Flaw Detection in Metal Additive Manufacturing Using Deep Learned Acoustic Features

Wentai Zhang*, Brandon Abranovic*, Jacob Hanson-Regalado, Can Koz, Bhavya Duvyrui, Kenji Shimada, Jack Beuth, Levent Burak Kara

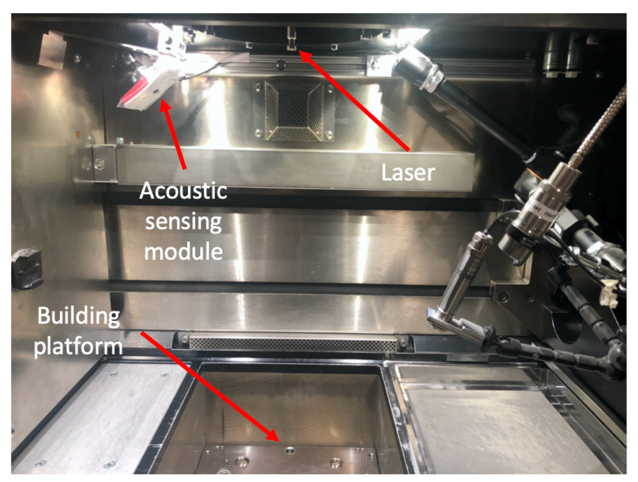
* Equal contributions.

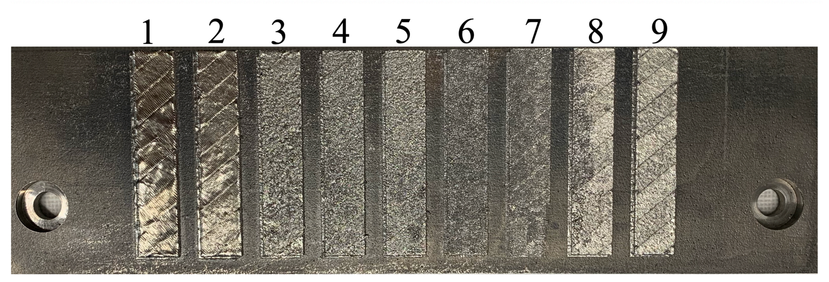
Overview

Various methods exist to monitor the quality of components built in metal additive manufacturing. We propose a novel pipeline for training two deep learning flaw formation detection techniques including convolutional neural networks and long short-term memory networks. Both approaches have yielded a classification accuracy over 99% on unseen test sets.

Our main contributions are:

- An efficient data generation method resulting in a large balanced dataset
- A novel, potentially in-situ, laser powder bed monitoring platform
- A convolutional neural network and a long short term memory neural network trained to classify the flaw types with high accuracy

