ICIC Express Letters
Part B: Applications
Volume **14**, Number **1**, January **2023**

ICIC International ©2023 ISSN 2185-2766
pp. **99–107**

CONVOLUTION NEURAL NETWORK BASED PREDICTION OF A QUALITY INDICATOR IN MIXING PROCESSES OF RUBBER PRODUCTS

Sujeong Son and Kwangyeol Ryu*

Industrial Data Science and Engineering, Department of Industrial Engineering
Pusan National University
63, Busandaehak-ro, Geumjeong-gu, Busan 46241, Korea
handsj@pusan.ac.kr; *Corresponding author: kyryu@pusan.ac.kr

Received May 2022; accepted July 2022

Abstract. *During the process for manufacturing rubber products, 'mixing' is a process of mixing by adding a raw material and a blending agent to a mixer. Most companies use the tacit knowledge of skilled workers to determine the order and time of input of materials, which hinders the uniformity of process reliability and quality. To systematize this, this study proposes a model that can predict rubber properties measured in the inspection process, the last process, in time series data, which are measurement data of sensors such as temperature, voltage, and RAM opening and closing of the mixer measured by the company. A model using CNN is constructed by grasping the characteristics of time series data, which is 1 dimension data. In the model study, in order to increase the prediction accuracy, a ResNet, which is advantageous for multi-layer stacking of convolution layers, was built, but there was a problem that the validation error was not lowered. Therefore, in this study, a model was developed to inversely reduce the number of layers, which is a one-dimensional Convolution Neural Network (CNN), and Symmetric-padding, which improves the problems of Zero-padding, was applied. The application of this method could improve the learning speed of the model combining CNN and ResNet34 and reduce the validation error by about 4%. Through the research results of this paper, it is expected that the basic research results on other approaches of the quality indicator prediction model of rubber mixing processes and similar processes can be provided.*

Keywords: CNN, Mixing process, Quality indicator prediction, Time series data, Rubber manufacturing, ResNet, Symmetric-padding

1. **Introduction.** The production of rubber products varies depending on the recipe of the product to be produced. The quality of rubber products is also affected by the environment, such as the amount of blending agent, the temperature and voltage of the mixing machine, and the weather around the factory. Therefore, it is difficult to obtain the physical properties required by the final product, and the process of inspecting the intermediate product is necessary. Even now, it is common in production sites to determine the input time and order of materials by the tacit knowledge of skilled workers. However, as interest in smart factories increases with the Industry 4.0, attempts to change from the existing work method are gradually increasing. Among them, the application of the methodology applying machine learning promotes process automation and is advantageous in discovering specific trends and patterns that are difficult for workers to do. Therefore, this study intends to develop a model that predicts inspection data with data from sensors generated during the process by applying machine learning. There are various analyses using time series data [1-3]. So this study develops the model that predicts inspection results by using sensor data occurring during the process. However, most of

DOI: 10.24507/icicelb.14.01.99

the previously performed rubber characteristic prediction models are studies focusing on rubber lifespan.

As a study dealing with the prediction of rubber physical properties, Jeong et al. [4] proposed a model to predict the properties of rubber compounding materials using an Artificial Neural Network (ANN). This study was able to directly apply the amount of raw material and the physical properties of rubber compounding material to the analysis model without synthesizing the actual material. However, in this paper study, it is difficult to use their model because it is necessary to make predictions using only data on changes in material input time and temperature within the same recipe. And since they used basic ANNs, it is difficult to apply to time series data of various channels in terms of accuracy. An example of using time series data in the manufacturing process is a study on the detection of abnormalities in the mold cylinder temperature cycle using the 1D Convolutional Neural Network (CNN) of Yu et al. [5]. This is a study that improved 3% more than the basic ANN accuracy by using CNN's effective pattern characteristic extraction ability. However, the purpose of this paper study is different in that it is not a prediction but a study on anomaly pattern detection.

Allamy and Koerich [6] applied 1D Residual CNN for continuous data to classify music genres by learning audio datasets. Their results showed 80.93% in average accuracy, which was higher than that of other models compared to other 1D CNN architectures. Based on this paper, we propose a ResNet model using time series data in one dimension.

