Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/33/20230246

Harnessing the power of AutoML: A comparative study of image recognition techniques for smoking detection

Anbang Wang[1,2] and Jianing Yu[1,3]

[1]School of Software, Shandong University, Jinan 250101, China

[2]202000300122@mail.sdu.edu.cn
[3]1508545752@qq.com

Abstract. As infrastructure continues to evolve, the significance of fire protection escalates. Many fires are caused by smoking in smoke-free areas, underscoring the necessity to promptly detect smoking activities in hazardous zones. In this scenario, image recognition emerges as a pivotal tool. The accuracy and efficiency of image recognition bear substantial implications for both academic and industrial sectors, and these aspects form the crux of our investigation. This study aims to compare the performance of image recognition techniques based on automatic machine learning with those of traditional methods such as YOLO. Our findings indicate that image recognition powered by automatic machine learning outperforms YOLO recognition in terms of efficiency and accuracy.

Keywords: machine learning, image recognition, yolo, smoking detection.

1. Introduction

The advent of the digital age has brought forth myriad advancements in various fields, one of which is the realm of image recognition. This technology has applications in diverse sectors, from academic research to industrial operations, and its significance cannot be overstated. However, the effectiveness of image recognition hinges on two key factors: accuracy and efficiency. These aspects form the cornerstone of our investigation.

In the fire protection context, image recognition's importance is further amplified. With the continuous development of infrastructure, fire safety has become a paramount concern. Many fires are ignited due to smoking in non-smoking areas, making it crucial to promptly detect such activities in hazardous zones. This is where image recognition comes into play. By accurately identifying smoking activities in real-time, potential fire incidents can be averted, thereby enhancing overall safety.

However, the question arises: what is the most effective method for image recognition? Traditional techniques have been employed for years, but a new contender has entered the arena with the advent of machine learning. With its ability to learn and improve over time, automatic machine learning presents a promising alternative to traditional methods.

This study aims to delve into this question by comparing the performance of image recognition techniques based on automatic machine learning with those of traditional methods. We employ image-processing techniques rooted in automatic machine learning to process datasets and evaluate various performance metrics and parameters. The performance of these techniques is then juxtaposed with the image recognition performance of You Only Live Once (YOLO), a widely-used traditional method.

Our research is not merely an academic exercise; it has far-reaching implications for the real world. If automatic machine learning proves to be more efficient and accurate than traditional methods, it could revolutionize the field of image recognition, leading to more effective fire protection measures and, ultimately, safer infrastructures.

In the following sections, we will present our methodology, discuss our findings, and explore the implications of our research. We hope our work will contribute to the ongoing discourse on image recognition and its applications in fire safety.

2. Literature review

2.1. Automated machine learning

Automated Machine Learning, often abbreviated as AutoML, is a relatively recent development in artificial intelligence. It aims to automate the typically complex and time-consuming process of machine learning model selection, hyperparameter tuning, and iterative modeling [1]. The concept of AutoML was first introduced by Thornton et al. Thornton, et al. [2] 2013 proposed an algorithm that could automatically select the best machine learning algorithm and its hyperparameters for a given dataset. This marked a significant shift in the machine learning paradigm, reducing the need for extensive domain knowledge and manual intervention.

Since then, numerous studies have been conducted to refine and expand upon this concept. For instance, Feurer, et al. [3] developed Auto-sklearn, an AutoML system based on the popular machine learning library, scikit-learn. Auto-sklearn uses Bayesian optimization, meta-learning, and ensemble construction to automatically select and tune the best machine-learning pipeline for a given task. In deep learning, AutoML has been used to automate the design of neural network architectures, a process known as Neural Architecture Search (NAS). Zoph and Le [4] presented a reinforcement learning-based approach to NAS, which was able to design network architectures that outperformed manually designed architectures on image recognition tasks.

AutoML has also found applications in image recognition. For instance, Real et al. used an evolutionary algorithm for NAS to design a convolutional neural network for image classification that achieved state-of-the-art performance on the CIFAR-10 dataset [5]. In the context of fire safety and smoking detection, however, the application of AutoML is still largely unexplored. Our study aims to address this gap by investigating image recognition performance based on AutoML compared to traditional methods.

2.2. Image recognition in smoking detection.

The recent advancements in image recognition for smoke and fire detection have been marked by innovative approaches and diverse applications. Xu, et al. [6] tackled the challenge of limited smoke image samples by introducing a deep domain adaptation method, generating synthetic smoke images and confusing feature distributions to improve recognition rates. Zell, et al. [7] took a step further in industrial fire safety by employing YOLOv4 object detection, significantly outperforming traditional smoke detectors in time efficiency. In the public health domain, Hellen and Marvin [8] developed an interpretable feature learning framework using VGG-16 and Layer-wise Relevance Propagation, capable of detecting smoking behavior and other smokable drugs. Xu, et al. [9] contributed to video smoke detection by proposing a method based on a deep saliency network, combining pixel-level and object-level salient convolutional neural networks to predict smoke existence. Lastly, Ayala, et al. [10] focused on portability and computational efficiency, creating a deep learning model for fire recognition that requires fewer floating-point operations.

