YOLO-KAN: Research on the Adaptability of the Kolmogorov-Arnold Networks to You Only Look Once Model

Xiaofeng Zhao, Supervised by Prof. Tehmina Amjad

Northeastern University
Khoury College of Computer Sciences

Linkedin Github

INTRODUCTION

- In computer vision, You Only Look Once (YOLO) is widely used for real-time object detection tasks. [1]
- Researchers continuously enhance YOLO's accuracy and speed.
- Kolmogorov-Arnold Networks (KANs) could potentially boost accuracy without increasing depth of the network. [2]

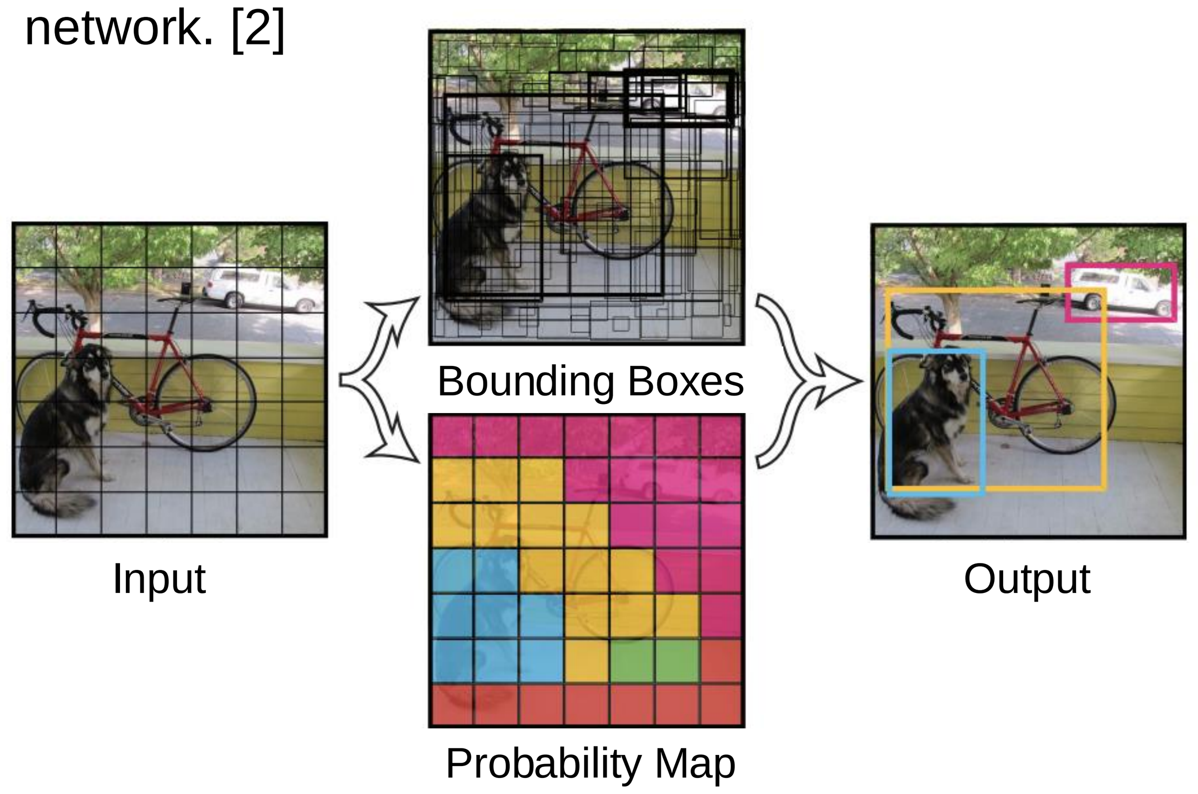
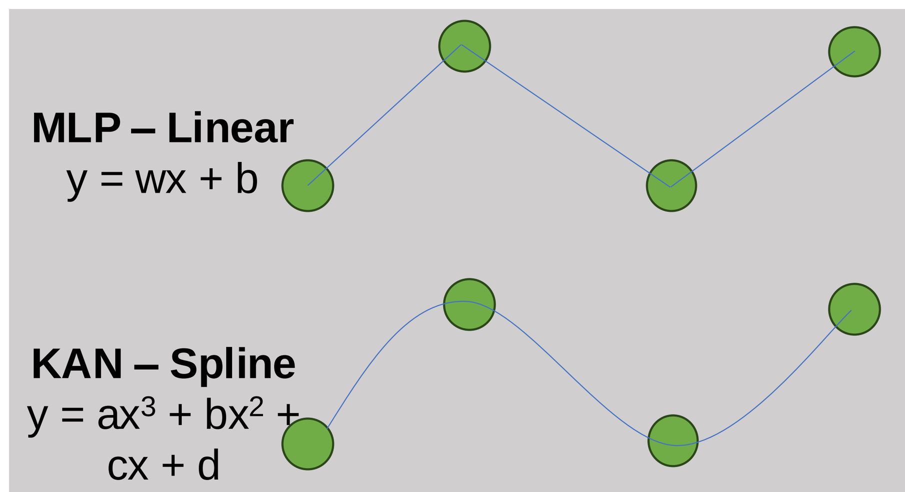