Therefore, in this paper, we propose a model that applies Residual Neural Network (ResNet), which is advantageous for multilayer stacking of convolutional layers, based on 1D CNN. And we propose a model that applies Symmetric-padding, which is shallower and simpler than deeply stacking the model, and improves the problems of Zero-padding. Through the research results of this paper, it is expected that the basic research results on the development of a quality indicator prediction model for kneading and similar processes of rubber products can be provided.

The background and related studies of this study are presented in Section 2. Section 3 describes the mixing process, the quality indicators to be predicted, and the data considered in this paper. Section 4 presents two models for predicting quality indicators. Section 5 describes the results and analysis of the experiment, and finally Section 6 concludes this paper and discusses possible future works.

2. **Literature Review.**

2.1. **Convolution Neural Network (CNN).**

CNN is an example of a CNN structure that is mainly used in analyzing images or image data [7]. Unlike conventional ANNs, spatial information is considered, so the number of parameters decreases and the amount of computation decreases. In the convolutional layer, the characteristics of the image are extracted through an operation that slides the characteristic map [8]. In order to preserve the dimensions of the input data in the output data, a process called padding is used for data reduction. Zero-padding refers to the process of symmetrically adding zeroes to the input matrix. Wu et al. applied Symmetric-padding to eliminating shift problem occurring in even-sized kernel convolutions. The result showed that Symmetric-padding achieved similar accuracy to a new compact model, which used much less memory and time during training [9]. In the pooling layer, the size of the feature map can be adjusted, and usually max pooling or average pooling is used [10]. The fully connected layer is called an FC layer, and flattens information so that it can produce the desired output by entering it.

CNN shows good performance for image classification problems. However, if the CNN layer is stacked above a certain level, there is a problem that the performance is rather poor [7]. A typical CNN model mainly used for image analysis is used for data of a two-dimensional structure. CNNs used in one-dimensional structures are called 1D CNN models and are mainly used for natural language processing or time series analysis. Here, the one-dimensional means that the kernel for composite product and the order of data to be applied have a one-dimensional shape.

2.2. **Residual Neural Network (ResNet).** ResNet is a model that can effectively learn deep layers on the 152nd floor by learning the residual presentation function [12]. As the basic CNN deepens, there is a problem that the gradient vanishes or explodes. In order to find the gradient of the neural network, the differentiation of the weighting loss function is obtained by back-propagation, in which the partial derivative of the activation function is obtained and multiplied by its value. In this process, if a small differential is multiplied several times, it approaches zero, and if a large differential is multiplied several times, it becomes very large. However, the ResNet model uses shortcut connection to transfer the input of the previous layer to the next layer (Figure 1). This shortcut connection enables deep neural network construction by avoiding gradient vanishing or exploding problems. In addition, the proposed ResNet is able to produce smooth output results from noise filtered images [13].

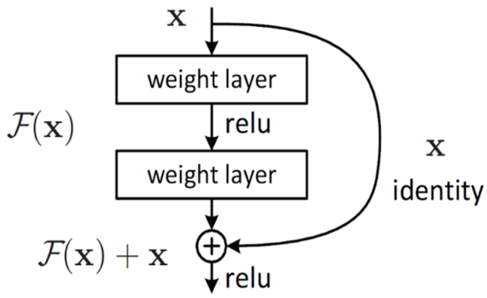

Figure 1. Shortcut connection of ResNet [7]

3. Mixing Process, Prediction of Quality Indicator, and Data.

3.1. **Rubber product manufacturing process and mixing process.** The basic process flow for producing rubber products proceeds with raw material procurement, material mixing, calendering (surface treatment), molding, vulcanization, and cutting (Figure 2). Although vulcanization may be omitted depending on the product to be produced, the mixing process is a common process. In the quality inspection for the final product, it is verified whether the rubber has properties to function properly, and whether the size, thickness, and shape are well molded. The mixing process is a part directly connected to physical properties, and whether it has been uniformly mixed is an important factor in quality. As illustrated in Figure 2, the mixing process proceeds with five detailed processes, such as mixing materials and blending agents, mixing them with a mixer, and cooling them. After that, it is divided by batch and physical property tests are performed. If there are no unusual outliers when checking the test results, the follow-up process is carried out.

