

Classification of Urease Activity in Full-Fat Soybean Production by Extrusion Using Machine Learning Algorithms

Ilyas Ozer

Abstract—Soybean is an important food source that is frequently preferred in animal feeds with its high protein value. However, soybeans contain many bioactive compounds that are antinutritional and/or poisonous. Urease is one of the most important of these. Processes such as extrusion is used to reduce these components' effect. Here, factors such as steam pressure and temperature affect the cooking level of the product. In the case of undercooked soybeans, components that harm animal health preserve their effect, while their nutritional value decreases in case of overcooking. The urease test has been used for many years to evaluate the cooking level of soybean. Here, according to the color change on the product as a result of the test, the cooking level is evaluated by an expert. This process is mostly done manually and is dependent on expert judgment. In this study, a machine learning-based approach has been proposed to evaluate the images of urease test results. Accordingly, samples were taken from the extruder during the processing of full-fat soybean. A data set consisting of overcooked, well-cooked and undercooked sample images was prepared by performing the urease test. A binary classification process as cooked and undercooked and a classification process with three classes was carried out with four different machine learning models on the data set. In this way, it is aimed to both automate the process and minimize the problems that may arise from expert errors. Classification achievements of 96.57% and 90.29% were achieved, respectively, for two and three class tests with the CNN-LSTM model in 10-fold cross-validation tests.


Index Terms—Convolutional neural network, Long short-term memory network, Soybean urease test.

I. INTRODUCTION

SOYBEAN MEAL is a very important part of all protein sources used in animal feed worldwide [1] due to its high protein concentration [2][3][4]. In addition, soybean meal extracted from oil (SBM) and full-fat (FFSB) are the major global raw materials for broiler diets [5]. The high protein content of soy and the wide availability of oil-extracted soybean by-products have made soybean a widely used alternative to animal protein sources.

Still, soybeans contain an exceptionally high concentration of bioactive compounds that are antinutritional and/or poisonous that have a detrimental effect on animals' metabolism [6].

İLYAS ÖZER, is with Department of Computer Engineering University of Bandırma Onyedi Eylül University, Balıkesir, Turkey, (e-mail: iozer@bandirma.edu.tr).

 <https://orcid.org/0000-0003-2112-5497>

Urease is one of these factors. The harmful effects of feeding urease-containing meals to animals have been reported in the literature. When unprocessed soybeans are combined with urea, ammonia is released due to the activity of urease, which is an unwanted result in the feed [2].

Ammonia reaches the bloodstream rapidly in ruminants and can cause a variety of negative effects, including decreased feed intake, decreased animal health, ammonia poisoning, and death [7]. Heat treatment is the most common procedure for removing or reducing the effects of antinutritional and/or harmful factors, such as urease, in soybeans [8][9][10][11][12]. On the other hand, as a result of Maillard (browning) reactions caused by overcooking raw soybean grain (high cooking temperature, excessive steam pressure or prolonged), lysine combines with carbohydrates to form a complex and its usefulness is greatly reduced [2][6][13].

Various processes such as microwave heating, fluidized-bed drying, spouted bed drying, extrusion, superheated steam and boiling can be used to inactivate the undesirable components of soybeans [10][14]. The high temperature of the extrusion and the screw speed effectively releases cellulosic microcrystals in the cell wall structure and, consequently, in the fragmentation of the wall. In addition, it is effective in reducing soybean-based anti-nutritional factors such as trypsin inhibitors [11][12]. Extrusion is also environmentally friendly due to its short processing time, low cost, protection of heat sensitive components, industrial ability, and the absence of hazardous chemical waste [11][12] has several advantages. Figure 1 shows an example extruder machine.

In summary, if soybeans are not adequately processed, undesirable conditions such as urease activity occur. On the other hand, its usefulness decreases significantly if it is overcooked. Biological experiments are the most effective method of determining processing performance and final soybean meal content (Real-Guerra et al., n.d.). However, the expense, time required, and difficulty of these tests limit their usage. Because of its rapidity and low laboratory equipment requirements, urease testing has been used as an indirect method of determining the heat treatment capacity of soybeans since the 1940s. An analysis (Yalcin and Basman, 2015) discovered a strong association between the activities of trypsin inhibitors, urease, and lectins, demonstrating that these analytical parameters could greatly predict the performance of soybean production. Many protocols have been established over the years to promote the calculation of urease behavior. These protocols measure the ammonia emitted either directly or indirectly. The Caskey-Knapp

process, one of the first developed (Caskey and Knapp, 1944), involves incubating the meal with urea in a buffered solution before adding phenol red. Unprocessed meals cause a rise in the pH of the solution with a color change (from red-orange to pink) after incubation, while adequately processed meals cause little or no color change.