These studies collectively represent a significant stride in the field of image recognition for smoke and fire detection. By leveraging deep learning, object detection, visual saliency, and computational efficiency, researchers are not only enhancing detection accuracy but also expanding the applicability of these technologies. Whether it's industrial safety, public health monitoring, or mobile device implementation, the innovations in this field are paving the way for more responsive and effective smoke

DOI: 10.54254/2755-2721/33/20230246

and fire detection systems. The integration of these technologies promises to contribute substantially to early detection, timely response, and overall public safety.

3. Research methods

3.1 EasyDL as AutoML tools

We utilized EasyDL, an automatic machine learning platform, in this study, for our image processing techniques. EasyDL is a user-friendly and efficient tool that allows users to build custom machine learning models without requiring extensive programming knowledge.

EasyDL is based on deep learning, and it excels in tasks related to image recognition. It provides a variety of pre-trained models, including Convolutional Neural Networks (CNNs), which have shown significant success in image recognition tasks. CNNs are designed to automatically and adaptively learn spatial hierarchies of features from the input images.

For our study, we used EasyDL to train a custom model specifically designed to recognize smoking activities in various settings. We selected the appropriate pre-trained model provided by EasyDL and fine-tuned it using our prepared datasets. This allowed the model to learn the specific patterns and features associated with smoking activities.

EasyDL's intuitive interface and powerful capabilities made it an ideal choice for our study. Its ability to handle large datasets and complex image recognition tasks allowed us to focus on analyzing and interpreting our results rather than the technical details of implementing and training the machine learning models.

3.2 Datasets

The dataset contains a total of 2400 raw images, where 1200 images are of smoking (smokers) category and remaining 1200 images belong to no-smoking (non-smokers) category [11]. These images were sourced from public databases and carefully selected to represent various scenarios, including indoor and outdoor environments, different lighting conditions, and various angles and distances. The images were then preprocessed to ensure optimal input for the machine learning models. This preprocessing involved resizing the images to a standard size, normalizing the pixel values, and augmenting the data to increase its size and variability. The images were then labeled and annotated to facilitate the training of the machine learning models.

3.3 Performance metrics

The performance of the automatic machine learning techniques and the YOLO method was evaluated using a comprehensive set of metrics. These metrics were chosen to provide a holistic view of the performance of the models, taking into account various aspects of their predictions. The following metrics were used:

Accuracy: This is the proportion of total correct predictions. It is calculated as (TP+TN)/(TP+FP+FN+TN), where TP is true positives, TN is true negatives, FP is false positives, and FN is false negatives. Accuracy is a fundamental metric in classification tasks, providing a straightforward measure of the model's overall performance.

Precision: This is the proportion of correct positive predictions. It is calculated as TP/(TP+FP). Precision is particularly important in situations where false positives are costly or undesirable. In the context of our study, high precision means that when our model predicts a smoking activity, it is highly likely to be correct.

Recall: This is the proportion of actual positives that were correctly identified. It is calculated as TP/(TP+FN). The Recall is crucial when missing a positive instance is particularly harmful. In our case, a high recall means that our model can identify most smoking activities.

F1 Score: This is the harmonic mean of precision and Recall. It provides a balanced measure of their performance and is particularly useful when the class distribution is uneven. It is calculated as

2*(Precision*Recall)/(Precision+Recall). The F1 score combines the precision and Recall of the model, and it will be high only if both precision and Recall are high.

4. Experimental result

Table 1. Experience results.

Tools	mAP	Precision	Recall
EasyDL	84.40%	85.70%	70.60%
YOLOv5[12]	83.34%	83.72%	80.21%
Efficientdet-d2[12]	81.13%	81.78%	79.31%

In our comparative study of image recognition techniques, as shown in Table 1, we evaluated three different tools: Baidu's EasyDL, YOLOv5, and EfficientDet-d2. The performance of these tools was assessed using three key metrics: mean Average Precision (mAP), Precision, and Recall.

The mAP is a comprehensive metric that considers both precision and recall across different thresholds. EasyDL achieved the highest mAP of 84.40%, followed closely by YOLOv5 at 83.34% and EfficientDet-d2 at 81.13%. This indicates that EasyDL was the most effective tool in our study regarding overall performance.

Precision measures the proportion of correct positive predictions. EasyDL also led in this metric, achieving a precision of 85.70%. YOLOv5 and EfficientDet-d2 followed with precisions of 83.72% and 81.78%, respectively. The higher precision of EasyDL suggests that it was more reliable in correctly identifying smoking activities.

Recall measures the proportion of actual positives that were correctly identified. In this metric, YOLOv5 achieved the highest Recall of 80.21%, followed by EfficientDet-d2 at 79.31% and EasyDL at 70.60%. This suggests that YOLOv5 was more effective in identifying most smoking activities, even though EasyDL had higher precision and mAP.

