DEEP LEARNING FOR MILLING CHATTER DETECTION: INTEGRATING SPINDLE AWARENESS IN TIME SERIES LIGHTWEIGHT ADAPTIVE NETWORK

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

FURKAN BAHÇELİ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MECHANICAL ENGINEERING

APRIL 2025

Approval of the thesis:

DEEP LEARNING FOR MILLING CHATTER DETECTION: INTEGRATING SPINDLE AWARENESS IN TIME SERIES LIGHTWEIGHT ADAPTIVE NETWORK

submitted by **FURKAN BAHÇELİ** in partial fulfillment of the requirements for the degree of **Master of Science in Mechanical Engineering Department, Middle East Technical University** by,

Prof. Dr. Naci Emre Altun
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Serkan Dağ
Head of Department, **Mechanical Engineering** _____

Assist. Prof. Dr. Hakan Çalışkan
Supervisor, **Mechanical Engineering, METU** _____

Assoc. Prof. Dr. Şeyda Ertekin
Co-supervisor, **Computer Engineering, METU** _____

Examining Committee Members:

Assoc. Prof. Dr. Buğra Koku
Mechanical Engineering, METU _____

Assist. Prof. Dr. Hakan Çalışkan
Mechanical Engineering, METU _____

Assoc. Prof. Dr. Şeyda Ertekin
Computer Engineering, METU _____

Prof. Dr. Hakkı Özgür Ünver
Mechanical Engineering, TOBB ETU _____

Assist. Prof. Dr. Orkun Özşahin
Mechanical Engineering, METU _____

Date:07.04.2025

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Furkan Bahçeli

Signature :

ABSTRACT

DEEP LEARNING FOR MILLING CHATTER DETECTION: INTEGRATING SPINDLE AWARENESS IN TIME SERIES LIGHTWEIGHT ADAPTIVE NETWORK

Bahçeli, Furkan
M.S., Department of Mechanical Engineering
Supervisor: Assist. Prof. Dr. Hakan Çalışkan
Co-Supervisor: Assoc. Prof. Dr. Şeyda Ertekin

April 2025, 95 pages

This thesis introduces the Spindle-Aware Time Series Lightweight Adaptive Network (SA-TSLANet) for improved chatter detection in milling operations. Chatter, a self-excited vibration that degrades machining quality and reduces tool life, remains challenging to detect reliably across varied conditions. The proposed model integrates physics-informed spectral processing with deep learning through the novel Spindle-Aware Adaptive Spectral Block (SA-ASB), which effectively separates spindle harmonics from chatter vibrations. A multi-objective loss function incorporating crest factor, spindle, and ratio components guides the model to learn critical signal characteristics. Experimental validation demonstrates exceptional performance with 0.99 accuracy, 0.98 precision, 0.99 recall, and 0.99 F1 score. SA-TSLANet detects chatter onset 0.19 seconds earlier than traditional FFT methods, providing crucial time for intervention. The model maintains robust generalization (0.98 F1 score) even with limited training data and shows computational efficiency suitable for real-time industrial implementation. The research also explores training exclusively with healthy cutting data to reduce implementation costs. This physics-informed deep learning

approach advances machining process monitoring by balancing theoretical rigor with practical applicability, providing a foundation for intelligent manufacturing systems that enhance productivity, quality, and tool utilization.

Keywords: Machine Learning, Chatter Detection, Signal Processing, Time Series Analysis, Manufacturing Automation

ÖZ

FREZEDE TIRLAMA TESPİTİ İÇİN DERİN ÖĞRENME: ZAMAN SERİLERİNDE İŞ MİLİ FARKINDALIĞIYLA ENTEGRE HAFİF ADAPTİF AĞ

Bahçeli, Furkan
Yüksek Lisans, Makina Mühendisliği Bölümü
Tez Yöneticisi: Dr. Öğr. Üyesi. Hakan Çalışkan
Ortak Tez Yöneticisi: Doç. Dr. Şeyda Ertekin

Nisan 2025 , 95 sayfa

Bu tez, frezeleme işlemlerinde tırlama tespitini iyileştirmek amacıyla İş Mili Farkındalıklı Zaman Serisi Hafif Adaptif Ağı'nı (SA-TSLANet) tanıtmaktadır. İşleme kalitesini düşüren ve takım ömrünü azaltan kendi kendine uyarılan bir titreşim olan tırlama, değişken koşullar altında güvenilir bir şekilde tespit edilmesi zor bir sorundur. Önerilen model, yenilikçi İş Mili Farkındalıklı Adaptif Spektral Blok (SA-ASB) aracılığıyla fizik tabanlı spektral işlemeyi derin öğrenmeyle entegre ederek, iş mili harmoniklerini titreşim sinyallerinden etkili bir şekilde ayırır. Tepe faktörü, iş mili ve oran bileşenlerini içeren çok hedefli bir kayıp fonksiyonu, modelin kritik sinyal özelliklerini öğrenmesini sağlar. Deneysel doğrulama, 0.99 doğruluk, 0.98 hassasiyet, 0.99 duyarlılık ve 0.99 F1 skoru ile başarılı bir performans göstermektedir. SA-TSLANet, geleneksel FFT yöntemlerine göre tırlama başlangıcını 0.19 saniye daha erken tespit ederek müdahale için kritik zaman kazandırır. Model, sınırlı eğitim verisiyle dahi güçlü genelleme yeteneğini korur (0.98 F1 skoru) ve gerçek zamanlı endüstriyel uygulamalar için uygun hesaplama verimliliği sunar. Araştırma ayrıca, yal-

nızca sağlıklı kesim verileriyle eğitim yapılarak uygulama maliyetlerinin azaltılmasını da ele alır. Bu fizik tabanlı derin öğrenme yaklaşımı, kuramsal titizlik ile pratik uygulanabilirliği dengeleyerek işleme süreci izlemeye katkı sağlar ve üretkenliği, kaliteyi ve takım kullanımını artıran akıllı üretim sistemleri için bir zemin oluşturur.

Anahtar Kelimeler: Makine Öğrenmesi, Tırlama Tespiti, Sinyal İşleme, Zaman Serisi Analizi, İmalat Otomasyonu

To my lovely family.

ACKNOWLEDGMENTS

I would like to express my heartfelt appreciation to Asst. Prof. Hakan Çalışkan for his invaluable mentorship and guidance during my thesis journey, as well as to Assoc. Prof. Şeyda Ertekin for her significant support.

I sincerely appreciate the support provided by Repkon Dynamics company and especially Mr. Ahmet Can Afatsun throughout this process.

I extend sincere thanks to Mr. Ömer Faruk Turhan, Mr. Furkan Enes Yıldırım, Mr. Cumhur Özbaş, Mr. Ekin Akdeşir, Mr. Yağız Güzelcan, Mr. Fırat Mert, and Mr. Cem Türksal for their valuable insights and solid encouragement.

Finally, I extend my sincere gratitude to my family, whose a lifelong support and encouragement made this achievement possible.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS

AI	Artificial Intelligence
ASB	Adaptive Spectral Block
CNN	Convolutional Neural Network
CNC	Computer Numerical Control
CPU	Central Processing Unit
CWT	Continuous Wavelet Transform
DFT	Discrete Fourier Transform
DMLP	Deep Multi-Layer Perceptron
EEMD	Ensemble Empirical Mode Decomposition
FFT	Fast Fourier Transform
GAN	Generative Adversarial Network
GELU	Gaussian Error Linear Unit
GPU	Graphics Processing Unit
HHT	Hilbert-Huang Transform
HSMP	Hot Strip Mill Operations
ICB	Interactive Convolution Block
IDFT	Inverse Discrete Fourier Transform
IMF	Intrinsic Mode Function
IoT	Internet of Things
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error

MICA	Modified Independent Component Analysis
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
ON-LSTM	Ordered-Neurons Long Short-Term Memory
PBT	Population-Based Training
RBF	Radial Basis Function
RMS	Root Mean Square
RMSE	Root Mean Square Error
RP	Recurrence Plots
RQA	Recurrence Quantitative Analysis
SA-ASB	Spindle-Aware Adaptive Spectral Block
SA-TSLANet	Spindle-Aware Time Series Lightweight Adaptive Network
SDAE	Stacked-Denoising AutoEncoder
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
TP	Tooth Passing
TPE	Tree-structured Parzen Estimators
TSLANet	Time Series Lightweight Adaptive Network
WPT	Wavelet Package Transform
WWPE	Weighted Wavelet Packet Entropy

LIST OF SYMBOLS

SYMBOLS

α	Weight parameter in forecasting loss function
a_s, a_r, a_c	Weighting factors for loss components
b	Bias term in SVM
d_h	Distance to closest harmonic
δ	Lagrange multipliers in SVM
ϵ	Small constant to prevent division by zero
$f(\mathbf{x})$	SVM classification function
f_s	Sampling rate, spindle frequency
$\phi_j(z)$	Angular position of tooth j at axial position z
γ	RBF kernel parameter controlling influence radius
$h_j(\phi, z)$	Instantaneous chip thickness
$K(\mathbf{x}_i, \mathbf{x}_j)$	Kernel function for SVM
$\mathcal{L}_{\text{total}}$	Total loss function
$\mathcal{L}_{\text{spindle}}$	Spindle loss component
$\mathcal{L}_{\text{ratio}}$	Ratio loss component
$\mathcal{L}_{\text{crest}}$	Crest factor loss component
$M_s(f)$	Spindle frequency mask
M_c	Chatter frequency mask
$\odot$	Hadamard (element-wise) product
σ	Sigmoid function
σ_d	Standard deviation
σ_s	Bandwidth parameter for spindle harmonics
τ	Threshold parameter in ASB
$\tau_{spindle}$	Threshold for spindle components
$\tau_{chatter}$	Threshold for chatter components

τ_{global}	Threshold for global components
$\mathbf{w}$	Weight matrix
W_s, W_c, W_g	Learnable complex weights
X_f	Frequency domain representation
$\mathcal{F}[\cdot]$	Fourier transform operator

CHAPTER 1

INTRODUCTION

1.1 Background of Metal Cutting Operations and Chatter

A basic machining technique that is often used in contemporary industrial systems is milling. It uses rotating cutting tools to remove material from a workpiece by combining the linear motion of the workpiece with the circular motion of the tool. Cutting speed, depth of cut, feed rate, and tool shape are some of the important characteristics that define the process. The material is removed by the intermittent interaction of the cutting tool's numerous cutting edges (teeth) with the workpiece in standard milling operations [5]. The main factors influencing the cutting conditions are feed rate, axial depth of cut, radial depth of cut, and spindle speed [6].

1.1.1 Types of Milling Operations

End Milling: A machining operation in which a cylindrical cutting tool rotates around its vertical axis to eliminate material from the workpiece. The instrument may traverse laterally over the workpiece to produce flat surfaces, slots, profiles, and curves. The cutting edges are positioned on both the end face and the sides of the tool. This work employs experimental equipment and empirical data from end milling operations to do thorough evaluations of chatter phenomena.

Face Milling: A machining procedure in which a large-diameter tool with several cutting blades rotates around an axis perpendicular to the surface of the workpiece. The tool mostly utilizes its face (bottom surface) to generate expansive, flat horizontal surfaces in a single operation. The cutting edges are mostly situated on the front of

the instrument rather than its sides.

Peripheral Milling: The cutting edges are located on the perimeter of the cutter, with the axis of rotation aligned parallel to the machined surface.

Milling might be also categorized into:

- Up Milling: The cutter rotation is oriented contrary to the feed direction.
- Down Milling: The cutter rotation direction aligns with the feed direction.

Machining processes might experience numerous vibrations that considerably affect machining quality and efficiency. The vibrations may be categorized into three primary types: free vibrations, forced vibrations, and self-excited vibrations [7]. Although both free and forced vibrations may be managed by appropriate machine design and maintenance, self-excited vibrations, especially chatter, provide a more intricate issue in machining processes [8].

Chatter is a critical issue of instability in the machining process, defined by undesirable excessive vibrations between the tool and the workpiece. This phenomena generates significant noise and leads to subpar surface finish quality. Moreover, it adversely impacts the longevity of machines and tools, along with the general dependability and safety of machining operations [9]. The manufacturing sector has historically been impacted by this issue, making it a vital subject for both scholarly and industrial investigation.

Self-excited chatter is the major sort of oscillation encountered in real machining operations [8]. While cycling at a frequency close to, but not equal to, the system's dominant structural mode, the maximum chip thickness might rise exponentially driven by the phase shift between successive wave surfaces. The rising vibrations cause heightened cutting pressures, ending in a subpar, undulating surface finish [10].

1.2 Significance of Chatter Detection in Manufacturing

In contemporary industrial environments, particularly with the advent of Industry 4.0 concepts, there is an increasing need for automated and intelligent machining pro-

cesses. In this context, the prompt detection of chatter is crucial, as it enables real-time process optimization and prevents quality issues from worsening [11]. Traditional methods of chatter detection relied heavily on operator experience and post-process assessment, which are neither efficient nor suitable for modern high-speed production requirements [12].

The detection of chatter has progressed owing to the development of various sensor technologies and signal processing methods. Monitoring systems often use sensors such as accelerometers, dynamometers, and acoustic emission sensors to identify physical events associated with chatter [13]. The selection of appropriate sensors and their placement is crucial, since different sensor types may demonstrate varying sensitivity to chatter conditions [14].

1.3 Chatter in Machine Tool Systems

Chatter is a notable kind of self-excited vibration in machine tool systems. Chatter mostly occurs when cutting forces excite a structural mode of the machine tool-workpiece system. This phenomenon is characterized by:

- Excessive vibration between the tool and the workpiece
- Surface polish lacking luster
- Decrease in the lifetime of machines and tools
- The reliability and security of operations have declined.

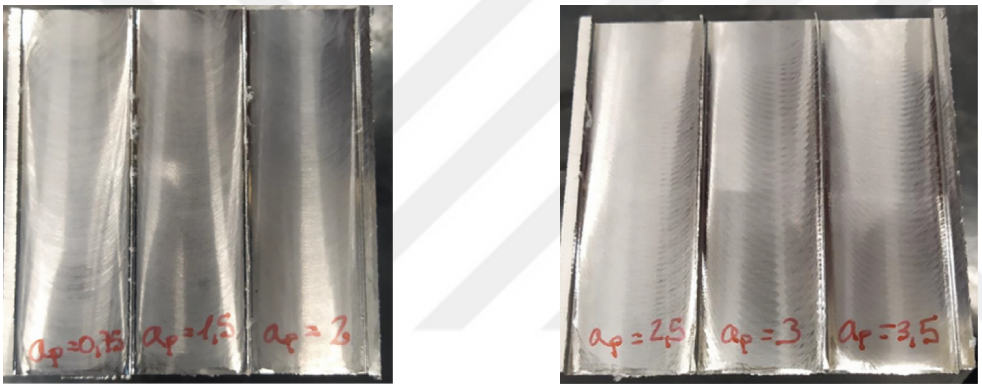
1.3.1 Types of Chatter

Three primary forms of self-excited chatter were recognized by Wierich and Groch [16].

1. **Arnold-Type:** This happens when changes in cutting velocity interact with changes in force. Given the name of its finder.

2. **Regenerative:** Surface abnormalities from previous cuts impact cutting forces that are applied to future cuts [8], resulting in:
    - Rapid increase in chip thickness
    - Main structural frequency vibrations
    - Rising cutting pressures
    - The surface's wave patterns
3. **Mode Coupling:** This happens when forces generate vibrations in different directions, which further reinforce each other, resulting in a feedback loop [17].

As seen in Figure 1.1, chatter on machined surfaces may be visually represented.



(a) Surface finish without chatter

(b) Surface finish with chatter

Figure 1.1: Comparison of surface finishes with and without chatter. a_p means depth of cut [1]

The mechanics of milling include intricate interactions between the cutting tool and the workpiece. The diagram illustrating milling dynamics is shown in Figure 1.2. In milling operations, material removal transpires via intermittent contact between the cutting blades of the spinning tool and the workpiece. Budak and Altintas [2] articulate that the instantaneous chip thickness $h_j(\phi, z)$ is represented as:

$$h_j(\phi, z) = [(y_c - y_c^p) - (y_w - y_w^p)] \cos\phi_j(z) + s_t \sin\phi_j(z) + [(x_c - x_c^p) - (x_w - x_w^p)] \sin\phi_j(z) \tag{1.1}$$

where:

- $\phi_j(z)$ is the angular position of tooth j at axial position z
- s_t is the feed per tooth
- (x_c, y_c) and (x_w, y_w) represent current cutter and workpiece displacements
- (x_c^p, y_c^p) and (x_w^p, y_w^p) represent displacements from the previous tooth pass

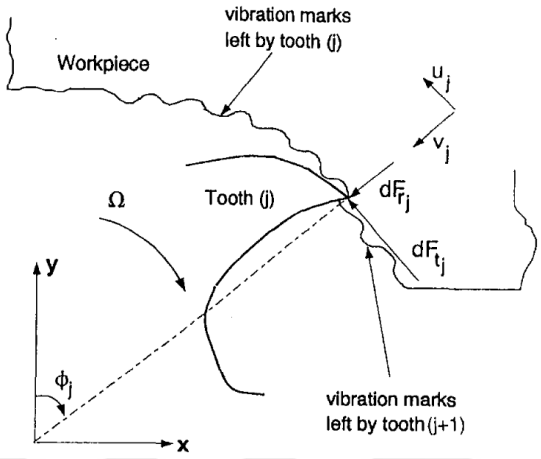


Figure 1.2: Milling diagram for dynamics [2]

This chip thickness formulation exemplifies a conventional mathematical method for understanding milling dynamics and chatter. Although these analytical approaches provide significant theoretical insights, their practical application may be difficult owing to the intricacies of actual machining circumstances. This encourages our investigation of AI-driven methodologies in subsequent chapters, which may effectively identify intricate communication patterns without necessitating explicit mathematical representation of all system variables.

CHAPTER 2

LITERATURE REVIEW

2.1 Literature review for chatter detection

Many methods exist for identifying chatter during milling operations. As seen in Figure 2.1, it may be arranged as a graph. One way to classify procedures is by whether they are conducted online or offline. Prior to the cutting process, offline approaches are used to forecast and avoid chatter. The inability to respond to changes occurring in real time during machining is a result of the need for pre-process analysis and modeling. Important features of non-online approaches:

- Center your efforts on using mathematical models to forecast stability [18].
- Rely on pre-determined system parameters [19].
- Used for process planning and parameter selection [20].
- Limited by model assumptions and idealizations.

Important features of online approaches:

- Real-time signal processing and analysis [21]
- Direct measurement of process variables [22]
- Immediate detection of chatter onset [23]
- Ability to implement real-time control actions [1]

Using a time domain simulation model, Li et al. established a novel milling chatter stability criteria that is based on the ratio of expected maximum dynamic cutting force

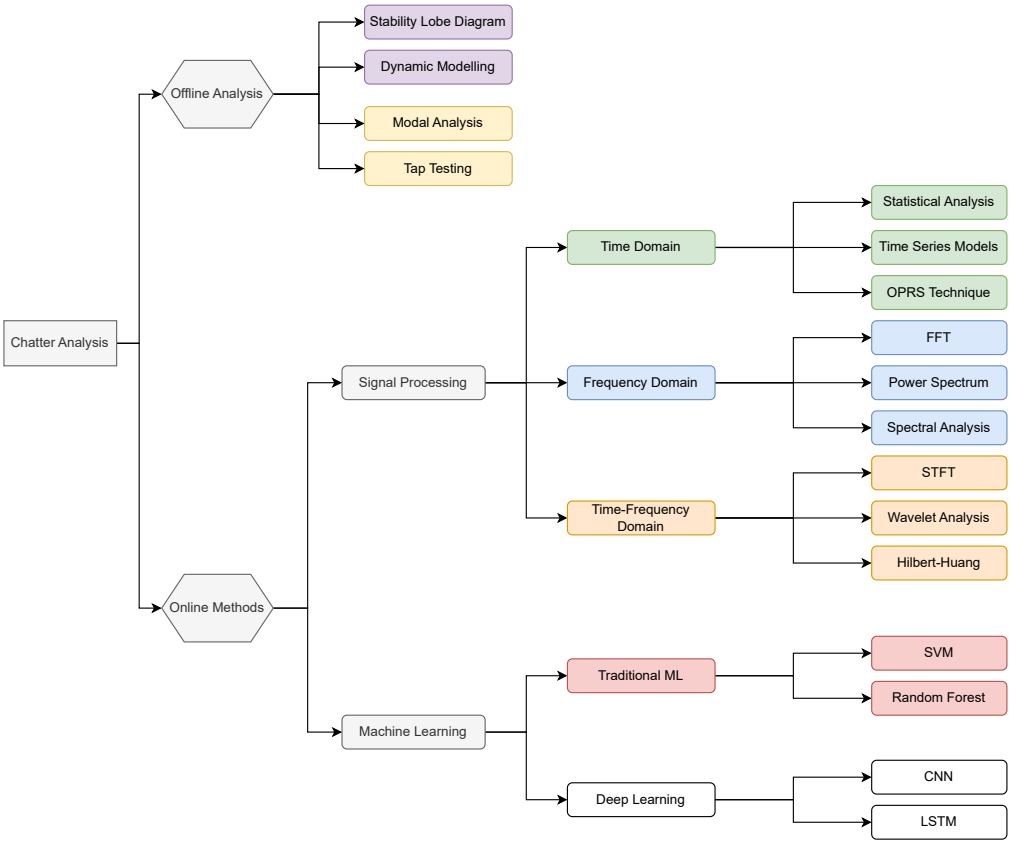


Figure 2.1: Chatter detection diagram

to predicted maximum static cutting force. The model integrates regeneration effects dynamically via estimates of immediate undeformed chip thickness and applies predictive machining theory to determine cutting force. The model's performance was confirmed by experimental validation [24]. For high-speed machining of thin-walled components, Bravo et al. suggested a 3-dimensional stability lobe diagram approach. This method takes into consideration the workpiece's and machine's dynamic behaviors, which is very useful in the aerospace sector.Experimental validation through thin wall machining tests demonstrated the method's effectiveness in preventing chatter and ensuring process stability, and their approach goes beyond simple superposition of machine and workpiece lobe diagrams by including a preliminary step of filtering out irrelevant resonance modes [25]. In order to measure cutting forces in three directions, Altintas and Park used a Spindle-Integrated Force Sensor system. This system included piezo-electric sensors that were installed in the spindle housing. To

avoid structural interference, they also employed a Kalman filter. Using this method, they were able to improve the accuracy of cutting force measurements in high-speed milling operations by achieving high-bandwidth force measurements up to 1000 Hz tooth passing frequency and successfully filtering out distortions from spindle structural modes.

Campa et al. developed a technique to eliminate noise during bull-nose end milling operations on thin floors. Their approach takes into consideration the fact that modal characteristics change as material is removed and employs a stability model based on an averaging method. Experimental testing supported their methods, and they presented a new way that machinists may use to choose the best spindle speeds for the cutting route by using stability diagrams that connect tool position to spindle speed [26]. Using Deep Neural Networks that were pre-trained on simulated data from receptance coupling and then fine-tuned using a limited experimental dataset, Postel et al. used a unique approach to update milling stability predictions. They achieve accurate stability predictions with five times fewer experimental samples than previous machine learning methods by using ensemble learning techniques [27], and their approach eliminates the need for cutting force coefficient measurements or extensive parameter identification. In an effort to achieve more stability with less machine weight, Tlusty and Polacek studied the effects of machine tool features on self-excited vibration (chatter). Using examples like boring bar orientation to support their claims, their theoretical work clarifies the basic features of self-excited vibration and shows how variables like machine rigidity in various directions and tool orientation greatly affect stability [28]. With an emphasis on orthogonal chatter stability theory and its practical applications, Altintas and Weck investigate the basic modeling of chatter vibrations in a variety of machining operations, including as turning, boring, drilling, milling, and grinding. In addition, the authors go over chatter suppression methods that are utilized in industry both offline and in real-time, compare different stability models to experimental findings, and point out where more research is needed to make commercial implementation more successful [29].