Fig.1. An example of an industrial extruder machine.

Urease test is also frequently preferred in the production of industrial-sized full-fat soybean by extrusion. Here, soybeans are classified as undercooked, well-cooked and over-cooked according to the color changes after testing by an expert. This process is entirely manual. It also depends heavily on the experience of the expert who made the assessment. In this study, an approach is presented on the classification of images obtained as a result of the urease test with machine learning techniques. In this way, it is aimed to both automate the process and eliminate the wrong evaluation results that may arise from human errors. Within the scope of his study, a data set consisting of 175 images, including at least 57 images from each class, was prepared. The prepared data set was classified with different machine learning methods and the results were reported. Many studies in the literature use soybean and machine learning approaches together, such as determination of legume type [15], yield estimation [16][17][18], evaluation of crop damage [19] and determination of seed and seedling quality [20]. According to the author's knowledge, this is the first study to analyze urease activity after extrusion on TYS with machine learning techniques.

II. MATERIALS AND METHODS

A. Data set

Periodic samples were taken from a soybean extruder to create the data set. The samples' contact with the appropriate solution was ensured, and the images of the outputs were labeled and recorded in accordance with the expert opinions. For this process, 25 ml of amber-colored urea-phenol red solution was added to the homogenized soybean meal and shaken gently. After this process, all samples were kept for 5 minutes and according to the red color formation on them, the class was decided by the experts. If there is no red color on the surface, there is no urease activity and the sample is over-cooked. If there are a few small scattered red particles, there is little urease activity and the product is considered well-cooked. If a large part of the sample surface is covered with red particles, the urease activity is high and the product is undercooked. In Figure 2, samples are belonging to all three classes.



Fig.2. Samples of urease test images (From left to right: well-cooked, undercooked and overcooked).

As a result of all these processes, a total of 175 pictures of 57 undercooked, 60 well-cooked and 58 over-cooked samples were prepared.

B. Fully Connected Neural Network

Fully Connected Neural Networks (FCNN) are among the most basic elements of artificial neural networks and have been used in numerous applications to date. Being structure agnostic is one of the main features of these networks [21]. These networks do not make any special assumptions about input. This makes them applicable to many different kinds of problems. However, the performance of FCNNs tends to be lower than networks adapted to solve a particular problem [21].

FCNN is created by placing fully connected layers one after the other. The fully connected layer can be defined as a function from $x \in \mathbb{R}^m$ to \mathbb{R}^n . Figure 3 shows a sample fully connected layer.

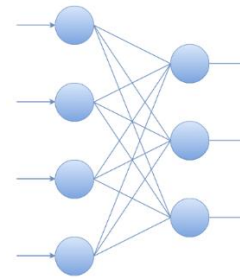


Fig.3. Representations of fully connected layer.

Where $x \in \mathbb{R}^m$ is the input, w_i is the learnable parameters, f is the non-linear function for the fully connected layer and $y_i \in \mathbb{R}^m$ is the i -th output of the fully connected layer, here y_i is calculated as follows:

$$y_i = \begin{pmatrix} f(w_{1,1}x_1 + \dots + w_{1,m}x_m) \\ \vdots \\ f(w_{n,1}x_1 + \dots + w_{n,m}x_m) \end{pmatrix} \quad (1)$$

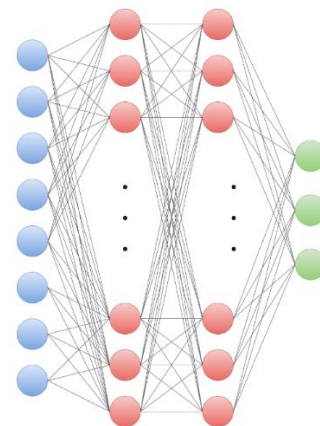


Fig.4. FCNN model architecture.

C. Convolutional Neural Network

CNN is a typical multilayer neural network structure, often used to analyze image-related applications [22][23][24]. The working principle of CNN architectures can be summarized as extracting the features of the image taken from the input layer and classifying the extracted features. Here fully connected layers are commonly used for the classification process. CNN's greatly adorn feed-forward neural networks. CNN-based machine learning models differ from traditional machine learning methods through convolution layers that can automatically extract features [25], [26]. CNN models generally consist of convolution, pooling and fully connected layers [25].

