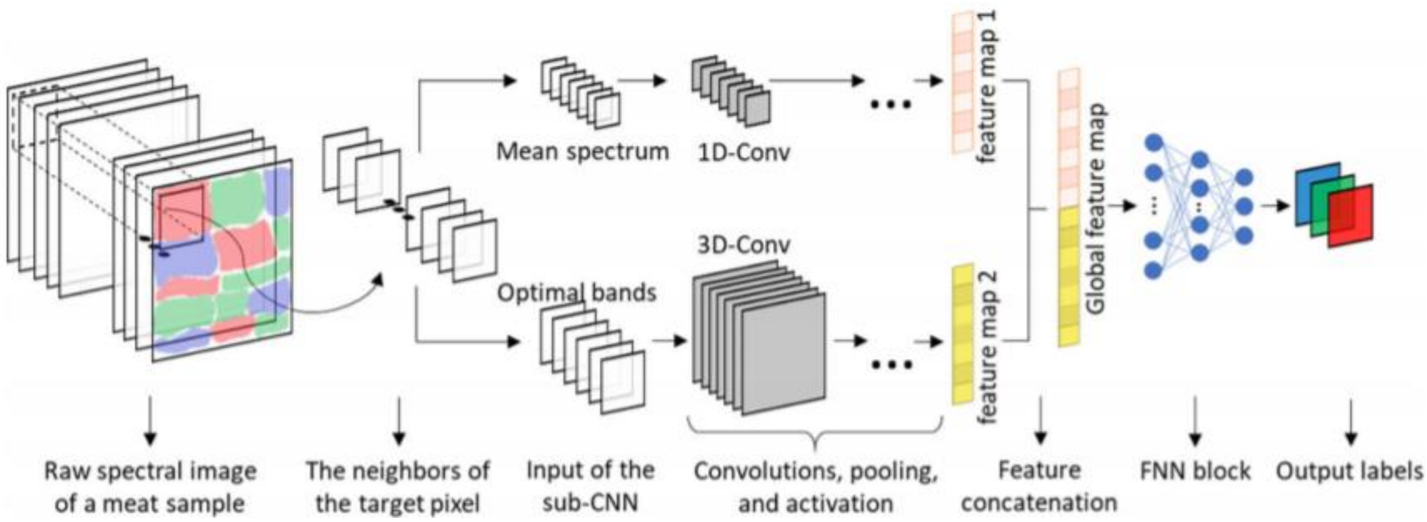

Fig.4.2 Layers in CNN

A deep CNN structure for spectral-spatial features extraction. The input of the network is the neighbours of the target pixel, a small region of a spectral image, which fed into two parallel subnetworks. The target region of a spectral image used to calculate the mean spectrum and treated by one-dimensional CNN in the first subnet while being reduced the number of channels to compressed input size and processed by three-dimensional CNN in another subnet. The abstracted features from the subnets were concatenated together and classified by an FNN block. After all the pixels traversed, the prediction result of a spectral image obtained.

The data with many images are to be convoluted to find the labels of the given input. The entire data have uploaded into the nanonets API server helps to train the data provided, and it can train any kind of given information of images more than 25 images for training.

4.2 Required Software Libraries :

4.2.1 Requests library:

It is a human-friendly HTTP library that can be used to send all kinds of HTTP requests to get information from various websites [7]. It is very easy to use and has features like passing parameters in the URL and passing custom headers as well.

Some of the features of request libraries are:

I. Keep-alive and connection pooling
II. Elegant key /value cookies
III. Automatic decompression
IV. Sessions with cookies persistence
V. Streaming Downloads
VI. Basic/Digest Authentication
VII. Chunked requests listen did you change points everywhere in this format?i didn't see can you do it now?

The main reason why we use python requests is we need not manually add query strings to URLs or form-encode post data [7]. We can use this library for GET and POST requests as well. With an image of the food as input, we are making use of the GET request for retrieving results of the classification from the API (nanonets server) which is stored on to a response object. The response data is converted into JSON type and this data is processed further for the prediction of food type.