The cutting process itself makes use of online techniques. Their primary function is to monitor and identify chatter in real-time. During machining, they can adjust to new circumstances and provide you input quickly so you can manage the opera-

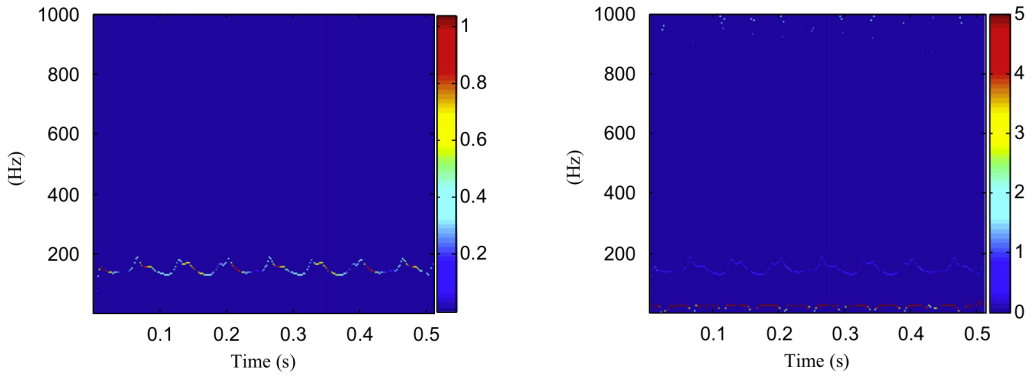
tion. They rely more on empirical evidence and scientific testing. Using both single- and multi-sensor configurations, Kuljanic et al. developed a strong chatter detection system for industrial uses by comparing a wide range of sensors, such as electrical power sensors, acoustic emission, accelerometers, and rotating dynamometers [12]. According to Kuljanic, the most dependable and strong chatter detection performance was achieved by using multi-sensor systems, namely combining an axial force sensor with accelerometers, but cutting torque was shown to be the best single-sensor indication. In order to improve the signal-to-noise ratio, Lamraoui designed a five-step chatter detection procedure for CNC milling machines that integrates artificial intelligence with vibratory signal analysis [30]. The procedure begins with multiband resonance filtering and continues with envelope analysis. Finally, entropy-based feature selection is used. According to Lamraoui, their method utilises RBF and MLP neural networks to distinguish between stable and unstable cutting situations. It was tested and shown effective on a Huron K2X10 milling machine with different cutting depths and speeds.

By analyzing cutting force signals using recurrence plots (RP) and recurrence quantitative analysis (RQA), Chen et al. developed a chatter monitoring approach for changeable cutting settings that overcomes the shortcomings of conventional time-frequency domain features. Their method effectively differentiates between stable and unstable cutting states under different situations by integrating affinity propagation clustering for feature identification with light gradient boosting for classification [31]. According to Schmitz's new chatter detection method, which is based on calculating the statistical variance of milling audio signals sampled once per revolution, stable cuts have low variance because they are synchronous, and unstable cuts have high variance because they are asynchronous. By comparing their variance-based strategy with analytical stability predictions and standard FFT-based chatter detection techniques, they were able to verify it [32]. By developing a chatter index (CI) and using a modified independent component analysis (MICA) technique, Jo et al. presented a thorough framework for tracking and analyzing noise in hot strip mill operations (HSMP), surpassing conventional methods of monitoring. According to Jo, a relative contribution plot approach was also used to enhance chatter cause diagnosis in HSMP, leading to clearer identification of the underlying factors responsible for

identified chatter [33]. For large-scale thin-walled part mirror milling, Wang et al. devised a novel chatter detection approach based on Q-factor analysis. This method extracts chatter-related components from coupled signals resulting from the supporting head's effect on system dynamics. Compared to conventional indicators and other techniques of frequency band analysis, their methodology, which combines Q-factor analysis with Support Vector Machine (SVM) classification, focused on frequency bands with largest Q-factor variation, and produced greater prediction accuracy [34].

Using the Local Outlier Factor (LOF) method for classification and relative wavelet packet energy entropy for high-frequency sensitivity, Yao created a chatter detection and suppression module for intelligent spindles that operates in real-time. They successfully prevented surface damage during end-milling operations by actively monitoring and suppressing noise using a LOF-trained model and a pre-obtained 3D stability lobe diagram [35]. Tangjitsitcharoen used wavelet transform analysis to develop a new method for detecting chatter in ball end milling processes. They introduced three parameters derived from cutting force variance ratios, and they successfully validated their method experimentally across different cutting conditions, demonstrating effective chatter detection [36]. Experimental validation showed that the hybrid chatter detection method that Cao et al. created for end milling, which combines Wavelet Package Transform (WPT) for signal denoising and Hilbert-Huang Transform (HHT) for time-frequency analysis, effectively identifies chatter levels [3]. As illustrated in Figure 2.2, the authors' suggested model and HHT output the results. After comparing various sensor types (sound, acceleration, and cutting force), Thaler achieved over 96% accuracy in chatter detection using a sound-based chatter detection method for band sawing. This method combines Short-Time Fourier Transform for signal processing and Quadratic Discriminant Analysis for classification [37]. It was validated through machining experiments that Sun and Xiong's real-time chatter detection method, which is based on weighted wavelet packet entropy (WWPE), achieves millisecond-level computational efficiency, and uses extreme value statistics for threshold determination. The authors optimized the weights to maximize distinction between stable and chatter states [38]. Chen et al. integrated machine learning with image processing to create a chatter detection system. They employed Support Vector Machine for classification after utilizing average FFT to identify dominant

frequency bands in STFT time-frequency images. The system outperformed wavelet transform and time-domain methods in micro-milling tests [39].



(a) Proposed technique with WPT

(b) Raw HHT

Figure 2.2: Comparison between articles proposed model and HHT [3]

By using a Kalman filter to eliminate forced vibrations and then implementing a nonlinear energy operator to detect chatter, Caliskan created a novel real-time approach to milling chatter detection. Experimental validation of their methodology demonstrated quicker chatter detection than conventional frequency-based approaches. Please refer to Figure 2.3 for their primary diagram. The authors used Extended Kalman Filters (EKF) to identify vibration characteristics, and they created a real-time algorithm that detects and eliminates chatter in milling automatically by adjusting spindle speed based on chatter energy feedback from axis encoder measurements. The algorithm was successfully tested on a CNC machine [40].

Alongside the above mentioned strategies, machine learning and deep learning may be used to identify chatter. Using a combination of a deep residual CNN for time-frequency spectrum analysis and a support vector machine classifier that processes entropy characteristics collected by variational mode decomposition, Gao et al. developed a unique machine learning strategy for spotting cutting chatter. Compared to standard chatter identification approaches, their hybrid method achieved much higher accuracy in recognizing various chatter states from vibration signals [41]. By measuring the structural vibrations of the machine headstock and isolating the modal components of structural vibrations using a shortpass lifter on the cepstrum, Jeong was able to create a real-time chatter detection system for milling operations, effectively

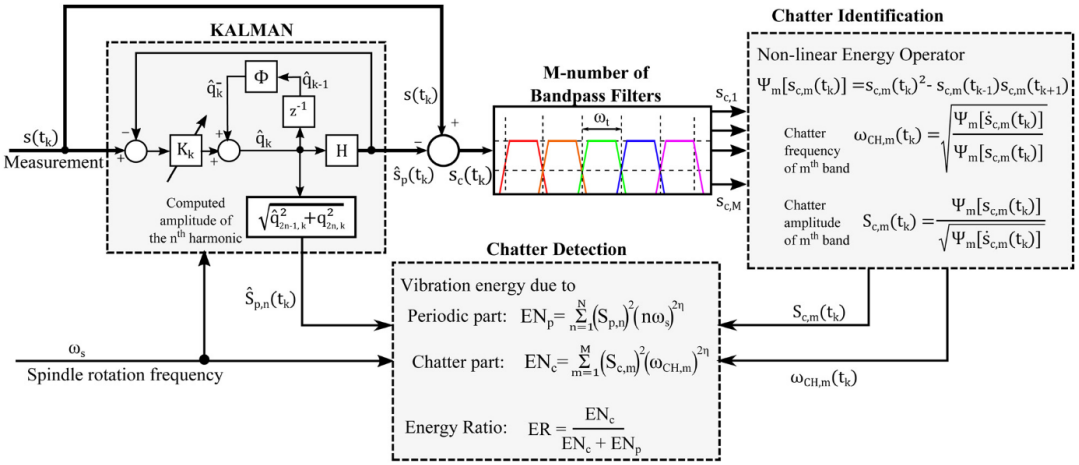


Figure 2.3: Transfer function identified experimentally

reducing the impacts of spindle rotation. According to Jeong, the classifiers used a 1D convolutional neural network to distinguish between stable and chatter states in the processed vibration data. They tested the algorithm with different cutting settings and compared the results to other classifiers.By integrating a deep convolutional generative adversarial network (GAN), Chen-Lee chaotic error mapping for data preprocessing, and a modified convolutional neural network (CNN), Kuo et al. came up with a novel approach to deal with the problem of insufficient chatter data in CNC machining. Their method outperformed traditional data augmentation techniques and achieved high accuracy in chatter detection using only 60 original training samples through leave-one-out cross-validation. This was demonstrated by their GAN-based approach [42]. Using audio data from CNC machining processes, Kvinevskiy created a viable chatter detection approach that integrates machine learning classification with autoencoder-based dimensionality reduction. According to Kvinevskiy, their method is easy to implement on the work floor since it simply asks machinists to classify audio samples as either chatter or non-chatter during training. A five-stage approach to chatter detection in CNC milling was devised by Lamraoui. It incorporates signal processing techniques such as multiband resonance filtering and envelope analysis, as well as entropy-based feature selection and neural network classification. The approach was tested with both RBF and MLP designs. The researchers successfully differentiated between stable and unstable cutting situations by testing their method on a Huron K2X10 milling machine with different cutting depths and rotating speeds

[43]. To solve the practical problem of costly data collecting in industrial settings, Rezaei created a new Encoder GAN (EGAN) method to forecast milling stability maps with very little experimental data (only 5-10 cutting attempts). Their methodology, which integrates networks for encoders, generators, and discriminators, outperformed previous methods and showed adaptability to various tool configurations, especially when faced with insufficient data on unknown parameters [44]. By combining surface roughness prediction with chatter vibration detection in machining operations, Rifai's machine vision-based solution leverages deep convolutional neural networks to do away with conventional feature extraction techniques. In comparison to conventional contact-based profilometer measurements, their non-contact method showed superior accuracy and faster processing speeds when tested on turned and milled surfaces using separate and combined models from four distinct deep learning architectures [45]. Using time-domain features extracted from vibration data collected by spindle-mounted accelerometers, Sener and Batihan created a chatter detection method for slot milling using a Deep Multi-Layer Perceptron (DMLP) neural network. These features include root mean square, clearance factor, skewness, crest factor, and shape factor. Their method was confirmed by identifying chatter occurrence in the training data using FFT analysis. This proved that their DMLP algorithm could successfully detect chatter using just time-domain characteristics [46]. In order to optimize the hyperparameters, Shi created a new chatter detection method that uses data gathered from four accelerometers during CNC milling. This method integrates population-based training (PBT) with ordered-neurons long short-term memory (ON-LSTM) networks. By comparing the RNN's performance with traditional STFT time-frequency analysis, their method showed that it could effectively detect chatter and that the learned patterns were consistent with conventional frequency-domain analysis, making it easy to understand and interpret [47]. An approach to real-time chatter detection that Tran suggested combines deep convolutional neural networks (CNN) with continuous wavelet transform (CWT). In this technique, cutting force signals are transformed into scalogram pictures, which record the energy shift from the frequency of tooth passage to the frequency of chatter. According to Tran, their method outperformed traditional machine learning approaches and eliminated the need for human feature engineering, all while achieving a classification accuracy of 99.67% in differentiating between stable, transitive, and unstable cutting

states. By integrating analytical stability lobe diagrams with AlexNet CNN architecture, Unver created a novel transfer learning framework. This framework uses numerical time-domain solutions for training and labeling, which eliminates the requirement for costly experimental data collecting. They achieved 82% to 100% accuracy across different situations using just analytical data for training instead of actual measurements, thanks to their methodology that combines an ensemble empirical mode decomposition (EEMD) preprocessing method with an IMF-based multi-band ensemble strategy [48]. By training an LSTM neural network on simulated control current sequences obtained from dynamic equations of the milling process, Vashisht created a new chatter detection approach that employs the current signals of the ball screw drive instead of vibration or sound sensors. By doing away with the need for extra sensors and expensive experimental training data, their sensorless method was able to reach a 98% accuracy rate in experimental validation, making it a more economical option for industrial use [5]. To eliminate the requirement for human threshold setting, Wan devised a machine learning-based chatter detection solution that combines an enhanced Adaptive Boosting (AdaBoost) algorithm employing SVM weak classifiers with stacked-denoising autoencoder (SDAE) for automated feature extraction. They show good accuracy across different milling settings and are resilient to perhaps mislabeled training samples using an approach that combines conventional dimensionless features from the time and frequency domains with automatically derived features from SDAE [23].

Although a lot of research has gone into this, there are still several challenges to achieving reliable chatter detection. The complex milling process is challenging because it uses many cutting blades that come into touch with the workpiece on occasion, resulting in complex and ever-changing interactions [49]. It is difficult to distinguish between steady cutting and early chatter phases due to noise in sensor data and variable cutting conditions [43].

2.2 Literature review for anomaly detection in time series data

When using deep learning to analyze time series data, chatter stands out as a major oddity. The use of deep learning techniques for anomaly detection has been the

subject of much study in both online and offline settings. Another name for this phenomena is novelty detection. In order to provide a unified, interdisciplinary strategy for chatter detection, this work reviews the literature on anomaly detection as well as novelty detection.

As an alternative to conventional manual feature generation methods, Fu and Yang presented a deep belief network–based automated feature construction methodology for monitoring machining states (idle movement, steady cutting, and chatter) using vibration data. Their approach safeguarded workpieces against chatter damage and achieved almost 100% accuracy in state identification, all while including a voting technique for dependable online detection [50]. In order to overcome the difficulty of choosing the best anomaly detection model for unlabeled time-series data, Goswami devised a strategy that uses robust rank aggregation to integrate three surrogate metrics: prediction error, model centrality, and performance on synthetic anomalies. A major practical issue in anomaly detection, where labels are usually few, was addressed by their method, which was just as successful as utilizing partly labeled data for model selection [51]. Through the integration of an autoencoder and a GAN architecture, AEGAN-AD was created by Jiang. This unique unsupervised method for machine audio anomaly detection uses the generator to rebuild input spectrograms and the discriminator to aid in both the training and detection processes. In the DCASE 2022 Challenge TASK 2 dataset, their approach obtained state-of-the-art results on five different kinds of machines and introduced a novel capacity for anomaly localization [52]. To tackle the problem of unbalanced time series data in industrial applications, Jiang created a GAN-based anomaly detection method using an encoder-decoder-encoder generator architecture that is trained only on normal samples. While testing on two rolling bearing datasets, their technique achieved 100% accuracy in differentiating abnormal samples from normal ones, without necessitating previous knowledge of abnormal samples, by calculating anomaly scores utilizing both apparent and latent losses [53]. Combining deep learning-based feature extraction with classical audio data (STFT, MFCC, and spectral centroid) and subsequent PCA and SVM classification, Mobtahej suggested a deep learning-based anomaly detection approach for midstream gas compressors. By analyzing audio sensor data, their method seeks to safeguard expensive compression systems (worth around $1–2 million USD)

from potential faults [54]. Using a two-step masking technique—random masking to learn normal data representation and an exclusive and entropy-based re-masking step to refine the model's understanding of uncertain parts—AnoFormer, a novel pure Transformer-based GAN framework for time series anomaly detection, was introduced by Shin et al. Their technique was the first to use a pure Transformer-based GAN for time series anomaly detection, and it outperformed state-of-the-art algorithms on many datasets [55]. Tan maintained and monitored the performance of deep learning models used in multivariate time series anomaly detection by developing an online data drift detection approach utilizing Variational Autoencoders (VAE). Using three real-world industrial datasets and four anomaly detection models, their three-step approach (data collection/statistical analysis, real-time drift detection, and drift interpretation) showed effectiveness, improving the F1 score by up to 0.16 and producing results that could be understood [56]. An innovative end-to-end method for fault-bearing diagnosis was put forward by Vu. It comprises a two-branch model, one based on transformers and cross-attention, and the other on metrics and Mahalanobis distance. The method also incorporates a multiscale big kernel feature extraction module. The ensemble method outperformed typical CNN-based models on the CWRU and PU datasets, even with little training data, and it tackled the problems of both actual and fake errors [57]. For multivariate time-series analysis, Xie et al. created MARINA, a flexible model based on MLP-attention that learns spatial and temporal correlations in the data at the same time. For real-time monitoring applications in AIOps and IoT systems that create enormous volumes of time-series data, their model proved to be very effective in both forecasting and anomaly detection tasks [58]. By building an association-based criteria for anomaly identification, Xu's Anomaly Transformer utilizes a new Anomaly-Attention mechanism. This mechanism takes use of the fact that anomalies tend to associate with nearby time points rather than the whole series. According to Xu, their model outperformed the state-of-the-art models in six benchmarks related to service monitoring, space and earth exploration, and water treatment applications. The model employs a minimax strategy to enhance the distinguishability between normal and abnormal points using Association Discrepancy. A new technique to time series anomaly detection, TGAN-AD, was developed by Xu. It combines a transformer-based generator for contextual feature extraction with a discriminator to aid in anomaly identification. Using both compo-

nents, it calculates anomaly scores. In three public datasets, their model achieved the greatest Recall and F1 scores, maintained high efficiency, and showed excellent hyperparameter selection, surpassing state-of-the-art approaches [59].

2.3 Introduction of TSLANet

Time series analysis is approached in a new way by TSLANet (Time Series Lightweight Adaptive Network), which offers a universal convolutional model meant to efficiently manage various time series activities. Originally created for broad time series analysis, its architecture has various features that make it especially ideal for chatter detection applications in machining operations [4].

2.3.1 Model Architecture

Figure 2.4 shows the whole architecture of TSLANet. Its components are:

- An embedding layer for initial feature extraction
- Multiple layers of ASB and ICB combinations
- A final classification layer adapted to the specific task

Input time series data is processed by the model progressively via various components, each of which contributes to the final representation utilized for detection or classification.

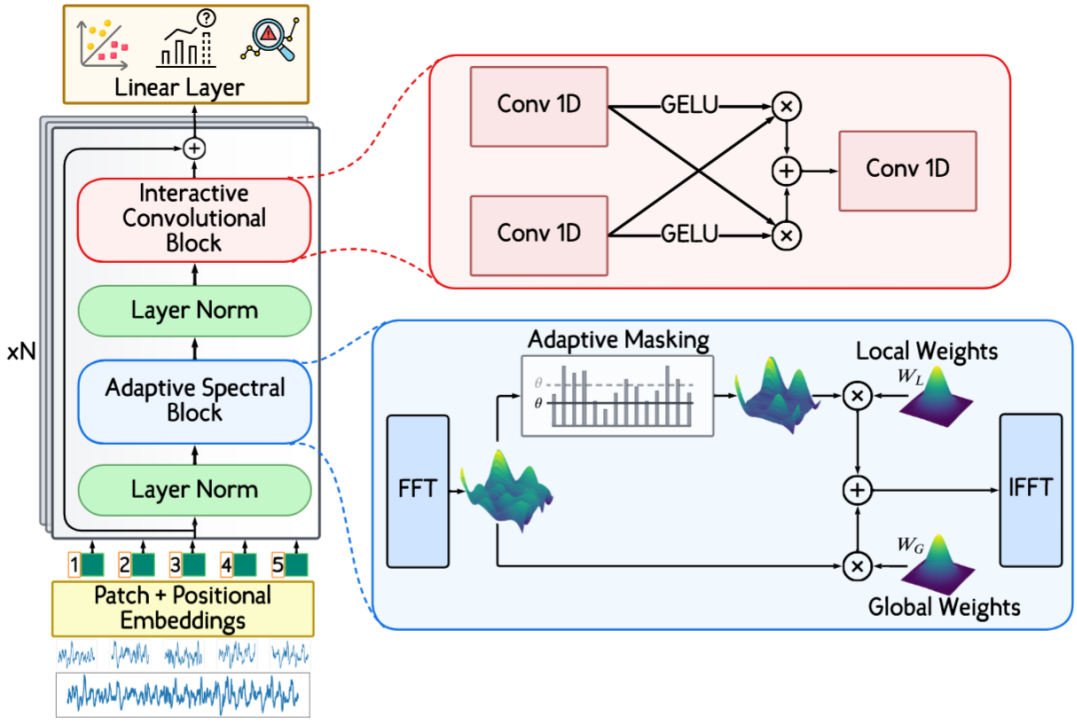


Figure 2.4: TSLANet model architecture

2.3.2 Core Components of TSLANet

Two main functional components form the architectural basis of TSLANet.

2.3.2.1 Adaptive Spectral Block (ASB)

The ASB uses Fourier analysis methods to improve feature representation by means of adaptive thresholding mechanisms lowering noise and by means of capturing both long-range and local interactions. Chatter detection systems especially benefit from this component as it:

- Process frequency domain representations effectively
- Adaptively filter high-frequency noise
- Maintain important spectral features characteristic of chatter phenomena

2.3.2.2 Interactive Convolution Block (ICB)

To catch local characteristics and longer-range relationships, the ICB uses a dual-layer convolutional architecture with varied kernel sizes. This approach helps in:

- Capturing complex temporal patterns
- Improving adaptability to temporal dynamics
- Enhancing feature extraction at multiple scales

2.4 Motivation for Applying TSLANet to Chatter Detection

The adaptation of TSLANet for chatter detection is motivated by many current problems in current detection methods as well as the particular qualities TSLANet offers to address these limitations. Traditional detection techniques might find false positives and lower reliability from the natural noise included in machining sensor data [3]. In industrial environments, where several noise sources overlap with signals of importance, this is particularly concerning. Many current methods rely mostly on human feature extraction and selection, which calls for great domain expertise and runs the danger of overlooking subtle but important signal traits indicating the start of chatter. Most production scenarios are limited by existing models' inability to change cutting conditions, tool geometries, and workpiece materials, thereby restricting their practical usage. Moreover, real-time chatter detection calls for low-latency sensor input analysis algorithms with high accuracy, a feat many sophisticated models fall short of achieving.