Filter size and the number of generated maps are used to define the convolution layer. Here, filters are a basic unit used to extract different features related to lines, corners and edges on the images [27]. These filters are shifted across the image matrix. During the shifting process, the values of the image matrix are multiplied by the values in the filter. The net result is found by summing the values obtained afterward. This process is applied to the entire image to generate feature maps. As a result, a new matrix is created [28]. Here y_l is the output vector, the number of elements in the input signal x_n , the filter h_{l-n} , the feature map values can be calculated as follows [29]:

$$y_l = \sum_{n=0}^{N-1} x_n h_{l-n} \quad (2)$$

To add non-linearity to the convolution layer, a rectified linear unit (ReLU) is generally used as the activation function. In this function, negative input values are removed by setting them to zero. To represent the x is input, the ReLU function can be expressed as:

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

In CNN architectures, there is usually the pooling layer after the convolution layer. These small rectangular blocks are used to reduce the size of the output of the convolution layer [29]. Thus, both computational costs are reduced and the problem of overfitting is minimized. In this study, the max-pooling method that takes the largest value in each rectangular block is used.

$$p_j = \frac{e^{x_j}}{\sum_1^k e^{x_k}} \quad j = 1, \dots, k \quad (4)$$

D. Long Short-Term Memory Network

LSTM is a special form of recurrent neural networks (RNN). It can learn long-term dependencies. First proposed in the mid-90s, this model is widely used today [30]. RNNs aims to store and transfer the state information of the artificial neural network while working on data in the sequences. However, status information is continuously processed and transmitted. For this reason, it is often not possible to transfer long-term dependencies without breaking them. That is, short-term dependencies can be transferred effectively. However, there is a problem in transferring long-term dependencies intact. LSTMs are models developed to deal with this problem.

All RNN-based network models consist of repetitive structures in the form of a chain. The main feature that distinguishes RNN models from each other is that their internal structures are different. In basic RNN models, these structures usually consist of a tanh layer or contain a similar function. The internal structures of LSTMs are different from the basic RNN models, as seen in Figure 5.

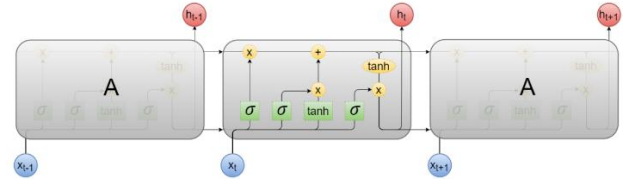


Fig.5. Internal structure of an LSTM module and its interaction with other modules

An LSTM module consists of three separate gates. These are the input, forgetting and output gates. The forgetting gate consists of a sigmoid function. The sigmoid function generates a value between 0 and 1. In case the generated value is 0, no information is transmitted. If it is 1, it means that all information must be transmitted. This process can be expressed mathematically as follows:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (5)$$

It is then decided what information needs to be updated. For this operation, the sigmoid function is used again. In addition, these two processes are combined to create a list of candidate values. \tilde{C}_t is a list of candidate values, and these two operations can be mathematically expressed as:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (6)$$

$$\tilde{C} = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (7)$$

After this step, the new state information of the memory cell is calculated. This process can be expressed as follows:

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \quad (8)$$

Finally, the output of the system h_t is calculated. This process can be expressed as follows:

$$h_t = o_t \tanh(C_t) \quad (9)$$

E. CNN-LSTM Network

In CNN architectures, the process can be summarized in two basic stages. In the first step, feature extraction is performed on the input data. Afterward, these features are commonly classified using an FCNN. On the other hand, different network structures may be more capable of capturing different relationships. For example, while CNN is successful in spatial relationships, LSTM models are more successful in temporal relationships. Therefore, hybrid models can be used to combine the capabilities of different network structures. In the CNN-LSTM model used in this study, the features extracted with the CNN architecture are applied to the LSTM layer. Details on model parameters are shown in Table 1.

F. Performance Evaluation Metrics

The metrics chosen for comparing machine learning models are very important. In this study, several different methods were used to compare the performance of different

machine learning algorithms. Classification accuracy (ACC), sensitivity (SENS), specificity (SPEC), precision (PREC), F-score, and k-fold cross-validation are the evaluation metrics used.

k-fold cross-validation is a widely used methodology for separating training and test data. In this method, the data is divided into k subgroups and each data in the data set is used for both testing and training. The classifier is trained with k-1 subsets and tested with the remaining subset to determine performance values. This process is repeated k times and the average of the performance values obtained for each subset gives the final performance of the model. 10-fold cross-validation was used in this study. In addition, the performance of the model was compared using the holdout test. For this process, the data set was set to be 80% training and 20% testing.

