A Study on Acoustic Signal Characterization for Al and Steel Machining by Audio Deep Learning

Tae-won Kim[*], Young Min Lee[**] and Hae-Woon Choi[***],[#]

[*]Graduate School of Mechanical Engineering, Keimyung Univ. Daegu, Korea
[**]Dept. of Robotics and Mechanical Engineering, Korea Polytechnics Univ. Kyungpook, Korea
[***]Dept. of Mechanical Engineering, Keimyung Univ. Daegu, Korea

오디오 딥러닝을 활용한 Al, Steel 소재의 절삭 깊이에 따른 오디오 판별

김태원*, 이영민**, 최해운***,#

*계명대학교 대학원 기계공학과, **한국폴리텍대학 로봇캠퍼스 로봇기계과, ***계명대학교 기계공학전공

ABSTRACT

This study reports on the experiment of using deep learning algorithms to determine the machining process of aluminium and steel. A face cutting milling tool was used for machining and the cutting speed was set between 3 and 4 mm/s. Both materials were machined with a depth to 0.5mm and 1.0mm. To demonstrate the developed deep learning algorithm, simulation experiments were performed using the VGGish algorithm in MATLAB toolbox. Downcutting was used to cut aluminum and steel as a machining process for high quality and precise learning. As a result of learning algorithms using audio data, 61%-99% accuracy was obtained in four categories: Al 0.5mm, Al 1.0mm, Steel 0.5mm and Steel 1.0mm. Audio discrimination using deep learning is derived as a probabilistic result.

Key Words : Deep Learning(딥러닝), Machining(기계 가공), Audio Discrimination(오디오 판별)

1. Introduction

In the 4th industrial era, deep learning or artificial intelligence has been applied to many industrial applications. In particular, research on neural networks similar to the human brain has been applied to various technologies from the year of 1940 to the present. In beginning of 2006, DBN(Deep Belief Network) and RBM(Restricted Boltzmann Machine) methods that can be trained even in multi-layer structures have begun to emerge[1~2].

Recently, many leading IT companies which are providing voice-based services through out smart phones, TV, SNS or portal sites are adopting voice-audio control methods.

Corresponding Author : hwchoi@kmu.ac.kr
  Tel: +82-53-580-5216, Fax: +82-53-580-6067

Tae-won Kim, Young Min Lee and Hae-Woon Choi
: Journal of the Korean Society of Manufacturing Process Engineers, Vol. 20, No. 7
===

In addition, many research and development in deep learning is being conducted not only in speech recognition but also in image recognition. Convolutional Neural Network (CNN)-based AlexNet was introduced in ILSVRC in 2012 in image classification technology, and deep learning-based research has been underway in various fields[3]. Deep learning is also being applied in various fields such as image classification, information search, and audio verification or recognition.

According to prior studies, researches were conducted on the development of automatic volume control technology based on audio content analysis using deep learning. A method has been reported that converts the input signal of audio content into Mel-Spectrogram and transmits it to a CNN-based deep learning network for analysis and results[4].

In this study, we applied audio deep learning technology to identify the material and machining depth for aluminium and steel. When machining the materials with different depth of cut, the generated acoustic signals were used as raw data for the audio deep learning. The generated data during machining is transferred to a pre-trained network to learn aluminium and steel materials according to the cutting depth.

In this paper, the cutting force signal applied to the cutting tool is detected in the cutting process, and the abnormal condition causing tool breakage is diagnosed through signal analysis through wavelet transformation, and the possibility of applicability to other machining signals as well as cutting is suggested.