TSLANet gets over these limitations with many novel characteristics that make it perfect for chatter detection uses. Combining ASB with ICB results in more dependable identification even under challenging sensing conditions as it conserves necessary signal qualities and provides natural resistance to noise. Reducing the need for significant human feature engineering, the TSLANet design lets critical features be learned end-to-end straight from raw or little processed sensor data, hence exposing subtle patterns that conventional methods would overlook. The design of the model

increases its adaptability across a broad variety of manufacturing environments by allowing it to automatically alter its internal representations to different cutting conditions and tool settings. The architecture of the model also enables effective knowledge transfer across different machining environments, hence reducing the amount of training data required for use in new applications.

2.5 Research Gap and Contribution

Although several deep learning techniques have been proposed for chatter detection, the usage of certain time series architectures like TSLANet is an underappreciated choice with significant potential benefits. By modifying and examining TSLANet, this work offers a novel time series architecture tailored for chatter detection. Compared to conventional methods, it builds a more strong detection system that reduces false positives and false negatives. The device is flexible and can maintain acceptable performance under several machining conditions without needing significant calibration. This work enhances real-time chatter detection features that might be used in industrial settings. These findings show notable promise to raise tool life, machining quality, and production in modern manufacturing processes and cover important holes in present knowledge.

Chatter detection was first done in this work using a direct model based on reconstruction loss; this approach fell short. Thus, Spindle Aware TSLANet (SA-TSLANet) is provided for chatter detection. Established chatter theory highlights the relevance of the relationship between spindle harmonics in the frequency domain [40]. Consequently, spindle awareness is added to the ASB block to effectively filter spindle and chatter harmonics.

CHAPTER 3

METHODOLOGY

3.1 Artificial Neural Networks

Artificial neural networks have arisen as potent computer models capable of learning intricate patterns and representations from input. Inspired by biological neural networks, these systems have linked layers of artificial neurons that interpret and change incoming data via non-linear processes. Each neuron calculates a weighted total of its inputs, applies a non-linear activation function, and transmits the result to neurons in succeeding layers. The essential component of contemporary neural networks is the feedforward neural network, which transmits information unidirectionally from input to output via one or more hidden layers. Every synapse between neurons has a corresponding weight that is modified during training to reduce the discrepancy between the network's predictions and the intended outputs. The backpropagation training procedure allows the network to autonomously acquire hierarchical feature representations from data. This foundation underpins essential deep learning elements such as multilayer perceptrons (MLPs) and attention processes, which will be analyzed comprehensively in the subsequent sections.

3.1.1 Multilayer Perceptron

The multilayer perceptron (MLP) is the most prevalent and adaptable kind of artificial neural network design. The core architecture comprises many tiers of artificial neurons, with signals traveling unidirectionally from input to output via one or more intermediary layers. This feed-forward design guarantees the absence of loops in the network, indicating that the output of each neuron does not influence its subsequent

states [60]. Each artificial neuron in the network fundamentally calculates a weighted sum of its inputs, then applying a non-linear activation function:

$$h_{ij} = f\left(\sum_{i=1}^{n} x_i \cdot w_{ij} - \theta_j\right) \tag{3.1}$$

where n denotes the number of inputs to neuron j, f represents the activation function, $x_i(p)$ signifies the input values, w_{ij} indicates the connection weights, and θ_j refers to the neuron's threshold. The multilayer perceptron structure can be seen in Figure 3.1.

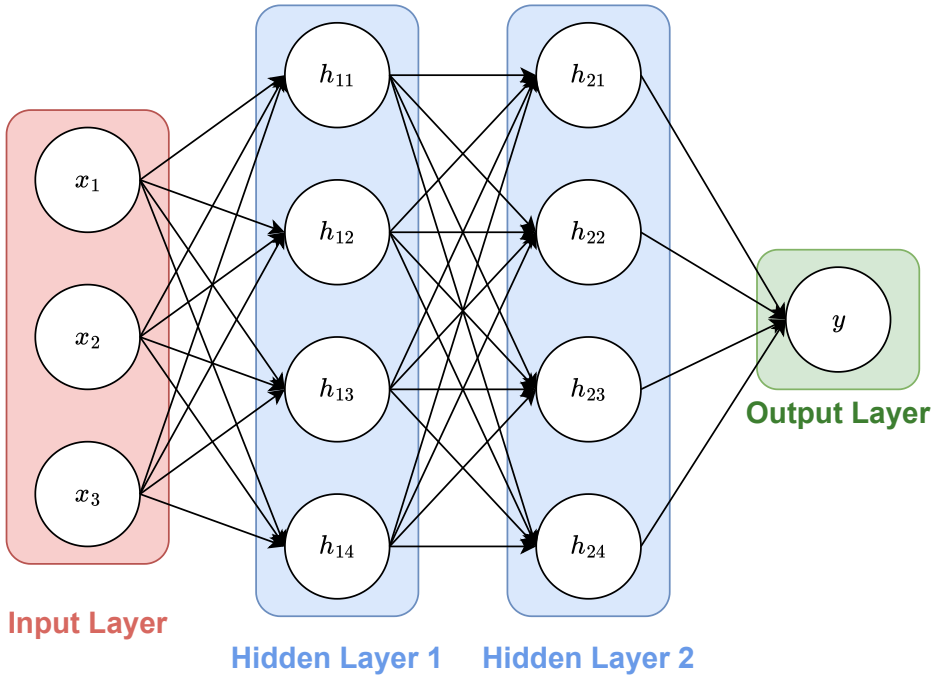


Figure 3.1: MLP Structure

The efficacy of MLPs derives from their non-linear activation functions, often realized using sigmoid functions. The predominant variation used is the unipolar (logistic) sigmoid:

$$f(s) = \frac{1}{1 + e^{-s}} \tag{3.2}$$

and the bipolar sigmoid (hyperbolic tangent):

$$f(s) = \frac{1 - e^{-as}}{1 + e^{-as}} \tag{3.3}$$

where a is a slope parameter. A widely used activation function is the Gaussian Error Linear Unit (GELU):

$$f(s) = \frac{s}{2}\left[1 + \text{erf}\left(\frac{s}{\sqrt{2}}\right)\right] \tag{3.4}$$

where erf is the error function. The activation functions are shown in Figure 3.2.

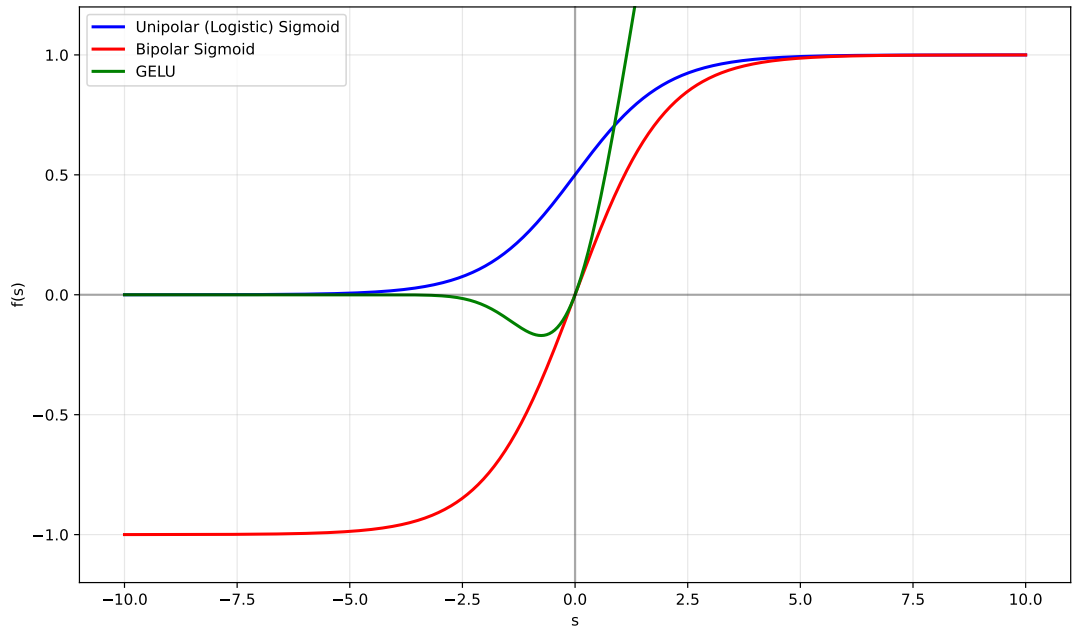


Figure 3.2: Activation functions

The network's capacity to learn intricate patterns arises from its layered architecture in conjunction with non-linear activation functions. Multilayer perceptrons with a solitary hidden layer has the capability to estimate any continuous function, hence qualifying them as universal approximators [60]. The attributes of this feature, together with their trainability via backpropagation, render MLPs very efficacious for many applications, including pattern recognition, function approximation, and time series forecasting. Training a multilayer perceptron entails modifying the network weights to reduce the discrepancy between the expected and actual outputs, usually achieved by the backpropagation technique, which effectively calculates the gradient of the error for each weight in the network. This training approach, albeit computationally demanding, allows the network to acquire intricate mappings directly from instances without necessitating explicit rule programming.

3.2 Proposed Method

3.2.1 Theoretical Foundation: Frequency Domain Transformation

Mathematically, this approach relies on spectrum analysis methods as the Discrete Fourier Transform (DFT). According to the TSLANet framework [4], the DFT is a vital component of the design. Moreover, given that their frequency-domain fingerprints essentially describe chatter occurrences, this model offers the optimal methodological approach for study and detection of chatter.

Think about a discrete time-domain sequence $s[n]$ with L samples, $n \in \{0, 1, ..., L-1\}$. DFT converts this sequence into the frequency domain, hence producing the spectral representation $S[k]$, defined as

$$S[k] = \sum_{n=0}^{L-1} s[n] e^{-j2\pi kn/L} = \sum_{n=0}^{L-1} s[n] W_L^{kn} \tag{3.5}$$

Here, j denotes the imaginary unit $\sqrt{-1}$, while $W_L = e^{-j2\pi/L}$ represents the main L-th root of unity. The transformation is L-periodic and the resultant spectrum $S[k]$ characterizes the frequency components of the signal at $\omega_k = 2\pi k/L$. The completeness of this transformation allows the Inverse Discrete Fourier Transform (IDFT) to perfectly recreate the original signal, expressed as

$$s[n] = \frac{1}{L} \sum_{k=0}^{L-1} S[k] e^{j2\pi kn/L} \tag{3.6}$$

Because frequency-domain representations conveniently capture the periodic patterns that define chatter occurrences, the spectral characterization of machining chatter is a key aspect in vibration analysis. Deep learning architectures with time-frequency domain transformations thus provide significant analytical advantages as they facilitate the extraction and understanding of the complex spectrum correlations stored in vibration signals. Such methodological approaches improve detection accuracy and interpretability across a broad variety of machining environments by enabling models to gain frequency-dependent correlations and resonant patterns that act as diagnostic indicators of chatter conditions.

3.2.2 Data Preprocessing

The effectiveness of our chatter detection method is greatly influenced by the appropriate preprocessing of acceleration data gathered during milling. A methodical preprocessing pipeline that handles the temporal and amplitude characteristics of the vibration data transforms raw signals into a format appropriate for our deep learning model.

3.2.2.1 Windowing Technique

The sliding window technique separates the ongoing acceleration signals into fixed-length segments suitable for analysis. This method allows us to manage data streams of arbitrary length while maintaining the input dimensions for the neural network model.

An acceleration signal $\mathbf{x} \in \mathbb{R}^n$ with n samples has w overlapping windows extracted:

$$\mathbf{X}_i = \mathbf{x}[i \cdot s : i \cdot s + l], \quad i \in \{0, 1, ..., w-1\} \tag{3.7}$$

l is the window length (1500 samples), s is the stride between subsequent windows (750 samples), and $w = \lfloor (n-l)/s \rfloor + 1$ is the total number of extracted windows. The 50% gap between adjacent frames allows for continuity in the research and offers sufficient temporal resolution to capture the transition from steady cutting to chatter situations.

3.2.2.2 Per-Window Normalization

An independent min-max normalization is applied to each window to reflect the different amplitude ranges that set different machining circumstances and sensor sites apart. While standardizing the input range for the neural network, this approach maintains the relative patterns within each window.

Each window $\mathbf{X}_i \in \mathbb{R}^{l \times c}$, where c is the number of channels (single-channel acceler-

ation data in our implementation), we calculate:

$$\mathbf{X}_i^{\text{norm}} = 2 \cdot \frac{\mathbf{X}_i - \min(\mathbf{X}_i)}{\max(\mathbf{X}_i) - \min(\mathbf{X}_i)} - 1 \tag{3.8}$$

This change avoids potential numerical problems during model training and preserves signal features by scaling each window to the range $[-1, 1]$. A default range value of 1.0 is used to avoid division by zero when a window contains constant values ($\max(\mathbf{X}_i) = \min(\mathbf{X}_i)$). Our SA-TSLANet model for chatter detection is then run on this preprocessed dataset.

3.2.3 Model Architecture

Figure 3.3 shows the suggested model architecture. Let the input signal be denoted as $\mathbf{X} \in \mathbb{R}^{B \times S \times C}$, where B is the batch size, S the sequence length, and C the number of channels linked with the sensor channels.

3.2.3.1 Patch and Positional Embedding

The input signal $\mathbf{X}$ is first segmented into overlapping patches with a stride of $P/2$, where P is the patch size. Given a patch size P, the signal is reorganized as $\mathbf{X}_p \in \mathbb{R}^{B \times N \times C \times P}$, with $N = \lfloor (S - P)/(P/2) \rfloor + 1$ indicating the total number of patches, where S is the sequence length. This overlapping strategy allows the model to capture transitions between adjacent signal segments more effectively. A linear transformation then maps every patch to a higher-dimensional space:

$$\mathbf{E} = \text{LayerNorm}(\mathbf{X}_p \mathbf{W}_E + \mathbf{b}_E) + \mathbf{P}_E \tag{3.9}$$

where $\mathbf{E} \in \mathbb{R}^{B \times N \times C \times D}$ and $\mathbf{W}_E \in \mathbb{R}^{P \times D}$ represents the embedding matrix, $\mathbf{b}_E$ denotes the bias term, D signifies the hidden dimension, and $\mathbf{P}_E$ refers to the positional embedding.

3.2.4 Spindle-Aware Adaptive Spectral Block

In machining processes, the Spindle-Aware Adaptive Spectral Block (SA-ASB) effectively separates and analyzes chatter and regular cutting vibrations. Figure 3.3 depicts the architecture of the suggested method. The embedded representation $\mathbf{E} \in \mathbb{R}^{B\times N\times C\times D}$ is reshaped to $x \in \mathbb{R}^{B\times L\times M}$ prior to spectral processing, where B is the batch size, L is the sequence length (derived from concatenating the patch embeddings), and M is the embedding dimension (which integrates the channel information from the original representation). This reshaping aggregates the patch-wise features while preserving the temporal structure needed for frequency analysis. The reshaped signal is then converted to the frequency domain:

$$X_f = \mathcal{F}[x] \in \mathbb{C}^{B\times L_f\times M} \tag{3.10}$$

The FFT process is represented as $\mathcal{F}[\cdot]$, and L_f is the count of frequency bins. The SA-ASB employs a parallel processing architecture including two key components: adaptive masking and tailored frequency filtering for both the spindle, chatter and global components.

3.2.4.1 Adaptive Masking

An adaptive thresholding technique is used, as detailed in [4], to find relevant frequency components in the signal spectrum.

$$A = \frac{|X_f|^2}{\text{median}(|X_f|^2) + \epsilon} \tag{3.11}$$

where A is the relative energy of every frequency in proportion to the median energy. Each processing route uses a different adaptive threshold.

$$M_{high} = \sigma(A - \tau) \tag{3.12}$$

$$X_{f,masked} = X_f \odot M_{high} \tag{3.13}$$

where τ is a learnable threshold value, σ is the sigmoid function that generates a soft mask, and $\odot$ denotes element-wise multiplication.

This adaptive masking process creates specialized high-pass filters for each processing route:

$$M_{high,spindle} = \sigma(A_{spindle} - \tau_{spindle}) \tag{3.14}$$

$$M_{high,chatter} = \sigma(A_{chatter} - \tau_{chatter}) \tag{3.15}$$

$$M_{high,global} = \sigma(A_{global} - \tau_{global}) \tag{3.16}$$

The respective masked frequency representations are then computed as:

$$X_{f_s,masked} = X_f \odot M_{high,spindle} \tag{3.17}$$

$$X_{f_c,masked} = X_f \odot M_{high,chatter} \tag{3.18}$$

$$X_{f_g,masked} = X_f \odot M_{high,global} \tag{3.19}$$

These masked representations isolate the significant frequency components in each route while suppressing noise and irrelevant frequencies, providing cleaner inputs for the subsequent filtering operations.

3.2.4.2 Spindle and Chatter Filtering

After adaptive masking, we apply tailored filters for the spindle and chatter components: The spindle route calls for a spindle-frequency mask built from the known spindle frequency f_s:

$$M_s(f) = \sum_{h=1}^{H} \exp\left(-\frac{(f - h \cdot f_s)^2}{2\sigma_s^2}\right) \tag{3.20}$$

where H is the largest harmonic number, and σ_s sets the bandwidth around harmonics. The complementing mask is used for the chatter path.

$$M_c = 1 - M_s \tag{3.21}$$

The adaptively masked frequency representation is then applied using these masks.

$$X_{f,s} = X_{f_s,masked} \odot M_s \tag{3.22}$$

$$X_{f,c} = X_{f_c,masked} \odot M_c \tag{3.23}$$

$$\tag{3.24}$$

This operation applies the frequency masks to the masked spectral representations, effectively separating the signal components. $X_{f,s}$ isolates the spindle-related frequency components by preserving only the harmonics of the spindle frequency, while $X_{f,c}$ isolates the chatter-related components by preserving everything except the spindle frequency harmonics.

3.2.4.3 Learnable Spectral Weighting

Every filtered component is handled with distinct learnable complex weights.

$$X'_{f,s} = W_s \odot X_{f,s} \tag{3.25}$$

$$X'_{f,c} = W_c \odot X_{f,c} \tag{3.26}$$

$$X'_{f,g} = W_g \odot X_{f_g,masked} \tag{3.27}$$

where W_s, W_c, and W_g denote learnable complex weights for spindle components, chatter components, and global spectrum, respectively.

3.2.5 Interactive Convolutional Block

After enhancing feature representation through the SA-ASB, the Integrated Convolutional Block (ICB) is implemented with a dual-path convolutional architecture as illustrated in the model diagram. The ICB design incorporates parallel convolutions with varying kernel sizes to extract both fine-grained patterns and broader contextual information.

The structure features two primary convolutional pathways: the first utilizes a smaller kernel (size 1) aimed at capturing precise, localized signal characteristics, while the second employs a larger kernel (size 3) to identify extended temporal dependencies. What distinguishes this approach is how these pathways interact—each pathway's output modulates the feature extraction of the other through element-wise multiplication, enabling rich cross-scale interactions.

Mathematically, with input x from the spectral processing, the ICB operations proceed as follows:

$$x_1 = \text{Conv}_1(x) \tag{3.28}$$

$$x_2 = \text{Conv}_3(x) \tag{3.29}$$

$$x_{1a} = \text{GELU}(x_1) \tag{3.30}$$

$$x_{2a} = \text{GELU}(x_2) \tag{3.31}$$

$$\text{out}_1 = x_1 \odot x_{2a} \tag{3.32}$$

$$\text{out}_2 = x_2 \odot x_{1a} \tag{3.33}$$

The resulting features are combined and processed through a final convolutional layer:

$$x_{\text{output}} = \text{Conv}_{\text{final}}(\text{out}_1 + \text{out}_2) \tag{3.34}$$

This interactive design encourages comprehensive feature extraction across different temporal scales, which is particularly valuable for detecting subtle patterns in vibration signals where chatter manifestations may appear at various time scales.

3.3 Training Objective

3.3.1 Multi-Objective Loss Function

A multi-objective loss function covering several aspects of chatter detection is suggested to correctly steer the learning process of the model.

$$\mathcal{L}_{\text{spindle}} = \frac{1}{B} \sum_{b=1}^{B} \left\| \frac{|X'_{f,g}|^2}{\sum |X'_{f,g}|^2} - \frac{|X'_{f,s}|^2}{\sum |X'_{f,s}|^2} \right\|^2 \tag{3.35}$$

The spindle loss $\mathcal{L}_{\text{spindle}}$ is derived from the energy ratio concept Kuljanic et al. [12] devised for chatter detection. While their method detected unstable cutting conditions using the ratio E_c/E, the suggested methodology measures the alignment of the global frequency representation and spindle frequency components during stable

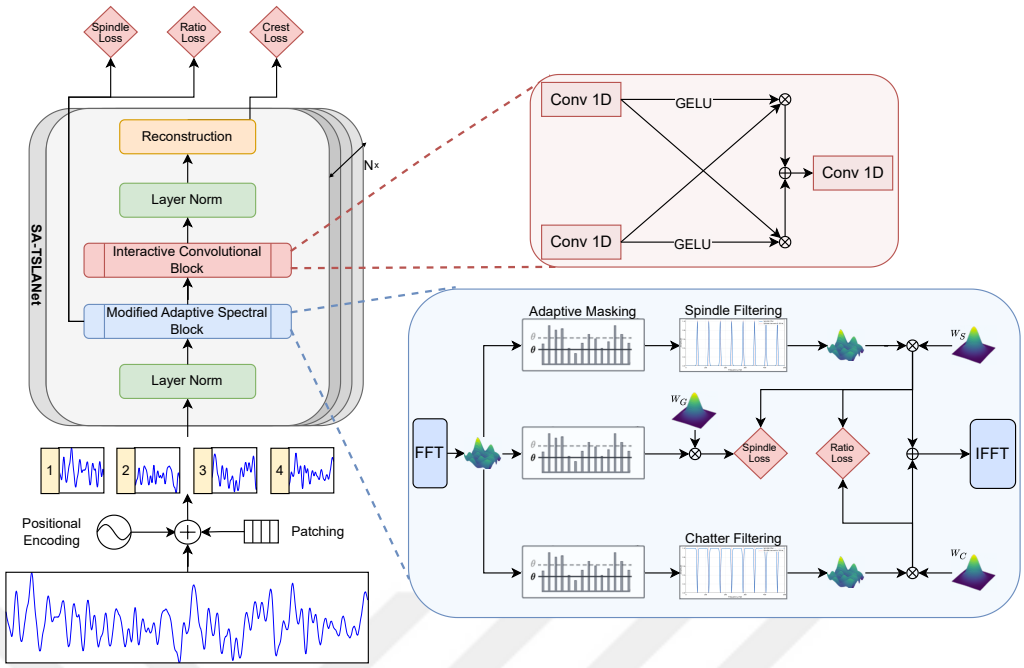


Figure 3.3: The design of the suggested SA-TSLANet is shown. Positional embeddings are applied to the input time series divided into segments. TSLANet layers then handle the output embeddings; each layer has two basic components. The first part of the Modified Adaptive Spectral Block is a frequency domain representation with spindle and chatter filtering that efficiently extracts features. Noise is also reduced by use of adaptive thresholding. The second part is the Interactive Convolution Block, which uses convolutional methods to seize sophisticated temporal patterns.

machining. To emphasize frequency distribution patterns over absolute magnitudes, both spectra are modified by their total energies. This method provides a different perspective to conventional chatter detection tools by helping the model to correctly identify and preserve the spectral fingerprints linked with normal spindle operation.

$$\mathcal{L}_{\text{ratio}} = \frac{1}{B} \sum_{b=1}^{B} \frac{\sum |X'_{f,c}|^2}{\sum |X'_{f,s}|^2 + \epsilon} \tag{3.36}$$

Ratio loss Caliskan's concept of chatter energy ratio, developed for real-time chatter detection in milling operations, is the basis of $\mathcal{L}_{\text{ratio}}$. Caliskan established that by comparing the energy in chatter frequencies to spindle-related frequencies, a valid

chatter detection metric could be built without the requirement for sophisticated system dynamics modeling. The suggested method converts this concept into a differentiable loss function for training neural networks. The ratio assesses the relative energy distribution throughout a batch of B samples between chatter components X_{c} and spindle-related frequency components X_{s}. Spindle frequencies transport most of the signal energy under constant cutting conditions, which leads to a low ratio value. Chatter causes energy to migrate toward non-spindle frequencies, hence greatly increasing this ratio. Including this physics-informed link into the training objective helps the model to differentiate between normal cutting vibrations and chatter conditions depending on their different spectral energy distributions.

$$\mathcal{L}_{\text{crest}} = \frac{1}{B} \sum_{b=1}^{B} \left(\frac{P_{\hat{x}}}{RMS_{\hat{x}} + \epsilon} - \frac{P_x}{RMS_x + \epsilon} \right)^2 \tag{3.37}$$

As a consistent indicator for chatter detection, the crest factor loss $\mathcal{L}_{\text{crest}}$ improves the statistical time-domain feature analysis method found by Mishra et al. [61]. Although their research indicated that peak-to-peak values were the most efficient characteristic (0.28), the crest factor—the ratio of peak amplitude to RMS value—also shown significant discriminative power for distinguishing between stable and unstable cutting conditions. This loss computes the squared difference between the crest factors of the reconstructed signal $\hat{x}$ and the original signal x over a series of B samples. Maintaining this vital time-domain characteristic helps the model to be sensitive to the impulsive character of chatter vibrations, which are usually shown as higher peak-to-RMS ratios than stable cutting conditions. This physics-based limitation ensures that the model captures the temporal patterns that experienced machinists perceive as early indicators of instability and prevents it from depending just on spectral properties.

$$\mathcal{L}_{\text{total}} = a_{\mathrm{s}} \cdot \mathcal{L}_{\text{spindle}} + a_{\mathrm{r}} \cdot \mathcal{L}_{\text{ratio}} + a_{\mathrm{c}} \cdot \mathcal{L}_{\text{crest}} \tag{3.38}$$

Overall Loss $\mathcal{L}_{\text{total}}$ balances the three components' relative importance by combining them with weighting factors a_{s}, a_{r}, and a_{c}. This multi-objective approach guarantees that the model improves time-domain signal characteristics, chatter component

energy distribution, and spindle component spectral alignment all at the same time. Under a range of machining conditions, the combined effect of these loss components lets the model effectively differentiate normal cutting vibrations from chatter vibrations.