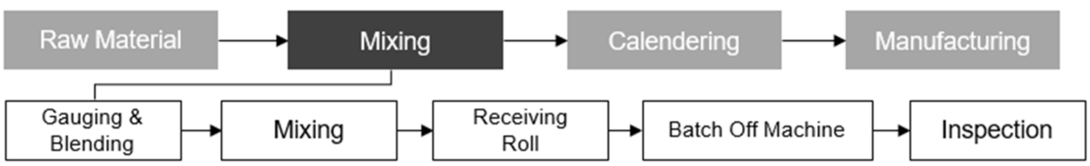

Figure 2. The order of the rubber product process

3.2. **Quality measurement indicator.** Tensile strength refers to the maximum stress until the material is broken by a tensile load. Tensile test is the most common industrial tests conducted to obtain basic data on material strength, such as Proportional Limit, Elastic Limit, Young's Modulus, Fracture Stress and Poisson's Ratio [11]. Moony viscosity is the most basic index for evaluating the physical properties of uncured rubber, indicating the degree of viscosity [14]. A rheometer usually refers to equipment for measuring a vulcanization reaction of rubber and is also referred to as a vulcanization tester. This shows whether the rubber is uniformly kneaded well and whether vulcanization can be performed well when performing the post-process vulcanization [15].

The indicators to be analyzed in this paper consist of simple codes and include M1T10, M1T90, M1ML, M1MH, and D2MN. The first alphabet of the code represents the manufacturer of the physical property measuring instrument. The second number is the order of the measuring instrument. Among the following alphabets, T10 and T90 represent the time it takes for the stress of tensile strength to reach a point corresponding to 10% and 90%. And ML and MH represent a minimum torque value and a maximum torque value of tensile strength stress. MN represents the pattern viscosity of rubber.

3.3. **The definition and purpose of the question.** These indicators are quality inspections of intermediate goods to move on to subsequent processes, not finished products, and are aimed at measuring good products and defects that determine the availability of delivery immediately, but at measuring physical properties to see if they are mixed well. Currently, intermediate products are sampled and physical properties are tested to determine the good or defect of the product. To simplify this measurement process, we develop a model that predicts the physical properties of rubber with the sensing data of the mixer. If physical properties can be predicted based on the mixing process state data described in Section 3.4, the physical properties of the product can be predicted with sensing data, so samples of physical properties can be selected small or the tests can be omitted. And you will be able to preemptively find defects in the product.

3.4. **Used data.** The data to be used for learning uses the temperature, voltage, and RAM opening and closing states obtained through the sensor of the mixer. In addition, additional data that can be used include the amount of carbon input, temperature and humidity at the factory location, and month. Figure 3 shows a graph showing the temperature, voltage, and RAM states of the mixer over time. RAM can be said to be the

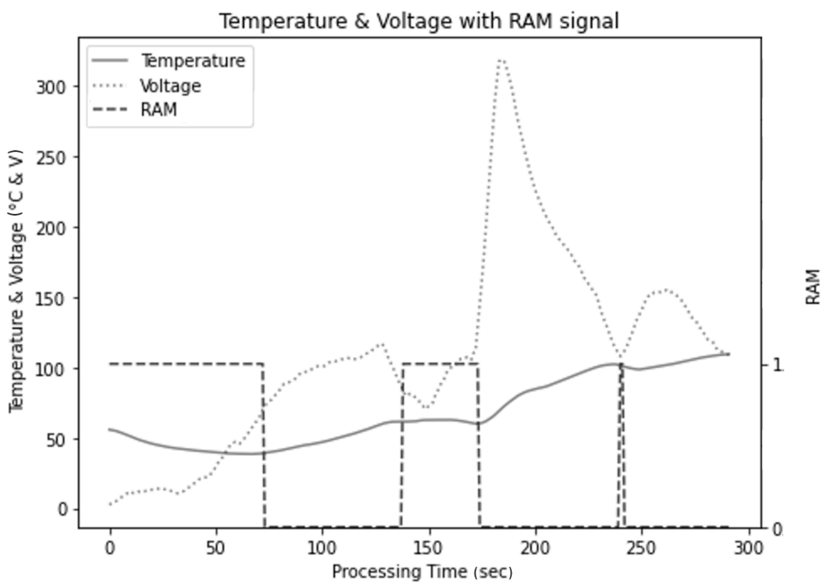

FIGURE 3. The state of the mixer by time series

cap of the mixer. The opening and closing of the RAM may include opening and closing for material input and opening and closing for internal pressure drop. The RAM has a discrete value of 1 in the open state and 0 in the closed state. One recipe randomly selected from among the recipes of various rubber products was used for learning, and a total of 358 time series data were used. Each data has an uneven size from a minimum of 262 to a maximum of 382.

