# MicroFiberDetect: An Application for the Detection of Microfibres in Wastewater Sludge Based on CNNs

**Félix Martí-Pérez [1], Ana Domínguez-Rodríguez [2], Cèsar Ferri [1], Carlos Monserrat [1]**

[1]Universitat Politècnica de València, Valencian Research Institute for Artificial Intelligence (VRAIN), València, Spain
[2]Universitat Politècnica de València, Instituto de Seguridad Industrial, Radiofísica y Medioambiental (ISIRYM), València, Spain
fmarper@upv.edu.es, andorod1@upv.edu.es, cferri@dsic.upv.es, cmonserr@dsic.upv.es

## Abstract

Microplastics and microfibres are now widespread in aquatic ecosystems, as oceans and rivers. A serious portion of these microplastics come from urban wastewater treatment plants. Traditional methods for detecting and quantifying them are labour-intensive and time-consuming. This paper introduces MicroFiberDetect, a novel application designed to enhance the detection and quantification of microfibres within sludge samples. Leveraging the power of deep learning, this innovative tool provides detection accuracy and insights into the size and colour of each identified fibre. Reducing time and manpower required for quantification and analysis, while increasing accuracy and throughput. The application has been deployed as a desktop application that allows field experts to quantify and analyse microfibres in sludge samples.

**Code —**
https://github.com/femartip/Detection_Microfibers_APP

## Introduction

The presence of microfibres and microplastics in waterways represents a global environmental issue (Andrady 2011), largely due to urban wastewater treatment plants (Zhou et al. 2023). These particles, often under five millimetres, present a significant thread to aquatic ecosystems and humans as they accumulate in the food chain (Eze et al. 2024). Traditional methods for detecting and quantifying microfibres in sludge, such as manual microscopy and chemical analysis, are time-consuming and lack scalability (Prata et al. 2019) (Löder and Gerdts 2015).

This paper introduces MicroFiberDetect, a deep learning application that optimises microfibre detection in sludge. Images of high resolution, captured using a stereo microscope, are subjected to processing by a Mask R-CNN (He et al. 2018) model that has been trained on data labelled by experts. The automated system enhances the precision of microfibre detection while reducing the necessity for manual intervention. Additionally, our architecture is capable of predicting the colour and length of each detected fibre, providing further valuable information for environmental analysis. Knowing the length and colour of the fibres can give additional information and understanding of their source and potential environmental impact.

Some works have explored the use of deep learning for automatic microplastics counting and classification. In (Lorenzo-Navarro et al. 2021), the authors propose an architecture based on deep learning networks to count and classify microplastics in the range of 1–5 mm, sourced from images of microplastics manually collected from beaches. The proposed architecture comprises a first stage for segmenting the particles of the image, implemented with the U-Net neural network (Ronneberger, Fischer, and Brox 2015), and a second stage based on the VGG16 neural network (Simonyan and Zisserman 2014), which classifies them into three types: fragments, pellets and lines. A similar approach to detect microplastics by scanning electron microscopy is presented in (Shi et al. 2022), again with the VGG16 neural network. Similar approaches have been employed in detecting microplastics in filtered water samples (Bianco et al. 2020; Zhu, Yeung, and Lam 2021) and farmland soil (Ai et al. 2023). Nevertheless, the aforementioned models do not account for length and colour and the utilisation of this model necessitates a certain degree of technical expertise in computer science. In this paper, we present an easy-use tool that is able to detect microfibres in sludge images, predict their colour and length, based on a modern Mask R-CNN (He et al. 2018) model.

## Data

For all the images, the same methods for sample extraction were applied. First, the peroxidised samples were filtered using either a glass fibre or cellulose acetate fibre. Later, they were dried in the oven for two and a half hours. We used a Leica S APO Zoom 8:1x stereo microscope to capture the images. The stereo microscope was then connected to a computer running the Leica application suite software. When a region of fibres was detected, a picture was captured, and the image was saved in a user-created folder with the selected magnification. These images can be taken at magnifications ranging from 10x to 80x.

This resulted in two distinct labelled datasets. The first dataset comprises of 1,203 images (4000x3000) extracted using a glass fibre filter (Glass dataset), where organic matter interferes with image clarity, restricting detection to coloured fibres. The second dataset comprises of 1,287 images (3840x2160) extracted using a cellulose acetate filter (CA dataset), which removes organic matter, enabling detection of transparent fibres. Both datasets contain equal numbers of images with light and dark backgrounds.