3.4 Hyperparameter Optimization

A methodical hyperparameter optimization approach utilizing Optuna, a Bayesian optimization framework, was used to optimize the performance of the chatter detection model. This method reduced computer resources use while allowing quick searches throughout a high-dimensional parameter space.

3.4.1 Optimization Setup

The optimization aimed at six important hyperparameters that greatly affect model performance:

- Embedding dimension (`emb_dim`): Controls the representational capacity of the model
- Patch size (`patch_size_detect`): Determines the temporal granularity of feature extraction
- Network depth (`depth`): Number of transformer encoder layers
- Dropout rate (`dropout`): Regularization strength to prevent overfitting
- Padding factor (`padding_factor`): Controls frequency resolution in spectral processing
- Learning rate (`lr`): Controls optimization step size
- Batch size: Number of samples processed simultaneously during training

Optuna was set up to employ the Tree-structured Parzen Estimator (TPE) sampler, which adaptively concentrates the search on promising areas of the parameter space.

Every trial assessed the performance of the model using validation loss as the goal function to reduce.

3.4.2 Optimization Results

The outcomes of the hyperparameter optimization project are shown in Table 3.1, which reveals the top 10 configurations found by Optuna.

Table 3.1: Optuna hyperparameter optimization results for the chatter detection model

Trial	Embedding Dimension	Patch Size	Depth	Dropout	Learning Rate	Padding Factor	Value
9	64	4	2	0.400	3e-4	8	0.008
8	64	1	6	0.002	1e-4	9	0.078
7	64	8	2	0.097	2e-4	32	0.077
6	64	4	4	0.125	2e-5	8	0.268
5	128	16	4	0.124	8e-4	8	0.030
4	64	4	3	0.169	8e-4	16	0.041
3	128	16	1	0.024	1e-4	4	0.029
2	96	4	2	0.290	3e-4	16	0.102
1	32	16	2	0.098	4e-5	32	0.487
0	256	8	3	0.257	4e-4	16	0.068

With embedding dimension 64, patch size 4, network depth 2, dropout rate 0.40, padding factor 0.000254, and batch size 8, Trial 9 produced the best validation loss of 0.007599. This is a significant increase over the first hand-tuned setup, which had a validation loss of 0.267658. The findings of the optimization show some intriguing trends. Consistently outperforming bigger ones, lower embedding dimensions (64) imply that the issue benefits from a more compact depiction. Moderate patch sizes (4) also outperformed very tiny or big patches, so striking a balance between local feature extraction and enough context. Especially, the comparatively high ideal dropout rate (0.40) suggests that robust regularizing was good for generalization. Among top performers, the learning rate varied significantly, hence stressing the need of precisely adjusting this parameter for any particular arrangement. The configuration from Trial

9 was chosen for the final model design based on these results as it showed better performance in both validation and later testing stages.

3.4.3 Model Implementation and Training Details

Architectural parameters chosen for chatter detection guided the SA-TSLANet model's construction in PyTorch. Comprising two stacked layers, the model architecture included the Interactive Convolution Block and the Spindle-Aware Adaptive Spectral Block on each layer. A 64-dimensional embedding was used across the network to balance representational capacity with computational performance. We used a patch size of 4 samples and a stride of 2 samples for patch-based processing of input signals, which provided sufficient overlap to capture transitional features. The FFT calculation's padding factor was set at 8 to provide reasonable frequency resolution and maintain computational performance. Training utilized a 0.1 dropout rate to guarantee regularization.

To prevent overfitting, the AdamW optimizer was used with a starting learning rate of 3×10^{-4} and weight decay of 1×10^{-6}. A ReduceLROnPlateau scheduler was used to automatically change the learning rate by 0.1 when validation loss plateaued for two consecutive epochs. Spectral energy normalization causes the spindle loss component to produce values several orders of magnitude lower than the other parts. Appropriate weighting factors (a_s, a_r, and a_c) were used to offset the contributions of each loss component, hence guaranteeing that no one term ruled the optimization process. With a batch size of 6, the model was trained for 20 epochs. 20% of the training data was kept for validation. With the whole training time of 9 minutes, all tests were conducted on an NVIDIA T4 GPU available via Google Colab.

3.5 Result with Benchmark Dataset for Deep Learning Models

The outcomes of the proposed model for the benchmark dataset are shown in this part. Detection model selects datasets independently. Anomaly detection makes use of Swat. Since it is a well-known benchmark in the domain of industrial control system security, the SWaT (Secure Water Treatment) dataset was selected for assess-

ing anomaly detection. The data came from a tiny water treatment testbed meant to mimic real water treatment processes, including both routine operations and cyber-attacks. Its multidimensional nature and the subtle representations of various anomalies that closely mimic normal operational variations cause the dataset to be rather difficult. Anomaly detection tasks are the focus of this study, which reimplements the TSLANet architecture [4]. The comparison results shown in Table 3.2 reveal the relative effectiveness of many approaches. With an F1-score of 94.23%, GPT4TS is followed closely by LightTS at 93.33% and FEDformer at 93.19%. The reimplemented model shows a fair performance with 91.30% precision, 92.05% recall, and 91.68% F1-score. Though there is a notable performance difference compared to the state-of-the-art models, this puts it competitively within transformer-based approaches. Especially, while the Anomaly Transformer and LogTransformer have great recall rates (97.32% and 97.31%, respectively), their accuracy ratings are much lower (72.51% and 68.67%), suggesting a tendency for false positives. By comparison, models like PatchTST and ETSformer have better accuracy but lower recall, suggesting a more conservative approach in anomaly detection. Often beneficial in actual scenarios where both false positives and false negatives generate operational expenses, the reimplemented model shows a more fair trade-off between accuracy and recall. Figure 3.4 shows the reconstruction error of the reimplemented model.

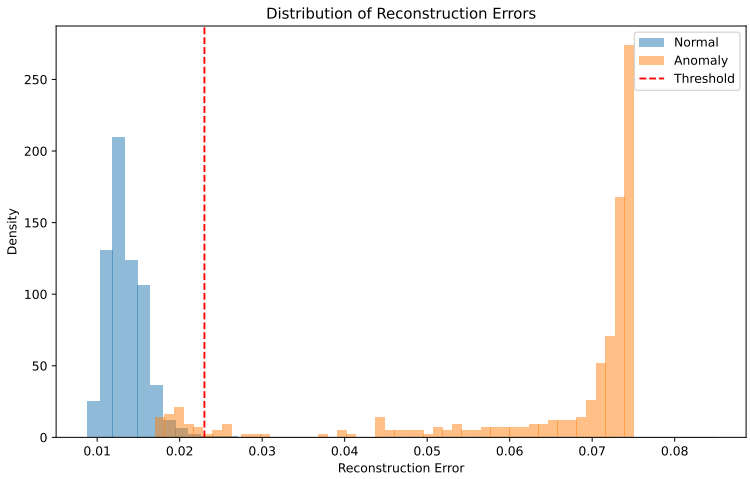


Figure 3.4: Reconstruction error distribution for SWAT dataset with reimplemented model

Table 3.2: Model performance comparison

Method	**Precision**	**Recall**	**F1-Score**
TSLANet	91.50	94.14	92.80
GPT4TS	92.20	96.34	94.23
TimesNet	90.75	95.40	93.02
PatchTST	91.10	80.94	85.72
ETSformer	90.02	80.36	84.91
FEDformer	90.17	96.42	93.19
LightTS	91.98	94.72	93.33
DLinear	80.91	95.30	87.52
Stationary	68.03	96.75	79.88
Autoformer	89.85	95.81	92.74
Pyraformer	87.92	96.00	91.78
Anomaly Transformer	72.51	97.32	83.10
Informer	70.29	96.75	81.43
Reformer	72.50	96.53	82.80
LogTransformer	68.67	97.31	80.52
Transformer	68.84	96.53	80.37
Reimplemented	**91.30**	**92.05**	**91.68**

CHAPTER 4

EXPERIMENTAL ANALYSIS

As seen in Figure 4.1, the experimental configuration included a modified Deckel FP5cc CNC machine tool. With spindle motor operations handled by a Beckhoff AX5125 (25A) driver, the control system's main controller was a Beckhoff PC C6930 industrial computer. Figure 4.1 depicts the placement of a single PCB 35C22 accelerometer on the machine to track vibration. To guarantee high-resolution signal collection, all experimental data were gathered at a sampling rate of 25600 Hz. The testing parameters, comprising spindle speeds ranging from 20.00 to 58.33 rpm, varying flute counts (3-4), and diverse depth settings, as stated in Table 4.1, were systematically executed to assess machining performance throughout numerous cutting situations. The dynamic response of the machine throughout the cutting process was accurately monitored by this comprehensive system.

4.1 Experimental Setup and Data Collection

4.1.1 Experimental Parameters

To investigate the occurrence of chatter, experimental data was collected under various milling conditions.

Table 4.1 displays the comprehensive dataset including the following parameters:

- Spindle Speed (SF): Operating frequency of the spindle in Hz
- Flute Count (N): Number of cutting teeth on the tool
- Tooth Passing Frequency: Events per revolution, calculated as SF × N

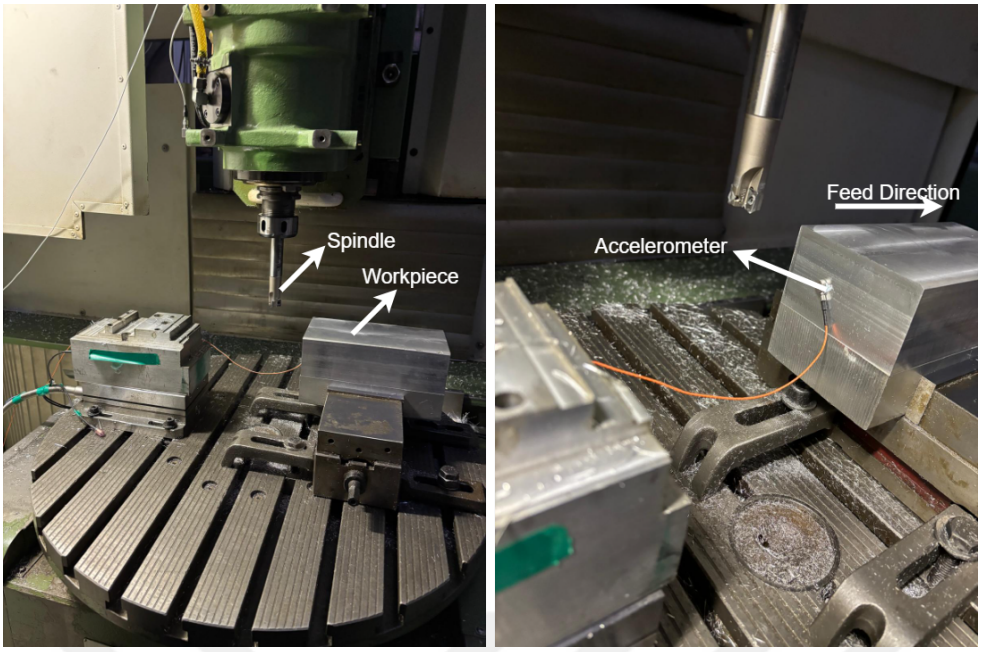


Figure 4.1: Experimental setup showing the modified Deckel FP5cc CNC machine tool

- Feed Rate: Feed Rate in mm/s
- Feed per Tooth: Distance a cutting tool advances per tooth
- Tool Diameter: Cutting tool diameter in mm
- Axial Depth: Cutting depth in axial direction in mm
- Radial Depth: Initial cutting depth in radial direction in mm
- Final Radial Depth: Final cutting depth in radial direction in mm

The last two columns denote:

- Chatter Exists: Binary indicator of chatter presence (Yes/No)
- Data Type: Binary indicators for dataset split (Training/Test)

Comprising 10 experimental trials under various cutting conditions. While four examples include the existence of chatter, six samples exhibit stable cutting conditions (no chatter). This fair allocation enables quick assessment and training of the suggested chatter detection method.

Table 4.1: Test parameters

Test No.	Speed (rpm)	Flute Count	TPF (Hz)	Feed (mm/min)	Feed/tooth (mm/tooth)	Diameter (mm)	Axial (mm)	Radial (mm)	Chatter Exists	Data Type
600	3500	3	175.0	1000	0.095	20	0.2	2.0/20	No	Train
601	3500	3	175.0	1000	0.095	20	0.3	2.0/20	No	Train
602	3500	3	175.0	1000	0.095	20	0.4	2.0/20	No	Train
603	3500	3	175.0	1000	0.095	20	0.5	2.0/20	No	Train
604	3500	3	175.0	1000	0.095	20	0.6	2.0/20	No	Train
605	3500	3	175.0	1000	0.095	20	0.7	2.0/20	Yes	Test
606	3500	3	175.0	1000	0.095	20	0.8	2.0/20	Yes	Test
625	3800	3	190.0	1000	0.095	20	1.0	0.1/20	Yes	Test
627	1200	3	60.0	700	0.194	20	0.5	1.0/19	No	Test
642	3300	4	220.0	1320	0.100	16	1.2	0.0/16	Yes	Test

4.1.2 Data Labeling Methodology

Frequency domain analysis classified the experimental data to determine chatter conditions. Acceleration data was processed using the Fast Fourier Transform (FFT), and the resulting spectra were analyzed in respect to spindle harmonics and tooth passing (TP).

Two example scenarios are used to show the labeling method in Figure 4.2. Stable cutting conditions (Experiment 602, Figure 4.2a) show the frequency content exactly matching the expected tooth passing and spindle harmonics, indicating suitable cutting behavior. With little energy between the harmonic frequencies, the spectrum shows clear peaks at the expected frequencies.

On the other hand, unstable cutting conditions (Experiment 606, Figure 4.2b) show significant frequency content different from the expected harmonics. Especially in the 800-900 Hz frequency range, the misalignment is suggestive of chatter vibrations. Harmonics with noticeable frequency components indicate self-excited vibrations out of rhythm with tool rotation or tooth passing events.

The experimental data was meticulously examined and classified using a comprehensive signal processing technique integrating frequency-domain translation with

advanced peak detection and harmonic analysis capabilities. By way of a measurement of the correlation between observed spectral peaks and predicted spindle-related harmonics, this approach produces an objective chatter score for classification. This automated approach ensures consistent labeling throughout all tests and records the physical characteristics distinguishing steady cutting from chatter conditions. The method operates a sequential pipeline of signal modification, frequency analysis, and harmonic comparison as shown in the following sections.

4.1.2.1 Fast Fourier Transform

The discrete Fourier transform is calculated for both channels for a segment of N samples.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-2\pi i k n/N} \tag{4.1}$$

where x_n designates the time-domain signal, while X_k represents its frequency-domain equivalent.

4.1.2.2 Frequency Analysis

The frequency vector f is calculated using:

$$f_k = k\frac{f_s}{N}, \quad k = 0, 1, ..., \frac{N}{2} - 1 \tag{4.2}$$

where f_s is the sampling rate, set at 25.6 kHz in this implementation.

4.1.2.3 Magnitude Computation

The magnitude spectra for channel is calculated and normalized:

$$|X_k| = \frac{2}{N}|X_k| \tag{4.3}$$

The aggregate magnitude spectrum is calculated using the RMS.

$$M_k = \sqrt{|X_k|^2} \tag{4.4}$$

4.1.2.4 Peak Detection

Peaks in the summed magnitude spectrum are determined using the following criteria:

- Minimum peak height: $h_{min} = 0.1 \max(M_k)$
- Minimum peak distance: $d_{min} = \max(1, \lfloor \frac{B}{\Delta f} \rfloor)$

where B denotes the bandwidth parameter (6 Hz) and Δf represents the frequency resolution.

4.1.2.5 Harmonic Analysis

For a specified spindle frequency f_s, harmonics are calculated up to a maximum frequency f_{max}:

$$f_{h,i} = i f_s, \quad i = 1, 2, ..., \lfloor \frac{f_{max}}{f_s} \rfloor \tag{4.5}$$

4.1.2.6 Chatter Score Computation

For each identified peak at frequency f_p with magnitude m_p, the distance to the closest harmonic is calculated:

$$d_h = \min_i |f_p - f_{h,i}| \tag{4.6}$$

If $d_h > B$, the peak influences the chatter score based on its relative magnitude:

$$s_p = \frac{m_p}{\max(M_k)} \tag{4.7}$$

The highest possible chatter score is calculated as follows:

$$S = \min(1, \max_{p \in P} s_p) \tag{4.8}$$

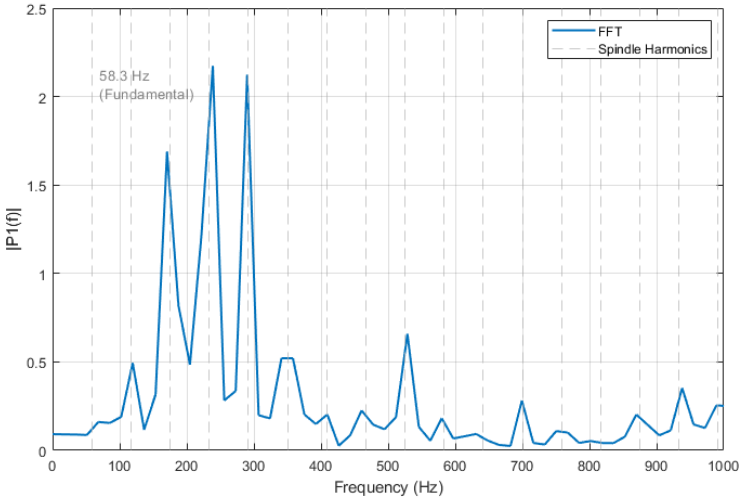
where P denotes the set of qualifying non-harmonic peaks.

All experimental datasets were meticulously analyzed using this frequency domain technique to provide ground truth labels. A threshold value of $S_{th} = 0.80$ for binary classification was set based on thorough prior testing.

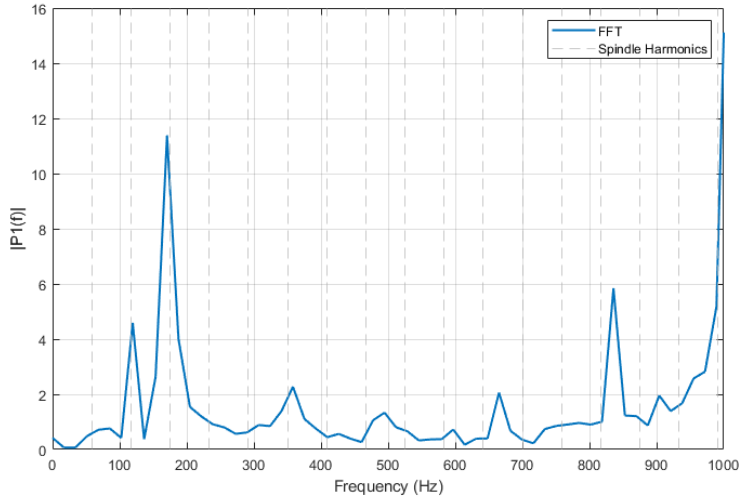
$$\text{Cutting Condition} = \begin{cases} \text{Chatter} & \text{if } S \geq S_{th} \\ \text{Stable} & \text{if } S < S_{th} \end{cases} \tag{4.9}$$

While those beyond the threshold showed significant spectrum deviations and were classed as chatter situations, those below the threshold showed clear agreement with theoretical harmonics and were classified as steady cutting conditions. The threshold was chosen to optimize discriminating between the two cutting states while preserving robustness across various machining settings. This quantitative labeling approach provides an objective basis for evaluating the effectiveness of the suggested detection system and ensures consistency in the classification of cutting scenarios throughout every test.

For the purposes of establishing ground truth labels, all data proceeding from this initial reference point has been classified as chatter.



(a) FFT spectrum for stable cutting conditions (Experiment 602)



(b) FFT spectrum for unstable cutting conditions (Experiment 606)

Figure 4.2: Frequency domain study contrasting stable and unstable cutting circumstances. The spectrum is examined for tooth passing (TP) harmonics (dashed green lines) and spindle harmonics (dotted green lines).

4.1.3 Model Training Analysis

The performance of the proposed SA-TSLANet model largely relies on the optimization of its specific loss functions during training. A comprehensive study of loss histories uncovers the training dynamics and convergence characteristics of the chatter detection system, which this section handles.

4.1.4 Loss Function Analysis

Aiming at certain areas of the signal qualities that separate normal cutting from chatter circumstances, the SA-TSLANet model uses many complimentary loss functions. Training kept an eye on three main loss components:

- **Crest Factor Loss**: Measures the model's ability to reconstruct the peak-to-RMS ratio of the original signal, which significantly differs between stable cutting and chatter conditions
- **Spindle Loss**: Quantifies the accuracy of spindle harmonic component extraction from the input signal
- **Ratio Loss**: Evaluates the energy separation between spindle and non-spindle (potential chatter) frequency components

These tailored loss functions help the model to acquire distinguishing characteristics that properly distinguish normal cutting circumstances from chatter vibrations. Tracking their convergence patterns allows us to evaluate not only the general training performance but also how effectively the model captures every particular feature of signal characteristics.

