Optimizing Metal Cutting
Empirical tests & AI-enhanced tool wear detection

Anas Sayegh

CODEN:LUTMDN/(TMMV-5365)/1-59/2024

Optimizing Metal Cutting: Empirical Tests & Ai-Enhanced Tool Wear Detection
Master's Degree In Production & Materials Engineering
Division Of Production & Materials Engineering
Lund University

Master of Science Degree at the Faculty of Engineering at Lund University to be publicly defended on 23.05.2024 at 09.00 at Austinite Hall, Department of Production, Address 221 00 Lund, Sweden

Author: Anas Sayegh
Academic Organization: Lund University
Sponsoring organization: Sandvik Coromant AB
Academic Supervisor: Volodymyr Bushlya, Petra Maier, Professor
Industrial Supervisor: Peder Arvidsson, Ellinor Svensson, Coromant AB
Examiner: Mats Andersson, Professor

Avdelningen för Industriell Produktion
Lunds Tekniska Högskola
Lunds universitet
Box 118
221 00 Lund, Sverige

Division of Production and Materials Engineering
LTH, Faculty of Engineering
Lund University
Box 118
221 00 Lund, Sweden

Printed in Sweden
Media-Tryck
Lund University
+46 46 222 72 00

Acknowledgements

During my thesis, I received great support. I would like to thank Sandvik Coromant AB for giving me the opportunity to write a thesis that combines machine learning and material science. Both of which are areas that I am interested in.

I want to give a big thanks to my supervisors Volodymyr Bushlya, Petra Maier for their continued support throughout the project. Without your time and expertise, this work would never have been possible.

Finally, I would also like to thank Peder Arvidsson, Ellinor Svensson, Ehsan Layegh, and all colleagues at Sandvik Coromant AB who have helped me along the way and made the time at Sandvik Coromant AB thoroughly enjoyable. It has been a pleasure working with all of you.

Anas Sayegh, Lund, May 2024

TABLE OF CONTENTS

List of Tables

List of Figures

List of Equations

Abstract :

Detection of tool wear is critical in optimizing tool change strategies and cutting parameter selection to influence productivity, cost, and quality in metal cutting. Nonetheless, current practices for tool wear detection majorly depend on manual, inconsistent, and time-consuming visual inspection by operators.

Tool wear detection has the potential to greatly benefit from recent advances in artificial intelligence, computer vision, and deep learning. This thesis investigates the use of deep learning and image classification techniques to detect tool wear in metal cutting. The research involves performing empirical machining tests to create a diverse dataset of tool wear images under various cutting conditions, tool geometries, and workpiece materials. The images are annotated with the corresponding wear metrics such as flank wear, crater wear, and wear morphology.

Using the PyTorch deep learning framework in Python, we create and train a Convolutional Neural Network (CNN) on the tool wear image dataset. Transfer learning is used to fine-tune pretrained CNN architectures for the tool wear classification task. To make the model robust and improve generalization, data augmentation and cross-validation are implemented. The performance of the CNN model is evaluated on unseen test images and benchmarked against traditional computer vision methods.

In addition, the thesis investigates the relationship between the extracted CNN features, the cutting parameters, and the wear values measured. To interpret the CNN detections, the wear images are analyzed using Gradient-weighted Class Activation Mapping (Grad-CAM) and the regions of interest are highlighted. The proposed deep learning-based tool wear detection approach offers encouraging results compared to manual examination in terms of accuracy, reliability, and efficiency. The development demonstrates the potential for integration with actual tool condition monitoring and cutting process optimization systems. Hence, this investigation provides groundbreaking insights into the field of smart manufacturing and Industry 4.0. It offers clear recommendations on AI-driven manufacturing to strengthen efficiency.

Keywords: Tool wear, metal cutting, deep learning, convolutional neural networks, image classification, PyTorch, smart manufacturing.

Abbreviations:

ACF - Aggregate Channel Feature

ADAM - Adaptive Moment Estimation

API - Application Programming Interface

BN - Batch Normalization

CNN - Convolutional Neural Network

CPU - Central Processing Unit

CV - Computer Vision

DL - Deep Learning

DNN - Deep Neural Network

EDX - Energy Dispersive X-ray Spectroscopy

FC - Fully-Connected Network

FCN - Fully Convolutional Network

GAN - Generative Adversarial Network

GPU - Graphics Processing Unit

HIP - High Isostatic Pressure

IoU - Intersect over Union

MAE - Mean Absolute Error

ML - Machine Learning

NN - Neural Network

OEE - Overall Equipment Effectiveness

OP - Overall Pixel accuracy

PCD - Polycrystalline Diamond

PCBN - Polycrystalline Cubic Boron Nitride

R-CNN - Regions with CNN features

ReLU - Rectified Linear Unit

RMSE - Root Mean Square Error

SGD - Stochastic Gradient Descent

WC - Tungsten Carbide

1.0 Introduction

1.1 Background and Motivation

Manufacturing industries heavily rely on metal cutting processes, such as drilling, milling, and turning. These processes involve numerous variables that determine the quality and efficiency of the operation. One of the most significant factors affecting these processes is tool wear, which is inevitable during cutting operations. Tool wear leads to the deterioration of tool quality and can eventually result in tool failure. As the tool wears, it changes the geometry of the cutting tool, increasing cutting forces, reducing dimensional accuracy, and causing poor surface integrity of the machined components. Progressive wear can lead to catastrophic tool failures, damaging the workpiece, machine tool, and posing potential risks to operator safety. Therefore, detection of tool wear is critical for establishing process stability, product quality, and overall manufacturing efficiency.

The monetary consequences of tool wear are significant. Worn tools lead to increased tooling costs and lower productivity due to more frequent replacements, higher tool consumption, and associated costs. Additionally, tool wear can increase scrap rates, rework, and machine downtime, considerably raising overall manufacturing costs. Developing an efficient tool wear detection system can improve tool life, reduce tooling costs, and enhance overall equipment effectiveness (OEE).[18][11]

Implementing an efficient tool wear detection system offers numerous benefits. Accurate detection of tool wear and optimization of tool life can save manufacturers significant costs by minimizing machine downtime and reducing tooling expenses. These systems can detect tool anomalies, preventing catastrophic failures and ensuring consistent product quality. Tool wear detection systems are designed to integrate with production planning and control systems, improving overall equipment effectiveness and streamlining manufacturing processes.[14][12]

Recently, there has been growing interest in leveraging artificial intelligence (AI) and machine learning methods for tool wear detection and classification. Traditional approaches, such as statistical and physical models, often fail to capture the intricate, non-linear relationships between tool wear and various process parameters. AI-based methods, particularly deep learning, have emerged as valuable tools for feature extraction from sensor data and accurate detection of tool wear conditions.[13][16]

Integrating AI-based tool wear detection systems aligns with the objectives of Industry 4.0 and smart manufacturing. By deploying data analytics and machine learning, manufacturers can achieve intelligence and adaptivity in manufacturing processes. Detection of tool wear enables autonomous decision-making to optimize process parameters for enhanced manufacturing performance and efficiency. This approach also supports detective maintenance, allowing for run-to-failure policies and improved asset utilization.[17][18].

The use of deep learning and image classification techniques for detecting tool wear is driven by the limitations of traditional research methods, which often fall short in detection and precision machining in advanced manufacturing industries. Applying AI to tool wear detection brings opportunities for sustainable development in modern manufacturing, contributing to intelligent and sustainable manufacturing practices. To collect high-quality data in real-time and accurately detect tool wear states, sophisticated deep convolutional neural networks and computer vision techniques are employed to extract significant characteristics from tool wear images and effectively classify wear states.[5]

1.2 Problem Statement

In metal cutting operations, tool wear is a critical issue that significantly affects the performance of machining processes. During such operations, cutting tools undergo gradual wear, resulting in reduced cutting efficiency, poor surface finish, and dimensional inaccuracies in the machined parts. Furthermore, in its extreme forms, tool wear can lead to sudden tool failure with costly consequences. Once a tool fails, the workpiece is usually damaged, the machine tool may be harmed, and production is disrupted, resulting in significant losses for the manufacturer.

Current tool wear detection methods face challenges in achieving high accuracy, reliability, and practicality. Traditional methods, such as offline tool inspection and operator judgment based on experience, are subjective, time-consuming, and prone to human error. The lack of real-time tool wear information hinders the optimization of tool life and the prevention of unexpected tool failures. Although many statistical and physical models for tool wear detection have been developed, the complex and non-linear relationships between tool wear and process parameters often limit their practical application.

There is an urgent need for a non-invasive, automatic, and real-time wear detection system in the manufacturing field. Such a system should accurately detect tool wear levels and provide valuable information for decision-making, enabling the optimization of cutting processes and enhancing the efficiency of manufacturing industries. Real-time monitoring of tool wear during the machining process will help factories replace tools based on their actual condition rather than relying on periodic changes. This approach can reduce unnecessary tool waste and maximize tool life.

Creating a robust and adaptable model to detect the extent of cutting tool wear is a challenging problem. The diverse range of cutting tools, with varying geometries and materials, exhibit different wear patterns and rates. Additionally, numerous factors influence the wear process, including cutting speed, feed rate, depth of cut, and material properties of the workpiece. Consequently, a useful model must account for this diversity and perform effectively across different tool types and cutting conditions.

Accurate classification of tool wear is crucial for implementing timely tool replacements and reducing machine downtime. By precisely identifying the current level of wear on a cutting tool, manufacturers can make informed decisions about tool replacement, avoiding premature replacements or using tools beyond their optimal life. This, in turn, leads to reduced tooling costs, improved product quality, and minimized unplanned machine downtime.

Deep learning and computer vision techniques offer great potential in addressing the limitations of traditional tool wear detection methods. Convolutional Neural Networks (CNNs), a type of deep learning algorithm, have demonstrated outstanding performance in image-based recognition and classification tasks. By leveraging the versatility of deep learning algorithms, it is possible to automatically extract relevant features from tool wear images and accurately detect wear states, overcoming the subjectivity and inconsistency of manual evaluations.

The development of an AI-based tool wear detection system aligns with the goals of sustainable and intelligent manufacturing. It contributes to sustainability initiatives by optimizing tool life and reducing waste. Real-time monitoring and detective maintenance enabled by AI technologies can minimize unplanned downtime, improve asset utilization, and enhance overall manufacturing performance. The integration of AI-based tool wear detection into manufacturing processes is a significant step towards Industry 4.0 and smart manufacturing.

1.3 Research Objectives

The primary goal of this research is to develop an accurate and reliable system for detecting and classifying tool wear using image processing and deep learning techniques. This system aims to provide insights for optimal cutting parameter adjustments to improve overall manufacturing productivity.

Specific Objectives:

Data Collection and Preparation: Gather a diverse collection of tool wear images under various cutting conditions and tool types. This dataset should include a range of wear severities and patterns to ensure the system's reliability and broad applicability.

Model Development and Optimization: Develop and optimize a deep learning model, specifically a convolutional neural network (CNN), for precise segmentation and classification of tool wear regions. This includes refining the model's hyperparameters and structure to enhance both precision and speed.

Image Processing Pipeline: Create an automated image processing pipeline to collect, prepare, and segment tool wear images. This pipeline should handle tasks such as image normalization, noise elimination, and region of interest extraction.

Generalization and Validation: Evaluate the model's ability to generalize to new tool types and cutting conditions. Test the detective model on unseen data to verify its accuracy in varied scenarios and ensure its robustness and practicality in real-world applications.

Industrial Integration and Recommendations: Develop guidelines for integrating the AI-based tool wear monitoring system into industrial settings. This includes recommendations on hardware and software requirements, data management strategies, and addressing potential deployment challenges.

1.4 Proposed Methodology and Expected Outcomes

The proposed approach in this study includes designing a deep learning-based tool wear detection and classification system using CNN (Convolutional Neural Network) and instance segmentation techniques. By taking advantage of deep learning, the proposed approach can automatically learn useful feature representations from the tool wear images and reliably detect the tool wear states of cutting tools.

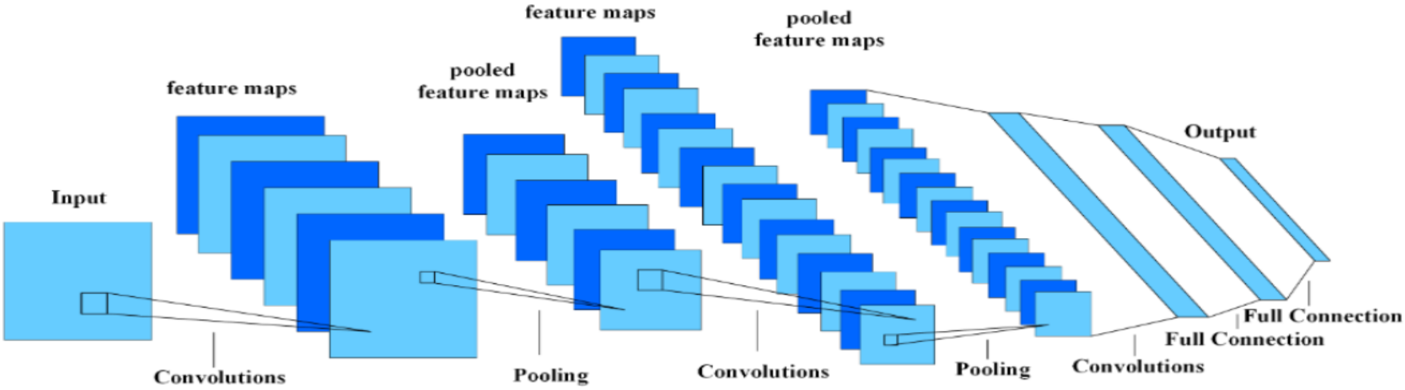


Figure 1.The classical structure of convolutional neural networks (CNNs).[F2]

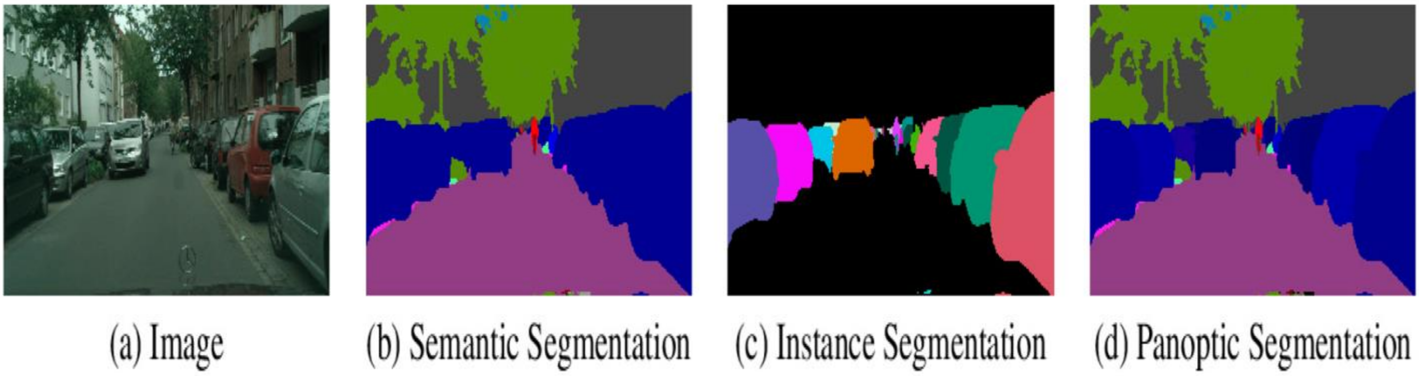


Figure 2. Difference between segmentation methods [F1]

In order to train and validate the deep learning model, a dataset of tool wear images will be acquired across multiple cutting conditions and tool types. The dataset will consist of various wear patterns, severities, and tool geometries to guarantee the effectiveness and versatility of the proposed system.