## Architecture

This section presents the architectural approach employed in the MicroFiberDetect application.

### Fibre Detection

Detectron2 (Wu et al. 2019), developed by Meta, provides high-quality implementations of detection and segmentation algorithms. We chose the Mask R-CNN (He et al. 2018) architecture for its balance between accuracy and speed, making it suitable for large-scale image analysis without requiring high-performance hardware such as GPUs or TPUs. This model was pretrained on ImageNet (Deng et al. 2009), a dataset with over 14 million labelled images across 20,000 categories.

The Mask R-CNN was fine-tuned on our data, with the glass dataset images resized to 1000x750 and the CA dataset images to 1280x720, maintaining essential details. Both datasets were split into subsets with light and dark backgrounds. Due to scarcity of data, we applied stratified five-fold cross-validation for hyperparameter tuning and performance evaluation, using average precision (AP) (Zhang and Zhang 2009) as our main metric.

Further insight into the training methodology and evaluation process can be found in "Detection of Microfibres in Wastewater Sludge with Deep Learning", which has been submitted to the journal Engineering Applications of Artificial Intelligence.

### Colour Prediction

The prediction of the colour is conducted in the following way. Firstly, the region containing the fibre is isolated by combining the output mask generated by the Mask R-CNN model with the original image, thereby extracting the fibre-specific region of interest (ROI). Once the ROI has been extracted, k-means clustering (Ikotun et al. 2023) is applied to the image's HSL colour space, using three clusters determined by the elbow method (Bholowalia and Kumar 2014). This clustering groups pixels based on colour attributes.

The primary colour of the fibre is identified from the centroid of the main cluster, representing the dominant colour profile. It is important to note that simpler methods, such as averaging pixel colours, are inadequate as they fail to capture colour variability across the image.

### Length Extraction

The accurate determination of a fibre's length is a challenging task due to the complex morphology of the fibre. Simply using the bounding box size is insufficient, as it overlooks the fibre's intricate structure. To address this issue, we used the Mask R-CNN model's predicted mask, but simplified it with a skeletonisation algorithm (Zhang and Suen 1984), which removes redundant information.

The process of skeletonisation reduces the mask to a one-pixel-wide representation, thereby preserving the essential structure of the fibre. This enables precise length measurement. By counting the pixels in the skeleton and applying a conversion factor based on real-world units, we can accurately estimate the fibre's length, accounting for both its complexity and image scale.

## Application

The MicroFiberDetect application was developed as a user-friendly tool for interacting with the aforementioned architectural model. The software has been constructed using the Python programming language, incorporating the PyQt6 and PySide6 libraries. The package, with an estimated size of approximately 1 GB, can be executed in Windows by simply extracting the compressed file. In our tests, the processing times averaged 4 seconds per image on an Intel i7-8750H CPU, indicating that the app is accessible for widespread use.

The user-friendly interface enables the selection of the appropriate model based on the filter type (glass or activated carbon) and the adjustment of the scale to align with the target image. User experts can feed all the samples collected from the target wastewater treatment plant into the application. After processing, users are presented with a tabular summary displaying the total number of fibres identified and comprehensive information for each image. This encompasses the fibre count, prediction confidence, colour profile, and fibre length in micrometres.

We hope this information helps in two distinct ways. Firstly, researchers can make a precise estimation of the number of fibres present in their samples, which can then be extrapolated to estimate the total number of fibres in larger systems. Secondly, additional information regarding size and length may provide insight into the source materials of the fibres, which could assist in identifying potential sources of pollution. We hope all this information providing researchers with invaluable data for further analysis.

## Conclusions

This paper presents an application to enhance the detection and quantification of microfibres in sludge. Microplastics and microfibres, primarily from urban wastewater, are pervasive pollutants in aquatic ecosystems, posing significant environmental and health risks. Traditional detection methods are labour-intensive and slow, highlighting the need for more efficient solutions.

The proposed architecture combines object detection, segmentation, colour prediction, and length extraction using the Mask R-CNN, k-means clustering, and skeletonisation for precise microfibre identification. This approach powers MicroFiberDetect, a user-friendly desktop application designed for easy use by field experts, allowing fast and accurate analysis without technical expertise. We believe this scalable solution will aid efforts to quantify and analyse microplastic pollution, contributing to effective environmental strategies.

