MEDICAL WASTE SORTING: A COMPUTER VISION APPROACH FOR ASSISTED PRIMARY SORTING

A. Bruno[1], C.Caudai[1], G.R. Leone[1], M. Martinelli[1], D. Moroni[1], F. Crotti[2]

[1] Institute of Information Science and Technologies, National Research Council, Pisa, Italy
[2] CISA, Lucca, Italy

ABSTRACT

Medical waste, i.e. waste produced during medical activities in hospitals, clinics and laboratories, represents hazardous waste whose management involves special care and high costs. However, this kind of waste contains a significant fraction of highly valued materials that can enter a circular economy process. To this end, in this paper, we propose a computer vision approach for assisting in the primary sorting of medical waste. The feasibility of our approach is demonstrated on a representative dataset we collected and made available to the community.

Index Terms— Medical waste management, Computer vision, Image classification, Annotated dataset, Automated sorting

1. INTRODUCTION

Medical Waste (MW) is a growing concern as healthcare facilities produce a large amount of hazardous waste daily. Such waste requires special handling and treatment that imposes dedicated management approaches. Generally, medical waste is disposed into special containers in the place where it is produced. Such containers are then sealed and should be disposed of within a rather short timeframe. Disposal (mainly through incineration or waste-to-energy) should be done in special treatment facilities, which are often located far away from where the waste has been produced. In addition, after disposal in the container, manual manipulation is not permissible due to the possible biological hazard. Humans can intervene in waste only in critical situations and should be equipped with special personal protection equipment. According to various sources, the production of medical waste has been increasing in recent years due to factors such as the growth of the healthcare industry and the increasing use of disposable medical equipment [1]; also, the outbreak of COVID-19 – which has lead to high peaks in the production of possibly dangerous and infectious waste – urged the creation of more scalable and sustainable ways to treat medical waste [2]. It should be observed that medical devices are made in many situations with first-grade virgin materials, especially polymers; also, there are significant fractions of glasses, textiles and metals. Therefore, sorting medical waste for recycling is a way to enable more sustainable waste management. However, such an operation should be done as primary sorting since it is impossible to intervene manually at a later stage due to the aforementioned hazards. Indeed, secondary sorting would be possible only after medical waste's infectious and hazardous nature has been neutralized, e.g. by sterilization. Notice that the manual primary sorting of medical waste is a labour-intensive and time-consuming process that might excessively increment the burden of operators in healthcare facilities, with a high potential for human error. To this end, a system for assisting them during sorting, facilitating the procedure and reducing possible errors should be devised. In this paper, we propose research in this direction by introducing an artificial intelligence model based on deep learning and computer vision that can classify several classes of items that compose standard medical waste. Based on our previous experience [3, 4], we propose a convolutional neural network based on the EfficientNet family [5]. To train the network, a special dataset has been collected mimicking a waste collection table endowed with a stereo camera. The dataset is still growing but, based on the current data, we were able to train a network architecture achieving encouraging results. Differently from other studies on MW in the literature, the dataset is made publically available to foster further research [6]. The paper is organized as follows; in Section 2 we briefly survey related works, while in Section 3, we introduce the contributed datasets and the network architecture. Current results are reported in Section 4, while Section 5 ends the paper with ideas for future work.

2. RELATED WORKS

In recent years, computer vision technology has emerged as a promising solution to improve the accuracy and efficiency of medical waste sorting. Computer vision involves using computer algorithms to analyze and interpret images or videos. By applying computer vision techniques to the sorting of medical waste, it is possible to automate the process, reducing the risk of human error and increasing the speed

---

Correspondence should be addressed to: massimo.martinelli@isti.cnr.it

and accuracy of the sorting. The use of machine learning and, more recently, deep learning has been established as a promising solution for automating the medical waste sorting process. A seminal work related to general waste is represented by the Trashnet dataset [7]; this relatively simple but public dataset has triggered much research in image classification for waste sorting , becoming the *de facto* benchmark on the subject, used for testing Convolutional Neural Network (CNN) architectures [8] as well as the fusion of deep features [9]. Images, shot by using mobile devices, represent standard classes such as glass, paper, cardboard, plastic, metal and garbage.

