Estimation of Extrusion Process Parameters in Tire Manufacturing Industry using Random Forest Classifier

Osman Onur Akırmak, Aytaç Altan

Abstract—**The extrusion process is a very complex process due to the number of process parameters involved. Throughout the workflow process, the process parameters are determined by trial-and-error method according to the recipe of materials. This technique causes loss of production and time as well as energy consumption. In extrusion, temperature and speed parameters are very important to obtain a homogeneous raw material product input and high-quality extruded products. It is necessary to monitor the temperature changes and process speed control during the flow of the molten raw material between the barrels of the extruder machine, which is the extrusion equipment. By monitoring the extruder in real time, estimating the extrusion process parameters according to the amount of product to be produced will make the extrusion process operations more efficient. In this study, a classification algorithm to process these parameters is developed in the "*Pycharm*" environment and the model is trained with the supervised learning method using the image processing algorithm outputs. The model is able to estimate the extruder 'speed and temperature parameters' and the 'ready to run' decision of the machine with 93% success for different production quantities entered by the operator.**

Index Terms— **Extrusion process, tire manufacturing, decision support estimation algorithm, surface inspection, random forest algorithm.**

I. Introduction

THE EXTRUSION process is a manufacturing technique in which materials are moved along a screw and pushed out of a die at a certain speed. It is widely used in different branches in the industry, and one of the most common areas where it is used is the tire manufacturing industries. In this manufacturing process, thermoplastic material (carbon reinforced synthetic rubber-SBR) is used as raw material. It becomes the final vehicle tire by going through the component preparation, building, curing and inspection production processes, respectively. Tire faces some potential problems throughout its manufacturing processes. Unlike techniques such as injection molding or blow molding, extrusion is continuous. This means that there is a constant change in extruder parameters. Input parameters should be defined, controlled and monitored for variables that have a non-negligible impact on product quality that will disrupt the extruder's stable running process. Control and monitoring are critical in estimating production quality. The focus of this study is the monitoring of the extruder parameters and the estimation of the parameter settings in the production phase of the tire belt, which is formed by the extrusion process of several rows of cord wire, which is an input of the tire making process.

Process parameters include melting temperature, screw speed and pressure. The parameters used in this study are the melting temperature and screw speed setting values and will be used as inputs in the classification model. These parameters indicate that the melt temperature changes in the extruder are important for determining the stability of the extruder, and fluctuations in other process parameters in the process, along with it, cause problems in tire production [1]. One of these problems is that the coating cannot be done properly, the other problem is that the black rubber coating material is not distributed on cord wires homogeneously on the surface, and the color fluctuation that occurs as a cluster. An example of coating defect after extrusion is shown in Fig. 1.

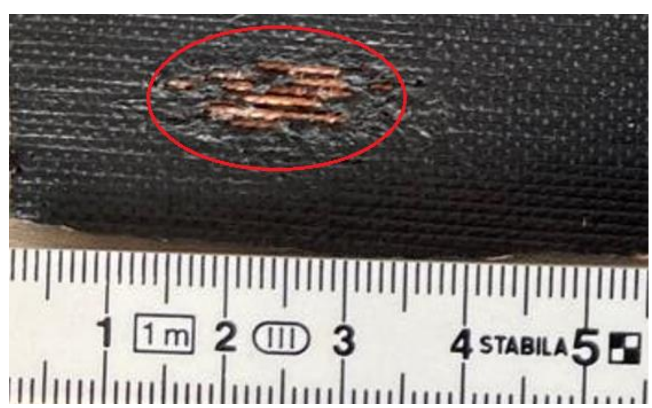

Fig.1. Coating defect after extrusion were shown with red colored marker

Some literature studies that we can pay attention to in extrusion processes regarding a certain product quality

OSMAN ONUR AKIRMAK, is with Department of Electrical Electronics Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Turkey, (e-mail: onur.akirmak@fbe.karaelmas.edu.tr).

https://orcid.org/0000-0003-0014-4680

AYTAÇ ALTAN, is with Department of Electrical Electronics Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Turkey, (e-mail: aytacaltan@beun.edu.tr).

https://orcid.org/0000-0001-7923-4528

Manuscript received January 12, 2023; accepted April 4, 2023.
DOI: 10.17694/bajece.1232811

situation; optimizing the flow rate of the plastic extrusion process, the characterization of rubber with a multilayer sensor neural network in the rubber extrusion process [2, 3], mapping neuro-fuzzy approach for the hot extrusion process using ANFIS [4], improving the quality of the final product in plastic extrusion with decision trees [5]. In order to detect bubbles and scratches, two-layer feed-forward neural network studies were carried out on the surface of aluminum profiles with co-occurrence matrices [6], temperature response control for the plastic extrusion system to avoid sudden input disturbances [7].

