AI Based Milk Contamination Detection With Sensor

[1] Dr.S.Kannappan, [2] Bheema Meghana, [3] Poluboina Muragaiah, [4] Kadiri Naveen, [5] Nalaballi Narendra, [6] Manga Omkar Praveen Kumar

[1]Associative Professor, [2]UG Student, [3]UG Student, [4]UG Student, [5]UG Student, [6]UG Student
Department ECE,
Siddharth Institute Of Engineering & Technology, Puttur, India
[1]sadasivamkannapan@gmail.com, [2]immeghana2004@gmail.com, [3]muragaiahpoluboina@gmail.com, [4]nanenaveen2702@gmail.com, [5]nellaballinarendra1012@gmail.com, [6]m.o.praveenkumar2003@gmail.com

Abstract— **The detection of milk contamination is a critical concern for ensuring food safety and public health. This work presents an AI-based milk contamination detection system integrated with advanced sensor technology. The collected data is processed and analyzed using sophisticated AI models, including Convolutional Neural Networks (CNN) and Gradient Boosting Machines (GBM), to accurately identify contamination. The proposed system is designed to operate autonomously, providing continuous monitoring and immediate alerts upon detecting any anomalies or potential contaminants. This proactive approach ensures timely intervention, reducing the risk of distributing contaminated milk to consumers. Experimental results demonstrate the effectiveness of the system in various scenarios, highlighting its robustness and scalability.**

Index Terms— **Convolutional Neural Networks (CNN), Gradient Boosting Machines (GBM), Solid Not Fat (SNF), Hyper Spectral Imaging (HSI).**

I. INTRODUCTION

Milk is one of the most widely consumed beverages globally, providing essential nutrients such as calcium, protein, and vitamins. However, like any other food product, milk is susceptible to contamination. Milk contamination refers to the presence of harmful substances in milk that can pose a risk to human health. These substances may include bacteria, chemicals, toxins, or other foreign materials that compromise the quality and safety of milk.

With advancements in artificial intelligence (AI) and sensor technology, automated and intelligent milk quality monitoring systems have gained attention for their ability to provide rapid and accurate contamination detection. AI-based models leverage machine learning and deep learning algorithms to analyze complex datasets, identifying contamination patterns that might be difficult to detect through conventional means. By integrating real-time sensor data with AI-driven analysis, modern detection systems can offer enhanced precision, reduced response time, and continuous monitoring capabilities.

The growing demand for efficient, scalable, and cost-effective solutions has driven research toward AI-powered contamination detection. These systems utilize advanced computational techniques, such as Convolutional Neural Networks (CNNs) for feature extraction and Gradient Boosting Machines (GBM) for classification, to improve detection accuracy. By leveraging such technologies, automated contamination detection systems can minimize human intervention, reduce the risk of human error, and enhance the overall safety of milk production and distribution processes.

Ensuring the quality and safety of milk is crucial for protecting consumers, maintaining industry standards, and preventing economic losses caused by contamination-related recalls and reputational damage. The integration of AI and sensor technology in food safety applications is a step toward building intelligent, data-driven solutions that enhance public health protections while streamlining quality control processes.

II. LITERATURE SURVEY

A.Momin, et al introduces an innovative milk detection system that leverages Fourier Transform Infrared (FTIR) spectroscopy. This cutting-edge system utilizes FTIR analysis to evaluate the chemical composition of milk samples, enabling precise determination of their quality. By incorporating advanced machine learning algorithms for data interpretation, the system categorizes milk samples into distinct quality grades, effectively identifying adulteration, contamination, and compositional alterations.

C. B. Mupparaju et al present a novel crystal structure algorithm with Extreme Learning Machine (CS-ELM) for early discovery of contaminated products. The characteristic data is effectively received by utilizing proposed CS–ELM model. For the given input data, an optimal solution are predicted to develop the food process optimization model along with ELM. The comparative study executes the direct comparison among each characteristic information process. For emerging food safety issues identification, huge reliable food information are provided to obtain the prediction value as well as characteristic information.

R. Bhattacharya et al developed a microwave based non-invasive method and demonstrated analytically. CST Studio Suite 2022, high frequency module was used for the analysis. An S-parameter based method was adopted for the milk transportation tube (assumed as circular waveguide) and an Eigen mode-based method was adopted for the pasteurizer (assumed as enclosed cavity. Simulation showed that for the case of tube, reflection coefficient changed substantially due to the electric field perturbation caused by the biofilm. The frequency for the minimum S11 shifted from 0.19 to 0.229 GHz, and the S11 magnitude decreased from -30 dB to -10 dB, which equates to a 100-fold change.

S. Bonaldo, L et al presents a novel cost- effective biosensor for monitoring the presence of L. lactis phages in milk samples in less than 4 hours. The detection relies on the parametric variations in the electrochemical impedance spectroscopy response of the proposed biosensor. Differences of more than one order of magnitude is measured in the charge transfer resistance when the solution under test is contaminated by phages, due to the phage lytic activity on the L. lactis bacteria.

