An internet of things (IoT)-based optimum tea fermentation detection model using convolutional neural networks (CNNs) and majority voting techniques

Gibson Kimutai1,2, Alexander Ngenzi1, Said Rutabayiro Ngoga1, Rose C. Ramkat3, and Anna Förster4

1African Centre of Excellence in Internet of Things, University of Rwanda, Kigali, Rwanda
2Department of Mathematics, Physics and Computing, Moi University, Eldoret, Kenya
3Department of Biological Sciences, Moi University, Eldoret, Kenya
4Department of Sustainable Networks, University of Bremen, Bremen, Germany

Correspondence: Gibson Kimutai (kimutaigibs@gmail.com)

Received: 11 October 2020 – Revised: 28 April 2021 – Accepted: 5 June 2021 – Published: 2 July 2021

Abstract. Tea (Camellia sinensis) is one of the most consumed drinks across the world. Based on processing techniques, there are more than 15 000 categories of tea, but the main categories include yellow tea, Oolong tea, Illex tea, black tea, matcha tea, green tea, and sencha tea, among others. Black tea is the most popular among the categories worldwide. During black tea processing, the following stages occur: plucking, withering, cutting, tearing, curling, fermentation, drying, and sorting. Although all these stages affect the quality of the processed tea, fermentation is the most vital as it directly defines the quality. Fermentation is a time-bound process, and its optimum is currently manually detected by tea tasters monitoring colour change, smelling the tea, and tasting the tea as fermentation progresses. This paper explores the use of the internet of things (IoT), deep convolutional neural networks, and image processing with majority voting techniques in detecting the optimum fermentation of black tea. The prototype was made up of Raspberry Pi 3 models with a Pi camera to take real-time images of tea as fermentation progresses. We deployed the prototype in the Sisibo Tea Factory for training, validation, and evaluation. When the deep learner was evaluated on offline images, it had a perfect precision and accuracy of 1.0 each. The deep learner recorded the highest precision and accuracy of 0.9589 and 0.8646, respectively, when evaluated on real-time images. Additionally, the deep learner recorded an average precision and accuracy of 0.9737 and 0.8953, respectively, when a majority voting technique was applied in decision-making. From the results, it is evident that the prototype can be used to monitor the fermentation of various categories of tea that undergo fermentation, including Oolong and black tea, among others. Additionally, the prototype can also be scaled up by retraining it for use in monitoring the fermentation of other crops, including coffee and cocoa.

1 Introduction

Tea (Camellia sinensis) is currently among the most prevalent and extensively consumed drinks across the world, with a daily consumption of more than 2 million cups. The high consumption is credited to its medicinal values, i.e. reducing heart diseases, aiding in weight management, preventing strokes, lowering blood pressure, preventing bone loss, and boosting the immune system, among others. Historical evidence indicates that the tea plant was indigenous to China and Burma, among other countries (Akuli et al., 2016). Table 1 presents the current leading tea-producing countries. Tea is a source of many types of tea, which includes Oolong tea, black tea, white tea, matcha tea, sencha tea, green tea, and yellow tea, among others. The processing techniques determine the category of tea produced. Globally, Kenya is the leading producer of black tea. Black tea is the most popular among the categories of tea, and it accounts for an estimate of 79 % (Mitei, 2011) of the entire global tea consumption.

As shown in Fig. 1, the processing steps of black tea are plucking, withering, cutting, tearing and curling, fermentation, drying, and sorting. The fermentation step is the most
Table 1. The countries leading in tea production globally.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Country</th>
<th>Estimates tonnes produced yearly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>China</td>
<td>2,400,000</td>
</tr>
<tr>
<td>2</td>
<td>India</td>
<td>1,320,000</td>
</tr>
<tr>
<td>3</td>
<td>Sri Lanka</td>
<td>439,000</td>
</tr>
<tr>
<td>4</td>
<td>Kenya</td>
<td>349,600</td>
</tr>
<tr>
<td>5</td>
<td>Vietnam</td>
<td>260,000</td>
</tr>
</tbody>
</table>

