ASME Journal of Tribology
Online journal at:
https://asmedigitalcollection.asme.org/tribology

Assessment of Tool Condition and Surface Quality Using Hybrid Deep Neural Network: CNN-LSTM-Based Segmentation and Statistical Analysis

K. Venkata Rao
Department of Mechanical Engineering,
Vignan's Foundation for Science,
Technology and Research,
Vadlamudi, Andhra Pradesh 522213, India
e-mail: kvenkatrama@gmail.com

(RMSE,R2)

Length: 0.219, 0.974
Width: 0.0018, 0.943

Effective and precise prediction of tool wear plays a key role in improving machining efficiency, and product quality and reducing production cost. The majority of earlier studies have depended on limited experimental data, which may not be sufficient to estimate tool wear and surface quality. Aiming at these issues, the present study proposed a convolutional neural network (CNN)-long short-term memory (LSTM) hybrid deep neural network model that directly utilizes heterogeneous data including timely captured tool images, working conditions, vibration data, surface roughness, flank wear length, and wear depth. First, experiments were conducted on AISI D2 steel at three levels of spindle speed and feed/tooth, and experimental results for wear length, wear depth, surface roughness, and vibration signals were collected. The time domain vibration signals were processed with a fast Fourier transformer and converted to the frequency domain, and 13 and 5 features were extracted from the time and frequency domain, respectively, and integrated with the heterogeneous data. Second, tool images were annotated using Roboflow software, and wear region information was collected using YOLOv8 and added to heterogeneous data. Third, the CNN-LSTM network was trained with heterogeneous data containing spatial and time-dependent features. The performance and accuracy of the proposed methodology were validated using experimental data collected at different working conditions. The results show that the CNN-LSTM model effectively predicted the tool wear length on the flank, with the root mean square error (RMSE) value of 0.219 mm, and the determination coefficient R^2 value of 0.974; wear depth with the RMSE value of 0.018 mm and R^2 value of 0.943; surface roughness with the RMSE value of 0.216 μm and R^2 value of 0.956. The proposed methodology has significance in metal-cutting applications and provides a solution to predict tool conditions and surface quality accurately.
[DOI: 10.1115/1.4067496]

Keywords: tool condition monitoring, tool vibration, surface roughness, machine learning, YOLO, abrasion, artificial intelligence, vibrations, wear

1 Introduction

Tool wear and surface roughness have a significant effect on the product quality and machining efficiency. In the machining process, 20% of the machining time is interrupted due to tool failure problems that greatly affect the machining efficiency and cost [1]. Continuous assessment of tool condition helps to replace the worn tool at the right time, which further improves the machining efficiency for sustainable manufacturing. It is necessary to develop intelligent manufacturing systems with appropriate tool condition monitoring (TCM) models to increase machining efficiency and economy [2]. Intelligent manufacturing systems are developed using online vibration-based tool condition modeling approaches to solve tool failure problems. However, there is still a need to develop online modeling techniques considering machining parameters along with the vibration data collected during the machining process [3]. In the present study, heterogeneous data including cutting parameters, tool wear features, captured tool image, surface roughness, and tool vibration features were generated to develop a hybrid deep neural network (HDNN) for tool condition and surface quality monitoring system.

Measurement of machining characteristics such as tool wear, surface roughness, cutting forces, tool vibration, and acoustic emissions plays a crucial role in TCM. There are two types of measurement methods such as direct and indirect methods, indirect measurement methods use contact and non-contact sensors to

Contributed by the Tribology Division of ASME for publication in the JOURNAL OF TRIBOLOGY. Manuscript received November 12, 2024; final manuscript received December 14, 2024; published online January 3, 2025. Assoc. Editor: Andreas Rosenkranz.

measure cutting forces, tool vibration, and noise emissions, which are used to detect tool conditions [4]. Since indirect and non-contact online measurement systems do not require machine downtime, it does not interrupt machining time. Prasad et al. [5] measured workpiece vibration using a non-contact-type laser Doppler vibrometer (LDV) and monitored tool condition based on the vibration signals in the face turning of AISI 1040 steel. Rao et al. [6] used LDV in the milling of Ti–6Al–4V alloy and monitored the tool wear condition and surface roughness using grey prediction model GM(1,N) and support vector machine. In the present study, LDV was used to measure the vibration of the cutter based on the advantage of the LDV over the contact-type sensors. Zhang et al. [7] studied the effect of flank wear on the tool vibration and surface roughness in the machining of NAK80 and reported that the flank wear caused tool vibration and roughness on the machined surface. However, it is required to develop a TCM system using deep machine learning techniques for accurate prediction of tool wear conditions.

In TCM, wear length and wear depth on the flank, crater wear, cutting edge wear, nose wear, and surface texture on the tool face and flank are directly quantified using an optical microscope, Talysurf surface profilometer, scanning electron microscope, and tool image capturing CMOS cameras [8]. In the direct measurement of tool wear, it is difficult to identify the complete wear region since the worn region appears in different colors. Non-contact-type image processing techniques perform feature extraction, segmentation, and morphological processing. Since image processing techniques have high computational efficiency and interpretability, it is widely used in TCM [9]. Recently, researchers have used image processing techniques in TCM and extracted the worn segment along with labeling from the captured tool image. Malhotra and Jha [10] developed a fuzzy c-means clustering-based wear monitoring system, extracted the region of interest from the captured image, and estimated the wear width using pixels in the worn segment with less than 5% error. Zhu et al. [9] captured tool images using a camera and identified the worn segment on the tool flank using the DC-UNet3+ segmentation model. Sun et al. [11] measured tool vibration, cutting forces, and flank wear using direct measurement systems and predicted tool conditions using a Semi-supervised TCM model. Manwar et al. [12] developed an image processing algorithm for a high-precision TCM system that aligns all the images with an initial image to capture the right image without error, converts the 2D image into 1D form, and predicts the tool wear. Makhesana et al. [13] developed machine vision-based TCM algorithms that extracted wear features from tool images and measured tool wear to replace the cutter at the right time.

In the above image processing techniques, the captured tool image was segmented into three regions such as background of tool, wear, and unworn regions. Different types of image segmentation methods and algorithms such as region-dependent segmentation, cluster-dependent segmentation, and threshold-dependent segmentation are used in image processing [14]. However, the aforementioned methods are occasionally very complex and call for a variety of tools, algorithms, and approaches. Consequently, there is a need to develop streamlined and efficient image processing methods that can immediately detect and quantify wear components. You Only Look Once (YOLO) was introduced in 2016 for object identification and image segmentation with high accuracy at impressive speed. Since then, several versions of the YOLO have been developed and used in different applications. Asadi et al. [15] developed convolutional neural network (CNN)based YOLOv5 and YOLOv8 segmentation models and monitored metal deposition in metal-based additive manufacturing using melt pool segmentation. Yu et al. [16] developed a maintaining the original dimension-YOLO algorithm and identified cracks in the civil infrastructure and improved the accuracy by 27.5% in detection. Wang et al. [17] developed a framework using YOLOv3 and detected welding slag in repairs of automotive bodies by machine vision and machine learning. Based on the high computational abilities in object identification and segmentation, the latest version YOLOv8 was used in the present study.

In metal cutting, traditional mathematical models are developed and predict machining characteristics such as surface roughness, tool wear, cutting forces, metal removal rate, and power consumption. However, various tools and machining parameters have nonlinear trends with machining characteristics and are also dynamically unstable. It is challenging for researchers to implement mathematical models in the field of machining and also difficult to handle complex problems. Traditional approaches to predicting tool life often fail due to uncertainty and device complexity, but machine learning techniques like artificial neural networks offer great potential for accurately predicting tool wear [18]. Agarwal et al. [19] developed a theoretical model taking the cutting speed and feed rate as input to the model in micro drilling of Inconel 718, the results indicating that the tool wear is caused by corner chipping and built-up edges. They found a good correlation between the model-predicted tool life and the experimental results. Researchers recognized the effectiveness and high computational capacity of the CNN, long short-term memory (LSTM), and gated recurrent unit (GRU) in handling the time series data and long-term dependent variables. CNNs are widely used in image processing and signal processing applications, they perform different tasks like classification, regression, and segmentation effectively [20].

Since the signal data like force, acoustics emission, and vibration signal are time-dependent, LSTM is added to CNN and recurrent neural networks [21]. LSTM is an alternative RNN network architecture, and it is particularly useful for solving time series problems because of its unique advantages for long-term knowledge retention [22]. Manwar et al. [12] developed an LSTM-based online TCM system using cutting force and tool vibration signals and predicted the tool condition in micro-milling. Zhou et al. [22] developed an LSTM-based prediction method and estimated the remaining useful life of the cutter at different working conditions in the milling of QT500-7GB/T1348-1988. Yang et al. [23] measured cutting force signals in X and Y directions, converted them into time series force maps, and trained an LSTM with the force maps to predict tool life and confirmed that the model accurately predicted tool life under different working conditions in milling of AISI 1045 metal. Compared to the LSTM method, CNN is a superior network in extracting the features from the image with low computational efforts. Lim et al. [24] developed a CNN-based tool wear prediction model using images of machined surfaces and predicted the surface roughness with an accuracy of 98.9%. It is well known that the tool is sharp at the beginning of machining, and the tool wear is fast in the initial stage, uniform in the middle stage, and faster in the final stage. However, tool wear is a gradual, stochastic, nonlinear process that heavily depends on timing parameters [25]. Therefore, HDNN networks like a combination of CNN and LSTM have become a new trend.

Recently, hybrid deep learning techniques such as a combination of CNN and LSTM and other networks have been developed in the manufacturing field to process spatial and temporal features and predict tool wear, remaining tool life, and surface roughness. Zhu et al. [9] developed a hybrid CNN-LSTM model and predicted surface roughness in the turning of $Cr_{12}MoV$ metal. They identified wear segments on the tool flank using the DC-UNet3+ technique, trained the prosed model, and predicted surface roughness. Marei and Li [26] developed a hybrid CNN-LSTM model and predicted the remaining useful life of the cutting tool in the machining of $Cr_{36}NiMo_4$. Wang and Shen [27] developed a flank wear prediction model called an auxiliary input-enhanced Siamese neural network and predicted flank wear at different working conditions in the milling of Hastelloy-X. In their proposed model, they gave an auxiliary input of cutting time to the fully connected layer and improved the prediction accuracy and precision. Peng and Li [28] developed a new wear monitoring framework called a temporal convolutional network-LSTM model using spindle current and cutting forces obtained at different working conditions and predicted tool conditions effectively. Qin et al. [29] developed a data-

driven model using tool wear image, vibration, and force signals measured in x, y, and z directions and accurately monitored tool wear conditions at different working conditions in the milling of TC4 titanium alloy. Li et al. [30] developed a one-dimensional CNN-LSTM model and predicted tool wear in the milling of carbon fiber-reinforced polymer at different working conditions, leading to a 25% improvement in prediction accuracy.

