



HAL
open science

Towards Real-Time Image Mining for Waste Detection in Deserts

Souheir Mehanna, Nazim Berchiche, Assitan B Niare, Abderrahmane
Maaradji, Hakim Hacid, Assia Soukane

► **To cite this version:**

Souheir Mehanna, Nazim Berchiche, Assitan B Niare, Abderrahmane Maaradji, Hakim Hacid, et al..
Towards Real-Time Image Mining for Waste Detection in Deserts. 2023 Innovation and Technological
Advances for Sustainable Development (ITAS), Mar 2023, Doha, Qatar. hal-04158769

HAL Id: hal-04158769

<https://hal.science/hal-04158769v1>

Submitted on 11 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Towards Real-Time Image Mining for Waste Detection in Deserts

Souheir Mehanna
Université Paris-Saclay (UVSQ)
ECE Engineering School
Versailles, France
souheir.mehanna@uvsq.fr

Nazim Berchiche
ECE Engineering School
Paris, France
nazim.berchiche@edu.ece.fr

Assitan B. Niare
ECE Engineering School
Paris, France
assitanboubacar.niare@edu.ece.fr

Abderrahmane Maaradji*
ECE Engineering School
Paris, France
abderrahmane.maaradji@ece.fr

Hakim Hacid
Technology Innovation Institute
Abu Dabhi, UAE
hakim.hacid@tii.ae

Assia Soukane
ECE Engineering School
Paris, France
assia.soukane@ece.fr

Abstract—As the world’s population has grown and mobility has become more accessible, human activities have expanded into remote areas and deserts. These activities (typically desert safaris) generate much left-behind waste. Untreated left-out wastes contribute directly to gas emissions that exacerbate the global warming, especially litter such as plastics, cans, or any other resistant material. We present a drone-based solution to find waste in extra-urban environments (mainly deserts). In this paper, we use an autonomous unmanned aerial vehicle (drone) that flies over a desert area to capture a video stream to locate and identify waste objects such as plastic bottles. Our solution is still a work in progress; hence, in this paper, we have explored an object detection algorithm using transfer learning (few-shot-learning), where a pre-trained classifier (YOLO) is retrained with our relatively small dataset DroneTrashNet. Our goal is to benchmark existing models and approaches for object detection and classification in order to provide a baseline for selecting the most appropriate models for detection and classification tasks in extra-urban contexts. Our DroneTrashNet dataset corresponds to three minutes of annotated video captured by a drone at a low altitude in a controlled experiment in a desert environment, where waste objects were manually placed, with some objects half-buried in the sand. Preliminary results show an achieved mAP of 42% in specific conditions that could be subject to a number of improvements. The first results are promising compared to the results obtained with the initial classifier. While the resulting model has been validated in a controlled desert environment, new challenges and opportunities remain to be addressed, such as full-scale validation and generalization to extra-urban environments.

Index Terms—Desert Environment, Drone, Image Mining, Waste Detection

I. INTRODUCTION

Environmental issues and waste pollution have significantly developed over the years as global warming worsens. One of the causes of global warming is the gas emissions from undetected wastes left out in nature. Waste is an object or substance that has undergone physical or chemical changes at the end of its life and no longer has a destination or intended disposal. Today, protecting the environment is a priority, and managing waste in urban and extra-urban areas is a primary concern of communities. Several solutions tackling the issue have been implemented, ranging from collection and sorting combined in a circular economy. However, no systematic solution has been proposed for isolated extra-urban areas.

As the world’s population continues to grow and mobility becomes more accessible, people can easily access areas deep in nature or in the desert, resulting in more waste being littered in these areas. For example, the 2022 FIFA World Cup, hosted by Qatar, and the 2022 Saudi Games, hosted by Saudi Arabia, are attracting many tourists from around the world. As a result, the deserts in these two countries are subject to massive littering. Man-made waste causes significant environmental damage, especially in remote areas that are difficult to access.

