**Working of previous models in object detection**

1. **YOLOv11**

Problems in previous models:

* **Inefficient feature extraction:** Previous YOLO models use a variety of feature extraction modules, but these modules are not always efficient. This can lead to reduced detection accuracy and increased computational cost.
* **Limited versatility:** Previous YOLO models are typically designed for a specific task, such as object detection. This limits their versatility and makes them difficult to adapt to other tasks.
* **Difficulty in detecting small objects:** Previous YOLO models often have difficulty in detecting small objects. This is because small objects have fewer pixels, which makes them more difficult to distinguish from the background.

How it address previous model problem:

* **C3k2 block:** A more efficient version of the C3 module used in previous YOLO models.

Solves: addresses the issue of **redundant gradient flow**

**"Redundant gradient flow" means that during neural network training, the information used to update the network's parameters is being repeated unnecessarily, leading to less efficient learning.**

* **SPPF:** A spatial pyramid pooling-fast module that can aggregate multi-scale information more efficiently.

Solves: helps the model see objects at different sizes more effectively

* **C2PSA:** A cross stage partial attention module that can help the model to focus on the most important features.

Solves: focuses on enhancing feature representation by **selectively emphasizing informative features** across different stages of the network

**"Selectively emphasizing informative features" means the model focuses more on the most important parts of an image for accurate detection.**

New problems arises then:

* **Inefficient attention mechanisms:** Previous YOLO models use attention mechanisms, but these mechanisms are not always efficient.

1. **YOLOv12**

Problems in previous model:

* **Inefficient attention mechanisms:** Previous YOLO models use attention mechanisms, but these mechanisms are not always efficient.
* **Limited ability to detect small objects:** Previous YOLO models often have difficulty in detecting small objects. This is because small objects have fewer pixels, which makes them more difficult to distinguish from the background.

New features added:

* **Area Attention Mechanism:** A novel attention mechanism that reduces computational complexity and improves speed.

Works:

* + **Reduced Computational Load:** Instead of calculating attention weights for every single pixel or feature point, "Area Attention" divides the feature map into smaller, non-overlapping regions ("areas"). It then calculates attention weights for these areas. This drastically reduces the number of calculations required.
  + **Focus on Relevant Regions:** By calculating attention at the area level, the model can quickly identify and focus on the most relevant regions of the image where objects are likely to be present. This helps to filter out irrelevant background information and improve detection accuracy.

Solves: Calculating attention weights for image regions instead of individual pixels, thus maintaining accuracy while significantly improving speed.

* **R-ELAN(Residual ELAN):** A residual efficient layer aggregation network that enhances feature aggregation.
* Terms:
  + Residual:
    - here, means Residual connections
    - Residual connections act like "shortcuts" that allow information to bypass some layers of the neural network.
    - Issue solves: **Preventing Information Loss:** information can sometimes get lost or degraded as it passes through many layers. Residual connections help to mitigate this problem.
    - residual connections solves: By providing these shortcuts, residual connections make it easier for the network to learn complex patterns
  + ELAN(**Efficient Layer Aggregation):**
    - aggregates features from different layers of the network. This allows the model to capture both fine-grained details and high-level semantic information.
* Works:
  + R-ELAN combines both residual connections and ELAN.
  + Residual connections improve information flow(prevent data loss), while ELAN efficiently aggregates features.
* Benefits: learn better features and improve its object detection accuracy.