**`Weekly report**

날짜 : **2022-11-03**

연구원 : **최윤석**

* **이전 수행계획**
* **차량 이미지 분류:**
  + YOLO 학습이 끝나면 2만개의 이미지 데이터를 Train : Test – 8 : 2의 비율로 나눠서 학습 진행
  + 학습된 모델을 사용해 이미지 검출할 때 발생하는 오류 해결 후 1.5만개의 데이터로 생성된 모델로 이미지에서 객체 검출
* **수행결과**
* **차량 이미지 분류 :** 
  + SSD 모델 학습 및 성능측정
    - 약 2만개의 이미지 데이터를 8 : 2로 나누어 15,000개의 학습 데이터와 3,800개의 테스트 데이터를 활용하여 Train / Evaluation 진행
      * 학습 epoch수는 2,000번과 20,000번 2가지를 진행해보았지만, 2,000번은 성능이 너무 좋지 않아 20,000번으로 진행한 결과만 넣었음

텍스트이(가) 표시된 사진

자동 생성된 설명

Figure 1. epoch 20,000번으로 학습한 모델의 Evaluation 결과 : 각 클래스의 average precision 값

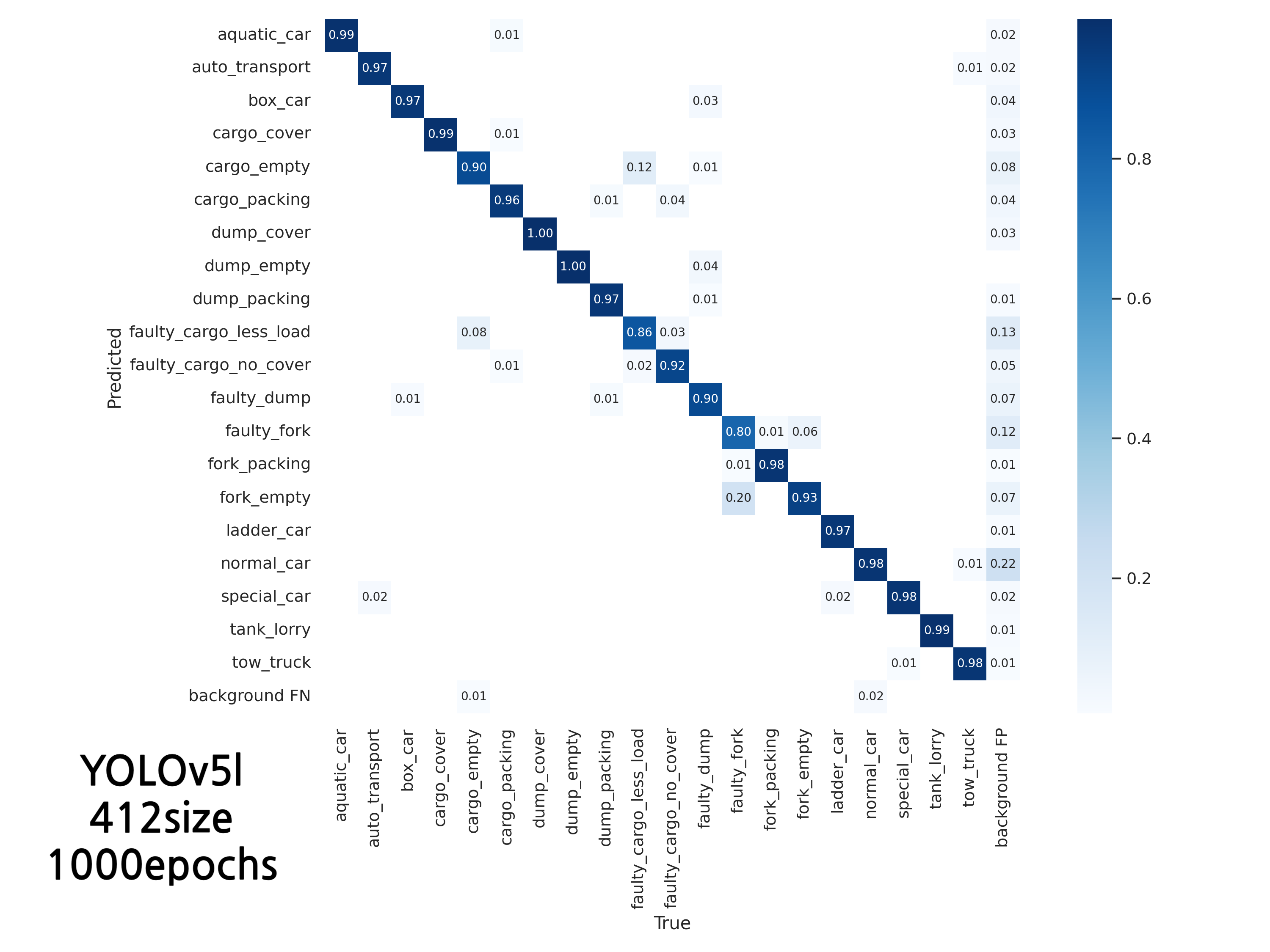


Figure 2. 김민중 연구원의 YOLO 모델 confusion matrix

* Average\_precision이 안 좋게 나온 라벨(4, 5, 9, 11, 12, 14)이 김민중 연구원의 confusion matrix에서 성능이 안 좋게 나온 라벨(4, 9, 10, 11, 12, 14)과 비슷함.

테이블이(가) 표시된 사진

자동 생성된 설명

Figure 3. SSD 모델에서 Averag\_precision이 안 좋게 나온 라벨의 국문명을 노란색으로 표시한 표



Figure 4. 전체 클래스들의 mean average precision 값

* + - * 15,000개의 학습데이터를 epoch를 2000으로 진행했을 때 mAP 값이 0.1이었고, epoch를 20000으로 진행했을 때 mAP 값이 0.77까지 오른 것으로 보아 epoch 수를 더 늘려서 학습하면 더 좋은 성능이 나올 것으로 기대됨.

텍스트이(가) 표시된 사진

자동 생성된 설명

Figure 5. 라벨 19번 – 견인차를 제대로 감지한 모습

실내, 컴퓨터이(가) 표시된 사진

자동 생성된 설명

Figure 6. 라벨 3번 – 카고 덮개형을 제대로 감지한 모습

* + - 학습된 모델을 사용해 이미지를 검출할 때 발생하는 오류 해결 후, 객체 검출을 각 라벨당 2개씩 뽑은 테스트 데이터 40개에 진행한 결과 bounding box를 제대로 잡아내는 것을 확인

Figure 7. 발생하는 오류 메시지

* **관계형 테이블 임베딩** :
  + SSD 오류 해결전까지 휴지
* **수행계획**
* **차량 이미지 분류** :
  + 모델의 성능 향상을 위해 epoch와 다른 요소를 변경하여 학습 진행
* **관계형 테이블 임베딩** :
  + 수치형 데이터를 이산화한 것을 범주형 데이터로 바꾸는 임베딩 방식을 보완하는 방법을 찾고 실험하여, 기존의 윤종찬 연구원의 연구와 성능 비교
* **기타사항**
  + **데이터 융합 프로세스**





혼합형 데이터셋 예시