This paper presents a solution based on drones to locate waste in these extra-urban environments. An autonomous drone flying over these spaces embeds a specially trained object classification algorithm to locate and identify waste objects such as plastic bottles. The tremendous volume of data generated from images and video footage captured by Unmanned

aerial vehicles (UAV)s requires a lot of time to process in order to produce insightful knowledge. Using the technique of *transfer learning*, a pre-trained classifier (*YOLO*) is re-trained with our dataset, which is made public. The first results show that the performances of the new re-trained classifier are superior to those of the initial classifier.

In the remaining of this paper, section II discusses existing works on object detection and image classification in different urban contexts. section III presents our current work and approach. section IV shows identified challenges and future perspectives. Preliminary results are then presented in section V. Finally, we conclude the paper in section VI.

II. RELATED WORKS

This section reviews several studies investigating image processing techniques for waste object detection and classification. These studies cover a wide range of topics, from convolutional neural networks and deep learning techniques to more traditional methods such as feature extraction and segmentation. They also explore different datasets and applications for waste object detection and examine ways to enhance the accuracy and efficiency of the methods. Overall, this research provides invaluable insight for information for further investigations in the field of waste object detection.

In [1], the authors propose a system that takes municipal solid waste (MSW) and then classifies plastic wastes into their appropriate subclass of plastics to help improve the effectiveness of the re-processing systems of waste materials. The authors tested two convolutional neural networks (CNN) for the classification with varying underlying architectures on 122×122 , and 277×277 input images. The first network consists of 23 layers, 9×9 convolution filters in the first convolution layer, six weave layers responsible for encoding the input image, and a three-layer fully connected layer. The second network consists of 15 layers, 64 convolution filters with 9×9 dimension in the first convolution layer, three layers of convolution to encode the input images, and a two-layer fully connected layer. The authors of [2] propose a waste detection system in urban cities with a camera embedded in a vehicle. The objectives of this work are first to localize all objects in an image frame and then to classify all detected objects into their correct litter categories. The authors used the Overfeat model [3] for the localization task where its classification layer is replaced by GoogLeNet [4] which will perform the classification task. The work of [5] presents a solution with a drone to

detect anthropogenic marine debris waste on a beach. To tackle this problem, the authors proposed combining an unmanned aerial vehicle (Drone) equipped with a high-definition RGB camera and software based on deep-learning convolutional neural networks. The software is a commercial software called PlasticFinder. According to the article, the algorithm achieved a sensitivity of 67%. However, this software is limited by climatic conditions (for example, on a sunny day, the sand shines too brightly, which prevents the detection), which may also be due to a camera that is not sufficiently developed for this style of use. To the best of our knowledge, this work was the first to confirm the effectiveness of drones for automatically detecting anthropogenic marine debris on the beach and is the closest study to our project.

You Only Look Once (YOLO) [6] is one of the most popular model architectures and object detection algorithms. It uses one of the best neural network architectures to produce high accuracy and overall processing speed, which is the main reason for its popularity. The YOLO algorithm aims to predict a class of an object and the bounding box that defines the location of the object on the input image. In addition to increased accuracy in predictions and better *Intersection on Union* in bounding boxes (compared to real-time object detectors), YOLO has the inherent advantage of speed. YOLO is a much faster algorithm than its counterparts, running at up to 45 FPS. The authors of [7] present a deep learning-based two-stage detector of waste localization and classification in natural and urban environments. The authors also benchmark existing waste littering open-source datasets with unified annotations across the datasets. The work of [8] proposes a construction waste semantic segmentation model in the context of automated waste composition recognition. Compared to image classification and object detection approaches, semantic segmentation works at a finer granularity in a pixel-wise manner. The authors adopted the DeepLabv3+ as the main architecture of the proposed construction waste segmentation model.