The literature reveals that there is still a knowledge gap in monitoring tool wear and its geometrical features in the machining process. Previous studies depended on inputting the image data into prediction models with limited specifications of the worn image. It can be observed that limited research was done on feature extraction, identification, and segmentation methods for flank wear. Furthermore, the image processing is alone not sufficient to estimate tool wear and surface roughness. To overcome these difficulties, vibration data was taken along with wear images in the present study to predict the wear length, wear depth, and surface roughness in the milling of AISI D2 steel. In the present study, an HDNN model was developed by combining CNN and LSTM to handle image and time-sequential data. This study contributes by (i) conducting experiments on AISI D2 steel as per L27 Taguchi's design of experiments and data acquisition, (ii) annotating the image using Roboflow software and collection of data set using YOLOv8s, (iii) data generation, developing an HDNN (CNN-LSTM) network and training the network using the image and time-sequential data, (iv) discussion of CNN-LSTM performance and its results, and (v) performing statistical analyses on wear length, wear depth and surface roughness.

2 Experimentation

The proposed methodology was experimentally implemented in the milling of AISI D2 steel using a cemented carbide mill cutter having a 10-mm diameter. Chemical elements with their percentages in the AISI D2 steel and its mechanical properties are presented in Tables 1 and 2, respectively. Tool geometry is presented in Table 3.

The experimental setup for the milling of AISI D2 steel is presented in Fig. 1. The workpiece was taken with a size of 150 × 75 × 15 mm and placed on the machine table and the experiments were conducted according to the L27 design of experiments. As presented in Table 4, 9 trials were conducted and 3 experiments were performed in each trial. Each trail was started with a fresh tool and machining was done at a length of 150 mm in each experiment. After 150-mm length of machining, the workpiece, and cutter were removed and surface roughness on the machined surface using the Mitutoyo 178-923E SJ210 type profilometer. Wear length and wear depth on the flank were measured using Metavis make prime 100 model stereo zoom microscope and also captured flank image. A Polytech 100-V-type LDV was placed in front of the machining center, and the tool vibration was measured in the time domain and processed using a fast Fourier transformer as shown in Figs. 2 and 3. The same cutter was used in the remaining two experiments of Trail 1 and surface roughness, wear length, and wear depth were measured. The same experimental procedure was followed in the remaining eight trials, and the experimental results for the wear length, wear depth, and surface roughness were collected and presented in Table 4.

In vibration-based TCM, the measurement of vibration plays a key role as the amplitude and frequency of vibration is a reflection of surface roughness and tool wear, according to ISO 10816, contact and non-contact-type sensors have been used in the measurement of tool vibration, and workpiece vibration such as accelerometer and LDVs, respectively [31]. Both of these devices are popular for measuring vibration, but accelerometers cannot measure the vibration of a rotating body as they are contact types. Non-contact-type LDVs are very popular in these situations because they can measure the vibration of any component on a rotating body. As shown in Fig. 1, the vibration is measured by focusing a laser beam on the targets. Based on the Doppler shift principle, the LDVs extract the cutter's vibration frequency and amplitude from the reflected laser beam frequency. During the measurement of vibration, a 100–1000-kHz bandpass filter was taken and filtered the vibration signals collected from the cutter. For signal processing, blocks of 6000 data points acquired at a sampling period of 1 ms with a sampling frequency of 4 MHz were used.

The vibration data collected from LDV was filtered by noise reduction using the Gaussian smoothing technique also known as Gaussian filtering. As shown in Figs. 2–4, the Gaussian smoothed vibration signals describe the reduced high-frequency noise in the vibration signal. The Gaussian smoothing technique performs two functions such as kernels or filter creation and convolution. The

Table 2 Properties of AISI D2 steel

Parameter	Hardness Rockwell C	Density	Modulus of elasticity	Poisson's ratio
Value	62	7696 kg/m^3	207 GPa	0.3

Table 3 Tool geometry

Parameter	Core diameter	Helix angle	Edge radius	Cutting edge radius	Radial rake angle	Radial relief angle
Value	6 mm	20 deg	1 mm	0.04 mm	5 deg	30 deg

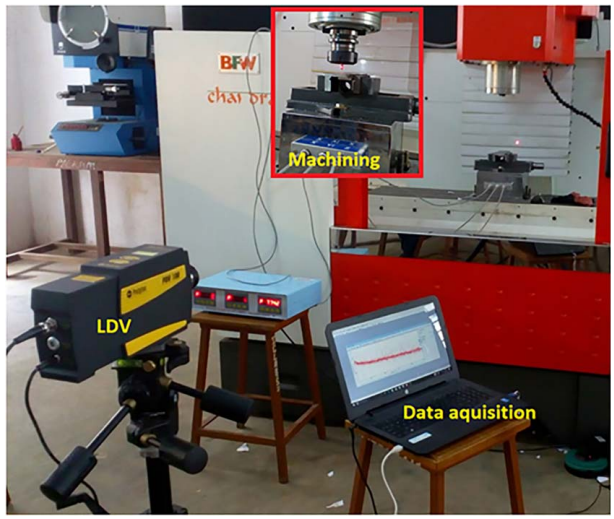

Fig. 1 Experimental setup for milling

Table 1 Chemical composition of AISI D2 steel

Element	Mn	Si	V	Cr	P	C	Mo	Co	S	Ni	Cu	Fe
wt %	0.6	0.6	1.10	12.25	0.03	1.5	0.95	1.0	0.03	0.3	0.25	81.39

Table 4 Experimental plan and results of experiments and prediction

Trail No.	Exp. No.	Spindle speed (rpm)	Feed (μm/tooth)	Machining length (mm)	Wear length (mm)		Wear depth (mm)		Surface roughness (μm)	
					Exp.	Pred.	Exp.	Pred.	Exp.	Pred.
1	1	5000	5	150	1.299	1.392	0.129	0.108	2.57	2.43
	2	5000	5	300	1.445	1.568	0.132	0.119	2.79	2.56
	3	5000	5	450	1.652	1.671	0.151	0.133	2.82	2.68
2	4	5000	7	150	1.801	1.714	0.166	0.144	3.17	3.25
	5	5000	7	300	2.178	1.945	0.177	0.16	3.38	3.47
	6	5000	7	450	2.249	2.162	0.185	0.165	3.73	3.59
3	7	5000	9	150	2.461	2.371	0.189	0.201	4.01	3.81
	8	5000	9	300	2.773	2.65	0.194	0.214	4.23	4.31
	9	5000	9	450	2.873	2.945	0.218	0.229	4.51	4.46
4	10	5500	5	150	2.287	2.469	0.164	0.151	2.31	2.18
	11	5500	5	300	2.563	2.673	0.191	0.175	2.37	2.19
	12	5500	5	450	2.89	3.047	0.187	0.215	2.46	2.21
5	13	5500	7	150	3.072	3.126	0.21	0.229	2.70	2.48
	14	5500	7	300	3.095	3.175	0.215	0.237	2.93	2.67
	15	5500	7	450	3.221	3.382	0.245	0.26	3.10	3.04
6	16	5500	9	150	3.362	3.471	0.272	0.251	3.25	3.04
	17	5500	9	300	3.405	3.378	0.276	0.257	3.26	3.51
	18	5500	9	450	3.44	3.326	0.302	0.291	3.35	3.97
7	19	6000	5	150	2.873	2.981	0.208	0.229	1.36	1.42
	20	6000	5	300	3.11	3.217	0.221	0.246	1.64	1.53
	21	6000	5	450	3.402	3.308	0.228	0.247	1.71	1.54
8	22	6000	7	150	3.584	3.492	0.247	0.246	1.80	2.05
	23	6000	7	300	3.62	3.614	0.248	0.257	2.14	2.38
	24	6000	7	450	3.892	3.925	0.263	0.275	2.25	2.13
9	25	6000	9	150	4.271	4.483	0.271	0.284	2.43	2.17
	26	6000	9	300	4.579	4.57	0.281	0.288	2.57	2.43
	27	6000	9	450	4.899	4.762	0.318	0.345	2.62	2.98

kernel is a matrix that was slid over the vibration data in the convolution step and reduced the noise in the vibration signals. The filtered signals were further processed using a fast Fourier transformer and converted the time domain signals into frequency domain signals. Figures 2–4 illustrate variations in the amplitude and frequency of cutter vibration at different working conditions. According to Figs. 3 and 4, the amplitude of cutter vibration increased at 9 μm/tooth feed and 450-mm machining length when the spindle speed increased to 5500 rpm and 6000 rpm, respectively.

From the time domain of vibration signals, 13 features such as standard deviation, minimum value, maximum value, average value, amplitude factor, peak-to-peak value, skewness, root mean square, kurtosis, Energy, impact factor, waveform factor, and margin factor, and 5 features from the frequency domain such as frequency domain amplitude average, variance frequency, frequency variance, mean square frequency and center of gravity frequency extracted as presented in Sec. 4.1. These features were added to working condition features and machining performance characteristics such as wear length, wear depth, and surface roughness to generate heterogeneous data to train the proposed CNN-LSTM model. The experimental plan and experimental results of wear length, wear depth, and surface roughness are presented in Table 4.

3 Image Processing and Labeling

In deep learning-based approaches, image processing, collection of data, preparation of data to train the model, and training process

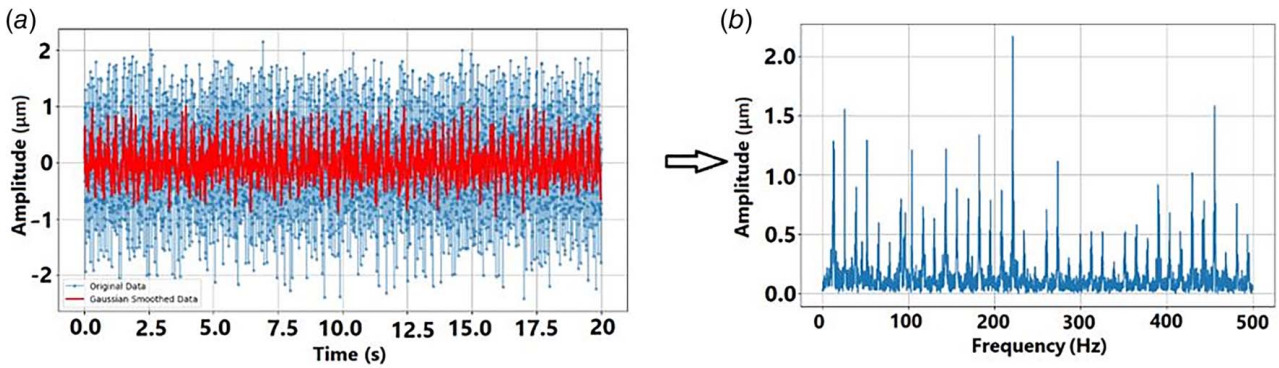

Fig. 2 Time and frequency domains of the tool vibration in the experiment conducted at 5000 rpm of spindle speed, 5 μm/tooth of feed, and 150 mm of machining length