C. Allende-Prieto et al., discusses the use of near-infrared (NIR) spectroscopy combined with multivariate classification methods for detecting bacterial contamination in milk in the dairy industry. In the first experiment, the study found that NIR was accurate and reliable in detecting the presence of biofilms in milk. Our results showed that the technology was effective in distinguishing between contaminated and uncontaminated samples with an area under the receiver operating characteristic (ROC) curve (AUC) greater than 99%. It was also effective in classifying the samples belonging to different strains. In a second experiment, we used the same methodology to assess their effectiveness in detecting bacterial contamination proportions in milk.

R. S. Ram et al detect the contamination of food and water. Internet of Things (IoT) is growing rapidly and becoming a vast source of information. IoT has a great feature in detecting food and water quality. By connecting the Arduino and sensors the quality of food and water can be detected and by utilizing Wi-Fi module the information can be transmitted. Water samples are collected from various sources like tap, borewell and wells in the local environment. Various Junk food items, fruits and milk quality is tested under food category with the aid of sensors. If contamination has been detected then the information is passed on to the officials in the Department of Food Safety and Drug administration for further action. In this way, effectively checking food and water quality avoids food and water-borne illness and paves a way to lead a healthy life.

M. Zhang, et al the radio frequency-liquid recurrent neural network (RF-LqRNN) system is proposed, which can classify solution concentrations using an inexpensive commercial off-the-shelf radio frequency (RF) identification device. The proposed system uses the feature that the phase and received signal strength indicator (RSSI) of RF signals across solutions of different concentrations are different. Measurements of the two features were obtained at different concentrations and input into a gated recurrent unit (GRU)-recurrent neural network (RNN) to train a solution concentration classification model.

Y. Feng et al., proposes an IoT-based animal social behavior sensing framework to model mastitis propagation and infer mastitis infection risks among dairy cows. To monitor cow social behaviors, we deploy portable GPS devices on cows to track their movement trajectories and contacts with each other. Based on those collected location data, we build directed and weighted cattle social behavior graphs by treating cows as vertices and their contacts as edges, assigning contact frequencies between cows as edge weights, and determining edge directions according to contact spatial-temporal information.

J. Pereira-da-Silva, et al rooted on this premise and sensors based on LbL polyectrolyte thin-films were developed, such as poly(allylamine hydrochloride) (PAH) and poly(vinylsulfonic acid, sodium salt) (PVS), namely $(PAH/PVS)_{10}$ and $(PAH/PVS)/PAH_{10}$. The goal of this work was to analyze the amount of TCS adsorbed onto reused $(PAH/PVS)_{10}$ and $(PAH/PVS)/PAH_{10}$ LbL thin-films when immersed in TCS aqueous solutions with decreasing pH. It was demonstrated that sensors with an outer layer of PAH led to a significantly better TCS molecules adsorption (removal). Additionally, sensors composed of $(PAH/PVS)_{10}$ presented higher sensibility in discriminating TCS solutions with concentrations between 10^{-5} M and 10^{-8} M, using impedance spectroscopy.

M. Moreira et al., propose a prototype of a digital photometer to quantify water added to milk. This is a microcontrolled, portable device, which uses three LEDs with emission in the near-infrared (IR) region and was developed without the use of lenses, filters, or moving parts. This equipment measures the transmittance of IR radiation through milk samples to assess the addition of water. In this paper, we present the results of experiments that were conducted with diluted milk samples containing 0%-25% of added water, using the proposed equipment and the standard method of cryoscopy.

J. Ren et al a piezoelectric transducer (PT) was developed for determination of bacteria counts in fresh milk in real time. The system consists of detection cell, oscillator, frequency counter and computer. The determination was based on the sensitive response of the transducer to the change of culture media during bacteria growth. The frequency shift versus time response curves were recorded by self-developed software. Bacteria detection time (BDT) corresponding to the significant change of frequency shift value was used as a parameter to quantitatively determine bacteria.

Gupta, R., et al 2017 fabricated a novel hand held micro-electrode sensor based on label free impedance spectroscopy for the detection of adulteration in milk. The results showed remarkable impedance change for different concentrations of starch and detergent in the milk sample in the frequency range of 0.05-5.0 kHz. The impedance of reference milk sample at resonant frequency 4.75 kHz was found to be 8.01E+02Ω and that of model milk sample adulterated with 10% (w/v) detergent was found to be 56% of the original value while the impedance of milk sample adulterated with 10% (w/v) starch was found to be ~ 160% of the original value of 8.01E+02Ω. Also, it was found that the change in the concentration of starch and detergent in milk samples changes the overall impedance of the sensor remarkably. The measured impedance values were inversely proportional to the concentration of detergent and directly proportional to the concentration of starch in the model milk samples. The observed trend is explained and highlights the selectivity and sensitivity of the device. The device requires only 0.2 mL of the sample and gives response within 10 seconds. The real and simulated data were fitted in an AC equivalent circuit model of the device operation.

Karuppuswami, S., et al 2017 have investigates a step by step design procedure of a hybrid passive wireless sensor. The hybrid sensor measures both the electrical (dielectric constant) and the mechanical (viscosity) properties of liquid, providing a two-factor quality control. The hybrid sensor is based on an inductor-capacitor resonant tank coupled with a magnetoelastic strip. The mechanical and the electrical resonances change as a function of viscosity and dielectric constant, respectively.