Figure 1. The basic processing steps of black tea.

crucial in deciding the final quality of the black tea (Saikia et al., 2015). During the process, catechin compounds react with oxygen during oxidation to produce two compounds, namely theaflavins (TF) and thearubigins (TR). These compounds determine the aroma and taste of the tea (Lazaro et al., 2018). Also, fermentation changes the tea colour to coppery brown and causes a fruity smell. Hence, the fermentation process must stop at the optimum point as fermentation beyond the optimum point destroys the quality of the tea (Borah and Bhuyan, 2005).

Presently, tea tasters estimate the level of fermentation of tea by monitoring change in colour, smelling the tea, and tasting an infusion of tea (Kimutai et al., 2020). These methods are biased, intrusive, consume a lot of time, and are inaccurate, which compromises the quality of the produced tea (Zhong et al., 2019).

The internet of things (IoT) has established itself as one of the greatest smart ideas of the modern day (Shinde and Shah, 2018), and its effects have been seen in each feature of human ventures, with huge possibilities for smarter living (Miazi et al., 2016). IoT has shown huge potential in many fields, including agriculture, medicine, manufacturing, sports, and governance, among others (Khanna and Kaur, 2019). Deep learning is currently shaping how machine learning is applied to various areas. Some of the prominent areas where deep learning has shown a lot of promise include machine vision, speech recognition, audio processing, health diagnosis, and fraud detection, among others (Too et al., 2019). With the complication of challenges in the 21st century, majority voting is being applied to machine learning to improve performance as it provides an extra layer of decision-making to the model. In (Kimutai et al., 2020), a deep learner dubbed “TeaNet” was developed based on image processing and machine learning techniques. The deep learner was trained, evaluated, and validated based on the dataset in (Kimutai et al., 2021). In this paper, the TeaNet model was deployed to classify real-time tea fermentation images in Sisibo Tea Factory. Additionally, a majority voting technique was applied to aid in decision-making by the prototype. The subsequent sections of this paper are presented as follows: Sect. 2 provides the related work, Sect. 3 provides the materials and methods, while Sect. 4 presents the evaluation results, and Sect. 5 gives the conclusion of the paper.

2 Related work

As mentioned in Sect. 1, the fermentation process is the most important step in determining the quality of the produced tea. Consequently, researchers have been proposing various methods of improving the monitoring of the fermentation process of tea. There are proposals to apply image processing, IoT, electronic nose, electronic tongue, and machine learning, among others. With the maturity of image processing, many researchers are presently proposing it for application in the detection of optimum tea fermentation. Saranka et al. (2016) have proposed the application of image processing and a support vector machine algorithm in the detection of the optimum fermentation of tea. Image processing has been proposed to classify black tea fermentation images (Borah and Bhuyan, 2003b; Chen et al., 2010). Additionally, Borah and Bhuyan (2003b) propose colour matching of tea during fermentation with neural networks and image processing techniques. Convolutional neural networks (CNNs) have shown great promise in image classification tasks across fields. Additionally, Krizhevsky et al. (2017) and Razavian et al. (2014) show the great capabilities of CNNs in image classification tasks. They have shown that the data-hungry nature of deep learning has been solved by the aspect of transfer learning. Consequently, CNNs are now being applied in monitoring tea processing, including the detection of optimum fermentation of tea (Kimutai et al., 2020; Kamrul et al., 2020). Furthermore, a neural-network-based model for estimating the basic components of tea theaflavins (TF) and thearubigins (TR) is proposed (Akuli et al., 2016). Chen et al. (2010) fused near-infrared spectroscopy with computer vision to detect optimum fermentation in tea. Borah and Bhuyan (2003a) proposed quality indexing of black tea as fermentation progresses. All these proposals are in the form
IoT is being applied in many fields, including agriculture. The tea sector is attracting attention from researchers, and the authors have proposed the application of IoT to monitor temperature and humidity during the fermentation of tea. Saikia et al. (2014) proposed a sensor network to monitor the relative humidity and temperature of tea during fermentation. Also, Uehara and Ohtake (2019) developed an IoT-based system for monitoring the temperature and humidity of tea during processing, while Kumar (2017) proposed an IoT-based system to monitor the temperature and humidity of tea during fermentation. The proposed works have been deployed in a tea factory to monitor temperature and humidity during tea processing. The models are, thus, a step in the right direction, but their scope was only on monitoring those physical parameters during tea processing.