The data were divided into training set 80%, validation set 15%, and test set 5%, and used for each purpose. In order to remove outliers before learning, the label, a quality indicator of data, was removed from the learning target if it became 1.5 times or more than the IQR (Inter-Quarter Range). In order to normalize the size of each variable, a min-max scaling technique was used to convert the number between the minimum and maximum values from 0 to 1. Finally, the time series length is increased to a constant value in consideration of the characteristics of the CNN that needs to be input with a predetermined size. The remaining values above the given data are proceeded with Zero-padding to fill 0 and Symmetric-padding to fill the value symmetrically. The length of the extended time series is greater than the maximum value of 382 and is set to 400 for convenience in calculation.

4. Model for Predicting Mixing Process Quality Indicators.

4.1. Changing parameters for model application. Prior to building the model, we can discuss the application of Symmetric-padding. Input data has time series characteristics and can be easily converted into 1 dimension data. However, each data has a different length of time series. We would like to build a model in two types of padding that can fill this length. First, the length of the data is adjusted with Zero-padding. Second, the length of the data is adjusted by symmetrically filling the end of the data with Symmetric-padding (Figure 4). Other parameters will be carried out the same. The model will be conducted in two ways to check the results of a total of four models that applied different padding.

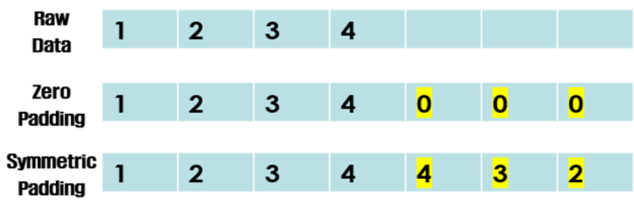

FIGURE 4. Feature of Zero-padding and Symmetric-padding

4.2. ResNet-42 model. As a result of step-by-step addition of convolution layers in the basic ResNet 34 model, the best-performing model was the ResNet-42 model. As shown in Table 1, the ResNet 42 model adds 8 convolutional layers to the ResNet34 structure to outline the structure of the 42-layer model. M1T10, M1T90, M1ML, M1MH, and D2MN values are output by receiving temperature, process time, and month information of the mixer as input. The convolutional layer of this model has parameters of kernel size = 5 and padding = 2. Among them, two models were conducted: a Zero-padding model and a Symmetric-padding model.

When the calculation is completed in each convolutional layer, Batch Normalization is performed and then the activation function ReLU is passed. Batch Normalization is effective in solving the problem of efficient learning and staying at Local Optical [16]. After passing through the convolution layer on the 41st floor, the number of output variables is adjusted to five to predict quality indicators in the last FC layer.

TABLE 1. Architecture for ResNet34 with the blocks stacked. Down-sampling is performed by conv3_1, conv4_1, and conv5_1 with a stride of 2.

Layer name	34-layer	42-layer (modified)
conv1	7×7, 64, stride 2	5, 64, stride 1
conv2_x	3×3 max pool, stride 2	3 max pool, stride 2
	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 5, 64 \\ 5, 64 \end{bmatrix} \times 3$
conv3_x	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 5, 128 \\ 5, 128 \end{bmatrix} \times 4$
conv4_x	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 5, 256 \\ 5, 256 \end{bmatrix} \times 6$
conv5_x	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 5, 512 \\ 5, 512 \end{bmatrix} \times 4$
conv6_x	no layer	$\begin{bmatrix} 5, 1024 \\ 5, 1024 \end{bmatrix} \times 3$
	average pool, 1000-d fc, softmax	

4.3. **3-layer model.** The 3-layer model is a shallow model consisting of 3 layers. It has the same parameters as the model to which ResNet is applied and two convolution operations are performed. Batch Normalization was not performed, and ReLU was used as the activation function. Five predicted values are output by the last FC layer. The convolution layer of this model has parameters of kernel size = 5 and padding = 2. As in the previous model, double parameter padding proceeds in two ways: Zero-padding model and Symmetric-padding model.