4.2.2 PyQt5 and PyQt5.QtWidgets:

PyQt5 is known for its ease and simplicity offered to a developer for the development of an interactive Graphical User Interfaces (GUIs). PyQt5 is written in python but, it's a wrapper around the popular Qt GUI toolkit that major companies have been using it. It is a cross-platform GUI toolkit, unlike .NET which works on only windows. But, PyQt5 is independent of the operating systems and hence it can be used on Mac, Linux or Windows. This can be used for building any kind of application or widget development using python. PyQt5 provides a tool (QtDesigner) for the front-end development of the GUI [8]. Widgets are considered to be the building block of an application. PyQt5 provides a wide range and different types of widgets like checkboxes, list boxes, pushbuttons etc.

4.2.3 gTTS (Google Text To Speech):

gTTS is a python library and a command line interface tool (CMI). This tool is interfaced with Google Translator's TTS (text-to-speech) API [9]. By using this library, it makes easy to convert text to speech as it provides a variety of languages for audio manipulation. It has various sets of parameters to customize audio such as lang [9] (Language, default: English), slow (Speed of spoken content) etc. This spoken content can be saved in terms of '.mp3' file format. To play sound files within the python we can use the playsound () module or run the file separately [9].

Some of the features of the gTTS library include:

I. Automatically all supported languages are retrieved.
II. To support pronunciation corrections, it provides customizable text pre-processors.
III. It supports customization of speech-specific sentence tokenizer for reading lengthy (or unlimited) texts, all while keeping proper intonation, decimals, abbreviations, and more.

4.3 FCNI Model Training :

Using the Food101 data set 5 classes were selected namely- Donuts, French Fries, Samosa, chocolate cake & pizza. The categories (classes of food) for choosing image classification should be done before the uploading of data and once the names are specified, then upload the image dataset. The dataset uploaded into the nanonets categories for the training of the model for further classification. The Image classification model was trained on the Nanonets API server with 1000 images of each class. The dataset with many images is to be convoluted to find the labels of the given input. The entire dataset has been uploaded into the nanonets API server helps to train the data provided, and it can train any kind of given information of images more than 25 images for training. The training time again depends on the number of images provided for each class.

This model having accuracy = 96.81 was made into an API, which was used to integrate with python. The incorrect data which the model would not be able to recognize will show to know the dependency of the model accuracy at the rate of decrease of confidence score of testing. Daily Label wise average confidence score of the model was plotted and is shown in the Fig.4.3.