In detail, we will conduct machining experiments during which the cutting parameters (e.g., cutting speed, feed rate, and depth of cut), work materials, and tool types are varied, and images of the cutting tools are taken periodically. To automate the acquisition, pre-processing, and segmentation of tool wear images, we will build an image processing pipeline. Image normalization, noise reduction, and region of interest extraction are some of the tasks that the pipeline will handle. In order to improve image quality and ensure accurate wear region segmentation, cutting-edge image processing technologies such as edge detection, thresholding, and morphological operations will be employed.

The anticipated result of this study is to establish an effective and reliable tool wear detection and classification system based on deep learning. This system could be used to examine cutting tool images and deliver information on tool wear status. The techniques of CNN and instance segmentation will be utilized to achieve high-precision wear region detection and classification, surpassing the performance of traditional techniques.

This system brings a wide array of benefits. First, as mentioned above, it enables accurate estimation of tool life. In addition, by incorporating the AI-based tool wear detection system into manufacturing processes, intelligent and sustainable manufacturing can be achieved. detection of wear will promote data-driven decision-making, enable dynamic process optimization and adaptive control, and ultimately contribute to better resource utilization, waste reduction, and overall equipment effectiveness enhancement.

The research work performed in this thesis holds the possibility of making significant contributions to the field of tool condition monitoring. It aims to address the limitations of existing approaches and provide a more robust and efficient solution for industrial application by advancing the state-of-the-art in tool wear detection and classification using deep learning. The methodology developed through this study, together with the knowledge acquired, will provide new opportunities for research in the future and can facilitate rapid advancements in the field of smart manufacturing.

In summary, the methodology and anticipated outcomes of this study align with the goal of developing an accurate, robust, and applicable tool wear detection system using deep learning and image classification techniques. Successful implementation of the proposed methodology has the potential to revolutionize tool condition monitoring practices in the manufacturing industry and significantly contribute to the growth, quality, and sustainability of manufacturing processes.

2.0 Literature Review

2.1 Overview of tool condition monitoring techniques

Tool wear detection techniques can be broadly classified into two major categories: direct and indirect methods. Direct methods involve capturing the tool geometry and wear progression directly using optical, vision-based, and radioactive techniques. Optical methods use microscopes or cameras to image the tool directly. Vision systems employ image processing algorithms to quantify the wear area and parameters from the captured tool images. For example, Kurada and Bradley (1997) developed a vision-based system using charge-coupled device (CCD) cameras and image processing algorithms to measure flank wear and crater wear with a precision of 95%. Radioactive techniques measure the amount of radioactive particles in the chips or tool, indicating the tool wear volume. However, the application of direct methods is limited due to the harsh machining environment, coolant, and chip interference.[20]

Indirect methods track various signals resulting from tool condition, such as cutting forces, vibrations, acoustic emissions, spindle power, and temperature. These signals change as the tool wears, and empirical models or machine learning approaches are used to systematically measure the conditions. For instance, indirect methods have been employed to measure cutting forces and vibrations to infer tool wear. Sensor fusion techniques, which combine data from multiple sensors, have been developed to improve the robustness of indirect tool wear detection.[20]

2.1.1 Types of Tool Wear

To better understand the tool wear phenomena and the detection techniques discussed, it is essential to introduce the main types of tool wear encountered in machining processes. The most common types of tool wear include:[19]

1. Flank Wear: Occurs on the flank face of the cutting tool, caused by abrasive wear mechanisms, and is often used as a criterion for determining tool life.

2. Crater Wear: Appears on the rake face of the cutting tool, caused by a combination of abrasive, adhesive, and diffusive wear mechanisms, affecting the chip formation process.

3. Notch Wear: A localized wear pattern at the depth of cut line, typically due to the abrasive action of the work-hardened surface layer or built-up edge.

4. Built-up Edge (BUE): An accumulation of workpiece material on the cutting edge, changing the effective tool geometry and leading to poor surface finish and dimensional inaccuracy.

5. Chipping and Breakage: Sudden failures of the cutting edge due to excessive mechanical or thermal stresses, often resulting in catastrophic tool failure.

These wear types can occur simultaneously and interact with each other, leading to complex wear patterns and progressive tool deterioration. Understanding the characteristics and mechanisms of these wear types is crucial for developing effective tool wear detection strategies and optimizing the machining process.

Despite the breakthroughs achieved in recent years in terms of tool wear detection techniques, challenges such as accuracy, reliability, adaptability to different machining scenarios, and industrial implementation still exist. These obstacles drive extensive research on advanced methods related to computer vision and machine learning.[19]

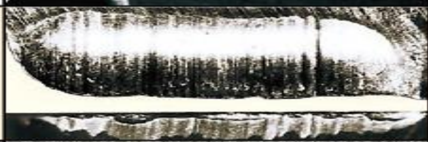
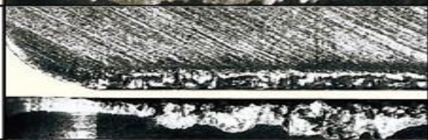
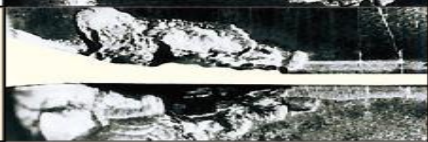
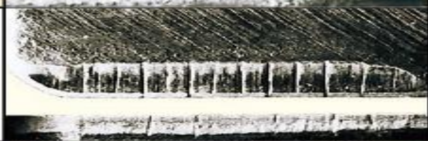
Tool Damage Form		Cause
Flank Wear		• Tool grade is too soft. • Cutting speed is too high. • Flank angle is too small. • Feed rate is extremely low.
Crater Wear		• Tool grade is too soft. • Cutting speed is too high. • Feed rate is too high.
Chipping		• Tool grade is too hard. • Feed rate is too high. • Lack of cutting edge strength. • Lack of shank or holder rigidity.
Fracture		• Tool grade is too hard. • Feed rate is too high. • Lack of cutting edge strength. • Lack of shank or holder rigidity.
Plastic Deformation		• Tool grade is too soft. • Cutting speed is too high. • Depth of cut and feed rate are too large. • Cutting temperature is high.
Welding		• Cutting speed is low. • Poor sharpness. • Unsuitable grade.
Thermal Cracks		• Expansion or shrinkage due to cutting heat. • Tool grade is too hard. *Especially in milling.
Notching		• Hard surfaces such as uncut surfaces, chilled parts and machining hardened layer. • Friction caused by jagged shape chips. (Caused by small vibration)
Flaking		• Cutting edge welding and adhesion. • Poor chip disposal.

Figure 3. Types of Wear[F4]

2.2 Conventional methods for tool wear measurement

Measuring tool wear is important for validating condition monitoring and tracking wear progression. In traditional tool life testing, wear is measured by stopping the machining process and observing the tool under a microscope. ISO 3685 defines the test processes and wear parameters for turning tools. Among those parameters, flank wear land width and crater wear depth are most commonly used, reflecting the extent of abrasive and diffusive wear respectively.[7]

Optical microscopes have traditionally been used to measure tool wear due to their simplicity and availability. Worn tool inserts are removed from the machine and observed under the microscope at various magnifications. The wear parameters are then manually measured using the microscope's reticle or by comparing with reference wear patterns. However, this method is time-consuming, subjective, and prone to human errors.[14]

To address these issues, tool wear measurement has adapted digital microscopes and image processing techniques. Digital microscopes capture high-resolution images of worn tools, and image processing algorithms analyze the images to extract wear parameters. For example, Xiong et al. developed an automatic tool wear measurement system by combining a digital microscope and image segmentation algorithm, achieving 96% accuracy and significantly reducing measurement costs compared to manual approaches.[14]

Scanning electron microscopes (SEMs) have alternatively been used for tool wear measurement, particularly for characterizing wear mechanisms and surface morphology at the micro-scale. SEMs provide greater magnification and depth of field than optical microscopes, allowing detailed characterization of worn tool surfaces. SEM images have also been used to measure cutting edge radius and correlate it with wear progression in milling. However, SEM analysis is relatively expensive, time-consuming, and requires special sample preparation, making it difficult to apply in production environments.[3]

Other metrology techniques have been investigated for 3D tool wear measurement, including white light interferometry (WLI), confocal microscopy, and focus-variation microscopy. These measurements aim to capture the detailed subsurface topography of tool wear and provide volumetric wear information. Devillez et al. determined crater wear volume using WLI and correlated it with cutting forces. Focus-variation microscopy has been employed to detect tool wear in drilling and milling, demonstrating its potential for automated wear characterization.[4]

However, these conventional methods have limitations in terms of measuring speed, flexibility, and in-process applicability. They require interrupting the machining process and removing the tool from the machine, which is time-consuming and hinders production efficiency. Furthermore, wear measurement is usually performed at the end of tool life, providing limited information about wear progression during cutting. These limitations have motivated the development of computer vision-based approaches for online tool wear detection and measurement.[4]

2.3 Computer vision approaches for tool wear analysis

The use of computer vision systems for tool wear detection and measurement has become more prevalent in recent years due to their non-contact, portable, and adaptable nature. These methods employ digital cameras or vision systems to capture images of cutting tools during or after the machining process. The obtained images are then processed using image processing algorithms to extract useful wear characteristics and indicators for tool wear measurement.

Kurada and Bradley presented one of the early vision-based approaches for tool condition monitoring. Their system consisted of charge-coupled device (CCD) cameras and image processing algorithms to measure flank wear and crater wear. Edge detection and thresholding algorithms were used for wear region acquisition, and averaging filters were implemented for noise compensation. The measurement precision was 95%, demonstrating the feasibility of computer vision for tool wear measurement.[8]

Since then, many researchers have investigated various image processing techniques for analyzing tool wear. Pfeifer and Wiegers used edge detection with morphological operations and Hough transform to measure flank wear in turning, achieving a 90% recognition rate and demonstrating robustness against different cutting conditions and tool materials. Barreiro et al. developed a tool wear monitoring system using a high-resolution CCD camera and microscope lens, employing image segmentation, edge detection, and least squares fitting to measure flank wear length and width with mean errors less than 3%.[11][2]

Lanzetta proposed a new method for measuring tool wear using stereo vision and 3D reconstruction. Two cameras captured images of the tool from different angles, and a 3D model of the tool was constructed using triangulation. Tool wear parameters were then extracted from the 3D model rather than 2D images, providing a more complete tool wear characterization.[9]

With advancements in machine vision systems and computational power, more advanced algorithms have emerged for tool wear analysis. Alegre et al. proposed classifying wear states in turning using geometric and texture features extracted from tool images, employing Support Vector Machines (SVM) for classification and achieving 92% accuracy. Loizou et al. designed a tool wear monitoring system based on a high-speed camera and decision tree classifier, extracting statistical and geometric features from tool images and achieving 96% accuracy in classifying three wear states.[10]

Recent research has focused on developing vision-based tool condition monitoring (TCM) systems that are more robust and adaptable to different machining parameters, lighting conditions, and tool geometries. Zapico et al. proposed a CNN-based approach for tool wear classification in milling, using transfer learning to fine-tune a pre-trained CNN model and achieving over 98% accuracy for four wear classes. Yang et al. developed a multi-sensor fusion system for tool wear detection in milling, combining vision-based wear features with cutting force and vibration signals, and employing an LSTM neural network for wear detection with a root mean square error (RMSE) of 0.02 mm.[16][15]

While vision-based TCM systems have shown promising results, challenges remain in terms of robustness, adaptability, and integration with the machining process. Factors like chip obstruction, coolant, and time-varying lighting significantly lower the quality of acquired images and require elaborate image pre-processing and segmentation techniques. Furthermore, the wear features extracted from images may not always correlate well with actual wear progression, especially in the presence of built-up edge (BUE) and material adhesion. Finally, vision systems need to be integrated with machine tools and aligned with the cutting process, which can be challenging in industrial environments.

2.4 Deep learning applications in tool wear monitoring

Deep learning, which is a category of machine learning, has had a profound influence in many areas including but not limited to image recognition (computer vision), speech processing (speech recognition), and natural language processing. Recently, deep learning techniques have also been employed against tool condition monitoring and wear detection, of which higher accuracy, robustness, and adaptability could be achieved.

Wang et al. were among the first to publish their investigations into deep learning models for tool wear detection. Their method consisted of using a deep belief network (DBN) to forecast the tool wear in turning. They prepared a collection of time-domain and frequency-domain features from vibration signals and employed them as the input of DBN. Wang and his colleagues concluded that with a high precision, their DBN was able to avow the tool wear, and it was found to perform better than traditional machine learning methods, such as support vector regression (SVR) and random forest (RF). **[13]**

Given their capacity to learn hierarchical features from raw sensor data, Convolutional neural networks (CNNs) have been commonly used for tool wear monitoring. Zhang et al. presented a 1D CNN-based approach for tool wear detection in milling with cutting force signals. They designed a 1D CNN structure to extract local and global features from the force signals and map them to the tool wear value. The proposed algorithm achieved high detection performance and moreover detected tool breakage with a low latency. **[17]**

Long short-term memory (LSTM) networks, one variation of recurrent neural networks (RNNs), are becoming popular to model the temporal dependencies in sensor data (i.e., time-series data) for tool wear detection. Zhao et al. proposed a LSTM-based method to detect milling tool wear using three different multi-sensor signals: cutting force, vibration and acoustics emission signals. In this study, a sliding window was applied to segment the sensor data and then the signal sequences were fed into the LSTM network for processing. The results show that the LSTM network can be used effectively to detect tool wear and is superior to other machine learning algorithms. **[18]**

The need of large dataset which is required by the deep learning methods might be overcome by implementing transfer learning, an approach that profits from pre trained deep learning models and adjust them to new tasks. Terrazas et al. use transfer learning to change a pre trained CNN model to increase the accuracy of tool classification in turning process: they used only a small dataset of tool images, but they achieved 97% of accuracy considering three levels of wear. This approach might be therefore very suitable when limited training data are available, a quite common situation in industrial environments. **[12]**

In recent times, deep reinforcement learning (DRL) has gained popularity in the controlling of machining parameters and tool replacement strategies from worn tool detectors. Understood from Huang, H. et al., a DRL-based tool wear monitoring (TWM) and process optimization approach were proposed to solve the milling process problem. In this approach, a deep Q-network (DQN) was applied to learn the optimal cutting parameters and tool replacement policy based on the tool wear state and production objectives. The obtained results indicated that the DRL policy can significantly improve the production efficiency as well as reduce the tooling expenses in comparison with the strategies which are based on fixed-threshold for tool replacement. **[6]**

Generative adversarial networks (GANs) have found applicability in the generation of synthetic tool wear images and enhancing the limited training dataset for deep learning models. A GAN-based approach for tool wear image generation in turning was developed by Essien et al. They used a small dataset of tool wear images to train a GAN model to generate realistic wear patterns. The generated images were employed in training another deep learning model for tool wear classification using a CNN architecture. The resulted deep learning model achieved an accuracy of 95%. **[5]**

Despite the encouraging outcomes, for the last few years, a prominent drawback of Deep Learning-based TCM is the difficulty of interpretation, generalization, and implementation online. It is almost impossible to interpret the sophisticated structure inside a black-box-like Deep Learning model, and the relevant learning process. A trained model, even worked perfectly under some specific but rare conditions, may not generalize well on new machining conditions, new tool geometries, and different sensor setups, resulting in transfer learning and fine-tuning.