In the tire industry, this study has been a cost-effective and applicable study in factories with more than one extrusion line with low hardware cost, and model training achievements are indicated in the literature studies, since there is no real-time work at the hardware level in the factory environment on calendaring production process on extruder machine, it will be one of the first contributors to work actively in the field in this sector.

The aim of this study is to significantly increase the adequacy value of the process encountered as a result of the studies on surface inspection in the tire industry to develop a decision support estimation algorithm that will contribute to the production of lower cost and high-quality products by reducing the waste significantly, reducing the human factor and related errors and physical inadequacies. In summary, the method is to analyze the extrusion parameters to estimate the product quality appropriate extruder machine operating setting values with a classification model based on these parameters. This study is based on a practical applied solution with the performance of the model estimation, the product quality, and the realization of the field application.

The rest of the paper is organized as follows. Section II describes the extrusion processes of the tire factory and the techniques used to create the forecast model. Field experiments are introduced in Section III, and the experimental results obtained are discussed in Section IV. Section V highlights the results of the study and offers suggestions for future work.

II. MATERIALS AND METHODS

In this part, respectively; the extrusion process was defined, then the Random Forest Algorithm (RFA) working principle and the most important feature of this algorithm, the Gini index, were defined. In addition, the method of preparation of the data set and the characteristics and success rate of the classifier algorithm used was mentioned.

A. Extrusion Process

In the extrusion process, the raw material to be processed is fed to the extruder with the help of a hopper. The motor system on the machine is responsible for producing the energy required for propulsion. The transmission, which must be provided for the rotation of the screw produced by the motor, is done with the help of the reducer. The screw system, which rotates with the drive from the motor and reducer, drags the raw materials poured into the hive forward by means of the barrel. Here, the plastic raw material is completely melted by heaters while passing through the feeding, melting and compression zones, respectively, and the extruder machine completes its function after it is taken out of the nozzle [2]. The steps of the injection process are shown in Fig. 2 [8].

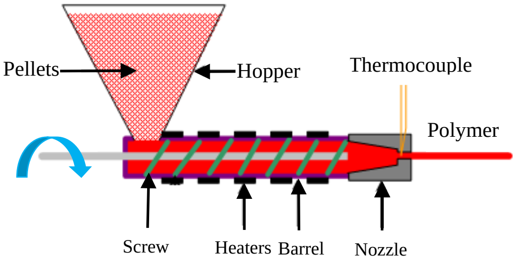

Fig.2. Extrusion process in extruder

B. Random Forest Algorithm

RFA is one of the supervised classification algorithms. It is used in both regression and classification problems. The algorithm aims to increase the classification value during the classification process by producing more than one decision tree. The RFA is the process of choosing the highest score among many decision trees that work independently of each other. The RFA determines classifications using the majority of multiple trees; so many decision trees must be built. Performance increases with the number of decision trees, requiring memory, and decreases if the number of decision trees is decreased [9].

There are two stages in the RFA, one is the random forest developer and the other is guessing through the random forest classifier responsible for the initial scenario. Pseudocode for generating RFA:

i. select the K features randomly from the total M features (must be $K < M$).

ii. calculate the d node using the best split point among the K features.

iii. split the node into child nodes using the best split.

iv. repeat steps ***i*** and ***iii*** until the L node number is reached.

v. repeat steps ***i*** and ***iv*** n times to create n number trees.

To know how a RFA works, we need to know decision trees, which are again a supervised machine learning algorithm used for classification and regression problems. Decision trees use a flowchart, such as a tree structure, to show the predictions resulting from a set of feature-based splits. It starts with a root node and ends with a decision made by the leaves are shown in Fig. 3.