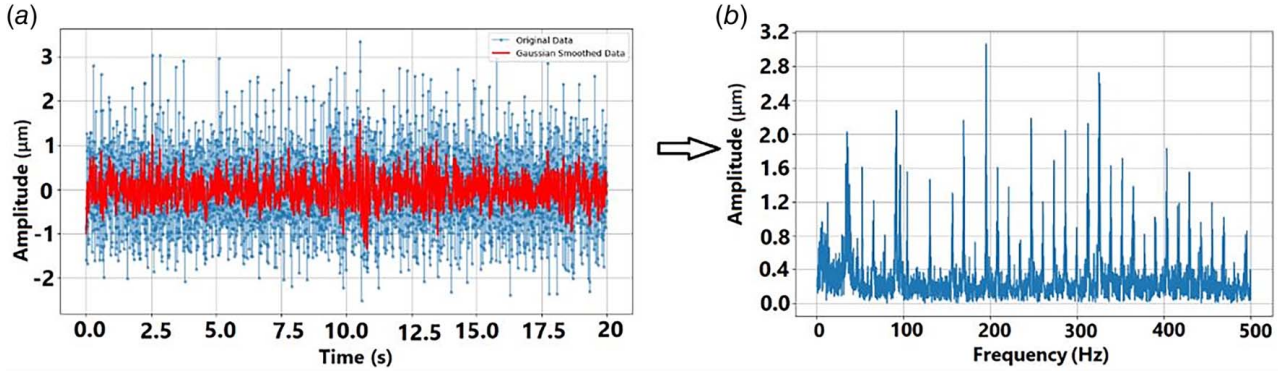

Fig. 3 Time and frequency domains of the tool vibration in the experiment conducted at 5500 rpm of spindle speed, 9 μm/tooth of feed, and 450 mm of machining length

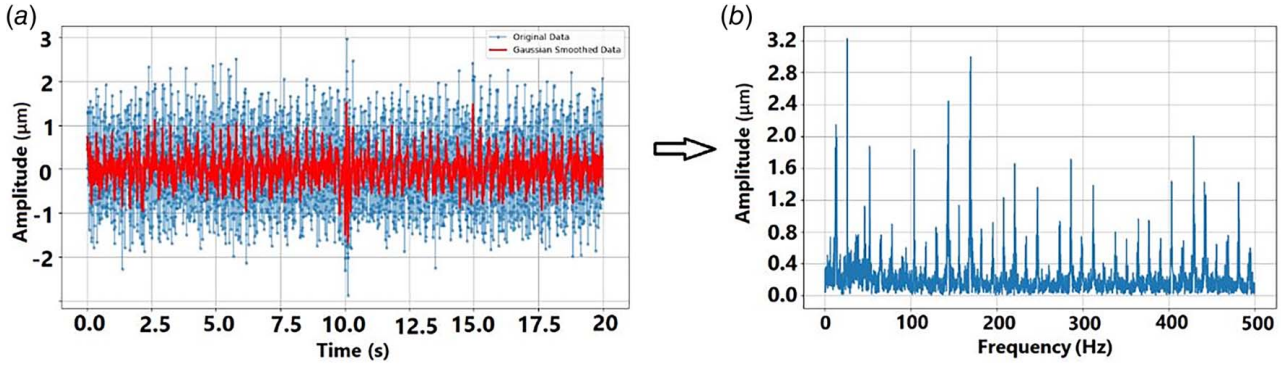

Fig. 4 Time and frequency domains of the tool vibration in the experiment conducted at 6000 rpm of spindle speed, 9 μm/tooth of feed, and 450 mm of machining length

are the key components. Among them, annotation of the image is a crucial component in image processing applications that segment the required region from the captured image [9,15]. Asadi et al. [15] captured images of the melt pool using a Cavitar camera in metal-based additive manufacturing and segmented the melt pool using YOLOv5 and YOLOv8 image processing approaches. Zhu et al. [9] have developed a segmentation model using DC-UNet3 + and identified the wear segment on the tool flank in turning of Cr12MoV. In their work, they have taken images of tool wear at intermediate stages using a camera. The processed images and the surface roughness data were collected, trained the Hybrid CNN-LSTM model, and used the tool wear segment. In the present study, a simple method was proposed to capture and process the image of wear on the flank. As presented in Table 4, 27 experiments were conducted, and a tool image was captured in each experiment using Metavis make prime 100 model stereo zoom microscope. Tool wear annotation and collection of wear data are presented in Fig. 5.

Captured images (27 images) were loaded into Roboflow online software for image processing. The 27 images of the flank wear were annotated to label the wear length and wear depth by drawing a polygon around the tool wear as shown in Fig. 6. The source images were classified into wear classes, and the data set was split into 70%, 20%, and 10% for training, validation, and test, respectively. In pre-processing of images, the annotated images were enhanced using augmented tools such as resize, rotation, brightness, color adjustments, crop, shear, exposure, etc.

In the next stage, the dataset was exported to the YOLOv8s training framework. The Roboflow generates an application programming interface (API) key to export the dataset to YOLOv8s. The YOLOv8s notebook (colab) was opened using the copy 'snippet

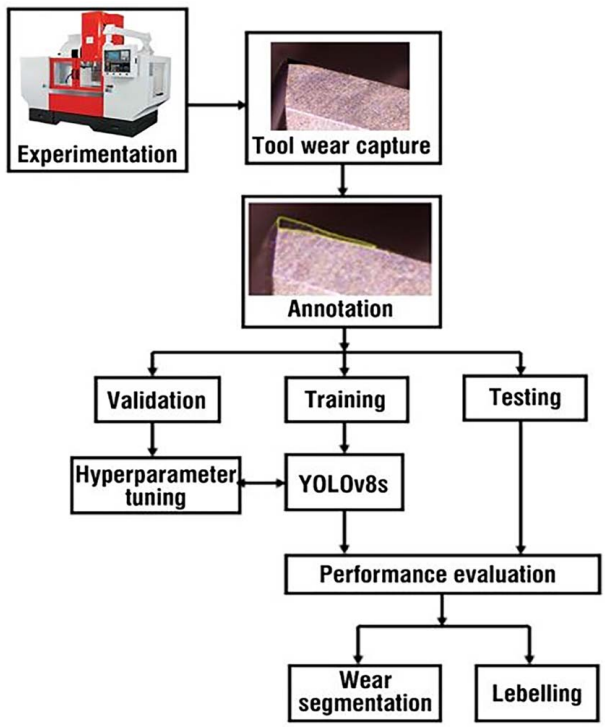

Fig. 5 Tool wear annotation and collection of wear data

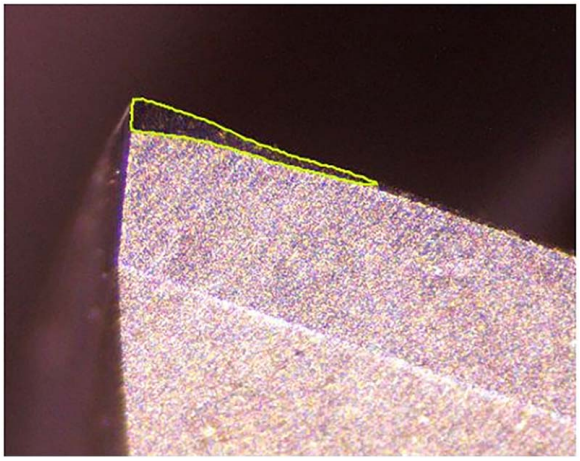

Fig. 6 Annotation of tool wear using Roboflow

and open YOLOv8s notebook' option, and the code generated by Roboflow was pasted into the YOLOv8 notebook.

YOLOv8s "You Only Look Once" version 8 is an expanded version of YOLOv5 developed by Ultralytics that can be used for object identification and image segmentation with high precision at impressive speed. In 2016, the first version of the YOLO was implemented [32], and then, several versions were developed considering more parameters to achieve accuracy in segmentation, classification, and identification. There are five variants in the YOLOv8s such as n (smallest model), s (small model), m (medium model), l (large model), and xl (extra-large model) based on the values of depth_multiple, width_multiple, and max_channels. In the present study, YOLOv8s was used for the wear segment, and the depth_multiple, width_multiple, and max_channels were taken as 0.33, 0.50, and 1024, respectively. As shown in Fig. 7, the YOLOv8s has three sections called backbone, neck and sections and five blocks such as convolutional, bottleneck, C2f, spatial pyramid pooling fast (SPPF), and detect blocks [15]. The convolutional block applies filters to extract features like textures, edges, and other patterns. The convolution black consists convolution layer that performs a mathematical operation on the input image with Keras or filters $(k) = 3$, stride $(s) = 2$, padding $(p) = 0$, bacthnorm2d layer that enhances the convergence speed and stability in the training of the model and SiLU (Sigmoid Linear Unit) activation function that allows smooth gradients in training of the model. The bottleneck block consists of a series of convolutions that reduce the dimensionality of the image, process the image, and restore it to its original size. C2f is a cross-stage partial fusion process consisting of a convolutional block that improves the network ability and integrates features obtained at different stages. SPPF block consists convolutional block and three maxpool layers that improve feature extraction. Detect block detects objects including class and segment. The size of the input image is defined by $H \times W \times C\_in$, where H and W are several pixels and the C_in is several input channels. The number of output channels is estimated using Eq. (1) as follows:[1]

$$C\_out = \min(64, mc)w \tag{1}$$

where 64, mc, and w are the base output channels, max_channel, and width_multiple, respectively. A model calculation for $C1\_out$ is given below:

$C1\_out = \min(64, 1024)\ 0.5 = 64 \times 0.5 = 32$ channels

The parameter n in the C2f block is estimated by Eq. (2)

$$n = 3 \times d \tag{2}$$

where d is the depth _multiple. As shown in Fig. 7, the size and number of channels of the output image were shown for all the blocks. The neck section upsamples (the 10 and 13 blocks in the YOLOv8s architecture in Fig. 7) the feature maps twice without changing the number of channels and combines the features extracted from the backbone. Concatenate block adds the information received from different blocks.

The dataset regenerated in Roboflow was exported to the YOLOv8s training framework using the API key. The dataset consists of annotated images and corresponding labels (x and y coordinates of the bounding box and polygon points constructed around the wear zone). YOLOv8s is an instance segmentation model that can identify and segment the wear zone within the input image. During training, the YOLOv8s-seg.pt model is selected and trained with a dataset of input image size $640 \times 640 \times 3$. The YOLOv8s has 261 layers of convolutional, pooling, and activation operations performed on the input image and identified the segment of wear zone from the image. The model was trained with an auto optimizer called AdamW optimizer that optimized hyperparameters such as 0.02 learning rate and 0.9 momentum, 35 batch size, 0.0005 wt decay, and training was conducted over 200 epochs. The images used for training the model are presented in Fig. 8.