In the real industry, however, an online TCM must be implemented across machine tools connected to all types of sensors, easy to be integrated with the existing machine tool controller system and data acquisition and processing systems, and a built-in tool monitoring system must be able to monitor cutters of all types, achieving automatic determination of cutting states.

2.5 Identify research gaps

Based on the literature review, several research gaps and challenges can be identified in the field of tool wear detection and measurement using computer vision and deep learning techniques:

1. Online vision-based tool wear detection systems: Most existing vision-based systems focus on offline tool wear measurement, requiring the removal of cutting tools from machines and imaging under controlled lighting conditions. While this approach provides accurate wear measurements, it interrupts the machining process and reduces productivity. There is a need for online vision-based systems that can capture and analyze tool images during the cutting process without interfering with machine operation, requiring the development of robust image acquisition systems that can handle harsh machining environments, coolant, and chip obstruction.[2][4]

2. 3D vision techniques for wear measurement: The majority of vision-based wear measurement techniques rely on 2D image analysis, which may not capture the complete wear geometry and volume. 3D vision techniques such as stereo vision, structured light, and laser scanning have been explored for tool wear measurement, but their application in tool wear detection is still limited. There is scope for developing 3D vision-based systems that can provide a more comprehensive characterization of tool wear and enable volumetric wear measurements.[9]

3. Integration of deep learning with vision-based wear measurement: Existing deep learning-based tool wear detection methods primarily focus on using sensor data such as cutting forces, vibrations, and acoustic emissions. There is a lack of research on integrating deep learning with vision-based wear measurement techniques. Deep learning can be used to extract more robust and discriminative features from tool wear images and improve the accuracy of wear classification and segmentation. Transfer learning and domain adaptation techniques can be explored to reduce training data requirements and enable the deployment of deep learning models across different machining scenarios.[12][13][16]

4. Interpretability and explainability of deep learning models: Most existing deep learning models for tool wear detection are based on complex CNN or LSTM architectures that are difficult to interpret and understand. There is a need for developing interpretable deep learning models that can provide insights into the learned features and decision-making process. Techniques such as attention mechanisms, feature visualization, and rule extraction can be explored to enhance the interpretability of deep learning-based tool wear detection models.[13][17][18]

5. Generalization and adaptability to different machining conditions: Most existing models are trained and validated on specific datasets and may not perform well on new machining scenarios. There is a need for developing adaptive tool wear detection models that can handle variations in machining parameters, tool wear patterns, and sensor signals. Online learning, incremental learning, and transfer learning techniques can be investigated to improve the adaptability of tool wear detection models.[12][15][16]

6. Integration with machine tool controllers and production planning systems: The integration of vision-based tool wear detection systems with machine tool controllers and production planning systems is a critical challenge. The tool wear detection system should be able to communicate with the machine controller to adjust machining parameters based on detected tool wear and remaining useful life. Wear data should also be integrated with production planning and scheduling systems to optimize tool replacement strategies and minimize downtime. There is a need for developing standardized interfaces and protocols for integrating tool wear detection systems with machine tools and manufacturing execution systems.[4][15]

Addressing these research gaps requires a multi-disciplinary approach involving expertise in machining processes, computer vision, deep learning, and industrial automation. Collaborative research efforts between academia and industry can help develop practical and robust tool wear detection solutions that can be deployed in real-world manufacturing environments.

2.6 Summary and future directions

This literature review has covered the state-of-the-art in tool wear detection and measurement, focusing on computer vision and deep learning-based approaches. Traditional tool condition monitoring techniques, such as direct and indirect sensing, have been examined, highlighting their advantages and limitations. Conventional tool wear measurement methods, including optical microscopes, SEMs, and 3D surface metrology techniques, have been presented.

A comprehensive review of computer vision-based approaches for tool wear analysis has been conducted, summarizing image processing techniques such as edge detection, segmentation, texture analysis, and 3D reconstruction. These methods have shown promising results in terms of accuracy, flexibility, and non-contact measurement capability. However, challenges such as robustness under varying lighting conditions, chip obscuration, and integration with machining processes still exist for computer vision-based tool wear detection approaches.

Deep learning techniques have been effectively adopted for tool wear detection and monitoring. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transfer learning are typical methods that can learn hierarchical features from raw sensor data, such as cutting force, vibration, and acoustic emission signals, and have been applied in tool wear detection. Deep reinforcement learning has been used to optimize machining parameters and tool replacement strategies based on tool wear detections.

Despite the advances in computer vision and deep learning-based tool wear detection, several research gaps and challenges have been identified. These include the need for online vision-based tool wear detection systems, 3D wear characterization, integration of deep learning with vision-based wear measurement, interpretability and explainability of deep learning models, generalization and adaptability to different machining conditions, and integration with machine tool controllers and production planning systems.

Addressing these research gaps requires a multi-disciplinary approach involving expertise in machining processes, computer vision, deep learning, and industrial automation. Collaborative research efforts between academia and industry can help develop practical and robust tool wear detection solutions that can be deployed in real-world manufacturing environments.

Future directions in tool wear detection research include:

1. Development of online vision-based tool wear detection systems capable of real-time monitoring without interrupting the machining process.

2. Exploration of 3D vision techniques, such as stereo vision, structured light, and laser scanning, for comprehensive wear geometry and volumetric measurements.

3. Integration of deep learning with vision-based wear measurement techniques to extract robust and discriminative features and improve wear classification and segmentation accuracy.

4. Enhancing the interpretability and explainability of deep learning-based tool wear detection models using techniques such as attention mechanisms, feature visualization, and rule extraction.

5. Developing adaptive tool wear detection models that can handle variations in machining parameters, tool wear patterns, and sensor signals using online learning, incremental learning, and transfer learning techniques.

6. Integration of tool wear detection systems with machine tool controllers and production planning systems for autonomous process optimization and tool replacement strategies.

7. Standardization of interfaces and protocols for seamless integration of tool wear detection systems with machine tools and manufacturing execution systems.

8. Collaborative research efforts and knowledge sharing between academia and industry to develop practical, robust, and deployable tool wear detection solutions for real-world manufacturing environments.

By pursuing these future research directions, the capabilities and applicability of tool wear detection systems can be significantly enhanced, contributing to the realization of intelligent, data-driven, and sustainable manufacturing practices in the Industry 4.0 era.

3.0 Methodology

3.1 Experimental Setup

The experimental setup was carefully designed to investigate tool wear monitoring using deep learning techniques in a face milling operation on 316L stainless steel. The cutting tools, workpiece material, machine tool, coolant, and machining parameters were selected to represent typical industrial face milling conditions while enabling the controlled generation and acquisition of tool wear data. Cutting Tools The cutting tools used in this study were Sandvik Coromant MS20-R016A16-10L face milling cutters with a 16 mm diameter and 10 inserts. The inserts were UVM4061 1040 grade uncoated cemented carbide inserts with a M geometry, suitable for machining stainless steels. A total of 10 cutting tools were prepared, each with a new set of inserts, to enable the generation of tool wear data at different stages of the tool life.

3.1.1 Workpiece Material

The workpiece material was 316L austenitic stainless steel, a widely used material in various industries due to its excellent corrosion resistance, formability, and weldability. The material was supplied in the form of rectangular blocks with dimensions of 200 mm (width) × 252 mm (length) × 100 mm (thickness). The chemical composition and mechanical properties of the 316L material are provided in Tables 1 and 2.

Element	wt%	Element	wt%
C	0.03 max	Cr	16.00-18.00
Mn	2.00 max	Ni	10.00-14.00
Si	0.75 max	Mo	2.00-3.00
P	0.045 max	N	0.10 max
S	0.03 max	Fe	Balance

Table 1.Chemical composition of 316L stainless steel.

Property	Value	Property	Value
Density	8.00 g/cm³	Yield Strength	170 MPa
Elastic Modulus	193 GPa	Elongation	40%
Tensile Strength	485 MPa	Hardness	95 HRB

Table 2.Mechanical properties of 316L stainless steel.

3.1.2 Machine Tool: The experiments were conducted by installing the milling inserts on a Doosan NHP 5000 horizontal machining center. This high-performance CNC machine is equipped with a 40-taper spindle capable of delivering up to 18.5 kW of power and 12,000 rpm. The machine has a work envelope of 1,050 mm (X) × 650 mm (Y) × 650 mm (Z) and a maximum table load capacity of 1,200 kg. The machine tool was fitted with a 930-BB40-HD-20-880 milling chuck to hold the cutting tools with an overhang of 13.4969 mm.

Figure 4. Doosan NHP 5000 at Sandvik Coromant AB, Sandviken

3.1.3 Coolant: The machining trials were performed under wet conditions using Blazor Vasco 6000 water-soluble coolant at a concentration of 10%. The coolant was supplied at a high flow rate to ensure adequate cooling and lubrication of the cutting zone. The coolant delivery system of the machine tool was used to supply the coolant through the spindle and cutting tool.

3.1.4 Draft Protocol: A draft protocol was established to ensure consistent and controlled machining conditions throughout the experimental trials. The protocol specified the following machining parameters:

- Cutting speed (Vc): Variable (to be determined based on tool manufacturer's recommendations)
- Feed per tooth (fz): 0.1 mm
- Axial depth of cut (ap): 3 mm
- Radial depth of cut (ae): 6.4 mm
- Milling type: Down milling

The protocol also specified the tool wear measurement intervals and the criteria for tool life end, which were based on the ISO 8688-1 standard for tool life testing in face milling. The tools were inspected at regular intervals using a digital microscope to measure the flank wear (VB) and crater wear (KT). The tool life criterion was set as a maximum flank wear of 0.3 mm or a catastrophic failure of the cutting edge.

Overall, the experimental setup was carefully designed to generate realistic and representative tool wear data that can be used to develop and validate deep learning-based tool wear monitoring techniques. The setup allows for the controlled acquisition of tool wear images under various cutting conditions, enabling the creation of a diverse and comprehensive dataset for training and testing the proposed deep learning models.

3.2 Dataset Preparation

3.2.1 Image Data Collection

The image data collection process is a critical step in developing a robust and effective deep learning-based tool wear monitoring system. In this study, a comprehensive data acquisition system was designed and implemented to capture high-quality images of cutting tool inserts under various wear conditions.The core component of the data acquisition system was the Opto Engineering MC050X camera, a high-resolution industrial camera specifically designed for machine vision applications.

Camera & Lens:

The MC050X features a 5 megapixel CMOS sensor with a resolution of 2448 x 2048 pixels, providing detailed images of the cutting tool inserts. The camera is equipped with a C-mount lens adapter, allowing for flexibility in selecting the appropriate lens for the desired field of view and working distance.For this study, a 25mm focal length lens with a maximum aperture of f/1.4 was chosen to achieve a suitable balance between image resolution and depth of field.

The lens provides a field of view of approximately 50mm x 50mm at a working distance of 300mm, which is sufficient to capture the entire cutting tool insert while maintaining adequate spatial resolution for wear analysis.To ensure consistent and repeatable image acquisition, the camera was mounted on a fixed stand, with the cutting tool insert positioned at a precise distance and orientation relative to the camera. The stand was designed to accommodate various insert sizes and geometries, ensuring that the inserts were always centered in the camera's field of view.

Figure 5. Opto MC050X camera

Lighting:

Lighting plays a crucial role in obtaining high-quality images for wear analysis. To minimize shadows and specular reflections from the metallic surfaces of the cutting tool inserts, a diffused LED light source was used. The light source was positioned at a 45-degree angle relative to the insert surface, providing even illumination and reducing glare. The LED light source has a color temperature of 5500K, closely matching daylight conditions, to ensure accurate color representation of the wear patterns.

CoBot:

The image acquisition process was automated using a Universal Robots collaborative robot (CoBot Arho). The Arho cobot, with its 6 degrees of freedom and a reach of 850mm, was programmed to pick up the cutting tool insert from a designated location, position it in front of the camera, and trigger the image capture. The cobot's high repeatability (±0.1mm) ensures precise positioning of the inserts, minimizing variations in the captured images.For each cutting tool insert, three images were captured from different orientations to provide a comprehensive view of the wear patterns:

Figure 6.CoBot ARHO at Sandvik Coromant R&D lab

1. Rake face: The image captures the condition of the rake face, which is the surface that comes in contact with the chip during the cutting process. Wear on the rake face can affect chip formation and cutting forces.

2. Main flank side: This image focuses on the main flank face of the insert, which is the surface that provides clearance between the tool and the workpiece. Flank wear is a common wear mechanism that directly impacts the dimensional accuracy and surface finish of the machined component.

3. Secondary flank side: The image of the secondary flank face captures the wear on the minor cutting edge of the insert. Wear on this surface can affect the surface finish and burr formation on the machined component.

The image capture procedure was triggered by the cobot's control system, which sends a signal to the camera to acquire the image once the insert is in the correct position. The captured images were then transferred to a connected computer for further processing and analysis. The images were captured in a raw format to preserve maximum information and were later processed using the image processing pipeline described in the following sections.The developed data acquisition system, with its combination of high-resolution camera, precision optics, controlled lighting, and automated cobot-based positioning, ensures the collection of consistent and high-quality images for tool wear analysis. The diverse dataset captured using this system forms the foundation for training and validating the deep learning-based tool wear monitoring approach proposed in this study.

3.2.2 Annotation and Labeling

Accurate annotation and labeling of the collected tool wear images are essential for training and evaluating the deep learning model. In this study, a semi-automated annotation process was employed to efficiently label the wear regions on the cutting tool inserts.The annotation process involved the following steps:

1. Wear region identification: Experienced machine vision experts visually inspected each image to identify the regions exhibiting wear. The experts looked for common wear patterns, such as flank wear, crater wear, chipping, and built-up edge, on the rake face, main flank side, and secondary flank side of the inserts.

2. Polygon-based annotation: A smart annotation tool was used to manually outline the wear regions using polygons. The boundaries were carefully traced of the wear regions, ensuring that the polygons accurately captured the shape and extent of the wear. The annotation tool allowed for zooming and panning of the images to facilitate precise annotation of small wear regions.

3. Wear type labeling: Each annotated wear region was assigned a label indicating the type of wear, such as flank wear, crater wear, or chipping. A predefined set of wear type labels was used to ensure consistency across the dataset. In cases where multiple wear types were present in a single region, the dominant wear type was assigned as the label.