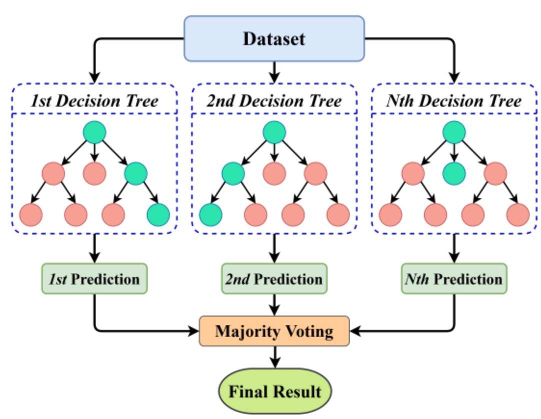

Fig.3. RFA structure

It consists of three components: root node, decision node and leaf node. The node at which the population begins to divide is called the root node. The nodes we obtain after dividing a root node are called decision nodes, and the nodes where further division is not possible are called leaf nodes. The decision of where to start the root node was made by calculating the Gini index.

C. Calculation of Gini Index

It is preferable to choose the feature with at least Gini index as the root node when constructing the decision tree. Another feature of this algorithm is that it can be used for feature selection. The feature evaluations were calculated as the mean and standard deviation of the impurity values in each tree. We defined P_i is the probability that an object will be classified into a particular class. Where P_{i+} is the probability of a positive class and P_{i-} is the probability of a negative class. When we take feature 1 as our root node, we get a pure split whereas when we take feature 2, the split is not pure. To know how much impurity this particular node has can be understood with the help of the “Gini index”. Basically, need to know the impurity of dataset and take that feature as the root node which gives the lowest impurity or say which has the lowest Gini index. Mathematically Gini index can be written as:

$$Gini\ index = 1 - [(P_{i+})^2 + (P_{i-})^2] \quad (1)$$

Suppose that try to split a new branch due to K feature which selected in screw speed window as 30 and temperature feature node value as 75℃ to create child nodes. Image processing techniques were used to classify the processing parameters as appropriate '+' and unsuitable '−' shown a sample on Table I. As seen in Table I, we have P_{i+} two good $(+)$ class element in row and one P_{i-} as bad $(-)$ in class and L number reached at once:

$$Gini\ index = 0.45 \text{ where } P_{i+} = 2/3\ , P_{i-} = 1/3 \quad (2)$$

TABLE I
DATASET STRUCTURE

Screw (rpm)	Hopper Temp. ℃	Barrel Temp. ℃	Class
30	30	69	−
30	69	85	+
30	70	83	+

Similarly, this algorithm will try to find the Gini index of all the splits possible and will choose that feature for the root node which will give the lowest Gini index. The lowest Gini index means low impurity. For the final decision, needs to calculate the weighted Gini index that is the total Gini index of this split. Final output is considered based on majority voting if it’s a classification problem each tree's classification is combined into a final classification through a "majority vote" mechanism and repeat steps are shown on pseudocode. By looking at the importance of the feature, it can be decided which features you want to reduce, low features do not contribute enough to the prediction process. This is important because a general rule in machine learning is that the more features you have, the more likely your model is to be overfitting and vice versa [10]. If a D data set is divided into two subsets, D_1 and D_2, on S, the Gini impurity index is defined as follows and the feature-dependent model classification output is given below for the first 4 nodes of the tree no. Let’s define P_- as class $+$ and P_+ as class $-$, and

$$GiniS(D) = \frac{D_1}{D} Gini(D_1) + \frac{D_2}{D} Gini(D_2) \quad (3)$$

Tree 1:
0 NODE: if feature [1] < 75.0 then next=1 else next=2
1 LEAF: return class=0
2 NODE: if feature [1] >75 then next=3 else next=4
3 LEAF: return class=1

D. Creating of the Dataset

The dataset for use in the model was obtained from the extruder on the production line of a vehicle tire manufacturing plant. The image processing device used is shown in Fig. 4. Due to the high flow rate of the line and the small surface coating defects, a profile sensor, an industrial image processing device, was positioned at the output of the extruder.

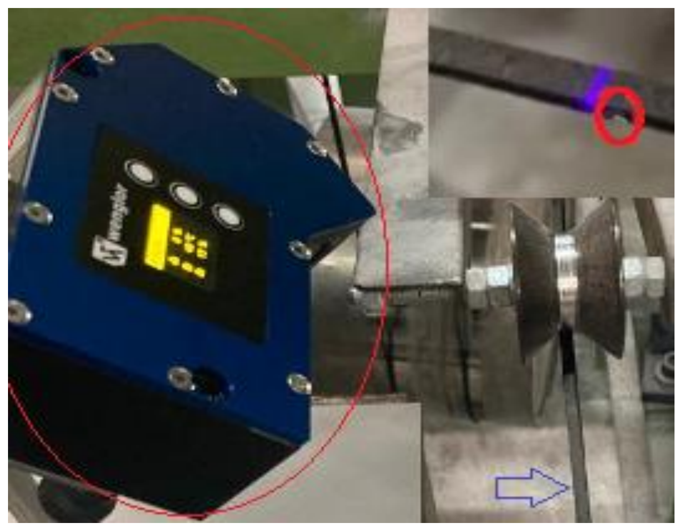

Fig.4. Profile sensor shown on left and defect detailed on top right corner, also can see blue laser light from profile sensor