However, as investigated in a literature review proposed in [7], the current studies are hardly comparable to ours since they focus on waste objects in well-controlled environments (such as waste management factories). This case is different from our target use case where we randomize the scattering of waste objects in a real desert, and we emulate a real-life situation where the objects are partially buried in the sand. Also, none of the existing

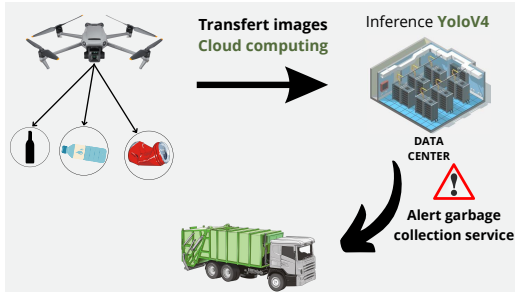


Fig. 1. Our general approach for waste detection and treatment.

works benchmark object detection and classification frameworks systematically and comprehensively for the waste pollution and littering context.

III. OUR APPROACH

This project aims to detect the presence of waste (bottles, glass, paper products, and food waste) using a drone equipped with a smart camera. The project aims to design an object detection algorithm embedded in a drone, specializing in detecting extra-urban waste. The real-time identification of litter is necessary to produce insights and statistics that will be sent to an operational center in order to decide on the intervention of the collection teams.

Figure 1 shows our general approach, which mainly consists of drones that will roam over the deserts to capture video footage of the desert from which we extract images covering the treatment area in the desert. And contrary to existing works that consider datasets taken in well-controlled environments where the images demonstrate a clear separation from the background and an angle facing the objects, our approach considers a practical, real-life situation where the images are taken from above using drones and in real-life circumstances where waste objects are actually in the desert and sometimes are partially covered with sand due to sandstorms. After the images are transferred to our data centers, our inference approach using YoloV4 is applied to localize and classify the waste objects to finally alert the local garbage collection entities of the computed insights about the existence of waste objects for their collection and treatment.

We started our studies with the pre-trained classifier (YOLO) [6]. Figure 2 shows an overview of our proposed approach. First, we train the pre-trained YoloV4 model on the DroneTrashNet dataset we have created, in order for the model to learn more about the different classes of waste in our context and to improve its detection accuracy on the specific images from the desert environment. The pre-

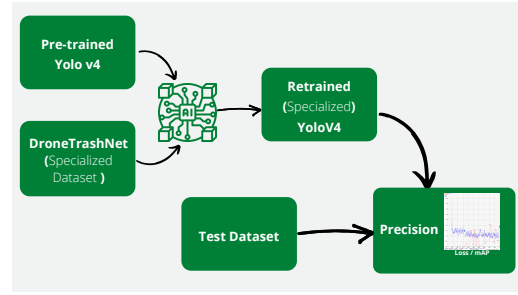


Fig. 2. Structure of our approach using YOLO v4 to detect waste objects taken in the desert by drones.

trained model is trained on a set of manually annotated images captured by our drone in the desert. This training of the model in different settings, and specifically to images from our context, significantly improved the model's detection accuracy.

IV. CHALLENGES AND FUTURE PERSPECTIVES

Many existing works address waste pollution and littering using computer vision frameworks to localize, detect, and learn the classes of the objects detected. The work we described in section III is only the beginning of our work on this subject, and it is still ongoing work that we are continuously working on enriching and improving. This section presents the research challenges we are currently addressing in our work and our future perspectives.

A. Benchmark the Existing Models

This is one of our primary goals in this work, as none of the existing works propose a benchmark that serves as a start point for an object detection and classification problem in the context of extra-urban environments and waste pollution. We are currently working on benchmarking existing algorithms and frameworks such as the Fast R-CNN (Detectron) by [9], the Single Shot multibox Detector (SSD) [10], the spatial pyramid pooling (SPP-net) [11], and others to have a solid basis for choosing the most suitable algorithm both to localize and classify the waste objects in real-life images and contexts. Also, it is in our future perspectives to test the Yolov6 algorithm with the best interference speed across the entire FPS spectrum. Currently, YOLOV4 is the model we trained and tested our data on. Our objective here is to explore the best-performing approaches of the proposed methods in terms of effectiveness, precision, and performance in order to provide a benchmark that allows readers to choose the most suitable object detection framework for their context.

