



УНИВЕРСИТЕТ ИТМО

# Детектирование и классификация тарелок.

Работу выполнили:  
Павлов Валентин

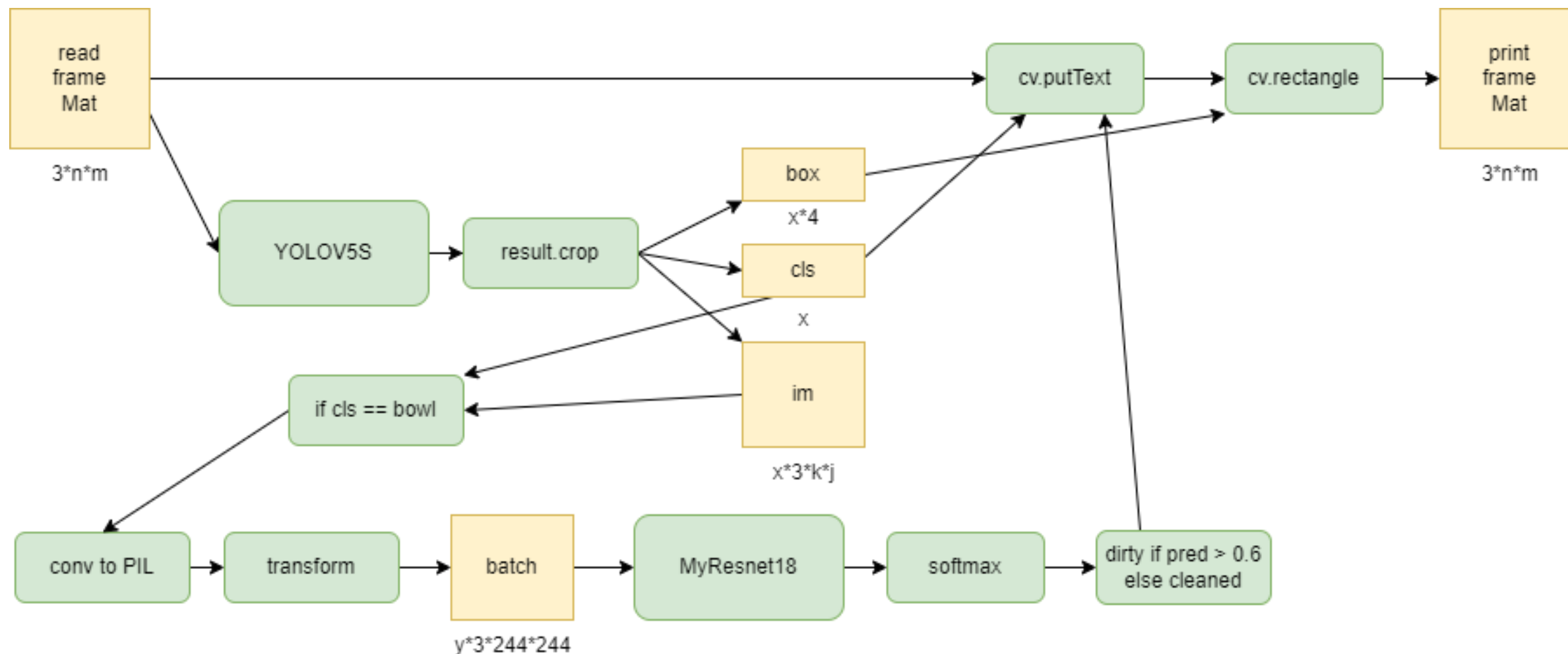
## Цель и применение

Цель: разработать нейросеть для детектирования тарелок и классификации на грязные и чистые.

Пример использования: по конвейеру едут тарелки после мойки и необходимо определить отмылись тарелки или нет.



