**Exploring Task Independence: Active Learning Approaches in Object Detection vs. Image Recognition**

**By Sanjana Kulkarni &** **Dr. Shabbeer Basha S. H.**

**Abstract—Computer vision has made significant advancements, particularly in the fields of object detection and image recognition. While these two tasks are interrelated, they each present distinct challenges that can be addressed through active learning. This paper explores the potential for task independence between object detection and image recognition, focusing on the unique opportunities and challenges each task offers. We examine how active learning strategies differ for both tasks and discuss the specific challenges associated with their implementation. Additionally, we highlight the benefits of independent optimization, including reduced annotation needs, improved performance, and the development of specialized architectures for more efficient systems. Understanding task independence is critical for advancing computer vision technologies and optimizing training processes for real-world applications.**

**Keywords—Object Detection, Image Recognition, Active Learning, Task Independence, Computer Vision**

## I. INTRODUCTION

Computer vision continues to advance rapidly, with significant developments in both object detection and image recognition systems. These two distinct tasks, while interconnected, present unique challenges and opportunities for optimization through active learning approaches. The relationship between these tasks and their potential for independent operation represents a crucial area of research in modern computer vision.

Object detection and image recognition serve different purposes within computer vision applications. Object detection combines localization and classification of multiple instances within images, whereas image recognition focuses on overall image classification. Active learning, which enables models to select informative samples for labeling, demonstrates varying effectiveness across these tasks.

Several key research questions emerge when examining task independence in this context:

1. The differentiation of active learning strategies between object detection and recognition

2. The specific challenges presented by each task during implementation

3. The potential benefits of independent optimization through active learning approaches

The significance of understanding task independence extends beyond theoretical interest. Practical applications benefit from optimized training processes, reduced annotation requirements, and enhanced performance metrics. Furthermore, insights into task independence contribute to the development of specialized computer vision architectures, leading to more efficient and effective systems.

## II. MAIN BODY

Recent research reveals several significant patterns and developments in the relationship between object detection, recognition, and active learning approaches. The literature demonstrates a clear progression in understanding task independence while highlighting persistent challenges and opportunities for optimization.

*A. Active Learning Strategies and Task Independence*

Ren et al. (2020) categorize the main types of active learning into uncertainty-based learning and diversity-based learning. Basically, uncertainty-based methods are effective in minimizing labels, while diversity-based methods provide comprehensive coverage, crucial for maintaining task independence. Hekimoglu et al. (2021) further developed this idea with multi-task networks, where they studied how to maintain task consistency and diversity when selecting data. Senzaki and Hamelain (2021) contributed by studying core sentence selection **strategies** that help maintain task independence even when resources are scarce. So this is not just theory, but something that can actually work in real-world conditions.

*B. Uncertainty Estimation in Multi-Task Learning*

Uncertainty estimation is very important in active multi-task learning because it helps set priorities for data labeling. Hekimoglu et al. (2022) tried to use task-specific uncertainty metrics to make labeling across tasks more efficient. Their work showed that using these metrics can reduce the need for labeling while keeping performance stable across tasks. Xiao et al. (2021) also developed an effective method to use uncertainty to maintain task independence, which is particularly useful when computing power is low.

*C. Architectural Choices and Performance Trade-offs*

Different architectures in multi-task learning give you trade-offs between speed and accuracy. Zhao et al. (2019) break it down like this: YOLO is super fast but less accurate, Fast R-CNN gives higher accuracy but is slower, and SSD sits somewhere in the middle. Hekimoglu et al. (2021) and Sun et al. (2024) showed that you can reach a solid balance with new architectures—Sun’s model hit 79.8% mean Average Precision at 89 FPS. Rotman and Reichart (2021) used transformers to boost task consistency, getting solid performance without needing a ton of labelled data.

*D. Sample Selection and Core-Sets*

Sample selection is a key focus of active multi-task learning, as it helps get the most value from each labelled data point. Hekimoglu et al. (2021) use core sets to keep the data representative but efficient for a multi-task setting, and Senzaki and Hamelain (2021) show how core sets work with limited computational resources. Rotman and Reichart (2021) rely on pre-trained models to make sample selection more efficient — this is important for high-dimensional data tasks such as natural language processing, where labelling is expensive.