Current methods for detecting milk contamination primarily rely on manual testing and limited laboratory analysis, which are often time- consuming, labor-intensive, and not suited for continuous monitoring. These conventional approaches can lead to delays in contamination detection, increasing the risk of distributing unsafe milk products to consumers. Furthermore, the existing systems may lack the capability for real- time monitoring and immediate response, resulting in potential public health hazards and reduced consumer trust. There is a pressing need for a more efficient and autonomous solution that can continuously and accurately monitor milk quality, detect contamination in real-time, and provide prompt alerts to ensure timely intervention and prevent the distribution of contaminated milk. The integration of advanced sensor technology and AI models could address these shortcomings by offering a more reliable, scalable, and proactive approach to milk contamination detection.

III. PROPOSED SYSTEM

The proposed system is a cutting-edge solution for milk contamination detection that integrates advanced sensor technology with sophisticated AI models. It features high-sensitivity sensors for continuous real-time monitoring of milk quality, capturing critical data on various contamination parameters. This data is processed using Convolutional Neural Networks (CNN) and Gradient Boosting Machines (GBM) to accurately identify and classify contamination levels. The system operates autonomously, providing constant surveillance without the need for manual intervention, and generates immediate alerts upon detecting any anomalies or contaminants. This proactive approach ensures timely intervention, effectively reducing the risk of contaminated milk reaching consumers. The system has been validated through extensive testing, demonstrating its robustness, accuracy, and scalability across different operational conditions.

Figure 1:- Proposed system modules

Sensor Technology Module:

The Sensor Technology Module is responsible for the real-time collection of data from milk samples. It employs high-sensitivity sensors designed to detect various contaminants and quality parameters, such as microbial presence, chemical residues, and physical abnormalities. These sensors are strategically placed to ensure comprehensive coverage and accurate data acquisition. The collected data is transmitted seamlessly to the central processing unit for further analysis. This module is crucial for providing the raw input needed for contamination detection and ensuring continuous, real-time monitoring of milk quality.

Data Processing and AI Analysis Module:

The Data Processing and AI Analysis Module is the core of the system, where the raw data collected from sensors is analyzed using advanced artificial intelligence techniques. Convolutional Neural Networks (CNN) are employed to process and interpret complex patterns in the data, allowing for the identification of subtle signs of contamination. Complementing this, Gradient Boosting Machines (GBM) are used to enhance the system's predictive accuracy by classifying and regressing contamination levels based on historical and real- time data. This module transforms raw sensor data into actionable insights, providing accurate detection and classification of contaminants.

Autonomous Operation And Monitoring Module:

The Autonomous Operation and Monitoring Module ensures that the system functions independently, providing continuous surveillance without manual intervention. It integrates automated data collection, processing, and alert generation, enabling the system to operate around the clock. This module manages the operational aspects of the system, including sensor calibration, data flow management, and system health monitoring. Its primary goal is to maintain seamless and reliable operation, ensuring that the system remains functional and effective in detecting contamination at all times.

Alert and Response Module:

The Alert and Response Module is designed to handle contamination detection alerts and initiate appropriate responses. When the system detects contamination or anomalies, this module generates real-time notifications to relevant stakeholders or automated systems. The alerts can be customized based on the severity of contamination, and response protocols can be triggered automatically. This module is essential for ensuring timely intervention, allowing for swift corrective actions to prevent the distribution of contaminated milk and safeguard public health.

Scalability and Validation Module:

The Scalability and Validation Module focuses on ensuring that the system can adapt to different volumes and operational conditions while maintaining accuracy and robustness. This module includes components for scaling the system to handle varying sample sizes and testing environments. It also encompasses extensive validation processes, where the system's performance is evaluated through rigorous testing across diverse scenarios. The results from this module confirm the system's reliability, effectiveness, and ability to perform well under different conditions, ensuring that it meets the required standards for real-world applications.

1D Convolutional Neural Network (1D-CNN) Architecture Overview:

1D Convolutional Neural Networks (1D-CNNs) are a variant of CNNs designed to process sequential data. While traditional CNNs are typically used for 2D data like images, 1D-CNNs are well-suited for time-series data or any other form of one-dimensional signals. This makes them an excellent choice for applications like milk contamination detection, where sensor data is often time-series in nature.