2. Experiments

2.1 Experimental setup and equipment

In order to proceed with deep learning using a milling machine, 66061-T6 and S45C material were selected for aluminium and steel specimens, respectively. The specimens were prepared as
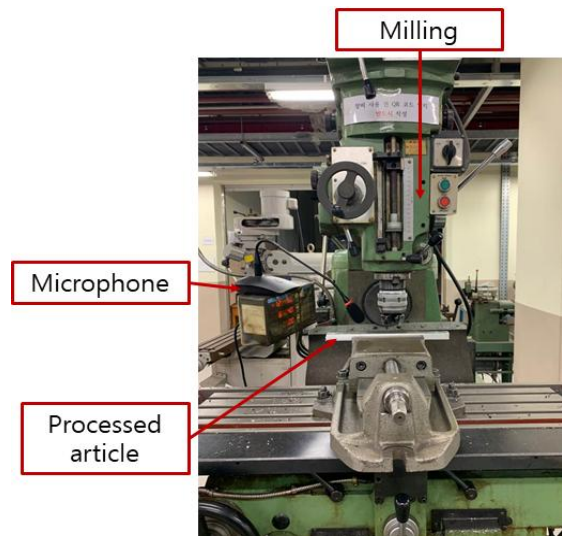

(a) Milling machine

(b) aluminium and steel specimen

Fig. 1 Experiment setup

dimension of 25mm wide and 300mm long. In the experiment, the cutting depth was set to 0.5mm and 1.0mm and the generated audio signal was recorded according to the each machining schedule or cutting depth.

As shown in Fig. 1, a full scale milling machine(HMTH-1100N, Hwacheon Machinery Corp.'s) was used for the experiment and audio data acquisition was conducted by using a high fidelity microphone(CM-7010, PILLAR Inc.).

Since aluminium material is softer than steel, the
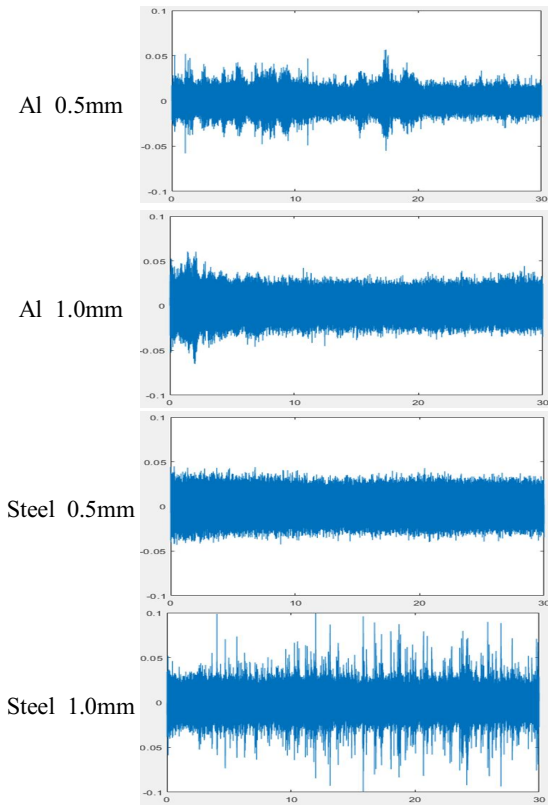
Fig. 2 Audio waveform (time scale = second)
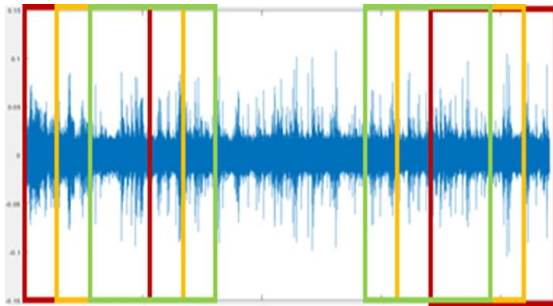
Fig. 3 Audio data preprocessing

feed rate for each material was set to 4mm/s for aluminium and 3mm/s for steel. The audio recording time was set to 1 min. for aluminium and 1.5 min. for the steel.

Figure 2 shows the audio generated when processing for each condition as a waveform graph.

While x-axis represents machining time, the y-axis represents amplitude. The intermittent disturbance signal in the waveform is a waveform that is generated by scattering chips from the workpiece and hitting the microphone during processing.

Based on the collected data, deep learning was performed using MATLAB and VGGish algorithms. For the transfer learning using pre-trained algorithms, we calculated the values using deep learning toolbox and parallel computing toolbox in MATLAB and CUDA tool kit and GPU processor(RTX-2070, NVIDIA Inc..

As shown in Fig. 3, the audio recording files for 4 cases were divided into 3 sec. of intervals, and the audio channel was set to Mono with the sampling frequency of 44.1kHz.