B. Real-time Detection

One of our primary goals is to develop a real-time detection approach that allows garbage collection services to act as quickly as possible after images have been taken. It takes a lot of time to take the videos, unload the drone and pre-process the data into frames to eventually detect and classify objects if each step is done manually and separately. However, if the detection and classification are done directly on the frames right after the drone has taken them, the whole process becomes faster. Then we are able to scan more areas, and collection services can act more quickly and efficiently. There are several challenges in designing a real-time approach to drones, as it is known that drones have limited capacities in terms of storage and computation abilities. The objective here is to propose a model that is light enough to be embedded and deployed on the drone and can compute the localization and classification of objects live while it is roaming over the area fast and efficiently. Also, it is crucial that the model can report its findings and communicate them to the appropriate authorities for action in real time. This is still a challenge because of the heavy weight of the detection models and the communication tasks, and the limited storage and computation capacities of the drones.

C. Data Collection and Sharing

To ensure the reliability and credibility of our results, we intend to collect more data in the near future, such that we have a sufficient amount of data to train and test our model. At the moment, we have a small video with a limited number of frames taken in the desert at a fixed altitude. The goal is to have more frames and broader coverage of the possible scenarios of waste objects in the desert. We will also train our model on other waste pollution public datasets to avoid the overfitting problem. Moreover, we will collect data with drones at different altitudes, such as 5, 10, and 20, above the surface of the desert and at different speeds of the drone and angles. Currently, we only have the frames taking footage of the waste objects from one angle, the top.

And finally, for the post-processing algorithm perspectives. We plan to contact waste collection companies to create a partnership. Then develop a system for sending signals to the collection companies indicating the location of the waste.

D. Model Development

The objective is to elaborate new methods for supporting litter identification in wild areas. While

developing the models, we should consider that we deal with real-time and heterogeneous data taken under different constraints. We will build on top of the knowledge learned from benchmarking existing approaches and the real-life constraints we will have in the wild environments. Also, we will secure dedicated machines to train and test the benchmark models on. Moreover, our current experiments solely relied on the mean average precision mAP of the bounding box detecting the objects for evaluation, while in the future, we will also add another evaluation metric which is the accuracy of the classification of the detected objects to facilitate the recycling process for the garbage treatment services. This project extensively uses experiments as a mean of validation of the results and the findings.

V. PRELIMINARY EXPERIMENTS AND EVALUATIONS

We present in this section the preliminary experiments that we conducted for our approach. The results of the experiments we present here were performed manually and in a non-systematic manner because this is an still ongoing work. However, conducting more systematic experiments with more data and a systematic design of the settings of the experiments is in our next steps for the project.

A. Dataset

Our DroneTrashNet dataset that was used for post-training the YOLOv4 model corresponds to 3 minutes of video footage taken by a manually piloted drone at a low altitude in a controlled desert environment. In this controlled experiment, the waste items were manually and randomly dispersed around the study area in the desert as depicted in Figure 3. To address other real-life use cases and to partially emulate realistic conditions, a part of these wastes was placed in a way that only 50% of the waste appears to the roaming drone recording footage while the rest half is buried in the sand. The reason behind this procedure is that, in reality, the longer time passes on dumped garbage in the desert, the more buried in the sand they will be due to the wind factor, for instance.

We first extracted 200 frames (images) from the 3-minutes captured video. In order to improve the performance of the model, we augment the size of our dataset by applying a 90° rotation to these 200 frames. The resulting dataset contained 400 frames. We also manually annotated our dataset with 5 specific classes/labels and a generic one. The specific classes are categories of waste, they are glass, plastic, bottle, cardboard, and metal. And the

generic class is just waste. Our dataset is available on a public repository on GitHub ¹.

B. Experiments Settings

For the experiments, we are using python on Jupyter Notebook on Google Colaboratory which is a cloud service offered by Google for training and research in machine learning. This platform makes it possible to train Machine Learning models directly on the cloud. We re-trained YOLO on our augmented dataset in order to achieve decent results that are comparable to the results achieved in [6].

