Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 18s (2025)
https://www.jneonatalsurg.com

Application of Deep Learning Algorithms in Studies of Cutting Tool Degradation using Image Processing

Govind Patidar[1], Dr. Lokendra Singh Songare[2]

[1]Research Scholar, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India govind.patidar@medicaps.ac.in
[2]Assistant Professor, Department of Computer Science and Engineering, Medi-Caps University, Indore, Madhya Pradesh, India lokendra.singh@medicaps.ac.in

Cite this paper as: Govind Patidar, Dr. Lokendra Singh Songare, et.al (2025) Application of Deep Learning Algorithms in Studies of Cutting Tool Degradation using Image Processing. *Journal of Neonatal Surgery*, 14 (18s), 113-119.

ABSTRACT
In the industrial industry, it is crucial to keep an eye on how cutting tools are deteriorating. High-quality products in terms of geometry, residual stress, and surface finish are not produced by tools that are heavily worn. Additionally, inefficient tool replacement might result in higher production costs and lost productivity. Thus, keeping an eye on the tool's health is crucial to preventing these extra expenses and guaranteeing high-quality output. In particular, VGG19, EfficientNetV2, and Vision Transformers are among the categorization models examined in this paper. These models use the tools' images to categorize their condition. The top-performing AI-based image analysis models are compared using transfer learning to determine which are best suited for cutting tool monitoring. They are compared in terms of explainability, performance, and generalizability. With an accuracy of 94%, VGG19 is the model that performs the best, followed by ViT and EfficientNetV2, both of which have accuracys of 87%. A thorough comparison of these findings is done.

Keywords: Cutting Tool; Degradation; Monitoring; Deep-Learning; Transfer Learning

1. INTRODUCTION

Tool wear results from the interaction between the tool and the workpiece during machining operations. Imperfect cutting from this wear may lower the quality of the parts that are produced. The state of the tool affects a number of variables, including the workpiece's geometry, residual stresses, and surface polish. It is estimated that the cost of cutting tools can account for anywhere from 3% to 12% of manufacturing expenses, and wear is a significant issue in machining operations. Up to 20% of production stoppages can be attributed to tool failure brought on by wear [1]. Regular tool replacement is necessary to prevent excessive tool wear. Testing at the beginning of production and taking safety margins to determine whether to replace the tool are standard industrial practices. Additionally, replacement is occasionally determined by the machine operator's judgment regarding whether the tool requires replacement. It is obvious that the existing solutions lack objective standards for changing the tool at the best time, which results in waste and expenses. Using decision support techniques to assess whether a tool is worn out or still functional is helpful in determining the best time to replace it.