5. **Experimental Results and Analysis.** The experimental environment of this study was executed on the Jupyter Notebook of the Linux environment. Python and Pytorch were used for network configuration. As a hyper parameter, the learning rate was set to 1*1e-4, the batch size was set to 8, and the number of epochs was set to 240. Adam was used as the optimizer and Root Mean Square Error (RMSE) was used as the loss function. Figure 5 shows the experimental results of ResNet-42 models. Finally, validation's RMSE did not reach 0.15. It is possible to confirm a relatively poorer RMSE than the 3-layer model. Symmetric-padding also could not confirm the difference in the results. Considering why ResNet-42 model outputs worse RMSE, CNN has the property of extracting overall features as the layer deepens. The input data is similar in overall appearance, even though it is temperature data of different batches in Figure 6. Therefore, the validation loss is maintained and only the training loss has increased. The deeper the layer, the better the result was sometimes not produced. Since the layer was relatively deeper than the 3-layer model, the advantage of Symmetric-padding, which can make use of the characteristics of the end of the data, was not highlighted.

Figure 7 shows that even for the 3-layer model, overfitting can be found that the validation loss is maintained and only the training loss increases. However, it can be seen that the RMSE of training and validation set converged below 0.15. Simple models showed about 15% reduction in RMSE in validation set compared to ResNet. There are two main assumptions that the shallow simple model performed better than the deep ResNet model.

First, it is the similarity of the overall data. As the temperature continues to rise, it has a similar overall shape, so it can be thought that sufficient prediction was possible even with a shallow layer. In this case, the focus should be on the characteristics of the part

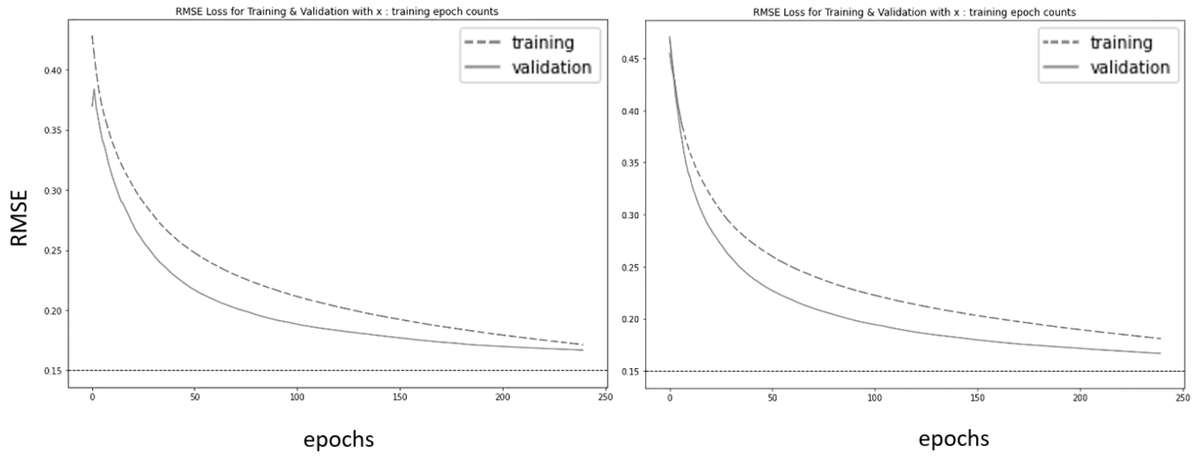

Figure 5. Results of ResNet model (Zero-padding on the left, Symmetric-padding on the right)