From the literature, it is evident that the tea fermentation process is receiving most of the attention from researchers due to its importance in the determination of the quality of tea, with many of the proposals being an ensemble of machine learning and image processing techniques. The IoT is presently gaining momentum in its application to monitoring temperature and humidity during tea processing. However, most of the proposals are in the form of simulation models and have not been deployed in real tea fermentation environments. Additionally, deep learning is gaining more acceptance compared to standard machine learning classifiers in monitoring tea processing due to its intelligence and the ability to use transfer learning to solve challenges across various domains.

3 Materials and methods

This section presents the following: the system architecture, resources, deployment of the prototype, image database, majority voting-based model, and the evaluation metrics.
3.1 System architecture

The architecture of the proposed model to monitor the fermentation of tea in real time is presented in Fig. 2. The system had a node containing a Pi camera attached to a Raspberry model. The system is connected through Wi-Fi technology to the edge and the cloud environments.

3.2 Resources

The following resources were applied in the implementation: Raspberry Pi 3 Model B+, a Pi camera, an operating system, a server, and programming languages. We discuss each of these in the next paragraphs.

1. **Raspberry Pi 3.** This study adopted Raspberry Pi 3 Model B+ with the following specifications: quad-core 1.2 GHz, a 64-bit central processing unit (CPU), 2 GB random-access memory (RAM), a 40-pin extended general-purpose input/output (GPIO), camera serial interface port, and micro secure digital (SD) card. The model was powered by 2A power supply (Marot and Bourennane, 2017; Sharma and Partha, 2018).

2. **Pi camera.** In this research, a Raspberry Pi camera of 8 MP was used. The board was chosen since it is tiny and weighs around 3 g, making it perfect for deployment with the Raspberry Pi.

3. **Operating System.** The Raspbian operating system (Marot and Bourennane, 2017) was used. It was chosen because it has a rich library, and it is easy to work with.

4. **Server.** The Apache server (Zhao and Trivedi, 2011) was adopted to obtain data and send the data to the edge environment for the local users and a cloud-based environment for remote users. For the cloud environment, Amazon Web Services (AWS) (Narula et al., 2015) was chosen as it provides a good environment for the deployment of IoT systems.

5. **Programming languages.** The Python programming language (Kumar and Panda, 2019) was used for writing programmes to capture the images using the Pi camera. It has various libraries (Dubosson et al., 2016) and is open source (Samal et al., 2017). Some of the libraries adopted included the following: TensorFlow (Tohid et al., 2019), Keras (Stancin and Jovic, 2019), Seaborn (Fahad and Yahya, 2018), Matplotlib (Hung et al., 2020), pandas (Li et al., 2020), and NumPy (Momma et al., 2020). Additionally, the Laravel PHP framework was adopted in writing application programming interfaces (APIs). HTML5 (the hypertext markup language) and CSS (cascading style sheets) were used in designing the web interfaces of the prototype.
Figure 7. The workflow of the region-based majority voting for TeaNet.

Figure 8. Application of Raspberry Pi 3 Model B+ with the Pi camera for taking pictures during the fermentation of black tea in the Sisibo Tea Factory, Kenya.