Cutting tools are increasingly being monitored using artificial intelligence (AI) as a decision support technique. There are a number of AI methods for tracking cutting tool deterioration, which are generally separated into two groups: direct and indirect. Indirect methods use machine-installed sensors to infer the tool's condition from cutting signals. The strategies that have been developed in this sector are presented in an extensive overview [2]. These indirect methods are challenging to adapt to more complicated industrial situations and are frequently unique to a particular experimental condition or setup. To determine whether the cutting tool can still execute the cutting action, the direct way is taking pictures of the tool. Although this approach is frequently more accurate, human interpretations and definitions of wear may lead to more ambiguous outcomes. As a result, there exist numerous tools and methods for picture analysis [3]. However, wear identification is difficult for conventional image analysis since these images are often captured in different lighting conditions and obscured by chips or cutting fluid. However, artificial intelligence techniques get around these problems by accurately detecting wear even under these circumstances [3]. Using pictures, numerical AI techniques have been used to recognize or classify the area of tool wear. For instance, Pagani et al. deduced the tool's state based on the color of the chips [4]. In order to differentiate between various wear categories, Wu et al. compared their own model with VGG16 in 2019. Although their model was faster, the two models' performances were similar [5]. To categorize the tool's condition, several researchers have additionally combined transfer learning models such as VGG16, LeNet, and others with pictures of the vibration spectrogram [6]. In both milling [7] and turning [8], other researchers have used U-Nets to directly segment the wear zone on tool pictures. Building models from scratch is generally not advised in the field of image classification or segmentation. In fact, to ensure stable and dependable performance, deep learning models typically need a lot of data and meticulous adjustment. For this reason, in this kind of application, transfer learning is frequently the recommended strategy. By offering models that are by nature more durable than those created from scratch, it provides a substantial benefit. Although there are many strategies for monitoring tool conditions, it is still unclear from the research how important transfer learning is and whether techniques created for image classification tasks are applicable. Therefore, the usage and comparison of three AI techniques for picture categorization utilizing transfer learning constitutes the novelty of this article. The three baseline architectures are Visual Transformers [11], Efficient-NetV2 [10], and VGG19 [9]. These methods are tested on altered photographs that are representative of those taken in an industrial setting, with brightness and contrast alterations as well as rotation. The efficacy of these approaches in adapting to these changes is examined. Explainability analysis is frequently used to draw attention to the difficulties these networks face. Lastly, a conversation regarding the situation that is most appropriate for industry is conducted. This method automatically detects and uses deep learning skills to robustly and automatically recognize the wear zone, in contrast to the manual wear studies currently performed in industry.

2. DATABASE

A turning database derived from experimental turning experiments on C45 steel bars is utilized to train the various methods discussed in this article. These bars were machined on a lathe, in order to minimize the amount of material used during the testing and to favor the appearance of wear, the tool has one of the lowest grades. Throughout the straight turning testing campaign, thirty instruments were utilized. Over the course of its life, each instrument was inspected six to seven times on average. Taking a photo of the insert is part of the inspection process. As a result, there are 180 pictures of the tool's flank face in the database. Although the cutting speeds of these tools varied, they were employed under identical cutting settings (Table 1). A Byameyee EU-1000X 3 digital portable microscope was used to examine the instrument every two minutes and forty seconds. A measurement of tool wear is taken from the 460 by 640-pixel color images that are obtained. Fig. 1 shows an example of a picture taken during the experimental campaign.

Table 1. Cutting conditions during experimental turning tests

Test	Cutting Speed [m/min]	Feed [mm/rev]	Depth of cut [mm]
1 to 10	250	0.2	1
11 to 20	240	0.2	1
21 to 30	275	0.2	1
31 to 40	Variable	0.2	1

In this article, the tool's condition is categorized using artificial intelligence. It has been decided to divide the tool's condition into three classes, each of which represents a distinct stage of its life (Fig. 2). A tool is considered new if its wear is between 0 and 150 microns, moderately worn if its wear is between 150 and 300 microns, and worn if its wear is greater than 300 microns. This classification scheme is designed to account for the many phases of a cutting tool's life. A new tool first degrades quickly in class 1 before entering class 2, where it stays for the majority of its service life. In contrast to class 1, the tool's rate of degradation decreases during this class 2 phase. The tool eventually experiences severe wear after a few minutes, which causes it to degrade quickly and fall into class 3. A tool is deemed "worn" in this last class if its flank wear is greater than 300 microns. The ISO 3685 standard sets this cutoff point [12].