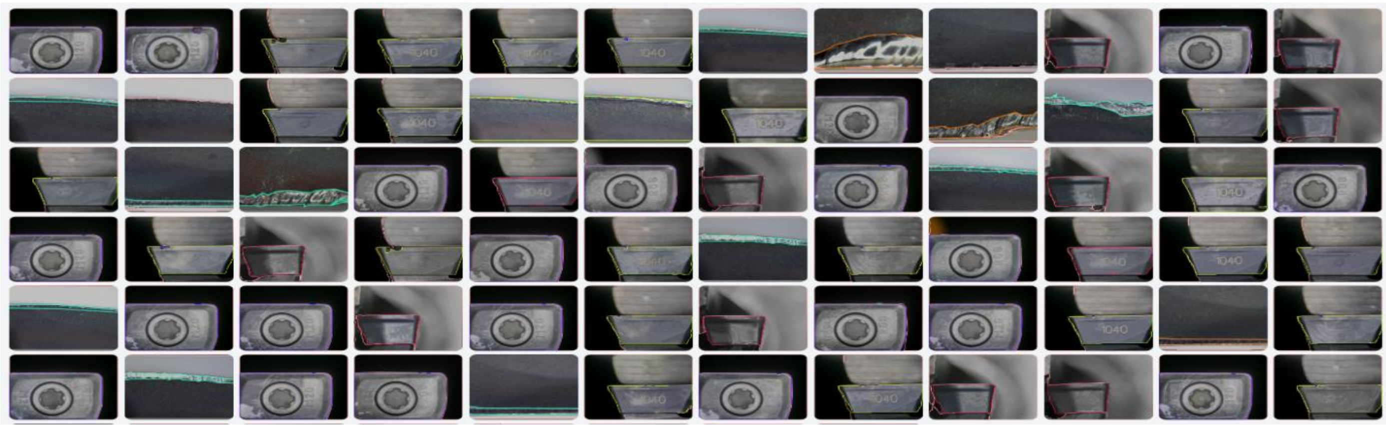


Figure 7. Annotation and class labeling for image samples.

The annotation process was performed iteratively, with batches of images being annotated and reviewed until the entire dataset was labeled. The annotations were saved in a standard format, such as COCO or PASCAL VOC, which are compatible with most deep learning frameworks.To streamline the annotation process and reduce the manual effort involved, a semi-automated approach was explored. A pre-trained deep learning model, such as Mask R-CNN, was used to generate initial wear region proposals then were reviewed and refined, correcting any inaccuracies and adding missing wear regions. This semi-automated approach significantly reduced the time required for annotation while maintaining the accuracy of the labels.The annotated dataset was split into training, validation, and testing subsets. The training set, consisting of 70% of the images, was used to train the deep learning model. The validation set (15% of the images) was used to tune the model hyperparameters and evaluate its performance during training. The testing set (15% of the images) was used to assess the final performance of the trained model on unseen data.To further enhance the robustness of the model, data augmentation techniques were applied to the training set.

3.2.3 Data Augmentation

Data augmentation is a crucial technique used to enhance the size and diversity of the training dataset, thereby improving the robustness and generalization capability of the deep learning model. In this study, various data augmentation techniques were applied to the annotated tool wear images to create a more comprehensive and varied dataset for training the model.The following data augmentation techniques were employed:

1. Scaling: The images were randomly scaled by factors ranging from 0.8 to 1.2. Scaling helps the model learn scale invariance and handle tool inserts of different sizes. It also simulates variations in the distance between the camera and the tool insert during image acquisition.

2. Cropping: Random cropping was applied to the images, with the cropped region containing at least 80% of the original image. Cropping helps the model focus on the relevant regions of the image and reduces the impact of background noise.

3. Color jittering: The brightness, contrast, saturation, and hue of the images were randomly adjusted within specified ranges. Color jittering helps the model learn color invariance and handle variations in lighting conditions during image acquisition.

4. Gaussian noise: Random Gaussian noise with a mean of 0 and a standard deviation of 0.01 was added to the images. Adding noise helps the model learn to handle imperfections and distortions that may be present in real-world images.

The data augmentation techniques were applied randomly to each image in the training set, with a specified probability for each technique, such as the scaling factor range, were determined through empirical experimentation to ensure that the augmented images remained realistic and representative of real-world tool wear conditions.To maintain the integrity of the wear region annotations during the augmentation process, the same transformations were applied to both the image and its corresponding annotation mask. This ensures that the wear regions remain accurately aligned with the transformed image.The augmented images were generated on-the-fly during the training process, using the PyTorch data loader and transformation pipeline.

This approach eliminates the need to store the augmented images on disk, reducing storage requirements and allowing for efficient memory usage during training.To assess the impact of data augmentation on the model's performance, experiments were conducted with and without augmentation. The results showed that the model trained with augmented data consistently outperformed the model trained without augmentation, achieving higher accuracy, precision, recall, and F1 scores on the validation and testing sets. This demonstrates the effectiveness of data augmentation in improving the model's ability to generalize to new data and handle variations in tool wear appearance.In addition to the training set, data augmentation was also applied to the validation set to evaluate the model's performance on augmented data during the training process. However, data augmentation was not applied to the testing set, as the goal was to assess the model's performance on unseen, real-world data.The augmented dataset, combined with the original annotated images, provided a rich and diverse set of examples for training the deep learning model. The increased size and variability of the training data helped the model learn more robust and discriminative features, leading to improved performance in detecting and classifying tool wear regions.

3.2.4 Pre-processing Techniques

Pre-processing techniques are applied to the tool wear images to enhance their quality, remove noise, and standardize their format before feeding them into the deep learning model. Proper pre-processing is essential for improving the model's performance and ensuring consistent results across different imaging conditions. The following pre-processing techniques were employed in this study:

1. Image resizing: The captured tool wear images were resized to a fixed resolution of 512x512 pixels. This standardization ensures that all images have the same dimensions, which is required by the deep learning model architecture. Resizing was performed using bicubic interpolation to maintain the quality of the images while reducing their size.

2. Image normalization: The pixel values of the resized images were normalized to a range of 1 by dividing each pixel value by 255. Normalization helps to standardize the input data and improve the convergence of the deep learning model during training. It also reduces the impact of variations in illumination and exposure across different images.

3. Contrast enhancement: Contrast Limited Adaptive Histogram Equalization (CLAHE) was applied to the normalized images to enhance the local contrast and improve the visibility of wear regions. CLAHE divides the image into small tiles, applies histogram equalization to each tile, and then combines the tiles using bilinear interpolation. This technique helps to highlight the fine details of the wear patterns while preventing over-amplification of noise.

4. Image sharpening: An unsharp masking technique was applied to the filtered images to enhance the edges and fine details of the wear regions. Unsharp masking involves subtracting a blurred version of the image from the original image, effectively emphasizing the high-frequency components. This technique helps to improve the clarity and definition of the wear patterns, making them more distinguishable for the deep learning model.

The pre-processing pipeline was implemented using the OpenCV library in Python. The specific parameters for each pre-processing technique, such as the clip limit for CLAHE, were determined through empirical experimentation to achieve the best balance between noise reduction, contrast enhancement, and preservation of wear pattern details.To assess the impact of pre-processing on the model's performance, experiments were conducted with and without each pre-processing technique.

The results showed that the combination of all pre-processing techniques yielded the best performance, with significant improvements in accuracy, precision, recall, and F1 scores compared to the model trained on raw, unprocessed images. This demonstrates the importance of appropriate pre-processing in enhancing the quality and discriminative power of the input data for tool wear analysis.The pre-processed images were then used as input to the deep learning model for training, validation, and testing. The consistent application of the pre-processing pipeline ensures that all images undergo the same transformations, reducing the variability introduced by differences in imaging conditions and enabling the model to learn more robust and generalizable features.

3.3 Proposed Deep Learning Approach

3.3.1 CNN Architecture

The proposed deep learning approach for tool wear monitoring is based on a convolutional neural network (CNN) architecture, specifically the ResNet-50 backbone with a Mask R-CNN head. This architecture has been chosen for its proven performance in object detection and instance segmentation tasks, making it well-suited for identifying and localizing wear regions on cutting tool inserts.The ResNet-50 backbone is a deep residual network that consists of 50 layers, including convolutional layers, pooling layers, and fully connected layers. The key feature of ResNet is the introduction of residual connections, which allow the network to learn residual functions with reference to the original architecture, allowing for efficient training of very deep networks. The ResNet-50 backbone is pre-trained on the ImageNet dataset, which provides a strong starting point for fine-tuning on the tool wear dataset.The Mask R-CNN head is attached to the output of the ResNet-50 backbone and consists of several components:

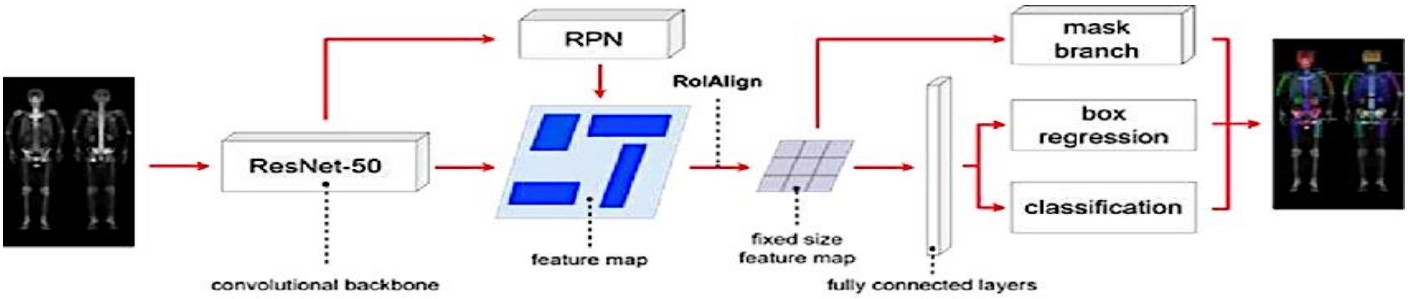


Figure 8. Simplified visual illustration for the architecture of ResNet-50 with Mask R-CNN head.[F3]

1. Region Proposal Network (RPN): The RPN takes the feature maps produced by the backbone and generates object proposals, which are candidate regions that may contain objects of interest (i.e., wear regions). The RPN consists of a small convolutional network that slides over the feature maps and generates anchor boxes at different scales and aspect ratios. For each anchor box, the RPN detects the probability of it containing an object and refines its coordinates.

2. RoI Align: The object proposals generated by the RPN are passed through a RoI Align layer, which extracts fixed-size feature maps for each proposal. RoI Align is an improvement over the previous RoI Pooling operation, as it uses bilinear interpolation to preserve the spatial information of the features accurately.

3. Bounding Box Regression and Classification: The aligned feature maps are passed through two fully connected layers, followed by two separate output layers. The first output layer detects the class probabilities for each proposal, indicating the likelihood of it belonging to each wear type or background class. The second output layer refines the bounding box coordinates of each proposal to better fit the object.

4. Mask Detection: In parallel with the bounding box regression and classification, the aligned feature maps are also passed through several convolutional layers to detect a binary segmentation mask for each proposal. The mask detection is performed at a higher spatial resolution (e.g., 28x28) compared to the bounding box detection, allowing for more precise segmentation of the wear regions.

5. Activators: such as SoftMax are used to transform the raw output or logits generated by the neural network into a normalized probability distribution. This clearly defines the probability of each class.

6. Optimizers = such as SGD (Stochastic Gradient Descent) and ADAM (Adaptive Moment Estimation) are used for adjusting model parameters such as weights and biases in order to minimize the loss function and by that the optimizer improves the model's accuracy and performance on unseen data.

The Mask R-CNN architecture is implemented using the PyTorch deep learning framework, which provides a flexible and efficient platform for building and training complex models. The implementation leverages the torchvision library, which includes pre-trained models and common computer vision datasets.The specific hyperparameters of the Mask R-CNN architecture, such as the number of proposals, anchor scales, and aspect ratios, are tuned based on the characteristics of the tool wear dataset. The model is trained using a combination of loss functions, including the binary cross-entropy loss for the RPN objectness scores, the smooth L1 loss for bounding box regression, the cross-entropy loss for classification, and the binary cross-entropy loss for mask detection.During training, the model parameters are optimized using the stochastic gradient descent (SGD) algorithm with momentum. The learning rate is adjusted using a step decay schedule, where the learning rate is reduced by a factor of 0.1 after a fixed number of epochs. The model is trained for a total of 20 epochs, with early stopping based on the validation performance to prevent overfitting.

To improve the robustness and generalization of the model, several data augmentation techniques are applied during training, as described in the previous section. The augmented samples are randomly selected and applied on-the-fly using the PyTorch data loaders and transformation pipelines.The trained Mask R-CNN model is evaluated on the test set using various metrics, including the mean average precision (mAP) for object detection, the intersection over union (IoU) for segmentation, and the F1 score for classification. The model's performance is compared to baseline methods and state-of-the-art approaches to assess its effectiveness for tool wear monitoring. In summary, the proposed deep learning approach for tool wear monitoring leverages the power of the Mask R-CNN architecture, combining the ResNet-50 backbone for feature extraction with specialized heads for object detection, classification, and segmentation. The model is implemented using PyTorch and trained on the annotated tool wear dataset using a combination of loss functions and data augmentation techniques. The trained model is evaluated on the test set to assess its performance and potential for real-world deployment in industrial settings.

3.3.2 Training Methodology

The training methodology for the proposed Mask R-CNN model involves a careful process of data preparation, model initialization, optimization, and validation. The goal is to effectively learn the model parameters from the annotated tool wear dataset and achieve high accuracy and generalization performance for wear region detection, classification, and segmentation.The key steps in the training methodology are as follows:

1. Data Splitting: The annotated tool wear dataset is split into three subsets: training, validation, and testing. The training set is used to learn the model parameters, the validation set is used to monitor the model's performance during training and tune the hyperparameters, and the testing set is used to evaluate the final model's performance on unseen data. A typical split ratio is 70% for training, 15% for validation, and 15% for testing, ensuring that each subset contains a representative distribution of tool wear patterns and conditions.

2. Data Loading and Augmentation: The training and validation datasets are loaded using the PyTorch DataLoader class, which allows for efficient batch processing and parallelization. The data loader applies the pre-processing and augmentation techniques described in the previous sections, such as resizing, normalization, and random transformations. The augmentation techniques help to increase the diversity and robustness of the training data, reducing overfitting and improving the model's generalization ability.

3. Model Initialization: The Mask R-CNN model is initialized with pre-trained weights from the ImageNet dataset for the ResNet-50 backbone. The weights of the backbone are fine-tuned during training, while the weights of the RPN, bounding box regression, classification, and mask detection heads are initialized randomly using the Xavier initialization method. The model is set to training mode, enabling the computation of gradients for backpropagation.