The crest factor loss history shown in Figure 4.3 shows a clear and efficient convergence trend during model training. Beginning with rather high beginning values—about 0.3 for validation and 0.14 for training—the loss drops quickly in the first five epochs, indicating the model's fast adaptation to capture the peak-to-RMS features of the signals. Early in the training phase, this sharp initial drop implies that the model efficiently learns to distinguish between the separate crest factor signatures of

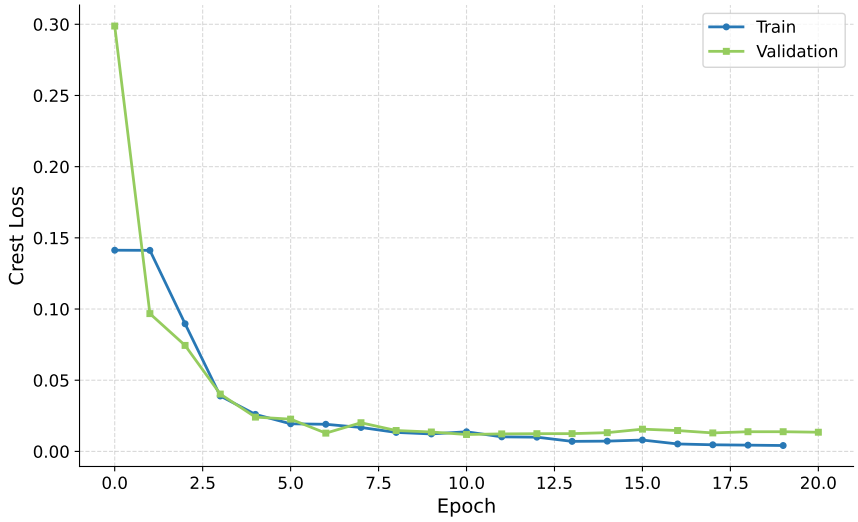


Figure 4.3: Training and validation crest factor loss history over 20 epochs.

normal cutting and chatter circumstances. Both training and validation losses stabilize at about 0.01 at epoch 7.5, suggesting the model has attained a strong solution that generalizes well to unknown data. The tight agreement between training and validation curves after epoch 5 reveals little overfitting, hence verifying that the model is acquiring relevant signal features instead than remembering training samples. Particularly noteworthy is this constant low inaccuracy in reconstructing the crest factor as this characteristic directly connects to the impulsive character of chatter vibrations in comparison to the more steady patterns of stable cutting. With few changes seen in the later part of the training phase, the convergence to a steady, low loss value by epoch 10 indicates that 20 epochs is more than enough for training. This quick convergence shows how well the crest factor works as a discriminating feature for chatter detection and how well the model can capture this important signal property.

Spindle loss during model training, which assesses the model's capacity to reliably extract spindle frequency components from the input signal, evolves as shown in Figure 4.4. While the training loss starts at a more reasonable 0.12, the validation loss starts at a somewhat high value of around 0.29. This first disparity implies that the model first finds it difficult to extend its knowledge of spindle harmonics to unknown data. Both curves indicate a steep drop in the first epoch, but the validation loss shows the most noticeable change. By epoch 2.5, both losses had dropped noticeably below 0.05, suggesting fast adaptation of the spindle component extraction capacity of the

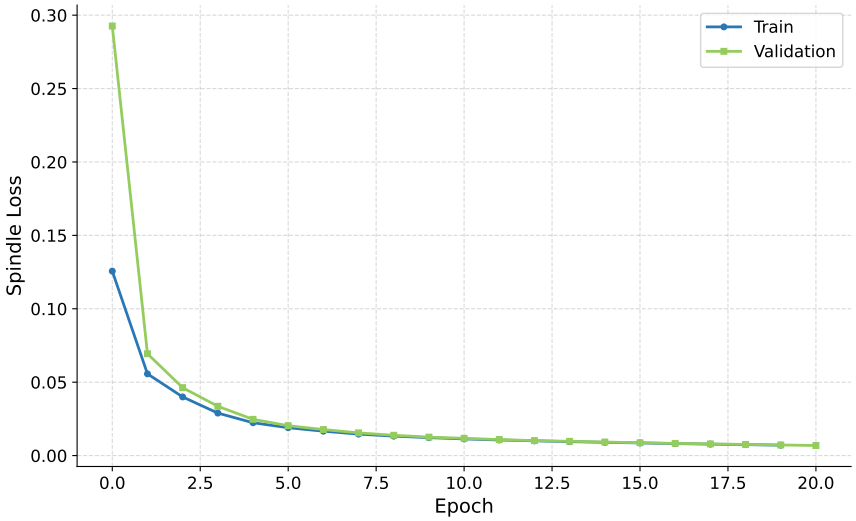


Figure 4.4: Training and validation spindle loss history over 20 epochs.

model. The training and validation curves almost completely match across epochs 2.5 and 7.5, showing that the model generalizes effectively across both datasets. The ongoing slow drop after epoch 7.5 reveals the model's fine-tuning of its spindle harmonic representation, finally converging to near-zero loss values at epoch 12.5. This convergence pattern verifies that the model effectively learns to segregate spindle frequency components—a vital skill for differentiating normal cutting vibrations from chatter signals. The extraordinary consistency of the model's knowledge of spindle harmonics across various operating situations, especially after the first sharp decrease, confirms that it is strong and constant throughout much of the training process. Accurately modeling spindle-related frequency components helps to provide a basis for spotting unusual frequency patterns suggesting chatter.

The ratio loss development during model training shown in Figure 4.5 evaluates the model's capacity to divide energy between spindle and non-spindle (potential chatter) frequency components. Compared to a more modest training loss of 0.4, the very high first validation loss of around 1.8 is the most notable characteristic. This significant early difference suggests that the model's most difficult area to generalize is initially energy separation between spindle and chatter components. Among all loss components, the ratio loss shows the most striking drop; the validation loss drops by an order of magnitude during the first period. Both training and validation losses drop below 0.1 by epoch 1.5, and the curves almost completely converge. Both curves keep

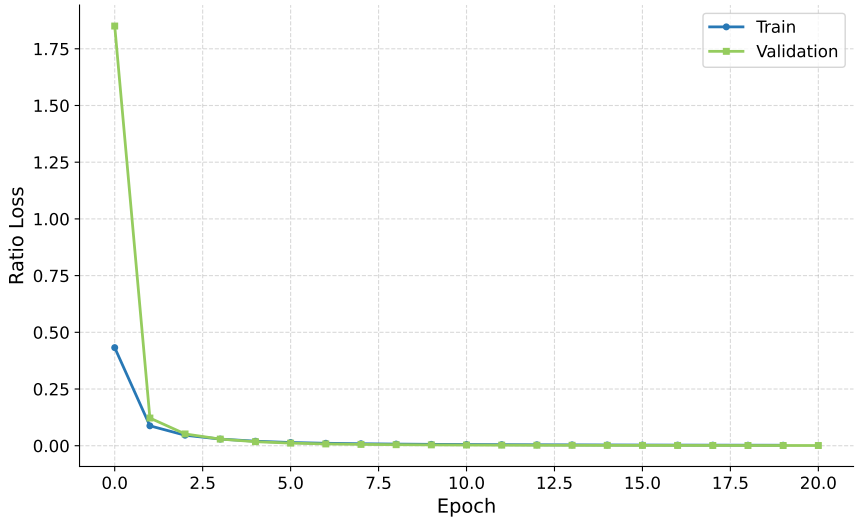


Figure 4.5: Training and validation ratio loss history over 20 epochs.

falling slowly from epoch 2.5 until they approach zero around epoch 5.0. This continuous convergence pattern shows that the model effectively learns to differentiate between the energy distributions typical of regular cutting vs chatter settings. Early in training, the very low loss values imply that the energy ratio between frequency components is a very strong discriminating characteristic for chatter detection. From the highest initial loss to near-perfect convergence with training loss, the significant increase in validation performance emphasizes the efficacy of the model's spectral decomposition approach in capturing generalizable patterns that consistently separate normal machining vibrations from chatter conditions across different operational parameters.

Through its specific loss functions, the SA-TSLANet model's training dynamics are collectively shown by Figures 4.3, 4.4, and 4.5. Every loss component shows unique convergence properties, hence representing various facets of the model's learning process. With the validation loss falling from 1.8 to under 0.1 after just 1.5 epochs, the ratio loss (Figure 4.5) exhibits the most striking early improvement. Though initially the most difficult thing to generalize, the model seems to rapidly understand the energy gap between spindle and chatter components. With both training and validation losses attaining great alignment by epoch 2.5, the spindle loss (Figure 4.4) shows a notable but more gradual improvement throughout early training. This suggests that although generalizing rather well across various cutting settings, proper modeling of

spindle harmonics needs somewhat longer training time than energy ratio characteristics. Among the three components, the crest factor loss (Figure 4.3) shows the most sluggish convergence trend, needing around 7.5 epochs to completely stabilize. This implies that, perhaps because of the difficulty of modeling transient signal qualities, correctly capturing the peak-to-RMS features of vibration signals calls for more thorough training. Across all loss components, some consistent patterns appear:

1. The validation loss consistently starts higher than the training loss but converges rapidly, indicating the model's strong generalization capabilities.

2. All losses show significant improvement within the first 5 epochs, with diminishing returns thereafter.

3. Near-perfect alignment between training and validation losses is achieved in the later stages of training, confirming minimal overfitting.

4. The extremely low final loss values (approaching zero) across all components demonstrate the model's comprehensive ability to capture the multifaceted characteristics that distinguish normal cutting from chatter conditions.

This study verifies that the multi-component loss method properly directs the model to acquire complementary signal traits essential for chatter identification. While the near alignment between training and validation performance confirms the strength of the learned features across different machining circumstances, the fast convergence across all loss components confirms the efficacy of the model architecture and its training procedure.

4.2 Test Data Analysis

The model was tested on test data after training to determine its ability to distinguish between normal cutting and chatter circumstances. The test dataset offered a strong assessment of the model's generalization capacity by including samples from trials not seen during training.

4.2.1 Feature Space Analysis

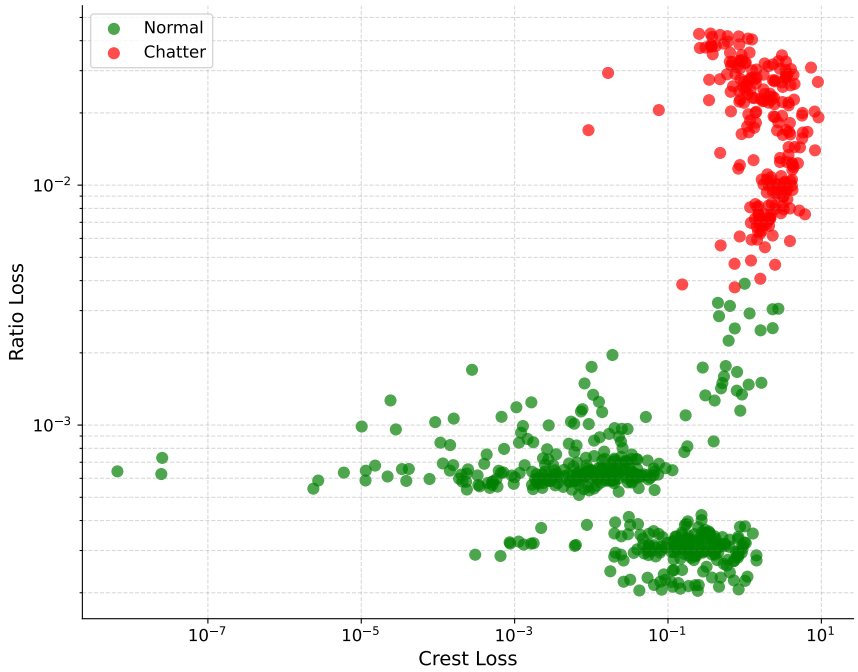


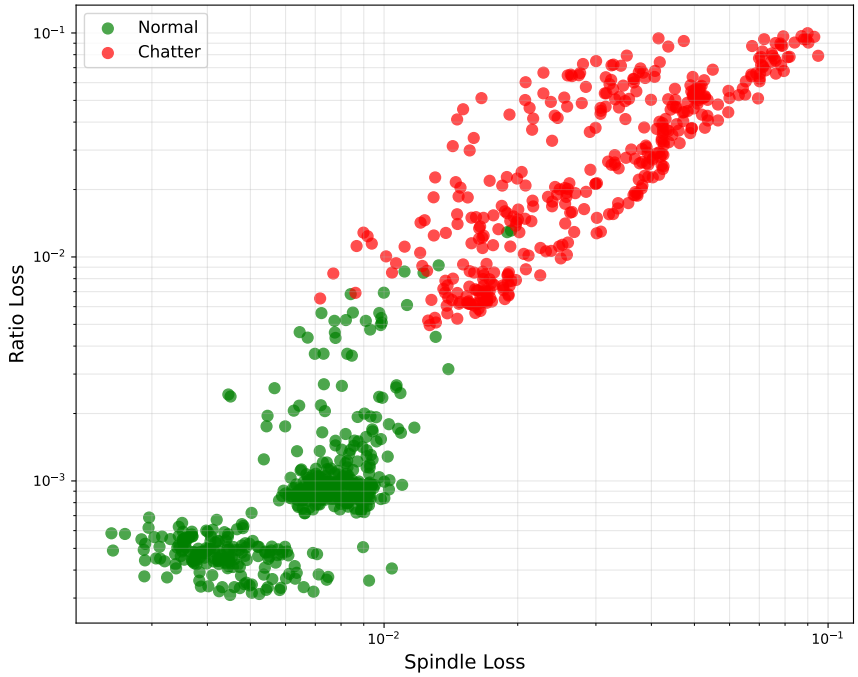
Figure 4.6: Feature space visualization showing the distribution of Crest Loss versus Ratio Loss for normal and chatter samples in the test dataset. Note the logarithmic scale on both axes.

Figure 4.6 presents the feature space distribution of two critical losses from the test dataset: Crest Loss (x-axis) and Ratio Loss (y-axis). This visualization reveals distinct clustering between normal cutting conditions (green) and chatter events (red), demonstrating the discriminative capacity of these learned features. Normal cutting samples exhibit Crest Loss values ranging from 10^{-7} to 10^{1} while maintaining low Ratio Loss values (2×10^{-4} to 2×10^{-3}). In contrast, chatter samples form a distinct cluster in the upper-right quadrant with Ratio Loss values approximately an order of magnitude higher ($> 5 \times 10^{-3}$) and generally elevated Crest Loss values. This vertical separation confirms that energy distribution across frequency components provides a robust discriminator for chatter detection, consistent with the physical understanding that chatter generates significant energy in unstable frequency bands. The logarithmic scale emphasizes the minimal overlap between classes, with the tenfold difference in Ratio Loss between normal and chatter samples providing substantial margin for reliable classification. This separation validates the efficacy of the model's learned repre-

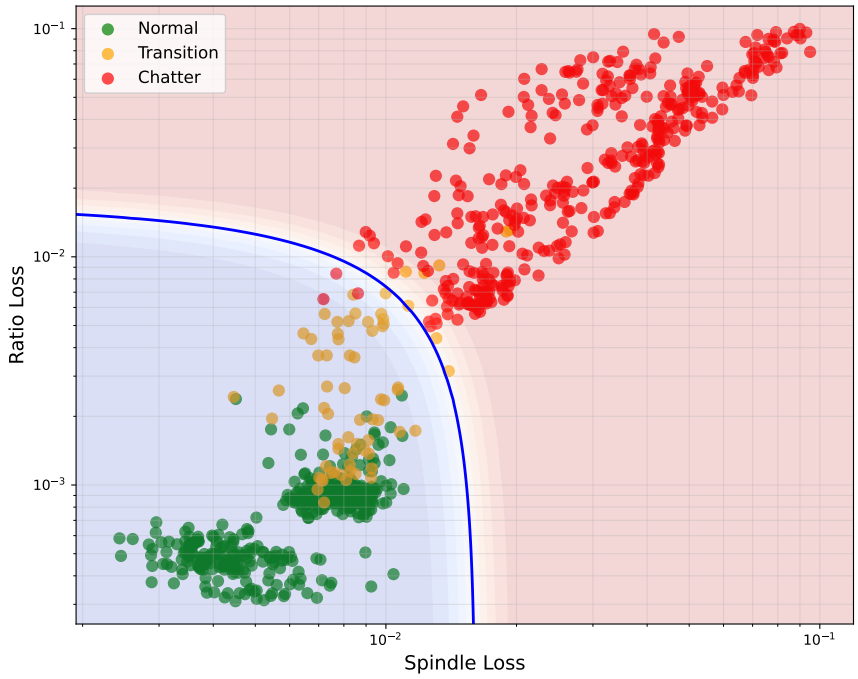
sentations in capturing the fundamental differences between stable cutting and chatter conditions. While Crest Loss offers discrimination in boundary cases, the distribution pattern suggests that Ratio Loss alone could provide an effective threshold-based classifier. This visualization demonstrates the complementary nature of time-domain (Crest) and frequency-domain (Ratio) feature extraction capabilities.

Figure 4.7a illustrates the Spindle Loss versus Ratio Loss feature space, which demonstrates superior discriminative properties. Normal cutting samples form two distinct subclusters in the lower-left region (Spindle Loss $\approx 10^{-2}$, Ratio Loss $\approx 2 \times 10^{-3}$), likely corresponding to different stable operating conditions across experimental parameters. Chatter samples exhibit a diagonal pattern in the upper-right region, indicating a positive correlation between Spindle Loss and Ratio Loss—suggesting that deterioration in cutting stability simultaneously compromises both harmonic extraction and frequency confinement. Figure 4.7b enhances this analysis by incorporating transition points (yellow) representing data captured approximately 0.5 seconds before chatter onset, and a decision boundary (blue curve). These transition points occupy the intermediate region between clusters, validating that the feature space captures the continuous progression from stability to instability rather than merely identifying discrete states. The nonlinear decision boundary effectively separates normal (blue-shaded) from chatter (red-shaded) regions with minimal misclassification. Its curvature suggests a complex relationship between features that reflects the underlying physics of chatter development. This clear separation and the ability to identify transition states highlight the practical utility of this feature space for both detection and prediction of impending chatter, enabling preventative intervention.

The logarithmic distribution of samples indicates that these features maintain discriminative power across multiple scales of vibration intensity, enhancing the method's robustness to varying cutting conditions and tool states. The relationship between Spindle Loss and Ratio Loss during chatter aligns with the physical understanding that chatter involves both degradation of normal spindle harmonics and emergence of self-excited vibrations at non-harmonic frequencies, providing complementary information that improves classification reliability compared to single-feature approaches.



(a) Distribution of normal and chatter conditions in the Spindle Loss versus Ratio Loss feature space



(b) Feature space with transition points and decision boundary

Figure 4.7: Analysis of the Spindle Loss versus Ratio Loss feature space for chatter classification, with logarithmic scales on both axes.

4.3 Support Vector Machine Classification for Chatter Detection

4.3.1 Methodology

A Support Vector Machine (SVM) classifier was used to convert the SA-TSLANet model's discriminative characteristics into an effective chatter detection solution. This methodology utilizes a three-dimensional feature space including Crest Loss, Ratio Loss, and Spindle Loss to provide a resilient classification border between normal and chatter circumstances. Support Vector Machines are very appropriate for this application because of their capacity to identify effective separation borders in high-dimensional feature spaces while minimizing the margin between classes. The SVM classification function may be articulated as:

$$f(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^{n} \delta_i y_i K(\mathbf{x}_i, \mathbf{x}) + b\right) \tag{4.10}$$

Here, $\mathbf{x}$ denotes the input feature vector of loss values, $\mathbf{x}_i$ signifies the support vectors, $y_i \in \{-1, 1\}$ represents the associated class labels, δ_i are the Lagrange multipliers established during training, b is the bias term, and $K(\mathbf{x}_i, \mathbf{x})$ denotes the kernel function. This work used a Radial Basis Function (RBF) kernel to elucidate the nonlinear interactions present in the feature space.

$$K(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\gamma ||\mathbf{x}_i - \mathbf{x}||^2\right) \tag{4.11}$$

where γ regulates the effect radius of each support vector. The likelihood of a sample being classified as chatter is determined by Platt scaling:

$$P(y = 1|\mathbf{x}) = \frac{1}{1 + \exp(Af(\mathbf{x}) + B)} \tag{4.12}$$

where A and B are values fitted on the training data using maximum likelihood estimation. Feature standardization was used before SVM training to provide equal weighting across the many loss dimensions:

$$\mathbf{x}_{scaled} = \frac{\mathbf{x} - \boldsymbol{\mu}}{\boldsymbol{\sigma_d}} \tag{4.13}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\sigma_d}$ are the mean and standard deviation vectors of the feature collection, respectively.

4.3.2 Implementation and Evaluation

A methodical approach to chatter beginning detection included temporal context into the SVM classifier's construction, hence incorporating it. A feature matrix $\mathbf{X}$ was built for every test dataset trial by combining the three loss values:

$$\mathbf{X} = \begin{bmatrix} \text{Crest Loss}_1 & \text{Ratio Loss}_1 & \text{Spindle Loss}_1 \\ \text{Crest Loss}_2 & \text{Ratio Loss}_2 & \text{Spindle Loss}_2 \\ \vdots & \vdots & \vdots \\ \text{Crest Loss}_n & \text{Ratio Loss}_n & \text{Spindle Loss}_n \end{bmatrix} \tag{4.14}$$

where every row denotes a time window from the machining process running sequentially. The posterior probabilities of chatter were recorded for every time frame after feature standardization and SVM classification, therefore generating a continuous measure of chatter probability throughout every trial. The first SVM-predicted chatter for each experiment was retrieved by:

$$t_{onset}(id) = \min\{t_i \mid \text{SVM prediction}(id, t_i) = 1\} \tag{4.15}$$

id is the experiment identification; t_i is the i-th window's timestamp. This method allows for precise temporal location of chatter beginning as well as categorization.

