**YOLOV9**

**Các cải tiến lớn của v9**

1. Information Bottleneck Principle: Nguyên lý nút thắt thông tin

Trong DL, dữ liệu sẽ mất dần thông tin khi đi qua nhiều lớp. Vì vậy, YOLOv9 giải quyết vấn đề này bằng cách sử dụng PGI để bào toàn dữ liệu xuyên suốt các lớp

* Giúp Gradient ổn định, hiệu suất, sự hội tụ của model tốt hơn

1. Hàm đảo ngược

Hàm đảo ngược là hàm có thể khôi phục lại đầu vào từ đầu ra mà KHÔNG MẤT DỮ LIỆU.

X = v\_zeta(r\_psi(X))

Yolov9 kết hợp hàm này để giảm thiểu mất mát thông tin, đặc biệt là ở các lớp sâu.

1. Tác động đến các mô hình nhẹ

Nhờ PGI và hàm đâor ngược đã giải quyết được vấn đề này. Vì data của lightweight model ít, sl tham số. Nên v9 vẫn giữ được các th.tin quan trọng.

1. PGI Programmable Gradient Information – Thông tin Gradient có thể lập trình

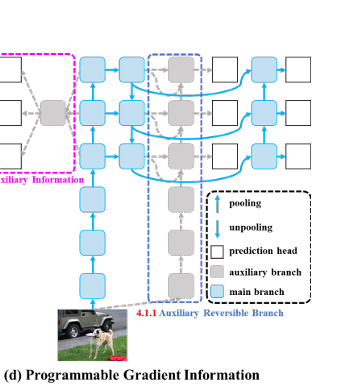
Là khái niệm mới do YOLOv9 giưới thiệu dùng để giải quyết nút thắt thông tin. Giúp ĐẢM BẢO BẢO TOÀN DỮ LIỆU trong các lớp.

Gồm 3 phần:

Main branch: dùng trong inference

Auxiliary reversible branch: dùng trong training, tạo ra grandient chính xác hơn

Multi-level auxiliary information:

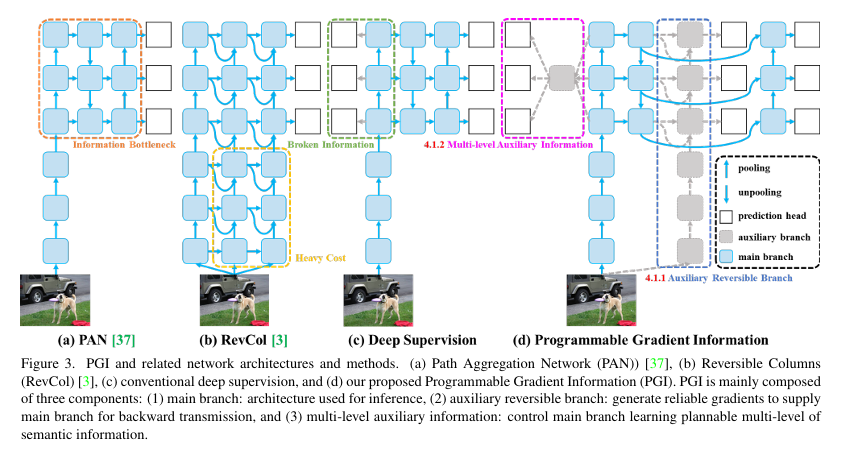


Ở hình 3.d Quá trình suy luận inference chỉ sử dụng ở nhánh chính ->Ko tạo nên chi phí khi tính toán. Hai phần auxiliary và prediction dùng để giải quyết hoặc làm chậm các vấn đề bên trong DL.

Auxiliary reversible branch (Nhánh phụ) Cung cấp thông tin ánh xạ từ dữ liệu đầu vào, hàm Loss có thể giúp tránh việc mô hình học sai từ những đặc trưng không liên quan trong quá trình FEEDFORWARD.

YOLOV9 đề xuất duy trì thông tin đầy đủ bằng cách sd kiến trúc đảo ngược reversible.

* Họ gắn reversible vào nhánh này nhằm tối ưu chi phí, vì nhánh chính tốn tiền, đại loại vậy.



MULTI LEVEL AUXILIARY INFORMATION

**Về nhận diện các vật thể nhỏ:**

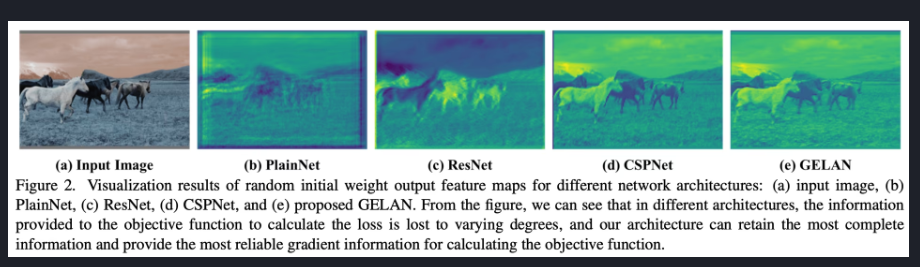
Mỗi layer học được tất cả thông tin từ các vật thể, kể cả nhỏ -> Ko bị lệch (4.1.2)

1. GELAN Generalized Efficient Layer Aggregation Network

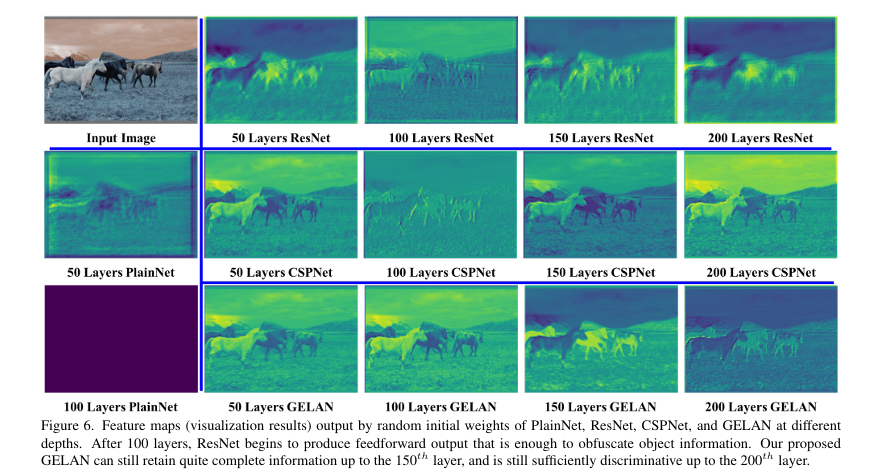
Biến thể dựa trên ELAN và CSPNet.

Là kiến trúc mới thuộc phần Backbone – giữ lại thông tin tốt hơn.

Thiết kế của GELAN cho phép tích hợp linh hoạt các khối tính toán khác nhau, đảm bảo cho việc YOLOv9 có thể thích ứng với nhiều ứng dụng mà kh ảnh hưởng tốc độ và độ chính xác.



Về môi trường mờ, thiếu ánh sáng, tối:



Trong bài báo có đề cập ở Fig6. GELAN có thể duy trì các thông tin quan trọng dù ở mốc 200 Layer.

Trong bài báo có so sánh GELAN với ELAN. Đại loại là ELAN nó có các khối cố định. Khó thay đổi.

GLAN cho phép đa dạng các loại block, không bị giới hạn như ELAN.

Rồi PGI tập trung vào vùng thực sự có chứa vật thể - có thể tránh phần bị chiếu sáng sáng.

TRONG MÔI TRƯỜNG VIDEO

Không có đề cập, chỉ test trên data COCO.