Design Exploration of CNN Parameters for Multi-Altitude UAV Object Detection

Michalis Piponidis[1,*], Theocharis Theocharides[1]

[1]KIOS Research and Innovation Center of Excellence, University of Cyprus, 1 Panepistimiou Avenue, 1678 Nicosia, Cyprus

Abstract

Object detection using Unmanned Aerial Vehicles (UAVs) introduces unique challenges compared to traditional methods, primarily due to the varying angles and altitudes from which images are captured. Conventional Convolutional Neural Network (CNN) implementations, while being the state-of-the-art in object detection, demand substantial memory and computational resources, posing difficulties for deployment on UAVs with limited onboard resources. Furthermore, these models typically generalize well within a limited distance range by training on images captured from various distances. However, UAVs capture images at a wide range of distances, complicating the model's ability to generalize effectively. In this work therefore, we present a design space exploration which aims to identify the effect of CNN parameters for UAV-based object detection at various altitudes by examining two critical parameters: input image resolution and network width (number of channels). We conduct extensive experiments to evaluate the effect of these parameters in terms of accuracy and computational efficiency across multiple altitudes. Lower resolutions reduce computational load but may compromise detection accuracy, while higher resolutions enhance accuracy at the expense of increased processing requirements. Similarly, varying the network width influences the balance between model complexity and detection performance. We showcase that the requirements vary significantly across different altitudes, demonstrating the potential of dynamic network structures that adjust parameters according to the altitude. Our findings provide insights into the optimal configuration of CNN parameters for UAV object detection across different altitudes, contributing to the development of more efficient and adaptable UAV vision systems. This research paves the way for more effective deployment of UAVs in various applications, from surveillance and search-and-rescue to environmental monitoring and beyond.

Keywords

Unmanned Aerial Vehicles, Object Detection, Convolutional Neural Networks, Dynamic Neural Networks, Design Space Exploration

1. Introduction

Unmanned Aerial Vehicles (UAVs), commonly known as drones, are increasingly employed across a broad spectrum of fields, including but not limited to search and rescue operations [1], emergency management [2], infrastructure inspection [3], agricultural monitoring [4], and environmental monitoring [5]. Their versatility, characterized by their unmanned nature, ease of deployment, cost-effectiveness, and capability to capture aerial imagery from various perspectives and altitudes, makes them an indispensable tool in modern technology.

Object detection on UAVs is a critical task due to its role in enhancing situational awareness, which is pivotal for safe navigation, obstacle avoidance, and the execution of missions that necessitate the identification of subjects of interest. However, several challenges are inherent to object detection using UAVs. Firstly, due to their nature, most UAV applications require the object detection task to be performed in real-time. On-board computational and power resources are limited, thereby requiring the deployment of models that are computationally efficient while maintaining the required accuracy for the task at hand. Secondly, unlike in traditional object detection tasks, UAV-based tasks encounter images from a wide variety of angles and altitudes, potentially impacting the generalization capabilities of the model. Furthermore, diverse environmental conditions such as reflections, smoke, and adverse

CPS Workshop 2024: September 16–20, 2024, Alghero, Sardinia (Italy)

*Corresponding author.
mpipon01@ucy.ac.cy (M. Piponidis); ttheocharides@ucy.ac.cy (T. Theocharides)
0000-0002-3171-8205 (M. Piponidis); 0000-0001-7222-9152 (T. Theocharides)

weather can obscure objects and degrade image clarity, further complicating the detection task.

Existing research in UAV-based object detection demonstrates significant advancements in handling various challenges posed by UAV imagery. These studies collectively emphasize the importance of efficient detection methods that satisfy the constraints of UAV platforms. Techniques such as adaptive feature extraction [6], multi-scale detection [7], and lightweight model designs have shown promise in improving detection accuracy and efficiency. Convolutional Neural Networks (CNNs) represent the state-of-the-art in object detection, which is why we have chosen to explore them. These networks typically generalize well within a limited distance range by training on images taken from various distances. However, UAVs capture images at a wide range of distances, complicating the model's ability to generalize effectively due to significant variations in optimal parameter values across large distance ranges. This highlights a gap in research: the need for dynamic parameter adjustments tailored for UAV object detection, which is crucial for enhancing real-time performance.