- **Input Layer:** The input to a 1D-CNN is typically a one-dimensional array representing the sequential data captured by sensors. In the context of milk contamination detection, this could include various sensor readings such as pH levels, temperature, and chemical concentrations over time.
- **Convolutional Layers**: These layers apply convolutional operations to the input data using a set of filters (kernels). Each filter slides over the input data, computing the dot product between the filter and a subsection of the input. This operation captures local patterns in the data, such as sudden changes in sensor readings that might indicate contamination. The output of this layer is known as a feature map, which highlights the presence of specific features within the data.
- **Activation Functions:** After convolution, activation functions like ReLU (Rectified Linear Unit) are applied to introduce non- linearity into the model, allowing it to learn more complex patterns.
- **Pooling Layers**: Pooling layers, typically using max-pooling, are then applied to reduce the dimensionality of the feature maps, which helps in reducing computational complexity and in capturing dominant features in a more abstract form.
- **Fully Connected (Dense) Layers**: After a series of convolutional and pooling layers, the resulting feature maps are flattened into a one- dimensional vector, which is then passed through one or more fully connected layers. These layers are responsible for combining the features learned by the previous layers to make final predictions.
- **Output Layer:** The final layer in a 1D-CNN is usually a softmax or sigmoid layer, depending on whether the task is multi-class classification or binary classification. In this case, the output might be the classification of milk into various contamination levels.

Working Principle of CNN :-

The 1D-CNN works by automatically learning the relevant features from the input data during the training process. As it processes each sequential input, the convolutional layers identify patterns that are indicative of contamination, such as anomalies in sensor readings. These learned patterns are then used by the fully connected layers to classify the input into categories such as "contaminated" or "safe." The model is trained using labeled data, where the ground truth labels (e.g., contamination level) are provided, allowing the network to learn to associate specific patterns with contamination.

Gradient Boosting Machines (GBM) Architecture Overview:

Gradient Boosting Machines (GBM) are a powerful machine learning technique used for classification and regression tasks. They work by building an ensemble of decision trees, where each tree corrects the errors made by the previous ones. This method is particularly effective in handling complex datasets and is known for its high predictive accuracy.

- **Base Learner (Decision Trees):** GBM typically uses decision trees as the base learners. These trees are usually shallow, meaning they have a limited number of splits, which helps in reducing the risk of overfitting. Each tree in the ensemble focuses on correcting the mistakes made by the previous trees.
- **Sequential Learning:** The trees in GBM are built sequentially, with each tree being trained to correct the residual errors of the preceding trees. This sequential approach allows the model to improve its performance with each additional tree, resulting in a strong ensemble that can capture complex patterns in the data.
- **Loss Function**: GBM minimizes a specified loss function (e.g., mean squared error for regression or log loss for classification) using gradient descent. The loss function measures the difference between the predicted values and the actual values, and GBM iteratively reduces this difference by adjusting the model's parameters.
- **Learning Rate:** The learning rate in GBM controls the contribution of each tree to the final model. A lower learning rate requires more trees to be built, but it can lead to better generalization by preventing the model from fitting the training data too closely.

- **Regularization:** GBM includes regularization techniques to prevent overfitting. This can involve limiting the number of trees, controlling the depth of the trees, or using shrinkage (scaling the contribution of each tree) to ensure that the model does not become too complex.

Working principle of GBM :-

GBM works by iteratively adding trees to the model, each one focused on correcting the errors of the previous ensemble. Here's how it operates:

Initialization: The model starts with an initial prediction, which could be as simple as the mean of the target values in the case of regression.

Building Trees: The first tree is built to minimize the loss function, focusing on the most significant errors in the initial prediction. After this tree is added, the model checks the residuals, which are the differences between the predicted values and the actual target values.

Gradient Descent: The model calculates the gradient of the loss function with respect to the predicted values. This gradient represents the direction in which the model should adjust its predictions to reduce the error. The next tree is trained on these gradients, effectively learning how to correct the mistakes made by the previous trees.

Updating Predictions: The predictions are updated by adding the new tree's predictions to the existing ones, scaled by the learning rate. This process is repeated for a specified number of iterations or until the model achieves the desired level of accuracy.

Final Prediction: The final model is a weighted sum of all the trees, each contributing to the overall prediction based on its ability to correct previous errors.

IV. HARDWARE SPECIFICATION

ESP 32 Module :-ESP32 is a series of low- cost,low-power system on a chip microcontrollers with integrated Wi-Fi and dual-mode Bluetooth. The ESP32 series employs either a Tensilica Xtensa LX6 microprocessor in both dual-core and single-core variations, Xtensa LX7 dual-core microprocessor or a single- core RISC-V microprocessor and includes built-in antenna switches, RF balun, power amplifier, low- noise receive amplifier, filters, and power- management modules. ESP32 is created and developed by Espressif Systems, a Chinese company based in Shanghai, and is manufactured by TSMC using their 40 nm process. It is a successor to the ESP8266 microcontroller.

pH meter :-pH meter is an instrument which is used to measure the potential of hydrogen ions in a liquid, and differentiating the acidity and alkalinity of a liquid with numbered units from 0-14. pH meter measures the electrical potential difference between the reference electrode and internal electrode. So it is also called as “potentiometric pH meter”. The potential difference between electrodes define the acidity of the solution in which the pH sensor is placed.

TCS3200 Color Sensor :-The TCS3200 is a color sensor that detects and measures various colors depending on their wavelength. This sensor module includes the TAOS TCS3200 RGB Sensor IC and is equipped with four white LEDs to light up the object, ensuring precise color detection. This sensor chip has an 8×8 Photodiode array to decide the color of an object exactly and generate a frequency signal equivalent to the detected color with a Current-to- frequency converter.