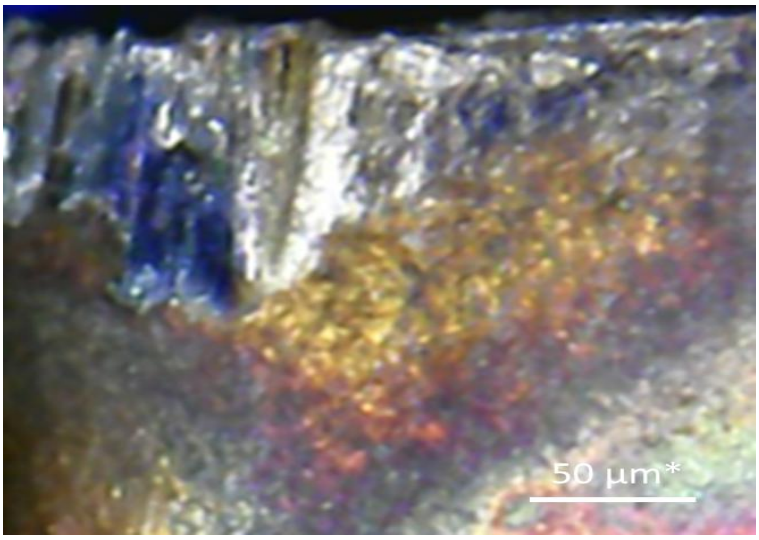

Fig. 1. Image of an insert at 280 µm magnification

However, the database is not uniform; there are more pictures of brand new instruments than old ones, which can be a drawback during AI training. The data are oversampled in order to correct for this imbalance. In order to have as many photographs in each class as possible, this strategy repeats the number of images in a class. This method's drawback is the potential repetition of the same image multiple times in the training database. To mitigate this issue, data augmentation is used, so that even 2 identical images are not augmented twice in the same way, thereby expanding the diversity of data for training. Data augmentation is frequently implemented in databases to compensate for a limited quan- tity of images and to generalize images under conditions that are unevenly represented in the database. Data augmentation is used here to generate new, unique images from the images in the database [13]. Two types of augmentation are considered: image manipulation and lighting modification During training, each image is randomly augmented using a combina- tion of the following modifications:

2.1. Image manipulation
Horizontal flip: There's a 50% probability that the image will be horizontally flipped, resulting in a mirrored version of the original. This modification applied on Fig 1 is shown in Fig. 3(a).
Image Rotation: This technique rotates the image. The rotation angle is randomly chosen between -20 to 20 degrees. This modification applied on Fig 1 is shown in Fig. 3(b)

2.2. Lighting modification
Contrast modification: This technique modifies the image contrast by a factor randomly chosen in the range -0.2 to 0.2. This modification applied on Fig 1 is shown in Fig. 3(c).
Brightness modification: This technique randomly changes the brightness of the image by a factor ran- domly chosen in the range -0.2 to 0.2. This modification applied on Fig 1 is shown in Fig. 3(d)

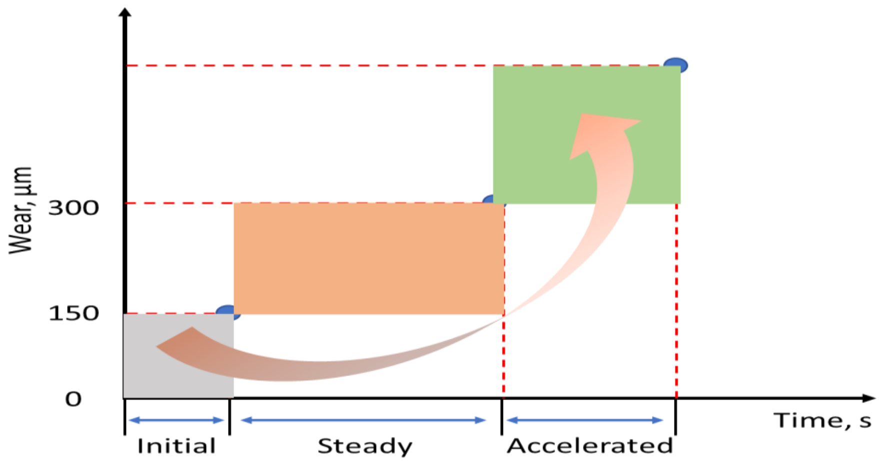

Fig. 2. Tool degradation classified

The selection of these modifications is driven by the objective to reproduce any distortion on image acquired indus- trially. Adjusting the image orientation is intended to replicate the inconsistencies and imperfections of real-world image cap- ture. Modifying the image by altering its contrast or brightness mimics the disruptive elements of the industrial environment which prevent measurements being taken under constant and controlled lighting conditions. The training domain of the net- works therefore consists of images of a single type of tool machining at different cutting speed and acquired under different lighting conditions. In order to test the ability of an AI approach to learn this training domain, a testing database needs to be defined. This database must represent all existing cases and it consist of 5 images of each class.

