

Development of a Customized Image Set for Object Detection in a Materials Synthesis Lab Using Convolutional Neural Networks

Alberto Esteban Reyes-Peralta*, David Eduardo Pinto-Avendaño, Francisco José López-Cortés, Rigoberto Cerino-Jiménez, Edgar Abidán Padilla-Luis, Darnes Vilariño-Ayala

Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación, Puebla,
Mexico

{alberto.reyesp, edgar.padilla}@alumno.buap.mx

Abstract. Convolutional Neural Networks (CNNs) have shown remarkable performance in object detection tasks, especially when trained on large and diverse datasets. However, in specialized domains such as materials synthesis laboratories, generic datasets may not capture the specific objects of interest or the unique challenges of the environment. This paper presents the development of a custom image dataset tailored for object detection in a materials synthesis laboratory. The dataset includes annotated images of equipment, chemicals, and other objects commonly found in such environments. We also describe the process of collecting and labeling the dataset, including the challenges faced and the strategies used to address them. To demonstrate the utility of the dataset, we trained a CNN model using the popular YOLO (You Only Look Once) architecture and evaluated its performance on a test set. The results show that our custom dataset enables the CNN model to accurately detect objects in materials synthesis laboratories, highlighting the importance of domain-specific datasets for enhancing the performance of object detection systems in specialized environments.

Keywords. Object detection, convolutional neural networks (CNNs), materials synthesis laboratory, custom image dataset.

1 Introduction

Computer vision focuses on training computers to acquire, process and analyze digital images to extract meaningful information and perform a wide range of tasks [11] such as image classification,

object recognition, object tracking, object location, image segmentation, image retrieval, pattern recognition to name a few.

To perform these tasks, various methods can be employed. Among them are convolutional neural networks (CNN), which have shown great success in object recognition. Their ability to capture relevant patterns and features at different spatial scales has been the key to their success and applicability in various areas [10].

In object recognition, one of the main challenges is the scarcity of labeled data. Obtaining large data sets can be costly and laborious. Also, objects can change in terms of size, shape, orientation, illumination, and background, which makes the task of recognition models difficult. In addition, models must be trained on data sets that are diverse and representative of real-world situations, which can be difficult to obtain.

Problems such as those described above can be encountered in a materials synthesis laboratory which has highly specialized equipment and materials, so the scarcity of labeled data limits the ability to train models to help within the laboratory. This paper presents the development of a set of images containing 7 particular objects which are: beaker, flask, vial, cuvette, sample holder, precision pipette and 3-mouth flask as seen in Figure 1.

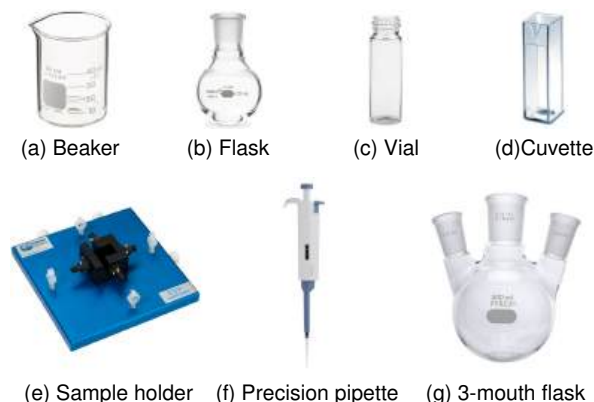


Fig. 1. Material synthesis laboratory instruments selected for the development of the customized image set



Fig. 2. Examples of images from the LabPics set

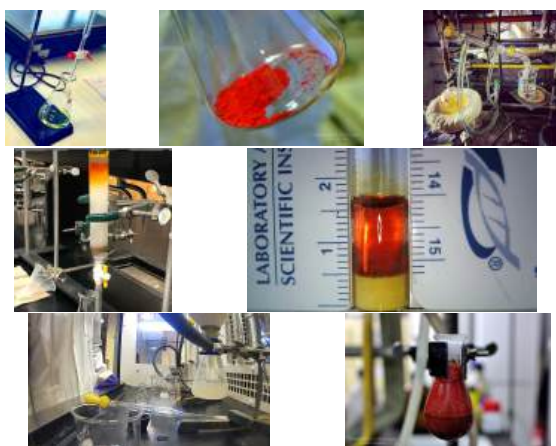


Fig. 3. Examples of images from the CLAD set

In addition, the usefulness of the image set is demonstrated by training a CNN with it, managing to identify the specialized objects.