In this study, we examine the effects of two critical parameters–input image resolution and network width–on the performance and efficiency of CNN models for UAV-based object detection. We hypothesize that higher input resolutions and wider networks can capture more details and features, respectively; however, these advantages come with an increased computational load. Our objective is to identify the optimal values for these parameters across different altitude ranges to optimize the accuracy-to-performance trade-off. We aim to use these findings to develop a dynamic network structure capable of adjusting these parameters based on a learnable dynamic decision point (rather than a static threshold) tailored to each input image. This will be implemented using reconfigurable network structures (either hardware or software), enabling us to enhance the accuracy-to-performance balance.

Specifically, the contributions of this research are twofold:

- We perform extensive evaluation of the effect of changing the two parameters, input resolution and network width (number of channels), on the accuracy, performance and memory requirements of the object detection CNN (tiny YOLOv7) at various altitude ranges.
- We demonstrate that the optimal values of these parameters varies significantly with altitude, highlighting the need for dynamic adjustments based on real-time altitude data.

The remainder of this paper is structured as follows: Section 2 reviews the existing related work. In Section 3, we detail our methodology. Section 4 presents the experimental results. Finally, Section 5 discusses some of the challenges and potential future work and Section 6 concludes our study.

2. Related Work

Over the years, numerous advancements have been made in the field of UAV-based object detection, particularly focusing on improving accuracy and efficiency. Several studies have aimed to address the unique challenges posed by UAV imagery, such as small object size, high density, and varying viewpoints.

A significant portion of the literature has focused on improving detection algorithms and network architectures. Tan et al. [6] proposed a multi-scale UAV aerial image detection method using adaptive feature fusion to better detect small target objects. They introduced an adaptive feature extraction module to the backbone network, enabling more accurate small target feature information extraction. Similarly, Xiaohu et al. [7] presented a dedicated object detector based on the FPN architecture, incorporating Deformable Convolution Lateral Connection Modules (DCLCMs) and Attention-based Multi-Level Feature Fusion Modules (A-MLFFMs) to enhance multi-scale object detection. The work by Liu et al. [8] also targeted small object detection in UAV imagery by optimizing YOLOv3 and the darknet structure to improve spatial information capture and receptive fields, leading to performance improvements on small object detection.

Comprehensive reviews of deep learning techniques applied to UAV-based object detection have highlighted the importance of lightweight models for deployment on UAVs with limited computational resources. Wu et al. [9] explored various CNN architectures like YOLO and Faster R-CNN, demonstrating

Table 1
Explored Parameter Values

Input Resolutions (pixels)						
1088x1088	896x896	768x768	640x640	512x512	384x384	256x256
Width Multipliers						
1x (Default)	0.75x	0.5x	0.25x	0.1x	0.01x	

the trade-offs between model complexity and computational efficiency. Furthermore, Mittal et al. [10] provided an extensive review of state-of-the-art deep learning-based object detection algorithms, particularly focusing on low-altitude UAV datasets, and discussed research gaps and challenges in the field.

Real-time object detection in UAV imagery has also gathered significant attention due to its importance in scenarios like emergency rescue and precision agriculture. Cao et al. [11] systematically reviewed previous studies on real-time UAV object detection, covering aspects such as hardware selection, real-time detection paradigms, and algorithm optimization technologies. They emphasized the importance of lightweight convolutional layers and GPU-based edge computing platforms to meet the demands of real-time detection.

Several innovative methodologies have been introduced to enhance UAV object detection. Bazi et al. [12] proposed a convolutional support vector machine (CSVM) network for UAV object detection, leveraging SVMs as filter banks for feature map generation. This approach was particularly useful for problems with limited training samples. Zhang et al. [13] introduced a global density fused convolutional network (GDF-Net) optimized for object detection in UAV images, using a Global Density Model to refine density features and improve detection performance in congested scenes.

Despite the extensive research in UAV object detection, our work is the first to examine dynamic parameters for UAV object detection, exploring the need for adaptable and flexible detection methods. By focusing on dynamic parameter adjustment, our approach aims to address the limitations of static models in varying UAV operational conditions, enhancing detection accuracy and efficiency across diverse environments and scenarios.