*E. Domain Adaptation and Cross-Task Learning*

Adapting to different domains while keeping tasks independent is tough. Xie et al. (2022) tackled this with energy-based sample selection, which improved cross-domain performance without a ton of extra data. Xiao et al. (2021) took a similar approach with an efficient method for multi-task learning, making it easier to move between tasks without losing independence. This is especially useful for applications where data is limited or where adapting across tasks quickly is crucial.

*F. Real-World Applications*

The theory is cool, but what about real-world use? Liu et al. (2023) reported 90% accuracy in specific applications with task-independent systems, which is solid evidence that this stuff actually works outside the lab. Schilling et al. (2023) went further by automating complex tasks in medical imaging using these task-independent frameworks. Hekimoglu et al. (2021) and Xiao et al. (2021) also tested their methods in real scenarios, showing how scalable multi-task active learning can be.

*G. Evaluating Task Independence*

When it comes to measuring task independence, researchers use different metrics to keep things consistent across tasks. Hekimoglu et al. (2022) suggest using task-specific accuracy and mean Average Precision (mAP), and Xiao et al. (2021) added metrics for processing time and resource use to measure efficiency. These give a full picture of performance, but as more research rolls in, we’ll probably need a set of standardized metrics to really compare results across different studies.

*H. Importance of Task Independence in active learning*

Task independence is an important optimization of active learning, especially for complex tasks like object detection and image recognition. This means that individual tasks can be optimized separately, and it results in more efficient data labeling, better generalization, and scalable models. The next section discusses why task independence is important and how it benefits active learning.

1. Efficient Data Labeling: Having models avoid focusing solely on task-specific data minimizes duplicate labeling. Ren et al. (2020) illustrate how uncertainty-based techniques decrease the cost of labels that still support performance across multiple tasks.

2. Stronger Generalization : Ensures that the tasks do not conflict with each other's objective. In this way, task independence supports stronger generalization. Hekimoglu et al. (2021) noted in a work that multi-task consistency should maintain tasks mutually beneficial such as, having more robust models .

3. Scalability and Resource Efficiency: Task independence allows models to scale horizontally without degrading performance. Rotman & Reichart (2024) illustrate a method where transformer-based models preserve task independence even during multiple tasks. Moreover, in resource scarce environments, techniques such as core-set selection optimize without overwhelming the system. Here, Senzaki & Hamelain et al.(2021) presents the efficiency of this over edge devices.

4. Domain Adaptation : Domain adaptation is based on the idea of task independence, which allows models to generalize across different environments without losing performance on other tasks. Xie et al. (2022) introduce the EADA algorithm, which reduces domain gaps while maintaining task independence.

In a nutshell, task independence is the need for improving efficiency in active learning, generalization, scalability, and resource optimization. Ensuring individual optimization of tasks enables active learning systems to work well even in complex, multi-task environments.

*I.Theoretical Foundations*

Sessa et al. (2023) bring in some solid theory with a mathematical model for balancing tasks without sacrificing too much performance. This complements Hekimoglu et al. (2021) and Xiao et al. (2021), who work on keeping task independence practical and efficient. But theory and practice don’t always line up, so there’s still a gap in making these ideas work smoothly in actual applications.

So, the research has come a long way. It’s gone from basic ideas about active learning to real-world applications that deal with task independence in multi-task setups. Early work from Ren et al. (2020) laid the groundwork, while Hekimoglu et al., Rotman and Reichart, and Xiao et al. have taken it to the next level, handling practical challenges. Architectural improvements like Zhao et al. (2019) and Sun et al. (2024), plus selection strategies from Xiao et al. (2021), show that this field is evolving fast, with a clear path toward scalable, practical solutions for different types of tasks.

## III. UNDERSTANDING OF EACH PAPER

This section delves into the methodology and results of each study, highlighting their contributions to task independence, active learning, object detection, and image recognition. The focus is on how these works address the challenges posed by these distinct yet interconnected tasks.