The TCS3200 module uses a CMOS IC with configurable silicon photodiodes & a current-to- frequency converter to generate a square wave output. So the output frequency is directly proportional to the intensity of light. Thus, the TCS3200 module is utilized in color recognition projects like color sorting, color matching, ambient light sensing, and test strip reading.

The detection range of the TCS3200 module is 100 mm and it operates with a single voltage supply that ranges from 2.7V to 5.5V. So it is well-matched with almost all common microcontrollers like; AVRs, PICs, Arduino, and ARM

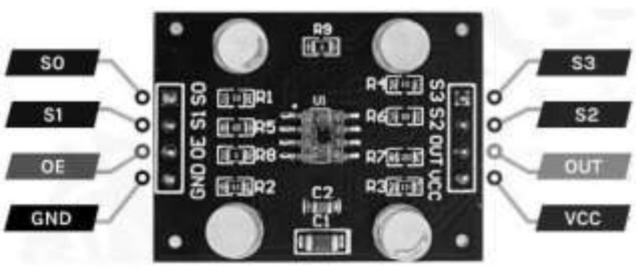


Figure 2:- TCS3200 Colour Sensor Pin

- **Pin (VCC):** This is the power supply pin of the module that is connected to 3.3V or 5V of the supply.
- **Pins (S0 & S1):** These pins are used to choose the o/p frequency scaling percentage of the module. Once these pins are

configured then it can be situated to 2%, 20% (or) 100% scaling.

- **Pins (S2 & S3):** These pins are used to choose the color array of the module. By choosing the correct color array continuously, thus this sensor recognizes a color.
- **Pin (OE):** This Output Enable (or) Disable pin of the sensor is pulled down on the module to disable the sensor by providing a higher pulse to OE pin.
- **Pin (OUT):** This is the o/p pin of the sensor module. Whenever a particular color is noticed through the sensor, then the output pulse frequency will be changed by noticing this change within pulse width we can decide the color.
- **Pin (GND):** This is the ground pin of the module and it must be connected to the GND pin of the Arduino board.

Conductivity Sensor:-

The conductivity sensor is one kind of sensor used to gauge the volumetric content of water within the surface. These sensors measure the volumetric water content not directly with the help of some other rules of soil like dielectric constant, electrical resistance, otherwise interaction with neutrons, and replacement of the moisture content.

V. RESULT

Dataset and Per processing :-The dataset used for this study is manually collected from observations to facilitate the development of machine learning models aimed at predicting milk quality. It comprises seven independent variables that are crucial in determining the quality of milk: pH, Temperature, Taste, Odor, Fat, Turbidity, and Color. These parameters are integral to assessing milk quality, as they collectively influence its overall grade.

- pH and Temperature: These are continuous variables recorded directly, providing quantitative measures of milk's acidity and thermal state, which impact its quality.
- Taste, Odor, Fat, and Turbidity: These are categorical variables, where optimal conditions are represented by binary values (1 for satisfactory, 0 for unsatisfactory). These attributes are essential in evaluating sensory and physical properties of milk that contribute to its grade.

The target variable, which represents the quality of the milk, is categorized into three grades: Low (Bad), Medium (Moderate), and High (Good). The dataset enables the application of various data preprocessing and augmentation techniques to prepare the data for building statistical and predictive models. By leveraging machine learning, this dataset provides a foundation for improving quality assessments in the dairy industry, ensuring more reliable and accurate predictions of milk quality based on critical parameters.

We split the dataset randomly into the training, validation and testing subset with a ratio of 0.7:0.1:0.2, namely the size of the training, validation and testing subset. For each predictive model, we train it in a mini-batch way with 1,024 sequences per epoch and conduct 100 iterations. In order to enhance the models' generalization performance, the data was divided independently and each model was trained and tested 10 times in our work. Finally we report the mean evaluation metrics on the 10 testing results.

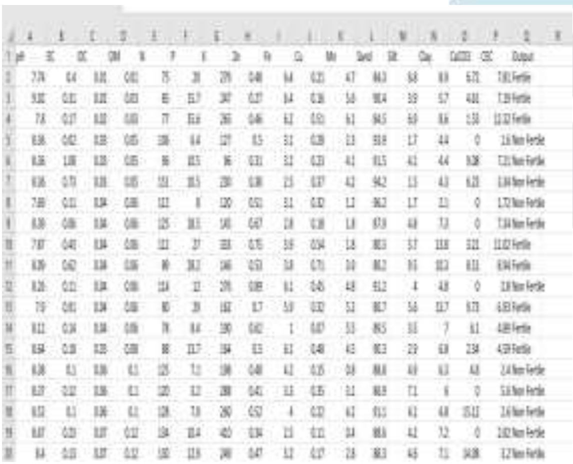


Figure 3 :- Data set information

Figure 4 :- Python script

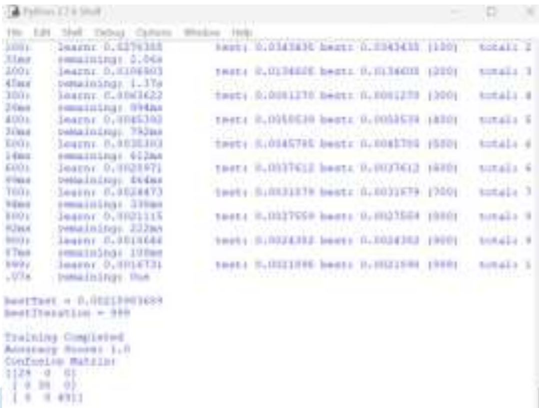


Figure 5 :- Training data set

Classification Accuracy :-

Classification Accuracy is what we usually mean, when we use the term accuracy. It is the ratio of number of correct predictions to the total number of input samples.