3. MODELLING FOR TOOL DEGRADATION STUDY
Artificial intelligence offers multiple strategies for image classification. The most straightforward approach involves building and training a deep neural network from scratch on the database. While this method is certainly viable, it's often more advantageous to utilize a pre-existing, pre-trained architecture that has been trained on extensive databases. This makes it pos- sible to take advantage of an architecture that is already capable of extracting valuable features from images thanks to its train- ing. This approach, known as transfer learning, involves using an existing network and modifying its output to adapt it to a new classification task [14].
In the following, different pre-existing models will be used. These models are selected based on their performance in dif- ferent classification tasks as well as the different approaches behind them. These models are:

3.1. VGG19

Developed in 2014, VGG19 is a deep convo- lutional neural network that aims at simplicity [9]. It is com- posed of 16 3x3 convolutional layers and 3 fully con- nected layers. It was trained on the ImageNet dataset (14 millions of images) to classify a thousand different classes. Its architecture is often used as a reference in image clas- sification.

3.2. EfficientNetV2-M

Introduced in 2019, it is an ex- tension of the initial EfficientNet [10]. The particularity of this model lies in its balance between model size, performance and computational efficiency. It was also trained on ImageNet. The EfficientNetV2 model family includes several variants, in this case the M stands for medium and is selected for its compromise between complexity and accuracy.

3.3. Vision Transformers (ViTs)

The Transformer model, originally designed for natural language processing, was adapted in 2020 for image data, leading to the creation of ViTs [11]. ViTs work by dividing images into patches, each of which is converted into a vector and processed by a transformer encoder. Unlike Convolu- tional Neural Networks (CNNs) that mainly capture lo- cal features, ViTs excel in identifying both local and global features, including long-range dependencies be- tween patches. This makes them especially useful for tasks requiring a comprehensive understanding of an im- age. Since 2020, ViTs have generated significant interest in the field of computer vision.