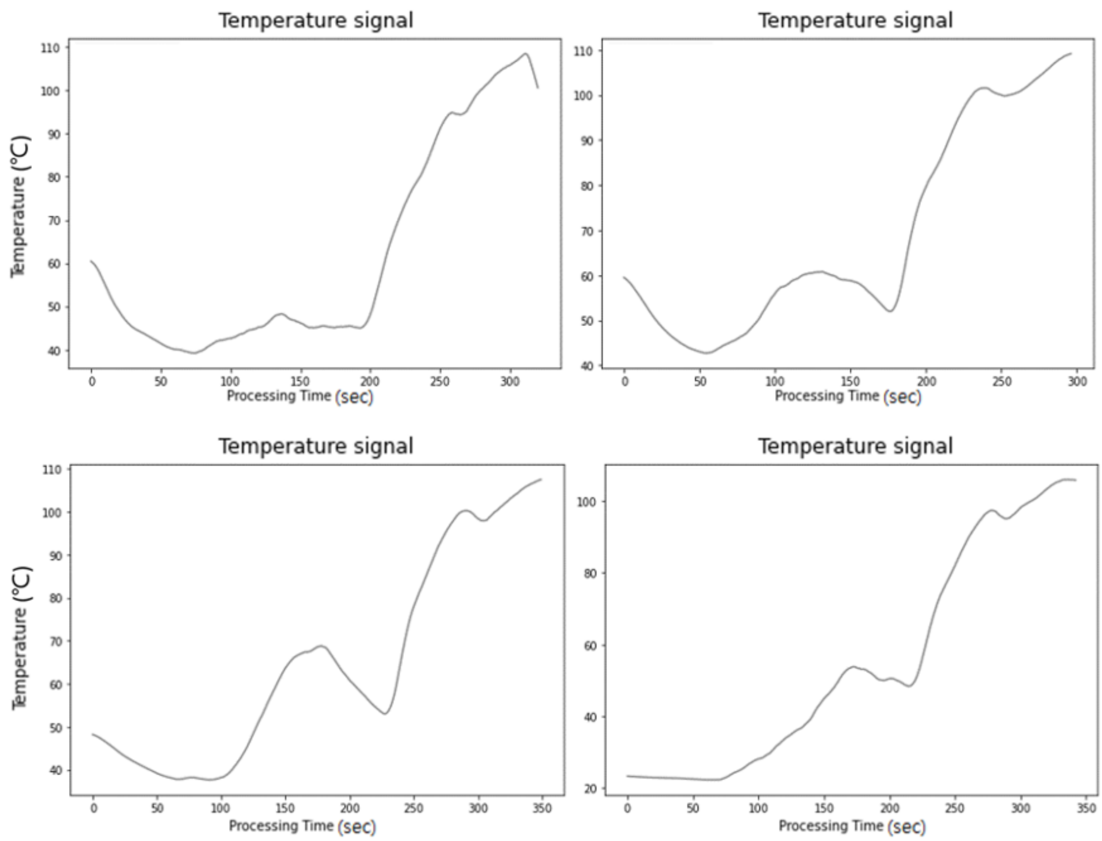

Figure 6. Temperature data of random different batches

rather than the whole. The second reason is the problem of the Zero-padding method of data preprocessing. It was judged that the method of filling with zero to match the length of the data could reduce the accuracy of the prediction by putting meaningless values in the last part. So, instead of Zero-padding, Symmetric-padding was used once more to improve the performance.

6. **Conclusions.** Unlike previous study [4], which used conventional ANNs to predict physical properties of rubber, this study proposed applying 1D CNNs to physical properties prediction. It is possible to play a positive role in productivity and profitability by efficiently inspecting intermediate products using the rubber physical property prediction

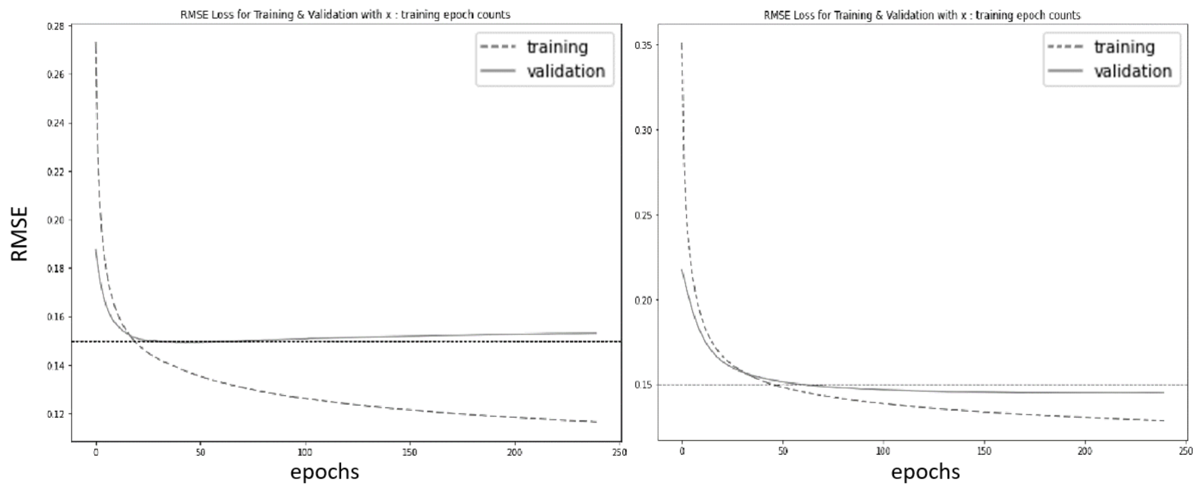

FIGURE 7. Results of 3-layer model (Zero-padding on the left, Symmetric-padding on the right)