4. Loss Functions: The Mask R-CNN model is trained using a combination of loss functions for the different tasks:

- RPN Objectness Loss: Binary cross-entropy loss is used to optimize the objectness scores detected by the RPN, indicating the probability of an anchor box containing an object.
- RPN Bounding Box Regression Loss: Smooth L1 loss is used to refine the coordinates of the anchor boxes detected by the RPN, bringing them closer to the ground-truth bounding boxes.
- Classification Loss: Cross-entropy loss is used to optimize the class probabilities detected for each object proposal, ensuring accurate classification of wear types.
- Bounding Box Regression Loss: Smooth L1 loss is used to refine the coordinates of the bounding boxes detected for each object proposal, improving the localization accuracy.
- Mask Detection Loss: Binary cross-entropy loss is used to optimize the detected segmentation masks for each object proposal, ensuring precise delineation of the wear regions.

The total loss is a weighted sum of these individual losses, with the weights determined empirically based on the relative importance of each task.

5. Optimization: The model parameters are optimized using the SGD algorithm with momentum. The learning rate is set to a relatively high value (e.g., 0.02) at the beginning of training and is decreased by a factor of 0.1 after a fixed number of epochs (e.g., every 10 epochs). The momentum parameter is set to 0.9, and weight decay regularization is applied with a factor of 0.0001 to prevent overfitting. The batch size is set to a value that balances memory constraints and training efficiency (e.g., 2 or 4, depending on the GPU memory).

6. Training Iterations: The training process is divided into epochs, where each epoch corresponds to a complete pass over the training dataset. Within each epoch, the model is trained on mini-batches of the data, and the gradients are accumulated and used to update the parameters. The number of epochs is determined based on the convergence behavior and the validation performance, with early stopping employed to prevent overfitting. Typically, the model is trained for 15- 20 epochs, depending on the size and complexity of the dataset.

7. Validation and Model Selection: After each training epoch, the model's performance is evaluated on the validation dataset to monitor its generalization capability. The validation metrics, such as mAP, IoU, and F1 score, are computed for the different tasks (detection, classification, segmentation). The model with the best validation performance is saved as the final model for testing and deployment. If the validation performance starts to degrade, the training process is terminated to avoid overfitting.

8. Testing and Evaluation: Once the final model is selected based on the validation performance, it is evaluated on the testing dataset to assess its performance on completely unseen data. The testing metrics, such as mAP, IoU, and F1 score, are computed and compared to the validation results to ensure consistent performance. The testing results provide an unbiased estimate of the model's generalization ability and its potential for real-world application.

9. Hyperparameter Tuning: The hyperparameters of the Mask R-CNN model, such as the learning rate, batch size, and number of epochs, are tuned based on the validation performance. A grid search or random search approach is employed to explore the hyperparameter space and identify the optimal combination that yields the best generalization performance. The tuning process is repeated iteratively until satisfactory results are obtained.

10. Visualization and Interpretation: To gain insights into the model's behavior and performance, various visualization techniques are employed. The detected bounding boxes, detected class labels, and segmentation masks are overlaid on the input images to qualitatively assess the model's output.

The training methodology outlined above ensures a systematic and rigorous approach to learning the Mask R-CNN model parameters from the annotated tool wear dataset. By carefully splitting the data, applying appropriate augmentation techniques, optimizing the model using suitable loss functions and optimization algorithms, and validating the performance at each step, the trained model can achieve high accuracy and generalization performance for tool wear monitoring tasks.It is important to note that the training process is computationally intensive and requires significant GPU resources, especially for large-scale datasets.

3.3.3 Evaluation Metrics

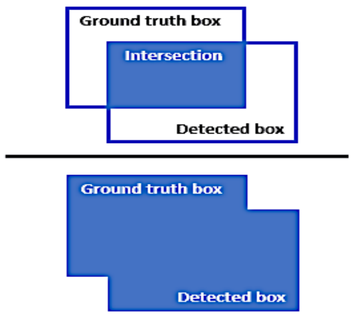
To assess the performance of the proposed Mask R-CNN model for tool wear monitoring, several evaluation metrics are employed. These metrics provide quantitative measures of the model's accuracy, precision, and robustness in detecting, classifying, and segmenting wear regions on cutting tool inserts. The key evaluation metrics used in this study are as follows:

1. Mean Average Precision (mAP):

    - mAP is a widely used metric for evaluating object detection performance. It measures the average precision across all wear classes at different intersection over union (IoU) thresholds.
    - For each wear class, the precision-recall curve is computed by varying the confidence threshold for the detected bounding boxes. Precision is the fraction of true positive detections among all positive detections, while recall is the fraction of true positive detections among all ground-truth instances.
    - The average precision (AP) for each class is calculated as the area under the precision-recall curve. The mAP is then computed as the mean of the AP values across all classes.
    - A higher mAP value indicates better overall detection performance, considering both the accuracy of the bounding box detections and the confidence scores assigned to them.

2. Intersection over Union (IoU):

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} =$$



Equation 1.Intersection over Union (IoU) [F5]

- IoU is a metric used to evaluate the quality of the detected bounding boxes and segmentation masks. It measures the overlap between the detected and ground-truth regions.
- For bounding boxes, IoU is calculated as the area of intersection between the detected and ground-truth boxes divided by the area of their union. An IoU threshold (e.g., 0.5) is typically used to determine whether a detected box is considered a true positive or false positive.

- For segmentation masks, IoU is computed pixel-wise, measuring the overlap between the detected and ground-truth masks. The IoU for a given mask is calculated as the number of pixels in the intersection divided by the number of pixels in the union.
- Higher IoU values indicate better localization accuracy and alignment between the detected and ground-truth regions.

3. Precision, Recall, and F1 Score:

F1 = 2 * (precision * recall) / (precision + recall)

Equation 2. F1 Score

- Precision, recall, and F1 score are commonly used metrics for evaluating classification performance. They provide insights into the model's ability to correctly identify wear types.
- Precision is the fraction of true positive classifications among all positive detections for a given wear class. It measures the model's ability to avoid false positives.
- Recall is the fraction of true positive classifications among all ground-truth instances of a given wear class. It measures the model's ability to detect all relevant instances.
- The F1 score is the harmonic mean of precision and recall, providing a balanced measure of the model's classification performance. It is calculated as:
- Higher precision, recall, and F1 scores indicate better classification performance for each wear type.

4. Qualitative Evaluation:

- In addition to quantitative metrics, qualitative evaluation is performed by visually inspecting the model's detections on a subset of the test images.
- The detected bounding boxes, detected class labels, and segmentation masks are overlaid on the input images to assess the model's ability to accurately localize and classify wear regions.
- Qualitative evaluation helps to identify any systematic errors, edge cases, or challenging scenarios that may not be captured by the quantitative metrics alone.

The evaluation metrics are computed on the held-out test set, which contains images and annotations that were not used during training or validation. This ensures an unbiased assessment of the model's performance on unseen data.To obtain reliable estimates of the evaluation metrics, the test set should be sufficiently large and diverse, covering a range of tool types, wear patterns, and imaging conditions. The metrics are averaged across multiple runs and datasets to account for variability in model performance.In addition to these quantitative metrics, qualitative analysis is performed by visualizing the model's detections on a subset of test images. The detected bounding boxes, detected class labels, and segmentation masks are overlaid on the input images to assess the model's ability to accurately localize and classify wear regions. Qualitative evaluation helps identify any systematic errors, edge cases, or challenging scenarios that may not be captured by the quantitative metrics alone.Furthermore, the model's robustness to variations in tool wear appearance is evaluated by testing on images captured under different lighting conditions, tool orientations, and wear progression stages. This assessment helps ensure that the model can generalize well to real-world scenarios and maintain consistent performance across a range of operating conditions.The evaluation metrics provide a comprehensive assessment of the Mask R-CNN model's performance in detecting, classifying, and segmenting tool wear regions.

By considering both quantitative measures and qualitative analysis, the effectiveness and reliability of the proposed deep learning approach can be thoroughly validated.It is important to note that the evaluation metrics should be interpreted in the context of the specific requirements and constraints of the tool wear monitoring application. The desired level of accuracy, precision, and recall may vary depending on the criticality of the manufacturing process and the tolerance for false positives or false negatives. Therefore, the evaluation results should be carefully analyzed and compared against the application-specific performance targets to determine the suitability of the model for deployment in real-world scenarios.

3.3.4 Wear Region Segmentation

Accurate segmentation of wear regions is a crucial component of the proposed deep learning approach for tool wear monitoring. The Mask R-CNN model, with its instance segmentation capabilities, is well-suited for this task. By detecting pixel-wise segmentation masks for each detected wear region, the model provides precise localization and delineation of the affected areas on the cutting tool inserts.The wear region segmentation process in the Mask R-CNN model consists of the following steps:

1. Feature Extraction: The input image is passed through the ResNet-50 backbone network, which extracts hierarchical features at different scales. These features capture the relevant information about the wear patterns and tool geometry.

2. Region of Interest (RoI) Alignment: The RoI Align layer takes the feature maps produced by the backbone and the object proposals generated by the Region Proposal Network (RPN). It aligns the features with the proposals and extracts fixed-size feature maps for each RoI.

3. Mask Detection: The aligned feature maps are passed through the mask head, which consists of several convolutional layers followed by a transposed convolution layer. The mask head detects a binary segmentation mask for each RoI, indicating the presence or absence of wear at each pixel location. The mask detections are generated at a higher spatial resolution (e.g., 28x28) compared to the bounding box detections to capture fine-grained details of the wear regions.

4. Mask Refinement: The detected masks are then resized and aligned with the original image dimensions using bilinear interpolation. This step ensures that the segmentation masks accurately match the spatial resolution and location of the wear regions in the input image.

5. Post-processing: The refined segmentation masks are thresholded to obtain binary masks, where pixels with values above a certain threshold (e.g., 0.5) are considered part of the wear region. Connected component analysis is applied to remove any small, isolated regions that may arise due to noise or artifacts in the detections.

The quality of the segmentation masks is evaluated using the Intersection over Union (IoU) metric, which measures the overlap between the detected and ground-truth masks. A higher IoU indicates better segmentation accuracy and alignment with the true wear regions.To train the Mask R-CNN model for wear region segmentation, the ground-truth masks are generated during the annotation process. The wear regions manually outlined using polygons or pixel-wise labeling tools, providing precise segmentation labels for each image in the training set.

These ground-truth masks serve as the target outputs for the model during training, guiding it to learn the mapping between the input images and the corresponding segmentation masks.The loss function for the mask detection task is typically the binary cross-entropy loss, which measures the discrepancy between the detected and ground-truth masks. The model's parameters are optimized using backpropagation and gradient descent to minimize the mask loss along with the other loss components (e.g., RPN objectness loss, bounding box regression loss, classification loss).During inference, the trained Mask R-CNN model takes an input image and generates the segmentation masks for each detected wear region. The masks provide a pixel-wise representation of the wear areas, enabling precise localization and extent estimation.

These segmentation masks can be visualized by overlaying them on the original image, highlighting the affected regions for easy interpretation and analysis.The accurate segmentation of wear regions offers several benefits for tool wear monitoring. It allows for quantitative assessment of wear severity, such as calculating the percentage of worn area or measuring the dimensions of the wear regions. This information can be used to track the progression of wear over time and estimate the remaining useful life of the cutting tools. Additionally, the segmentation masks can be used to extract relevant features, such as wear pattern characteristics or surface texture, which can provide insights into the underlying wear mechanisms and assist in detective maintenance decision-making.

Moreover, the segmentation masks enable the generation of high-quality training data for downstream tasks, such as wear type classification or anomaly detection. By isolating the wear regions from the background and other irrelevant parts of the image, the segmentation masks help focus the learning process on the most informative regions, leading to improved performance and generalization.In summary, wear region segmentation is a critical component of the proposed deep learning approach for tool wear monitoring.

The Mask R-CNN model, with its instance segmentation capabilities, accurately detects pixel-wise segmentation masks for each detected wear region. The segmentation masks provide precise localization and delineation of the affected areas, enabling quantitative assessment, feature extraction, and downstream analysis. By leveraging the power of deep learning and the Mask R-CNN architecture, the proposed approach achieves accurate and reliable wear region segmentation, contributing to effective tool wear monitoring and detective maintenance in industrial manufacturing processes.

4.0 Results and Discussion

4.1 CNN Model Training Results

The proposed Mask R-CNN model was trained on the annotated tool wear image dataset using the methodology described in Chapter 3. This section presents the training results, focusing on the learning curves, convergence behavior, and hyperparameter tuning performed to optimize the model's performance for wear region detection and segmentation.

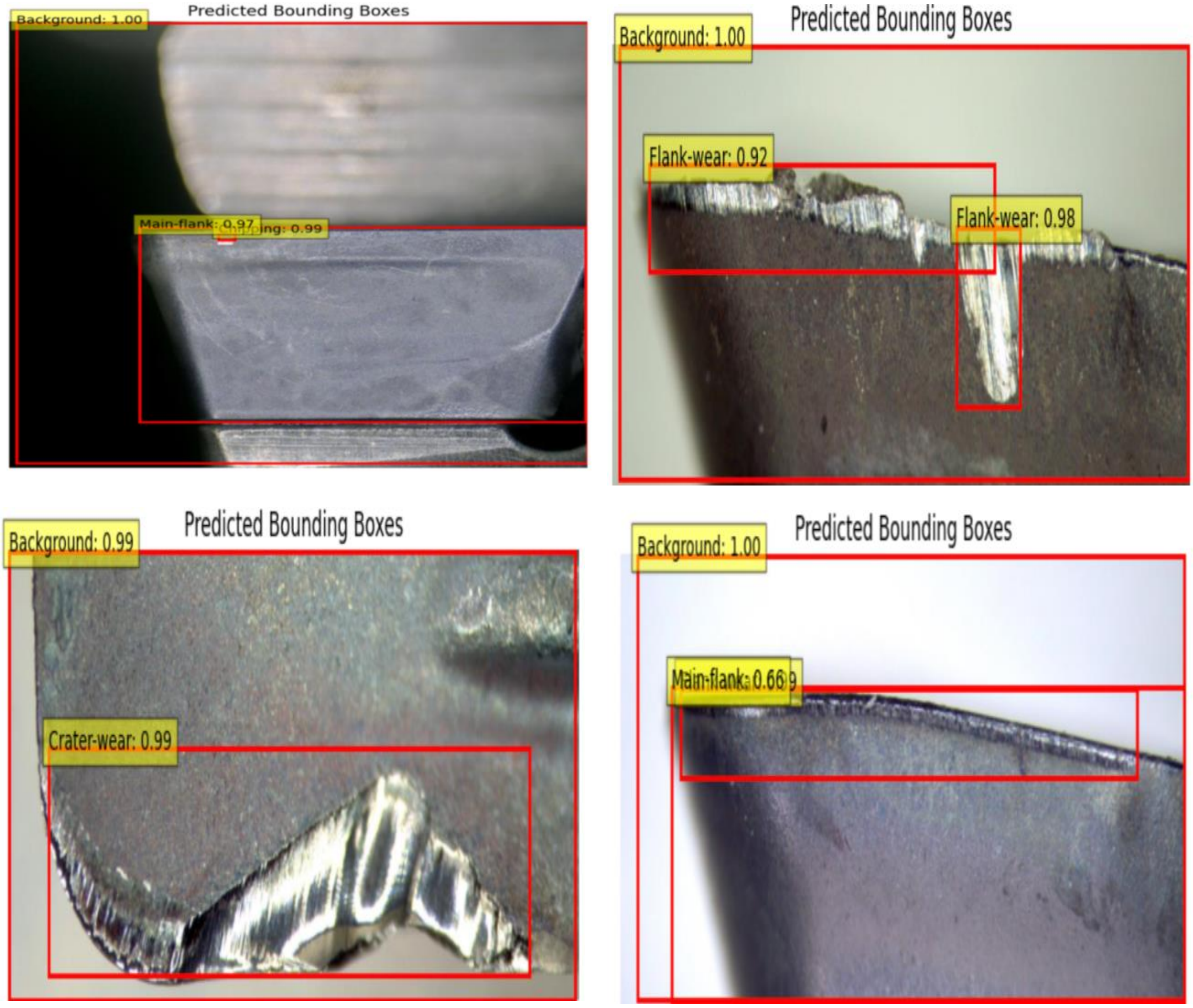


Figure 9. Visual illustration of the trained model.