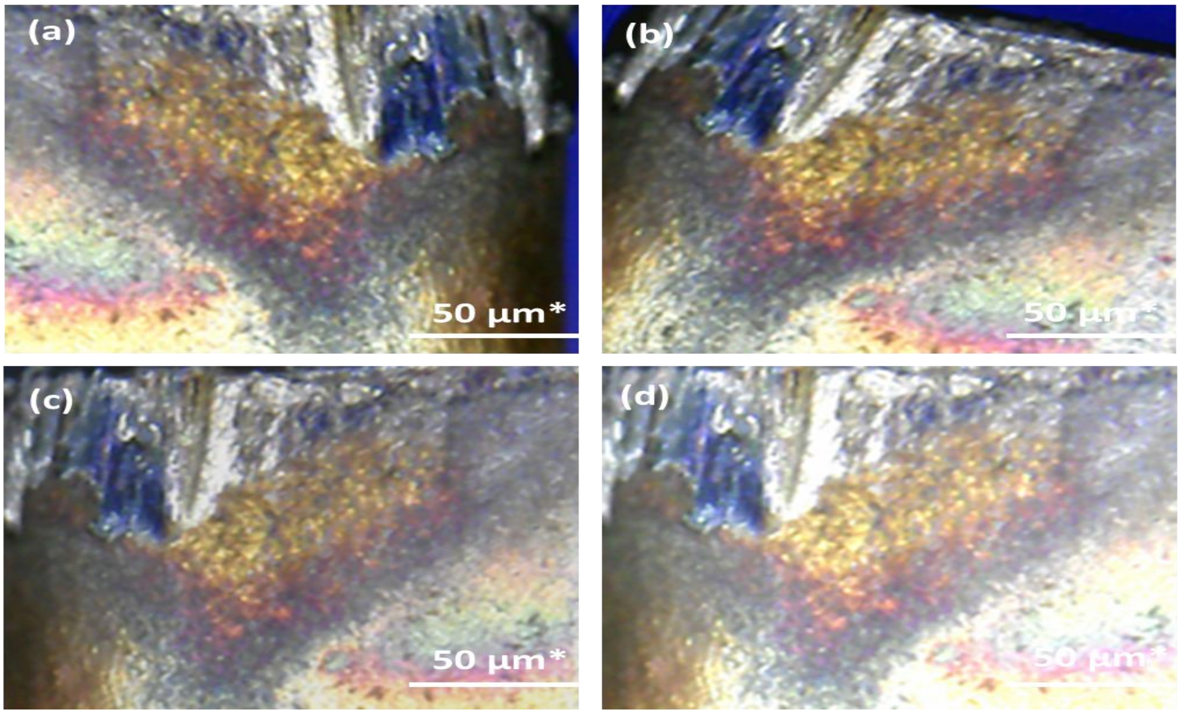

Fig. 3. Image transformations applied on image from Fig. 1 (a) Horizontal flip (b) Image rotation (c) Modification of contrast (d) Modification of brightness

The choice of these models is driven by the intent to contrast three distinct approaches: an initial “classic” model, an im- proved version of this type of model, and finally an innovative approach to vision through transformers. This provides a com- parative perspective on the different image processing methods. As these models have all been pre-trained on ImageNet, the en- tries for these networks are colour images of 224 by 224 pixels. The images in the database are therefore resized to be compati- ble with this shape.

To adapt these models to classify the state of the tool, the last layers are modified to classify the tool state. The layers that are added consist of a Maxpooling2D layer, a layer con- taining 1024 neurons with ReLu activation function, a dropout layer and finally a classification layer representing the 3 possible states of the tool.

All the approaches are trained under comparable conditions with an AMD Ryzen 9 7950 X3D CPU. The input for each model is a color image with dimensions of 224 by 224 pix- els. The Adam optimizer is employed during the training pro- cess. The loss function” Categorical crossentropy” is used. The metrics for evaluation is” precision”,” recall”,” F1-score” and” accuracy”. A maximum of 300 epochs is set. To avoid overfit- ting and eliminate unnecessary calculations, an early stopping mechanism is implemented. This mechanism stops the learning process if there is no improvement in network performance over a period of 60 epochs. To obtain optimal results, the learning rate is progressively reduced as the model approaches convergence.

4. RESULTS AND DISCUSSION

Table 2 presents a comparative analysis of the performance across the three network architectures previously discussed. It provides detailed results, segmented by class and the type of image modification applied. Each type of modification is com- posed of 5 images per class. Therefore, there are 15 test images per modification. The following observations are drawn from this table.

VGG19 is the architecture that obtained the best results in this article. It obtained an overall accuracy of 94% and F1 scores of 0.95, 0.92 and 0.97 for classes 1 to 3 respectively. The model reached convergence after 120 epochs, taking approximately 22 minutes. Once trained, this model is capable of making a prediction, known as inference time, in 40 ms. The data presented in the table indicates that the architecture retains its ability to accu- rately identify tool wear, regardless of the changes made to the image.

The EfficientNetV2-M model achieved an overall accuracy of 87%, making it almost as good as VGG19. This model can perfectly identify class 1, but it seems to be less accurate for classes 2 and 3. However, even with the original image (referred to as 'Initial' in Table 2), it makes mistakes in identifying class 2 and 3. This model is quicker to train than VGG19, taking only 7 minutes. It's worth noting that this model reached its best performance quite fast, in just 65 cycles of training. To put it in per- spective, the EarlyStopping function waits for 60 cycles without improvement before stopping, which means the network converges in just 5 cycles. The inference time, is similar to VGG19, at 80 milliseconds. The ViT model achieved the same overall accuracy as EfficientNetV2. However, it tended to classify class 3 less well than the other architectures. This architecture is also the longest to train, with a total training time of around 3.5 hours for 300 epochs. The inference time is also the longest with 2 s per image.