3.1 Performance of the Model. The performance of the model is estimated by finding the train loss, segmentation loss, and mAP50-95. As shown in Fig. 9, the train segmentation and class losses were estimated to be 0.015 and 0.018, respectively and this indicates that the training is done well. The mean average performance (mAP) is a key evaluation matrix used in image instance segmentation algorithms. The mAP is estimated by plotting a curve using the basic metrics of precision versus recall [15]. The precision and the recall are separately calculated for the tool wear using Eqs. (3) and (4), respectively. The estimated precision represents the accuracy of the model that correctly predicted the targets, and the estimated recall represents the model's ability to detect relevant results [33].

$$\text{Precision} = \frac{\text{True Positives}}{\text{True positives} + \text{False Positives}} \tag{3}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True positives} + \text{False Negetives}} \tag{4}$$

In the estimation of mAP, average values of the precision were taken over different levels of recall as shown in Fig. 10, and each average value is associated with an Intersection over Union (IoU) threshold. The IoU estimates the overlap of the predicted mask and true segmentation mask, with perfect overlap represented by the value of 1. The threshold value is often set to 0.5 or above, and the mAP above the threshold value indicates prediction is positive. The mAP is estimated using Eqs. (5) and (6) [34].

$$mAP = \frac{\sum_{i=1}^{N} AP_1}{N} \tag{5}$$

$$AP = \int_0^1 P(R)dR \tag{6}$$

where AP, P, R, and N are the average performance, precision, recall, and number of classes, respectively. Figure 10 shows the precision–recall curve along with mAP values of the YOLOv8s of the wear segment with a value of 0.913 that indicates the proposed Robolfow-YOLOv8s model exhibits superior performance in prediction of the wear segment as well as labeling of the segment. mAP@50-95 is the estimated average accuracy of all detected IOU values between 0.5 and 0.95. In the present study, the value of mAP@0.5:0.95 was found to be 0.61 as shown in Fig. 10(*c*), and this indicates that the predicted object based on the threshold value (0.5) is real.

[1] https://abintimilsina.medium.com/yolov8-architecture-explained-a5e90a560ce5 viewed on July 21, 2024.

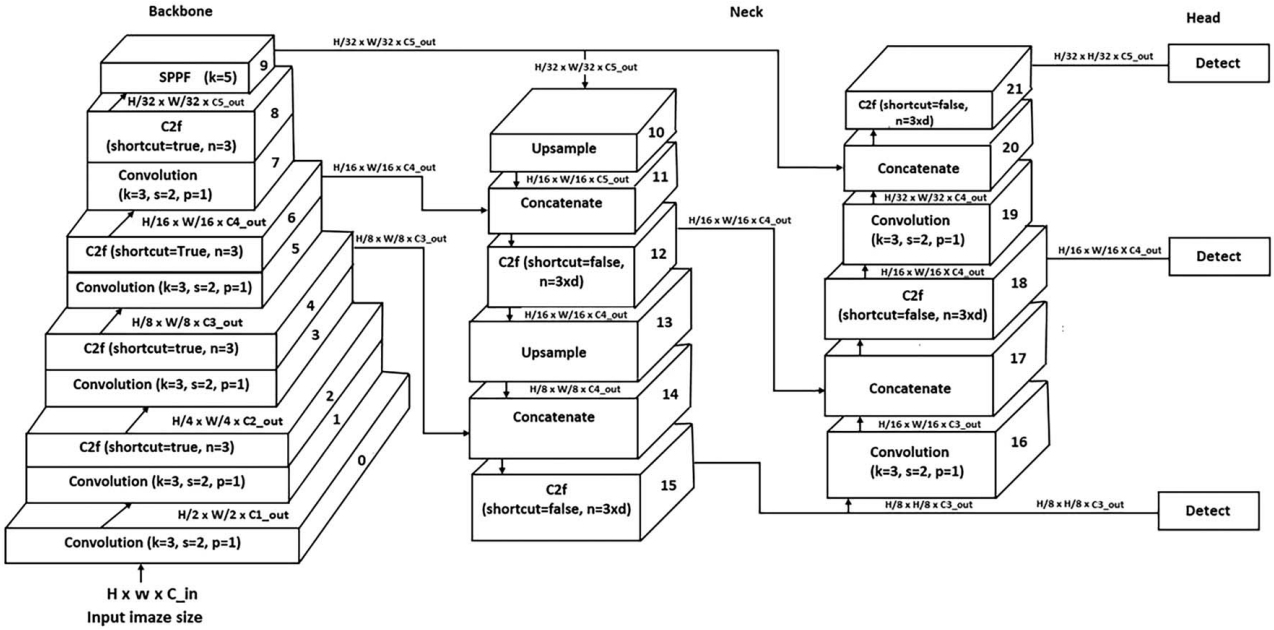

Fig. 7 YOLOv8s architecture [37]

The trained model was tested with three samples (10% of the image samples) which were split from the 27 image samples. The model predicted the wear segment along with labeling for the three image samples with confidence levels of 0.92, 0.87, and 0.94, respectively, and this indicates that the segmentation of the wear was done accurately for the 27 input images. One of the tested images is depicted in Fig. 11, showing the predicted wear segment at a 0.92 confidence level.

The size of the wear segment was defined in the study as the depth of wear and length of the wear segment. As shown in Fig. 12, the depth of the wear segment is the distance between the point (x_1, y_1) and (x_2, y_2) and the length of the wear segment is the distance between the points (x_1, y_1) and (x_3, y_3). In the present study, the images with a size of 40×30 mm were taken for processing in the Roboflow. During the processing, the images were stretched to 640×640 pixels. The wear length and the depth were estimated as 22.96 and 2.16 pixels, respectively, and also estimated in real-world units as 1.312 and 0.108 mm, respectively. Similarly, the wear length and wear depth were estimated for the 27 images. This wear data along with annotated wear images were used to train the CNN-LSTM model.

4 Data Acquisition and CNN-LSTM Architecture

In this study, 25 features of heterogeneous data were extracted from the experiments and trained the proposed CNN-LSTM model to predict wear length, wear depth and surface roughness.

4.1 Data Acquisition. As per the experimental plan, experiments were conducted and the wear data, vibration data, and surface roughness data were collected as presented in the experimentation section. The heterogeneous data are a collection of independent and dependent variables features like working condition features (three such as spindle speed, feed/tooth, and length of cutting), wear data features (three such as wear length and wear depth plus an image of tool wear obtained from YOLOv8s), surface roughness, time domain features (13 such as standard deviation, minimum value, maximum value, average value, amplitude factor, peak-to-peak value, skewness, root mean square, kurtosis, Energy, impact factor, waveform factor, and margin factor), frequency domain features (five such as frequency domain amplitude average, variance frequency, frequency variance, mean square frequency, and center of gravity frequency) as shown in Fig. 13. Before extracting the features from the vibration signals, the noise was removed from the vibration signals using Gaussian smoothing operator. Time domain features were extracted from the smoothened vibration signals using Eqs. (7)–(19).

$$\text{Standard deviation} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{7}$$

$$\text{Minimum value} = \min(y_i,\ y_i,\ \cdots,\ y_n) \tag{8}$$

$$\text{Maximum value} = \max(y_i,\ y_i,\ \cdots,\ y_n) \tag{9}$$

$$\text{Average value} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{10}$$

$$\text{Amplitude factor} = \frac{y_{\max} - y_{\min}}{2} \tag{11}$$

$$\text{Peak} - \text{to} - \text{peak value} = y_{\max} - y_{\min} \tag{12}$$

$$\text{Skewness} = \frac{1}{n}\sum_{i=1}^{n}\frac{(|y_i| - \bar{y})^3}{\text{Standard deviation}^3} \tag{13}$$

$$\text{Root mean square} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i)^2} \tag{14}$$

$$\text{Kurtosis} = \frac{1}{n}\sum_{i=1}^{n}\frac{(|y_i| - \bar{y})^4}{\text{Standard deviation}^4} \tag{15}$$

$$\text{Energy} = \frac{1}{n}\sum_{i=1}^{n}(y_i)^2 \tag{16}$$

$$\text{Impact factor} = \frac{\text{Root mean square}}{\text{Average value}} \tag{17}$$

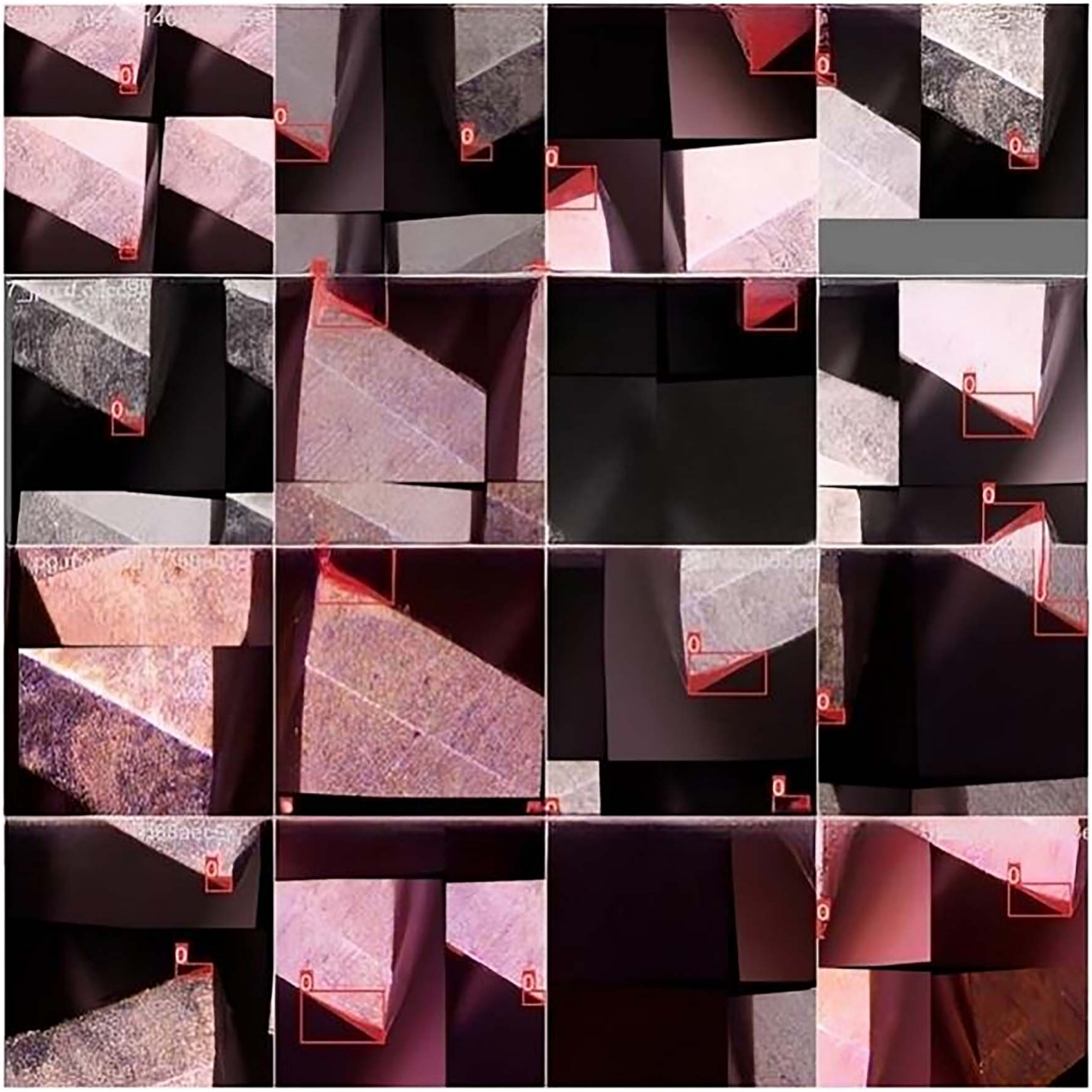

Fig. 8 Training samples in YOLOv8s

$$\text{Waveform factor} = \frac{\text{Root mean square}}{\frac{1}{n}\sum_{i=1}^{n}|y_i|} \tag{18}$$