2 Related Work

Datasets with large amounts of images that are commonly used for CNN training such as Common Objects in Context (COCO) [8] which contains 330 000 images with 80 object categories, another set used is Objects365 [9] which contains 365 object categories with more than 2 million images.

On the other hand Roboflow 100 [2] is a dataset that covers a wide range of domains and contains real world images representing everyday scenarios. ImageNet [3] is a massive dataset containing over 14 million carefully selected and annotated images, the images cover a wide range of categories. While these sets offer a wide range of categories, they do not specifically cover laboratory instruments used in materials synthesis. There are specialized databases in the literature that focus on these types of instruments.

“LabPics V1” contains 2187 images in 61 categories of chemical experiments with materials inside mostly transparent containers in various laboratory settings and under everyday conditions. Each image in the data set has an annotation of the region of each material phase and its type [5]. For its second version the “LabPics V2” dataset contains 10 528 annotated images.

The images are divided into two sets: a training set with 8 422 images and a test set with 2 106 images. The images in the “LabPics V2” dataset were obtained from a variety of sources, including images from real laboratories, images from laboratory simulations among others. [6], some examples of the images contained in the set are shown in the figure 2.

A more recent work is “Chemistry Laboratory Apparatus Dataset” (CLAD) which contains 21 different types of chemical laboratory instrument types, with no less than 200 images of each type, mainly glass instruments are carefully labeled with chemistry student information. Each image may contain one or more images of chemical instruments [4] as the images shown in the Figure 3.

Table 1. Type of images contained in each set as well as the number of categories

Laboratory equipment imaging datasets			
Name	Type of images	Number of images	Number of categories
COCO	Animals, people, vehicles, furniture, kitchen utensils, musical instruments, sports, cities, landscapes, fields, food, wild animals and domestic objects.	330000	80
Objects365	Real-world objects in various conditions of illumination, pose, size, shape, and background	2 millones	365
Roboflow 100	Images of airplanes, drones, cells, tissues, sea creatures, documents, radar fields, animals, plants, objects, people, places, events, transportation, art, fashion, food and text.	232000	828
Imagenet	Wide variety of images, including animals, plants, objects, people, places, events, transportation, art, fashion, food, text and much more.	14 million	22000
LabPics V1*	Materials and vessels in chemistry laboratories.	2187	61
LabPics V2*	Materials and vessels in chemistry and medical laboratories.	10528	61
CLAD*	Chemistry laboratory instruments.	2246	21

*These are sets specialized in laboratory objects

Table 2. Objects that are recognized by the sets of images found in the literature

Image set	Selected objects of a materials synthesis laboratory						
	Flask	Vial	Sample holder	Precision pipette	Cuvette	Beaker	3-mouth flask
COCO*	55	55	55	55	55	55	55
Objects365*	55	55	55	55	55	55	55
Roboflow 100	55	55	55	55	55	55	55
Imagenet	55	55	55	55	55	51	55
LabPics V2	51	51	55	55	55	51	55
CLAD	51	55	55	55	55	51	51

After reviewing the image sets found in the literature, it has been found that there is no set that includes the 7 selected objects from a materials synthesis laboratory as shown in the table 2. Therefore, the creation of an own image set becomes necessary.

3 Development

The elaboration of a set of customized images involves a series of steps ranging from image collection and labeling, to data augmentation and final evaluation of the set.

The first task performed was to capture color images of the selected objects using a 2D camera with a capture resolution of 1080×1920 pixels obtaining a total of 673 images. However, the set of images generated was very small, since only 673 images for CNN training do not ensure optimal results.

Therefore, we resorted to using data augmentation techniques, the study conducted by [7] shows the different data augmentation techniques and that the use of these techniques can improve the accuracy between 2.83% and 95.85%.

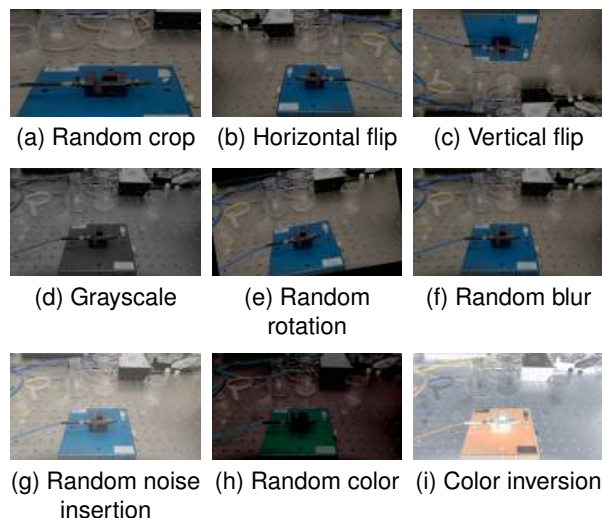


Fig. 4. Images generated using geometric and color space transformations from the original image