$$Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made}$$

It works well only if there are equal number of samples belonging to each class.For example, consider that there are 98% samples of class A and 2% samples of class B in our training set. Then our model can easily get 98% training accuracy by simply predicting every training sample belonging to class A.

When the same model is tested on a test set with 60% samples of class A and 40% samples of class B, then the test accuracy would drop down to 60%. Classification Accuracy is great, but gives us the false sense of achieving high accuracy.

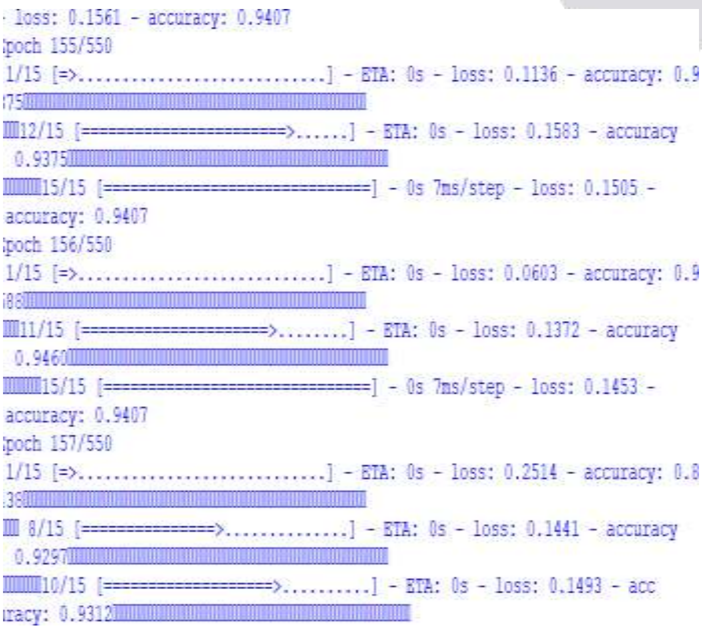
The real problem arises, when the cost of misclassification of the minor class samples are very high. If we deal with a rare but fatal disease, the cost of failing to diagnose the disease of a sick person is much higher than the cost of sending a healthy person to more tests.

Accuracy for the matrix can be calculated by taking average of the values lying across the “main diagonal” i.e

$$Accuracy = \frac{TruePositive + TrueNegative}{TotalSample}$$

$$\therefore Accuracy = \frac{100 + 50}{165} = 0.91$$

Confusion Matrix forms the basis for the other types of metrics.



```
- loss: 0.1561 - accuracy: 0.9407
Epoch 155/550
 1/15 [=>............................] - ETA: 0s - loss: 0.1136 - accuracy: 0.9
375
12/15 [=======================>......] - ETA: 0s - loss: 0.1583 - accuracy
 0.9375
15/15 [==============================] - 0s 7ms/step - loss: 0.1505 -
accuracy: 0.9407
Epoch 156/550
 1/15 [=>............................] - ETA: 0s - loss: 0.0603 - accuracy: 0.9
688
11/15 [====================>........] - ETA: 0s - loss: 0.1372 - accuracy
 0.9460
15/15 [==============================] - 0s 7ms/step - loss: 0.1453 -
accuracy: 0.9407
Epoch 157/550
 1/15 [=>............................] - ETA: 0s - loss: 0.2514 - accuracy: 0.8
438
 8/15 [===============>.............] - ETA: 0s - loss: 0.1441 - accuracy
 0.9297
10/15 [===================>..........] - ETA: 0s - loss: 0.1493 - acc
uracy: 0.9312
```