Moving to the specific class of medical waste, the interest seems to be more recent. In [10], for instance, the authors used a two-step approach. First, they identify the object of interest in an image by resorting to standard image processing procedures based on background subtraction. Then, they use a CNN consisting of an optimized version of SSD-MobileNet [11] to identify medical waste. Their system is able to classify syringes, hemostatic forceps, infusion bags and gloves with a recognition accuracy of more than 98.5% and an average recognition speed of 52 milliseconds, making it suitable for real-time analysis and prompt feedback to operators.

In [12], the authors focus on the classification of medical waste with images acquired with complex backgrounds. To deal with this issue, they introduce a deep network with segmentation capabilities, named EnSegNet, and they concatenate a set of deep features with additional handcrafted features extracted from the segmented images (e.g. textural features based on Gray-Level Co-occurrence Matrix (GLCM), Multi-level Local Binary Pattern (MLBP), Local Derivative Pattern (LDP) and Local Ternary Pattern (LTP)). Concatenaed features are then classified by a fully connected layer. Their hybrid feature vector has a stronger discriminant ability compared to the single feature vector and the overall pipeline shows that the system attains a 93.7% of classification accuracy for 100 trash images. The dataset in any case seems to be small and not sufficiently representative. A more comprehensive work is presented in [13], where they applied a fine-tuned version of ResNext [14] on a dataset of 3480 images and succeeded in correctly identifying eight kinds of medical waste with an accuracy of 97.2%. Unfortunately, the dataset has not been disclosed.

After the COVID-19 outbreak, several implications for medical waste management emerged. For instance, possible infectious waste has been produced also in houses and in facilities where there is not usually such a production. This urged the creation of vision-based systems to detect different streams of waste. For instance, in [15], they proposed a waste type classification based on the image-texture-dependent features, which provided an accurate sorting and classification before the recycling process starts. The classification model based on the support vector machine (SVM) classifier performs best (with 96.5% accuracy, 95.3% sensitivity, and

95.9% specificity) in classifying waste types in the context of circular manufacturing and exhibiting the ability to manage COVID-related medical waste mixed.

In addition to the use of computer vision, some researchers have also explored the use of alternative technologies, such as infrared imaging and hyperspectral imaging [16] in waste sorting. Indeed, in such a way, it is possible to achieve a *chemical imaging* and sort the waste depending on its material composition. This might turn very useful in medical waste sorting, separating plastic materials made of different kinds of polymers. While these technologies have shown promising results, they also face challenges, such as the high equipment cost and the need for specialized training make them unsuitable at present, at least in primary sorting.

3. MATERIALS AND METHODS

Primary sorting of medical waste refers to the initial stage of the waste management process, where waste is separated into different categories based on its composition and properties. The primary sorting process is crucial for properly disposing of waste and reducing waste-related environmental and health hazards. To provide effective support in this process, we have schematised a use case in which operators convey waste to a special disposal station consisting of a flat surface where they deposit garbage. Such a surface is under active surveillance by an imaging system which is able to classify acquired images into a set of predefined classes and provide feedback to operators and/or send the waste to the right disposal/recycling pathway. Based on such a use case, we have collected a representative dataset (Section 3.1) and designed a CNN architecture (Section 3.2) for solving the computer vision task.