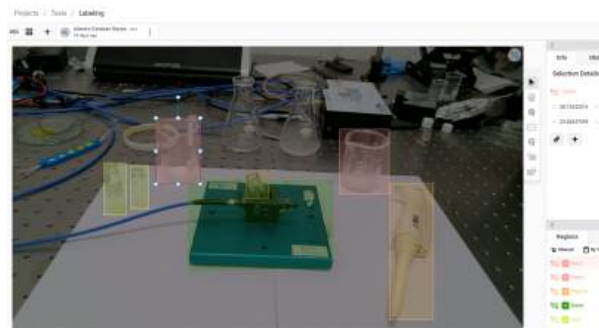


Fig. 5. Labeled with the software Label Studio

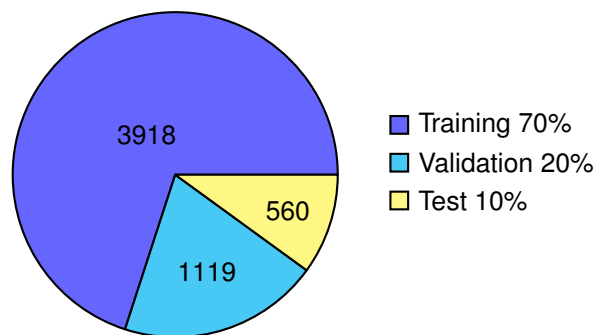


Fig. 6. Distribution of images for training

Of the techniques shown in this study we use geometric transformations and color space transformations.

These techniques were implemented in Python with the help of the albumentations library [1], the operations that were performed are random crop, horizontal and vertical flip, random rotation, grayscale, random blur, random noise insertion, random color and finally color inversion. With this implementation, 5597 images were obtained; Figure 4 shows the transformations used. Once the images were captured, they were manually labeled using the open source software Label Studio [13], a very helpful tool for labeling images in a simple way as shown in figure 5.

4 Experiments and Discussion

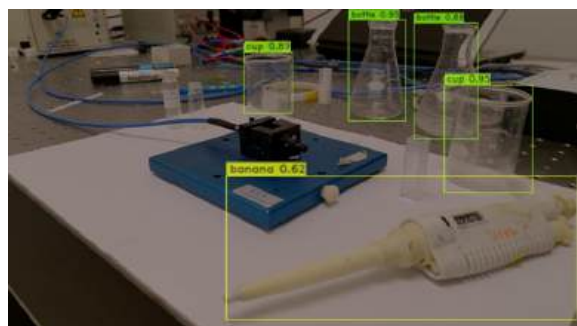
To demonstrate the usefulness of the set of images generated, the retraining process of an existing CNN was carried out in this case YOLO-NAS, which is a new YOLO model launched in 2023, YOLO-NAS developed by Deci manages to improve the speed and accuracy of the previous versions [12]. For training, the set was divided with the distribution that is normally used, 70% for training, 20% for validation and 10% for test.

The distribution of the images is shown in graph 6. Later, the obtained detection was compared between the model without the retraining process and the retrained model using our custom dataset. There was a noticeable difference, as seen in figure 7, where some predictions did not match the actual objects despite having a high confidence level. In contrast, our model correctly identifies the selected objects and achieves a mean Average Precision mAP@0.50 of 0.93.

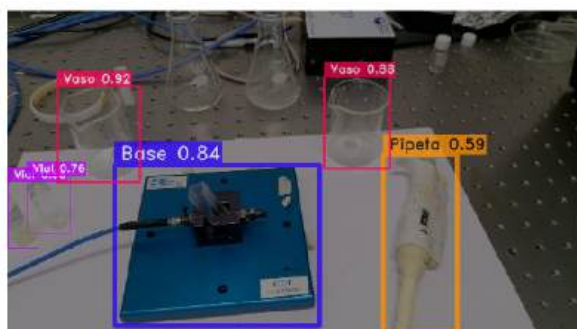
The confusion matrix shown in figure 8 reveals values greater than 0.75 on the diagonals, indicating high classification accuracy. Furthermore, the values of false positives (FP) and false negatives (FN) are very low, which confirms the reliability of our model.

5 Conclusions and Future Work

In this work, a set of images was generated that allows a CNN to identify particular objects such as chemical instruments used in materials synthesis laboratory environments.