Table 2. Comparison of results for the different methods segmented by class and the type of image modification applied

Technique	Wear Class	Initial	Flip	Rotated	Contrast	Bright	Precision	Recall	F1	Accuracy
VGG19	Initial wear	100%	80%	100%	100%	100%	94%	98%	0.96	96%
	Seady state	100%	100%	80%	80%	100%	90%	94%	0.94	
	Accelerated wear	100%	100%	80%	100%	100%	100%	92%	0.96	
Efficient NetV2	Initial wear	100%	100%	100%	100%	100%	83%	99%	0.93	84%
	Seady state	80%	80%	60%	80%	80%	80%	82%	0.82	
	Accelerated wear	80%	80%	80%	80%	80%	100%	79%	0.88	
ViT	Initial wear	100%	100%	100%	100%	100%	86%	99%	0.93	89%
	Seady state	100%	80%	60%	100%	80%	79%	88%	0.81	
	Accelerated wear	80%	80%	60%	80%	60%	100%	72%	0.84	

A straightforward comparison of results serves as a useful performance indicator, but it does not fully determine whether one method outperforms another. This is particularly true when the accuracies achieved by the architectures are quite similar, hence, more detailed analysis are needed. The goal of classify- ing the tool's condition means that detecting class 3, or a worn tool, is the only classification that can influence the tool's re- placement. The wear limit is set at 300 microns, so in a strict case, a tool with 299 microns of wear would still be consid- ered as moderate wear being only 1 micron away from being worn. In practice this is not the case, the boundary between a usable tool and a worn one is not always clear and needs to be considered. To illustrate this fact, Table 3 shows the po- sition of errors made by the different architectures in different classes. A buffer zone of 50 microns, 25 for each class, is added between the classes, creating a transition zone between classes 1 and 2, and between classes 2 and 3. The first zone covers wear from 125 to 175 microns (transition from class 1 to 2), and the second covers wear from 275 to 325 microns. Analysis of the results in this zone reveals where the boundary between a us- able tool and a worn one becomes unclear. Table 3 indicates that the EfficientNetV2 architecture made 6 errors in the tran- sition zone between class 2 and 3. This is higher than the other approaches. However, these errors are of little consequence in practical applications. Indeed, this error is less than 25 microns which have almost negligible impact on the quality of produc- tion. Therefore, although EfficientNetV2 has the same accuracy as ViT, its errors have less impact in practice.

In addition to the position of errors, it is also necessary to understand the reason for a correct or incorrect classification. In order to explain and visualise the cause of the classification made by the architectures, a Grad-CAM (Gradient-weighted

Class Activation Mapping) method is used [15]. Grad-CAM is a technique utilized for understanding and explaining the de- cisions made by a CNN in image classification. By analyzing the gradients in the last convolutional layer of the CNN, this technique determines the importance of each region of the im- age. In other words, it indicates the areas of focus of the net- work during its classification process. Grad-CAM serves as a crucial tool in explaining the workings of the CNN and verifies that the network has successfully learned the desired patterns. Fig. 4(a) shows an example of an attention map obtained us- ing Grad-CAM on the VGG19 architecture. This attention map highlights the region of the image used to predict the condi- tion of the tool. In particular, the focused area is located on the tool's wear zone, enabling the network to correctly categorise this image. The area of focus can differ based on the architecture employed. In this study, all accurately classified images have a concentrated attention located on the area of wear. This proves that the methods have identified that the distinction be- tween a new tool and a worn tool is attributed to the amount of flank wear. Fig. 4(b) shows the VGG19 attention map for the classification of the same image as Fig. 4(a) but this image has undergone a change in contrast. In this case, the network wrongly classified the image as representing a new tool. This change in contrast appears to have misled the network, causing it to focus on an incorrect part of the tool for its prediction. The highlighted area clearly indicates that the network has focused on a part of the tool that shows minimal signs of wear. The same conclusions can be drawn for all the errors made by the different approaches. In general, misclassification is due to an inability to detect the area of wear on the image. This is either because the image has been modified or because the original image contains features that make it difficult for the AI to detect.