3.1. Dataset acquisition

Medical waste was collected using a lab table with an adjustable camera stand depicted in Figure 1. As a camera device, we opted for a high-resolution camera with depth vision and on-chip machine learning powered by an Intel Movidius-X chip. Specifically, we used a Luxonis OAK-D device [17], equipped with an Intel Movidius Myriad-X VPU, USB 3.1, gen2 10gps, 2 x global shutter mono cameras (1Mp - 1280 x 800px) Omnivision OV9282 for stereo vision and 1x centre RGB camera (12MP - 4056 x 3040px) Sony IMX378. Considering the standard input layer size of the state-of-the-art convolution neural network, the full resolution achievable by the OAK-D camera was not used. Instead, each sample is composed of a triplet of images with the following characteristics: 1 RGB image of size 1920 x 1080 and 2 grayscale of size 640 x 400, from which depth information can be inferred. All the images can be aligned using the factory calibration of the OAK-D to compute depth and produce RGBD images. The inclusion of 3D information might be useful for refined spatial AI algorithms and for estimating the volume of waste

at the moment of first disposal. In addition, in principle, some artificial intelligence algorithms can be executed directly on the OAK-D device, fully embracing an edge computing perspective. The acquisition setup was completed by an NVIDIA Jetson Xavier AGX connected to the OAK-D camera for executing CNN inference and/or sending data and results through the network as depicted in Figure 2. Variable natural and artificial illumination was used during the acquisitions which took place in two different rooms located at the Area delle Ricerca CNR in Pisa and CISA premises in Lucca, Italy.
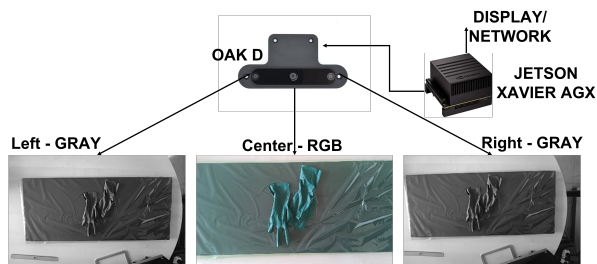

Fig. 1. Image acquisition setup.

Fig. 2. Overview of the acquisition using OAK-D camera [17].

Due to safety issues, medical waste was simulated using brand-new medical devices after a screening activity of the most common typologies used in hospitals. At current, a dataset for the following distinct classes has been collected: gauzes, latex gloves (single and pair), nitrile gloves (single and pair), surgery gloves (single and pair), medical caps, medical glasses, shoe covers (single and pair), test tubes, and urine bags. Sample images are reported in Figure 3. Over 1400 image triplets were already published in the public dataset [6], while a more structured dataset with over 2100 image triplets

will be made available in the near future.

3.2. Model definition and training

Since the image in the dataset have a standard background and most of the image area is occupied by an object of interest, the computer vision task to be solved has been represented as an image classification pipeline. In the future, a detector of the area of interest might be included to exclude areas irrelevant to the classification. Similarly, an object detection method might be used to replace the simple image classification, thus providing g also information on the location of the object of interest in the scene. Based on our previous experience [3, 4], it has been evaluated to assess the performance of convolutional neural networks belonging to the EfficientNet family [5]. As the name suggests, EfficientNet improves the classification quality with lower complexity compared to models having similar classification performances. This is possible since EfficientNet performs optimised network scaling, given a predefined complexity. Eight EfficientNet networks named b0-b7 are available with different levels of complexity. Since we aimed at a sustainable and on-the-edge analysis of the images in the place where primary sorting is done, the less complex network b0 has been considered, having only 5.5M parameters with respect to b7 coming with 66M parameters. For the time being, 3D information has not been included and some of the classes were merged for the classification task. Namely, the six classes in the dataset corresponding to gloves of different materials and as single items or in pairs were merged. Similarly, single items and pairs of shoe covers were merged, resulting in a total of 7 macro-classes. Therefore, the final fully connected layers of EfficientNet b0 were changed to accommodate the desired output size.