(a)



(b)

Fig. 7. Predictions using the model without retraining (a) and using the retrained model (b)

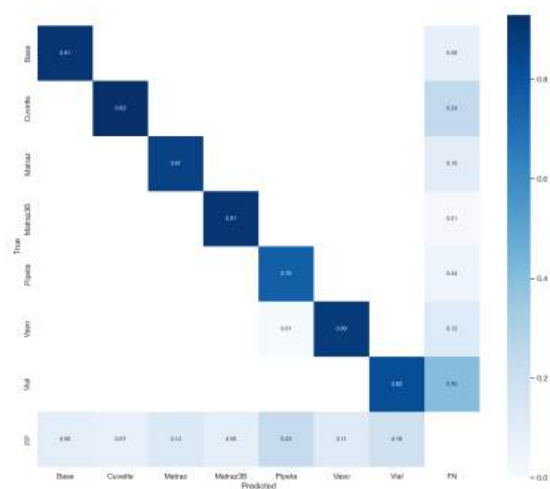


Fig. 8. Confusion matrix obtained from the evaluation of the model generated with the customized data set

Initially, 673 images were obtained, but with the implementation of data augmentation techniques, 5597 were obtained.

Images, thus managing to adequately retrain a CNN and obtaining a mean Average Precision mAP@0.50 of 0.93. The importance of creating custom image sets for application to specific problems should be highlighted. CNNs trained with custom image sets perform better than CNNs trained with generic image sets. This advantage is due to the greater precision in the identification of specific instruments, avoiding confusion such as identifying a beaker and a flask as a “bottle” using generic sets of images.

A methodology has been proposed for creating customized image sets that adapt to the specific needs of the problem. As future work, it is proposed to create an identification system with the generated set and test it in real environments. It is worth mentioning that this set will be a fundamental piece if at some point it is necessary to automate processes within a materials synthesis laboratory, for example the spectroscopy process, since to achieve this these objects must first be identified.

References

1. Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A. A. (2020). Albumentations: Fast and flexible image augmentations. *Information*, Vol. 11, No. 2, pp. 125. DOI: 10.3390/info11020125.
2. Ciaglia, F., Zuppichini, F. S., Guerrie, P., McQuade, M., Solawetz, J. (2022). Roboflow 100: A rich, multi-domain object detection benchmark. DOI: 10.48550/ARXIV.2211.13523.
3. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., Li, F. F. (2009). ImageNet: A large-scale hierarchical image database. pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.
4. Ding, Z. (2023). Chemistry laboratory apparatus dataset.
5. Eppel, S., Xu, H., Bismuth, M., Aspuru-Guzik, A. (2020). Vector-LabPics dataset for images of materials in vessels in the chemistry lab. DOI: 10.5281/zenodo.3697452.

6. **Eppel, S., Xu, H., Bismuth, M., Guzik-Aspuru, A. (2021).** LabPics dataset for computer vision for autonomous chemistry labs and medical labs. Zenodo.
7. **Khosla, C., Saini, B. S. (2020).** Enhancing performance of deep learning models with different data augmentation techniques: A survey. *Proceedings of the International Conference on Intelligent Engineering and Management*, pp. 79–85. DOI: 10.1109/ICIE M48762.2020.9160048.
8. **Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C. L. (2014).** Microsoft COCO: Common objects in context. *Computer Vision – European Conference on Computer Vision*, pp. 740–755. DOI: 10.1007/978-3-319-10602-1_48.
9. **Shao, S., Li, Z., Zhang, T., Peng, C., Yu, G., Zhang, X., Li, J., Sun, J. (2019).** Objects365: A large-scale, high-quality dataset for object detection. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 8429–8438. DOI: 10.1109/ICCV.2019.00852.
10. **Shelhamer, E., Long, J., Darrell, T. (2017).** Fully convolutional networks for semantic segmentation. *Vol. 39, No. 4*, pp. 640–651. DOI: 10.1109/TPAMI.2016.2572683.
11. **Szeliski, R. (2022).** *Computer vision: Algorithms and applications*. Springer International Publishing. DOI: 10.1007/978-3-030-34372-9.
12. **Terven, J., Córdova-Esparza, D. M., Romero-González, J. A. (2023).** A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Machine Learning and Knowledge Extraction*, Vol. 5, No. 4, pp. 1680–1716. DOI: 10.3390/make5040083.
13. **Tkachenko, M., Malyuk, M., Holmanyuk, A., Liubimov, N. (2020-2022).** Label Studio: Data labeling software. github.com/heartexlabs/label-studio.

Article received on 12/04/2024; accepted on 19/06/2024.

**Corresponding author is Alberto Esteban Reyes-Peralta.*