*1. A Survey of Deep Active Learning (Ren et al., 2020)*

Ren et al. conducted a comprehensive review of deep active learning techniques, emphasizing uncertainty-based and diversity-based strategies. These strategies were applied to object detection and image recognition tasks, both of which demand significant amounts of labelled data. The study explored methods to reduce this labelling requirement.

The findings indicated that uncertainty-based strategies excel at selecting informative data points, effectively reducing labelling efforts without diminishing overall performance. In contrast, diversity-based approaches were shown to enhance recall rates, thereby improving task independence by allowing models to generalize effectively across different visual tasks.

*2. Active Learning for Deep Neural Networks on Edge Devices (Senzaki & Hamelain, 2021)*

This research by Senzaki and Hamelain applied active learning techniques, such as core-set selection and stream-based batch learning, to resource-constrained environments like edge devices. These environments typically face limitations in computational power and dataset size.

The results demonstrated that core-set selection significantly improved the efficiency of edge-based models, enabling real-time task handling despite constrained resources. Stream-based learning further enhanced performance, supporting task independence by optimizing the management of diverse tasks in real-time conditions.

*3. Multitask Learning with No Regret (Sessa et al., 2023)*

Sessa et al. developed an adaptive no-regret algorithm aimed at managing multiple tasks simultaneously, without compromising overall performance. The study was validated using real-world drug discovery data, an area known for its complexity and need for precision across different tasks.

The algorithm's success lay in its ability to improve confidence bounds across multiple tasks, thus achieving task independence. The ability to balance task performance without degradation is crucial, particularly in multitasking environments where detection and recognition tasks must be performed concurrently.

*4. Object Detection with Deep Learning: A Review (Zhao et al., 2019)*

Zhao et al. reviewed key object detection methods, including YOLO, SSD, and Fast R-CNN. These methods were analysed for their speed, accuracy, and applicability to real-time detection tasks.

The results highlighted trade-offs between different architectures: YOLO demonstrated superior speed, Fast R-CNN excelled in accuracy, and SSD balanced the two. This comparison underscored how different detection models manage task independence, offering varying strengths depending on whether speed or accuracy is prioritized in object detection tasks.

*5. The Evolution of Object Detection Methods (Sun et al., 2024)*

Sun et al. examined the progression of object detection techniques, focusing on improvements in SSD and the introduction of transformer-based architectures. The study utilized standard benchmarks to measure performance across tasks, with a particular focus on speed and accuracy.

The improved SSD achieved notable performance, with 79.8% mAP and 89 FPS. Transformer architectures were identified as promising for real-time applications, demonstrating potential for enhancing task independence by enabling efficient task-switching in environments requiring both detection and recognition.

*6. Object Detection using YOLO (Diwan et al., 2022)*

Diwan et al. analysed the evolution of the YOLO framework, with particular emphasis on YOLOv3. The study focused on YOLO’s ability to handle real-time object detection through its single forward-pass design.

The results demonstrated that YOLOv3 achieved 57.9% mAP, highlighting its efficiency in real-time detection tasks. However, the study noted limitations in detecting small objects, which can affect task independence when the system must recognize and detect a variety of object sizes in complex environments.

*7. An Improved Deep Learning-based Optimal Object Detection (Yadav et al., 2023)*

Yadav et al. compared the performance of three popular object detection architectures: YOLO, SSD, and Faster R-CNN. The study assessed their capabilities in handling complex environments where both speed and accuracy are critical.

Faster R-CNN exhibited the highest accuracy, while YOLO led in speed. SSD struck a balance between these metrics. The comparison highlighted how different architectures handle task independence, each offering unique advantages depending on the specific requirements of object detection tasks.

*8. Active Learning for Domain Adaptation (Xie et al., 2022)*

Xie et al. introduced the Energy-based Active Domain Adaptation (EADA) algorithm to improve sample selection for domain adaptation in visual recognition tasks. The algorithm was tested on cross-domain data to assess its effectiveness in reducing domain gaps.

The results demonstrated that the EADA algorithm effectively minimized domain gaps, which in turn improved task independence. By enabling better generalization across different domains, the study showed how task performance could be enhanced, particularly in multitask environments requiring both detection and recognition.

*9. Image Recognition Technology (Liu et al., 2023)*

Liu et al. investigated the application of CNN and SVM approaches to image recognition, evaluating their effectiveness across various tasks and image resolutions.