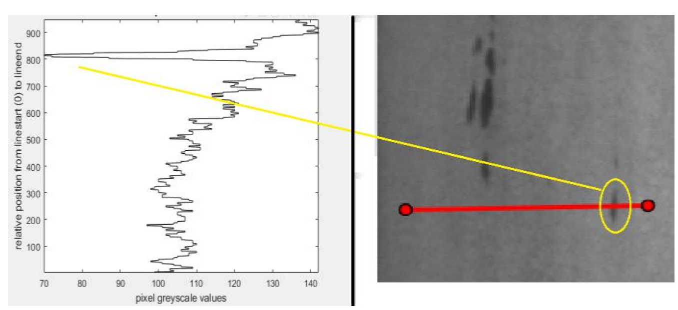

Fig.5. The image values are inverted in grayscale in the (0-255) band, showing the peak value of the laser light's reflection intensity value from the band at the defect point

While the profile sensor was being placed, different camera and laser angles were tried to be mounted at the exit of the extruder according to the layouts shown in Fig. 4. Finally, the

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 11, No. 2, April 2023

profile sensor is positioned at a distance of 80 mm from the rubber belt piece that reaches with the coating of 3 cord wires with a width of 5 mm to be controlled in the process and inclined 35 degrees horizontally. The two-dimensional images of the area scanned by the profile sensor on the belt were taken and transferred to the computer. The obtained data were analyzed by the developed software and graphics and numerical outputs were created. From the graph created, the peak value of the points with surface defects in the belt was determined and shown in Fig.5.

The temperature values read from hopper, barrel, and nozzle were taken with DS18B20 which connected to microcontroller board. The DS18B20 is a small temperature sensor with a built in 12bit analog to digital convertor (ADC). It can be easily connected to a microcontroller digital input. The sensor communicates over a one-wire bus and requires little in the way of additional components. The sensors have a quoted accuracy of $\mp 0.5^{\circ}C$ in the range $-10^{\circ}C$ to $+85^{\circ}C$. In the order of preparing the dataset, the data titled '+' and '−' taken from the profile sensor is divided into two classes under the title of "*Class"* in Table I and the temperature sensors placed on the extruder hopper, barrel, and nozzle respectively. The dataset was created by recording real-time data in *.xls* format via "*Data Streamer*" property of excel.

E. Classifier Structure

Model parameter optimization is the process of hyperparameter tuning to determine optimal values for a particular random forest classifier. The performance of a model is highly dependent on the value of the hyperparameters. Since there is no way to know the best values in advance for the hyperparameters, ideally, all possible values are determined after the relevant algorithm has been run to know the optimum values, and the model with supervised learning structure has emerged, which is configured as 20 trees and learning coefficient 0.8.

The correlation relationship between the 1369 pieces of data, the mold and hive temperature values increase in direct proportion to the production amount. Another linear relationship was also found to exist between the screw speed data and the amount of product coated as the extruder output. The confusion matrix of the model is shown in Table II. The general accuracies of the applied methods are given in the upper left and lower right corners of the matrix. In addition, the values in the upper right and lower left corners indicate incorrect estimates.

TABLE II
CONFUSION MATRIX

Positive	True Positive (TP) 397 samples	False Positive (FP) 6 samples
Negative	False Negative (FN) 2 samples	True Negative (TN) 232 samples

III. FIELD EXPERIMENTS

The system was developed for a low-budget application, and the hardware was selected accordingly. The experimental setup shown on Fig.6 block structure designed to control the mathematical model of an unknown system. The system consists of three parts, classification model, extruder and image processing system. Depending on the hardware, the limiting criterion in choosing the RFA of the classification model is that it gives higher accuracy 93% than the k-Nearest Neighbor (kNN) algorithm 88%, which is among the model algorithms that can be run on the microcontroller hardware.