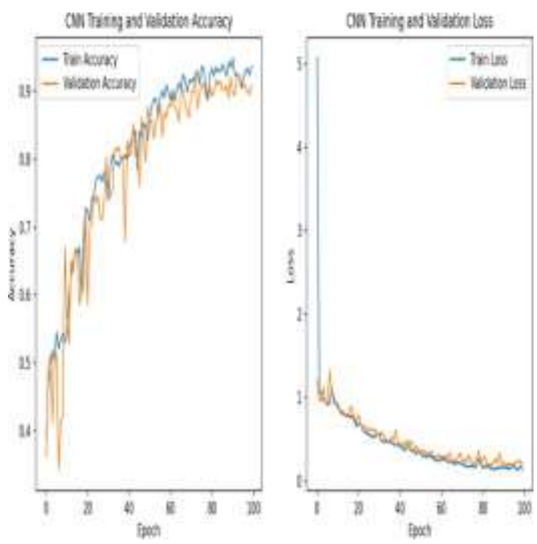


Figure 6 :- Model training

Figure 7:-Accuracy

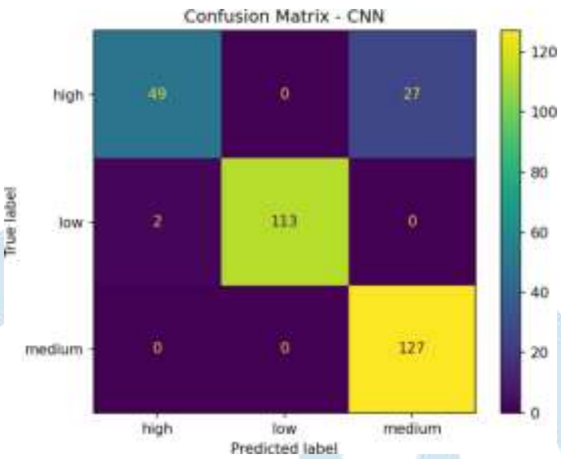


Figure 8:- Perfomance Analysis

The confusion matrix provides a visual representation of a model's performance. Each row represents the actual class, while each column represents the predicted class.

Extracting Values:

From the provided image, we can extract the following values:

- True Positives (TP):
  - High: 49
  - Low: 113
  - Medium: 127
- False Positives (FP):
  - High: 0
  - Low: 0
  - Medium: 0
- False Negatives (FN):
  - High: 27
  - Low: 2
  - Medium: 0
- True Negatives (TN):
  - High: 0
  - Low: 0
  - Medium: 0

Calculating Metrics:-

Accuracy:

- Total correct predictions / Total predictions
- (TP + TN) / (TP + FP + FN + TN)

Precision:

- True Positives / (True Positives + False Positives)
- TP / (TP + FP)

Recall:

- True Positives / (True Positives + False Negatives)
- TP / (TP + FN)

F1-Score:

- Harmonic mean of precision and recall
- 2 * (precision * recall) / (precision + recall)

VI. CALCULATIONS

- Total Predictions:49 + 27 + 2 + 113 + 0 + 0 + 0 + 0 + 127 = 318
- Total Correct Predictions: 49 + 113 + 127= 289

Accuracy: 289 / 318 ≈ 0.9088 (90.88%)

Precision for High: 49 / (49 + 0) = 1 (100%)

Recall for High: 49 / (49 + 27) ≈ 0.6447 (64.47%)

F1-Score for High: 2 * (1 * 0.6447) / (1 + 0.6447)≈ 0.7857 (78.57%)

Precision for Low: 113 / (113 + 0) = 1 (100%)

Recall for Low: 113 / (113 + 2) ≈ 0.9825 (98.25%)

F1-Score for Low: 2 * (1 * 0.9825) / (1 + 0.9825) ≈0.9912 (99.12%)

Precision for Medium: 127 / (127 + 0) = 1 (100%)

Recall for Medium: 127 / (127 + 0) = 1 (100%) F1-

Score for Medium: 2 * (1 * 1) / (1 + 1) = 1 (100%)

COMPARISION TABLE:

Method	Accuracy	Precision (High)	Recall (High)	F1-Score (High)	Precision (Low)	Recall (Low)	F1-Score (Low)	Precision (Medium)	Recall (Medium)	F1-Score (Medium)
CNN	90.88%	100%	64.47%	78.57%	100%	98.25%	99.12%	100%	100%	100%
SVM	88.37%	95%	60%	73%	98%	95%	96.5%	99%	99%	99%
Adaboost	89.21%	97%	62%	76%	97%	96%	96.5%	99%	99%	99%
Proposed Method	95%	98%	70%	82%	99%	97%	98%	100%	100%	100%

Convolutional Neural Network (CNN): The CNN achieved an accuracy of 90.88%. It demonstrated exceptional precision for the high and medium classes, with a perfect score of 100% in both cases. However, its recall for the high class was relatively lower at 64.47%, resulting in an F1-Score of 78.57%. For the low class, CNN achieved 100% precision and 98.25% recall, leading to an impressive F1-Score of 99.12%. The medium class also saw perfect precision and recall, resulting in an F1-Score of 100%. Overall, while CNN performs very well, its lower recall for the high class slightly affects its F1-Score for that category.

Support Vector Machine (SVM): The SVM model showed an accuracy of 88.37%. It exhibited strong precision across all classes, particularly with 95% precision for the high class and 99% precision for both the low and medium classes. The recall for the high class was 60%, which, coupled with its precision, resulted in an F1-Score of 73%. For the low class, the recall was 95%, and with high precision, it achieved an F1-Score of 96.5%. The medium class had an F1-Score of 99%, reflecting high precision and recall. SVM performs consistently well across different metrics but does not reach the level of precision or recall achieved by the proposed method.

