

Распознавание текста

Герасимов Илья

Содержание

- Датасет
- Классический ML
- VLM

Датасет

RusTitW*



train: 24366p

test:

3795
unique values



true
3098 82%
false
697 18%



true
1311 35%
false
2484 65%



true
1316 35%
false
2479 65%

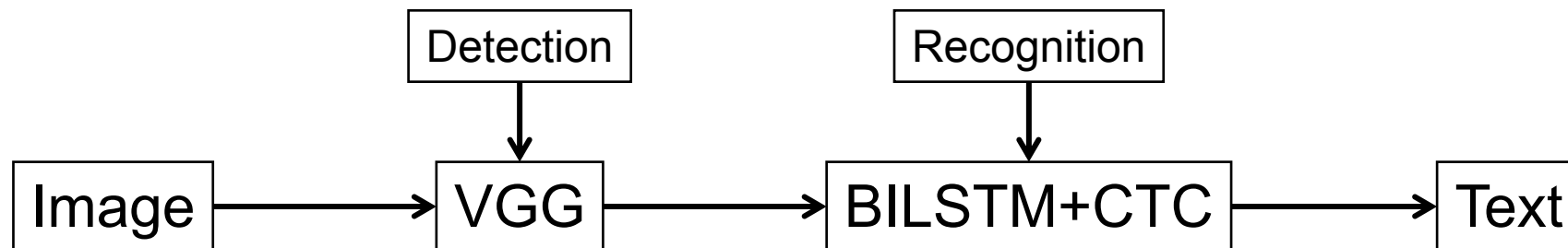


true
1567 41%
false
2228 59%

*<https://www.kaggle.com/datasets/hardtype/rustitw-russian-language-visual-text-recognition>