Although the model results with 94% trained with the Support Vector Machine (SVM) algorithm whose parameters were optimized on a laptop with a Windows operating system in the "*Pycharm*" environment, however it was not used in the final tests due to the hardware limitation. One advantage of the RFA is that it can be used for both regression and classification tasks and it is easy to see the relative importance it gives to input properties. It is considered a very useful and easy-to-use algorithm because default hyper parameters usually produce a good prediction result. The number of hyperparameters is also not that high and is easy to understand. The main limitation of the RFA is that a large number of trees can make the algorithm slow and ineffective for real-time predictions. In general, these algorithms can be trained quickly, but are extremely slow to generate predictions once trained [11].

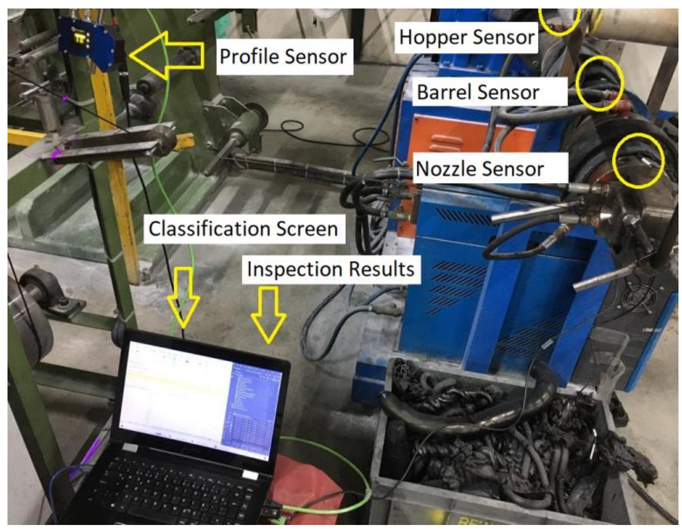

Fig.6. Experimental setup

A more accurate prediction requires more trees, resulting in a slower model. The RFA is fast enough for most real-world applications, but there may be situations where runtime performance is important and other approaches may be preferable. The RFA is a simple algorithm for early training in the model development process and to see how it performs, and it is difficult to create a "*bad*" RFA due to its simplicity. Overall, the RFA is a fast, simple and flexible tool despite its limitations. The Streaming Random Forest (SRF) algorithm handles multi-class problems successfully as opposed to many stream classification algorithms that are designed and/or tested on only two-class classification algorithms. It is fast enough to handle stream rates of up to 2.8×104 records/sec, when executed on a fairly small machine such as Pentium 4 machine with 3.2 GHz processor and 512 MB RAM [12].

The success of the classification methods used in the study are shown in Table III. All algorithms are trained with the same training size. The aim is using a low-cost hardware. Due to its small memory size, RFA algorithm was chosen as it has the highest accuracy and most feasible on the hardware. The F1-score which is machine learning evaluation metric that measures a model's accuracy was calculated. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

TABLE III
PERFORMANCES OF ALGORITHMS USED IN THE STUDY

Algorithm	Class	F1-score	Accuracy
Decision Tree	0	0.86	0.83
	1	0.78	
kNN	0	0.90	0.88
	1	0.85	
RFA	0	0.93	0.93
	1	0.92	
SVM	0	0.95	0.94
	1	0.93	

A. *Experiment 1*

According to the prepared the first experiment scenario, without model using, the extruder operator set the barrel and die temperature values as the first initial set value at 85°C and the frequency inverter controlling the screw speed as 20 Hz at the beginning of the workday shift while the system was cold, and loaded the raw material into the extruder system at ambient temperatures without preheating. The cold raw material input to the extruder, whose barrel and mold temperatures reached set values, decreased the average of the barrel and mold temperature value of the single screw extruder with a screw length of 1.5 meters and caused the first 75-meter product to be produced incorrectly. Faulty production was detected by the profile sensor system.

B. *Experiment 2*

According to the prepared the second experiment scenario, the extruder operator set the barrel and die temperature values at 85°C at the beginning of the shift, the frequency inverter controlling the screw speed as 20 Hz, while the system was cold, and preheated the raw material by circulating the product. Meanwhile, the model is running. The operator only manually entered the desired production amount as 42 kg/hour. Nozzle, barrel and hopper temperature information is given to the model in real time by the microcontroller hardware.