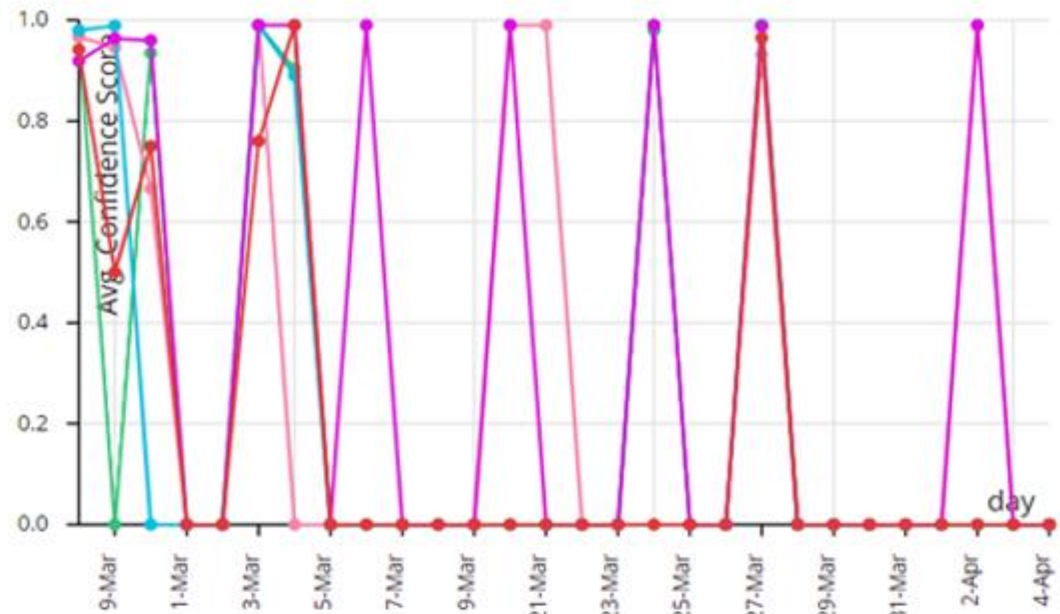

Fig 4.3 Confidence score of the FCNI Model

Once the training process is completed the confusion matrix of all the classes is shown in the Table 4.1

PREDICTIONS / TRUE	Donuts	French_Fries	Samosa	chocolate_cake	pizza
Donuts	101 / 106	0 / 106	2 / 106	2 / 106	1 / 106
French_Fries	1 / 107	105 / 107	0 / 107	0 / 107	1 / 107
Samosa	1 / 107	0 / 107	105 / 107	1 / 107	0 / 107
chocolate_cake	2 / 108	0 / 108	3 / 108	103 / 108	0 / 108
pizza	1 / 106	0 / 106	1 / 106	1 / 106	103 / 106

Table 4.3.1 Confusion matrix obtained after the training of the FCNI model

4.4 FCNI Architecture :

The architecture of the proposed FCNI is given in Fig. 4.4

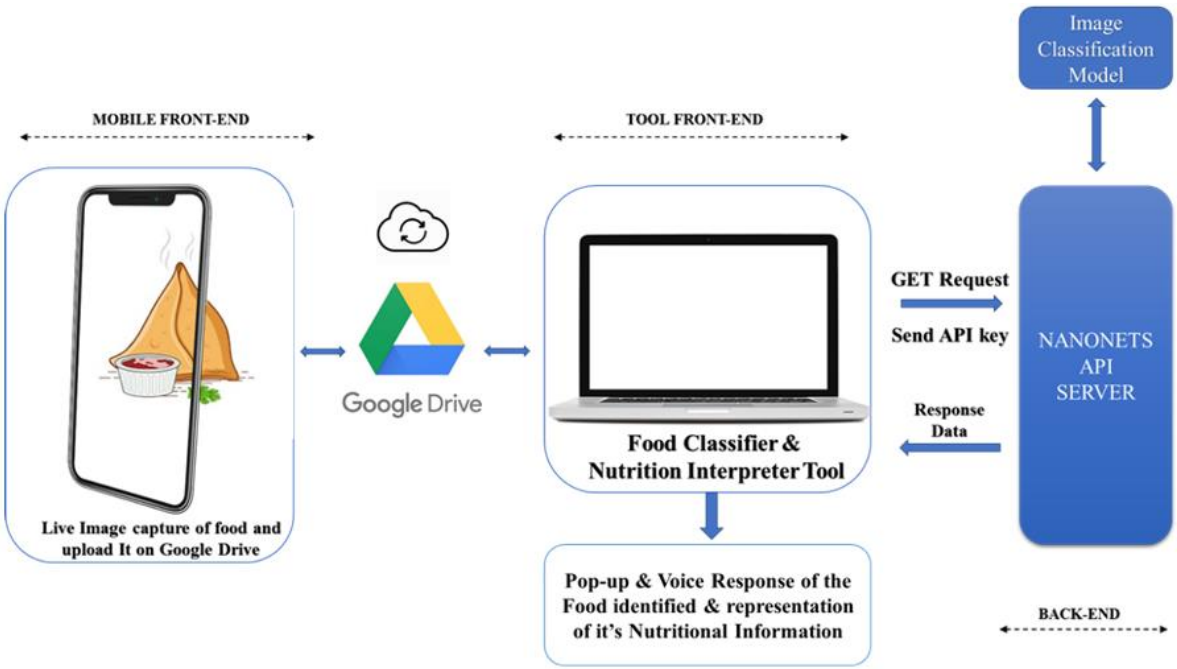

Fig 4.4 Proposed FCNI architecture

The various blocks of the proposed FCNI architecture are:

I. **Mobile Front-End:** This part of the block is in which the user takes a live picture of the food to be classified and uploads it on Google Drive. In this part of the block, Google drive is acting as a front-end on the mobile device, which provides the user to directly capture the picture.

