

Описание экспериментов

В рамках исследования было проведено 6 экспериментов: обучены 3 модели YOLO11 “nano”, “small” и “medium” в режимах 30 и 50 эпох.

Блюдам были присвоены классы:

- 0: "ribs",
- 1: "chop",
- 2: "salad1",
- 3: "salad2",
- 4: "bortsch",
- 5: "pumpkin soup"

Ниже приведены результаты валидации моделей:

“nano”

30 эпох

```
Ultralytics 8.3.162 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210M)
YOLO11n summary (fused): 100 layers, 2,583,322 parameters, 0 gradients, 6.3 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.737	0.795	0.857	0.817
ribs	40	40	0.771	1	0.995	0.995
chop	40	40	0.77	1	0.995	0.995
salad1	40	40	0.98	1	0.995	0.995
salad2	40	40	0.603	1	0.995	0.986
bortsch	40	40	0.875	0.351	0.756	0.671
pumpkin soup	40	40	0.421	0.417	0.406	0.259

50 эпох

```
Ultralytics 8.3.162 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210M)
YOLO11n summary (fused): 100 layers, 2,583,322 parameters, 0 gradients, 6.3 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.975	0.863	0.932	0.887
ribs	40	40	0.945	1	0.995	0.995
chop	40	40	0.958	1	0.995	0.995
salad1	40	40	0.967	1	0.995	0.995
salad2	40	40	0.979	1	0.995	0.995
bortsch	40	40	1	0.781	0.978	0.891
pumpkin soup	40	40	1	0.398	0.635	0.451

100 эпох

```
Ultralytics 8.3.162 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210M)
YOLO11n summary (fused): 100 layers, 2,583,322 parameters, 0 gradients, 6.3 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.914	0.866	0.952	0.893
ribs	40	40	0.747	1	0.995	0.995
chop	40	40	0.954	1	0.995	0.995
salad1	40	40	0.96	1	0.995	0.995
salad2	40	40	0.922	1	0.995	0.995
bortsch	40	40	1	0.722	0.978	0.896
pumpkin soup	40	40	0.904	0.472	0.752	0.482

200 эпох

```
Ultralytics 8.3.162 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11n summary (fused): 100 layers, 2,583,322 parameters, 0 gradients, 6.3 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.96	0.902	0.962	0.904
ribs	40	40	0.854	1	0.995	0.995
chop	40	40	0.957	1	0.995	0.995
salad1	40	40	0.967	1	0.995	0.995
salad2	40	40	0.98	1	0.995	0.995
bortsch	40	40	1	0.938	0.993	0.908
pumpkin soup	40	40	1	0.474	0.796	0.536

“small”

30 эпох

```
Ultralytics 8.3.162 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11s summary (fused): 100 layers, 9,415,122 parameters, 0 gradients, 21.3 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.939	0.874	0.961	0.927
ribs	40	40	0.759	1	0.995	0.995
chop	40	40	0.958	1	0.995	0.995
salad1	40	40	0.985	1	0.995	0.989
salad2	40	40	0.935	1	0.995	0.995
bortsch	40	40	1	0.714	0.993	0.957
pumpkin soup	40	40	1	0.53	0.792	0.632

50 эпох

```
Ultralytics 8.3.162 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11s summary (fused): 100 layers, 9,415,122 parameters, 0 gradients, 21.3 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.963	0.936	0.981	0.942
ribs	40	40	0.977	1	0.995	0.995
chop	40	40	0.955	1	0.995	0.992
salad1	40	40	0.875	1	0.995	0.995
salad2	40	40	0.978	1	0.995	0.995
bortsch	40	40	1	0.867	0.98	0.948
pumpkin soup	40	40	0.996	0.75	0.924	0.728

100 эпох

```
Ultralytics 8.3.162 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11s summary (fused): 100 layers, 9,415,122 parameters, 0 gradients, 21.3 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.976	0.908	0.984	0.937
ribs	40	40	0.965	1	0.995	0.995
chop	40	40	0.976	1	0.995	0.995
salad1	40	40	0.977	1	0.995	0.995
salad2	40	40	0.976	1	0.995	0.994
bortsch	40	40	1	0.828	0.99	0.968
pumpkin soup	40	40	0.961	0.62	0.932	0.673

200 эпох

```
Ultralytics 8.3.162 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11s summary (fused): 100 layers, 9,415,122 parameters, 0 gradients, 21.3 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.894	0.929	0.97	0.928
ribs	40	40	0.585	1	0.995	0.995
chop	40	40	0.945	1	0.995	0.995
salad1	40	40	0.976	1	0.995	0.995
salad2	40	40	0.977	1	0.995	0.948
bortsch	40	40	0.94	1	0.99	0.948
pumpkin soup	40	40	0.942	0.575	0.848	0.688

“medium”

30 эпох

```
Ultralytics 8.3.162 🚀 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11m summary (fused): 125 layers, 20,034,658 parameters, 0 gradients, 67.7 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.901	0.871	0.92	0.868
ribs	40	40	0.725	1	0.995	0.995
chop	40	40	0.961	1	0.995	0.995
salad1	40	40	0.979	1	0.995	0.995
salad2	40	40	0.98	1	0.995	0.954
bortsch	40	40	0.857	1	0.992	0.946
pumpkin soup	40	40	0.901	0.229	0.547	0.321

50 эпох