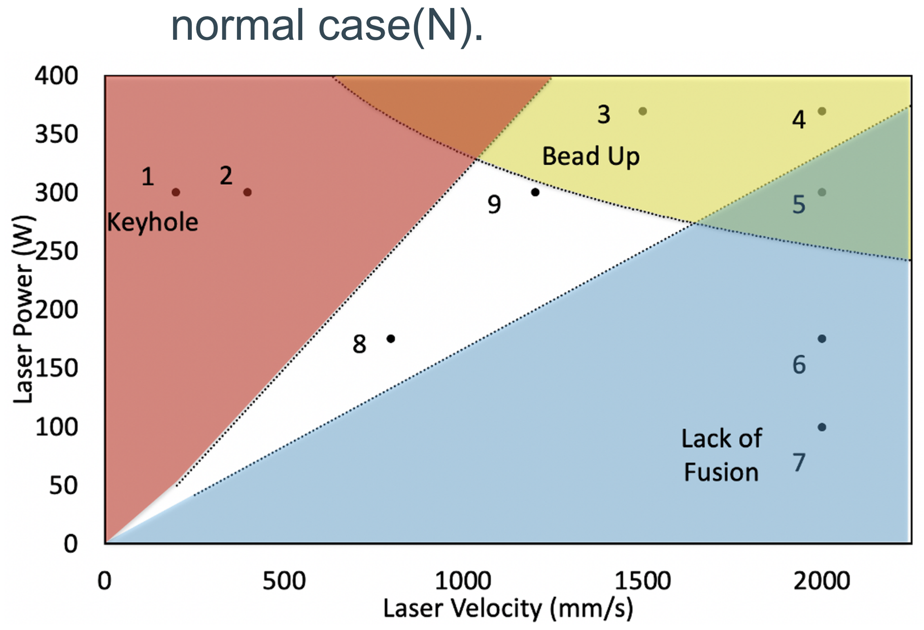




9 build samples in our experiments. Keyhole: 1 & 2. Bead Up: 3, 4, & 5. Lack of Fusion: 6 & 7. Normal: 8 & 9.

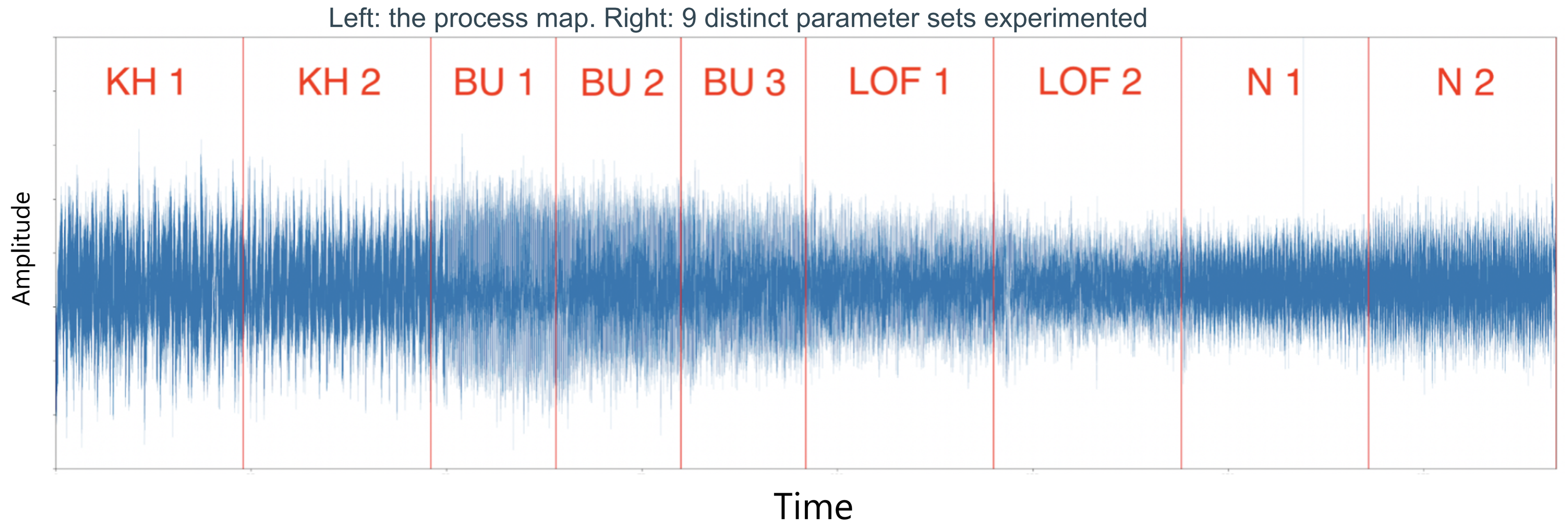
Data Collection

- Audio data was recorded by mounting a microphone inside of a machine during the manufacturing of a simple geometry.
- Using the method of process mapping, we periodically varied the laser power and velocity to create 9 distinct parameter sets.
- These 9 parameter sets corresponded to 3 flaw formation mechanisms, including Keyhole(KH), Bead Up(BU) and Lack of Fusion(LOF), and 1 normal case(N).