The results indicated that both approaches achieved high accuracy rates, with CNN and SVM reaching up to 90% accuracy. Furthermore, higher image resolution contributed to improved performance. These findings suggest that task independence in image recognition can be maintained across different tasks, provided that resolution is appropriately optimized.

*10. Automated High-throughput Image Processing (Schilling et al., 2023)*

Schilling et al. focused on the development of automated pipelines for image processing in oncology, utilizing CNN-based segmentation and assisted annotation techniques. The study aimed to optimize the efficiency and accuracy of medical image analysis.

The results showed that CNN segmentation effectively supported task independence in medical imaging, while assisted annotation significantly enhanced processing efficiency. These findings underscore the importance of task independence in high-throughput image processing, particularly in medical applications where accuracy and speed are essential.

*11. Multi-Task Consistency for Active Learning (Hekimoglu et al., 2024)*

Hekimoglu et al. introduce a \*multi-task consistency\* approach to active learning (AL), focusing on aligning multiple tasks—like object classification and boundary detection—to improve model efficiency with less labelled data. Unlike traditional uncertainty- or diversity-based AL, this method identifies instances where task predictions diverge, prioritizing these for labelling. Results show that multi-task consistency reduces labelling requirements while maintaining accuracy, making it ideal for complex applications (e.g., autonomous driving) where models must manage layered tasks. By enhancing sample efficiency and generalization across tasks, this approach enables faster, more robust learning, with potential for broader AL frameworks that integrate multi-task and uncertainty strategies.

*12. Active Learning with Task Consistency and Diversity in Multi-Task Networks (Hekimoglu et al., 2024)*

Hekimoglu et al. propose an active learning (AL) framework that combines \*task consistency\* and \*diversity\* to improve multi-task networks. By maintaining consistency across tasks (e.g., object classification and segmentation) and maximizing data diversity, this method reduces labelling needs while enhancing model performance. Task consistency ensures that related tasks are mutually reinforcing, while diversity in sample selection promotes broader generalization. Experimental results show this approach significantly improves sample efficiency and robustness, making it suitable for complex, multi-task applications. This framework suggests potential advancements by integrating both consistency and diversity into AL for multi-task networks, providing a strong basis for future work in scalable AL.

*13. Multi-task Active Learning for Pre-trained Transformer-based Models (Rotman & Reichart, 2024)*

Rotman and Reichart present a multi-task active learning (AL) framework tailored for pre-trained transformer models. By simultaneously leveraging multiple tasks (e.g., classification, sequence labelling) during AL, this approach selects samples that enhance model performance across tasks, thereby optimizing labelling efforts. The framework capitalizes on transformers’ transfer learning capabilities, enabling efficient sample selection that improves generalization and reduces labelled data requirements. Results show that multi-task AL not only lowers labelling costs but also enhances task-specific and cross-task learning, underscoring its effectiveness for resource-constrained NLP applications. This work advances AL by demonstrating how multi-task learning can boost transformers’ sample efficiency and adaptability.

*14. An Efficient Active Learning Method for Multi-task Learning (Xiao, Chang & Liu, 2024)*

Xiao, Chang, and Liu introduce an active learning (AL) method optimized for multi-task learning, focusing on maximizing sample efficiency across related tasks. This approach selects samples that jointly benefit multiple tasks (e.g., classification, regression), thus reducing overall labelling costs. By identifying data points with high uncertainty across tasks, the method targets examples that can improve shared representations and generalization. Experimental results demonstrate that this multi-task AL method minimizes labelled data needs while maintaining strong performance across tasks, making it ideal for resource-limited, multi-task scenarios. This work highlights a scalable AL solution for effectively training multi-task models with minimal labelled data.

IV. EXPERIMENTAL SETUP

This experiment is designed to evaluate the role of task independence in multi-task learning for object detection and image recognition. Two models will be designed and compared to analyze how task-specific optimization may be achieved when combining these tasks. The experimental setup includes two different models, both of which use active learning techniques to minimize the amount of labeled data required while maintaining or even improving performance on both tasks.

*1. Model 1: Object Detection Model with Image Classification Layers Model Architecture*.