$$\text{Margin factor} = \frac{\text{Root mean square}}{\text{Peak} - \text{to} - \text{peak value}} \tag{19}$$

where y_i, $\bar{y}$, and n are then individual amplitude data points, average of the data points, and number of vibration data points, respectively. Five features were collected from the frequency domain using Eqs. (20)–(24)

$$\text{Amplitude average} = \frac{1}{N}\sum_{i=1}^{N} M_i \tag{20}$$

$$\text{Frequency variance} = \frac{\sum_{i=1}^{N}(f_i - \text{center of gravity frequency})^2 M_i^2}{\sum_{i=1}^{N} M_i^2} \tag{21}$$

$$\text{Variance frequency} = \sqrt{\text{Mean square frequency}} \tag{22}$$

$$\text{Mean square frequency} = \frac{\sum_{i=1}^{N} f_i^2 M_i^2}{\sum_{i=1}^{N} M_i^2} \tag{23}$$

$$\text{Center of gravity frequency} = \frac{\sum_{i=1}^{N} f_i M_i^2}{\sum_{i=1}^{N} M_i^2} \tag{24}$$

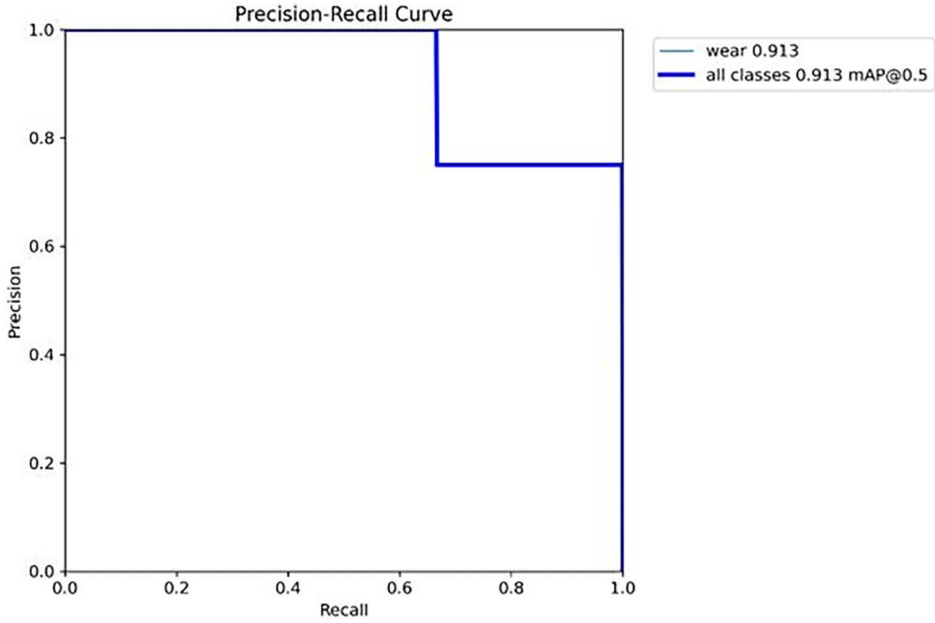

Fig. 9 Precision–recall curve shows mAP

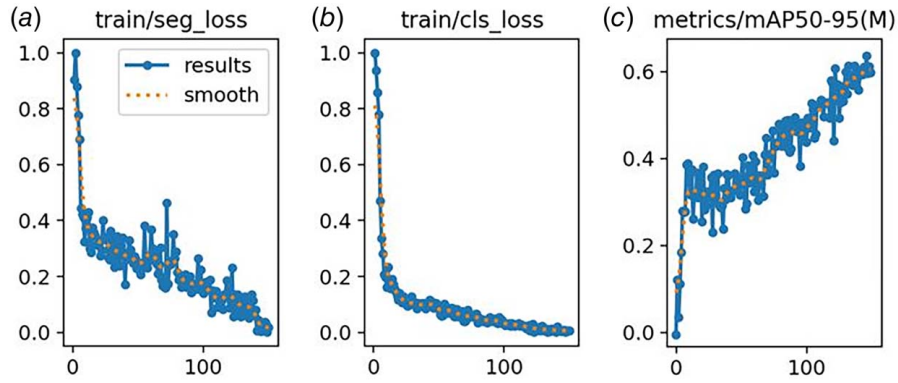

Fig. 10 (*a*) Train/seg loss, (*b*) train/class loss, and (c) mAP50-95

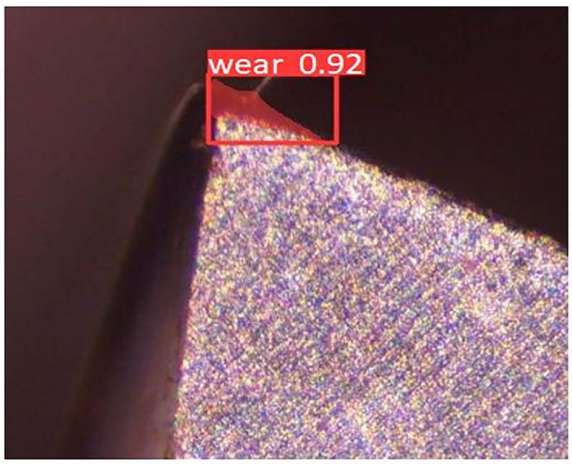

Fig. 11 Predicted wear length and wear depth with confidence level by YOLOv8s

where N, f_i, and M_i are number of frequency bins, frequency bins, and magnitude of frequency bin, respectively.

4.2 CNN-LSTM Architecture. The hybrid deep neural network is an inventive class of deep learning models that combines different machine learning architectures to improve the accuracy of prediction. In the present study, a CNN-LSTM hybrid deep neural network was developed to predict wear length (WL), wear depth (VB), and surface roughness (Ra) in the milling of AISI D2 steel. Figure 13 shows the integrated architecture of the developed HDNN model linked to assess and predict the tool condition and surface quality to replace the worn tool at the right time. Before

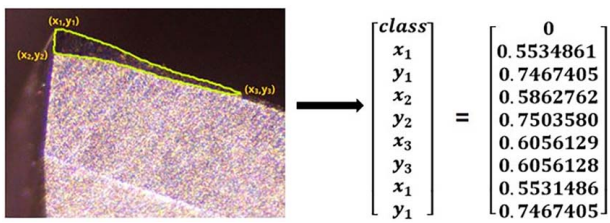

Fig. 12 Wear segment labeling in Roboflow-YOLOv8s

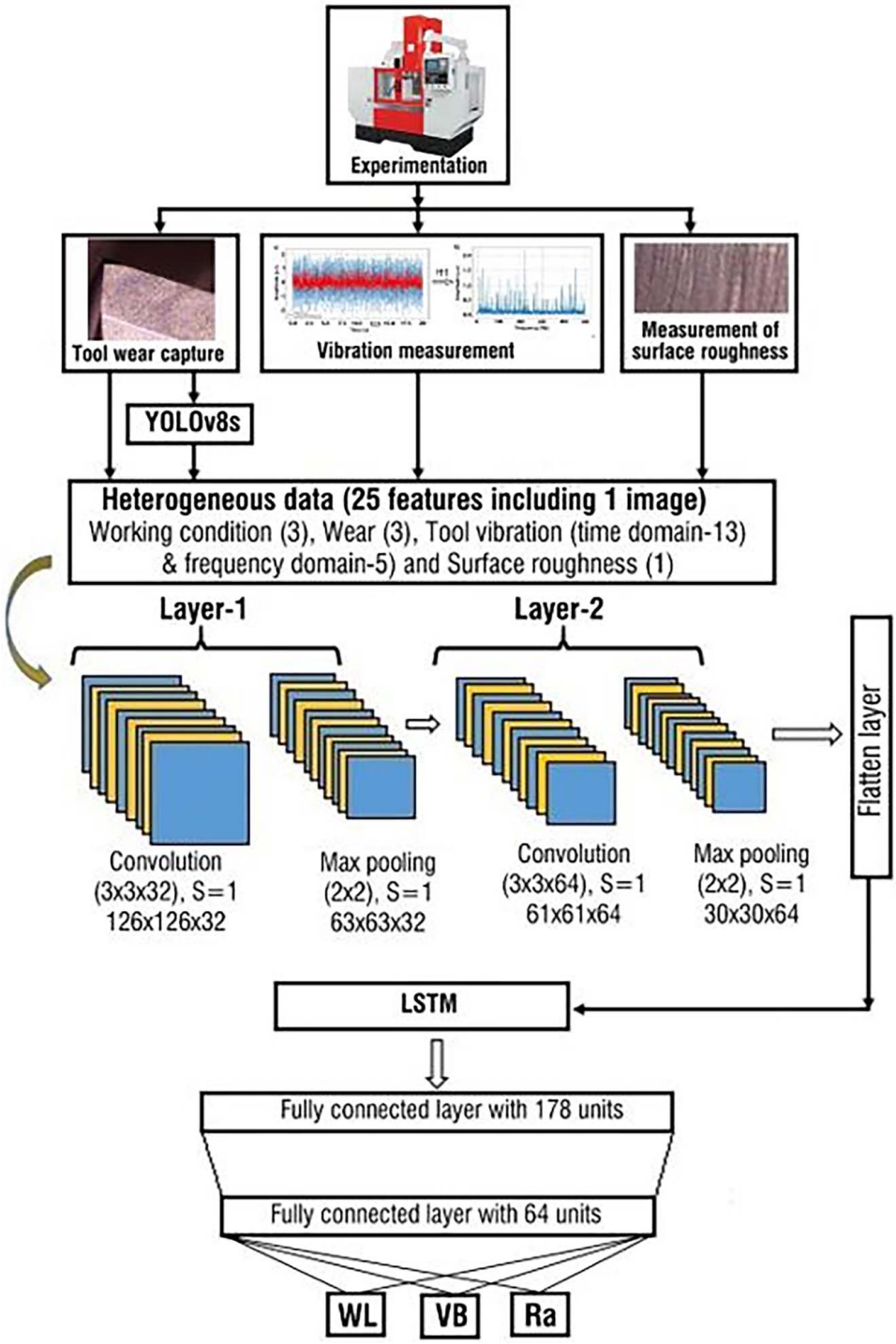

Fig. 13 Data acquisition and CNN-LSTM architecture

implementing the HDNN, the model was trained with the 25 features of heterogeneous data collected during the experimentation and YOLOv8s. Among the 25 features, wear length, wear depth, and surface roughness were dependent features, and the remaining are independent features. During the implementation, the captured tool image along with the remaining dependent features were fed to the HDNN and predicted the weal length, wear depth, and surface roughness. The condition of the tool is judged by the surface roughness predicted by the HDNN. Uniqueness of the developed HDNN in the present study is that the model is able to predict the wear length and wear depth using the raw image of the tool captured during the machining process. In these hybrid models, the CNN and LSTM models handle spatial and temporal responses, respectively.