2.2 Deep learning algorithm structure

In order to proceed with audio deep learning data, transfer learning of the pre-trained CNN structure VGGish was used. As widely used in deep learning, CNN structures are used for image classification, audio detection and classification. Especially, VGGish is a modified VGG model of a CNN (Convolutional Neural Network) structure and adopted in the experiment.

As explained in Fig. 4 VGGish has a Mel-Spectrogram input data size or array was set as 94 x 64. As described in the example of Fig. 4, it has 24 layers for the deep learning. The layers are composed of 9 layers of training, 6 convolution layers and 3 fully connected layers[5~7].
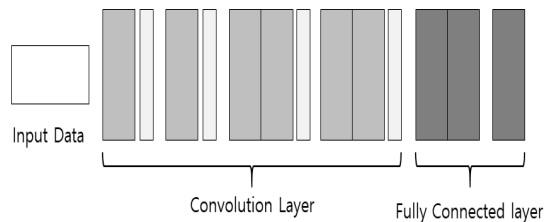
Fig. 4 VGGish algorithm

2.3 Audio deep learning proces

A total of 11,813 audio data were acquired, and 80% of the total data were classified as training data and 20% as validation data.

The optimized activation function for the deep learning was Adam function and max. Epoch was set to 30 times. The size of the mini batch was set as 128 and the initial learning rate was set to 0.001.

The deep learning was carried out by decreasing the learning rate every 10 epochs. In deep learning, three experiments were conducted to discriminate audio signal. Audio data were collected according to the cutting depth and material classification. In the first experiment, deep learning was carried out by dividing the data into 9,450(80%) training data and 2,363(20%) validation data out of 11,813 pieces of data. As shown in Table 1, it took 1 hour, 45 minutes and 21 seconds, and the accuracy was 97.48%.

Figure 5 shows the progress of deep learning. The x-axis is the time for deep learning, and the y-axis represents the accuracy when the algorithm is trained and verified.

In the second experiment, an experiment was conducted to discriminate Al and Steel using audio processed with the same cutting depth. 4,504 (80%) of the 5,630 audio data processed by 0.5mm and
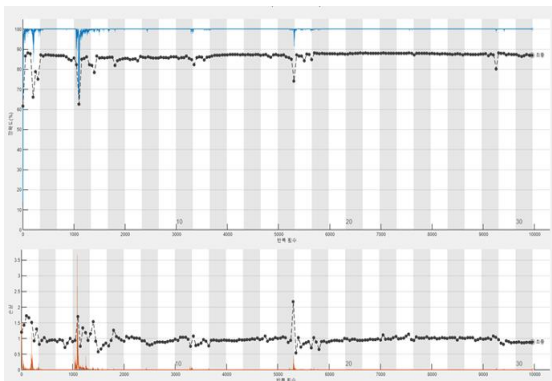

Fig. 5 Deep learning process

Table 1 Deep learning process running time

Epoch	Iteration	Time since start	Training accuracy	Base learn rate
1	550	00:03:28	24.60%	0.001
5	2900	00:17:41	97.86%	0.001
10	5800	00:35:14	97.92%	0.001
15	8700	00:52:42	100.00%	0.0005
20	11600	01:10:12	98.44%	0.0005
25	14500	01:27:40	95.30%	0.0003
30	17430	01:45:21	97.48%	0.0003

1.0mm of aluminum were divided into 1,126 (20%) training data. As same manner 4,946 (80%) was used for training and 1,237(20%)were used for the validation data for the steel.

The third experiment was conducted to discriminate audio that was processed into 0.5mm and 1.0mm of aluminium and steel, respectively. Total of 7,308 audio data were processed with aluminium and steel 0.5mm cases. The collected data, total of 4,872(80%) were classified as training and 2,436(20%) were classified as validation data.

3. Results

3.1 Classification of audio data by deep learning

Deep learning was carried out by the collected audio data in aluminium and steel machining with cutting depths of 0.5mm and 1.0mm. Each audio was pre-processed with a 96 x 64 Mel-Spectrogram, followed by deep learning. The test results are summarized in Fig. 6 to Fig. 8.