## Acknowledgements

This work was funded by the Norwegian Research Council grant 329745 Machine Teaching for Explainable AI, CIPROM/2022/6 (FASSLOW) funded by Generalitat Valenciana, the EC H2020-EU grant agreement No.952215 (TAILOR), and Spanish grant PID2021-122830OB-C42 (SFERA) funded by MCIN/AEI/10.13039/501100011033 and "ERDF A way of making Europe". Authors thank Grant PID2021-127468OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU. Authors also thank Business Chair FACSA-FOVASA from Universitat Politècnica de València.

## References

Ai, W.; Chen, G.; Yue, X.; and Wang, J. 2023. Application of hyperspectral and deep learning in farmland soil microplastic detection. *Journal of Hazardous Materials*, 445: 130568.

Andrady, A. L. 2011. Microplastics in the marine environment. *Marine Pollution Bulletin*, 62(8): 1596–1605.

Bholowalia, P.; and Kumar, A. 2014. EBK-means: A clustering technique based on elbow method and k-means in WSN. *International Journal of Computer Applications*, 105(9).

Bianco, V.; Memmolo, P.; Carcagnì, P.; Merola, F.; Paturzo, M.; Distante, C.; and Ferraro, P. 2020. Microplastic identification via holographic imaging and machine learning. *Advanced Intelligent Systems*, 2(2): 1900153.

Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; and Fei-Fei, L. 2009. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255.

Eze, C. G.; Nwankwo, C. E.; Dey, S.; Sundaramurthy, S.; and Okeke, E. S. 2024. Food chain microplastics contamination and impact on human health: a review. *Environmental Chemistry Letters*, 22(4): 1889–1927.

He, K.; Gkioxari, G.; Dollár, P.; and Girshick, R. 2018. Mask R-CNN. arXiv:1703.06870.

Ikotun, A. M.; Ezugwu, A. E.; Abualigah, L.; Abuhaija, B.; and Heming, J. 2023. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622: 178–210.

Löder, M. G. J.; and Gerdts, G. 2015. Methodology Used for the Detection and Identification of Microplastics—A Critical Appraisal. In *Marine Anthropogenic Litter*.

Lorenzo-Navarro, J.; Castrillón-Santana, M.; Sánchez-Nielsen, E.; Zarco, B.; Herrera, A.; Martínez, I.; and Gómez, M. 2021. Deep learning approach for automatic microplastics counting and classification. *Science of the Total Environment*, 765: 142728.

Prata, J. C.; da Costa, J. P.; Duarte, A. C.; and Rocha-Santos, T. 2019. Methods for sampling and detection of microplastics in water and sediment: A critical review. *TrAC Trends in Analytical Chemistry*, 110: 150–159.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, 234–241. Springer.

Shi, B.; Patel, M.; Yu, D.; Yan, J.; Li, Z.; Petriw, D.; Pruyn, T.; Smyth, K.; Passeport, E.; Miller, R. D.; et al. 2022. Automatic quantification and classification of microplastics in scanning electron micrographs via deep learning. *Science of The Total Environment*, 825: 153903.

Simonyan, K.; and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Wu, Y.; Kirillov, A.; Massa, F.; Lo, W.-Y.; and Girshick, R. 2019. Detectron2. https://github.com/facebookresearch/detectron2.

Zhang, E.; and Zhang, Y. 2009. *Average Precision*, 192–193. Boston, MA: Springer US. ISBN 978-0-387-39940-9.

Zhang, T. Y.; and Suen, C. Y. 1984. A fast parallel algorithm for thinning digital patterns. *Commun. ACM*, 27(3): 236–239.

Zhou, Y.; Li, Y.; Yan, Z.; Wang, H.; Chen, H.; Zhao, S.; Zhong, N.; Cheng, Y.; and Acharya, K. 2023. Microplastics discharged from urban drainage system: Prominent contribution of sewer overflow pollution. *Water Research*, 236: 119976.

Zhu, Y.; Yeung, C. H.; and Lam, E. Y. 2021. Microplastic pollution monitoring with holographic classification and deep learning. *Journal of Physics: Photonics*, 3(2): 024013.