CNN network is considered as an image classification algorithm for robust extraction of image features. As shown in Fig. 13, the CNN architecture has three layers, the first layers have two sub-layers called convolutional (Conv) and max-pooling layers and one flattened layer connected to LSTM. The output of the LSTM is inputted to fully connected layers to predict wear length, wear depth, and surface roughness. The Conv layer in the first layer uses 32 filter sizes of 3×3 on the input image and applies the activation function of Rectified Linear Unit (ReLU) to introduce non-linearity in the data. The max-pooling layer in the first layer reduces the image size or spatial dimensions by selecting the maximum value over the region covered by 2×2 filters. The Conv layer in the second layer applies 64 filter sizes of 3×3 and

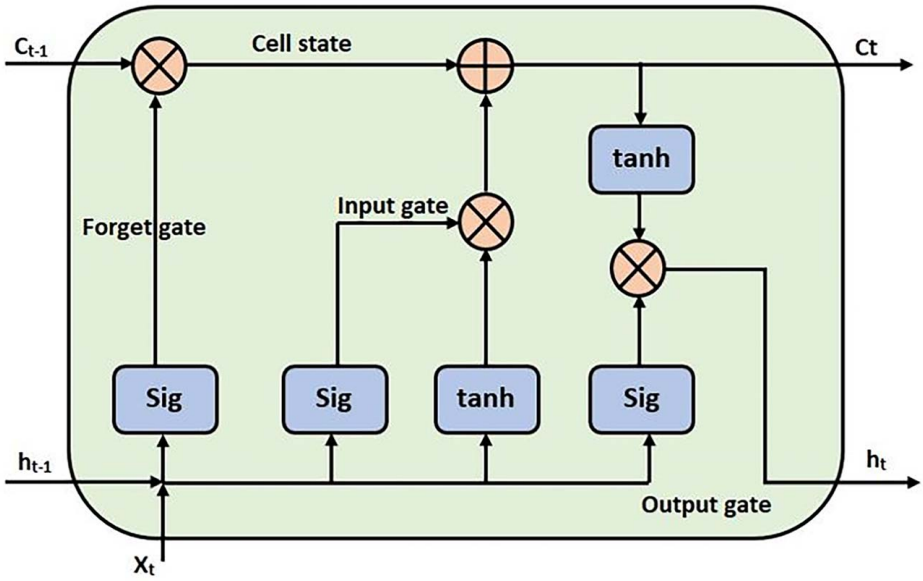

Fig. 14 LSTM Architecture

applies the activation function of Rectified Linear Unit (ReLU) on the input image. The second max-pooling layer again reduces the spatial dimensions of the image by selecting the maximum value over the region covered by 2×2 filters. There is a flattened layer in the CNN that flattens the output of the second max-pooling layer (3D tensor) into a one-dimensional vector.

LSTM is an effective tool for sequential dataset analysis and is adept at processing sequences by embracing long-term dependencies and nonlinear dynamics in sequential data [26]. LSTM has the advantage of remembering the historical data of the network during the prediction process [23]. The CNN-LSTM model fuses the multiple signals and response collected during experimentation and automatically allocate weights to the connections between the layers in the network [35]. As shown in Fig. 14, the LSTM architecture has three gates called as forget gate, the input gate, and the output gate.

Forget gate has a sigmoid layer that takes the output value of the previous step or cell and the input value of the present step as shown in Fig. 14 and takes the decision to retain or forget the information. The output value of the forget gate is estimated using Eq. (25)

$$f_t = \sigma(W_f.[h_{t-1}, \quad X_t] + b_f \tag{25}$$

where f_t is the output vector of the forget gate in the range of 0 and 1, W_f is the weight, h_{t-1} is the output value of the previous cell, X_t is the input value to the present cell, and b_f is the bias value of the forget gate.

The input gate has two layers called sigmoid and *tanh* layers that update the information of the forget gate. In this gate, new information is added to the forget information using Eqs. (26) and (27)

$$i_t = \sigma(W_i.[h_{t-1}, X_t] + b_i \tag{26}$$

$$\tilde{C}_t = \tanh(W_c.[h_{t-1}, X_t] + b_c \tag{27}$$

where i_t is the output value of the sigmoid layer in the input gate in the range of 0 and 1, $\tilde{C}_t$ is the output value of the candidate input gate, W_i, b_i, W_c and b_c are weight of input gate, bias of input gate, weight of the candidate gate, and bias of candidate gate, respectively.

The output again has two layers similar to the input gate that decides the output value of the LSTM network. The sigmoid layer estimates the output value of the output gate using Eq. (28) as follows:

$$O_t = \sigma(W_o.[h_{t-1}, X_t] + b_o \tag{28}$$

where O_t is the output vector of output gate in the range of 0 and 1, W_o and the b_o are weight and bias values of the output gate, respectively. Using the output value of the output gate and the memory cell state (C_t), the output value (h_t) of the LSTM is estimated using Eq. (29).

$$h_t = O_t \cdot \tanh(C_t) \tag{29}$$

Memory cell state is estimated using Eq. (30)

$$C_t = f_t \cdot C_{t-1} + i_t + \tilde{C}_t \tag{30}$$

where C_{t-1} is the memory cell state of the previous step. Output of the LSTM is inputted to the fully connected (dense) layers, and the wear length, wear depth, and surface roughness were predicted. The fully connected layer having 64 units applied the ReLU activation function on the concatenated features of the CNN-LSTM. The output layer has three units corresponding to wear length, wear depth and surface roughness, and a linear activation predicts the responses. The model was compiled with an Adam optimizer to estimate mean square error.

5 Results and Discussion

The hybrid CNN-LSTM model combines the strengths of both models to process both spatial and temporal dimensions. The hybrid model is inputted with 25 features including one image, the CNN handles the image, and LSTM handles the sequence data with 24 features. The CNN-LSTM model was built in Google Colab using PYTHON coding. In the input layer of the CNN-LSTM model, any number of image data and corresponding 24 features of the data were fed to the model. The size of the input image is 128×128 pixels with three colors such as red, green, and blue. The Conv layer in the first layer applied 32 Convolutional filters on the $128 \times 128 \times 3$ image using a 3×3 filter or kernels and 1 step of stride. The spatial dimensions of the input image were reduced to 126×126 with 32 feature maps or output channels as shown in Fig. 14. The total number of weights was estimated by $3 \times 3 \times 3$ (input channels) $\times 32$ filters is equal to 864. The total parameter is estimated as 896 $(864 + 32)$ as the 32 filters include a bias term which helps the Conv layer to learn the

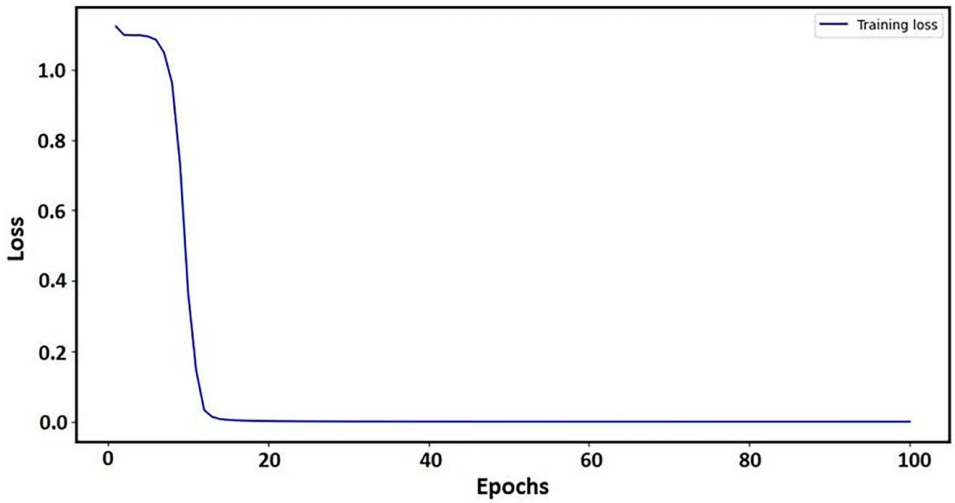

Fig. 15 CNN-LSTM Model training loss

relationship in the input data. The max-pooling layer in the first layer reduced the size of the image to $63 \times 63 \times 32$ from $126 \times 126 \times 32$ using a 2×2 pool size with a stride of one step.

In the second layer, the Conv layer applied 64 filters with 3×3 kernel size and one-step stride and reduced the image size to $61 \times 61 \times 64$. The total number of parameters was estimated to be 18,496 after the convolutional operation. The max-pooling layer reduced the size of the image to $63 \times 63 \times 32$ from $30 \times 30 \times 64$ using a 2×2 pool size with a stride of 1 step. After two layers in the CNN model, the flattened layer flattens the $30 \times 30 \times 64$ 3D tensor into a one-dimensional vector containing 57,600 elements. The LSTM processed the temporal (sequential) data produced 50 units of output and established relationships among the input features using 10,400 parameters. The output layer has three units and predicts the wear length, wear depth, and surface roughness.