As discussed earlier, our dataset comprises 200 images that are augmented to 400 images after applying data augmentation. Some of the experiments are done on the full dataset and others on a subset of the data that comprises only the initial 200 images extracted from the video.

C. Evaluation Metrics

In this section, we present the evaluation metrics we are using to assess the performance of the approach on our data. Our approach detects and classifies wastes in the desert context. This means that we will need metrics to assess the appropriate trade-off between precision and recall. Hence, we chose the mean average precision (mAP), commonly used by object detection algorithms for the evaluations such as COCO ². The mean average precision (mAP) metric is used to analyze object detection. It compares the ground truth bounding box of detected objects to the actual detected box by our approach and returns a score. The higher the score, the more accurate the model is in detecting. The mAP is the mean of the average precision (AP); it is a trade-off between precision and recall. It is computed as follows: $mAP = \frac{1}{N} \sum_{i=1}^N AP_i$, where N is the number of classes of the classification. AP is the average precision of each class.

Average precision (AP) values are averaged over recall values from 0 to 1. The mAP formula is based on precision and recall which are on the measures from the confusion matrix: the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The precision is computed as: $Precision = \frac{TP}{TP+FP}$, and the recall is computed as: $Recall = \frac{TP}{TP+FN}$. An actual detected bounding box is assumed to be a true positive if the detected part of the frame is above a certain threshold called intersection on Union (IoU). Intersection on Union (IoU) indicates the overlap

of the coordinates of the predicted bounding box with the ground truth box. A higher IoU indicates that the predicted bounding box coordinates closely resembles the ground truth box coordinates.

D. Preliminary results

In this section, we report our preliminary results obtained from a limited evaluation as an exploratory step towards a more systematic and exhaustive evaluation in the future.

Before training the classifier on our dataset, we ran the pre-trained YOLO classifier during the exploratory process on our test dataset and obtained a *mAP* of about 13%. This first result indicates the irrelevance of the pre-trained YOLO classifier for the task at hand. This measure provides a baseline to evaluate the usefulness of our *few-shot-learning* approach.

We then retrained the YOLO classifier with our training dataset to test the retrained *specialized* YOLO classifier on our test data. This retraining operation was performed with different settings. At times we included the augmented data obtained from 90° rotation, and at other times we didn't; also, by varying the number of waste object classes. For example, the YOLO classifier was retained on the training dataset annotated with only two annotation labels, i.e., *Bottle* and *Waste* (for any other waste object). The resulting *specialized* YOLO classifier achieved a higher *mAP* of 42%. Figure 3 shows an example of waste detection using the specialized YOLO classifier that was re-trained on our data on a photo taken by a drone in our controlled stage.

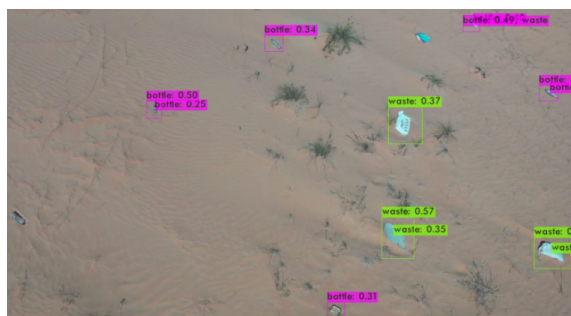


Fig. 3. An example of the inference of the specialized-YOLO-model on a drone image capture.

Despite being only a preliminary step, we believe these results are encouraging on several aspects. Not only this specialized classifier could learn only with relatively few annotated examples, but it could also keep the same precision order of magnitude when tested in the very specific use case it was retrained for. Indeed, in [6], the authors reported that YOLO

¹<https://github.com/ECE-StP/waste-desert-dataset>

²Details available at <https://cocodataset.org/#detection-eval>

classifier trained and tested on the COCO dataset scored a 44% of precision. This same classifier performed poorly when tested on our specific use case of waste detection in a desert environment. However, when it was first retrained with relatively few annotated images from our controlled stage, this same classifier performed almost as well as its original reported performance. These preliminary results could be further validated in a systematic and comprehensive evaluation.