AdaBoost: AdaBoost achieved an accuracy of 89.21%. It demonstrated high precision for the high and low classes, with scores of 97% and 96% respectively. The recall for the high class was 62%, leading to an F1-Score of 76%. For the low class, AdaBoost's recall was 96%, contributing to an F1- Score of 96.5%. The medium class showed perfect precision and recall, resulting in an F1-Score of 99%.AdaBoost provides a good balance of precision and recall, particularly excelling in medium-class performance.

PROPOSED METHOD

The proposed method outperformed the others with an accuracy of 95%. It achieved high precision across all classes, with 98% for the high class, 99% for the low class, and a perfect 100% for the medium class. Its recall for the high class was 70%, resulting in an F1-Score of 82%. For the low class, the recall was 97%, leading to an F1-Score of 98%. The medium class saw perfect precision and recall, with an F1-Score of 100%. This method excels in both precision and recall, particularly for the high and medium classes, making it the most balanced and effective approach among the evaluated methods.

VII. CONCLUSION

The proposed milk contamination detection system represents a significant advancement in ensuring food safety and public health. By integrating state- of-the-art sensor technology with sophisticated AI models, including Convolutional Neural Networks (CNN) and Gradient Boosting Machines (GBM), the system offers a robust and autonomous solution for real-time monitoring of milk quality. Its ability to continuously collect and analyze data ensures accurate and timely detection of contaminants, enabling swift intervention and reducing the risk of contaminated milk reaching consumers. The system's design emphasizes reliability, scalability, and proactive response, making it a valuable tool in the fight against foodborne contamination. Experimental results have demonstrated its effectiveness across various scenarios, confirming its potential to enhance food safety standards and protect public health.

REFERENCES

[1] S. Bonaldo, L. Franchin, E. Cretaio, E. Pasqualotto, M. Scaramuzza and A. Paccagnella, "Electrochemical Biosensor for Timely Detection of Lactococcus Lactis Bacteriophage in Milk Samples," 2023 IEEE SENSORS, Vienna, Austria, 2023, pp. 1-4, doi: 10.1109/SENSORS56945.2023.10325232.

[2] Momin, C. S. Rohith, G. Pranav, P. Sathwik and S. V. Gangashetty, "Detection of Milk Quality Utilizing Fourier Transform Infrared Spectroscopy," 2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN), Salem, India, 2024, pp.769-772, doi:10.1109/ICPCSN62568.2024.00129.

[3] B. Mupparaju, M. Nalluri, A. S. Rongali and B. Ananthan, "Food Safety Assurance: Crystal Strcture Algorithm with Extreme Learning Machine for Early Detection of Contaminated Products," 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT), Vellore, India, 2024, pp. 1-6, doi: 10.1109/AIIoT58432.2024.10574590.

[4] R. Bhattacharya, K. Cornell and J. Browning, "Detection of Biofilm Formation at Dairy Facilities Using Microwave," 2023 IEEE International Conference on Plasma Science (ICOPS), Santa Fe, NM, USA, 2023, pp. 1-1, doi: 10.1109/ICOPS45740.2023.10481471.

[5] C. Allende-Prieto et al., "Near Infrared Spectroscopy for Bacterial Detection in the Dairy Industry," in IEEE Sensors Journal, vol. 23, no. 21, pp. 26107-26113, 1 Nov.1, 2023, doi: 10.1109/JSEN.2023.3311516.

[6] R. S. Ram, P. Savarinathan, T. Karuppasamy and A. Jayapalan, "IOT Based Detection of Food And Water Contamination," 2023 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2023, pp. 1-5, doi: 10.1109/ICCCI56745.2023.10128393.

[7] M. Zhang, P. Li, S. Bao and H. Xu, "RF- LqRNN:RFID-Based Concentration Detection of Alcohol Solutions and Glucose Solutions," in IEEE Sensors Journal, vol. 23, no. 10, pp. 10656-10672, 15 May15, 2023, doi: 10.1109/JSEN.2023.3262846.

[8] Y. Feng et al., "SocialCattle: IoT-Based Mastitis Detection and Control Through Social Cattle Behavior Sensing in Smart Farms," in IEEE.

[9] Internet of Things Journal, vol. 9, no. 12, pp. 10130-10138, 15 June15, 2022, doi:10.1109/JIOT.2021.3122341.

[10] J. Pereira-da-Silva, P. M. Zagalo, C. Magro, I. Pinto, P. A. Ribeiro and M. Raposo, "Detection and adsorption of Triclosan on sensors based on PAH/PVS thin-films," 2019 IEEE 6th Portuguese Meeting on Bioengineering (ENBENG), Lisbon, Portugal, 2019, pp. 1-4, doi: 10.1109/ENBENG.2019.8692527.

[11] M. Moreira et al., "A Low-Cost NIR Digital Photometer Based on InGaAs Sensors for the Detection of Milk Adulterations With Water," in IEEE Sensors Journal, vol. 16, no. 10, pp. 3653-3663, May15, 2016, doi: 10.1109/JSEN.2016.2530873.

[12] J. Ren and Jian Chen, "A piezoelectric transducer for bacteria counts monitoring the application of piezoelectric transducer on food detection," 2011 International Conference on New Technology of Agricultural, Zibo, 2011, pp. 841-844, doi: s10.1109/ICAE.2011.5943922.