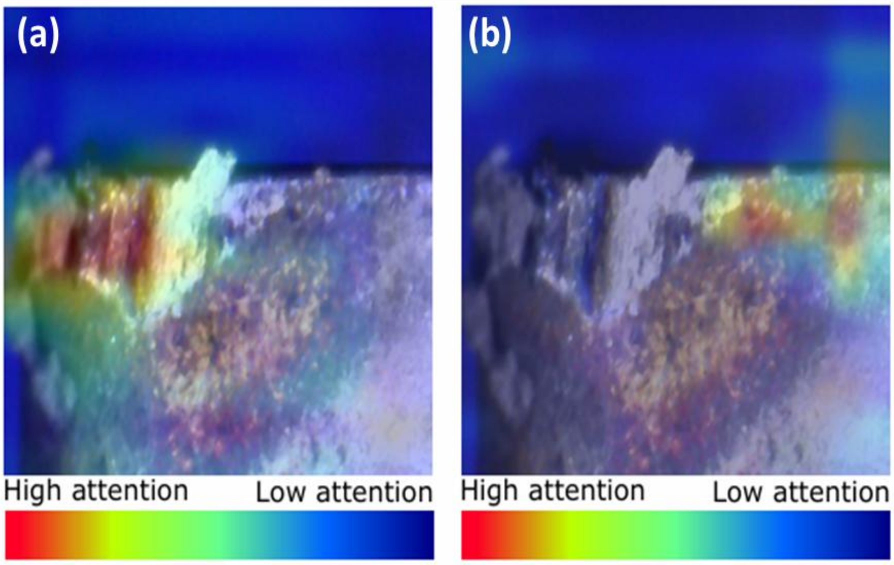

Fig. 4. Attention map obtain with a Grad-CAM analysis of an image classified with VGG19. (a) Attention map of a correctly classified images, the attention map is located on the wear (b) Attention map of an incorrectly classified image with VGG19.

This article explores the use of transfer learning to classify the state of cutting tools based on their images. Three architectures pre-trained on the ImageNet database are utilized for this purpose: VGG19, EfficientNetV2, and Vision Transformers. These models are employed to categorize the state of the tools into three classes: new, moderate, or worn. The images used for training the model are derived from experimental tests conducted during turning operations. The original dataset is augmented and balanced through data augmentation and over- sampling. This is done to assess the robustness of the different architectures against variations in brightness, contrast, and image orientation that may occur in an industrial environment.

Table 3. Comparison of the position of errors for the different architectures. A transition between classes is added.

Approach	Position of errors in the classification				
	Class 1: 0 to 125 µm	Transition 1- 2: 125 to 175 µm	Class 2: 175 to 275 µm	Transition 2-3: 275 to 325 µm	Class 3: 325+ µm
VGG19	1	0	2	1	1
EfficientNetV2-M	0	0	6	6	0
ViT	0	1	3	1	7

Among the approaches, VGG19 yielded the best results with an accuracy of 94%. It was closely followed by EfficientNetV2 and Vision Transformers (ViT), both achieving an accuracy of 87%. All models demonstrated robustness against image modi- fications, showcasing the strength of transfer learning.

In terms of speed, EfficientNetV2 was the fastest model to train and query, with a training time of just 7 minutes and an in- ference time of 80 ms. VGG19, while slightly faster in querying (70 ms), took more than three times longer to train, with a time of 22 minutes. The ViT approach was the slowest of all, with a training time of 3 hours and an inference time of 2 seconds.