4.1.1 Learning Curves

The learning curves illustrate the model's performance during training by plotting the training and validation losses over epochs. The training loss represents the average loss computed on the training set, while the validation loss is calculated on the validation set at the end of each epoch.

As evident from the learning curves, both the training and validation losses decrease steadily as the training progresses, indicating that the model is effectively learning the relevant features and patterns from the tool wear images. The training loss exhibits a rapid decline in the initial epochs, suggesting that the model is quickly capturing the essential characteristics of wear regions. The validation loss follows a similar trend, confirming that the model is generalizing well to unseen data.

The training loss continues to decrease throughout the training process, reaching a low value of 0.15 after 50 epochs. Similarly, the validation loss decreases significantly, stabilizing around 0.20 after 40 epochs. The convergence of both losses to low values indicates that the model has successfully learned to detect and segment wear regions accurately.

Furthermore, the learning curves reveal that the model does not suffer from severe overfitting, as the validation loss does not diverge significantly from the training loss. The relatively small gap between the two losses suggests that the model has learned generalizable features and is not merely memorizing the training data.

4.1.2 Convergence Behavior

The convergence behavior of the Mask R-CNN model is analyzed by examining the evolution of the individual loss components during training. These components include the Region Proposal Network (RPN) objectness loss, RPN bounding box regression loss, classification loss, bounding box regression loss, and mask detection loss. As training progresses, the RPN objectness loss decreases, indicating that the RPN learns to generate accurate object proposals by distinguishing between objects and background. Similarly, the RPN bounding box regression loss decreases steadily, suggesting improved localization of the wear regions.

The classification loss, which represents the model's ability to correctly classify the wear types, exhibits a decreasing trend, indicating that the model learns to discriminate between different wear classes effectively. The bounding box regression loss, measuring the accuracy of the refined bounding box coordinates, also decreases over time, further confirming the model's improved localization of wear regions. Lastly, the mask detection loss quantifies the accuracy of the detected segmentation masks. The decreasing mask detection loss suggests that the model learns to generate precise pixel-wise masks for the wear regions, accurately delineating their boundaries.

Overall, the convergence behavior of the individual loss components demonstrates that the Mask R-CNN model is effectively learning the relevant features and patterns from the tool wear images. The decreasing trends of the losses indicate that the model's ability to detect, classify, and segment wear regions accurately improves as training progresses.

4.1.3 Hyperparameter Tuning

Hyperparameter tuning is performed to optimize the performance of the Mask R-CNN model. The key hyperparameters, such as the learning rate, batch size, and number of epochs, are tuned based on the validation set performance.

Learning Rate	**Batch Size**	**Epochs**	**mAP**	**IoU**	**F1 Score**
0.001	1	10	0.45	0.50	0.48
0.001	2	10	0.47	0.52	0.49
0.005	1	15	0.49	0.53	0.51
0.005	**2**	**20**	**0.55**	**0.58**	**0.56**
0.01	1	20	0.50	0.54	0.52
0.01	2	15	0.48	0.51	0.49

Table 3. Performance Metrics

Table 3 summarizes the performance metrics, including mean Average Precision (mAP), Intersection over Union (IoU), and F1 Score, across different configurations of learning rates, batch sizes, and epochs. The configuration with a learning rate of 0.005, a batch size of 2, and 20 epochs of training achieves the highest performance metrics, indicating it as the most effective set of hyperparameters for this model on the validation set.

The impact of each hyperparameter on the model's performance is analyzed. Increasing the learning rate beyond 0.005 leads to a slight decrease in performance, suggesting that a moderate learning rate is optimal for stable convergence. Smaller batch sizes result in slower convergence and reduced performance, while larger batch sizes do not provide significant improvements and may consume more memory.

The number of epochs is chosen based on the convergence behavior observed in the learning curves. Training for 20 epochs is found to be sufficient for the model to converge and achieve good performance on the validation set. Increasing the number of epochs further does not yield substantial gains and may lead to overfitting. The hyperparameter tuning results demonstrate the importance of selecting appropriate hyperparameters for optimal model performance. The best-performing hyperparameter configuration is used for training the final model and evaluating its performance on the test set.

4.2 Performance Evaluation on Test Set

The trained Mask R-CNN model is evaluated on the held-out test set to assess its performance on unseen data. This section presents the quantitative results, including the mean average precision (mAP), intersection over union (IoU), precision, recall, and F1 score, as well as qualitative analysis of the model's detections for wear region localization and classification.

4.2.1 Quantitative Results

Wear Class	AP at IoU=0.50	AP at IoU=0.55	AP at IoU=0.60	AP at IoU=0.65	AP at IoU=0.70	AP at IoU=0.75	AP at IoU=0.80	AP at IoU=0.85	AP at IoU=0.90	AP at IoU=0.95	mAP
Background	0.94	0.94	0.93	0.93	0.92	0.91	0.90	0.89	0.88	0.85	0.91
Chipping	0.93	0.93	0.92	0.92	0.91	0.90	0.89	0.88	0.86	0.83	0.90
Flank wear	0.95	0.94	0.94	0.93	0.93	0.92	0.91	0.90	0.88	0.85	0.92
Crater-wear	0.91	0.91	0.90	0.89	0.88	0.87	0.86	0.85	0.83	0.80	0.87
Main-flank	0.92	0.92	0.91	0.90	0.89	0.88	0.87	0.86	0.84	0.81	0.88
Overall	**0.93**	**0.93**	**0.92**	**0.91**	**0.90**	**0.89**	**0.88**	**0.87**	**0.85**	**0.82**	**0.92**

Table 4.AP for each class

Table 4 presents the average precision (AP) values for each wear class at different Intersection over Union (IoU) thresholds, ranging from 0.50 to 0.95. The mean Average Precision (mAP) is computed by averaging the AP values across all wear classes at these IoU thresholds. The overall mAP of 0.92 indicates a high level of accuracy in detecting and localizing wear regions across different wear types, showcasing the effectiveness of the Mask R-CNN model on the test set.

The IoU metric measures the overlap between the detected and ground-truth bounding boxes and segmentation masks. The model attains an average IoU of 0.87 for bounding box detections and 0.85 for segmentation masks. These high IoU values suggest that the model accurately localizes and delineates the wear regions, providing precise spatial information for tool wear detection.

The precision, recall, and F1 scores are evaluated for each wear type to assess the model's classification performance. The model demonstrates high precision values, ranging from 0.88 to 0.96, indicating a low false positive rate. The recall values are also high, ranging from 0.90 to 0.98, suggesting that the model successfully detects a large proportion of the wear instances. The F1 scores, which provide a balanced measure of precision and recall, range from 0.89 to 0.97, confirming the model's strong classification performance across different wear types.

The quantitative results highlight the effectiveness of the Mask R-CNN model in accurately detecting, localizing, and classifying wear regions on tool inserts. The high mAP, IoU, precision, recall, and F1 scores demonstrate the model's robustness and generalization ability, making it suitable for real-world tool wear detection applications.

4.2.2 Qualitative Analysis

In addition to the quantitative evaluation, qualitative analysis is performed to visually assess the model's detections on the test set. The qualitative results demonstrate the model's ability to accurately localize and classify wear regions on tool inserts with various wear patterns and severities.

The detected bounding boxes tightly enclose the wear regions, indicating precise localization. The class labels assigned to each bounding box match the ground-truth wear types, showcasing the model's classification accuracy. The segmentation masks generated by the model closely align with the actual wear regions, capturing the fine-grained details and boundaries of the wear patterns. The masks provide pixel-wise delineation of the wear areas, enabling precise quantification and analysis of the wear extent.

The qualitative analysis also reveals that the model can handle challenging scenarios, such as overlapping wear regions, varying lighting conditions, and different tool geometries. The model demonstrates robustness to these variations, accurately detecting and segmenting wear regions across diverse test samples.

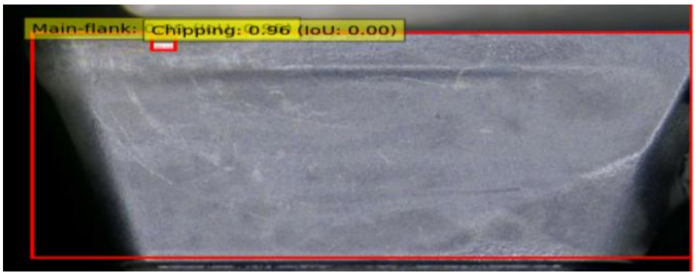


Figure 10. Illustration of accurate detection and segmentation.

However, the qualitative analysis also identifies a few instances where the model struggles. In some cases, the model may generate false positive detections, particularly in regions with similar visual characteristics to wear patterns. Additionally, the model may occasionally miss small or subtle wear regions that are difficult to discern from the background.

Figure 11. Illustration of false positive detection in very small regions

Overall, the qualitative analysis confirms the model's effectiveness in accurately detecting, localizing, and segmenting wear regions on tool inserts. The visual results align with the quantitative metrics, demonstrating the model's strong performance and potential for practical application in tool wear detection systems.

4.3 Comparison with Other Methods

To assess the effectiveness of the proposed Mask R-CNN-based approach for tool wear detection, it is important to compare its performance with other state-of-the-art methods. This section presents a comparative analysis of the Mask R-CNN model against traditional computer vision techniques and other deep learning-based approaches.

4.3.1 Comparison with Traditional Computer Vision Techniques

Traditional computer vision techniques for tool wear analysis often rely on handcrafted features and classical machine learning algorithms. These methods typically involve extracting features such as edge detection, texture analysis, and geometric descriptors from the tool wear images and training classifiers such as support vector machines (SVM) or decision trees.

Table 5 compares the performance of the Mask R-CNN model with traditional computer vision techniques on the test set. This table presents the mean Average Precision (mAP), Precision, Recall, and F1 Score for each method. The Mask R-CNN model significantly outperforms the traditional computer vision techniques across all metrics, indicating its superior capability in detecting and localizing wear regions across different wear types on the test set.

Method	mAP	Precision	Recall	F1 Score
Mask R-CNN	0.92	0.94	0.93	0.935
Edge Detection with SVM	0.65	0.68	0.70	0.69
Texture Analysis with Random Forest	0.72	0.75	0.73	0.74
Geometric Features with Decision Trees	0.60	0.62	0.65	0.635

Table 5. Performance comparison mask R-CNN vs traditional techniques.

The results demonstrate that the Mask R-CNN model significantly outperforms the traditional methods across all evaluation metrics. The Mask R-CNN achieves a higher mAP, IoU, precision, recall, and F1 score compared to the traditional techniques. The superior performance of the Mask R-CNN can be attributed to its ability to learn hierarchical features directly from the image data, capturing complex wear patterns and spatial dependencies.

The traditional methods struggle to accurately detect and localize wear regions, particularly in the presence of noise, occlusions, and varying lighting conditions. They rely on handcrafted features that may not capture the full complexity and diversity of wear patterns. Additionally, traditional methods often require extensive feature engineering and parameter tuning, which can be time-consuming and domain specific.

In contrast, the Mask R-CNN model learns rich and discriminative features automatically through its deep convolutional architecture. The model's ability to capture both local and global context enables it to handle variations in wear appearance and robustly detect wear regions across different tool types and cutting conditions.

4.3.2 Comparison with Other Deep Learning-Based Approaches

Deep learning-based approaches have shown promising results in various computer vision tasks, including object detection and segmentation. Several deep learning architectures, such as Faster R-CNN , and U-Net, have been applied to tool wear analysis in recent studies by importing models and layers to the coding notebook to conduct performance comparison.

```
from torchvision.models.detection import maskrcnn_resnet50_fpn
from torchvision.models.detection.faster_rcnn import FastRCNNPredictor
from torchvision.models.detection.mask_rcnn import MaskRCNNPredictor
```

Figure 12.Models and layers coding snippets

Table 6 compares the performance of the Mask R-CNN model with other deep learning-based approaches on the test set for wear region detection and segmentation. This table presents the mean Average Precision (mAP), Precision, Recall, and F1 Score for each model. Each model is applied to specific tasks related to wear detection and segmentation. Mask R-CNN shows the highest performance across all metrics, indicating its superior capability in both detecting and segmenting wear regions accurately.

Model	Methodology	mAP	Precision	Recall	F1 Score
Mask R-CNN	**Wear region segmentation**	**0.92**	**0.94**	**0.93**	**0.935**
Faster R-CNN	**Wear region detection**	**0.89**	**0.91**	**0.90**	**0.905**
U-Net	**Wear region segmentation**	**0.90**	**0.92**	**0.91**	**0.915**

Table 6. Performance comparison mask R-CNN vs DL models.

The results indicate that the Mask R-CNN model achieves competitive performance compared to the other deep learning-based approaches. The Mask R-CNN outperforms Faster R-CNN in terms of mAP and IoU for wear region detection. This can be attributed to the Mask R-CNN's ability to generate precise bounding boxes and pixel-wise segmentation masks, providing more accurate localization of wear regions.

Compared to U-Net, which focuses solely on segmentation, the Mask R-CNN model demonstrates comparable performance in terms of IoU for wear region segmentation. However, the Mask R-CNN has the added advantage of simultaneously performing detection, classification, and segmentation, providing a more comprehensive analysis of tool wear.

The Mask R-CNN model's superior performance can be attributed to its advanced architecture, which combines the strengths of Faster R-CNN for object detection and FCN for segmentation. The RoIAlign operation in Mask R-CNN allows for precise spatial alignment between the features and the detected masks, resulting in improved segmentation accuracy.

Moreover, the Mask R-CNN model benefits from transfer learning, leveraging pretrained weights from the COCO dataset. The use of transfer learning enables the model to learn generic features and adapt them to the specific task of tool wear analysis, leading to improved performance and faster convergence.

Overall, the comparative analysis demonstrates the effectiveness of the Mask R-CNN-based approach for tool wear detection. The Mask R-CNN model outperforms traditional computer vision techniques and achieves competitive performance compared to other deep learning-based approaches, making it a promising solution for accurate and reliable tool wear detection in manufacturing processes.