II. **Tool Front-End:** An interactive graphical user interface (GUI) has been designed with the help of PyQt5 for providing the user the ease in use of this Tool.

III. **Back-End:** In the back-end, the nanonet API sever contains our saved model API and it receives requests from the tool and hence contains the various logics that send the appropriate response data back to the tool.

4.5 Sample Scenario Of FCNI Tool :

First the picture is captured on a mobile device and directly uploaded on to a specific directory on Google drive. The "Food Classifier & Nutrition Interpreter Tool" is also synched with Google drive over the cloud, hence it downloads this image and saves this image as the input to the tool for further classification process. When the user runs the tool by clicking on the 'CHECK' button, the input image downloaded is loaded on to the tool.

Once the image is loaded, the Tool uses the request library to send the loaded image along with the API key to the Nanonets API server. This server containing the image classification model verifies the API key received. In case the API key is not authentic the response code is returned as 404. If the verification was successful with a valid API key, the model processes the image, classifies the image and sends back response code (200) along with a response data.

This response data contains the label followed by the probability of each class. The data obtained is of the type response and cannot be processed directly. So, we convert this response data into a JSON file and then this is processed to abstract the class with the highest probability. Hence, we obtain the class of captured food. Based on the class the tool displays a pop-up message followed by an audio response of the Food identified. Following this the tool also displays the complete nutritional information of the predicted food with a graphical and tabular representation. This kind of representation makes it easier for the user to get a quick view of the various nutritional contents and some facts about the predicted type of food.

The 'CLEAR' button erases and removes the previous input image entirely from the directories. The user can again take another picture from their mobile device and use the tool for the next classification.

4.6 FCNI Tool Integration and Real-Time Testing :

4.6.1 Image Capture On Mobile Device :

Once the user selects the upload button on the Google drive application, they have the choice to either capture a live picture from their camera or upload an existing image from their gallery. Once the picture is clicked, this image is supposed to upload on to a specific directory. The screenshots are shown in fig. 4.5 and fig. 4.6

Fig.4.5 Real-time image capture of Food on a mobile device

Fig.4.6 Captured image uploaded on Google Drive

4.7 FCNI Tool GUI:

Basically, GUI is very much required to satisfy user convenience. The user need not have any prior knowledge about programming but it still would be easy to learn and use the application/tools. One of the main advantages of this Tool is that it provides users with immediate audio and visual feedback with required information about the captured food. The food classifier and nutritional identifier tool's home screen is shown in fig.4.7