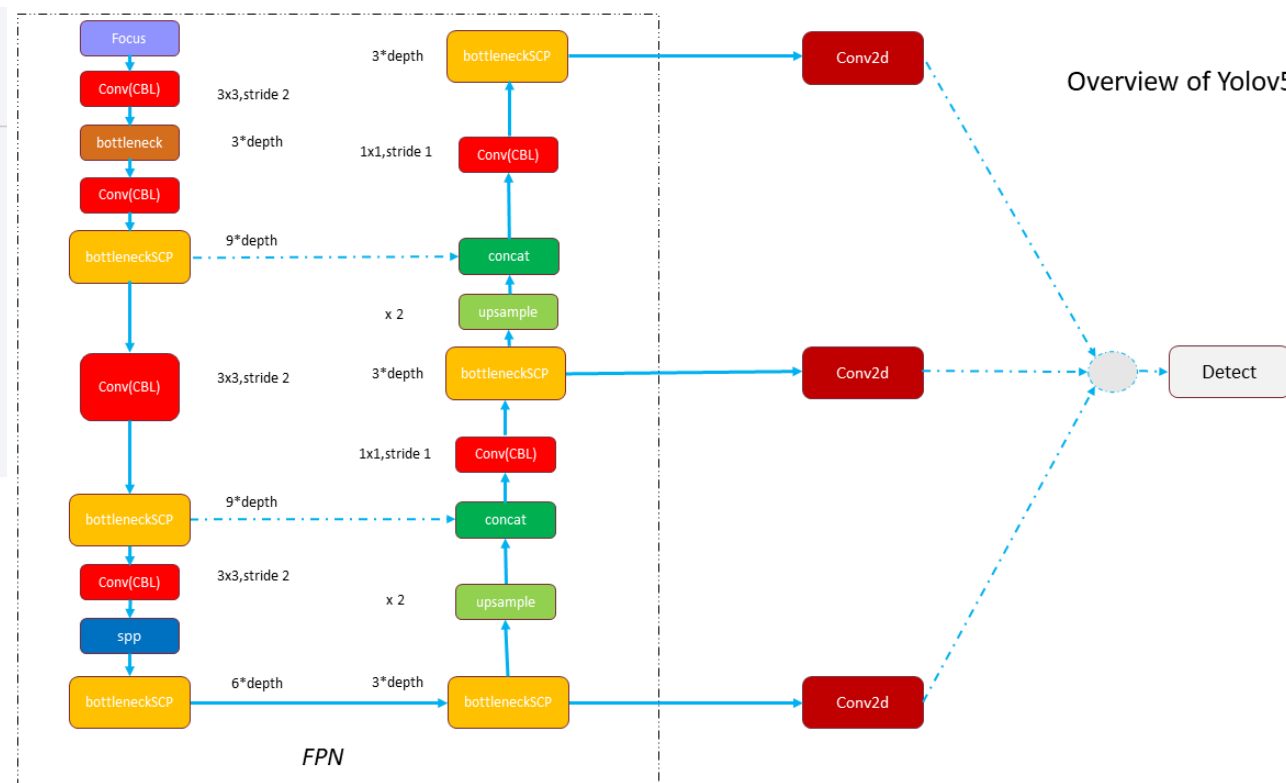
# Архитектура приложения



# Модели и их характеристики моделей

Model	size (pixels)	mAP <sup>val</sup> 0.5:0.95	mAP <sup>test</sup> 0.5:0.95	mAP <sup>val</sup> 0.5	Speed V100 (ms)	params (M)	FLOPS 640 (B)
YOLOv5s6	1280	43.3	43.3	61.9	<b>4.3</b>	12.7	17.4
YOLOv5m6	1280	50.5	50.5	68.7	8.4	35.9	52.4
YOLOv5l6	1280	53.4	53.4	71.1	12.3	77.2	117.7
YOLOv5x6	1280	<b>54.4</b>	<b>54.4</b>	<b>72.0</b>	22.4	141.8	222.9
YOLOv5x6 TTA	1280	<b>55.0</b>	<b>55.0</b>	<b>72.0</b>	70.8	-	-

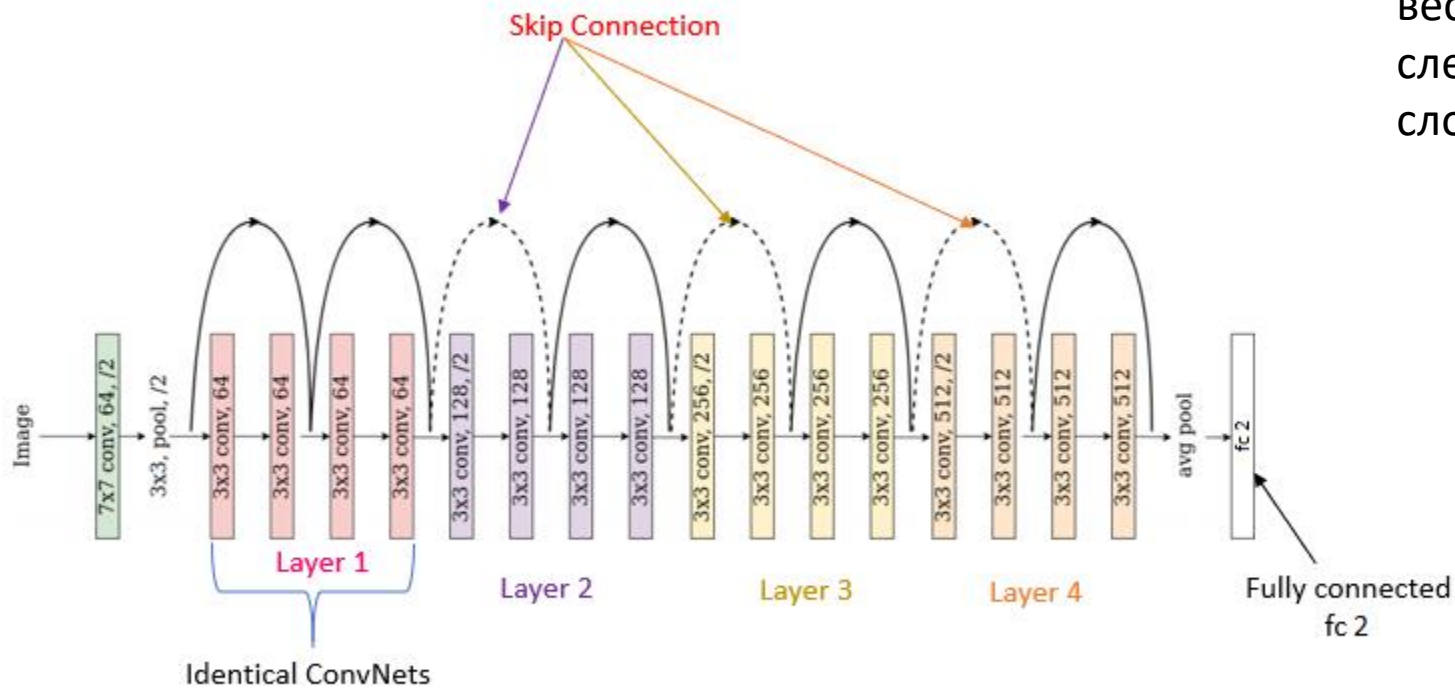
Для детектирования и первичной классификации объектов используется обученная модель YOLOV5, а именно версия YOLOV5(small) поскольку это самая быстрая модель если верить данным разработчика, а для работы в реальном времени это самый важный параметр. Сама модель взята с хаба моделей pytorch.