3.3 Image database

As discussed in Sect. 1, TeaNet was trained, validated, and evaluated using data set from Kimutai et al. (2021). The data set contains 6000 images of three classes of tea, i.e. underfermented, fermented, and overfermented. Fermentation experts provided the ground truths of all the images, which enabled the classification of all the images into the three classes. From the experts’ judgement, fermentation degrees of tea with time depends on the following factors: time (Obanda et al., 2001), temperature and humidity level at which fermentation takes place (Owuor and Obanda, 2001), the clones of the tea, nutrition levels of the tea, age of tea, stage of growth of tea, plucking standards, and post-harvesting handling. Presently, more than 20 clones of tea are grown in Kenya (Kamunya et al., 2012). Figure 3 shows sample images of three classes of tea, i.e. unfermented, fermented and overfermented.

A Pi camera attached to a Raspberry Pi was adopted for capturing the images. The images were collected in the Sisibo Tea Factory in Kenya. Underfermented tea is tea for which the fermentation is below optimum, making it of low quality, while fermented tea is optimally fermented and is considered to be the ideal tea. For overfermented tea, the fermentation cycle has exceeded the optimum, making the tea harsh and poor in quality.

3.4 Majority voting for TeaNet

We developed a deep learning model that is dubbed TeaNet (Kimutai et al., 2020) and is based on CNNs. AlexNet (Sun et al., 2016), the widely used network architecture in CNNs, inspired the development process. We designed the model for simplicity and to reduce the computational needs. TeaNet was chosen over the traditional methods since it outperformed the traditional methods in the simulation experiments reported in Kimutai et al. (2020). Additionally, TeaNet as a deep-learning-based model is trained rather than programmed; thus, it does not require much fine-tuning. TeaNet is flexible as it be can be retrained using other data sets for
other domain applications, unlike OpenCV algorithms that are domain-specific (O’Mahony et al., 2020). With transfer learning (Sutskever et al., 2013), TeaNet can be applied to solve challenges in other fields. Figure 4 shows the architecture of the developed TeaNet model.

Gaussian distribution was adopted in initializing the weights of the network. A stochastic gradient descent (Sutskever et al., 2013) technique, with a batch size of 16 and a momentum value of 0.9, was chosen. The rate of learning was 0.1, with a threshold minimum of 0.0001. The network learning iterations were set at 50, with a weight decay of 0.0005. The model registered a steady increase in accuracy with increasing epoch numbers, registering a perfect precision of 1.0 at epoch 10. The accuracy of the model increased with each iteration as the weights of the neurons were turned after every iteration. The validation accuracy of the model was 1.0 at epoch 10 (Fig. 5).

Figure 6 presents the loss of the model during training and validation. The loss steadily reduces, with epoch increases up to epoch 10 where there is no loss. These values are promising, as they depict a stable loss rate.

In this paper, we have proposed a region-based majority (RBM) voting for TeaNet (RBMTeaNet) composed of three steps. The steps were image segmentation, the training of the TeaNet model, and majority voting (Fig. 7). After the image collection, as discussed in Sect. 3.3, the images were prepared for input into the CNN network by resizing them to 150 \( \times \) 150 pixels. The semantic segmentation annotation method was followed to annotate the images according to Tylecek and Fisher (2018). Some of the common types of noise in the images include photon noise, readout noise, and dark noise (Khan and Yairi, 2018). To perform denoising, the linear filtering method was adopted.

For the region-based majority voting for TeaNet (RBM-TeaNet), each region was labelled by voter data generated by the region majority voting system. Each of the patches had three voters. One of the voters was in the centre of the patch, with the others being generated randomly within the patch. Finally, the classification results were arrived at from the candidate label that had the highest vote numbers.