```
Ultralytics 8.3.162 🚀 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11m summary (fused): 125 layers, 20,034,658 parameters, 0 gradients, 67.7 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.824	0.898	0.927	0.894
ribs	40	40	0.613	1	0.995	0.995
chop	40	40	0.954	1	0.995	0.995
salad1	40	40	0.977	1	0.995	0.995
salad2	40	40	0.976	1	0.995	0.995
bortsch	40	40	1	0.878	0.991	0.933
pumpkin soup	40	40	0.421	0.51	0.592	0.453

100 эпох

```
Ultralytics 8.3.162 🚀 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11m summary (fused): 125 layers, 20,034,658 parameters, 0 gradients, 67.7 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.97	0.853	0.935	0.907
ribs	40	40	0.891	1	0.995	0.995
chop	40	40	0.953	1	0.995	0.995
salad1	40	40	1	1	0.995	0.994
salad2	40	40	0.977	1	0.995	0.995
bortsch	40	40	1	0.997	0.995	0.948
pumpkin soup	40	40	1	0.122	0.637	0.514

200 эпох

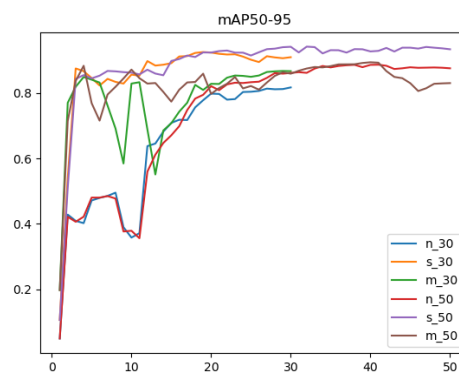
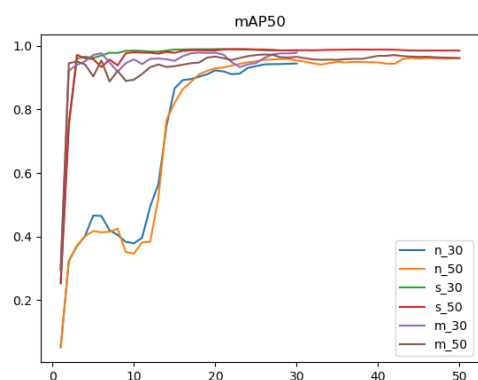
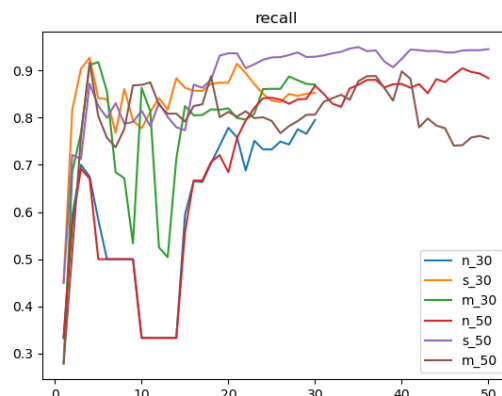
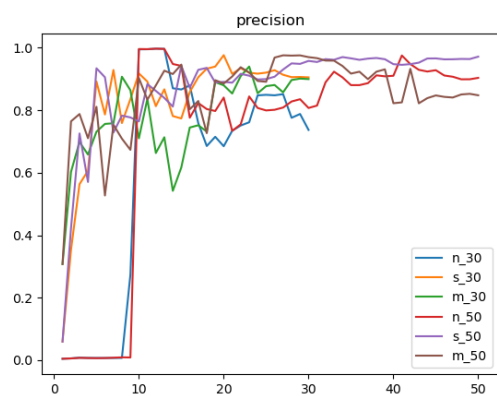
```
Ultralytics 8.3.162 🚀 Python-3.12.2 torch-2.5.1 CUDA:0 (NVIDIA GeForce RTX 4090, 24210MiB)
YOLO11m summary (fused): 125 layers, 20,034,658 parameters, 0 gradients, 67.7 GFLOPs
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	77	240	0.897	0.961	0.985	0.918
ribs	40	40	0.57	1	0.995	0.995
chop	40	40	0.959	1	0.995	0.995
salad1	40	40	1	0.928	0.995	0.984
salad2	40	40	0.984	1	0.995	0.995
bortsch	40	40	0.897	1	0.995	0.88
pumpkin soup	40	40	0.971	0.838	0.937	0.659

Видно, что несмотря на большую емкость модели “medium” не помогает ей в работе с нашим тривиальным датасетом – похоже, она начинает выучивать специфические признаки, мешающие обобщению. Усиление аугментаций не улучшает качество.

Для всех моделей удлинение обучения приводит к улучшению, но на 200 эпохах происходит насыщение, и далее улучшение не наблюдается. Эти результаты получены со стандартным режимом обучения YOLO11.

Также были отслежена динамика основных метрик в процессе обучения:



Видно, что модель “nao” ввиду малой информационной емкости нуждается в большем количестве итераций, чтобы начать показывать хороший результат, но затем догоняет: все-таки задача тривиальная.

Выводы:

1. Задача в данной постановке слишком легкая даже для самой маленькой модели.
2. Блюда, которые практически не меняют “вид сверху” в процессе еды представляют наибольшую сложность.
3. После определенного уровня увеличение емкости модели не улучшает результат.