4. EXPERIMENTAL RESULTS

A version of EfficientNet b0 pretrained on the COCO Dataset has been used in the experiments adapting the final layer to fit the classification in the 7 classes described before. The initial dataset, named dataset $\mathcal{A}$, has been split into the train (4416 images), validation (551 images) and test (553 images) sets. Images were not preprocessed in any way; no augmentation has been used in the present paper. The training and validation losses and the relative accuracies plots (Figure 4) show that in 10 steps the maximum result is obtained, and an early stopping set to ten avoids overfitting. After training, the network was used to perform inference on the test set: no error in classification was observed achieving a 100% accuracy. An additional test was executed by acquiring new data and by building a new dataset named $\mathcal{B}$. The dataset $\mathcal{B}$ was acquired on different days from the dataset $\mathcal{A}$, in order to avoid any possible duplicated images between datasets $\mathcal{A}$ and $\mathcal{B}$. The results of inference on dataset $\mathcal{B}$ are reported in Table 1, showing an accuracy of 99.54%, demonstrating the
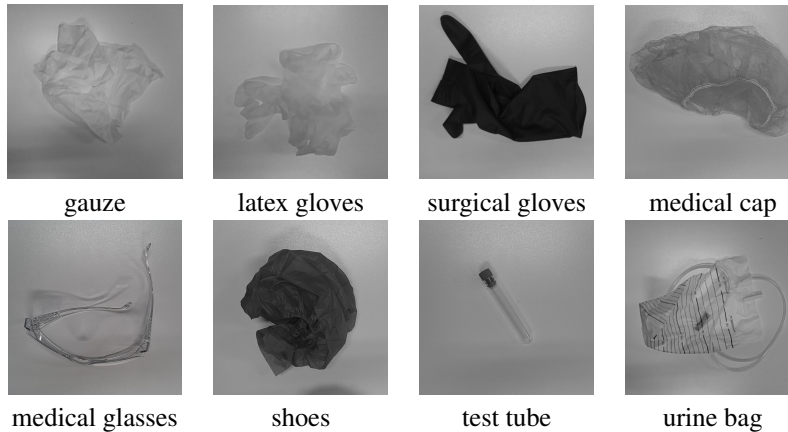
| gauze | latex gloves | surgical gloves | medical cap |
| medical glasses | shoes | test tube | urine bag |

Fig. 3. Examples from the classes of the dataset; for the class `gloves` we reported two different typologies.
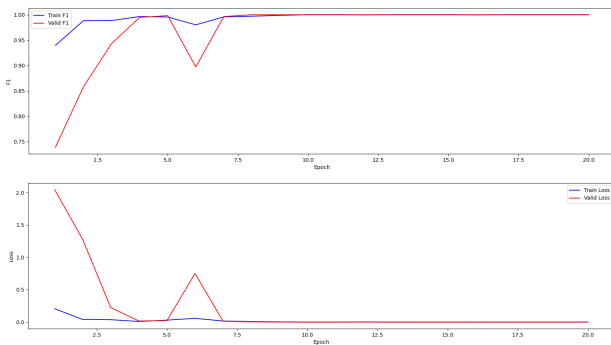

Fig. 4. Training and validation loss on top and training and validation accuracy curves at bottom for model training on dataset $\mathcal{A}$

generalisation capabilities of the trained model.

5. CONCLUSIONS AND FUTURE WORKS

In this paper, we have introduced the importance of primary sorting in medical waste management and contributed to the definition of approaches for assisting operators in such a task. To this end, we introduced a computer vision system based on deep learning. The approach was tested on datasets we collected ad hoc. In the future, we plan to extend the datasets so that they can serve as a valuable resource for the research community and provide a benchmark for evaluating medical waste sorting algorithms using computer vision. To make our dataset widely accessible, we have decided to release it and make it available for research purposes publicly. We hope this contribution will encourage further research in the field and help advance the development of medical waste sorting using computer vision.

Table 1. Performance of the model trained on dataset $\mathcal{A}$ during testing in the dataset $\mathcal{B}$.