3.5 Deployment of the prototype

We deployed the developed tea fermentation monitoring system in a tea fermentation bed in the Sisibo Tea Factory, Kenya (Fig. 8). The Pi camera was attached to a Raspberry Pi model and used to take images of the tea in the fermentation bed at an interval of 1 min. The learned model developed in Kimutai et al. (2020) was trained and deployed in the Sisibo Tea Factory for validation between 1 and 30 July 2020 and thereafter evaluated between 10 and 16 August 2020. In the tea fermentation bed, a Pi camera was deployed to take tea fermentation images in real time. The server side contained the Raspberry Pi, a Wi-Fi router, and internet wall. Every collected image was sent to the Raspberry Pi through the jumper wires. In Raspberry, the TeaNet model predicted the classes of images based on the knowledge gained during training. The image was sent to the cloud servers by the Raspberry Pi with the use of Wi-Fi technology for use by remote users. The internet wall was used to secure the connection to cloud servers. A copy of each image was then sent to the edge servers locally for use by local users. Each real-time image was displayed on the web page, which is accessible through both mobile phones and computers. Additionally, the web page displayed images alongside their predicted classes.

3.6 Evaluation metrics

The following metrics were used in the evaluation of the performance of the developed model when deployed in the tea factory: precision, accuracy, and confusion matrix.

3.6.1 Precision

Precision is the degree of refinement with which a classification task is undertaken (Flach, 2019). It can be represented by Eq. (1). 

\[
\text{Precision} = \frac{TP}{TP + FP}, \tag{1}
\]

where TP is the correct classification of a positive class, and FP is the incorrect classification of a positive class.

3.6.2 Accuracy

Accuracy is the measure of how a classifier accurately predicted the total number of input samples (Flach, 2019). Equation (2) shows its representation.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \tag{2}
\]

where TP is when a positive class is classified accurately, and FN is when a negative class is correctly classified, and FP is when a positive class is incorrectly classified, and TN is when a negative class is accurately classified.

3.6.3 Confusion matrix

A confusion matrix is adopted in classification tasks to evaluate a classification model based on the true values (Flach, 2019). Sensitivity highlights the number of true positives that are accurately classified. Equation (3) shows the representation of sensitivity.

\[
\text{Sensitivity} = \frac{TP}{TP + FN}, \tag{3}
\]

where TP is when a positive class is accurately predicted by a classifier, and FP is when a negative class is incorrectly predicted by a classifier as a positive class. FN denotes the incorrect classification of a negative class.
4 Evaluation results

The model developed in this research was evaluated based on precision, accuracy, and confusion matrix, as discussed in Sect. 3.6. Figure 9 shows the evaluation results of the model based on average precision. The evaluation results showed that TeaNet produced a perfect precision of 1.0, with an average of between 0.8485 and 0.9589 in real deployment. The model produced an average precision of between 0.9260 and 0.9737 when TeaNet and majority voting were combined. From the results, TeaNet performed better in terms of precision when evaluated offline compared to when it was evaluated in real deployment.

Figure 10 shows the evaluation results of the TeaNet model based on accuracy. Also, TeaNet showed high effectiveness when evaluated offline, based on the achieved average accuracy of 1.0 across the scanning days. When TeaNet was evaluated in a real deployment environment, it achieved an average accuracy of between 0.7179 and 0.8646 across the scanning days. Additionally, when majority voting was employed to aid in the decision-making process of TeaNet, performance improved to an average ranging between 0.8372 and 0.8916.

Table 2 presents the performance of the RBM-TeaNet model in terms of sensitivity. In the tea factory, overfermented tea was not found, since such tea is low in quality.
Table 2. Confusion matrix of RBM-TeaNet model during the fermentation of tea.