Figure 1. The one-stage object detection process of YOLO (Adapted from J Redmon et al., 2015)

Methods

1. Mathematical Principles: Spline Function



2. Ablation Experiments on Serialization

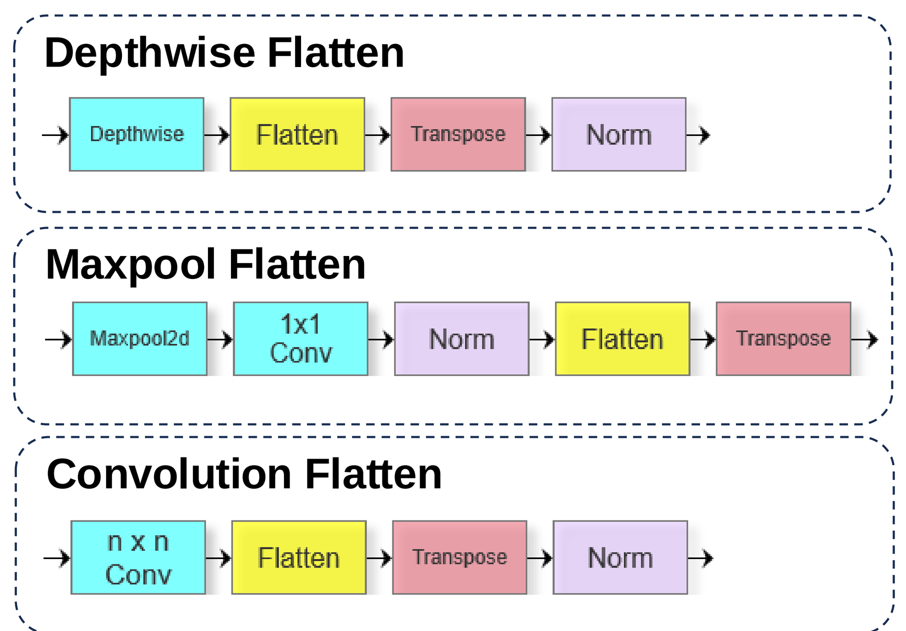


Figure 3. Flatten block structure

3. Overall YOLO – KAN Architecture

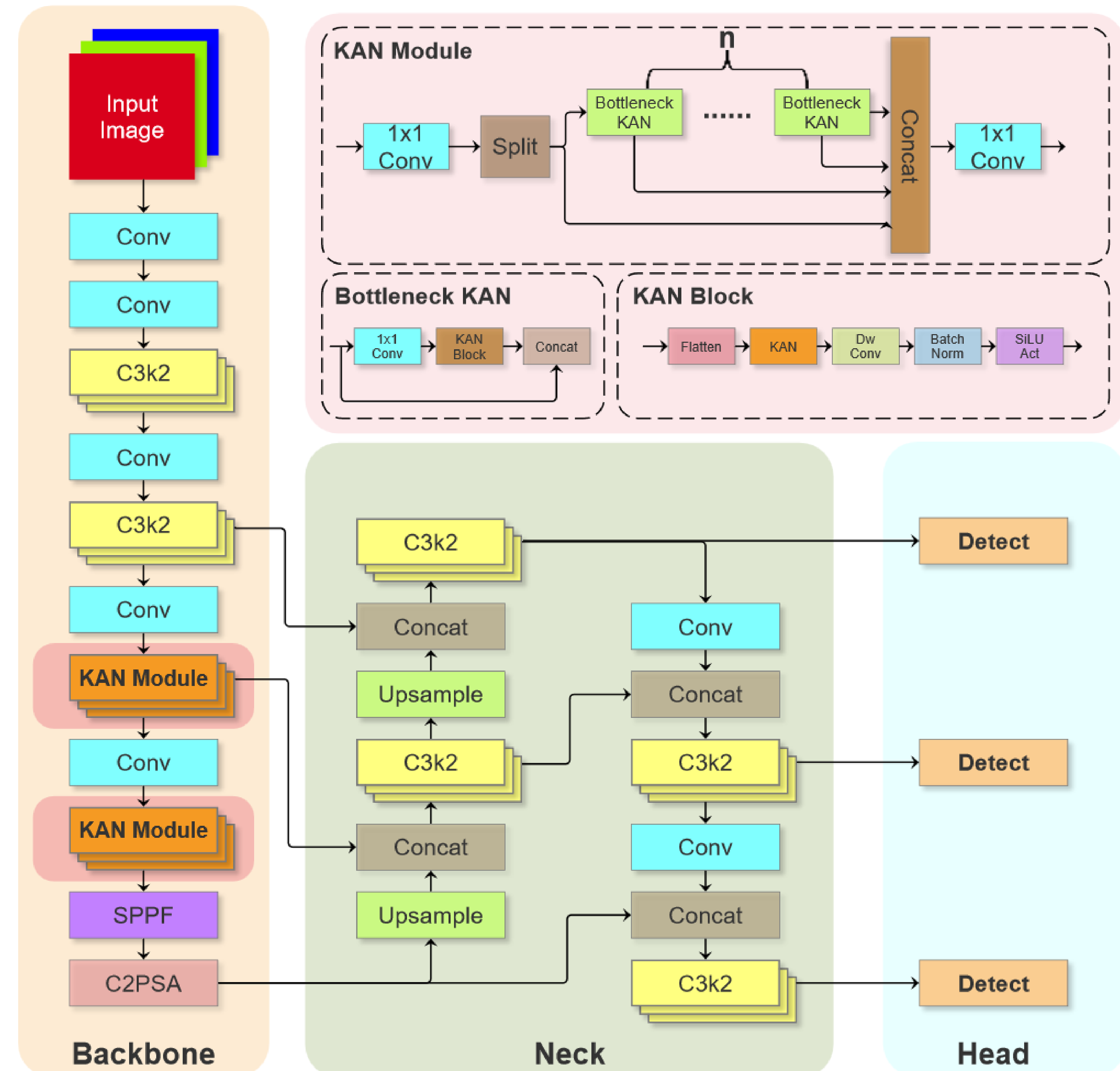


Figure 4. YOLO – KAN network architecture

4. Simplified Network Architecture

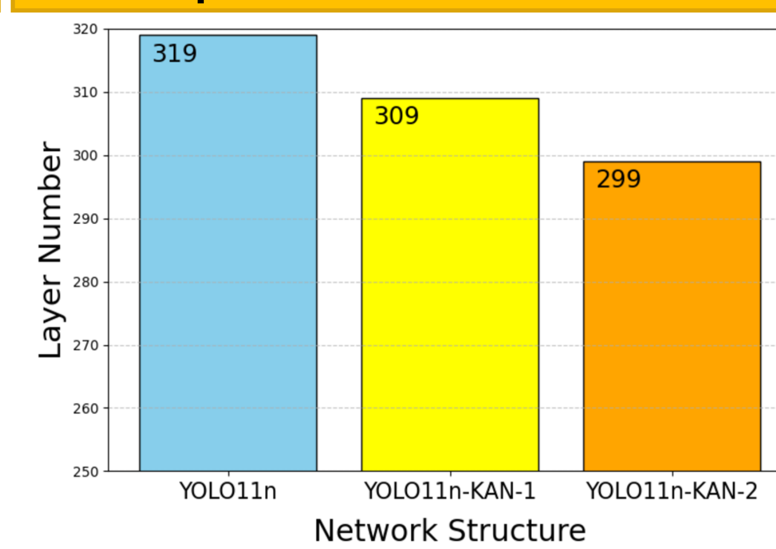


Figure 5. Comparison of layer numbers

Layer Number ↓

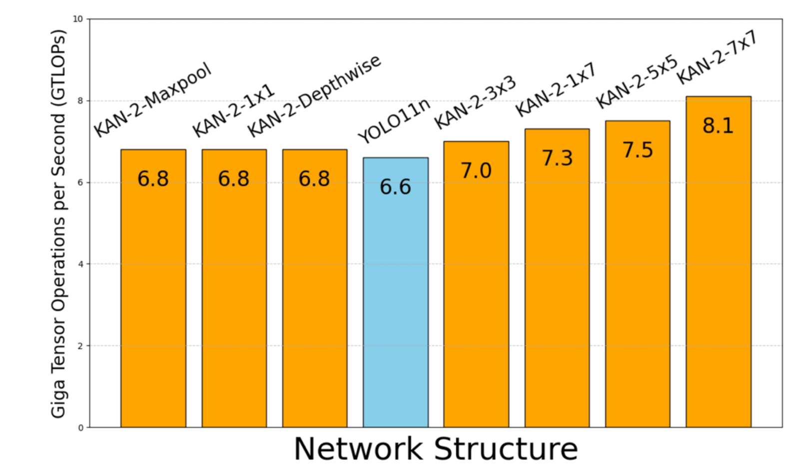


Figure 6. Comparison of GTLOPs