Классический ML

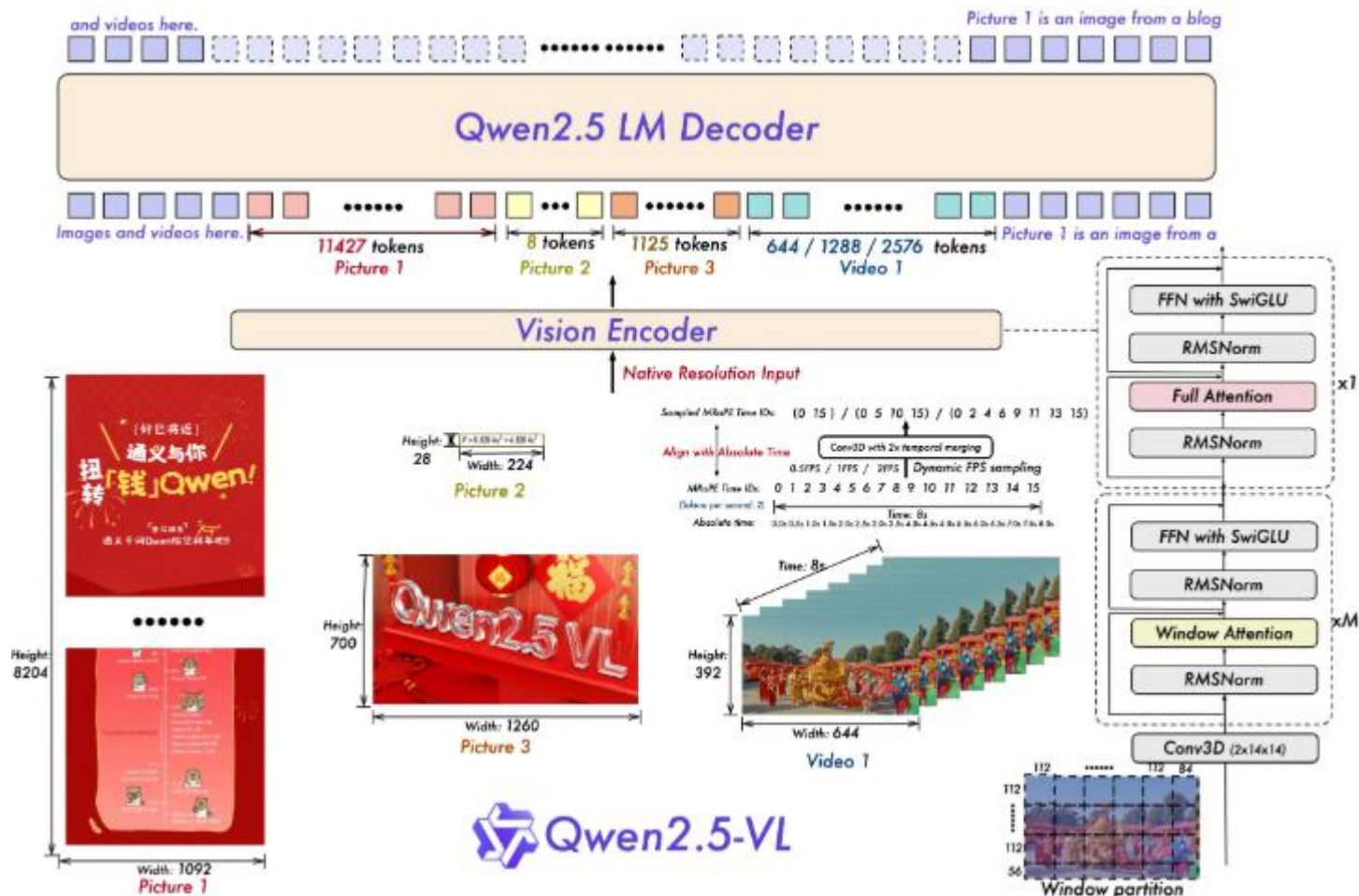
EasyOCR*



*<https://github.com/JaidedAI/EasyOCR>

VLM

QWEN2.5-VL*

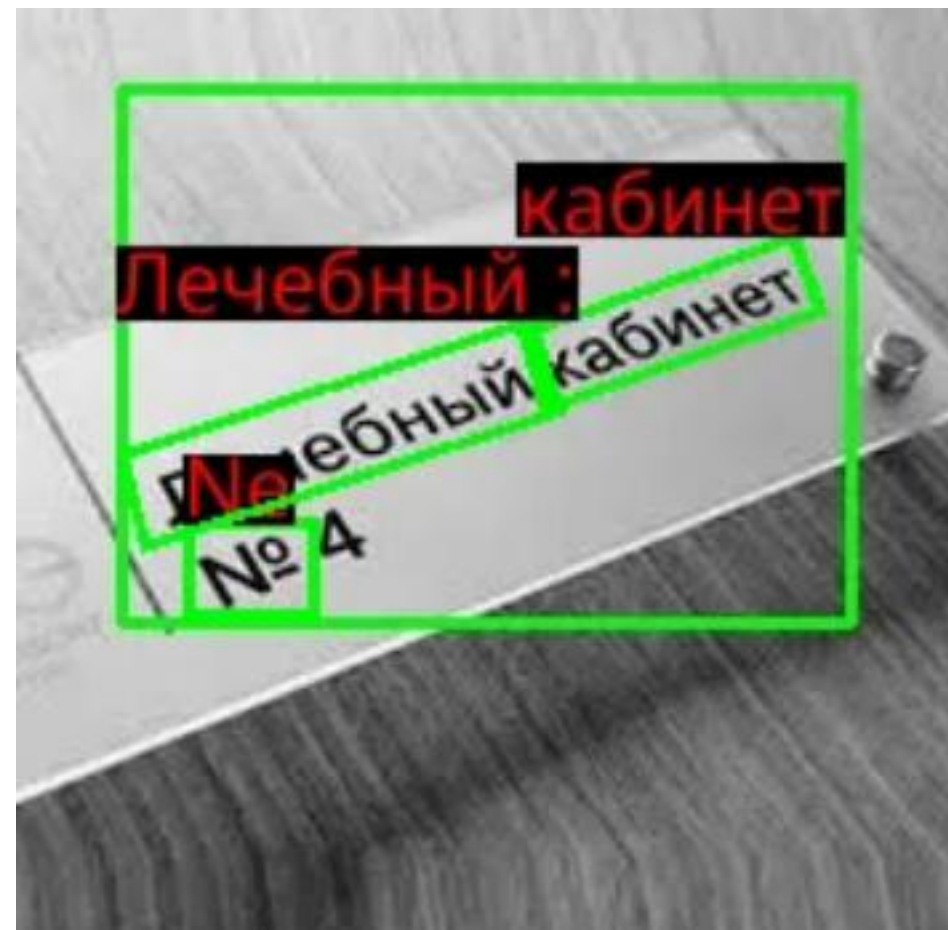


*<https://github.com/QwenLM/Qwen2.5-VL>

Базовые случаи



QWEN-VL2.5



EasyOCR

Базовые случаи



QWEN-VL2.5



EasyOCR

Базовые случаи



QWEN-VL2.5



EasyOCR

Метрика

$$\text{CER}(x, y) = \frac{\textit{Leventstain_distance}(x, y)}{y.\textit{size}()}$$

| | EasyOCR | Qwen2.5-VL |
|-----|---------|------------|
| CER | 1.0116 | 0.94 |