5.1 Performance Evaluation of the Model. The training process of the CNN-LSTM model is visualized as shown in Fig. 15, which evaluates the performance of the model. As illustrated in Fig. 15, training loss was gradually reduced in the initial stage of the training process, simultaneously accuracy of the model increased to 0.97 after 15 epochs and continued till the end of the training process. The results were verified with the confusion matrix presented in Fig. 16, where the predicted labels are taken on the *x*-axis and the true labels are taken on the *y*-axis. Class 0, class 1, and class 2 represent the wear length, wear depth, and surface

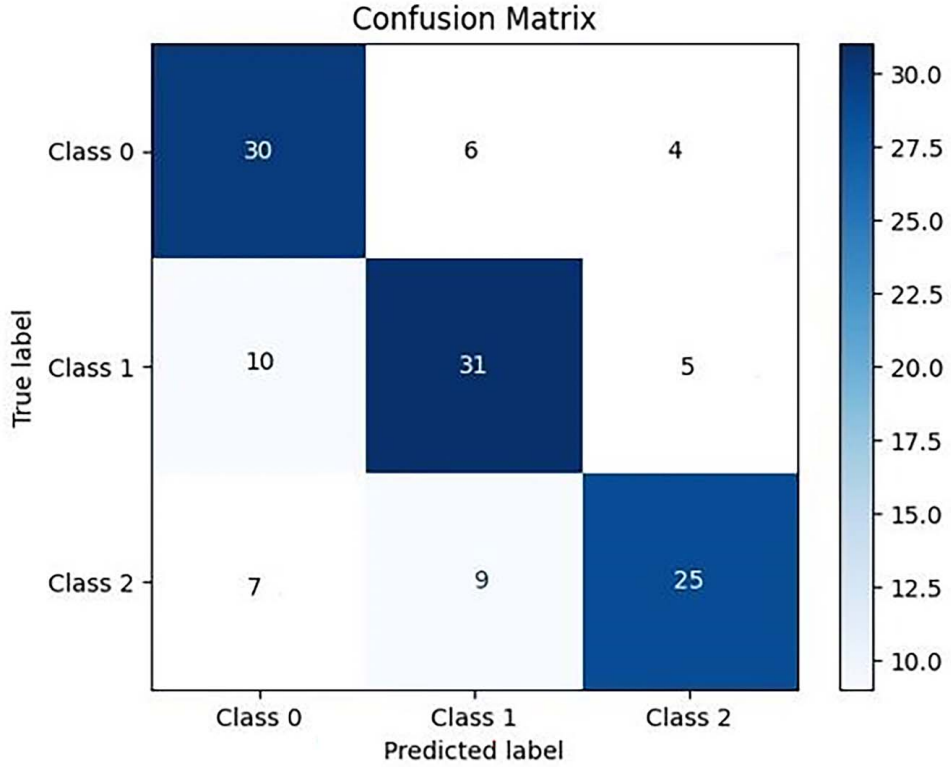

Fig. 16 Confusion matrix for the wear length, wear depth, and surface roughness

Table 5 Performance of the CNN-LSTM model

	RMSE	*R*-squared
Wear length	0.219 mm	0.974
Wear depth	0.018 mm	0.943
Surface roughness	0.216 μm	0.956

roughness, respectively. The confusion matrix ensures that the CNN-LSTM model shows good classification performance for the wear length, wear depth, and surface roughness.

The prediction performance of the CNN-LSTM model is quantified by estimating the evaluation indexes such as root mean square error (RMSE) and *R*-squared (R^2). The RMSE and the R^2 values estimate the performance of the proposed model, smaller values

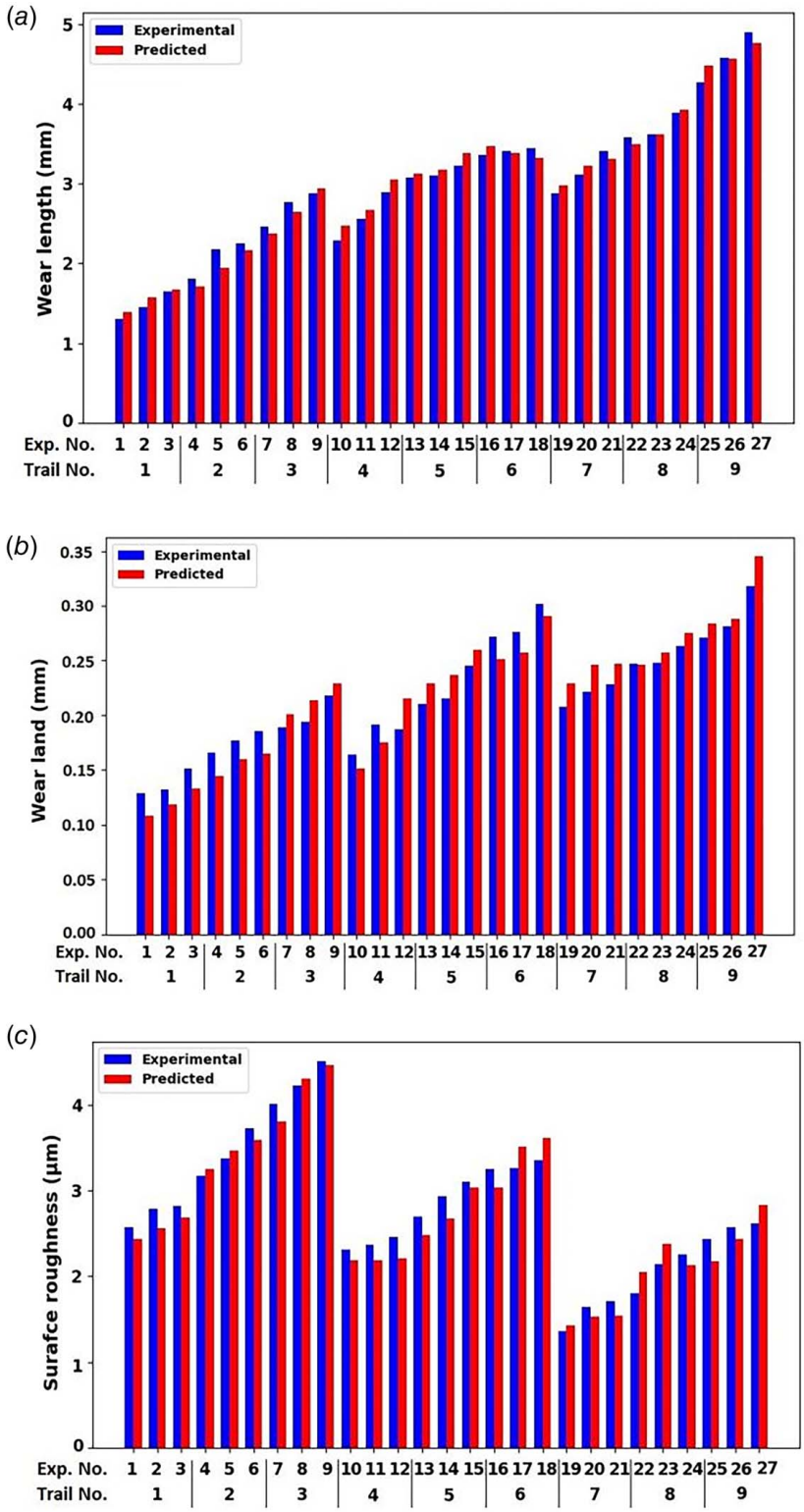

Fig. 17 Experimental and predicted results of (*a*) wear length, (*b*) wear depth, and (c) surface roughness

of RMSE, and the R^2 value close to 1 indicate the model predicts the responses at higher precision and accuracy. These values are estimated using Eqs. (31)–(32), respectively

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - p_i)^2} \tag{31}$$

$$R^2 = 1 - \frac{\text{SSE}}{\text{SST}} \tag{32}$$

where x_i, p_i, SSE, and SST are experimental value, predicted value total sum of squares, and sum of square errors, respectively. The predicted values of the wear length, wear depth, and surface roughness were presented in Table 4. The RMSE and the *R*-squared were estimated between the experimental and predicted values for wear length, wear depth, and surface roughness as presented in Table 5.

The trained CNN-LSTM model was tested using the experimental data presented in Table 4. Figure 17 illustrates the experimental and predicted results for wear length, wear depth, and surface roughness of the CNN-LSTM model. From Table 5 and Fig. 17, it can be observed that the three dependent variables have the smallest value of RMSE and *R*-square close to 1, which indicates that the proposed CNN-LSTM model showed good accuracy in prediction. Thus, it has been verified that the model can predict wear length, wear depth, and surface roughness under different working conditions. Time series data like spatial and temporal dimensions have effectively characterized wear length, wear depth, and surface roughness.

5.2 Analysis of Tool Wear and Surface Roughness. Figures 17(*a*)–17(*c*) describe the effect of the process parameters on the wear length, wear depth, and surface roughness, respectively. On the *x*-axis, trail numbers and the experiment numbers were taken in the three figures. As presented in Table 4, there are three experiments in each trial, and the machining length increased in the experiments by 150 mm at constant spindle speed and feed. The spindle speed and the feed were changed in each trail as per the design of the experiments. The wear length and wear depth are measured on the cutter flank showing the same trend as presented in Figs. 17(*a*) and 17(*b*). According to ISO 3685:1993, wear depth has a significant effect on tool life or remaining useful tool life, and a tool is considered to have failed when wear depth reaches 0.3 mm. The spindle speeds are 5000, 5500, and 6000 rpm in 1–3, 4–6, and 7–9 trails, respectively, and the feed was taken as 5, 7, and 9 μm/tooth in each group. It was observed that the spindle rotation speed has a significant effect on the wear length and wear depth followed by the length of machining and feed. At high cutting speeds and feeds, the cutter flank is subject to different wear mechanisms such as diffusion, oxidation, adhesion, and abrasion, and tool wear progresses as the cutting speed and feed increase [36]. In the milling of Hastelloy-X, Wang and Shen [27] reported that the machining time/length has a significant effect on the flank wear. Peng and Li [28] found that the flank wear, cutting forces, and spindle current increased when the spindle speed increased in the milling of carbon structural steel. Airao et al. [36] studied the effect of process parameters on the tool wear and surface roughness in the milling of wrought Ti–6Al–4 V and SLM Ti–6Al–4 V and found that these two increased when the cutting speed and feed increased.

Tool vibration is greatly affected by feed and tool wear, causing surface roughness. As shown in Fig. 17(*c*), it can be observed that the surface roughness decreased by increasing the cutting speed, but the roughness was found to increase as the feed increased. With high spindle speed, small chips are formed efficiently, and cutting forces are uniformly distributed during the machining process which is responsible for reducing surface irregularities. Larger feeds cause larger chip size and introduce more tool vibration leading to increased surface roughness. Zhang et al. [7] also found that tool vibration reduced when the feed decreased in the machining of NAK80. During the machining process, when the flank wear exceeds 71 μm, the tool vibration increases and forms a chattering pattern on the machined surface, resulting in surface roughness on the machined surface [7]. The same trend was observed in the present study also.