Computational Cost ↑

*KAN-2-1 means YOLO-KAN architecture using 2 KAN Module and 1 x 1 Convolutional Flatten Layer

CONCLUSIONS

- Here, We successfully introduced the KAN structure into the YOLO model and trained it on the large graphical dataset Microsoft COCO, obtaining better accuracy than the original model and also reduced network depth.
- The ablation experiments on several KAN structures shows that the input flatten layer structure has a significant effect on the effectiveness of the KAN.

MOTIVATIONS

KAN has gained widespread attention since its release due to its interpretability and efficient feature extraction capabilities. Some studies have demonstrated its effectiveness in fields such as scientific computation.

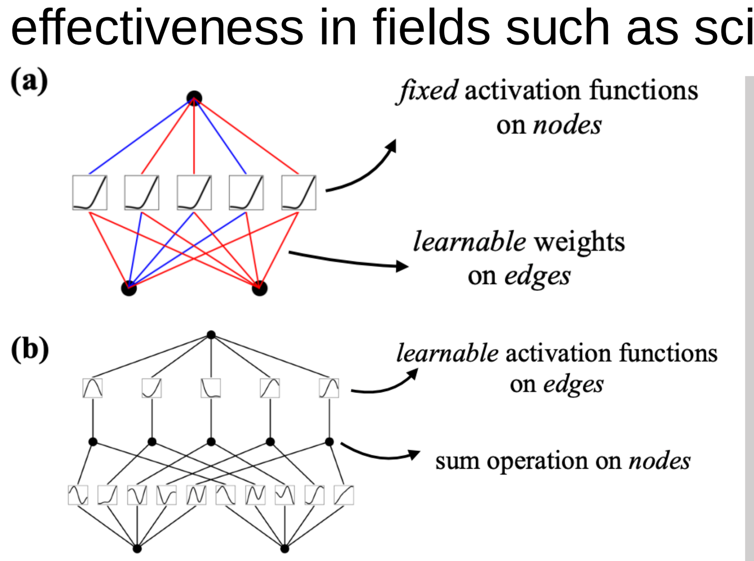


Figure 2. Comparison of MLPs and KANs structure

Compared to **MLP**, **KAN is potentially**

More Accurate

Fewer Parameters

Table 1. Comparison of previous studies and our approach

Previous Works	Ours
Small Dataset (Toy Dataset)	**Large** Dataset (Microsoft COCO)
Small Models (Simple Architecture)	**Large** Model (YOLO Framework)
Limited Evaluation	Comprehensive Evaluation

Object:
To enhance accuracy while reducing the network depth.