4.3.3 Three-Dimensional SVM Classification

A three-dimensional SVM classifier was built and shown in Figure 4.8 to fully use the complimentary character of all three loss components. Providing insightful analysis of the classification capacity of the model, this all-encompassing picture shows the

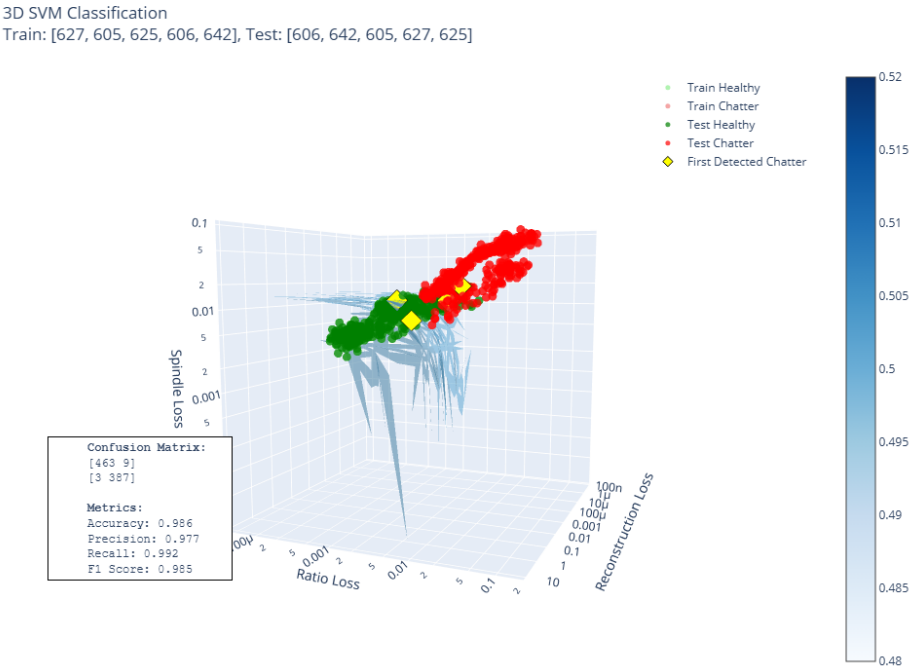


Figure 4.8: Three-dimensional SVM feature space visualization displaying the distribution of training and test data points. While red dots indicate chatter samples, green points denote healthy samples. Yellow diamonds show the first identified chatter spots for every experiment. The blue surface is the SVM decision border; the decision confidence is shown by the colour intensity.

geographic distribution of data across Ratio Loss, Spindle Loss, and Reconstruction Loss dimensions on logarithmic scales. With healthy samples (green) occupying a separate clear area apart from chatty samples (red), the feature space shows obvious clustering of data by condition type. Effectively separating the feature space with least misclassification, the SVM decision boundary—shown by the blue semi-transparent surface with variable color intensity according to decision confidence. With 463 true negatives (correctly identified healthy samples), 387 true positives (properly classified chatter samples), and only 12 misclassifications (9 false negatives and 3 false positives), the confusion matrix shown in the inset reveals remarkable classification performance.

Especially important are the yellow diamond dots indicating the first identified chatter

spots for every experiment. Predominantly at the decision boundary, these spots indicate the important shift from stable to unstable cutting circumstances. Their placement confirms the SVM's capacity to detect the early indications of chatter, which is vital for proactive action before major chatter arises. Unlike two-dimensional projections, the three-dimensional space exposes structural information. The image shows the cross-validation method used to guarantee resilience across many machining circumstances by identifying the particular experiment numbers used for training ([627, 605, 625, 606, 642]) and testing ([606, 642, 605, 627, 625]). With an accuracy of 0.986, a precision of 0.977, a recall of 0.992, and an F1 score of 0.985, the inset performance metrics reflect the remarkable classification skills. This steady performance across all measures suggests balanced categorization with no bias toward either class. In machining applications, the strong recall is especially crucial as it reduces the likelihood of overlooking real chatter occurrences that can harm the tool or workpiece. The discriminative ability of the chosen characteristics is shown by the obvious separation between healthy and chatter clusters across the three loss dimensions. Though the SVM efficiently combines this data into a consistent decision boundary, the distinct logarithmic scaling of the axes highlights the various sizes and distributions of these features. For applications of real-time monitoring and control, this flexibility and capacity to detect early chatter at the border area make the method very useful.

4.3.4 Window Size Analysis

Inherently, the effectiveness of time-series analysis for chatter detection depends on the choice of a suitable window size balancing temporal resolution with feature stability. A thorough examination across many window widths from 750 to 3000 samples—corresponding to 29.3 to 117.2 milliseconds at the 25600 Hz sampling rate—was done to methodically assess this vital factor. Under two training conditions—(1) using the whole dataset to train the SVM classifier and (2) using a limited training set made up only of experiments 627, 606, and 625, which reflects a more difficult generalization situation—Table 4.2 shows the performance metrics achieved with various window sizes.

This study reveals some notable tendencies. As the window size rises from 750 to

Table 4.2: SA-TSLANet performance measures with different window sizes: SVM decision boundary formation: Comparison between whole dataset and limited training set (Experiments 627, 606, and 625)

Window Size	SVM Train Data	Accuracy	Precision	Recall	F1
750	All	0.96	0.92	0.95	0.94
	627,606,625	0.73	0.99	0.50	0.66
900	All	0.97	0.94	0.97	0.96
	627,606,625	0.91	0.98	0.85	0.91
1200	All	0.98	0.96	0.97	0.97
	627,606,625	0.94	0.99	0.89	0.94
1500	All	0.99	0.98	0.99	0.99
	627,606,625	0.97	0.97	0.99	0.98
2000	All	0.99	0.98	0.98	0.98
	627,606,625	0.98	0.97	0.99	0.98
2500	All	0.99	0.98	0.98	0.98
	627,606,625	0.98	0.98	0.98	0.98
3000	All	0.98	0.98	0.98	0.98
	627,606,625	0.98	0.98	0.99	0.98

1500 samples, all performance indicators show first consistent improvement. In the limited training situation, where the F1 score rises from 0.66 with a 750-sample window to 0.98 with a 1500-sample window, this improvement is very noticeable. This suggests that bigger windows collect more complete spectral data required for strong feature extraction and generalization to unobserved machining settings. The limited training situation exposes important information about the generalization capacity of the model. Smaller windows (750-1200 samples) reveal a considerable performance difference between the full and limited training situations; the restricted model demonstrates sufficient accuracy but much reduced recall (as low as 0.50 for 750-sample windows). This disparity shows that the model trained on limited data tends to be unduly cautious, successfully recognizing obvious chatter cases (high accuracy) while ignoring more nuanced chatter circumstances (low recall).

At the 1500-sample window size (58.6 ms), a significant threshold emerges where the limited training model performs almost identically to the full training model. With accuracy, accuracy, recall, and F1 scores all at or above 0.98, both situations provide outstanding outcomes. This convergence implies that at this window size, the feature extraction capacity of the SA-TSLANet model becomes sufficiently strong to capture basic chatter features that generalize efficiently across many cutting situations. Curiously, in each case, more window size beyond 1500 samples produces no more benefit. Though they do not exceed those obtained with 1500-sample windows, windows of 2000, 2500, and 3000 samples maintain continuously good performance measures. This plateau effect implies that the 1500-sample window catches an ideal amount of information for chatter detection; bigger windows might include superfluous information at the expense of temporal resolution. The outstanding generalization performance with 1500-sample windows (0.98 F1 score with little training) has notable practical consequences. It shows that while trained on a small number of reference tests, the SA-TSLANet model can nevertheless perform well under various machining settings. In industrial environments where gathering thorough training data across all conceivable machining conditions can be unrealistic, this capacity is very useful.

The window size analysis offers insight into the temporal features of chatter phenomena from a theoretical point of view. Roughly 58.6 ms of signal data, which covers many revolutions of the cutting tool at normal spindle speeds, correlates to the better performance of 1500-sample windows. While keeping responsiveness for practical real-time monitoring purposes, this period seems enough to record the cyclic patterns and frequency characteristics separating chatter from regular cutting vibrations.

4.3.5 Temporal Analysis of Detection Performance

The temporal behavior of the detection system was extensively analyzed to evaluate its real-time performance characteristics. Figure 4.9 illustrates the SVM probability outputs across time for all test experiments, with a 0.5 threshold (dashed line) for chatter classification. For practical implementation, a low-pass filter was applied to the loss signals to simulate real-time processing conditions and reduce false posi-

tives from transient fluctuations. This approach mimics how the system would function in an actual machining environment, where signal stability is critical for reliable decision-making. Experiments 642 and 625 demonstrate particularly clear transitions from stable to unstable cutting conditions, with the SVM probability rapidly increasing from near-zero to above the threshold at the onset of chatter. Notably, in Experiment 625, some minor fluctuations in the probability curve prior to definitive chatter detection are effectively managed by the low-pass filtering, preventing premature triggering. Experiment 627, which contains only healthy cutting conditions, correctly maintains low probability values throughout the entire duration, with occasional minor spikes remaining well below the detection threshold. This demonstrates the robustness of the method against false positives. In Experiment 605, the chatter occurs relatively late in the process, and the model detects this transition with exactly. Similarly, Experiment 606 shows a sharp probability increase at approximately 5.63 seconds, coinciding precisely with the true onset of chatter conditions.

The vertical yellow markers indicate the exact time points of first detected chatter, which align closely with the boundaries between green (healthy) and red (chatter) regions representing ground truth. The consistent performance across diverse experimental conditions demonstrates the system's ability to provide timely detection without excessive false alarms. This temporal analysis confirms that the SA-TSLANet framework, combined with SVM classification and appropriate signal filtering, provides reliable early detection of chatter conditions while maintaining stability against noise and transients in real-time operational scenarios.

4.3.6 Generalization Capacity Analysis

The model's ability to generalize beyond its training data is a critical aspect for practical deployment in diverse machining environments. To evaluate this capability, we implemented a cross-validation strategy by training the SVM classifier exclusively on data from experiments 627, 606, and 625, then testing its performance on unseen experiments.

Figure 4.10 showcases the temporal detection performance on experiments 642 and 605, which were completely excluded from the training process. Despite having no

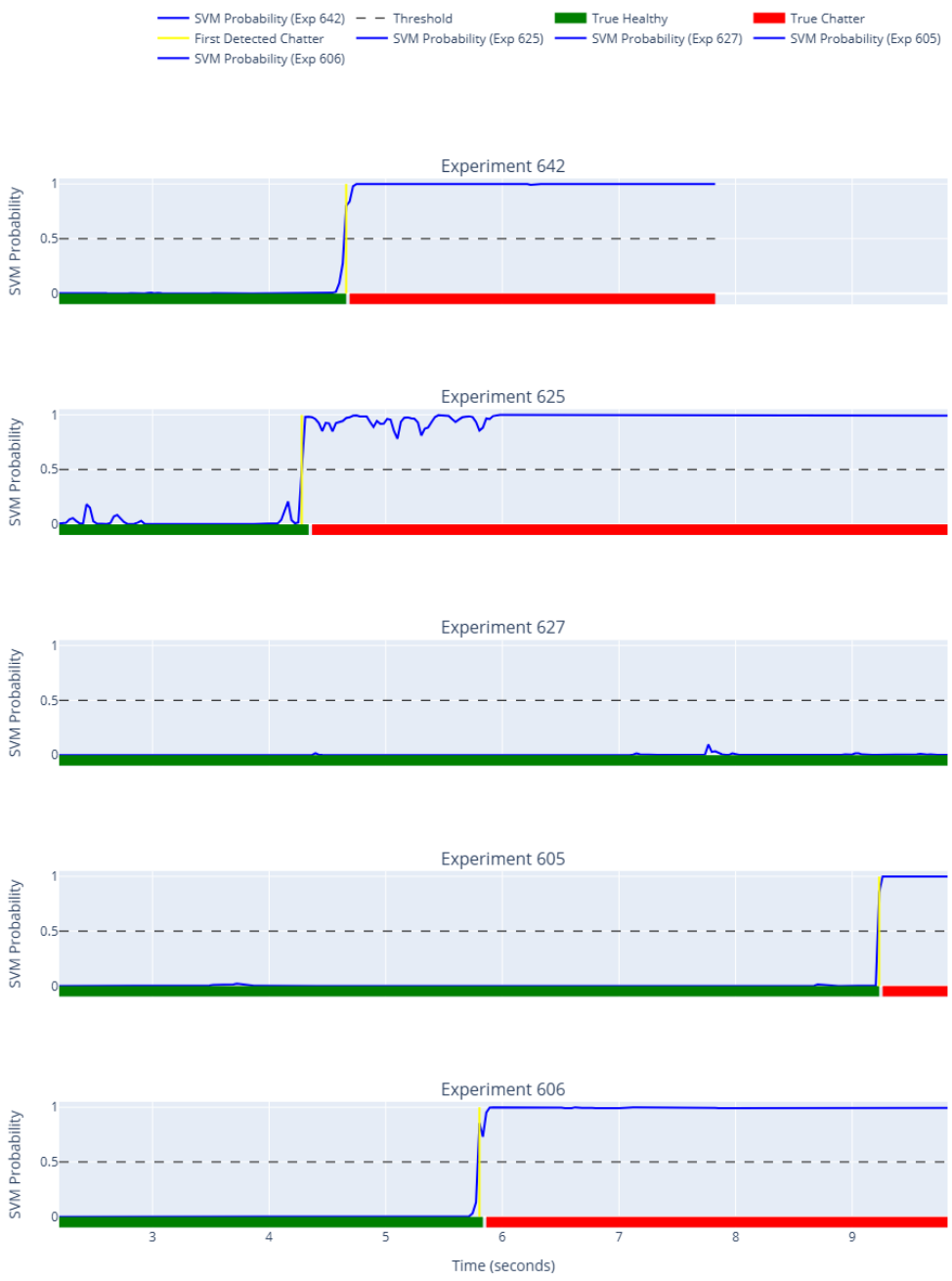


Figure 4.9: Temporal analysis of SA-TSLANet detection efficacy over five cutting tests. The blue lines denote the SVM probability scores of the SA-TSLANet model, while the green and red bars signify real healthy and chatter zones. The dashed horizontal line at 0.5 signifies the categorization threshold, whereas yellow markers denote the first detected chatter point in each trial.

prior exposure to these specific cutting conditions, the model demonstrates remarkable generalization capabilities. In Experiment 642, the classifier identifies the transition to chatter conditions with exceptional precision, where the SVM probability rises sharply from near-zero to nearly 1.0 at the exact moment when the true chatter begins. Similarly, for Experiment 605, which exhibits a much later onset of chatter compared to other tests, the classifier correctly maintains low probability values throughout the extended stable cutting phase before accurately detecting the chatter transition at approximately 9.2 seconds. The consistent performance on these unseen experiments validates that the SA-TSLANet framework extracts fundamental physical features that characterize chatter conditions across different machining scenarios rather than merely memorizing patterns specific to the training data. This generalization capability is particularly significant for industrial applications, where the system must reliably detect chatter under varying operating conditions, tool geometries, and workpiece materials without requiring extensive retraining. The successful generalization observed in these experiments suggests that the multi-objective loss function and the specialized frequency-domain processing implemented in the model effectively capture the universal physical signatures of chatter phenomena, independent of the specific experimental conditions under which they occur.

4.3.7 Impact of Adaptive Filtering in Spectral Processing

The Spindle-Aware Adaptive Spectral Block (SA-ASB) represents a key innovation in the SA-TSLANet architecture, with its adaptive filtering mechanism playing a crucial role in enhancing spectral feature extraction. To quantify the contribution of this component, it is conducted comparative experiments with the adaptive filter disabled, maintaining all other architectural elements identical. Figure 4.11 presents the SVM classification results for Experiments 625 and 642 when the model lacks adaptive filtering capabilities. The most notable difference compared to the standard implementation is the increased instability in the probability curves. In Experiment 625, significant fluctuations are visible throughout the detection timeline, with the probability temporarily dropping below the threshold even after chatter onset has been correctly identified. Similarly, in Experiment 642, a false positive situation occurs. These fluctuations represent potential false negatives and positives in a practical im-

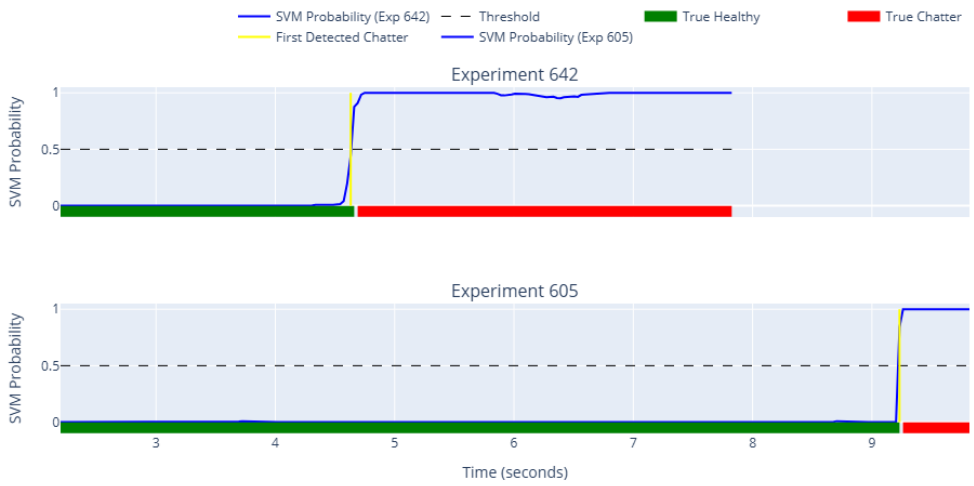


Figure 4.10: SVM classification performance on unseen test data (Experiments 642 and 605) when trained exclusively on Experiments 627, 606, and 625. The green and red bars indicate ground truth labels (healthy and chatter conditions), while the blue curves show the SVM probability over time. The dashed horizontal line represents the 0.5 classification threshold, and yellow vertical markers indicate the first detected chatter point.

plementation, where the system might intermittently report that chatter conditions have ceased when they are actually continuing. Such inconsistent classification could lead to ineffective control responses in an automated intervention system. The adaptive filtering mechanism provides two critical functions that explain its contribution to detection stability:

1. **Dynamic Frequency Emphasis**: By adaptively adjusting thresholds based on the spectral energy distribution, the filter emphasizes the most relevant frequency components while suppressing noise.

2. **Spectral Separation Enhancement**: It improves the separation between spindle and chatter components by adaptively highlighting their distinct energy patterns.

These experimental results confirm that the adaptive filtering component significantly enhances the robustness of chatter detection, particularly for maintaining stable clas-

sification after initial detection. While both variants can successfully identify the onset of chatter, the complete model with adaptive filtering provides more reliable and consistent monitoring throughout the cutting process, which is essential for practical industrial applications where detection stability directly impacts control decisions.

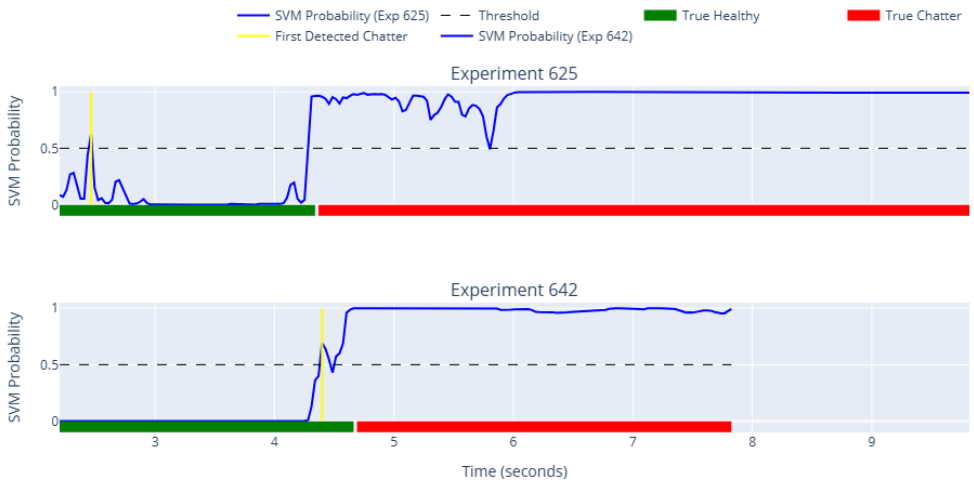


Figure 4.11: SVM classification performance on Experiments 625 and 642 with the adaptive filter disabled in the SA-ASB. Note the increased fluctuations in probability values compared to the standard implementation with adaptive filtering enabled. The green and red bars indicate ground truth labels (healthy and chatter conditions), while the blue curves show the SVM probability over time. The dashed horizontal line represents the 0.5 classification threshold, and yellow vertical markers indicate the first detected chatter point.

4.3.8 Visual Validation of Detection Accuracy

To validate the detection accuracy of the SA-TSLANet model against physical evidence, surface quality analysis was performed on the machined workpieces. Figure 4.12 presents a representative example from Test 606, where the transition from stable cutting to chatter conditions is clearly visible on the workpiece surface. The machined surface exhibits two distinct regions: a relatively smooth surface texture before the 9.35 cm mark (indicated by the green line), followed by a region with characteristic chatter patterns beyond this point. The SA-TSLANet detection system

identified the chatter onset at approximately 9.38 cm (indicated by the red dashed line), demonstrating exceptional accuracy with only a 0.3 mm difference from the visually identified transition point. This slight difference between visual identification and algorithmic detection can be attributed to the gradual nature of chatter development. Chatter typically initiates as low-amplitude vibrations that progressively intensify as cutting depth increases. The SA-TSLANet model's detection point aligns remarkably well with the earliest visible manifestation of surface texture degradation, confirming the model's sensitivity to incipient chatter conditions. The detection capability demonstrated here holds significant practical value, as the algorithm reliably identifies the transition zone between stable and unstable cutting conditions. This precise detection is particularly critical in high-precision machining operations where even minor surface imperfections are unacceptable.

These visual validation results corroborate the temporal detection improvements presented in Table 4.3, confirming that the SA-TSLANet model accurately identifies the physical onset of chatter. This accuracy would enable timely intervention in practical applications, potentially preventing further deterioration of workpiece surface quality and reducing scrap rates in industrial settings.

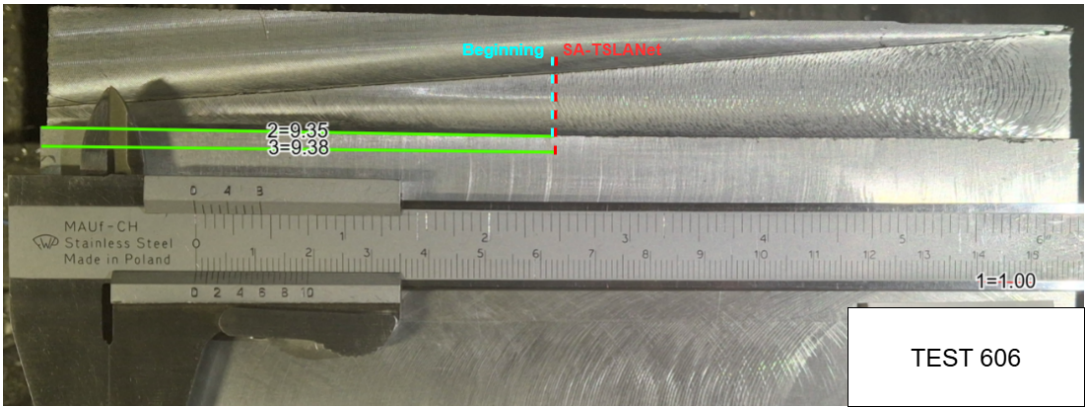


Figure 4.12: Surface quality analysis of Test 606 workpiece showing the transition from stable cutting to chatter conditions. The green line indicates the visually apparent beginning of surface deterioration at 9.35 cm, while the red dashed line shows the SA-TSLANet detection point at 9.38 cm. The measurement scale is shown in centimeters.