6 Conclusions

The present study proposed a CNN-LSTM hybrid deep neural network and assessed tool condition and surface roughness in the milling of AISI D2 steel. The CNN-LSTM network was trained with heterogeneous data including timely captured tool images, working conditions, vibration data, surface roughness, flank wear length, and wear depth, and the model was validated with the experiments. YOLO-based algorithms were used for annotation and segmentation of wear regions. The uniqueness of the developed HDNN in the present study is that the model can predict the wear length and wear depth using the raw image of the tool captured during the machining process. Statistical analysis of tool wear and surface roughness was carried out. The following conclusions can be from this work:

- The capture images were annotated using Roboflow software and the wear region was segmented using a YOLOv8s, with the mAP values of 0.913 and the mAP@0.5:0.95 value of 0.61. This indicates that the YOLOv8s model exhibits superior performance in the prediction of the worn segment as well as labeling of the segment.
- The CNN-LSTM model effectively predicted the length of flank wear, with the RMSE value of 0.219 mm, and the determination coefficient R^2 value of 0.974; wear depth with the RMSE value of 0.018 mm and R^2 value of 0.943; surface roughness with the RMSE value of 0.216 μm and R^2 value of 0.956.
- The proposed CNN-LSTM methodology is very simple, and versatile and provides a solution to accurately predict tool wear as well as surface roughness in metal cutting applications. Heterogeneous training data included 18 features extracted from the vibration signals improved the accuracy of prediction.
- From the analysis of experimental results, it was observed that the spindle rotation speed has a significant effect on the wear length and wear depth followed by the length of machining and feed. At high cutting speeds and feeds, the cutter flank was subject to different wear mechanisms such as diffusion, oxidation, adhesion, and abrasion.
- The surface roughness decreased by increasing the cutting speed, but the roughness was found to increase as the feed increased. At high cutting speed, formation of small chips and uniform distribution of cutting forces caused increased surface roughness. Larger feeds caused larger chip size and introduced more tool vibration leading to increased surface roughness [37].

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The authors attest that all data for this study are included in the paper.

References

[1] Yang, Y., Guo, Y., Huang, Z., Chen, N., Li, L., Jiang, Y., and He, N., 2014, "Research on the Milling Tool Wear and Life Prediction by Establishing an Integrated Predictive Model," Measurement, **145**, pp. 178–189.

[2] Cheng, Y., Lu, M., Gai, X., Guan, R., Zhou, S., and Xue, J., 2024, "Research on Multi-Signal Milling Tool Wear Prediction Method Based on GAF-ResNext," Robot. Comput. Integr. Manuf., **85**, p. 102634.
[3] Yan, S., Sui, L., Wang, S., and Sun, Y., 2023, "ON-Line Tool Wear Monitoring Under Variable Milling Conditions Based on a Condition-Adaptive Hidden Semi-Markov Model (CAHSMM)," Mech. Syst. Signal. Process., **200**, p. 110644.
[4] Li, X., Liu, X., Yue, C., Liang, S. Y., and Wang, L., 2022, "Systematic Review on Tool Breakage Monitoring Techniques in Machining Operations," Int. J. Mach. Tools Manuf., **176**, p. 103882.
[5] Prasad, B. S., Sarcar, M. M. M., and Ben, B. S., 2010, "Development of a System for Monitoring Tool Condition Using Acousto-Optic Emission Signal in Face Turning—An Experimental Approach," Int. J. Adv. Manuf. Technol., **51**(1-4), pp. 57–67.
[6] Rao, K. V., Kumar, Y. P., Singh, V. K., Raju, L. S., and Ranganayakulu, J., 2021, "Vibration-Based Tool Condition Monitoring in Milling of Ti-6Al-4 V Using an Optimization Model of GM(1,N) and SVM," Int. J. Adv. Manuf. Technol., **115**(5–6), pp. 1931–1941.
[7] Zhang, G., Wang, Y., Huo, Z., Zheng, J., and Zhang, W., 2024, "Tool Wear Induced Multimode Vibration and Multiscale Patterns in Precision Turning NAK80," Wear, **554–555**, p. 205467.
[8] Dutta, S., Pal, S. K., Mukhopadhyay, S., and Sen, R., 2013, "Application of Digital Image Processing in Tool Condition Monitoring: A Review," CIRP J. Manuf. Sci. Tech., **6**(3), pp. 212–232.
[9] Zhu, X., Chen, G., Ni, C., Lu, X., and Guo, J., 2024, "Hybrid CNN-LSTM Model Driven Image Segmentation and Roughness Prediction for Tool Condition Assessment With Heterogeneous Data," Rob. Comput. Integr. Manuf., **90**, p. 102796.
[10] Malhotra, J., and Jha, S., 2021, "Fuzzy c-Means Clustering Based Colour Image Segmentation for Tool Wear Monitoring in Micro-Milling," Precis. Eng., **72**, pp. 690–705.
[11] Sun, Y., He, J., Gao, H., Song, H., and Guo, L., 2024, "A New Semi-Supervised Tool-Wear Monitoring Method Using Unreliable Pseudo-Labels," Measurement, **226**, p. 113991.
[12] Manwar, A., Varghese, A., Bagri, S., and Joshi, S. S., 2023, "Online Tool Condition Monitoring in Micromilling Using LSTM," J. Intell. Manuf.
[13] Makhesana, M. A., Bagga, P. J., Patel, K. M., Patel, H. D., Balu, A., and Khanna, N., 2024, "Comparative Analysis of Different Machine Vision Algorithms for Tool Wear Measurement During Machining," J. Intell. Manuf.
[14] Ramadan, H., Lachqar, C., and Tairi, H., 2020, "A Survey of Recent Interactive Image Segmentation Methods," Comp. Vis. Med., **6**(4), pp. 355–384.
[15] Asadi, R., Queguineur, A., Wiikinkoski, O., Mokhtarian, H., Aihkisalo, T., Revuelta, A., and Ituarte, I. F., 2024, "Process Monitoring by Deep Neural Networks in Directed Energy Deposition: CNN-Based Detection, Segmentation, and Statistical Analysis of Melt Pools," Rob. Comput. Integr. Manuf., **87**, p. 102710.
[16] Yu, B., Li, Z., Cao, Y., Wu, C., Qi, J., and Wu, L., 2024, "YOLO-MPAM: Efficient Real-Time Neural Networks Based on Multi-Channel Feature Fusion," Expert Syst. App., **252**, pp. 124282.
[17] Wang, D., Han, C., Wang, L., Li, X., Cai, E., and Zhang, P., 2023, "Surface Roughness Prediction of Large Shaft Grinding via Attentional CNN-LSTM Fusing Multiple Process Signals," Int. J. Adv. Manuf. Technol., **126**(11-12), pp. 4925–4936.
[18] Shah, R., Pai, N., Thomas, G., Jha, S., Mittal, V., Shirvni, K., and Liang, H., 2025, "Machine Learning in Wear Prediction," ASME J. Tribol., **147**(4), p. 040801.
[19] Agrawal, V., Gajrani, K. K., Mote, R. G., Barshilia, H. C., and Joshi, S. S., 2022, "Wear Analysis and Tool Life Modeling in Micro Drilling of Inconel 718 Superalloy," ASME J. Tribol., **144**(10), p. 101706.
[20] Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2017, "ImageNet Classification With Deep Convolutional Neural Networks," Commun. ACM, **60**(6), pp. 84–90.
[21] Burkov, A., 2019, *The Hundred-Page Machine Learning Book*, Vol. 1, Andriy Burkov Canada, Quebec City, Canada.
[22] Zhou, J. T., Zhao, X., and Gao, J., 2019, "Tool Remaining Useful Life Prediction Method Based on LSTM Under Variable Working Conditions," Int. J. Adv. Manuf. Technol., **104**(9-12), pp. 4715–4726.
[23] Yang, C., Zhou, J., Li, E., Zhang, H., Wang, M., and Li, Z., 2022, "Milling Cutter Wear Prediction Method Under Variable Working Conditions Based on LRCN," Int. J. Adv. Manuf. Technol., **121**(3-4), pp. 2647–2661.
[24] Lim, M. L., Derani, M. N., Ratnam, M. M., and Yusoff, A. R., 2022, "Tool Wear Prediction in Turning Using Workpiece Surface Profile Images and Deep Learning Neural Networks," Int. J. Adv. Manuf. Technol., **120**(11–12), pp. 8045–8062.
[25] Liang, J. H., Gao, H. L., Xiang, S., Chen, L., You, Z., and Lei, Y. 2022, "Research on Tool Wear Morphology and Mechanism During Turning Nickel-Based Alloy GH4169 With PVD-TiAlN Coated Carbide Tool," Wear, **508-509**, pp. 204468.
[26] Marei, M., and Li, W., 2022, "Cutting Tool Prognostics Enabled by Hybrid CNN-LSTM With Transfer Learning," Int. J. Adv. Manuf. Technol., **118**(3-4), pp. 817–836.
[27] Wang, C., and Shen, B., 2024, "Auxiliary Input-Enhanced Siamese Neural Network: A Robust Tool Wear Prediction Framework With Improved Feature Extraction and Generalization Ability," Mech. Syst. Signal Process., **211**, pp. 111243.
[28] Peng, D., and Li, H., 2024, "Intelligent Monitoring of Milling Tool Wear Based on Milling Force Coefficients by Prediction of Instantaneous Milling Forces," Mech. Syst. Signal Process., **208**, pp. 111033.
[29] Qin, Y., Liu, X., Yue, C., Wang, L., and Gu, H., 2025, "A Tool Wear Monitoring Method Based on Data-Driven and Physical Output," Rob. Comput. Integr. Manuf., **91**, pp. 102820.
[30] Li, B., Lu, Z., Jin, X., and Zhao, L., 2024, "Tool Wear Prediction in Milling CFRP With Different Fiber Orientations Based on Multi-Channel 1DCNN-LSTM," J. Intell. Manuf., **35**(6), pp. 2547–2566.
[31] Rao, K. V., Murthy, B. S. N., and Mohan Rao, N., 2015, "Experimental Study on Surface Roughness and Vibration of Workpiece in Boring of AISI 1040 Steels," Proc. Inst. Mech. Eng. Part B: J. Eng. Manuf., **229**(5), pp. 703–712.
[32] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A., 2016, "You Only Look Once: Unified, Realtime Object Detection," Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016, Las Vegas, NV, IEEE, pp. 779–788.
[33] Zhou, C., Wang, C., Sun, D., Hu, J., and Ye, H., 2025, "An Automated Lightweight Approach for Detecting Dead Fish in a Recirculating Aquaculture System," Aquaculture, **594**, pp. 741433.
[34] Gu, W., Bai, S., and Kong, L., 2022, "A Review on 2D Instance Segmentation Based on Deep Neural Networks," Image Vis. Comput., **120**, pp. 104401.
[35] Wang, S., Yu, Z., Xu, G., and Zhao, F., 2023, "Research on Tool Remaining Life Prediction Method Based on CNN-LSTM-PSO," IEEE Access, **11**, pp. 80448–80464.
[36] Airao, J., Kishore, H., and Chandrakant, K. N., 2023, "Measurement and Analysis of Tool Wear and Surface Characteristics in Micro Turning of SLM Ti6Al4 V and Wrought Ti6Al4 V," Measurement, **206**, pp. 112281.