3. Methodology

The parameter values explored in this study are shown in Table 1. The input image resolutions were varied from 1088x1088 pixels to 256x256 pixels, allowing us to investigate the impact of different levels of detail and computational demands on the object detection performance. Additionally, the network width was systematically adjusted using multipliers ranging from 1x, which represents the original model structure, to 0,01x. This adjustment helped us evaluate how the model's number of channels in the convolutional layers influenced its detection capabilities and efficiency.

For evaluation, we employed several metrics to provide a comprehensive assessment of the model's performance. Mean average precision (mAP) was utilized to quantify the accuracy of object detection across the different parameter settings. To evaluate computational efficiency, we measured the multiply-accumulate (MAC) operations, which reflect the computational load required for processing, and the network size in megabytes (MB), which indicates the model's memory requirements.

For the object detector, we selected the tiny YOLOv7 model [14], which is known for its balance between detection accuracy and computational efficiency. This model is part of the YOLO (You Only Look Once) family, which is renowned for real-time object detection capabilities. The tiny YOLOv7 variant is specifically designed to be lightweight, making it particularly suitable for on-board implementation on UAVs where computational resources and power are limited.

As the dataset for our study, we utilized the Multi-Altitude Aerial Vehicles Dataset [15], which focuses on single-class object detection specifically targeting cars. This dataset was chosen due to its unique composition of images captured at various altitudes, ranging from 50 meters to 500 meters, with increments of 50 meters. This range provides a comprehensive platform for experimenting with different parameter values across diverse altitude levels.

To comprehensively analyze the impact of varying parameters at different altitudes, we segmented the original dataset into five distinct subsets, each corresponding to specific altitude ranges:

- 50–100 meters
- 150–200 meters
- 250–300 meters
- 350–400 meters
- 450–500 meters

In addition to these altitude-specific subsets, we also utilized the entire dataset, which includes images from all altitude ranges. This dataset, referred to as *mix_alt*, serves as a baseline for comparison against the individual altitude-specific subsets.

For each of these six datasets (the five altitude-specific subsets and the *mix_alt* dataset), we conducted extensive training experiments. We systematically varied the input image resolution and network width across all possible combinations while maintaining other parameters at their default values. This approach allowed us to evaluate the influence of the two parameters on the performance metrics.

4. Experimental Results

4.1. Performance

The results in terms of mean Average Percision (mAP) by setting the threshold of Intersection over Union (IoU) to 0.5 for each dataset and parameter combination are illustrated in Figure 1. As expected, lower altitude datasets require less computationally intensive models, with less parameters, to achieve high accuracy. This is due to the larger size, greater clarity and detail available in lower altitude images, which simplifies the task.

Conversely, as the altitude increases, the detection accuracy tends to diminish unless higher parameter values are utilized. This can be attributed to the increased complexity of identifying objects from higher vantage points, where objects appear smaller and less distinct. Thus, higher altitudes necessitate models with greater capacity to maintain comparable levels of accuracy, reflecting the need for more detailed feature extraction and processing capabilities.

Interestingly, *mix_alt* demonstrates an intermediate performance, performing better than the higher altitude-specific models. This is because *mix_alt* is trained on the whole dataset, including the easier, lower-altitude images, whereas the higher altitude-specific models are trained only on the hardest images of the dataset.

4.2. Computational Efficiency

MAC operations are a key indicator of computational efficiency in neural network models, representing the total number of multiplications and additions needed to process an image. By reducing the number of MAC operations, we can lower the computational load on the UAV, extending its operational time and enabling real-time processing.

In addition to MAC operations, the size of the model is another important factor affecting computational efficiency. Model size refers to the amount of memory required to store the model parameters, including weights, biases, and other configurations. This size directly influences the amount of onboard storage needed for the UAV to operate.

Figure 2 presents the computational efficiency results for each model configuration. As anticipated, both MAC operations and model size increase with larger parameter values. Notably, MAC operations scale uniformly with both parameters, whereas model size is more significantly affected by the width multiplier compared to the resolution. This suggests that for prioritizing a smaller model size, a lower width combined with a higher resolution is more effective than the reverse.