model. The model proposed in this paper is significant in that it discussed the reasons why it is difficult to use deep ResNet for time series data and attempted a direction to improve it. In this study, a 1D CNN-based ResNet-42 model and a simple model consisting of three layers were constructed, and Symmetric-padding was used instead of Zero-padding. The reason why ResNet with a deep layer performs worse than a simple model is presumed to be due to the specificity of time series data with different input lengths. Symmetric-padding seemed to solve some of these problems, but it is not a fundamental solution. Therefore, the prediction accuracy will be improved by applying a model that combines RNN regardless of the input length of the data and CNN showing high accuracy, or a methodology that can re-scale the length of time series data with different lengths later.

Acknowledgment. This work is partially supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A2C2009 984). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] D. Cheng, F. Yang, S. Xiang and J. Liu, Financial time series forecasting with multi-modality graph neural network, *Pattern Recognition*, vol.121, 108218, 2022.

[2] Z. Yue, S. Ding, L. Zhao, Y. Zhang, Z. Cao, M. Tanveer, A. Jolfaei and X. Zheng, Privacy-preserving time-series medical images analysis using a hybrid deep learning framework, *ACM Transactions on Internet Technology (TOIT)*, vol.21, no.3, pp.1-21, 2021.

[3] Z. Yang, J. Keung, M. A. Kabir, X. Yu, Y. Tang, M. Zhang and S. Feng, AComNN: Attention enhanced compound neural network for financial time-series forecasting with cross-regional features, *Applied Soft Computing*, vol.111, 107649, 2021.

[4] N. Jeong, K. Baek, J. Cho, B. Yoon, M. Kim, J. Nam, J. Seo and M. Seo, Prediction of rubber compounding material properties using machine learning, *The 2017 Fall Conference of the Korean Society of Automotive Engineers (KSAE)*, pp.1236-1237, 2017.

[5] G. Yu, H. Yang and J. Wang, Abnormal detection of the mold cylinder temperature cycle using 1D CNN, *The 2019 Korean Institute of Industrial Engineers Joint Spring Conference*, pp.5873-5878, 2021.

[6] S. Allamy and A. L. Koerich, 1D CNN architectures for music genre classification, *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp.1-7, 2021.

[7] S. Albawi, T. Mohammed and S. Al-Zawi, Understanding of a convolutional neural network, *International Conference on Engineering and Technology (ICET)*, pp.1-6, 2017.

[8] A. Krizhevsky, I. Sutskever and E. H. Geoffrey, ImageNet classification with deep convolutional neural networks, *Communications of the ACM*, vol.60, no.6, pp.84-90, 2017.

[9] S. Wu, G. Wang, P. Tang, F. Chen and L. Shi, Convolution with even-sized kernels and symmetric padding, *Advances in Neural Information Processing Systems*, vol.32, 2019.
[10] S. Han, *A Study on the Standardization of QFP Pull Test with Lead-Free Solder*, Master Thesis, The Mechanical Engineering, Chung-Ang University, 2006.
[11] E. Jang, Introducing the rubber test machine, *Rubber Technology*, vol.20, pp.78-93, 2019.
[12] K. He, X. Zhang, S. Ren and J. Sun, Deep residual learning for image recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.770-778, 2016.
[13] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, *The 32nd International Conference on Machine Learning (ICML)*, pp.448-456, 2015.
[14] Z. Guo, Y. Sun, M. Jian and X. Zhang, Deep residual network with sparse feedback for image restoration, *Applied Sciences*, vol.8, no.12, 2018.
[15] D. Yu, H. Wang, P. Chen and Z. Wei, Mixed pooling for convolutional neural networks, in *Rough Sets and Knowledge Technology. RSKT 2014. Lecture Notes in Computer Science*, D. Miao, W. Pedrycz, D. Ślęzak, G. Peters, Q. Hu and R. Wang (eds.), Springer, Cham, 2014.
[16] S. Lee and J. Bae, Effect of processing additives on vulcanization and properties of EPDM rubber, *Journal of the Oil & Applied Science*, vol.35, no.1, pp.173-185, 2018.