4.4 Comparison with Benchmark

Often regarded as the conventional standard for frequency domain analysis in chatter detection, the Fast Fourier Transform (FFT) is the basis for many monitoring systems in milling operations. Although FFT-based techniques have shown fair efficacy in controlled contexts, their susceptibility to noise, need for professional interpretation, and difficulties in early identification of incipient chatter conditions typically restrict them in actual industrial applications. This part offers a methodical comparison between the suggested SA-TSLANet model and the FFT-based labeling approach detailed in part 4.1, which generated ground truth labels for the experimental data. The comparison seeks to evaluate several important performance aspects: detection accuracy under different cutting conditions, time-to-detection measures for early chatter identification, resilience to signal noise and changes in machining parameters, and computational efficiency for real-time application. Establishing quantitative performance differentials between these approaches helps to show the better detection capabilities of the suggested method and offers insights on the particular situations where deep learning-based detection provides the most notable benefits over conventional spectral analysis methods. A suitable baseline is provided by the frequency domain analysis techniques used for data labeling, which let one directly assess how the SA-TSLANet model stacks against traditional harmonic analysis and peak detection approaches.

4.4.1 Frequency Domain Analysis for Detection Time Comparison

The time needed to spot the beginning of unstable cutting circumstances is a key measure for assessing chatter detection systems. Across four test instances, Table 4.3 shows a direct comparison between the proposed SA-TSLANet methodology and the standard FFT-based method. The findings indicate that SA-TSLANet reliably identifies chatter circumstances sooner than the conventional FFT-based method in all test scenarios. Although the enhancement seems little in absolute terms, ranging from 0.07 to 0.31 seconds, these temporal discrepancies signify crucial gaps in high-velocity machining processes. The mean detection time for SA-TSLANet (5.85 seconds) is 0.19 seconds swifter than the FFT-based technique (6.04 seconds), equat-

ing to almost 3% sooner detection. The early detection capacity is notably important in test cases 625 and 642, where the time benefit surpasses 0.14 seconds, possibly averting more serious damage to tools and workpieces in actual applications. The constant performance superiority under various cutting settings suggests that the enhanced detection capabilities arises from the model's proficiency in recognizing tiny indicators of fully evolved chatter that conventional frequency analysis may miss.

Table 4.3: Comparison of chatter detection times (in seconds) between FFT-based analysis and SA-TSLANet

Test No.	**FFT**	**SA-TSLANet**
605	9.28	9.21
606	5.85	5.63
625	4.40	4.25
642	4.62	4.31
Average	6.04	5.85

4.4.2 Time-Frequency Domain Analysis

To further validate the spectral characteristics that inform our detection approach, time-frequency analysis was performed on the acceleration data. Figure 4.13 presents a spectrogram for Test 625, which provides crucial insights into the frequency content evolution throughout the machining process. The spectrogram was generated using Short-Time Fourier Transform (STFT) with a window size of 1500 samples and 50% overlap (750 samples), allowing for both adequate frequency resolution and temporal tracking. For enhanced visual interpretation, the magnitude was converted to decibels, and the analysis was focused on the frequency range of 0-1500 Hz. At first glance, the spectrogram appears visually complex and potentially noisy, making direct interpretation challenging. The horizontal banding patterns corresponding to spindle harmonics (marked with blue dashed lines) are present but embedded within a rich spectral structure that may obscure subtle transitions between cutting states. By merely looking at these spectrograms, it would be difficult to precisely identify the onset of chatter or distinguish between normal process variations and actual instability.

This inherent complexity of the time-frequency representation highlights why more sophisticated analysis methods are necessary. The visual ambiguity in the raw spectrogram underscores the limitations of conventional frequency analysis approaches and motivates to subsequent detailed spectral energy investigations. As shown in the following sections, SA-TSLANet model effectively addresses these challenges by systematically decomposing the complex spectral patterns and extracting meaningful features that might be missed through visual inspection alone.

This time-frequency visualization, despite its visual complexity, provides the foundational data for model's more detailed computational analyses that follow, where it is quantitatively examined of the energy distribution between spindle-related and chatter-related frequency components.

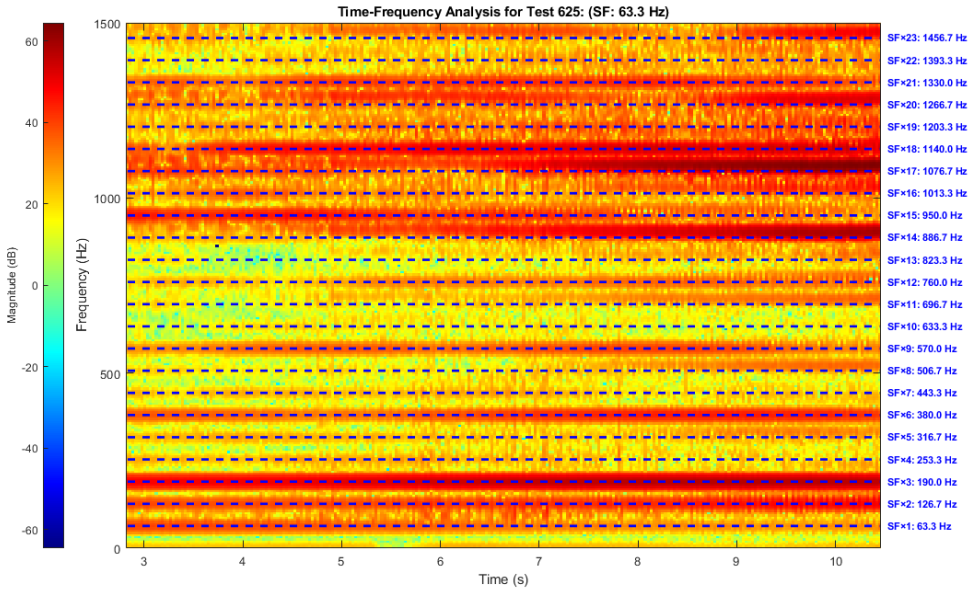


Figure 4.13: Time-frequency analysis for Test 625 with spindle frequency of 63.3 Hz. The spectrogram shows the evolution of frequency content over time with spindle harmonics marked by blue dashed lines. The color scale represents magnitude in decibels, with warmer colors indicating higher energy. Twenty-three distinct harmonics of the spindle frequency are clearly visible throughout the cutting process.

4.4.2.1 Energy Ratio Analysis

To provide a comparative benchmark for the SA-TSLANet approach, conventional frequency-domain analysis was performed using an energy ratio method. This technique, implemented through spectrogram analysis, quantifies the relationship between energy concentrated at spindle frequency harmonics versus energy distributed across other frequency bands. The energy ratio calculation was based on spectrograms generated with parameters matching our time-frequency visualization (window size of 1500 samples, 50% overlap). A bandwidth of ±6 Hz was applied around each spindle harmonic to capture the associated energy as same as selected for SA-TSLANet, with all remaining spectral content considered as non-spindle energy. The ratio of non-spindle to spindle energy serves as an indicator of chatter, with higher values suggesting unstable cutting conditions.

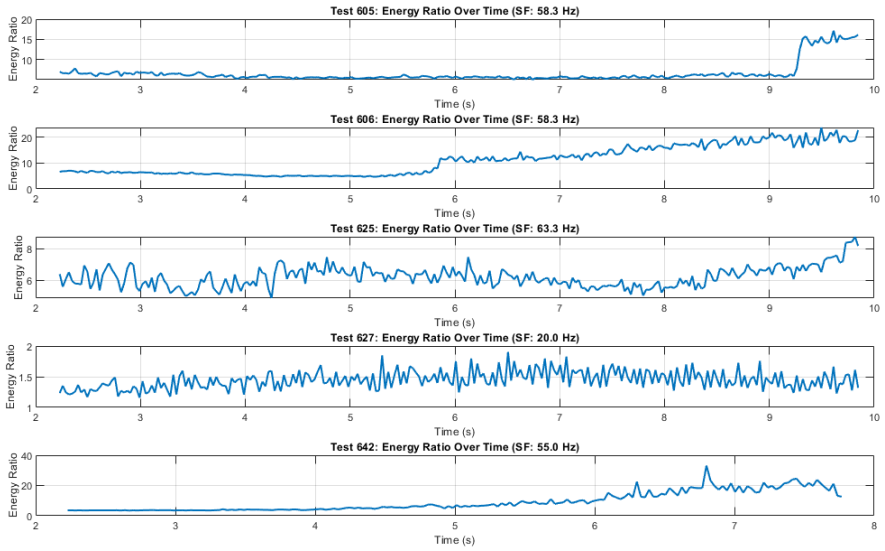


Figure 4.14: Energy ratio (non-spindle to spindle energy) versus time for all test cases. Each subplot shows the temporal evolution of the energy ratio, calculated from spectrogram data with a ±6 Hz bandwidth around spindle harmonics. Note the varying y-axis scales across different tests, reflecting the range of energy ratio values characteristic of each cutting scenario.

Figure 4.14 presents the temporal evolution of energy ratio across all test cases. Sev-

eral key observations emerge from this analysis. Test 605 exhibits a stable, low energy ratio until approximately 9.3 seconds, after which a sharp increase occurs, providing a clear indication of chatter onset. Test 606 shows a distinct transition at approximately 5.8 seconds, marked by a gradual but consistent increase in energy ratio, which aligns well with the SA-TSLANet detection point. Test 627, containing only stable cutting conditions, maintains a consistently low energy ratio with minor fluctuations throughout the process, confirming the absence of chatter. Test 642 demonstrates a progressive increase in energy ratio starting from around 4.5 seconds, but the ratio only reaches significant values after 6.0 seconds, substantially later than the actual chatter onset identified by the SA-TSLANet model. This delayed response highlights a key limitation of the conventional approach. Test 625 presents a more challenging case, with an initially elevated but relatively stable energy ratio that increases more notably after 9 seconds. This less pronounced transition explains why conventional methods might struggle with reliable detection in this experiment. While the energy ratio method provides valuable insights into the frequency characteristics of the cutting process, it exhibits limitations when compared to SA-TSLANet. Specifically, tests 625 and 642 lack the definitive and timely transition points evident in other cases, making threshold-based detection more challenging with conventional approaches. The substantial delay in Test 642's energy ratio response is particularly concerning for practical applications, where early detection is crucial. Additionally, the energy ratio method requires careful tuning of bandwidth parameters and lacks the adaptive capabilities of the proposed neural network architecture. The comparative analysis substantiates the effectiveness of the SA-TSLANet approach, which leverages multi-objective loss functions and adaptive spectral processing to capture more nuanced transitions that might be missed by simpler energy ratio calculations. This is particularly evident in cases where chatter development is more gradual or where baseline vibration patterns exhibit greater complexity.

4.5 Implementing Forecast Module

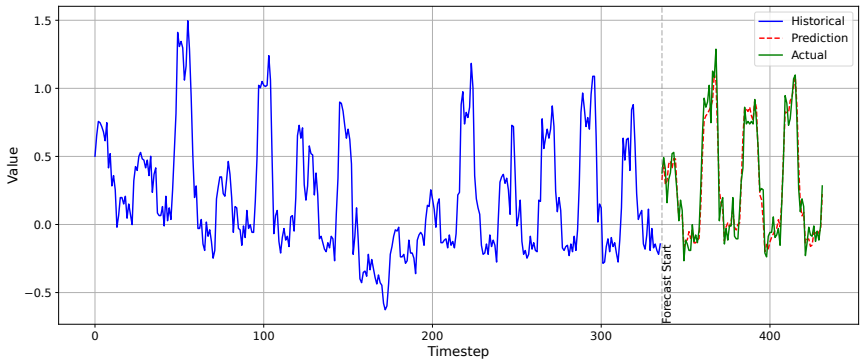
The original TSLANet architecture, as illustrated in Figure 2.4, serves as the foundation for the development of a predictive forecasting module. This extension enables the system to anticipate future accelerometer signal patterns in machining operations with high fidelity. The forecasting capabilities were systematically evaluated to assess their potential utility in preemptive process monitoring and early anomaly detection scenarios.

4.5.1 Result with Benchmark Dataset for Forecasting

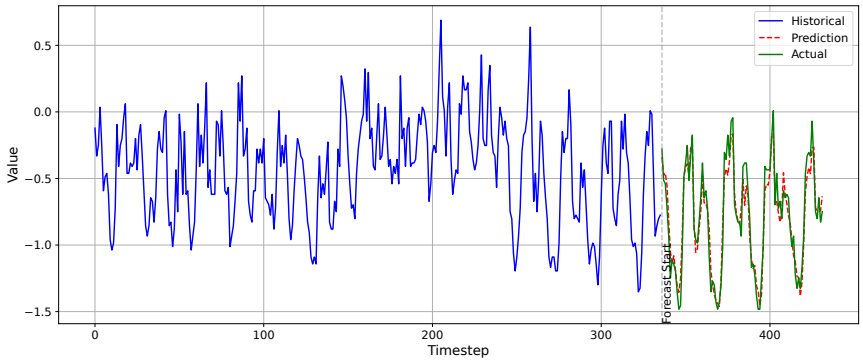
The forecast model's findings for the benchmark dataset are shown in this part. Consisting of hourly temperature measurements of electricity transformers over a two-year period, the ETTh1 (Electricity Transformer Temperature) dataset is utilized for predicting jobs. Comprising various factors including oil temperature, load, and climatic conditions, the dataset is well suited for the evaluation of multivariate time series forecasting. The comparative performance analysis in Table 4.4 shows the effectiveness of reimplemented approach in relation to prominent time series forecasting techniques. With a mean squared error (MSE) of 0.378, the reimplemented model is competitive among modern designs. With a mean squared error of 0.362, the Time-LLM shows improved performance likely due to its sophisticated language model-based approach. Strong performance is shown by conventional models like RLinear and DLinear, with mean squared errors of 0.386 and 0.375, respectively, highlighting the ongoing effectiveness of linear approaches. Proposed model outperforms other complicated designs, including Autoformer (0.449) and SCINet (0.654), suggesting that recommended architectural choices strike a good balance between complexity and forecasting effectiveness. Figure 4.15 shows examples of forecasts.

Table 4.4: Forecast model performance comparison for Etth1 dataset

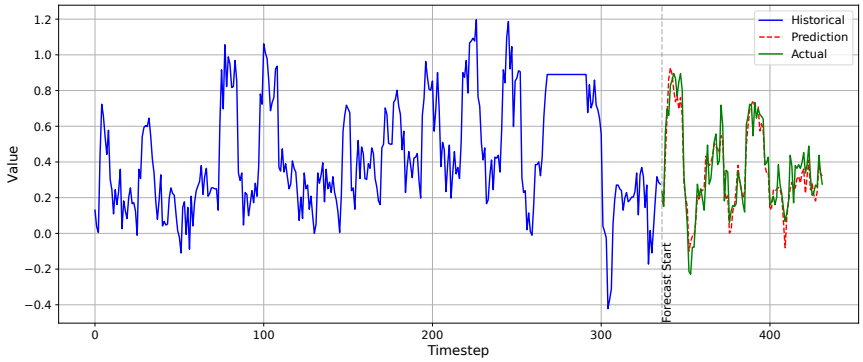
Method	**MSE**
TSLANet	0.370
Time-LLM	0.362
iTransformer	0.386
PatchTST	0.382
Crossformer	0.423
FEDformer	0.376
Autoformer	0.449
RLinear	0.386
Dlinear	0.375
TimesNet	0.384
GPT4TS	0.376
SCINet	0.654
Reimplemented Model	**0.378**



(a) Feature HUFL



(b) Feature LUFL



(c) Feature MUFL

Figure 4.15: Forecast examples for Etth1 Benchmark Dataset with reimplemented forecast model [4].

4.5.2 Training Methodology

Trained on accelerometer data gathered under steady cutting settings, the forecast module could understand the regular patterns of machine tool dynamics. The following method was used to optimize the model:

1. **Window Size Selection**: A window size of 1500 samples–58.6 ms at 25.6 kHz sampling rate—was chosen as ideal after thorough investigation (see Table 4.2), as it offered enough temporal context and preserved computational efficiency.

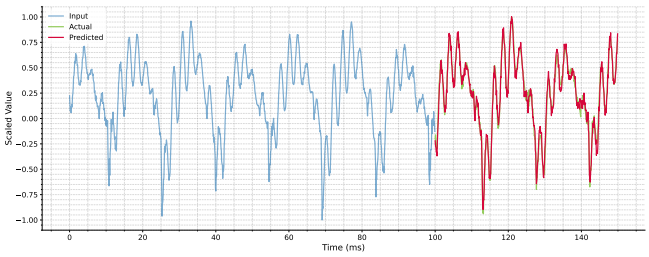
2. **Sequence Splitting**: A target prediction sequence of 500 samples and an input sequence of 1000 samples made up every 1500-sample window.

3. **Loss Function**: With further spectral consistency criteria to guarantee frequency-domain integrity, the model was trained on mean squared error (MSE) between expected and actual signal values.

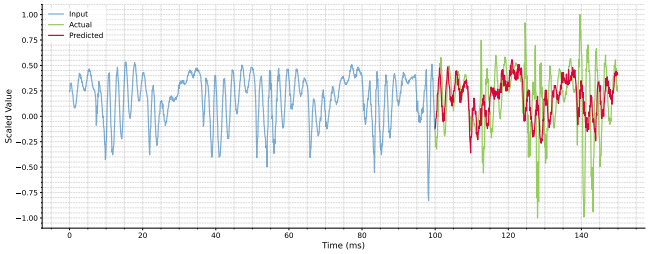
4. **Optimization**: Using the AdamW optimizer with an initial learning rate of 1×10^{-4} and weight decay of 1×10^{-6}, a ReduceLROnPlateau scheduler is used for adaptive learning rate modification.

4.5.3 Forecast Performance Evaluation

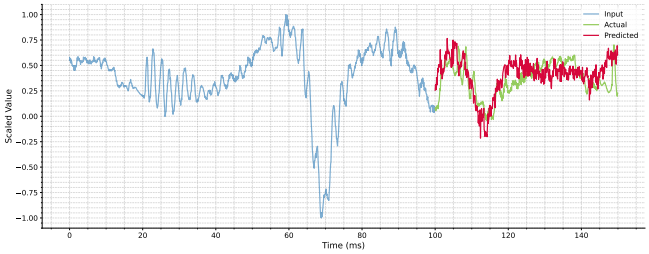
Different tests were used to assess the performance of the forecast module. Sample predictions from many experimental configurations are shown in Figure 4.16.



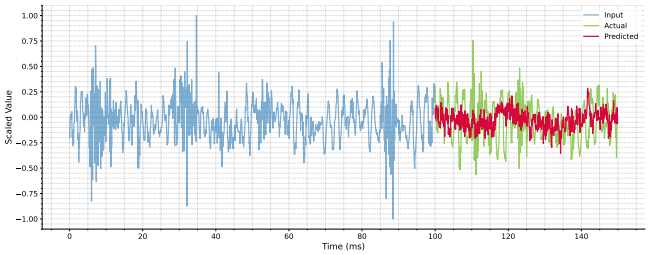
(a) Experiment 605 (Stable)



(b) Experiment 605 (Established Chatter)



(c) Experiment 627 (Stable)



(d) Experiment 642 (Stable)

Figure 4.16: Illustrations of signal predicting under several experimental settings. Blue areas are input data (1000 samples), green lines depict the actual future signal, and red lines indicate the model's forecast (500 samples). Observe how prediction accuracy changes over experiments and between steady cutting and chatter situations.

Table 4.5 and Table 4.6 show the quantitative performance metrics under various experimental settings. The model maintains more accuracy for stable cutting circumstances (label 0) as predicted; for chatter situations (label 1) it shows more prediction error—a logical result of the naturally chaotic and less predictable character of chatter vibrations.

Table 4.5: Forecast performance metrics by cutting condition

Cutting Condition	MSE	MAE	MAPE (%)
Stable (Label 0)	0.0737	0.1761	238.34
Chatter (Label 1)	0.1570	0.3090	451.84

Table 4.6: Forecast performance metrics by experiment

Experiment ID	MSE	MAE	MAPE (%)
605	0.0120	0.0758	99.71
606	0.0568	0.1646	299.01
625	0.2245	0.3866	510.97
627	0.1314	0.2769	353.12
642	0.1443	0.3011	456.20

Especially, Experiment 605 outperforms all others with the least error criteria; Experiment 625 has the greatest forecast inaccuracy. Different spindle speeds, cutting depths, and chatter characteristics found in each experiment account for this difference. Particularly during the transition between steady cutting and chatter conditions, the relatively high MAPE values across all studies show the difficulty of accurately predicting high-frequency vibration signals. The difference in prediction accuracy under steady vs chatter situations is very clear, as seen in Figure 4.16. While displaying typical variations when predicting chatter vibrations (Figure 4.16b), the model generates very accurate forecasts for steady cutting segments (Figures 4.16a and 4.16c).

4.5.4 Integration with Detection Framework

For the forecasting task, the 1500-sample window size was very useful as it provided the best compromise between preserving computing efficiency for real-time implementation and collecting enough temporal information for precise prediction. This window size lets the system forecast 19.5 milliseconds into the future (500 samples), about 58.6 milliseconds at the 25.6 kHz sampling rate utilized in our trials, which is a vital time frame for preventative action before chatter may cause harm. The forecast module, as shown in Figure 4.16, has more accuracy during steady cutting settings but exhibits typical deviation patterns when chatter starts to appear.

4.5.5 SVM Classification with Forecasting Module

A comparative study employing both active and passive forecast integration methods was done to assess the effect of integrating forecasting features with the chatter detection system. While the passive method treats the forecast module as a parallel system whose outputs are analyzed separately, the active integration adds forecast mistakes directly into the classification characteristics.

4.5.5.1 Performance Analysis

Table 4.7: Performance metrics of SA-TSLANet with various window sizes: Comparison between complete dataset and restricted training set (Experiments 627, 606, and 625) for SVM decision boundary formation for module active and passive scenarios

Window Size	Forecast	SVM Train Data	Accuracy	Precision	Recall	F1
1500	Passive	All	0.99	0.98	0.99	0.99
		627,606,625	0.97	0.97	0.99	0.98
1500	Active	All	0.97	0.95	0.98	0.96
		627,606,625	0.68	0.99	0.45	0.62

Table 4.7 offers a comparison study of the system performance using active and pas-

sive forecast integration techniques. These findings allow one to make many important observations: Using the whole training dataset, the passive integration method first shows better performance across all measures, attaining near-perfect classification with accuracy, precision, recall, and F1 scores all at or above 0.97. More significantly, this strategy maintains strong performance even with the limited training set (Experiments 627, 606, and 625), with only a little drop in accuracy (0.99 to 0.97) and precision (0.98 to 0.97). By comparison, the active integration method performs well when trained on the whole dataset but suffers a significant decline when confined to the small training sample. Especially remarkable is the sharp drop in recall from 0.98 to 0.45, suggesting that over half of the real conversation occurrences are overlooked. With a precision of 0.99, the system seems too cautious, almost never flagging talk except in the most clear-cut situations. The basic difference in how forecast information is used explains these opposing outcomes. The active integration method adds forecast error as a feature straight, hence adding more complexity to the feature space. Although this may improve discriminating ability with enough training data, it also raises the danger of overfitting to the particular traits of the training trials. The model finds it difficult to transfer this complicated feature connection to new cutting settings when confronted with insufficient training data. By contrast, the passive integration strategy takes use of the complementing detection capacity of the forecasting module while preserving the strong three-dimensional feature space of the primary classifier. With the SVM concentrating on spectral and temporal qualities and the forecast module using prediction deviations during unstable cutting, this division of responsibilities lets each component specialize in collecting various aspects of chatter events. The forecasting module was evaluated at critical transition points where chatter initiates; however, it demonstrated inadequate performance during these dynamic phase shifts. Nevertheless, its robust performance under stable cutting conditions suggests significant potential if the training corpus were expanded to incorporate a more comprehensive representation of chatter phenomena.