ACC is one of the most widely used methods to compare the performance of machine learning models. Here, N shows the test set, cn implies the class of the value of n , $Estimate(n)$ is classification result of n , k is the k-fold validation parameter, ACC can be expressed as follows:

$$Accuracy(N) = \frac{\sum_{i=1}^{|N|} estimate(n_i)}{|N|}, \quad n_i \in N \quad (10)$$

$$Estimate(n) = \begin{cases} 1, & \text{if } estimate(n) = cn \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

$$Classification\ Accuracy(ML) = \frac{\sum_{i=1}^{|k|} Accuracy(N_i)}{|k|} \quad (12)$$

TABLE I
PARAMETERS OF ALL MACHINE LEARNING MODELS USED IN THE STUDY

	LSTM	CNN	CNN-LSTM	FCNN
1. Layer	LSTM Layer Node Count: 150 Activation: tanh	Convolutional Layer Filter Count: 16 Filter Size: 3x3 Activation: ReLU	Convolutional Layer Filter Count: 16 Filter Size: 3x3 Activation: ReLU	Fully Connected Layer Node Count: 100 Activation: RELU
2. Layer	Dropout Layer Value: 0.2	Dropout Layer Value: 0.2	Dropout Layer Value: 0.2	Fully Connected Layer Node Count: 100 Activation: RELU
3. Layer	Fully Connected Layer Node Count: 2 or 3 Activation: Softmax	Max-pooling Size: 2x2	Max-pooling Size: 2x2	Dropout Layer Value: 0.2
4. Layer		Convolutional Layer Filter Count: 16 Filter Size: 3x3 Activation: ReLU	Convolutional Layer Filter Count: 16 Filter Size: 3x3 Activation: ReLU	Fully Connected Layer Node Count: 2 or 3 Activation: Softmax
5. Layer		Dropout Layer Value: 0.2	Dropout Layer Value: 0.2	
6. Layer		Max-pooling Layer Size: 2x2	Max-pooling Layer Size: 2x2	
7. Layer		Fully Connected Layer Node Count: 100 Activation: RELU	LSTM Layer Node Count: 150 Activation: tanh	
8. Layer		Dropout Layer Value: 0.2	Dropout Layer Value: 0.2	
9. Layer		Fully Connected Layer Node Count: 2 or 3 Activation: Softmax	Fully Connected Layer Node Count: 2 or 3 Activation: Softmax	

ACC alone may not be a sufficient parameter for performance comparison. Therefore, using the terms SPEC, SENS, PREC and F-score, which are statistical performance measures, is useful in many cases for accurate performance comparison. The term SPEC refers to the proportion of correctly predicted true negatives. On the other hand, the term SENS indicates the proportion of correctly predicted true positives. The ratio of correct positive predictions is expressed as PREC. The F score is the harmonic mean of SENS and PREC values. The formulas used to calculate SPEC, SENS, PREC and F-score are as follows:

$$SPEC = \frac{TN}{TN + FP} \quad (13)$$

$$SENS = \frac{TP}{TP + FN} \quad (14)$$

$$PREC = \frac{TP}{TP + FP} \quad (15)$$

$$F\text{-score} = \frac{2TP}{2TP + FP + FN} \quad (16)$$

Here, true positive (TP) refers to the number of correctly classified positive samples and true negative (TN) refers to the number of negative samples correctly classified. On the other hand, false-positive (FP) is the number of negative samples classified as positive, and false-negative (FN): the number of positive samples classified as negative.

G. Model Parameters

Four different machine learning models were used for classification processes. Table 1 shows the parameters of the machine learning models used. The weights of the trainable

layers in all models were started with Xavier [31]. In addition, cross-entropy is used as a loss function in all models. Also, Adam optimizer was used in all models [32]. Adam is an optimization algorithm widely used in deep learning studies in recent years.

The color values of the images were normalized from [0-255] to [0-1] before they were applied to the models. All images were resized to 80x160x3 dimensions before they were applied to the models to reduce computational costs. It has been applied to CNN and CNN-LSTM models with these dimensions. All images are flattened and applied to the FCNN network. In addition, it has been reshaped to 80x480 dimensions in the LSTM model. In addition to these, the

III. RESULTS AND DISCUSSION

Soybean is an important food source for animals. However, problems such as urease activity are encountered if soybeans are not adequately processed. On the other hand, overcooking also reduces its benefits. The urease test is one of the commonly used methods to evaluate the cooking level of soybeans. According to the color change on the soybean, the cooking level of the product is evaluated by a expert. Automating the process will both minimize the risk of operator error and reduce time losses.

In this section, firstly, a binary classification process is performed on soybean images that have been tested for urease. In this context, the data were primarily evaluated as cooked and undercooked. The over-cooked and well-cooked class samples in the data set were evaluated in the same category. Classification performances were tested using four different machine learning models.