As shown in Fig. 6, a result of the first experiment is summarized. The x-axis is the categories of aluminium 0.5mm, 1.0mm, Steel 0.5mm, and 1.0mm and the y-axis is the probability of each category.

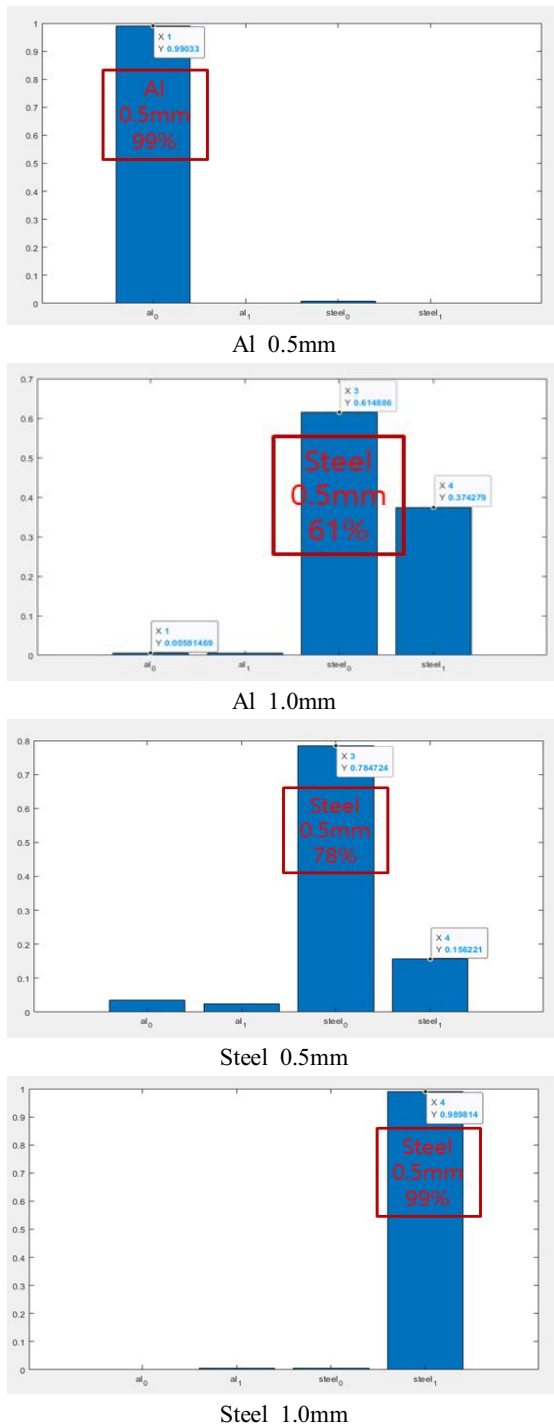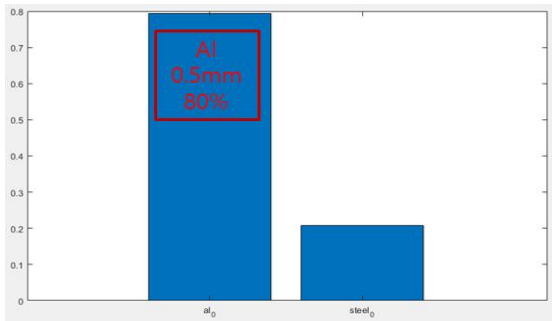The test was conducted using the data of each

A Study on Acoustic Signal Characterization for Al and Steel Machining by Audio Deep Learning
: Journal of the Korean Society of Manufacturing Process Engineers, Vol. 20, No. 7
═══

Al 0.5mm

Al 1.0mm

Steel 0.5mm

Steel 1.0mm

Fig. 6 Deep learning result (summary)

condition in the trained network. As a result, it can be seen that the audio processed by 0.5mm aluminium is discriminated as the audio processed for 0.5mm cutting depth with a 99% probability. However, in the case of audio processed with aluminium of 1.0mm cutting depth, there was a 61% probability that it was judged to be audio processed with steel 0.5mm. The audio processed by 0.5mm and 1.0mm steel was judged as an audio signal processed by steel 0.5mm and 1.0mm with 78% and 99% probability, respectively.