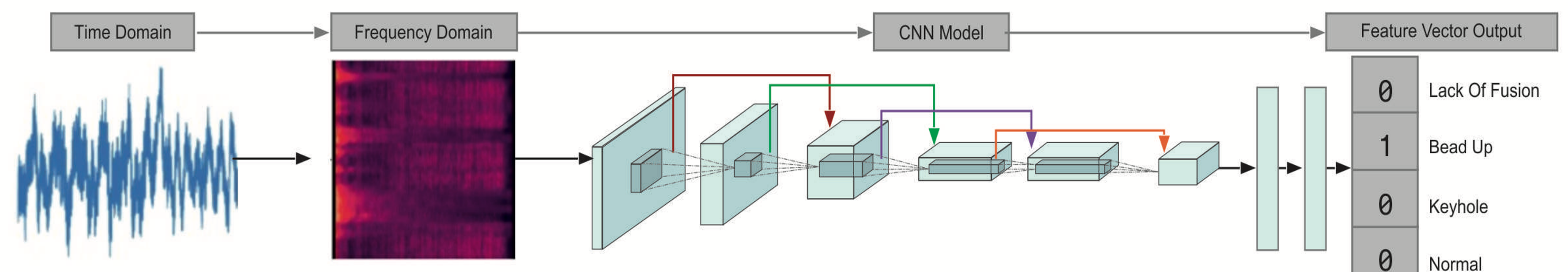
Index	Power (W)	Velocity (mm/s)	Class
1	300	200	Keyhole
2	300	400	Keyhole
3	370	1500	Bead Up
4	370	2000	Bead up
5	300	2000	Bead up
6	175	2000	Lack of Fusion
7	100	2000	Lack of Fusion
8	175	800	Normal
9	300	1200	Normal

Left: the process map. Right: 9 distinct parameter sets experimented



CNN Approach

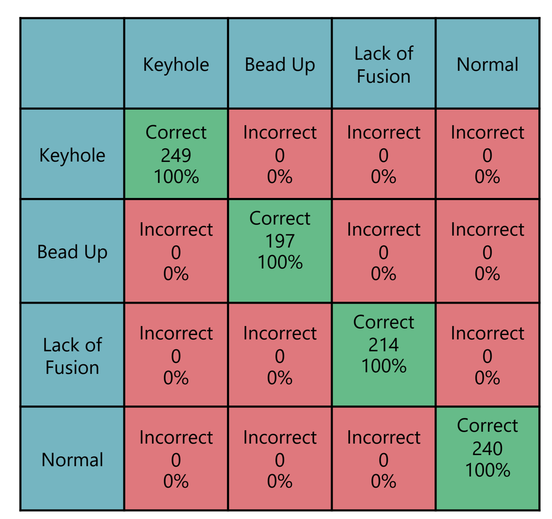
Transform the acoustic signals into spectrograms and analyze using ResNet18, ResNet34 and DenseNet