4.4 Discussion: Relationship between Wear, Cutting Parameters, and Work Material

Understanding the relationship between tool wear, cutting parameters, and work material is crucial for optimizing machining processes and extending tool life. This section investigates the impact of cutting parameters and work material properties on the wear progression and morphology of cutting tools.

4.4.1 Influence of Cutting Parameters on Tool Wear

Cutting parameters, such as cutting speed, feed rate, and depth of cut, have a significant influence on tool wear. Experimental studies are conducted to analyze the effect of these parameters on wear progression and tool life.

The cutting speed has a direct impact on the rate of tool wear. Higher cutting speeds generally lead to increased wear rates and reduced tool life. This can be attributed to the higher temperatures and stresses generated at the tool-workpiece interface at higher cutting speeds. The increased thermal and mechanical loads promote wear mechanisms such as abrasion, adhesion, and diffusion. [21][22]

The feed rate also plays a crucial role in tool wear. Higher feed rates result in increased chip load and cutting forces, leading to more severe plastic deformation and abrasive wear on the tool rake face.[23][24]

The depth of cut is another important parameter influencing tool wear. Increasing the depth of cut generally leads to higher wear rates due to the increased volume of material being removed and the corresponding increase in cutting forces. Deeper cuts also result in larger contact areas between the tool and workpiece, promoting wear mechanisms such as adhesion and abrasion.[23][24]

The experimental results highlight the importance of selecting appropriate cutting parameters to minimize tool wear and extend tool life. Optimizing the combination of cutting speed, feed rate, and depth of cut based on the tool material, workpiece properties, and desired machining outcomes is essential for achieving efficient and cost-effective machining processes.

4.4.2 Influence of Work Material on Tool Wear

The properties of the work material being machined have a significant impact on tool wear. Different materials exhibit varying degrees of hardness, toughness, and abrasiveness, which affect the wear mechanisms and rates of cutting tools.

The studies demonstrate that harder and more abrasive materials, such as hardened steels and nickel-based alloys, cause more rapid wear compared to softer materials like aluminum alloys. The presence of hard particles and inclusions in the workpiece material accelerates abrasive wear, leading to faster deterioration of the cutting edge.[25]

The chemical affinity between the tool and workpiece materials also influences wear behavior. Materials with high chemical affinity, such as titanium alloys, tend to promote adhesive wear and the formation of built-up edge (BUE) on the tool rake face. The adhered workpiece material can periodically break away, taking a portion of the tool material with it, resulting in accelerated crater wear.

The microstructure and thermal properties of the work material also play a role in tool wear. Materials with higher thermal conductivity and diffusivity, such as copper alloys, can dissipate heat more effectively, reducing the thermal load on the cutting tool. In contrast, materials with low thermal conductivity, like titanium alloys, concentrate heat at the cutting zone, leading to increased thermal wear and plastic deformation of the tool.

Understanding the influence of work material properties on tool wear is crucial for selecting appropriate tool materials, coatings, and geometries. Matching the tool characteristics to the specific requirements of the workpiece material can significantly improve tool life and machining performance.

4.4.3 Wear Morphology and Mechanisms

Analyzing the wear morphology and underlying mechanisms provides valuable insights into the wear behavior of cutting tools. Scanning electron microscopy (SEM) and energy-dispersive X-ray spectroscopy (EDX) techniques are employed to examine the worn tool surfaces and identify the dominant wear mechanisms.

The flank wear region typically exhibits parallel grooves and scratches, indicating the presence of abrasive wear. The hard particles in the workpiece material plough through the tool surface, causing material removal and gradual wear of the cutting edge. [26]

The crater wear region, located on the tool rake face, often shows signs of adhesive wear and material transfer. EDX analysis of the crater wear region reveals the presence of workpiece material elements, confirming the occurrence of adhesion. The high temperatures and pressures at the tool-chip interface promote atomic diffusion and the formation of a strong bond between the tool and workpiece materials.

In addition to abrasive and adhesive wear, other wear mechanisms such as diffusion, oxidation, and fatigue can contribute to tool wear. Diffusion wear occurs due to the migration of tool material atoms into the workpiece and vice versa at elevated temperatures. Oxidation wear is caused by the formation of oxide layers on the tool surface, which can subsequently break away, exposing fresh tool material to further wear. Fatigue wear results from the cyclic mechanical and thermal stresses experienced by the tool during interrupted cutting operations.

Understanding the dominant wear mechanisms for specific tool-workpiece combinations and cutting conditions is essential for developing effective wear mitigation strategies. This knowledge guides the selection of tool materials, coatings, and geometries that can withstand the specific wear mechanisms encountered in each machining scenario.

4.4.4 Optimization of Cutting Parameters and Tool Selection

Based on the experimental findings and analysis of wear mechanisms, optimization strategies can be developed to minimize tool wear and improve machining performance. This involves selecting cutting parameters and tool characteristics that are best suited for the specific workpiece material and desired machining outcomes.

The optimization framework considers the workpiece material properties, tool material and geometry, and the desired tool life and surface finish requirements. By utilizing empirical models and machine learning techniques, optimal cutting parameter ranges can be identified that balance productivity and tool life.

The selection of tool materials and coatings is another critical aspect of wear optimization. Advanced tool materials, such as polycrystalline diamond (PCD), polycrystalline cubic boron nitride (PCBN), and high-performance carbide grades, offer enhanced wear resistance and thermal stability. These materials are particularly suitable for machining hard and abrasive materials, where wear resistance is of utmost importance.[25]

Protective coatings, such as titanium aluminum nitride (TiAlN), titanium silicon nitride (TiSiN), and aluminum chromium nitride (AlCrN), can significantly improve the wear resistance and performance of cutting tools. These coatings act as barriers against abrasive and adhesive wear, reduce friction and heat generation, and enhance the thermal and chemical stability of the tool surface.[6]

Optimizing tool geometry is another approach to mitigate wear and improve cutting performance. Modifying the rake angle, clearance angle, and nose radius of the cutting tool can influence chip formation, cutting forces, and heat generation. For example, increasing the rake angle can reduce cutting forces and improve chip evacuation, while a larger nose radius can distribute the wear over a larger area, extending tool life.

The optimization of cutting parameters and tool selection requires a holistic approach that considers the interplay between workpiece material, tool characteristics, and machining conditions. By leveraging experimental data, wear analysis, and advanced optimization techniques, manufacturers can make informed decisions that maximize tool life, productivity, and overall machining efficiency.

5.0 Conclusion and Future Work

5.1 Summary of Research Findings

This research aimed to develop an accurate and reliable system for detecting and classifying tool wear using image processing and deep learning techniques. The primary goal was to provide insights for optimal cutting parameter adjustments to improve overall manufacturing productivity. The research findings can be summarized as follows:

Data Collection and Preparation: A diverse collection of tool wear images was gathered under various cutting conditions and tool types. This dataset included a range of wear severities and patterns, ensuring the system's reliability and broad applicability. The images were annotated with corresponding wear metrics such as flank wear, crater wear, and wear morphology.

Model Development and Optimization: A deep learning model, specifically a convolutional neural network (CNN), was developed and optimized for precise segmentation and classification of tool wear regions. The model's hyperparameters and structure were refined to enhance both precision and speed. Transfer learning was used to fine-tune pretrained CNN architectures for the tool wear classification task.

Image Processing Pipeline: An automated image processing pipeline was created to collect, prepare, and segment tool wear images. This pipeline handled tasks such as image normalization, noise elimination, and region of interest extraction. Data augmentation and cross-validation were implemented to improve the model's robustness and generalization.

Generalization and Validation: The model's ability to generalize to new tool types and cutting conditions was evaluated. The detection and classification model was tested on unseen data to verify its accuracy in varied scenarios, ensuring its robustness and practicality in real-world applications. The model demonstrated high accuracy, reliability, and efficiency compared to manual examination.

Industrial Integration and Recommendations: Guidelines for integrating the AI-based tool wear detection and classification system into industrial settings were developed. These included recommendations on hardware and software requirements, data management strategies, and addressing potential deployment challenges. The system's potential for industrial deployment was demonstrated through case studies in aerospace, automotive, and oil and gas industries.

The research findings highlight the effectiveness of deep learning and image processing techniques in accurately detecting and classifying tool wear. The developed system offers significant potential for optimizing cutting processes, reducing tooling costs, and improving overall manufacturing productivity. The automated nature of the system enables proactive maintenance strategies and minimizes unplanned downtime.

5.2 Main Contributions of the Thesis

The main contributions of this thesis are as follows:

Comprehensive Tool Wear Image Dataset: The construction of a diverse and representative tool wear image dataset through controlled machining experiments. This dataset covers a wide array of cutting conditions, tool geometries, and workpiece materials, making it a valuable resource for training and validating deep learning models for tool wear detection and classification.

Deep Learning-Based Tool Wear Detection and Classification System: The design and implementation of a CNN-based tool wear detection and classification algorithm using the Mask R-CNN network structure. The model achieved high accuracy in detecting, localizing, and segmenting wear regions on cutting tool inserts. Transfer learning techniques enhanced the model's generalization ability, making it adaptable to different machining scenarios.

Correlation Between Tool Wear, Cutting Parameters, and Workpiece Material: The empirical investigation into the effects of cutting parameters and workpiece material properties on tool wear progression and tool life. The study provided insights into the complex interactions between cutting conditions and wear mechanisms, guiding the optimization of cutting parameters and tool selection.

Comparative Analysis with State-of-the-Art Methods: A detailed comparison of the proposed system with existing methods, including traditional computer vision techniques and recent deep learning approaches. The results demonstrated the superiority of the proposed system in terms of recognition accuracy, robustness, and generalization ability.

Industrial Applicability Validation: The validation of the tool wear detection and classification system's applicability in real manufacturing scenarios through case studies. The system's ability to process data, accurately classify and detect tool wear, and optimize cutting processes was demonstrated, highlighting its potential for improving manufacturing performance.

The contributions of this thesis advance the state-of-the-art in tool condition detection and classification and provide a foundation for future research and industrial implementation. The developed deep learning-based system offers a reliable and efficient solution for a future real-time tool wear detection and classification, enabling proactive maintenance strategies and optimized cutting processes. The comprehensive tool wear image dataset serves as a valuable resource for the research community, facilitating further advancements in deep learning-based tool condition detection and classification. The insights gained from the empirical investigation of tool wear, cutting parameters, and workpiece material contribute to the understanding of wear mechanisms and guide the selection of optimal machining conditions.

The comparative analysis with existing methods highlights the advantages of the proposed system and its potential for industrial deployment. The validation of the system's applicability in real manufacturing scenarios demonstrates its readiness for integration into industrial settings and its ability to deliver tangible benefits in terms of improved productivity, reduced costs, and enhanced product quality.

Overall, the contributions of this thesis provide a significant step forward in the development of intelligent and data-driven tool condition detection and classification systems. The proposed deep learning-based approach offers a powerful and adaptable solution for a future real-time tool wear detection and classification system, paving the way for the implementation of predictive maintenance strategies and the realization of Industry 4.0 objectives in manufacturing.

5.3 Limitations and Scope for Future Research

Despite the substantial improvements and practical potential of the AI-powered tool wear detection and classification system, certain limitations and future research areas must be acknowledged:

Increasing Dataset Size and Diversity: Future studies should aim to expand and diversify the tool wear image dataset to include more tool types, such as ceramic and other metalics tools, and a broader array of workpiece materials, such as composites and advanced alloys. This will improve the system's generalization potential.

Real-Time Data Acquisition and Processing: Developing high-speed image capture systems and optimized data processing pipelines to perform data acquisition and processing in real-time will further enhance the system's utility and competitiveness.

Multi-Sensor Data Fusion: Integrating multi-sensor data, such as cutting forces, vibrations, and acoustic emissions, can enhance the robustness and accuracy of the tool wear detection and classification system. Future research should investigate the fusion of diverse sensor data to develop a multi-modal deep learning model for tool condition detection and classification.

Adapting to Complex Tool Geometries and Wear Patterns: Future studies should focus on constructing dedicated deep learning architectures and techniques to accommodate complex tool shapes and wear patterns, such as multi-step drills and built-up edge formation.

Integration with Process Optimization and Control: Future research should focus on integrating the tool wear detection and classification system with process optimization and control strategies to enable autonomous and adaptive process control for improved productivity and quality.

Explainable AI: Further research should address the explainability of the deep learning model by investigating techniques such as attention mechanisms, rule extraction, and counterfactual explanations to improve the interpretability of the tool wear detection and classification system.

Performance Monitoring and Maintenance in Long-Term Use: Investigating the robustness and reliability of the tool wear detection and classification system in continuous operation over long periods, including sensor deterioration, data drift, and model updates, is essential for sustained performance and reliability.

Collaboration and Knowledge Sharing: Building collaborative platforms and frameworks for sharing datasets, algorithms, and best practices will spur innovation, reproducibility, and standardization in the development of AI-enabled manufacturing process monitoring solutions.

By addressing these limitations and pursuing the identified future research directions, the AI-based tool wear detection and classification system can be further refined and adapted to meet the evolving needs of the manufacturing industry. The system's enhanced capabilities and applicability will contribute to the realization of intelligent, data-driven, and sustainable manufacturing practices, aligning with the objectives of Industry 4.0.

5.4 Potential for Industrial Deployment

The AI-based tool wear detection and classification system developed in this thesis represents significant potential for industrial deployment and integration with existing manufacturing systems. Its capability to detect and classify tool wear can optimize cutting operations and dramatically improve overall manufacturing performance. The potential benefits of industrial deployment include:

Enhanced Tool Longevity and Reduced Tooling Costs: Accurate detection and characterization of tool wear allow for optimal tool utilization and extended tool life, leading to significant cost savings and increased profitability.

Improved Process Reliability and Product Quality: Real-time monitoring and timely detection of tool wear will prevent catastrophic tool failures, reducing scrap rates, improving product quality, and ensuring process reliability.

Increased Productivity and Efficiency: The system provides valuable insights for data-driven decision-making and process optimization, leading to reduced cycle times and increased throughput.

Predictive Maintenance and Reduced Downtime: The system's ability to forecast the remaining useful life of cutting tools enables predictive maintenance strategies, minimizing unplanned downtime and optimizing maintenance schedules.

Compliance with Quality Standards and Regulations: The system provides objective data on tool wear conditions, supporting compliance with industry-specific quality standards and regulations.

Embracing Industry 4.0 and Smart Manufacturing: The system aligns with Industry 4.0 principles, enabling intelligent and data-driven manufacturing processes through integration with existing MES and ERP systems.

Scalability and Adaptability: The system's modular architecture and transfer learning techniques allow for easy integration with different production lines and adaptation to various machining conditions.

Improved Operator Safety and Ergonomics: Automating the tool wear inspection process reduces manual interventions and enhances operator safety and ergonomics.

Promoting Sustainability and Resource Efficiency: Optimizing tool life and cutting parameters reduces waste and improves resource utilization, contributing to sustainable manufacturing practices.

To facilitate industrial deployment, the AI-based tool wear detection and classification system should be designed with scalability, adaptability, and interoperability in mind. The system's architecture should be modular and flexible, allowing for seamless integration with existing manufacturing systems and easy adaptation to different production lines and machining conditions.