4.6 Computational Complexity Analysis

The SA-TSLANet chatter detection system's computing needs were evaluated to gauge its viability for real-time deployment. The detection module's overall computing need is about 1.74 billion FLOPs whereas the prediction module's is around 138.8 million FLOFs each forward pass. This means around 4.74 GFLOPs/s for the prediction module and 59.51 GFLOPs/s for the detection module under real-time operation at the experimental sample rate of 25.6 kHz. The whole system would use around 64% of available resources with a conventional desktop CPU providing about 100 GFLOPs of processing power, therefore it would be feasible for practical use. Accounting for around 20.5% of the total computational burden, the FFT/IFFT computations in the Adaptive Spectral Block are the most computationally demanding. The prediction module by itself (at less than 5% CPU use) might be implemented to provide early warnings for resource-constrained settings; more robust equipment would be required for the whole detection capabilities.

4.6.1 Inference Time vs. Forecasted Time Analysis

The solution produced efficient execution speeds given the quite large theoretical FLOP needs. Using just 0.20% of the 58.59 ms available processing window (1500 samples at 25.6 kHz), benchmarking showed the detection module requires only 0.116 ms for inference on GPU hardware. While the prediction module was initially considered as part of the system architecture, the final implementation focused exclusively on the detection module for real-time chatter identification. Due to improved tensor operations and efficient hardware use, these recorded inference times are far quicker than the theoretical projections. The computational efficiency demonstrates that the system may either be implemented on less powerful embedded systems while preserving real-time performance or manage multiple machining channels concurrently on conventional hardware. These findings provide strong evidence that the method is not only theoretically valid but also practically feasible for real-time chatter identification in industrial machining applications.

Table 4.8 presents the execution times of the SA-TSLANet system on a standard CPU

Table 4.8: Computational Performance of TSLANet-based Chatter Detection System

Module	**FLOPs**	**CPU Time**	**GPU Time**
Detection	1.74 G	17.4 ms	0.116 ms
Prediction	138.8 M	1.388 ms	0.00925 ms

(100 GFLOPs/s) versus GPU (15 TFLOPs/s). The GPU implementation achieves 150× speedup, enabling efficient real-time processing at 25.6 kHz sampling rate with minimal resource utilization.

CHAPTER 5

CONCLUSION AND DISCUSSION

5.1 Summary of Findings

This thesis presents a novel approach to chatter detection in milling operations using the Spindle-Aware Time Series Lightweight Adaptive Network (SA-TSLANet). By integrating frequency domain knowledge with deep learning techniques, the proposed model demonstrates significant improvements over traditional methods in terms of both detection accuracy and early warning capabilities.

The primary findings of this research indicate that the SA-TSLANet model achieves exceptional classification performance with accuracy of 0.99, precision of 0.98, recall of 0.99, and F1 score of 0.99 when using a window size of 1500 samples, surpassing conventional FFT-based approaches. Early detection capabilities are demonstrated across all test cases, with the model identifying chatter onset on average 0.19 seconds earlier than traditional frequency domain analysis methods, which represents a critical advantage for preventing tool damage and workpiece surface degradation. The model exhibits strong generalization capabilities, maintaining high performance (F1 score of 0.98) even when trained on a limited subset of experimental data, suggesting robust feature extraction that captures fundamental physical characteristics of chatter phenomena.

Furthermore, the multi-objective loss function incorporating crest factor, spindle, and ratio components effectively guides the model to learn complementary signal attributes essential for distinguishing normal cutting from chatter conditions. The Spindle-Aware Adaptive Spectral Block (SA-ASB) provides a physics-informed approach to frequency domain analysis by effectively separating spindle-related har-

monics from chatter components, enhancing the model's interpretability and performance. Computational efficiency analysis confirms the model's suitability for real-time implementation, with inference times significantly below the available processing window, enabling practical deployment in industrial settings.

These findings collectively demonstrate that the SA-TSLANet approach represents a significant advancement in the field of chatter detection for machining operations.

5.2 Discussion

5.2.1 Implications for Industrial Applications

The demonstrated performance of the SA-TSLANet model has several important implications for industrial machining applications. The ability to detect chatter conditions up to 0.19 seconds earlier than conventional methods provides a critical time advantage for implementing corrective actions. In high-speed machining operations, this early detection could prevent substantial damage to cutting tools and workpieces, potentially saving thousands of dollars in manufacturing costs. The model's robust generalization capabilities across different cutting conditions suggest that once trained, the system could be deployed across various machining scenarios with minimal recalibration. This versatility is particularly valuable in manufacturing environments where tool geometries, workpiece materials, and cutting parameters frequently change. The adaptive nature of the model's spectral processing is especially beneficial for addressing the variable conditions encountered in real-world machining. Unlike fixed-threshold approaches that may require manual adjustment when operating conditions change, the SA-TSLANet model can automatically adapt to different spindle speeds and cutting configurations.

5.2.2 Training with Healthy Data Only

One particularly promising aspect that emerged from our research is the potential for training the model using only data from healthy cutting operations. This approach addresses several significant practical challenges in implementing machine learning

solutions for chatter detection. Cost reduction represents a primary advantage, as collecting comprehensive datasets that include chatter conditions often requires sacrificing tools and workpieces, as chatter can cause significant damage. Training with exclusively healthy data eliminates these costs. Additionally, this approach enhances practical data collection, as in production environments, deliberately inducing chatter to generate training data is rarely acceptable due to the associated risks and costs. A model that can be trained on readily available healthy cutting data overcomes this limitation.

The ability to train with only one class of data (healthy) opens possibilities for unsupervised learning approaches where the model learns the normal behavior patterns and identifies deviations as potential chatter conditions. Furthermore, as the system operates in a production environment, it can continuously refine its understanding of normal cutting conditions without requiring explicit examples of failure states. Our preliminary experiments with training on limited datasets suggest that the model captures fundamental signal characteristics that distinguish normal cutting from chatter. This indicates that a training strategy focused exclusively on healthy data, coupled with appropriate anomaly detection frameworks, could yield effective chatter detection systems while significantly reducing implementation barriers.

5.2.3 Limitations and Challenges

Despite the promising results, several limitations and challenges warrant discussion. The model's performance is sensitive to the chosen window size, with optimal results obtained at 1500 samples (58.6 ms at 25.6 kHz). Different machining operations with substantially different spindle speeds may require adjustments to this parameter. Signal noise sensitivity presents another challenge; while the adaptive filtering mechanism improves noise rejection, extremely noisy environments might still challenge the system's reliability. Additional signal preprocessing might be necessary in such cases.

An important methodological limitation is that the model results were taken after the initial spindle engagement to ensure that transient responses during the initial tool entry did not affect classification performance. The initial spindle engagement cre-

ates significant transient vibrations that the model is not robust enough to classify accurately. This limitation necessitates a buffer period in practical implementations, which may slightly delay the overall monitoring process in situations where immediate classification following tool engagement is required.

Transfer learning requirements also merit consideration, as although the model generalizes well across the tested conditions, application to radically different machining setups (e.g., different machine tools or completely different operations like turning) would likely require some level of transfer learning or retraining. Sensor placement considerations remain important, as the current study utilized a specific accelerometer placement. The sensitivity of detection performance to sensor location and orientation requires further investigation. Additionally, the current implementation relies on a single accelerometer channel. Expanding to multi-channel or multi-sensor configurations could enhance robustness but would increase computational requirements.

5.2.4 Forecasting Module Potential

The forecasting module developed alongside the detection system shows promise but also highlights challenges in predicting chaotic vibration patterns. While the forecasting accuracy was high for stable cutting conditions, the inherently unpredictable nature of chatter vibrations limits prediction performance during unstable cutting. The passive integration approach, which uses the forecasting module as a complementary system rather than directly incorporating forecast errors into the feature space, demonstrated better generalization capabilities. This suggests that forecasting and detection should be treated as complementary rather than integrated tasks for optimal performance. The significant difference in forecasting accuracy between stable and chatter conditions itself provides a potential indicator for chatter detection. A sudden increase in prediction error could serve as an early warning signal, even before traditional frequency-domain indicators become apparent.

5.3 Future Research Directions

Based on the findings and limitations identified in this work, several promising directions for future research emerge. Developing fully unsupervised variants of the SA-TSLANet model that can be trained exclusively on healthy data would address the practical limitations of collecting chatter examples in production environments. Creating mechanisms for dynamically adjusting the window size based on spindle speed and other operating parameters would enhance the model's adaptability across diverse machining conditions. Extending the model architecture to incorporate data from multiple sensors (e.g., accelerometers, acoustic emission, current sensors) could improve detection reliability and provide redundancy.

Developing interfaces between the detection system and machine control systems would enable automatic intervention when chatter is detected, implementing optimal parameter adjustments in real-time. Investigating efficient transfer learning methods to adapt pre-trained models to new machining setups with minimal additional data collection presents another valuable research direction. Expanding the binary classification approach to a multi-class system that can distinguish between different levels of chatter severity would provide more nuanced monitoring capabilities.

Exploring the relationship between tool wear status and chatter detection performance could lead to integrated systems that simultaneously monitor both critical aspects of machining operations. Optimizing the model architecture for deployment on edge computing devices attached directly to machine tools would enable wider industrial adoption without requiring extensive infrastructure changes.

5.4 Final Remarks

The SA-TSLANet approach presented in this thesis represents a significant advancement in the field of chatter detection for milling operations. By combining physics-informed spectral processing with deep learning techniques, the model achieves superior performance in terms of both accuracy and early detection capabilities. The demonstrated generalization abilities and computational efficiency make this approach

particularly promising for industrial implementation. The potential for training with exclusively healthy data addresses a critical practical barrier to the adoption of machine learning solutions in manufacturing environments. This approach aligns with the industrial preference for non-destructive implementation and continuous improvement methodologies.

As manufacturing continues to embrace Industry 4.0 concepts and technologies, intelligent monitoring systems like SA-TSLANet will play an increasingly vital role in optimizing machining processes, reducing costs, and enhancing product quality. The physics-informed machine learning approach demonstrated in this work provides a template for similar advancements across various manufacturing domains where complex physical phenomena must be monitored and controlled in real-time.

REFERENCES

[1] X. Zheng, P. Arrazola, R. Perez, D. Echebarria, D. Kiritsis, P. Aristimuño, and M. S. de Buruaga, "Exploring the effectiveness of using internal cnc system signals for chatter detection in milling process," *Mechanical Systems and Signal Processing*, vol. 185, p. 109812, 2023.

[2] E. Budak and Y. Altintas¸, "Analytical prediction of chatter stability in milling—part i: General formulation," *Journal of Dynamic Systems, Measurement, and Control*, vol. 120, pp. 22–30, 03 1998.

[3] H. Cao, Y. Lei, and Z. He, "Chatter identification in end milling process using wavelet packets and hilbert–huang transform," *International Journal of Machine Tools and Manufacture*, vol. 69, pp. 11–19, 2013.

[4] E. Eldele, M. Ragab, Z. Chen, M. Wu, and X. Li, "Tslanet: Rethinking transformers for time series representation learning," 2024.

[5] R. Vashisht and Q. Peng, "Online chatter detection for milling operations using lstm neural networks assisted by motor current signals of ball screw drives," *Journal of Manufacturing Science and Engineering*, vol. 143, pp. 1–29, 08 2020.

[6] L. Zhu and C. Liu, "Recent progress of chatter prediction, detection and suppression in milling," *Mechanical Systems and Signal Processing*, vol. 143, p. 106840, 2020.

[7] M. Shaw, *Metal Cutting Principles*. Oxford series on advanced manufacturing, Oxford University Press, 2005.

[8] S. Tobias, *Machine-tool Vibration*. Blackie, 1965.

[9] M. Wiercigroch, "Chaotic Vibration of a Simple Model of the Machine Tool-Cutting Process System," *Journal of Vibration and Acoustics*, vol. 119, pp. 468–475, 07 1997.

[10] Y. Altintas, *Manufacturing Automation: Metal Cutting Mechanics, Machine Tool Vibrations, and CNC Design*. Cambridge University Press, 2012.

[11] Z. Li and K.-S. Wang, "Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario," *Advances in Manufacturing*, vol. 5, pp. 1–11, 12 2017.

[12] E. Kuljanic, M. Sortino, and G. Totis, "Multisensor approaches for chatter detection in milling," *Journal of Sound and Vibration*, vol. 312, no. 4, pp. 672–693, 2008.

[13] E. Kuljanic, G. Totis, and M. Sortino, "Development of an intelligent multisensor chatter detection system in milling," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1704–1718, 2009.

[14] Y. Chen, H. Li, X. Jing, H. Liang, and X. bu, "Intelligent chatter detection using image features and support vector machine," *International Journal of Advanced Manufacturing Technology*, vol. 102, 06 2019.

[15] H. E. Merritt, "Theory of self-excited machine-tool chatter: Contribution to machine-tool chatter research—1," *Journal of Engineering for Industry*, vol. 87, pp. 447–454, 11 1965.

[16] M. Wiercigroch, "Chaotic vibration of a simple model of the machine tool-cutting process system," *Journal of Vibration and Acoustics, Transactions of the ASME*, vol. 119, no. 3, pp. 468–475, 1997.

[17] N. H. Cook, "Self-excited vibrations in metal cutting," *Journal of Engineering for Industry*, vol. 81, pp. 183–186, 05 1959.

[18] Y. Altintaş and E. Budak, "Analytical prediction of stability lobes in milling," *CIRP Annals*, vol. 44, no. 1, pp. 357–362, 1995.

[19] T. Insperger, G. Stépán, P. Bayly, and B. Mann, "Multiple chatter frequencies in milling processes," *Journal of Sound and Vibration*, vol. 262, no. 2, pp. 333–345, 2003.

[20] G. Quintana, F. J. Campa, J. Ciurana, and L. N. de Lacalle, "Productivity improvement through chatter-free milling in workshops," *Proceedings of the Insti-*

tution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, vol. 225, no. 7, pp. 1163–1174, 2011.

[21] S. Tangjitsitcharoen and N. Pongsathornwiwat, "Development of chatter detection in milling processes," *International Journal of Advanced Manufacturing Technology*, vol. 65, pp. 919–927, Mar. 2013.

[22] L. Zhu and C. Liu, "Recent progress of chatter prediction, detection and suppression in milling," *Mechanical Systems and Signal Processing*, vol. 143, p. 106840, 2020.

[23] S. Wan, X. Li, Y. Yin, and J. Hong, "Milling chatter detection by multi-feature fusion and adaboost-svm," *Mechanical Systems and Signal Processing*, vol. 156, p. 107671, 2021.

[24] H. Z. Li, X. P. Li, and X. Q. Chen, "A novel chatter stability criterion for the modelling and simulation of the dynamic milling process in the time domain," *The International Journal of Advanced Manufacturing Technology*, vol. 22, no. 9–10, pp. 619–625, 2003.

[25] U. Bravo, O. Altuzarra, L. López de Lacalle, J. Sánchez, and F. Campa, "Stability limits of milling considering the flexibility of the workpiece and the machine," *International Journal of Machine Tools and Manufacture*, vol. 45, no. 15, pp. 1669–1680, 2005.

[26] F. Campa, L. Lopez de Lacalle, and A. Celaya, "Chatter avoidance in the milling of thin floors with bull-nose end mills: Model and stability diagrams," *International Journal of Machine Tools and Manufacture*, vol. 51, no. 1, pp. 43–53, 2011.

[27] M. Postel, B. Bugdayci, and K. Wegener, "Ensemble transfer learning for refining stability predictions in milling using experimental stability states," *The International Journal of Advanced Manufacturing Technology*, vol. 107, no. 9, pp. 4123–4139, 2020.

[28] J. Tlusty and M. Polacek, "The stability of machine tools against self excited vibrations in machining," in *International Research in Production Engineering*, pp. 465–474, ASME, 1963.

[29] Y. Altintas and M. Weck, "Chatter stability of metal cutting and grinding," *CIRP Annals*, vol. 53, no. 2, pp. 619–642, 2004.

[30] M. Lamraoui, M. Barakat, M. Thomas, and M. El badaoui, "Chatter detection in milling machines by neural network classification and feature selection," *Journal of Vibration and Control*, vol. 2, pp. 1251–1266, Aug. 2013.

[31] K. Chen, X. Zhang, Z. Zhao, *et al.*, "Milling chatter monitoring under variable cutting conditions based on time series features," *International Journal of Advanced Manufacturing Technology*, vol. 113, pp. 2595–2613, Apr. 2021.

[32] T. L. Schmitz, K. Medicus, and B. Dutterer, "Exploring once-per-revolution audio signal variance as a chatter indicator," *Machining Science and Technology*, vol. 6, no. 2, pp. 215–233, 2002.

[33] H.-N. Jo, B. E. Park, Y. Ji, D.-K. Kim, J. E. Yang, and I.-B. Lee, "Chatter detection and diagnosis in hot strip mill process with a frequency-based chatter index and modified independent component analysis," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7812–7820, 2020.

[34] Y. Wang, Q. Bo, H. Liu, *et al.*, "Mirror milling chatter identification using q-factor and svm," *International Journal of Advanced Manufacturing Technology*, vol. 98, pp. 1163–1177, 2018.

[35] Y. C. Yao, Y. H. Chen, C. H. Liu, *et al.*, "Real-time chatter detection and automatic suppression for intelligent spindles based on wavelet packet energy entropy and local outlier factor algorithm," *International Journal of Advanced Manufacturing Technology*, vol. 103, pp. 297–309, July 2019.

[36] S. Tangjitsitcharoen, T. Saksri, and S. Ratanakuakangwan, "Advance in chatter detection in ball end milling process by utilizing wavelet transform," *Journal of Intelligent Manufacturing*, vol. 26, pp. 485–499, June 2015.

[37] T. Thaler, P. Potočnik, I. Bric, and E. Govekar, "Chatter detection in band sawing based on discriminant analysis of sound features," *Applied Acoustics*, vol. 77, pp. 114–121, 2014.

[38] Y. Sun and Z. Xiong, "An optimal weighted wavelet packet entropy method with application to real-time chatter detection," *IEEE/ASME Transactions on Mechatronics*, vol. 21, no. 4, p. 2004 – 2014, 2016. Cited by: 65.

[39] Y. Chen, H. Li, X. Jing, *et al.*, "Intelligent chatter detection using image features and support vector machine," *International Journal of Advanced Manufacturing Technology*, vol. 102, pp. 1433–1442, June 2019.

[40] H. Çalışkan, "Real-time milling chatter detection and control with axis encoder feedback and spindle speed manipulation," *Journal of Manufacturing and Materials Processing*, vol. 8, no. 4, p. 173, 2024.

[41] H. N. Gao, D. H. Shen, L. Yu, and W. C. Zhang, "Identification of cutting chatter through deep learning and classification," *International Journal of Simulation Modelling*, vol. 19, pp. 667–677, Dec. 2020.

[42] P.-H. Kuo, P.-C. Luan, Y.-R. Tseng, and H.-T. Yau, "Machine tool chattering monitoring by Chen-Lee chaotic system-based deep convolutional generative adversarial nets," *Structural Health Monitoring*, vol. 22, no. 6, pp. 3891–3907, 2023.

[43] M. Lamraoui, M. Barakat, M. Thomas, and M. Badaoui, "Chatter detection in milling machines by neural network classification and feature selection," *Journal of Vibration and Control*, vol. 21, no. 7, pp. 1251–1266, 2015.

[44] S. Rezaei, A. Cornelius, J. Karandikar, *et al.*, "Using GANs to predict milling stability from limited data," *Journal of Intelligent Manufacturing*, Jan. 2024.

[45] A. P. Rifai, R. Fukuda, and H. Aoyama, "Surface roughness estimation and chatter vibration identification using vision-based deep learning," *Journal of the Japan Society for Precision Engineering*, vol. 85, pp. 658–666, July 2019.

[46] B. Sener, G. Serin, M. U. Gudelek, A. M. Ozbayoglu, and H. O. Unver, "Intelligent chatter detection in milling using vibration data features and deep multilayer perceptron," in *2020 IEEE International Conference*, IEEE, 2020.

[47] F. Shi, H. Cao, Y. Wang, *et al.*, "Chatter detection in high-speed milling processes based on ON-LSTM and PBT," *International Journal of Advanced Manufacturing Technology*, vol. 111, pp. 3361–3378, Dec. 2020.

[48] H. O. Unver and B. Sener, "A novel transfer learning framework for chatter detection using convolutional neural networks," *Journal of Intelligent Manufacturing*, vol. 34, pp. 1105–1124, Mar. 2023.

[49] Y. Altintas and S. Park, "Dynamic compensation of spindle-integrated force sensors," *CIRP Annals*, vol. 53, no. 1, pp. 305–308, 2004.

[50] Y. Fu, Y. Zhang, H. Gao, *et al.*, "Automatic feature constructing from vibration signals for machining state monitoring," *Journal of Intelligent Manufacturing*, vol. 30, pp. 995–1008, Mar. 2019.

[51] M. Goswami, C. Challu, L. Callot, L. Minorics, and A. Kan, "Unsupervised model selection for time-series anomaly detection," 2023.

[52] A. Jiang, W.-Q. Zhang, Y. Deng, P. Fan, and J. Liu, "Unsupervised anomaly detection and localization of machine audio: A gan-based approach," 2023.

[53] W. Jiang, Y. Hong, B. Zhou, X. He, and C. Cheng, "A gan-based anomaly detection approach for imbalanced industrial time series," *IEEE Access*, vol. 7, pp. 143608–143619, 2019.

[54] P. Mobtahej, X. Zhang, M. Hamidi, and J. Zhang, "Deep learning-based anomaly detection for compressors using audio data," in *2021 Annual Reliability and Maintainability Symposium (RAMS)*, pp. 1–7, 2021.

[55] A.-H. Shin, S. T. Kim, and G.-M. Park, "Time series anomaly detection using transformer-based gan with two-step masking," *IEEE Access*, vol. 11, pp. 74035–74047, 2023.

[56] G. Tan, P. Chen, and M. Li, "Online data drift detection for anomaly detection services based on deep learning towards multivariate time series," in *2023 IEEE 23rd International Conference on Software Quality, Reliability, and Security (QRS)*, pp. 1–11, 2023.

[57] M.-H. Vu, V.-Q. Nguyen, T.-T. Tran, V.-T. Pham, and M.-T. Lo, "Few-shot bearing fault diagnosis via ensembling transformer-based model with mahalanobis distance metric learning from multiscale features," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–18, 2024.

[58] J. Xie, Y. Cui, F. Huang, C. Liu, and K. Zheng, "Marina: An mlp-attention model for multivariate time-series analysis," in *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22, (New York, NY, USA), p. 2230–2239, Association for Computing Machinery, 2022.

[59] L. Xu, K. Xu, Y. Qin, Y. Li, X. Huang, Z. Lin, N. Ye, and X. Ji, "Tgan-ad: Transformer-based gan for anomaly detection of time series data," *Applied Sciences*, vol. 12, no. 16, 2022.

[60] M.-C. Popescu, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," *WSEAS Transactions on Circuits and Systems*, vol. 8, no. 7, pp. 579–588, 2009.

[61] R. Mishra, M. S. Kiran, M. Maheswaram, A. Upadhyay, and B. Singh, "Investigation of optimal feature for milling chatter identification using supervised machine learning techniques," *Journal of Engineering Research*, vol. 12, no. 4, pp. 950–962, 2024.