The model works on the microcontroller. When the ready information of the extruder is shown on the touch screen, 10 Hz appears to be sufficient according to the "*rpm*" estimation on the same screen, the operator is asked to adjust the extruder speed, the temperature values, the raw material, barrel, and hopper temperatures 60, 75, and 79°C as soon as the "*Machine Ready*" information is received. The operator accordingly reduced the heater set value from 85°C to 80°C and terminated the circulation and sent the product to the mold end with the wire for coating. Meanwhile, the image processing system checked the error status of the coated product and no faulty product production was observed. The flow chart of the system is shown on Fig.7.

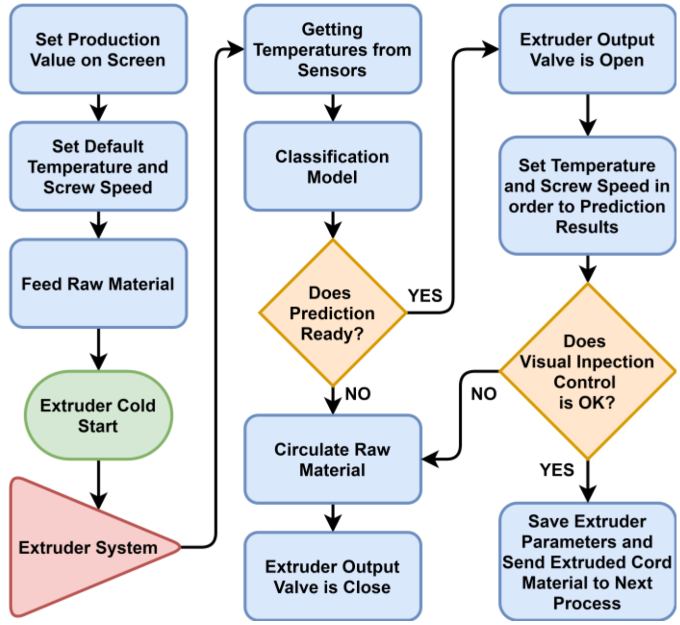

Fig.7. Flow chart of the system

IV. Experimental Results and Discussions

As seen in the results of Experiment 2, it is possible to set the equipment temperature set values below 85°C in order to produce faultless products. Again, when the results were evaluated, different screw speed rpm values were estimated in order to produce the desired 42 kg/min product amount. By connecting a wattmeter, the instantaneous power was recorded periodically every 5 minutes, and then after 1 hour, the motor consumption was calculated by finding how much power was with an arithmetic average. Here, setting the temperature value as 80°C and setting the inverter speed value as 10 Hz decreased the hourly energy consumption by 25-30% respectively, while the electricity total consumption amount per two workday shifts was 56 Turkish Liras in the previous setting. In addition, since the amount of production counted as faulty waste in the Experiment 1 will decrease/can disappear, an improvement will be achieved in Overall Equipment Efficiency (OEE) calculations in the company's production processes and in quality problems with the real-time image processing system used. Extrusion quality and product quantity at the extruder outlet are directly related to the raw material temperature in the machine feed, machine sleeve temperatures and screw speed data. Accurate estimation of the extruder optimum operating data range increases efficiency in work and energy, and reduces waste.

Although we obtained the result that we achieved an energy gain with a low screw speed in the Experiment 2, we saw that there are reasons that limit our screw speed. It is not possible to reduce to much the speed of the tire winding machine which is connected end of the line and working with the extruder's output material. If we reduce the speed of the tire winding machine, there will be a time-related loss of production. An unpredictable output of this study is that when we set the speed of the tire winding machine to 65 m/min instead of between 25 to 45 m/min at Experiment 2, it is seen in the image processing software output that there is not enough coating on the product which faulty. Extruder output speed cannot reach this production speed; therefore, the production line speed balance is not properly planned, however an optimum balance can be found with our proposed model. This is necessary for line balancing and stable production planning.

As can be seen from this study, the model can predict the extrusion process parameters with an accuracy of 93%. In future studies, the results can be improved by using different classification algorithms. Experimental hardware has been selected from inexpensive and open-source software available in the market, open-source software, and appropriate hardware. The possibilities of using the neural network in the extrusion process are endless because it is also used for the control of the entire system. When the neural network controller is combined with an extruder, it is sufficient to select the appropriate control input for the plant behavior that can be optimized, the future performance of the system will become predictable. The use of advanced artificial neural networks can eliminate the loss of production due to the individual, the need for learning by trial and error, which is a negative outcome of the process, and improvement can be achieved with more advanced parameter settings by making error classification.