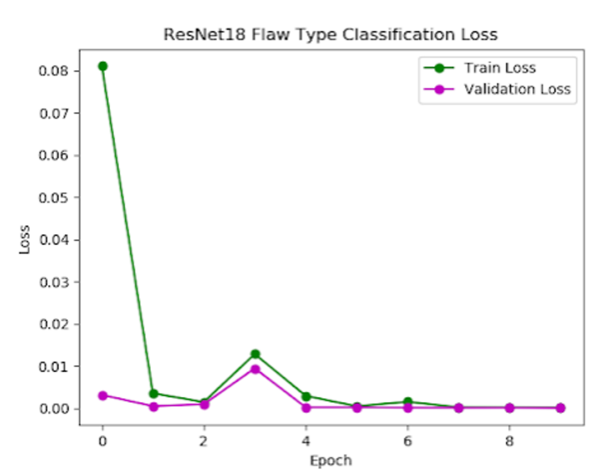


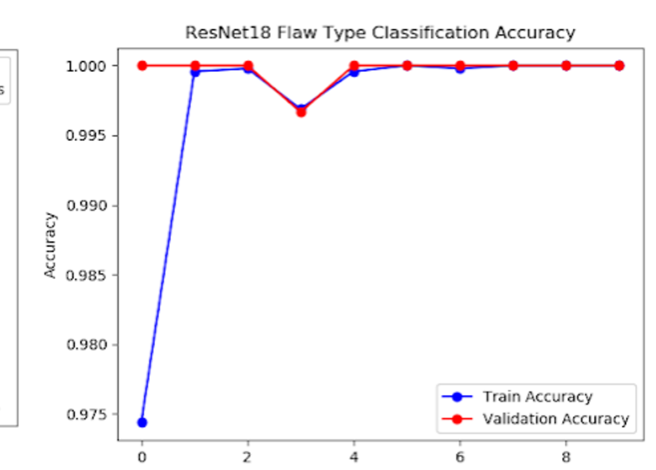
A schematic graph of our proposed pipeline

Our first approach utilized a Convolutional Neural Network to classify ~2 second segments from the clipped data. Each segment was converted into a spectrogram, a visual representation of audio signals that plots the intensity of different frequencies over time. Each spectrogram was normalized, resized, and saved as an image before being passed into our model and classified.

We experimented with various well-known CNN architectures, such as the Residual Neural Network1 (ResNet 18 & 34), and the Dense Neural Network2 (DenseNet121). All models showed strong results in classification, so we used ResNet18 as our final model due to its smaller size. All models were previously trained on ImageNet to decrease training time.







Left: the confusion matrix of our best model(ResNet18). Mid: Loss curves. Right: Accuracy curves

Ground Truth Class	Average Probability	Standard Deviation
Keyhole	0.99997	0.00005
Bead Up	0.99964	0.004
Lack of Fusion	0.99999	0.00001
Normal	0.99998	0.00003

Model	Average Prediction Speed (seconds)	Number of Trainable Parameters
ResNet18	0.004	11689512
ResNet34	0.007	21797672
DenseNet121	0.02	7978856

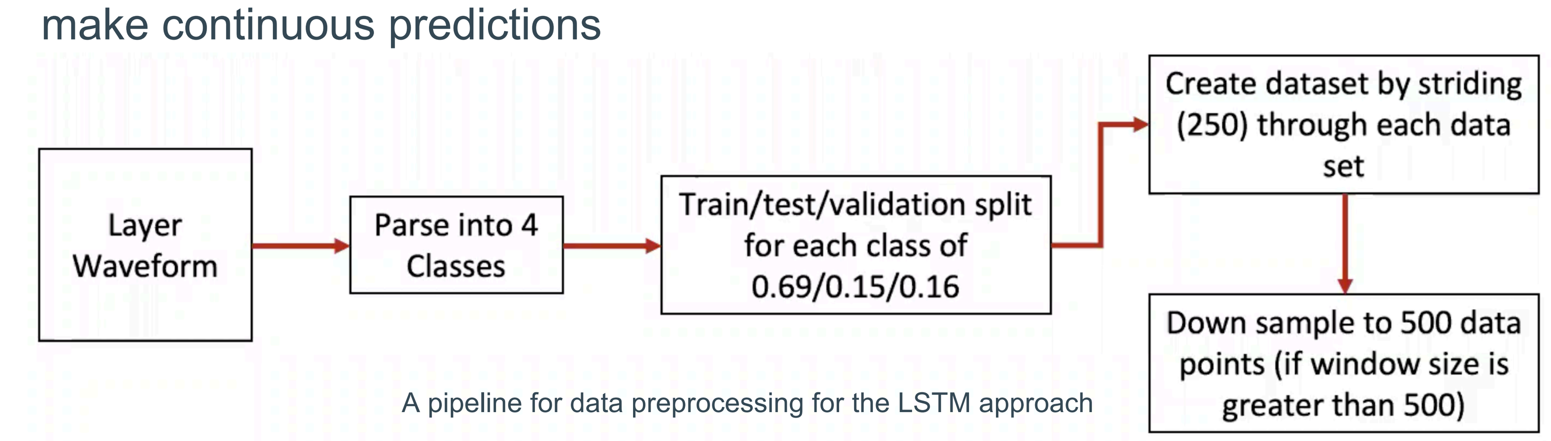
Utilizing ResNet and DenseNet models pretrained on ImageNet, we were able to build a strong classifier for 3 flaw formations mechanisms and 1 normal case. All models were able to classify the dataset very quickly, so we chose ResNet18 as the final model due to its small size.