The results of the second experiment is summarized in Fig. 7 where the x-axis is the category of audio cutting Al and Steel to the same depth. The y-axis shows the probability of each category. As a result of testing audio processed by aluminium and steel with 0.5mm on the trained network. As shown in Fig. 7, the probability is 80% and 99% for the aluminium test case.

When testing aluminium and steel with 1.0mm, the result is that the audio processed with aluminium processed with 1.0mm has a 97% probability, and the audio processed with Steel with 1.0mm has a 99% probability of being Steel. Finally, the result of the third experiment is described in Fig. 8. The x-axis of Fig. 8 is set in the categories of Al 0.5mm, 1.0mm, and steel 0.5mm, 1.0mm with the same material but different cutting depths. The y-axis shows the test probabilities of each category.

As a result of the test, when the audio data processed by 0.5mm and 1.0mm of aluminium was discriminated in the trained network, it was classified into the categories of aluminium 0.5mm and aluminium 1.0mm with 93% and 92% probability, respectively.
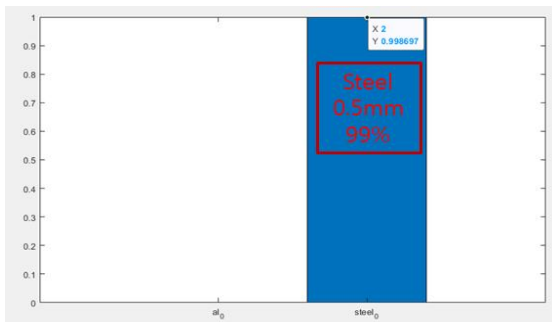
In the same way, when steel audios with different cutting depths were discriminated, each category was identified with a probability of 63% for audio processed with 0.5mm and 89% for audio processed with 1.0mm.
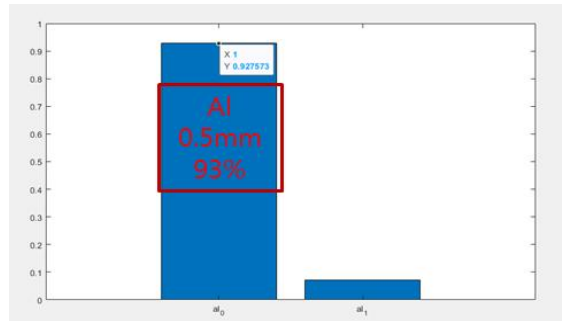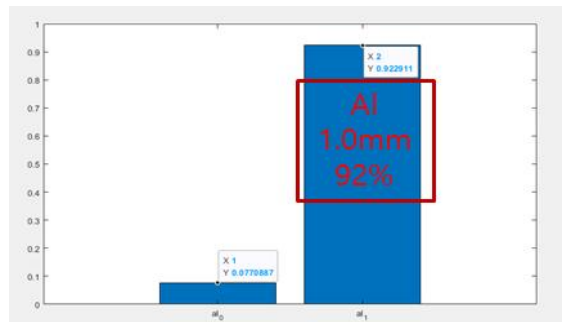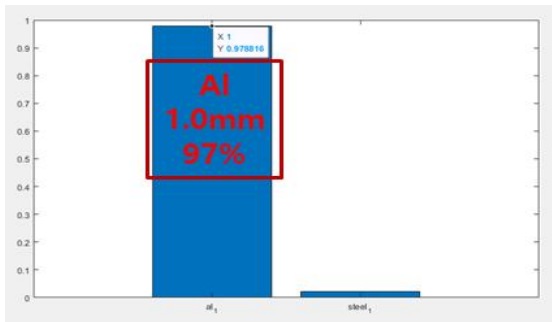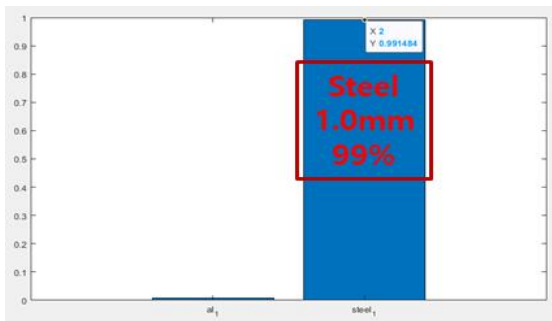
Al 0.5mm