[https://pytorch.org/hub/ultralytics\\_yolov5/](https://pytorch.org/hub/ultralytics_yolov5/)

# Модели и их характеристики моделей

Для классификации тарелок на чистые и грязные используется отдельная модель. Это предобученная модель ResNet18 с измененными полно-связными (FC) слоями, веса всех слоёв кроме FC фиксируются и следовательно необходимо обучить только FC слои.



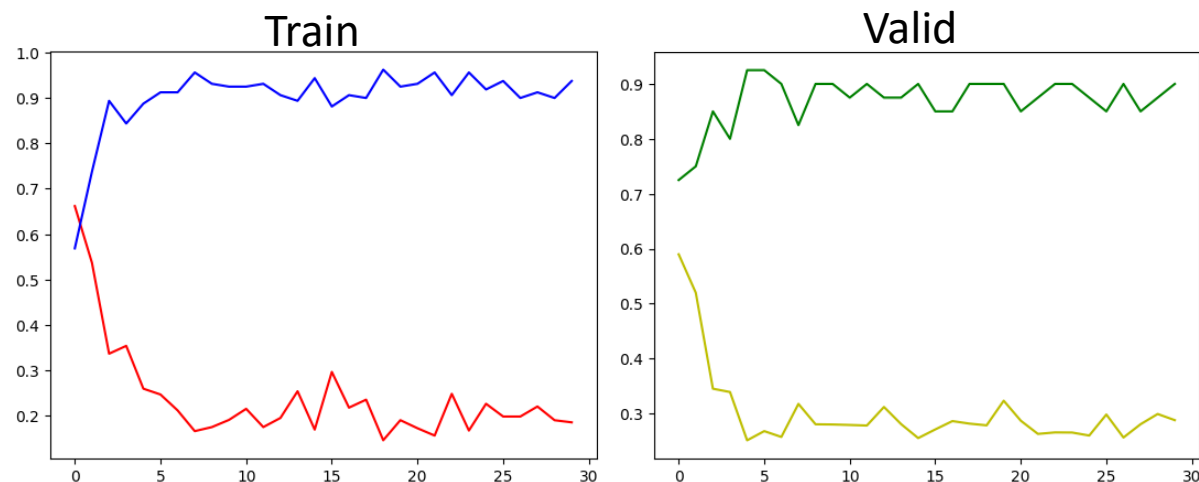
ResNet-18 Architecture

Input Image size=224 \* 224px

# Модели и их характеристики моделей

Для обучения используется датасет тарелок с сайта kaggle, для тренировок и валидации размечены по 100 грязных и чистых тарелок, каждое 5 изображение используется для валидации, остальное для тренировки. К изображениям для тренировки применяется набор случайных аугментаций для большей эффективности обучения. Обученная модель сохраняется как jit скрипт и загружается уже в основной программе, jit скрипт позволяет не определять модель при загрузке.

Loss фикция и ассурасу при обучение модели:



Также был проведён тест на 100 новых изображениях из того же дата сета, и получен уровень достоверности предсказания 92 – 93%, в сравнение с ResNet50 у которого 94 - 96%

# Тестирование





# Тестирование

Модель плохо классифицирует тарелки с сложным узором или рисунком:



Также замечена особенность что при близком цвете тарелки и фону модели сложно детектировать тарелку:







# Тестирование

Потребление памяти всей программой около 1800 MiB, установлено с помощью утилиты `memory_profiler` с соответствует моим приблизительным расчётам и данным из мониторинга ресурсов windows.

Быстродействие программы можно оценить в FPS (кадров, которые программа обрисовывает за секунду).

1. Без использования модели FPS счётчик показывает 29 – 27 кадров, поскольку входящий поток с камеры установлен на 30 кадров.
2. При использовании только модели YOLO программа показывает приблизительно 22 - 20 кадров.
3. При использовании обеих моделей программа показывает приблизительно 16 кадров, в кадре 2 тарелки.
4. При использовании обеих моделей программа показывает приблизительно 10 - 9 кадров, в кадре 7 тарелки.