V. Conclusions and Future Works

It can be observed from this study that the use of an classifiers with a low-cost hardware can accurately predict the extrusion process parameters. This can improve the output quality and increase the production rate of tire raw materials. Production workers in tire industries can be equipped with the appropriate tools which can enable them to produce quality tire material while eradicating the need to perform long experiments which can lead to waste of materials and increase the cost of production. The prospects of utilizing the artificial neural network in near future instead of using classifier, will be used for the decision support of the entire system. The neural network controller integrated with an extruder, which enables it to be able to predict future plant behaviors and select appropriate control input can optimize future performance.

References

[1] V. García, J.S. Sánchez, L.A. Rodríguez-Picón, L.C. Méndez-González, H.D.J. Ochoa-Domínguez. "Using regression models for predicting the product quality in a tubing extrusion process." Journal of Intelligent Manufacturing, vol. 30, 6, 2019, pp. 2535-2544.

[2] C.Y. Wu, Y.C. Hsu. "Optimal shape design of an extrusion die using polynomial networks and genetic algorithms." The International Journal of Advanced Manufacturing Technology, vol. 19, 2, 2002, pp. 79-87.

[3] S.A. Oke, A.O. Johnson, O.E. Charles-Owaba, F.A. Oyawale, I.O. Popoola. "A neuro-fuzzy linguistic approach in optimizing the flow rate of a plastic extruder process." International Journal of Science & Technology, vol. 1, 2, 2006, pp. 115-123.

[4] R.S. Sharma, V. Upadhyay, K.H. Raj. "Neuro-fuzzy modeling of hot extrusion process." Indian Journal of Engineering & Materials Sciences, vol. 16, 2009, pp. 86-92.

[5] S.H. Hsiang, Y.W. Lin, J.W. Lai. "Application of fuzzy-based Taguchi method to the optimization of extrusion of magnesium alloy bicycle carriers." Journal of Intelligent Manufacturing, vol. 23, 3, 2012, pp. 629-638.

[6] A. Chondronasios, I. Popov, I. Jordanov. "Feature selection for surface defect classification of extruded aluminum profiles." The International Journal of Advanced Manufacturing Technology, vol. 83, 1, 2016, pp. 33-41.

[7] S. Ravi, P.A. Balakrishnan. "Temperature response control of plastic extrusion plant using MATLAB/Simulink." International J. of Recent Trends in Engineering and Technology, vol. 3, 4, 2010, pp. 135-140.

[8] URL: https://www.substech.com/dokuwiki/doku.php?id=extrusion_of polymers (Access: Jan 9, 2023).

[9] Y. Mishina, R. Murata, Y. Yamauchi, T. Yamashita, H.H. Fujiyoshi. "Boosted random forest." IEICE Transactions on Information and Systems, vol. 98, 9, 2015, pp. 1630-1636.

[10] URL: https://web.stanford.edu/class/archive/cs/cs221/cs221.1186/lectures/learning3.pdf (Access: Jan 5, 2023).

[11] G. Leshem, Y.A. Ritov. "Traffic flow prediction using Adaboost algorithm with random forests as a weak learner." International Journal of Mathematical and Computational Sciences, vol. 1, 1, 2007, pp. 1-6.

[12] H. Abdulsalam. "Streaming random forests." Ph.D. Thesis, Queen's University, Canada, July 2008.

Biographies

OSMAN ONUR AKIRMAK received his B.Sc. degrees in the department of Electrical Electronics and Computer Engineering in 2015 and 2016, respectively. He is studying his M.Sc. degree in the department of Electrical Electronics Engineering from Zonguldak Bülent Ecevit University since 2021. He is currently a Project Engineer at Wenglor Sensoric GmbH. His research interests including software development of sensors and image processing solutions for digitalization of factories.

AYTAÇ ALTAN received his B.Sc. and M.Sc. degrees in the department of Electrical Electronics Engineering from Anadolu University in 2004 and 2006, respectively. He is received his Ph.D. degree in the department of Electrical Electronics Engineering from Zonguldak Bülent Ecevit University in 2018. He is currently an Associate Professor at the department of Electrical Electronics Engineering at the Zonguldak Bülent Ecevit University in Turkey. His research interests include signal processing, image processing, optimization techniques, artificial intelligence, data mining, system identification, model-based control, and robotic systems.