The first model begins with the architecture of object detection, such as YOLO, Faster R-CNN, or SSD, and adds some layers that are specifically designed for image classification, like some convolutional layers or some fully connected layers for classification. This hybrid model is learned using active learning strategies, specifically uncertainty-based and diversity-based sampling, to select data points most likely to be informative for being labeled.

Training Process: The model first learns to do object detection (localize and classify multiple objects in an image). Added image recognition layers enable the model also to classify the overall image, so a multi-task learning framework is created. Active learning will be applied to optimize the performance of the model while keeping the number of labeled samples as small as possible for both tasks.

*2. Model 2: Image Recognition Model with Object Detection Layers Model Architecture*

The model architecture is a picture recognition architecture, such as CNN or ResNet applied to image classification, but extends it with object detection layers, such as region proposal networks or even bounding box prediction layers. Active learning strategies can be explored to improve labelling efficiency and model performance.

Training Process: The first part of training is to train to classify entire images. The object detection layers added will enable the model to localize and detect objects within images, making it a multi-task learning system. Both models will use active learning in order to reduce the amount of labeled samples for each task.

*3. Comparison between the two models*

Objective: The two models shall be compared in key performance metrics such as accuracy, mAP, and labeling efficiency. The comparison shall establish:

* Each model's performance over both tasks of object detection and image classification.
* The extent to which task independence can be induced by adding layers from one task on top of another.
* The success of active learning techniques in bringing out the best for both task performances.

V. RATIONALE FOR THE EXPERIMENTAL SETUP

This experimental setup was supposed to determine whether task independence could be preserved when combining object detection and image recognition tasks in one model. The rationale for each model and approach is as follows:

*1. Object Detection Model with Image Recognition Layers (Model 1)*

Why This Approach?: Object detection models are inherently more complex because they are both localization and classification. Adding recognition layers to an object detection model aims to test whether the model learns both tasks independently without being compromised in performance. It will be used to establish if task independence can be attained when the model first focuses on detection and then learns to classify. This model will also be useful in determining the degree to which adding the layers of image recognition to an object detection model affects the independence of the task. It further provides the insights into how active learning optimizes the hybrid model.

*2. Model with object detection layers incorporated into image recognition model (Model 2)*

Why This Approach?: Image recognition models are often much less complex because they aim to classify whole images. Adding the object detection layers enables it to perform both tasks. What the approach does here is checking whether adding localization capabilities on top of the classification results in a loss of task independence.

Contribution to the Goal: This model will determine whether adding object detection layers to an image recognition model will preserve task independence while allowing the model to perform both tasks. Active learning strategies will be used to optimize the training process of both tasks.

*3. Comparison of the Two Models*

Why Compare the Two?: A comparison of these two models will indicate whether one approach - starting from object detection, versus one starting from image recognition-is superior in preserving task independence. Both of these models incorporate active learning and thus allow the trade-offs between task-specific optimization and multi-task generalization to be evaluated.

VI. TIMELINE

Complete Literature Search  
Date: September 4th  
Task: Identify and collect all relevant research papers focusing on task independence, active learning, object detection, and image recognition.

1. Summarize Key Papers  
   Date: September 9th  
   Task: Read each selected paper and document key findings, methodologies, and insights. Summarize each paper’s contribution to understanding task independence and active learning.
2. Identify Research Gaps and Synthesize Findings  
   Date: September 30th  
   Task: Analyze patterns, gaps, and emerging trends in the literature. Synthesize findings to highlight how task independence in multilevel learning contributes to the field of active learning and object detection.
3. Draft Literature Review Section  
   Date: October 30th  
   Task: Write the literature review section, organizing content by themes, challenges, methodologies, or other relevant categories based on your analysis.
4. Revise and Finalize Literature Review  
   Date: November 14th  
   Task: Edit the literature review for coherence, clarity, and adherence to the IEEE format. Ensure that the review effectively emphasizes the importance of task independence in active learning within computer vision applications.