Results

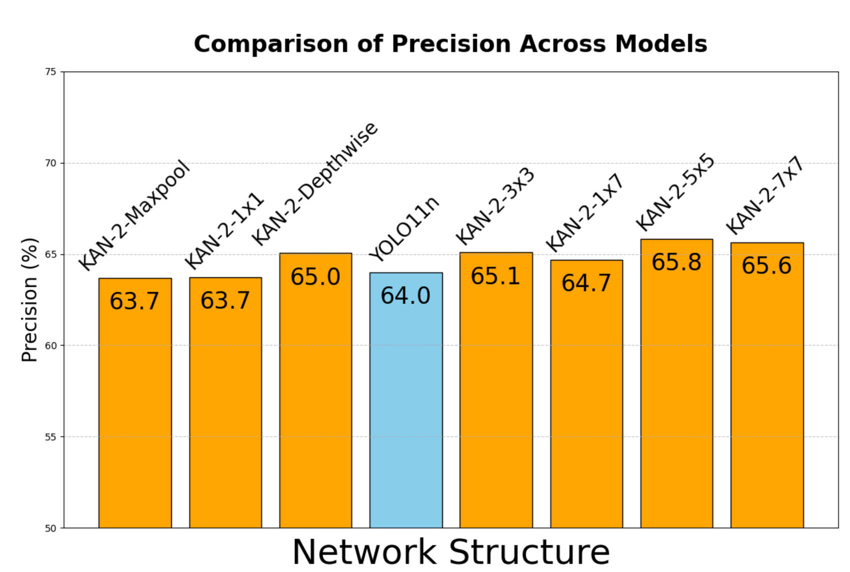


Figure 7. Comparison of Precision

Figure 7 demonstrates that replacing the KAN Module generally leads to an improvement in overall accuracy, with the highest increase reaching **1.84%.** This highlights the enhanced **feature extraction capability** of the KAN Module.

The structure of the flatten layer significantly influences the curve. While
a) KAN-2-5 achieves the highest accuracy,

b) KAN-2-7 demonstrates better balance on the AOC curve.