VI. CONCLUSION AND DISCUSSION

In this paper, we presented our ongoing work and the results we have achieved so far. We also elaborated on some relevant approaches from the literature. We presented the current challenges we are facing and our future perspectives and goals for our project. One of our primary goals is to find the model and settings that work best in the waste pollution context. We made our data publicly available on Github³. We managed to identify and implement an object detection algorithm based on YOLOV4 for waste localization. We showed our preliminary experiments and results with 42% mAP, that is comparable to the results achieved by [6]; which proves that our approach can deliver promising results.

Currently, we have limitations such as the amount of data for the accuracy of the results because our experiments would be more reliable with a bigger dataset for training and testing. Also, in climatic conditions that are not favorable to the flight of the drone (heat), sandstorms can quickly hide the waste. The lack of networks and antennas for better control can affect the detection capacity of the drone. We also lack a setting that makes the detection and classification happen in real-time.

While our proposed solution might address the problem of waste management in desert areas and more generally in extra-urban areas, it is important to study the economical [12] and the carbon footprint of such a solution [13]. On the one hand, the economic value of a drone-based waste detection system would depend on the cost of its implementation compared to the potential cost savings and improved waste management it can provide. For example, by using drones to efficiently and effectively survey landfills and bodies of water for waste, the system can help reduce cleanup costs and prevent illegal dumping. On the other hand, it is also important to consider ways to minimize

the carbon footprint of this technology, such as by using energy-efficient hardware, implementing green energy in the data centers, and optimizing the frequency and duration of drone flights to minimize energy consumption.

REFERENCES

- [1] J. Bobulski and M. Kubanek, "Waste classification system using image processing and convolutional neural networks," in *International work-conference on artificial neural networks*. Springer, 2019, pp. 350–361.
- [2] M. S. Rad, A. v. Kaenel, A. Droux, F. Tietche, N. Ouerhani, H. K. Ekenel, and J.-P. Thiran, "A computer vision system to localize and classify wastes on the streets," in *International Conference on computer vision systems*. Springer, 2017, pp. 195–204.
- [3] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," in *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2014. [Online]. Available: <http://arxiv.org/abs/1312.6229>
- [4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," *CoRR*, vol. abs/1409.4842, 2014. [Online]. Available: <http://arxiv.org/abs/1409.4842>
- [5] L. Fallati, A. Polidori, C. Salvatore, L. Saponari, A. Savini, and P. Galli, "Anthropogenic marine debris assessment with unmanned aerial vehicle imagery and deep learning: A case study along the beaches of the republic of maldives," *Science of The Total Environment*, 2019.
- [6] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.
- [7] S. Majchrowska, A. Mikołajczyk, M. Ferlin, Z. Klawikowska, M. A. Plantykowski, A. Kwasigroch, and K. Majek, "Deep learning-based waste detection in natural and urban environments," *Waste Management*, vol. 138, pp. 274–284, 2022.
- [8] W. Lu, J. Chen, and F. Xue, "Using computer vision to recognize composition of construction waste mixtures: A semantic segmentation approach," *Resources, Conservation and Recycling*, vol. 178, 2022.
- [9] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448.
- [10] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *European conference on computer vision*. Springer, 2016, pp. 21–37.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, 2015.
- [12] H. Wilts, B. R. Garcia, R. G. Garlito, L. S. Gómez, and E. G. Prieto, "Artificial intelligence in the sorting of municipal waste as an enabler of the circular economy," *Resources*, vol. 10, no. 4, p. 28, 2021.
- [13] J. Busch, T. J. Foxon, and P. G. Taylor, "Designing industrial strategy for a low carbon transformation," *Environmental Innovation and Societal Transitions*, vol. 29, pp. 114–125, 2018.

³<https://github.com/ECE-StP/waste-desert-dataset>