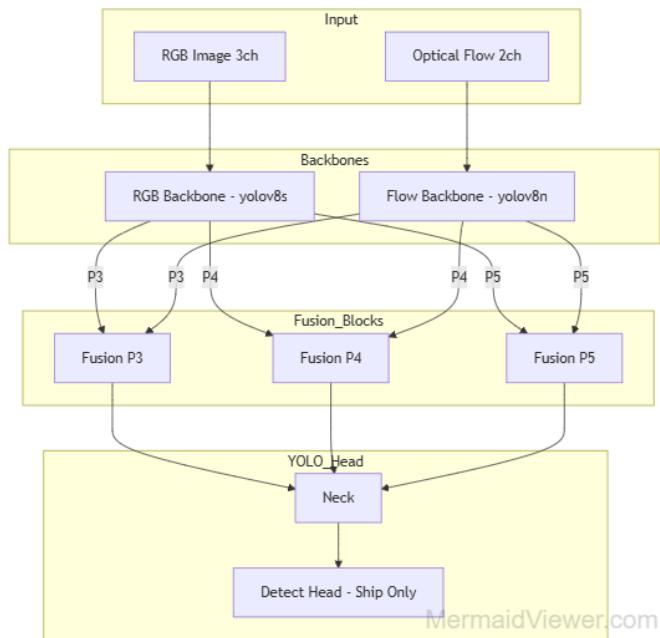


Architecture

1. 전체 구조도



2. 세부 컴포넌트 설계

A. Dual-Backbone (RGB & Flow Stream)

- RGB Stream (Spatial Encoder):** YOLOv8s의 백본을 활용하여 해수면, 선박의 선체, 구조물 등의 정적 시각 특징을 추출합니다.
- Flow Stream (Temporal Encoder):** YOLOv8n의 백본을 활용합니다. 입력 채널을 2채널(dx, dy)로 수정하여, 노이즈를 뚫고 움직이는 선박의 동역학적 경로 및 윤곽 특징을 추출합니다. 백본을 n(Nano) 모델로 선정하여 계산 부하를 최소화하면서도 핵심적인 움직임 정보를 보존했습니다.

B. 중간 계층 특징 융합 (Multi-scale Fusion)

단순히 마지막 결과만 합치는 것이 아니라, 신경망의 서로 다른 해상도 단계(P3, P4, P5)에서 정보를 결합합니다.

- P3 (Scale 1/8):** 작은 배나 먼 거리의 객체를 위한 고해상도 특징 융합.
- P4 (Scale 1/16):** 중간 크기 선박 추적용.

- **P5 (Scale 1/32):** 근거리 대형 선박 및 강인한 의미적(Semantic) 특징 융합.

C. Fusion Block의 내부 매커니즘

각 융합 지점에서는 다음과 같은 연산이 수행됩니다:

1. **Concatenation:** RGB 특징맵과 Flow 특징맵을 채널 방향으로 결합합니다.
2. **1×1 Convolution:** 두 정보 소스 사이의 상관관계를 학습하여 최적의 채널 가중치를 재 할당합니다.
3. **Stability Layer:** 배치 정규화(Batch Normalization)와 SiLU 활성화 함수를 통해 폭 풍우 노이즈로 인한 그래디언트 폭주를 방지하고 학습 안정성을 확보합니다.

D. 통합 탐지 헤드 (Unified Detection Head)

- 융합된 특징맵은 YOLOv8의 탐지 헤드로 입력됩니다.
- 본 아키텍처는 **End-to-End**로 학습되므로, 탐지 헤드에서 발생한 오차가 RGB와 Flow 양쪽 백본으로 역전파(Backpropagation)되어, 모델이 노이즈 환경에서 어떤 스트림을 더 신뢰해야 할지 스스로 학습하게 합니다.