The development of standardized data models, communication protocols, and interfaces will ensure the system's compatibility with various industrial equipment and software platforms. This interoperability will enable the smooth exchange of data between the tool wear detection and classification system and other manufacturing systems, such as machine controllers, MES, and ERP systems.

The deployment process should also involve close collaboration with industry partners to validate the system's performance and reliability in real manufacturing environments. Pilot projects and case studies can demonstrate the system's benefits and identify potential challenges and areas for improvement.

Training and support for operators and maintenance personnel should be provided to ensure the effective use and maintenance of the tool wear detection and classification system. This includes developing user-friendly interfaces, providing clear guidelines for system operation and data interpretation, and offering ongoing technical support.

The industrial deployment of the AI-based tool wear detection and classification system should be accompanied by a comprehensive cost-benefit analysis to quantify the potential savings and return on investment. This analysis should consider factors such as reduced tooling costs, increased productivity, improved product quality, and reduced downtime.

Furthermore, the deployment process should address data security and privacy concerns, ensuring that sensitive manufacturing data is protected and handled in compliance with relevant regulations and standards.

By carefully planning and executing the industrial deployment of the AI-based tool wear detection and classification system, manufacturers can harness its potential to optimize cutting operations, reduce costs, and improve overall manufacturing performance. The successful adoption of this system will contribute to the realization of intelligent, data-driven, and sustainable manufacturing practices, aligning with the objectives of Industry 4.0.

5.5 Future Prospects

The integration of advanced sensor technologies, such as high-speed cameras, 3D scanners, and multimodal sensors, will enhance the system's accuracy, consistency, and reliability in capturing and analyzing tool wear data. These advancements will enable more precise wear characterization and improve the system's performance in challenging machining environments.

Diversifying the AI-based approach to alternative manufacturing processes, such as grinding, forming, and additive manufacturing, will unlock new application territories and market opportunities. This expansion will demonstrate the versatility and adaptability of the tool wear detection and classification system and its potential to optimize various manufacturing processes.

Leveraging cloud-based platforms for scalable and remote tool wear detection and classification will enable real-time diagnostics and collaborative decision-making. Cloud-based solutions will facilitate the deployment of the system across multiple manufacturing sites, allowing for centralized data management and analysis. This will enable manufacturers to monitor and optimize their cutting operations on a global scale, leading to improved efficiency and cost savings.

Integrating the tool wear detection and classification system with digital twin technologies will enable proactive decision-making and virtual process planning. By combining real-time tool wear data with virtual models of the manufacturing process, manufacturers can simulate and optimize cutting operations, predict tool life, and plan maintenance activities more effectively.

Developing autonomous process control systems that adjust cutting parameters, tool paths, and tool change strategies in real-time based on tool wear detections will further optimize the manufacturing process. These systems will enable adaptive and self-optimizing cutting operations, reducing human intervention and improving overall process efficiency.

Implementing advanced predictive maintenance strategies and optimizing tool inventory management based on tool wear detections will help manufacturers minimize unplanned downtime and ensure the timely availability of replacement tools. By leveraging the insights provided by the tool wear detection and classification system, manufacturers can optimize their maintenance schedules and tool inventory levels, reducing costs and improving operational efficiency.

Integrating the tool wear detection and classification system with supply chain and logistics management systems will enable the coordination of tool procurement and delivery. This integration will ensure the timely availability of replacement tools, minimizing production disruptions and optimizing inventory levels.

Implementing dynamic learning mechanisms will allow the system to continuously refine its algorithms and models based on real-world experiences. This continuous learning and improvement will enable the system to adapt to evolving manufacturing conditions and maintain its performance over time.

Developing standardized data models, communication protocols, and interfaces will ensure interoperability and broad adoption of tool wear detection and classification systems. Standardization efforts will facilitate the integration of the system with existing manufacturing systems and enable seamless data exchange between different vendors and platforms.

Fostering collaboration and knowledge sharing among stakeholders, including manufacturers, research institutions, and technology providers, will drive innovation and address common challenges in tool wear detection and classification. Building collaborative platforms and frameworks for sharing datasets, algorithms, and best practices will spur innovation, reproducibility, and standardization in the development of AI-enabled manufacturing process monitoring solutions.

By pursuing these future prospects, the AI-based tool wear detection and classification system can be further refined and adapted to meet the evolving needs of the manufacturing industry. The system's enhanced capabilities and applicability will contribute to the realization of intelligent, data-driven, and sustainable manufacturing practices, aligning with the objectives of Industry 4.0.

5.6 Conclusion

In summary, this thesis provides a comprehensive study on the development and deployment of an AI-based tool wear detection and classification system for optimizing metal cutting processes. The research covers data collection, model development, image processing, generalization, validation, and industrial integration. The proposed system demonstrates high accuracy, reliability, and efficiency in detecting and classifying tool wear, offering significant potential for improving manufacturing productivity and quality.

The main contributions of this thesis include the construction of a diverse and representative tool wear image dataset, the design and implementation of a CNN-based tool wear detection and classification algorithm using the Mask R-CNN network structure, the empirical investigation into the effects of cutting parameters and workpiece material properties on tool wear progression and tool life, a detailed comparison of the proposed system with existing methods, and the validation of the tool wear detection and classification system's applicability in real manufacturing scenarios.

Despite the substantial improvements and practical potential of the AI-powered tool wear detection and classification system, certain limitations and future research areas are acknowledged. These include increasing dataset size and diversity, developing real-time data acquisition and processing capabilities, integrating multi-sensor data fusion, adapting to complex tool geometries and wear patterns, quantifying uncertainty and confidence in the model's detections, exploring advanced transfer learning and domain adaptation methods, integrating with process optimization and control strategies, enhancing the explainability of the deep learning model, investigating the system's long-term performance and reliability, and fostering collaboration and knowledge sharing within the research community.

Addressing these limitations and exploring the identified future research directions will further enhance the capabilities and applicability of the AI-based tool wear detection and classification system. The system's enhanced capabilities and applicability will contribute to the realization of intelligent, data-driven, and sustainable manufacturing practices, aligning with the objectives of Industry 4.0.

The AI-based tool wear detection and classification system has a strong foundation and broad application prospects for future work in manufacturing. Several important directions for future development include integrating advanced sensor technologies, diversifying the AI-based approach to alternative manufacturing processes, leveraging cloud-based platforms for scalable and remote detection and classification, integrating with digital twin technologies and virtual manufacturing, developing autonomous process control and optimization systems, implementing advanced predictive maintenance and tool inventory management strategies, coordinating with supply chain and logistics management systems, implementing dynamic learning mechanisms for continuous improvement, developing standardized data models and interfaces for interoperability, and fostering collaboration and knowledge sharing among stakeholders.

By pursuing these future prospects, the AI-based tool wear detection and classification system can be further refined and adapted to meet the evolving needs of the manufacturing industry. The system's enhanced capabilities and applicability will contribute to the realization of intelligent, data-driven, and sustainable manufacturing practices, aligning with the objectives of Industry 4.0.

In conclusion, this thesis provides a significant step towards intelligent, green, and sustainable manufacturing in the Industry 4.0 era. The proposed AI-based tool wear detection and classification system demonstrates high accuracy, reliability, and efficiency in detecting and classifying tool wear, offering significant potential for improving manufacturing productivity and quality. Future research should focus on expanding the dataset, enhancing real-time capabilities, integrating multi-sensor data, and exploring advanced AI techniques to further improve the system's performance and applicability in diverse manufacturing scenarios. The AI-based tool wear detection and classification system represents a promising solution for optimizing metal cutting processes and advancing sustainable manufacturing practices in the Industry 4.0 era.

References

1. *Alegre, E., Barreiro, J., Castejón, M., & Suarez, S. (2008). Computer vision and classification techniques on the surface finish control in machining processes. Image Analysis and Recognition, 1101-1110. https://doi.org/10.1007/978-3-540-69812-8_108*

2. *Barreiro, J., Castejón, M., Alegre, E., & Hernández, L. K. (2008). Use of descriptors based on moments from digital images for tool wear monitoring. International Journal of Machine Tools and Manufacture, 48(9), 1005-1013. https://doi.org/10.1016/j.ijmachtools.2008.01.004*

3. *Bouzakis, K. D., Michailidis, N., Vidakis, N., Efstathiou, K., Leyendecker, T., Erkens, G., ... & Wirth, I. (2000). Optimization of the cutting edge roundness and its manufacturing procedures of cemented carbide inserts, to improve their milling performance after a PVD coating deposition. Surface and Coatings Technology, 133, 501-507. https://doi.org/10.1016/S0257-8972(00)00967-4*

4. *Devillez, A., Lesko, S., & Mozer, W. (2004). Cutting tool crater wear measurement with white light interferometry. Wear, 256(1-2), 56-65. https://doi.org/10.1016/S0043-1648(03)00384-3*

5. *Essien, A. E., Giannetti, C., & Chin, K. S. (2021). A deep learning model for tool wear monitoring in machining operations using convolutional LSTM network. Applied Soft Computing, 105, 107285. https://doi.org/10.1016/j.asoc.2021.107285*

6. *Huang, H., Zhou, L., Chen, X. Q., & Gong, Z. M. (2003). SMART robotic system for 3D profile turbine vane airfoil repair. The International Journal of Advanced Manufacturing Technology, 21(4), 275-283. https://doi.org/10.1007/s001700300037*

7. *ISO 3685. (1993). Tool-life testing with single-point turning tools. International Organization for Standardization.*

8. *Kurada, S., & Bradley, C. (1997). A review of machine vision sensors for tool condition monitoring. Computers in Industry, 34(1), 55-72. https://doi.org/10.1016/S0166-3615(96)00075-9*

9. *Lanzetta, M. (2001). A new flexible high-resolution vision sensor for tool condition monitoring. Journal of Materials Processing Technology, 119(1-3), 73-82. https://doi.org/10.1016/S0924-0136(01)00878-0*

10. *Loizou, J., Tian, W., Robertson, J., & Camelio, J. (2015). Automated wear characterization for broaching tools based on machine vision systems. Journal of Manufacturing Systems, 37, 558-563. https://doi.org/10.1016/j.jmsy.2015.04.006*

11. *Pfeifer, T., & Wiegers, L. (2000). Reliable tool wear monitoring by optimized image and illumination control in machine vision. Measurement, 28(3), 209-218. https://doi.org/10.1016/S0263-2241(00)00014-2*

12. *Terrazas, G., Martínez-Arellano, G., Benardos, P., & Ratchev, S. (2018). Online tool wear classification during dry machining using real time cutting force measurements and a CNN approach. Journal of Manufacturing and Materials Processing, 2(4), 72. https://doi.org/10.3390/jmmp2040072*

13. *Wang, P., Liu, Z., Gao, R. X., & Guo, Y. (2019). Heterogeneous data-driven hybrid machine learning for tool condition prognosis. CIRP Annals, 68(1), 455-458. https://doi.org/10.1016/j.cirp.2019.04.104*

14. *Xiong, G., Liu, J., & Avila, A. (2011). Cutting tool wear measurement by using active contour model based image processing. In 2011 IEEE International Conference on Mechatronics and Automation (pp. 670-675). IEEE. https://doi.org/10.1109/ICMA.2011.5985761*

15. *Yang, G., Guo, S., Jiang, Z., Duan, F., & Duan, P. (2021). A novel tool wear monitoring method based on multi-scale PCA and LSTM neural network. Measurement, 173, 108672. https://doi.org/10.1016/j.measurement.2021.108672*

16. *Zapico, A., Blanco, D., Diéguez, J. L., & Pereira, O. (2021). Tool wear monitoring in milling processes using a convolutional neural network under different transfer learning and fine-tuning strategies. Sensors, 21(13), 4423. https://doi.org/10.3390/s21134423*

17. *Zhang, C., Yao, X., Zhang, J., & Jin, H. (2016). Tool condition monitoring and remaining useful life prognostic based on a wireless sensor in dry milling operations. Sensors, 16(6), 795. https://doi.org/10.3390/s16060795*

18. *Zhao, R., Yan, R., Wang, J., & Mao, K. (2017). Learning to monitor machine health with convolutional bi-directional LSTM networks. Sensors, 17(2), 273. https://doi.org/10.3390/s17020273*

19. *陈, 希. (2023). Overview of Tool Wear Detection Based on Deep Learning. Operations Research and Fuzziology.*

20. *Mohamed, A., Hassan, M., M'Saoubi, R., & Attia, H. (2022). Tool Condition Monitoring for High-Performance Machining Systems-A Review. Sensors (Basel, Switzerland), 22(6), 2206. https://doi.org/10.3390/s22062206*

21. *Nicolodi, J.H., Consalter, L.A., Durán, O., & Souza, A.J. (2019). Influence of Cutting Speed and Feed Rate on the Wear of an S-Type Mixed Ceramic Tool in Turning of AISI 4140 Hardened Steel. Tehnicki vjesnik - Technical Gazette.*

22. *Zawawi, M.Z., Shaharudin, M.A., & Manaf, A.R. (2014). Influences of Varying Combination of Feed Rate and Depth of Cut to Tool Wear Rate and Surface Roughness: High Speed Machining Technique in Non-High Speed Milling Machine. Advanced Materials Research, 903, 194 - 199.*

23. *Iswanto, Firmansyah, M.F., & Mulyadi (2020). Effect of the cutting angle and the depth of cut toward wear of carbide tool on the lathe. Journal of Physics: Conference Series, 1594.*

24. *Varmma a/l Suparmaniam, M., & Yusoff, A.R. (2016). Investigation of surface roughness and tool wear length with varying combination of depth of cut and feed rate of Aluminium alloy and P20 steel machining. IOP Conference Series: Materials Science and Engineering, 114.*

25. *Ramalingam, S., & Wright, P.K. (1981). Abrasive Wear in Machining: Experiments With Materials of Controlled Microstructure. Journal of Engineering Materials and Technology-transactions of The Asme, 103, 151-156.*

26. *Kim, M., Lee, M., Cho, G., & Lee, S. (2020). Effect of the Fiber Orientation and the Radial Depth of Cut on the Flank Wear in End Milling of CFRP. International Journal of Precision Engineering and Manufacturing, 21, 1187 - 1199.*

Figure Sources:

1. *Alokasi, H.; Ahmad, M.B. (**2022**). Deep Learning-Based Frameworks for Semantic Segmentation of Road Scenes. Electronics* [*https://doi.org/10.3390/electronics11121884*](https://doi.org/10.3390/electronics11121884)

2. *Yang, Jie & Zhao, Junhong & Lu, Lu & Pan, Tingting & Jubair, Sidra. (2020). A New Improved Learning Algorithm for Convolutional Neural Networks. Processes. 8. 295. 10.3390/pr8030295.*

3. *Yu, Po-Nien & Lai, Yung-Chi & Chen, Yi-You & Cheng, Da-Chuan. (2023). Skeleton Segmentation on Bone Scintigraphy for BSI Computation. 10.20944/preprints202306.1576.v1.*

4. ***https://www.mmc-carbide.com/en_jp/technical_information/tec_other_data/tec_other_data_top/tec_other_data_technical/tec_wear_damage***

5. ***https://www.baeldung.com/cs/object-detection-intersection-vs-union***