<table>
<thead>
<tr>
<th>Class</th>
<th>Unfermented</th>
<th>Fermented</th>
<th>Overfermented</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 August 2020 (unfermented)</td>
<td>41</td>
<td>4</td>
<td>0</td>
<td>91.1 %</td>
</tr>
<tr>
<td>10 August 2020 (fermented)</td>
<td>3</td>
<td>39</td>
<td>5</td>
<td>83.0 %</td>
</tr>
<tr>
<td>11 August 2020 (unfermented)</td>
<td>45</td>
<td>6</td>
<td>0</td>
<td>88.2 %</td>
</tr>
<tr>
<td>11 August 2020 (fermented)</td>
<td>6</td>
<td>29</td>
<td>2</td>
<td>78.4 %</td>
</tr>
<tr>
<td>12 August 2020 (unfermented)</td>
<td>42</td>
<td>5</td>
<td>0</td>
<td>89.4 %</td>
</tr>
<tr>
<td>12 August 2020 (fermented)</td>
<td>3</td>
<td>23</td>
<td>4</td>
<td>76.7 %</td>
</tr>
<tr>
<td>13 August 2020 (unfermented)</td>
<td>48</td>
<td>3</td>
<td>0</td>
<td>94.1 %</td>
</tr>
<tr>
<td>13 August 2020 (fermented)</td>
<td>6</td>
<td>26</td>
<td>5</td>
<td>70.3 %</td>
</tr>
<tr>
<td>14 August 2020 (unfermented)</td>
<td>42</td>
<td>5</td>
<td>0</td>
<td>89.4 %</td>
</tr>
<tr>
<td>14 August 2020 (fermented)</td>
<td>3</td>
<td>28</td>
<td>4</td>
<td>80.0 %</td>
</tr>
<tr>
<td>15 August 2020 (unfermented)</td>
<td>48</td>
<td>4</td>
<td>0</td>
<td>92.3 %</td>
</tr>
<tr>
<td>15 August 2020 (fermented)</td>
<td>0</td>
<td>28</td>
<td>5</td>
<td>84.8 %</td>
</tr>
<tr>
<td>16 August 2020 (unfermented)</td>
<td>45</td>
<td>2</td>
<td>0</td>
<td>95.7 %</td>
</tr>
<tr>
<td>16 August 2020 (fermented)</td>
<td>4</td>
<td>27</td>
<td>3</td>
<td>79.4 %</td>
</tr>
</tbody>
</table>

and no tea factory allows the fermentation of tea to reach that level. Generally, the model had good sensitivity across days, with a minimum of 70.3 % being achieved on 13 August 2020 where the model classified six fermented images as underfermented and five of the fermented images were classified as overfermented. More promisingly, the model could not confuse unfermented and overfermented tea. This is because of the clear distinction in the two classes in terms of colour.

5 Conclusions

This research has proposed a tea fermentation detection system based on IoT, deep learning, and majority voting techniques. The IoT components were Raspberry Pi 3 Model B+ and a Pi camera. The deep learner model was composed of three convolutional layers and three pooling layers. The model developed was deployed to monitor tea fermentation in real time in a tea factory in Kenya. The capabilities of the system were assessed based on the ground truths provided by tea experts. The results from the evaluation are promising and signify a breakthrough in the application of IoT, CNNs, and majority voting techniques in the real-time monitoring of tea fermentation. The same technique can be applied to monitor the processing of other categories of tea that undergo fermentation, including Oolong tea. Additionally, the prototype can be used for monitoring the fermentation of coffee and cocoa, since all of them have a distinction in colour based on the fermentation degrees. It is recommended that future studies monitor the physical parameters (temperature and humidity) of tea during fermentation to find their effect on the quality of the made tea.

Data availability. All relevant data presented in the article are stored according to institutional requirements and as such are not available online. However, all data used in this paper can be made available upon request to the corresponding authors.

Author contributions. The formal analysis and methodology were done by GK and AF, who also reviewed and edited the paper after GK wrote it. GK developed the software, and AF, SRN, and AN supervised the project. GK visualized the projected and validated it with AF and RCR.

Competing interests. The authors declare that they have no conflict of interest.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. We thank the chief executive officer and the entire staff of the Sisibo Tea Factory, Kenya, for allowing us to deploy the developed system in their factory and for providing insight and expertise that greatly assisted the research.

Financial support. This research received financial support from the Kenya Education Network (KENET) through the CS/IT mini-grants and the African Centre of Excellence in the Internet of Things (ACEIoT).

Review statement. This paper was edited by Qingquan Sun and reviewed by two anonymous referees.
References