50-100m

Resolution

mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,996	0,997	0,996	0,994	0,994	0,992
896	0,997	0,996	0,995	0,995	0,991	0,973
768	0,998	0,994	0,994	0,994	0,989	0,969
640	0,997	0,996	0,995	0,992	0,961	0,936
512	0,997	0,991	0,995	0,985	0,984	0,898
384	0,998	0,991	0,982	0,954	0,928	0,971
256	0,996	0,954	0,926	0,879	0,688	0,731

Width Multiplier

350-400m

Resolution

mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,951	0,917	0,83	0,739	0,584	0,372
896	0,935	0,892	0,849	0,649	0,383	0,143
768	0,922	0,845	0,766	0,314	0,45	0,0168
640	0,571	0,017	0,111	0,139	0,0427	0,157
512	0,774	0,00547	0,0569	0,0054	0,00205	0,0113
384	0,314	0,0109	4,45E-05	0,000342	7,72E-05	0,000101
256	5,73E-06	6,54E-06	2,69E-05	1,15E-05	2,11E-05	0,000834

Width Multiplier

150-200m

Resolution

mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,997	0,995	0,994	0,989	0,991	0,978
896	0,997	0,996	0,994	0,984	0,976	0,983
768	0,998	0,994	0,982	0,985	0,967	0,957
640	0,997	0,973	0,984	0,969	0,932	0,943
512	0,996	0,99	0,967	0,929	0,869	0,848
384	0,969	0,144	0,0239	0,00934	0,00826	0,000339
256	0,223	0,00145	0,000104	0,00142	3,53E-05	9,12E-05

Width Multiplier

450-500m

Resolution

mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,935	0,878	0,801	0,628	0,454	0,39
896	0,897	0,714	0,69	0,296	0,25	0,271
768	0,807	0,494	0,511	0,35	0,189	0,288
640	0,417	0,463	0,121	0,339	0,116	0,176
512	0,0954	0,232	0,00675	0,00652	0,0777	0,0422
384	0,00102	0,001	0,000114	0,00905	0,0123	0,00513
256	4,97E-06	0,000209	3,66E-05	1,47E-05	2,16E-05	9,17E-06

Width Multiplier

250-300m

Resolution

mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,992	0,986	0,987	0,967	0,883	0,834
896	0,991	0,983	0,988	0,938	0,794	0,685
768	0,989	0,978	0,959	0,807	0,689	0,408
640	0,981	0,956	0,911	0,704	0,438	0,00322
512	0,938	0,00962	0,499	0,465	0,384	0,0185
384	0,483	0,4	0,294	0,00335	0,000375	0,000369
256	0,00547	0,133	7,62E-05	8,03E-05	4,55E-05	7,07E-05

Width Multiplier

mix_alt

Resolution

mAP@.5	1	0.75	0.5	0.25	0.1	0.01
1088	0,961	0,935	0,916	0,808	0,652	0,551
896	0,939	0,826	0,845	0,677	0,559	0,441
768	0,943	0,859	0,881	0,708	0,594	0,453
640	0,919	0,879	0,513	0,533	0,482	0,264
512	0,802	0,318	0,26	0,0591	0,0569	0,0701
384	0,0211	0,274	0,0324	0,00188	0,000165	0,0308
256	0,0154	0,0174	0,0157	0,000104	0,0025	0,00687

Width Multiplier

Figure 1: Performance Results: mAP@.5. The data is presented in heatmap format, where green represents high values and red represents low values.

MAC Operations (M)

Resolution

	1	0.75	0.5	0.25	0.1	0.01
1088	45,42	34,13	22,84	11,55	5,81	2,38
896	31,99	24,04	16,09	8,14	4,09	1,68
768	22,71	17,07	11,42	5,78	2,9	1,19
640	16,35	12,29	8,22	4,16	2,09	0,86
512	11,03	8,29	5,55	2,81	1,41	0,58
384	6,75	5,07	3,39	1,72	0,86	0,35
256	2,92	2,19	1,47	0,74	0,37	0,15

Width Multiplier

Size (MB)