TABLE II
10-FOLD CROSS-VALIDATION RESULTS FOR BINARY CLASSIFICATION

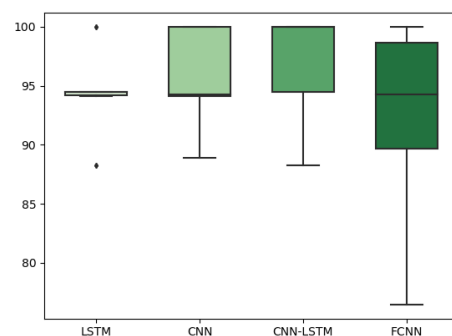
	ACC	SENS	SPEC	PREC	F-Score
LSTM	94.29	89.47	96.61	92.72	91.07
CNN	95.43	91.23	97.46	94.55	92.86
CNN-LSTM	96.57	92.98	98.31	96.36	94.64
FCNN	92.57	87.72	94.92	89.29	88.50

Table 2 shows the binary classification results obtained with 10-fold cross-validation. Here, undercooked samples were considered as positive classes and evaluation metrics were created accordingly. The best results in all evaluation metrics belong to the CNN-LSTM model. In this model, ACC, SENS, SPEC, PREC and F-Score values were obtained as 96.57%, 92.98%, 98.31%, 96.36% and 94.64%, respectively. The CNN model has an ACC value of 95.43% and an F-Score value of 92.86%. ACC and F-Score values of the LSTM model were 94.29% and 91.07%, respectively. On the other hand, the lowest performance in all evaluation metrics belongs to the FCNN model. In this model, ACC was achieved as 92.57%. As stated in the sections above, FCNNs can be applied to many problems due to their structure agnostic nature. However, their performance is generally inferior to networks adapted to solve a specific problem [21]. For this reason, its performance is considered to be lower compared to other models. On the other hand, CNN's are very successful in detecting spatial dependencies. Consequently, their performances in image-related applications are quite good. An important limitation of CNN architectures is their

image augmentation process was applied. Augmented images are used for validation purposes only. It was not used in the training process. The validation set was created by applying scaling on the images, turning the images vertically and horizontally, rotating and zooming randomly. Before each training process, images in the training data were randomly selected and augmented by the number of images in the test set. Finally, 140 images were used for training and 35 images were used for testing in the holdout test. In each part of the 10-fold cross-validation tests, 157 or 158 images were used for training and 18 or 17 images for testing. All models are trained for 200 epochs. Testing was carried out with the weights with the highest ACC value in the validation set.

weakness in learning sequential dependencies [33]. The combination of CNN and LSTM layout feature greatly improves classification [34]. Thus, the CNN-LSTM model's performance is considered to be higher than the other models. In addition, SPEC values are higher than SENS values in all models. It is considered that this is because machine learning models tend to the majority class.

In Graphic 1, box plot representation of ACC values obtained with 10-fold cross-validation in the binary classification process can be seen. It is seen that the lowest values of the FCNN model are different from other models. On the other hand, it is seen that CNN and CNN-LSTM models behave close to each other.



Graph 1. Box plot representation of 10-fold cross-validation results of the binary classification.

Table 3 shows the holdout test results for binary classification. SPEC and PREC values were obtained as 100% for all models. This means that no model mistakenly evaluates cooked grade samples as undercooked. On the other hand, models differ from each other depending on their SENS values. For other evaluation metrics, the results are very similar to the 10-fold cross-validation results. The best ACC value belongs to the CNN-LSTM model. In this model, ACC was realized at 94.29%. For the CNN-LSTM model, the F-Score value is 90.0%. The lowest ACC and F-Score values belong to the FCNN model as in the 10-fold cross-validation results. In the FCNN model, ACC and F-Score values are 82.86% and 62.50%, respectively.

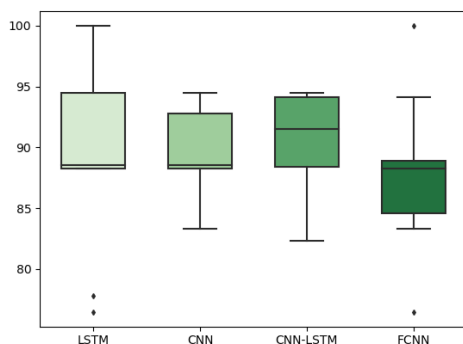
TABLE III
HOLDOUT TEST RESULTS FOR BINARY CLASSIFICATION