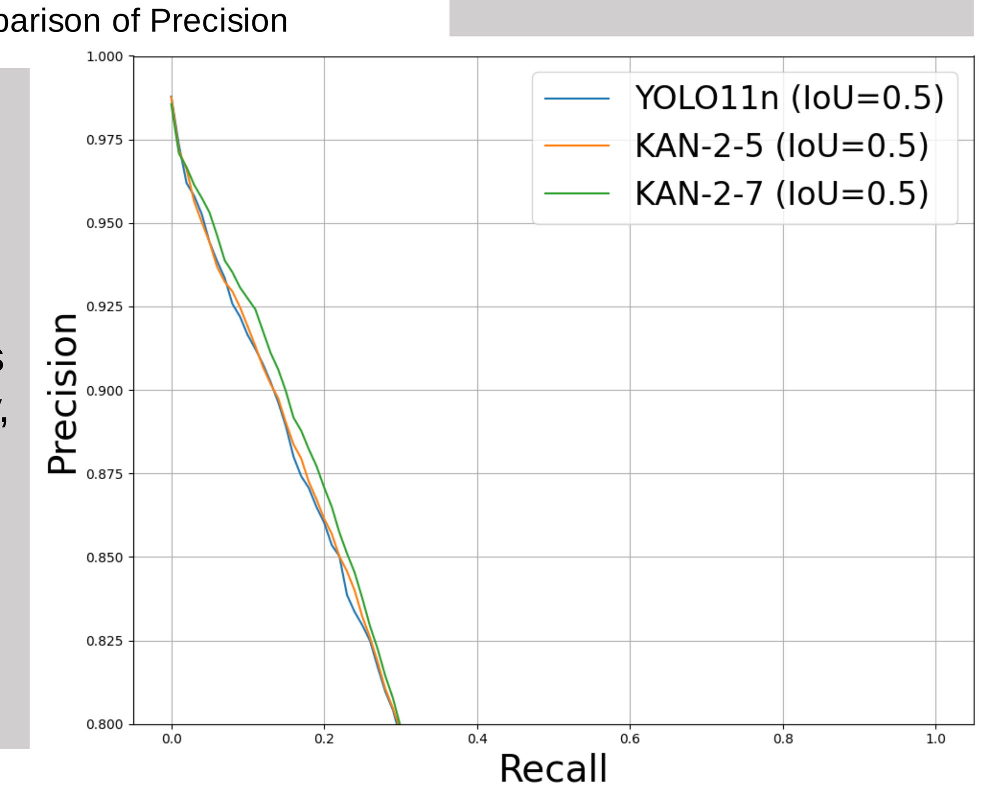
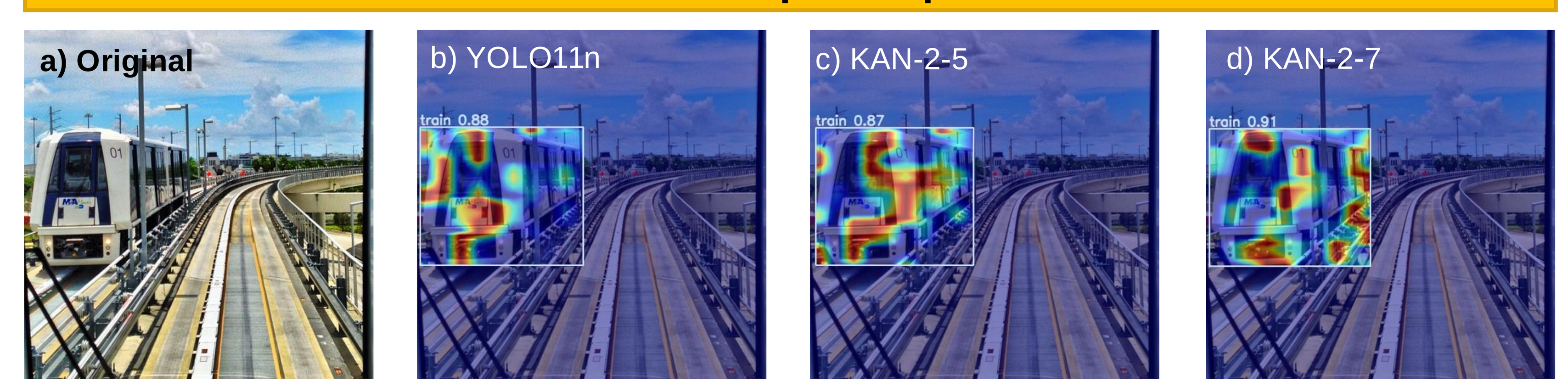


Figure 8. Comparison of PR curve

Dataset: Microsoft COCO
Base Model: YOLOv11n
Software Environment: Pytorch 12.4, CUDA 12.4
Hardware Environment: H100 x 1, 16-core i7 CPU, 64GB Memory

5. Heatmap Examples



Comparison of heatmap (a: Original Image; b: YOLO11n; c: KAN-2-5; d: KAN-2-7)

b) YOLO11n model, focusing on **key features** like the front and windows;
c) KAN-2-5 model, which concentrates on both the train and **background** regions;
d) KAN-2-7 model, covering the train and surroundings for **better contextual understanding**.
We also observe that as the kernel size increases, the expanded receptive field causes the model, especially KAN-2-7, to focus more on non-primary features. In our experiments, KAN-2-5 achieved the best balance.

Table 2. The overall comparison of all ablation experiments

Model	Precision (%)	Recall (%)	mAP@50 (%)	map@50-95 (%)
KAN-2-Maxpool	63.69	48.96	53.19	37.93
KAN-2-1	63.73	49.59	53.65	38.16
KAN-2-Depthwise	65.05	49.40	53.97	38.32
YOLO11n	63.99	49.49	53.85	38.38
KAN-2-3	65.08	48.94	54.00	38.52
KAN-1-7	64.69	**50.37**	**54.51**	**38.88**
KAN-2-5	**65.83**	49.24	54.36	38.87
KAN-2-7	65.64	49.44	54.48	38.77

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016.

[2] Z. Liu et al., "KAN: Kolmogorov-Arnold Networks," arXiv.org, 2024.

[3] Blealtan, "Blealtan/efficient-kan," GitHub, May 18, 2024. https://github.com/Blealtan/efficient-kan

[4] C. Li et al., "U-KAN Makes Strong Backbone for Medical Image Segmentation and Generation," arXiv.org, 2024.

CS 7980 Research Capstone