On average, the CNN approach takes ~0.06 seconds to process and classify a raw sample, making real-time monitoring systems possible and effective.

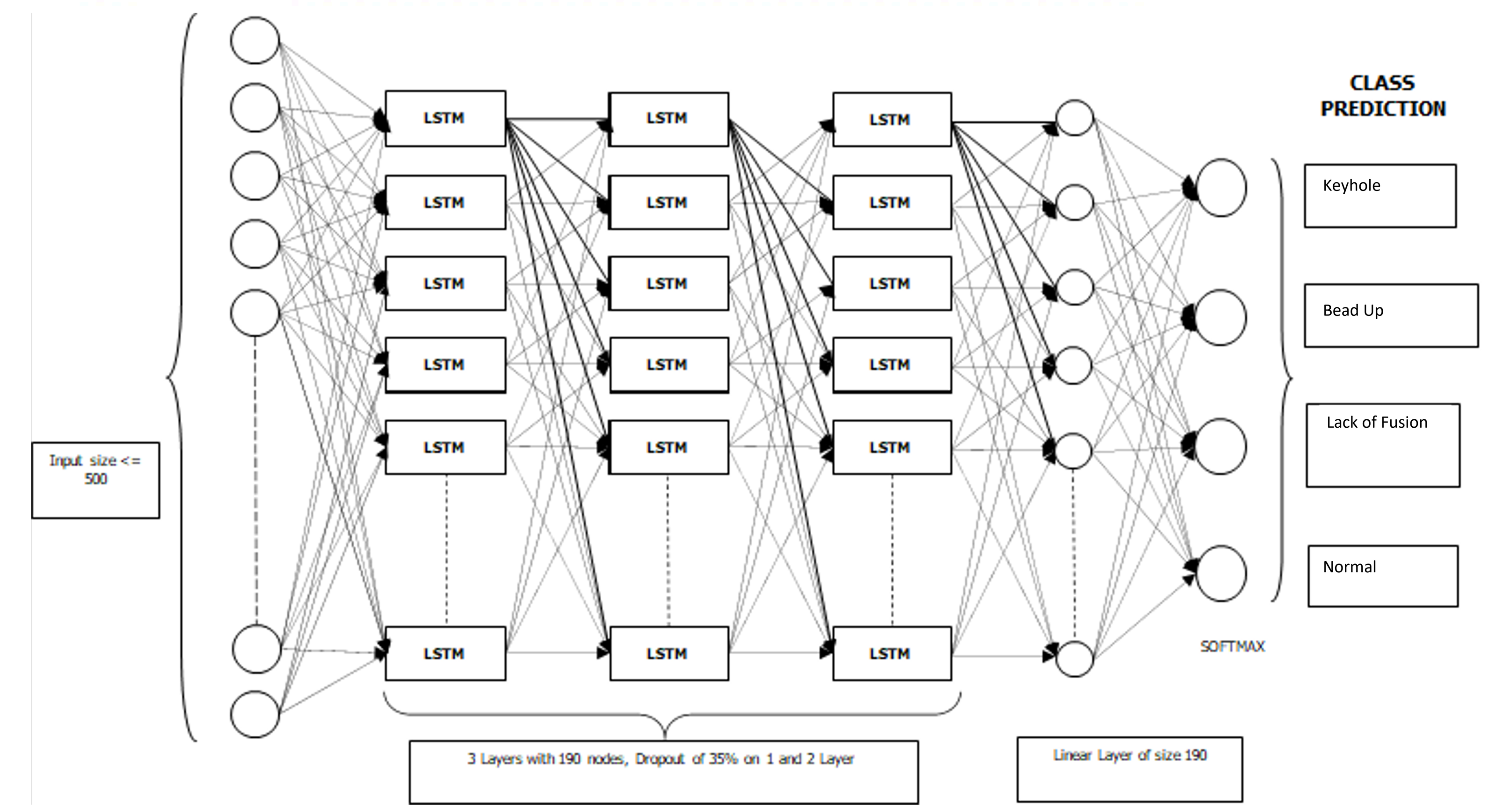
LSTM Approach

Explore the long-term dependencies in the acoustic data

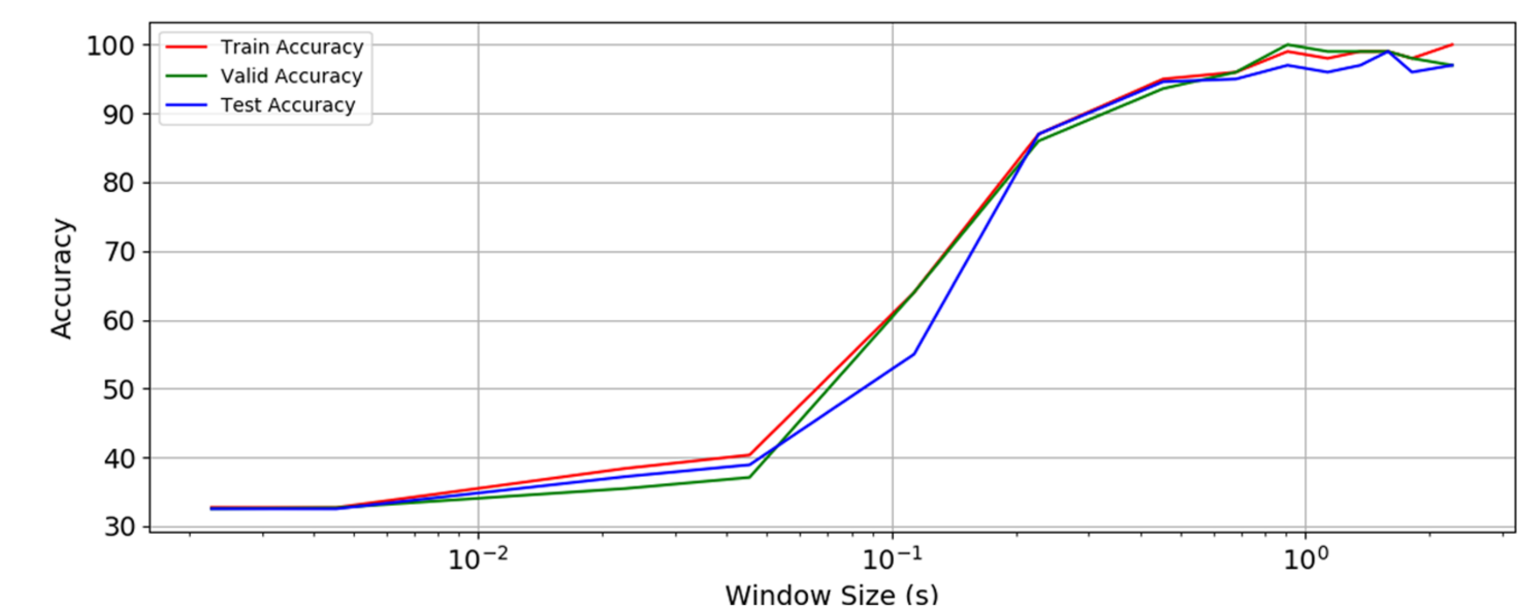
The acoustic signal generated during the manufacturing process is processed as sequence of numbers (as time series) and then fed to the LSTM models to make continuous predictions



A pipeline for data preprocessing for the LSTM approach



Our experimented LSTM network architecture



The model performance boost when the window size is larger than 0.5 seconds and reaches 95% when the window size is larger than 1 second. We can conclude that we found a characteristic scale to reliably identify the flaw type. The best model can also achieve 99% test accuracy and takes ~0.1 sec to make a prediction, which is slightly longer than the CNN models.

References

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).