Resolution

	1	0.75	0.5	0.25	0.1	0.01
1088	17,41	12,15	8,39	6,13	5,55	5,4
896	15,83	10,57	6,81	4,55	3,97	3,82
768	14,73	9,47	5,71	3,45	2,87	2,72
640	13,98	8,72	4,96	2,7	2,12	1,97
512	13,35	8,09	4,33	2,07	1,49	1,34
384	12,85	7,59	3,83	1,57	0,99	0,84
256	12,4	7,14	3,38	1,12	0,54	0,39

Width Multiplier

Model	Parameters	Parameter Size (MB)
1	12.015.192	12,05
0.75	6.768.120	6,79
0.5	3.016.600	3,03
0.25	760.632	0,77
0.1	190.680	0,19
0.01	36.488	0,04

Figure 2: Computational Efficiency Results: MAC Operations, Model Size, and Parameters Size. The data is presented in heatmap format, where green represents high values and red represents low values.

5. Challenges and Future Work

The most significant challenge is the seamless transition between models during inference. Currently, switching between models that are optimized for different altitude ranges can be computationally intensive and may introduce latency, which is detrimental to real-time processing requirements. To address this, we will explore the potential of dynamic networks that can reconfigure themselves on-the-fly to adapt to changing altitudes.

Reconfigurable hardware presents a promising solution to this challenge. By utilizing hardware that can adapt its configuration based on the UAV's altitude, it is possible to optimize the processing pipeline for speed and efficiency. Field Programmable Gate Arrays (FPGAs) that support dynamic reconfiguration could be investigated for this purpose [16].

Additionally, exploring algorithmic, non-manual methods for selecting the optimal parameters for each altitude range will streamline the process and reduce the large computational time and resources typically required for manual exploration, which scales exponentially as the network size and number of parameters increase. By employing automated approaches, the system could efficiently determine the best parameters based on the specific UAV's capabilities and the task's required accuracy. These methods would minimize human intervention, allowing UAVs to dynamically adjust to varying conditions and mission requirements while achieving an optimal accuracy-performance trade-off. This would enable the UAVs to adapt in real time without requiring labor-intensive manual tuning, enhancing operational efficiency.

In addition to technical advancements in model adaptability, there is a critical need to expand the datasets used for training and evaluation. Our current dataset has fixed altitudes and limited environmental diversity, which does not fully capture the complexities encountered in real-world scenarios. Therefore, we aim to create a comprehensive dataset that includes images captured at varying altitudes and under diverse environmental conditions such as different weather patterns, times of day, locations etc. This dataset would not only improve the robustness of the detection models but also provide a more rigorous benchmark for future research in UAV-based object detection.

Addressing the challenges of parameter optimization, seamless model transition, and dataset diversity will be crucial for advancing the field of UAV-based object detection. Through a combination of innovative algorithms, adaptable hardware solutions, and comprehensive data collection, we aim to significantly enhance the performance and applicability of UAV object detection systems.

6. Conclusions

Our study confirms that different altitude ranges necessitate distinct parameters to achieve optimal accuracy levels in UAV-based object detection. We have demonstrated that input image resolution and network width are critical factors that must be tuned according to the altitude from which images are captured.

The findings indicate that dynamic network structures, which adjust their parameters based on real-time altitude data, can substantially enhance both the efficiency and performance of object detection systems deployed on UAVs. Such an approach would not only optimize the accuracy-to-performance trade-off but also ensure that the computational resources of UAVs are utilized more effectively. This is particularly important given the limited processing power and battery life of most UAVs.

In summary, our research highlights the importance of considering altitude-specific parameter optimization in the design of UAV object detection systems. The use of dynamic network structures that can adapt to altitude variations presents a promising avenue for developing more flexible, efficient, and effective UAV vision systems.

This approach paves the way for advancements in UAV technology, enabling more accurate and efficient object detection, and ultimately enhancing the capabilities and applications of UAVs across various fields.

Acknowledgments

This work is supported by the European Union (Horizon 2020 Teaming, KIOS CoE, No. 739551) and from the Government of the Republic of Cyprus through the Deputy Ministry of Research, Innovation, and Digital Policy.

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

References

[1] S. Yeom, Thermal image tracking for search and rescue missions with a drone, Drones 8 (2024). URL: https://www.mdpi.com/2504-446X/8/2/53. doi:10.3390/drones8020053.