Fig.4.7 Home Screen of FCNI Tool

4.8 FCNI Tool Testing And Results :

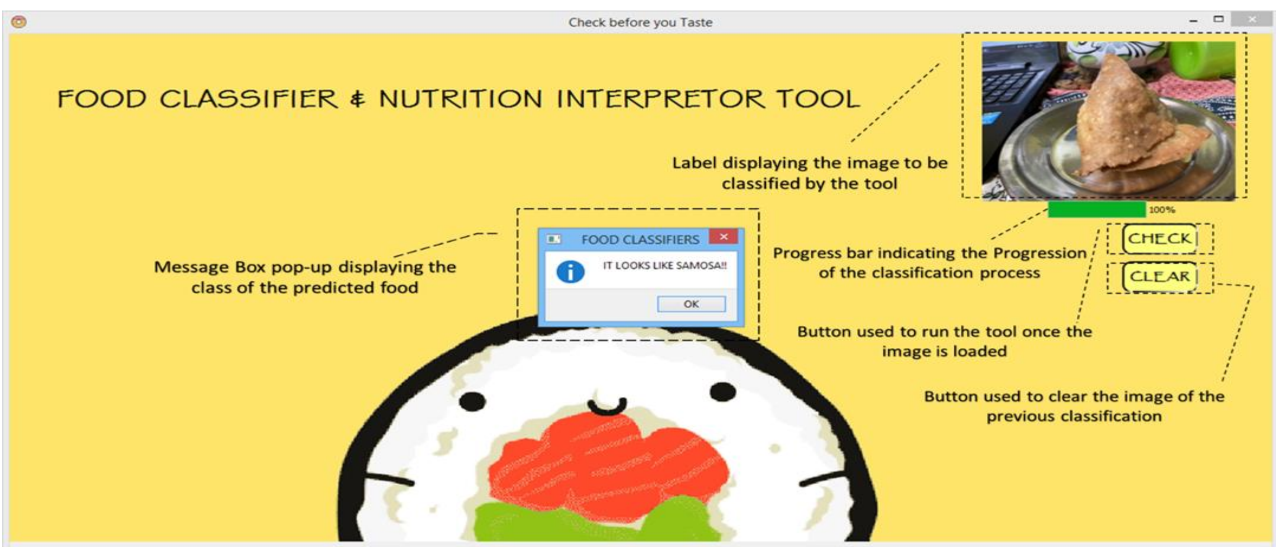

Fig.4.8 Working of the tool once the image has been captured on the mobile device by the user

4.9 Nutrient Facts :

The facts of various nutrients have been specified such as calories, fats, sodium, potassium, cholesterol, carbohydrates and proteins which gives a piece of specific information to certain kind of vitamins present in the particular Food which helps to maintain the diet and consume the Food depending on the need.