The results of our study demonstrate the effectiveness of EasyDL in terms of mAP and precision, indicating its potential as a powerful tool for image recognition in fire safety applications. However, the lower Recall of EasyDL compared to YOLOv5 and EfficientDet-d2 highlights potential improvement areas. The close performance of YOLOv5 and EfficientDet-d2 in terms of Recall, and their competitive performance in mAP and precision, also provide valuable insights into the capabilities of these traditional methods. These findings contribute to our understanding of the strengths and weaknesses of different image recognition techniques and provide a foundation for further research and development in this field.

5. Discussion

The comparative study of image recognition techniques involving Baidu's EasyDL, YOLOv5, and EfficientDet-d2 yielded insightful findings. EasyDL demonstrated superior performance in terms of mean Average Precision (mAP) at 84.40% and Precision at 85.70%. These metrics indicate that EasyDL was highly effective in overall performance and in correctly identifying positive instances of smoking activities.

However, the Recall metric revealed a different aspect of performance. YOLOv5 achieved the highest Recall of 80.21%, followed by EfficientDet-d2 at 79.31% and EasyDL at 70.60%. This suggests that while EasyDL was more precise, YOLOv5 was more effective in identifying most smoking activities.

These results highlight the nuanced performance characteristics of different image recognition techniques. While EasyDL excelled in overall accuracy and precision, its lower Recall compared to traditional methods like YOLOv5 indicates an area for potential improvement. The close performance of YOLOv5 and EfficientDet-d2 in terms of Recall, and their competitive performance in mAP and precision, underscores the continued relevance and effectiveness of these traditional methods.

6. Conclusion

This study aimed to compare the performance of image recognition techniques, specifically focusing on Baidu's EasyDL, YOLOv5, and EfficientDet-d2 in the context of fire safety. The findings demonstrate the strengths and weaknesses of these tools in different aspects of performance.

EasyDL's superior mAP and precision suggest its potential as a powerful tool for image recognition in fire safety applications. Its ability to accurately identify smoking activities could play a vital role in preventing potential fire incidents and enhancing overall safety. However, the lower Recall of EasyDL compared to traditional methods highlights the complexity of image recognition tasks and the need for a balanced approach that considers various performance metrics.

The study contributes valuable insights to the field of image recognition, emphasizing the importance of a comprehensive evaluation that considers different aspects of performance. It also opens up new avenues for future research, technological development, and practical applications in fire safety and beyond.

The potential of tools like EasyDL to revolutionize the field of image recognition is an exciting prospect. However, the nuanced performance characteristics revealed in this study remind us that continued research, innovation, and critical evaluation are essential to fully realize this potential and fully develop effective and efficient solutions.

References

[1] J. Waring, C. Lindvall, and R. Umeton, "Automated machine learning: Review of the state-of-the-art and opportunities for healthcare," *Artificial intelligence in medicine,* vol. 104, p. 101822, 2020.

[2] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2013, pp. 847-855.

[3] M. Feurer, A. Klein, K. Eggensperger, J. Springenberg, M. Blum, and F. Hutter, "Efficient and robust automated machine learning," *Advances in neural information processing systems,* vol. 28, 2015.

[4] B. Zoph and Q. V. Le, "Neural architecture search with reinforcement learning," *arXiv preprint arXiv:1611.01578,* 2016.

[5] E. Real *et al.*, "Large-scale evolution of image classifiers," in *International conference on machine learning*, 2017: PMLR, pp. 2902-2911.

[6] G. Xu, Y. Zhang, Q. Zhang, G. Lin, and J. Wang, "Deep domain adaptation based video smoke detection using synthetic smoke images," *Fire safety journal*, vol. 93, pp. 53-59, 2017.

[7] O. Zell, J. Pålsson, K. Hernandez-Diaz, F. Alonso-Fernandez, and F. Nilsson, "Image-Based Fire Detection in Industrial Environments with YOLOv4," *arXiv preprint arXiv:2212.04786,* 2022.

[8] N. Hellen and G. Marvin, "Interpretable feature learning framework for smoking behavior detection," *arXiv preprint arXiv:2112.08178,* 2021.

[9] G. Xu *et al.*, "Video smoke detection based on deep saliency network," *Fire Safety Journal,* vol. 105, pp. 277-285, 2019.

[10] A. Ayala, B. Fernandes, F. Cruz, D. Macêdo, A. L. Oliveira, and C. Zanchettin, "KutralNet: A portable deep learning model for fire recognition," in *2020 International Joint Conference on Neural Networks (IJCNN),* 2020: IEEE, pp. 1-8.

[11] A. Khan, "Dataset containing smoking and not-smoking images (smoker vs non-smoker)," *Mendeley Data*, vol. 1, 2020.

[12] H. Yin, M. Chen, W. Fan, Y. Jin, S. G. Hassan, and S. Liu, "Efficient Smoke Detection Based on YOLO v5s," *Mathematics,* vol. 10, no. 19, p. 3493, 2022.