Class	Precision	Recall	F1-Score	Support
Gauzes	0.98913	0.97849	0.98378	93
Gloves	0.99209	0.99603	0.99406	252
Med Cap	1.00000	1.00000	1.00000	57
Med Glasses	1.00000	1.00000	1.00000	93
Shoe Cover	0.99265	1.00000	0.99631	135
Test Tube	1.00000	1.00000	1.00000	54
Urine Bag	1.00000	0.97778	0.98876	45
Accuracy			0.99451	729
Macro avg	0.99627	0.99319	0.99470	729
Weighted avg	0.99452	0.99451	0.99450	729

Acknowledgement

6. REFERENCES

[1] YW Cheng, FC Sung, Y Yang, YH Lo, YT Chung, and K-C Li, "Medical waste production at hospitals and associated factors," *Waste Management*, vol. 29, no. 1, pp. 440–444, 2009.

[2] Wen-Tien Tsai, "Analysis of medical waste management and impact analysis of covid-19 on its generation in taiwan," *Waste Management & Research*, vol. 39, no. 1_suppl, pp. 27–33, 2021.

[3] Antonio Bruno, Davide Moroni, and Massimo Martinelli, "Efficient adaptive ensembling for image classification," *arXiv preprint arXiv:2206.07394*, 2022.

[4] A Bruno, M Martinelli, and D Moroni, "Exploring ensembling in deep learning," *Pattern Recognition and Image Analysis*, vol. 32, no. 3, pp. 519–521, 2022.

[5] Mingxing Tan and Quoc Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.

[6] Antonio Bruno, Davide Moroni, and Massimo Martinelli, "Medical Waste 4.0 Dataset," https://doi.org/10.5281/zenodo.7643417, 2022, Last reviewed March 6, 2023.

[7] Mindy Yang and Gary Thung, "Classification of trash for recyclability status," *CS229 project report*, vol. 2016, no. 1, pp. 3, 2016.

[8] Umut Ozkaya and Levent Seyfi, "Fine-tuning models comparisons on garbage classification for recyclability," *arXiv preprint arXiv:1908.04393*, 2019.

[9] Kashif Ahmad, Khalil Khan, and Ala Al-Fuqaha, "Intelligent fusion of deep features for improved waste classification," *IEEE access*, vol. 8, pp. 96495–96504, 2020.

[10] Xiaoxiao Bian, Yuchao Chen, Shuyi Wang, Fengyi Cheng, and Haotong Cao, "Medical waste classification system based on opencv and ssd-mobilenet for 5g," in *2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 2021, pp. 1–6.

[11] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg, "Ssd: Single shot multibox detector," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, 2016, pp. 21–37.

[12] T Mythili and A Anbarasi, "A concatenation of deep and texture features for medicinal trash image classification using ensegnet-dnn-based transfer learning," *Materials Today: Proceedings*, vol. 62, pp. 4691–4698, 2022.

[13] Haiying Zhou, Xiangyu Yu, Ahmad Alhaskawi, Yanzhao Dong, Zewei Wang, Qianjun Jin, Xianliang Hu, Zongyu Liu, Vishnu Goutham Kota, Mohamed Hasan Abdulla Hasan Abdulla, et al., "A deep learning approach for medical waste classification," *Scientific reports*, vol. 12, no. 1, pp. 2159, 2022.

[14] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1492–1500.

[15] Nallapaneni Manoj Kumar, Mazin Abed Mohammed, Karrar Hameed Abdulkareem, Robertas Damasevicius, Salama A Mostafa, Mashael S Maashi, and Shauhrat S Chopra, "Artificial intelligence-based solution for sorting covid related medical waste streams and supporting data-driven decisions for smart circular economy practice," *Process Safety and Environmental Protection*, vol. 152, pp. 482–494, 2021.

[16] Rosalba Calvini, Giorgia Orlandi, Giorgia Foca, Alessandro Ulrici, et al., "Developmentof a classification algorithm for efficient handling of multiple classes in sorting systems based on hyperspectral imaging," *Journal of Spectral Imaging*, vol. 7, pp. 1–15, 2018.

[17] Luxonis, "OAK-D," https://tinyurl.com/OAKDLUXO, 2022, Last reviewed March 6, 2023.