A detailed analysis of the errors made by the networks re- vealed that even though EfficientNetV2 and ViT have the same overall accuracy, the errors committed by EfficientNetV2 oc- cur at the transition between classes and it makes fewer errors in the last class. Consequently, the errors made by Efficient- NetV2 have less negative impact on tool replacement compared to those made by ViT.

In addition to the accuracy of each technique, the Grad-CAM method provides additional information on the ability of the networks to detect the area of flank wear. The analysis reveals that the networks successfully located the region of interest in the images, which corresponds to the flank wear region. In ad- dition, this examination highlights the reasons for the misclas- sification of certain images, illustrating the challenges faced by the networks in recognising the area of wear.

5. CONCLUSION

In conclusion, for databases similar to the one presented in this article, we recommend using CNN approaches such as VGG19 and EfficientNetV2 to classify the state of cutting tools from their images. Thanks to transfer learning, it is possible to detect excessive tool wear and therefore replace it at the most optimal time. Future studies could explore detecting and classifying tool defects during machining, like tool's plastic deformation, un- usual damage, etc. Another direction of research is to automati- cally identify the wear zone and damages and measure the wear based on this identification. The attention map obtained in this article indicates that the networks can automatically identify the wear zone. An image segmentation network would make it possible to measure the wear zone and thus predict a remaining useful life of the tool. In this study, the database is augmented and limited to a single type of tool. An analysis on a more diversified database could also help industries to better understand the implementation and limitation of these techniques.

REFERENCES

1. S. Javed, M. Khaisar, M. Shakoor, and P. Raja. 2024. Development of an ANN model for prediction of tool wear in turning EN9 and EN24 steel alloy. Adv. Mech. Eng., vol. 13, no. 6. p.16878140211026720.
2. Equeter, L., Dehombreux, P., & Ducobu, F. (2021). A systematic literature review of cutting tool wear monitoring in turning by using artificial intelligence techniques. Machines, 9(12), 351.
3. K. Shunmugesh 2023. Artificial intelligence-based tool condition monitoring for digital twins and industry 4.0 applications. Int. J. Interact. Des. Manuf. IJIDeM, vol. 17, no. 3, pp. 1067–1087.
4. Parenti, P.; Cataldo, S.; Scott, P.J.; Annoni, M. Indirect cutting tool wear classification using deep learning and chip colour analysis. Int. J. Adv. Manuf. Technol. 2020, 111, 1099–1114.
5. Liu, Y., Zhou, X., & Mou, A. (2022). Automatic identification of tool wear based on convolutional neural network in face milling process. Sensors, 19(18), 3817.
6. Pardeshi, S. S., & Patange, A. D. (2023). Health Monitoring of Milling Tool Inserts Using CNN Architectures Trained by Vibration Spectrograms. CMES-Computer Modeling in Engineering & Sciences, 136(1).
7. Holst, C., Gupta, P., & Augspurger, T. (2022). Digital image processing with deep learning for automated cutting tool wear detection. Procedia Manufacturing, 48, 947-958.
8. Zhao, Z., Sun, C., Li, B., & Yan, R. (2022). A U-Net-based ap- proach for tool wear area detection and identification. IEEE Transactions on Instrumentation and Measurement, 70, 1-10.
9. Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
10. Le, Q. (2021, July). Efficientnetv2: Smaller models and faster training. In International conference on machine learning (pp. 10096- 10106). PMLR.
11. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
12. ISO 3685—Tool Life Testing with Single-Point Turning Tools. 1993. Available online: https://www.iso.org/fr/standard/9151.html (accessed on 27 may 2024).
13. Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre- processing and data augmentation techniques. Global Transitions Proceedings, 3(1), 91-99.
14. Iman, M., Arabnia, H. R., & Rasheed, K. (2023). A review of deep transfer learning and recent advancements. Technologies, 11(2), 40.
15. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2019). Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision (pp. 618-626).