	ACC	SENS	SPEC	PREC	F-score
LSTM	88.57	63.64	100.00	100.00	77.78
CNN	91.43	72.73	100.00	100.00	84.21
CNN-LSTM	94.29	81.82	100.00	100.00	90.00
FCNN	82.86	45.46	100.00	100.00	62.50

In addition to binary classification, a triple classification process was carried out as over-cooked, well-cooked and undercooked. Table 4 shows the 10-fold cross-validation results of the triple classification process. Here, very close scores were obtained in all comparison metrics in CNN and LSTM models. Both models have an ACC value of 89.14%. Like the binary classification problem, the lowest scores in all evaluation metrics belong to the FCNN model. ACC and F-Score values of this model are 88.0% and 87.85%, respectively. It is evaluated here that using a model in structure agnostic structure negatively affects its performance. On the other hand, the CNN-LSTM model has the best scores in all evaluation metrics. The performance value of this model was obtained as 90.29%. Here, it is evaluated that the sequential processing of the features extracted by CNN with LSTM increases the performance. In Graphic 2, box plot representation of ACC values of triple classification tests can be seen. Here, it is seen that CNN and CNN-LSTM models exhibit similar behaviors as in the binary classification tests.

TABLE IV
10-FOLD CROSS VALIDATION RESULTS FOR THREE-CLASS CLASSIFICATION

	ACC	SENS	SPEC	PREC	F-score
LSTM	89.14	89.08	94.59	89.14	89.09
CNN	89.14	89.14	94.59	89.19	89.08
CNN-LSTM	90.29	90.24	95.14	90.46	90.29
FCNN	88.00	87.92	94.00	87.98	87.85



Graph 2. Box plot representation of 10-fold cross-validation results of the three-class classification.

Finally, in Table 5, the holdout test results of the triple classification process can be seen. The CNN-LSTM model performed better here compared to the other models as well. ACC and F-Score values of the CNN-LSTM model are 91.43% and 90.82%, respectively. The performance of the FCNN model is significantly lower than other models. FCNN model has an ACC value of 77.14% and an F-Score value of 77.09%.

TABLE V
HOLDOUT TEST RESULTS FOR THREE-CLASS CLASSIFICATION

	ACC	SENS	SPEC	PREC	F-score
LSTM	82.86	84.18	89.38	85.93	84.73
CNN	85.71	87.21	92.15	84.87	85.79
CNN-LSTM	91.43	92.09	95.50	90.05	90.82
FCNN	77.14	80.30	88.80	76.80	77.09

IV. CONCLUSION

Soybean meal has a high protein concentration. For this reason, it constitutes a very important part of the protein sources used in animal feed throughout the world. On the other hand, soybeans contain many bioactive compounds that are antinutritional and/or poisonous, such as urease. For this reason, it is desired to reduce the effect of these components by cooking with machines such as extruders. However, many factors such as steam pressure, cooking time, temperature and screw speed affect the cooking level. Less cooking of soybeans negatively affects animal health, while overcooking negatively affects nutritional values.

The urease test has been used for many years to evaluate the cooking level of soybeans. Here, the cooking level of soybeans is evaluated by an expert according to the color change of the samples. In this study, a machine learning-based approach is proposed to evaluate urease test results to automate the process and minimize possible expert errors. In this context, a data set consisting of over-cooked, well-cooked and undercooked samples were prepared. The prepared samples were tested in both two classes and three classes. In the tests performed with four different machine learning models, it was observed that the highest performance values were obtained by the CNN-LSTM model, in which the features extracted by CNN are processed sequentially with the LSTM model. In 10-fold cross-validation tests, ACC was 96.57% for two classes and 90.29% for three classes.

According to the author's knowledge, it is the first study to evaluate images of urease tests in soybeans with machine learning models. The results obtained are very promising for future studies.