## VII. CONCLUSION

The reviewed papers reveal several key trends in the advancement of task independence, active learning, and object detection within the broader context of computer vision. Across all studies, active learning emerges as a central strategy for reducing data labeling requirements while maintaining performance across diverse tasks. Techniques such as uncertainty-based sampling, diversity-based methods, and core-set selection have proven effective in balancing the varying demands of object detection and image recognition. These approaches help address the challenges of task independence by allowing models to prioritize and select the most informative samples for learning, ultimately reducing the need for vast labelled datasets.

In terms of object detection methods, the comparison of popular frameworks like YOLO, SSD, and Faster R-CNN highlights a consistent trade-off between speed and accuracy. While YOLO remains favored for its real-time capabilities, Faster R-CNN excels in precision, and SSD provides a middle ground. The evolution of these architectures reflects an ongoing trend toward more efficient, task-independent systems that can switch between detection and recognition tasks without significant performance loss. The integration of multi-task active learning further supports this goal by ensuring task consistency and enhancing performance across different tasks in a unified model, as highlighted in studies such as those by Hekimoglu et al. (2020) and Rotman & Reichart (2021).

Moreover, several studies (e.g., Rotman & Reichart, 2021; Xia et al., 2021) point to the promise of transformer-based architectures in enhancing multi-task learning. These architectures are shown to improve task-switching capabilities, offering potential breakthroughs in task independence and real-time performance across diverse and complex tasks. Transformer models, as discussed in these studies, could offer novel solutions for real-time adaptation in active learning systems, addressing the challenges of task independence in dynamic environments.

Another significant development is the adaptation of these technologies to resource-constrained environments, such as edge devices. Research on active learning for edge computing, like that presented by Xia et al. (2021), demonstrates the potential for real-time task handling despite computational limitations, suggesting a broader scope for deploying these models in practical applications beyond high-performance computing environments. This aligns with the findings of Hekimoglu et al. (2020), who explored lightweight architectures for edge devices, showing that task independence can be maintained even with reduced resources.

In conclusion, the continued exploration of task independence, active learning, and optimized architectures will drive significant advancements in object detection and recognition. The integration of transformer-based models, multi-task consistency, and adaptation to resource-constrained environments is expected to push the boundaries of what is achievable in modern computer vision, with wide-reaching implications for both academic research and practical applications.

## VIII.. FUTURE RESEARCH DIRECTIONS

Several avenues for future research emerge from this review:

1. Task Independence Optimization: There remains room for further optimization in balancing object detection and recognition tasks. Future research should focus on developing architectures that can dynamically adapt to multiple tasks, minimizing trade-offs between speed and accuracy while maintaining task independence across various domains.
2. Transformer Architectures: Early studies show the promise of transformer-based models in improving real-time performance and task-switching capabilities. Exploring their application in more complex multitask environments could lead to breakthroughs in task independence and efficiency.
3. Small Object Detection: Despite advances in object detection, challenges remain in handling small objects across diverse environments. Future work should aim to refine models like YOLO and SSD to better accommodate smaller object sizes while maintaining overall task performance.
4. Edge Computing and Resource Efficiency: As active learning becomes more integrated into edge devices, further research should explore lightweight model architectures that maintain high performance with minimal computational resources. This includes real-time adaptation to changing task requirements in edge environments.
5. Domain Adaptation and Generalization: While active learning and domain adaptation have shown promise in improving task independence, future studies should focus on reducing domain gaps across more diverse datasets. Research into energy-based and other adaptive algorithms can help models generalize better across new environments and tasks.
6. Task Consistency in Multi-Task Networks: Research should focus on improving task consistency in multi-task active learning, balancing the performance of each task while maintaining overall efficiency in model training.
7. Diversity in Sample Selection: Future research could investigate strategies that promote diversity in sample selection within multi-task active learning systems, improving the robustness and efficiency of models in complex tasks.
8. Real-Time Adaptation to Task Requirements: The development of models that can adapt in real-time to evolving task demands is a key area for future work, particularly for applications requiring continuous updates to model behavior.

In conclusion, the continued exploration of task independence, active learning, and optimized architectures will drive significant advancements in object detection and recognition, with implications for both academic research and practical applications. The identified trends suggest that future developments will likely focus on refining models for greater adaptability, efficiency, and generalization, pushing the boundaries of what is achievable in modern computer vision.