Steel 0.5mm

Al 1.0mm

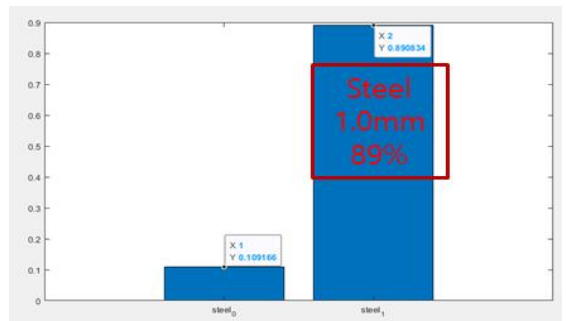
Steel 1.0mm

Fig. 7 DL result (same depth, different mat'l)

Al 0.5mm

Al 1.0mm

Steel 0.5mm

Steel 1.0mm

Fig. 8 DL result (same mat'l, different depth)

4. Conclusion

In this paper, we conducted audio deep learning using audio data generated during machining for the different material and cutting depth. The study on how to determine the type of material and the depth of machining by the audio data was successfully demonstrated with the developed deep learning process. The developed program used transfer learning with VGGish algorithm with verification accuracy of 97.48%.

A total of 3 experiments were carried out for the different material and the cutting depth. The best results were found at 80% of training data and 20% of validation data. As a result of the test, the probability of discriminating between the audio processed by aluminium of 1.0mm thick and steel of 0.5mm thick is lower than that of other categories. This reason is considered to be caused by the fact that the sampling rate of the data processed by aluminium of 1.0mm thick and the data processed by steel of 0.5mm thick when the waveform is checked is not clear. For accurate deep learning, it is judged that the more high-quality data with accurate features for each category, the more the network accuracy increases.

Also, it can be seen that the reason for the low result of deep learning is that the waveform of the audio rises rapidly due to the chips scattered during the process of recording audio while cutting.

5. Discussion

Through this study, it was found that information on material and cutting depth using audio signals can be obtained by using artificial intelligence. It is expected that by expanding the research, it is possible to detect the signal of the cutting force applied to the cutting tool, and to diagnose the abnormal condition in which the tool breakage occurs through signal analysis through wavelet transformation. It is also expected that additional data of audio data may result in better accuracy for this research.

Acknowledgement

REFERENCES

1. Lee, Y. H., "Speech/Auido Processing based on Deep Learning," Journal of Broadcasting and Media Magazine, Vol. 22, No. 1, pp. 46-57, 2017.

2. Hinton, G., et al., "Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups," Journal of IEEE Signal Processing Magazine, Vol. 29, No. 6, pp. 82-97, 2012.

3. Krizhevsky, A., Sutskever, I. and Hinton, G. E., "Imagenet classification with deep convolutional neural networks," Journal of Communications of the ACM, Vol. 60, No. 6, pp. 84-90, 2017.

4. Lee, Y. H., Cho, C. S. and Kim, J. W., "Development of Automative Loudness Control Tecgnique based on Audio Contents Analysis using Deep Learning," Journal of Broadcasting and Media Magazine, pp. 42-43, 2018.

5. Hershey, S., Chaudhuri, S., Ellis, D., Gemmeke, J., Jansen, A. and Moore, R., "CNN architectures for large-scale audio classification," In Acoustics, Speech and Signal Processing (ICASSP), pp. 131-135, 2017.

6. Suh, S., Lim, W., Jeong, Y., Lee, T., & Kim, H. Y. "Dual CNN Structured Sound Event Detection Algorithm Based on Real Life Acoustic Dataset," Journal of Broadcast Engineering, Vol. 23, No. 6, pp. 855‒865, 2018.

==

7. Lee, S., Kim, G., Choi, S. "A Machine Learning Program for Impact Fracture Analysis," The Korean Society of Manufacturing Process Engineers, Vol. 20, No. 1, pp. 95-102, 2021.