REFERENCES

- [1] G. L. Cromwell, "Soybean Meal-The 'Gold Standard,'" 1999. Accessed: Apr. 25, 2021. [Online]. Available: https://www.nutritime.com.br/arquivos_internos/artigos/soybeanmeal-thegoldstandard.pdf.
- [2] R. Real-Guerra, ... F. S.-A. C. S., and 2013, "Soybean urease: over a hundred years of knowledge," *books.google.com*, Accessed: Apr. 25, 2021. [Online]. Available: <https://books.google.com/books?hl=tr&lr=&id=87WiDwAAQBAJ&oi=fnd&pg=PA317&dq=Real-Guerra,+Rafael,+Fernanda+Stanisçuaski,+and+Célia+Regina+Carlini.+%22Soybean+urease:+over+a+hundred+years+of+knowledge.%22+A+Comprehensive+Survey+of+International+Soybean+Research>.
- [3] K. Zhang, Q. Wu, and Y. Chen, "Detecting soybean leaf disease from synthetic image using multi-feature fusion faster R-CNN," *Comput. Electron. Agric.*, vol. 183, p. 106064, Apr. 2021, doi: 10.1016/j.compag.2021.106064.
- [4] Y. Ni *et al.*, "Computational model and adjustment system of header height of soybean harvesters based on soil-machine system," *Elsevier*, Accessed: Apr. 25, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0168169920331124>.
- [5] E. Clarke and J. Wiseman, "Effects of extrusion conditions on trypsin inhibitor activity of full fat soybeans and subsequent effects on their nutritional value for young broilers," *Br. Poult. Sci.*, vol. 48, no. 6, pp. 703–712, Dec. 2007, doi: 10.1080/00071660701684255.
- [6] I. E. Liener, "Implications Of Antinutritional Components In Soybean Foods," *Crit. Rev. Food Sci. Nutr.*, vol. 34, no. 1, pp. 31–67, Jan. 1994, doi: 10.1080/10408399409527649.
- [7] G. B. Huntington, D. L. Harmon, N. B. Kristensen, K. C. Hanson, and J. W. Spears, "Effects of a slow-release urea source on absorption of ammonia and endogenous production of urea by cattle," *Anim. Feed Sci. Technol.*, vol. 130, no. 3–4, pp. 225–241, Nov. 2006, doi: 10.1016/j.anifeedsci.2006.01.012.
- [8] G. Qin, E. R. Ter Elst, M. W. Bosch, and A. F. B. Van Der Poel, "Thermal processing of whole soya beans: Studies on the inactivation of antinutritional factors and effects on ileal digestibility in piglets," *Anim. Feed Sci. Technol.*, vol. 57, no. 4, pp. 313–324, Mar. 1996, doi:

- 10.1016/0377-8401(95)00863-2.
- [9] C. Luanga Ouédraogo, E. Combe, J.-P. Lallès, R. Toullec, S. Trèche, and J.-F. Grongnet, "Nutritional value of the proteins of soybeans roasted at a small-scale unit level in Africa as assessed using growing rats." Accessed: Apr. 25, 2021. [Online]. Available: https://rmd.edpsciences.org/articles/rmd/pdf/1999/02/RND_0926-5287_1999_39_2_ART0005.pdf.
- [10] S. Yalcin and A. Basman, "Effects of infrared treatment on urease, trypsin inhibitor and lipoxigenase activities of soybean samples," *Food Chem.*, vol. 169, pp. 203–210, Feb. 2015, doi: 10.1016/j.foodchem.2014.07.114.
- [11] F. S. Tabibloghmany, M. Mazaheri Tehrani, and A. Koocheki, "Optimization of the extrusion process through response surface methodology for improvement in functional and nutritional properties of soybean hull," *J. Food Sci. Technol.*, vol. 57, no. 11, pp. 4054–4064, Nov. 2020, doi: 10.1007/s13197-020-04439-w.
- [12] Y. Jing and Y. J. Chi, "Effects of twin-screw extrusion on soluble dietary fibre and physicochemical properties of soybean residue," *Food Chem.*, vol. 138, no. 2–3, pp. 884–889, Jun. 2013, doi: 10.1016/j.foodchem.2012.12.003.
- [13] N. Şenköylü, H. Akyürek, H. Ersin ŞAMLI, and A. Ağma, "Tam Yağlı Soyanın Metabolik Enerji Değerinin Broiler Performansından Tahmini," 2004. Accessed: Apr. 25, 2021. [Online]. Available: <https://dergipark.org.tr/en/pub/uluvfd/issue/13529/163661>.
- [14] K. Krishnamurthy, H. K. Khurana, J. Soojin, J. Irudayaraj, and A. Demirci, "Infrared heating in food processing: An overview," in *Comprehensive Reviews in Food Science and Food Safety*, Jan. 2008, vol. 7, no. 1, pp. 2–13, doi: 10.1111/j.1541-4337.2007.00024.x.
- [15] G. L. Grinblat, L. C. Uzal, M. G. Larese, and P. M. Granitto, "Deep learning for plant identification using vein morphological patterns," *Comput. Electron. Agric.*, vol. 127, pp. 418–424, Sep. 2016, doi: 10.1016/j.compag.2016.07.003.
- [16] R. A. Schwalbert, T. Amado, G. Corassa, L. P. Pott, P. V. V. Prasad, and I. A. Ciampitti, "Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil," *Agric. For. Meteorol.*, vol. 284, p. 107886, Apr. 2020, doi: 10.1016/j.agrformet.2019.107886.
- [17] M. Yoosefzadeh-Najafabadi, H. J. Earl, D. Tulpan, J. Sulik, and M. Eskandari, "Application of Machine Learning Algorithms in Plant Breeding: Predicting Yield From Hyperspectral Reflectance in Soybean," *Front. Plant Sci.*, vol. 11, Jan. 2021, doi: 10.3389/fpls.2020.624273.
- [18] M. Herrero-Huerta, P. Rodriguez-Gonzalvez, and K. M. Rainey, "Yield prediction by machine learning from UAS-based multi-sensor data fusion in soybean," *Springer*, doi: 10.1186/s13007-020-00620-6.
- [19] J. Zhang, Y. Huang, K. N. Reddy, and B. Wang, "Assessing crop damage from dicamba on non-dicamba-tolerant soybean by hyperspectral imaging through machine learning," *Wiley Online Libr.*, vol. 75, no. 12, pp. 3260–3272, Dec. 2019, doi: 10.1002/ps.5448.
- [20] A. D. de Medeiros, N. P. Capobianco, J. M. da Silva, L. J. da Silva, C. B. da Silva, and D. C. F. dos Santos Dias, "Interactive machine learning for soybean seed and seedling quality classification," *Sci. Rep.*, vol. 10, no. 1, p. 11267, Dec. 2020, doi: 10.1038/s41598-020-68273-y.
- [21] J. Xia, S. Pan, M. Yan, G. Cai, J. Yan, and G. Ning, "Prognostic model of small sample critical diseases based on transfer learning," *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*, vol. 37, no. 1, pp. 1–9, Feb. 2020, doi: 10.7507/1001-5515.201905074.
- [22] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.
- [23] I. OZER, "Pseudo-colored rate map representation for speech emotion recognition," *Biomed. Signal Process. Control*, vol. 66, p. 102502, Apr. 2021, doi: 10.1016/j.bspc.2021.102502.
- [24] J. Ma, F. Wu, J. Zhu, D. Xu, and D. Kong, "A pre-trained convolutional neural network based method for thyroid nodule diagnosis," *Ultrasonics*, vol. 73, pp. 221–230, Jan. 2017, doi: 10.1016/j.ultras.2016.09.011.
- [25] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, Nature Publishing Group, pp. 436–444, May 27, 2015, doi: 10.1038/nature14539.
- [26] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, 2013, doi: 10.1109/TPAMI.2013.50.
- [27] I. Ozer, Z. Ozer, and O. Findik, "Noise robust sound event classification with convolutional neural network," *Neurocomputing*, vol. 272, pp. 505–512, Jan. 2018, doi: 10.1016/j.neucom.2017.07.021.
- [28] L. Wen, X. Li, L. Gao, and Y. Zhang, "A New Convolutional Neural Network-Based Data-Driven Fault Diagnosis Method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018, doi: 10.1109/TIE.2017.2774777.
- [29] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Comput. Biol. Med.*, vol. 100, pp. 270–278, Sep. 2018, doi: 10.1016/j.compbiomed.2017.09.017.
- [30] I. Ozer, S. Efe, and H. Ozbay, "A combined deep learning application for short term load forecasting," *Alexandria Eng. J.*, vol. 60.4, pp. 3807–3818, 2021, Accessed: Apr. 25, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S111001682100137X>.
- [31] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *JMLR Workshop and Conference Proceedings*, Mar. 2010. Accessed: Apr. 25, 2021. [Online]. Available: <http://www.ijcnn.org/ijcnn/papers/2010/Glorot.pdf>.
- [32] J. B. Kingma, Diederik P., "Adam: A method for stochastic optimization," *International Conference on Learning Representations (ICLR)*, 2015. .
- [33] C. Feng, A. Mehmani, and J. Zhang, "Deep Learning-Based Real-Time Building Occupancy Detection Using AMI Data," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4490–4501, Sep. 2020, doi: 10.1109/TSG.2020.2982351.
- [34] M. Z. Islam, M. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Informatics Med. Unlocked*, vol. 20, p. 100412, Jan. 2020, doi: 10.1016/j.imu.2020.100412.

BIOGRAPHIES



İLYAS ÖZER was born in Yozgat, Turkey, in 1984. He received a B.Eng. degree from the Department of Electric & Electronic Engineering at the University of Dumlupınar of Turkey, in 2006. He received a M.Sc. degree from the Department of Computer Engineering at the University of Sakarya of Turkey in 2011. Also, He received a Ph. D. degree from the Computer Engineering at the University of Karabük of Turkey in 2018. His current research interests include signal processing, sound event detection and advanced machine learning techniques.