[2] O. Bushnaq, D. Mishra, E. Natalizio, I. Akyildiz, Unmanned aerial vehicles (UAVs) for disaster management, 2022, pp. 159–188. doi:10.1016/B978-0-323-91166-5.00013-6.

[3] L. Figuli, V. Gattulli, K. Hoterová, E. Ottaviano, Recent Developments on Inspection Procedures for Infrastructure Using UAVs, 2021. doi:10.3233/NICSP210003.

[4] A. Rejeb, A. Abdollahi, K. Rejeb, H. Treiblmaier, Drones in agriculture: A review and bibliometric analysis, Computers and Electronics in Agriculture 198 (2022) 107017. URL: https://www.sciencedirect.com/science/article/pii/S0168169922003349. doi:https://doi.org/10.1016/j.compag.2022.107017.

[5] R. Rani Hemamalini, R. Vinodhini, B. Shanthini, P. Partheeban, M. Charumathy, K. Cornelius, Air quality monitoring and forecasting using smart drones and recurrent neural network for sustainable development in chennai city, Sustainable Cities and Society 85 (2022) 104077. URL: https://www.sciencedirect.com/science/article/pii/S221067072200395X. doi:https://doi.org/10.1016/j.scs.2022.104077.

[6] S. Tan, Z. Duan, L. Pu, Multi-scale object detection in uav images based on adaptive feature fusion, PLOS ONE 19 (2024) 1–21. URL: https://doi.org/10.1371/journal.pone.0300120. doi:10.1371/journal.pone.0300120.

[7] D. Xiaohu, Y. Qin, Y. Gao, R. Fu, S. Liu, Y. Ye, Attention-based multi-level feature fusion for object detection in remote sensing images, Remote Sensing 14 (2022) 3735. doi:10.3390/rs14153735.

[8] M. Liu, X. Wang, A. Zhou, X. Fu, Y. Ma, C. Piao, Uav-yolo: Small object detection on unmanned aerial vehicle perspective, Sensors 20 (2020). URL: https://www.mdpi.com/1424-8220/20/8/2238. doi:10.3390/s20082238.

[9] X. Wu, W. Li, D. Hong, R. Tao, Q. Du, Deep learning for unmanned aerial vehicle-based object detection and tracking: A survey, IEEE Geoscience and Remote Sensing Magazine 10 (2022) 91–124. doi:10.1109/MGRS.2021.3115137.

[10] P. Mittal, R. Singh, A. Sharma, Deep learning-based object detection in low-altitude uav datasets: A survey, Image and Vision Computing 104 (2020) 104046. URL: https://www.sciencedirect.com/science/article/pii/S0262885620301785. doi:https://doi.org/10.1016/j.imavis.2020.104046.

[11] Z. Cao, L. Kooistra, W. Wang, L. Guo, J. Valente, Real-time object detection based on uav remote sensing: A systematic literature review, Drones 7 (2023). URL: https://www.mdpi.com/2504-446X/7/10/620. doi:10.3390/drones7100620.

[12] Y. Bazi, F. Melgani, Convolutional svm networks for object detection in uav imagery, IEEE Transactions on Geoscience and Remote Sensing 56 (2018) 3107–3118. doi:10.1109/TGRS.2018.2790926.

[13] R. Zhang, Z. Shao, X. Huang, J. Wang, D. Li, Object detection in uav images via global density fused convolutional network, Remote Sensing 12 (2020). URL: https://www.mdpi.com/2072-4292/12/19/3140. doi:10.3390/rs12193140.

[14] C.-Y. Wang, A. Bochkovskiy, H.-Y. M. Liao, YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.

[15] R. Makrigiorgis, C. Kyrkou, P. Kolios, How high can you detect? improved accuracy and efficiency at varying altitudes for aerial vehicle detection, in: 2023 International Conference on Unmanned Aircraft Systems (ICUAS), 2023, pp. 167–174. doi:10.1109/ICUAS57906.2023.10156376.

[16] F. Manca, F. Ratto, F. Palumbo, Onnx-to-hardware design flow for adaptive neural-network inference on fpgas, 2024. URL: https://arxiv.org/abs/2406.09078. arXiv:2406.09078.