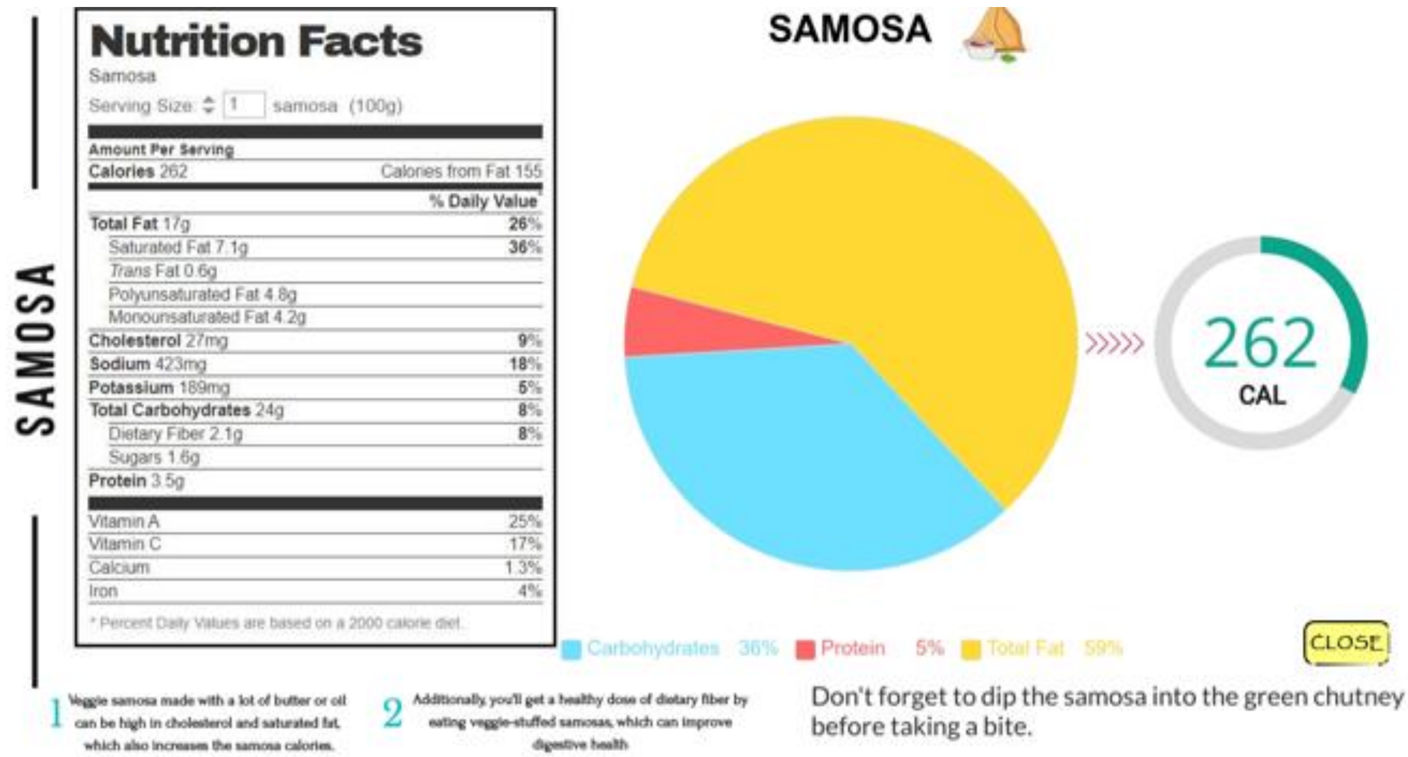

Fig.4.9 Tabular and graphical representation of Nutritional information of the predicted type of food

4.10 Features of FCNI :

I. Ability to predict the type of Food with good accuracy and gives an Audio & Graphical response.

II. It also helps in giving a good idea about the nutritional contents and respective calories of the Food predicted.

III. The Tool gives the user, information/warning regarding, what kind of people are restricted from eating the predicted type of Food.

IV. The captured image stored on to the google drive in which both mobile and laptop synced over the cloud

V. These are stored In a specific path in which the images to be tested is to be stored.

Chapter-5

CONCLUSION

The rapid progressions in technology over the recent years can be attributed to how Neural Networks has evolved and implemented in many fields and industries – automation, finance, healthcare etc. In this proposed model we have addressed the effectiveness of Convolutional Neural Networks (CNNs) in Deep Learning for food image classification and detection. The accuracy of CNNs for these kinds of classification is quite remarkable and can be applied to various real-time applications. The current hot topic among the different levels of people is diet control, to keep our body fit. Hence obesity is influenced by our diet, physical activity and lifestyle. By keeping this in mind as a primary goal and to overcome this, we successfully designed, developed and tested a compact and cost-effective Food Classifier and Nutrition Interpreter Tool (FCNI) using a deep learning approach. An accuracy of 96.81 is achieved. This FCNI tool alleviates the process of food intake estimation and dietary assessment, hence beneficial for the user.

REFERENCES

[1] https://imagga.com/blog/the-top-5-uses-of-image-recognition/

[2] https://towardsdatascience.com/the-4-convolutional-neural-network-models-that-can-classify-your-fashion-images-9fe7f3e5399d

[3] Mohammed Ahmed Subhi, Sawal Hamid Ali, Mohammed Abulameer Mohammed. "Vision-Based Approaches for Automatic Food Recognition and Dietary Assessment: A Survey", IEEE Access, 2019

[4] https://in.mathworks.com/solutions/deep-learning/convolutional-neural-network.html

[5] https://nanonets.com/

[6] https://www.tensorflow.org/datasets/catalog/food101

[7] https://pypi.org/project/requests/

[8] https://pypi.org/project/pyqt5-tools/

[9] https://gtts.readthedocs.io/en/latest/

COURSE OUTCOMES

CO1: Students will be able to explain emerging technologies and recent developments in the field.

CO2: Students will be able to appreciate latest inventions and also analyse different technologies and recognize innovative methods for further developments for the need of mankind.

CO3: Students will be able to exhibit effective communication and presentation skills.

CO-PO mapping table

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1														
CO2														
CO3														

Signature of the guide