

СИСТЕМА РАСПОЗНАВАНИЯ ДОРОЖНЫХ ЗНАКОВ НА БАЗЕ ОДНОПЛАТНОГО КОМПЬЮТЕРА LUCKFOX PICO

Шендик Д.В., студент

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Вашкевич М.И. – д-р техн. наук, профессор

Работа посвящена разработке программно-аппаратного комплекса для распознавания дорожной обстановки в реальном времени. В качестве вычислительной платформы используется одноплатный компьютер с интегрированным нейронным процессором (NPU). Рассмотрен процесс подготовки специализированного набора данных, обучение модели на легковесной архитектуре YOLOv5n и многоступенчатой конвертации модели для аппаратного ускорения инференса. Приведены результаты распознавания дорожной обстановки и оценка метрик эффективности нейросетевой модели. Разработанное решение ориентировано на интеграцию в современные системы помощи водителю (ADAS).

Введение. Современные системы помощи водителю (ADAS) требуют высокопроизводительной и энергоэффективной обработки визуальных данных в режиме реального времени. Классические подходы, основанные на мощных графических ускорителях (GPU), не применимы во встраиваемых решениях из-за жестких ограничений по габаритам, тепловыделению и энергопотреблению. Решением данной проблемы является переход к использованию специализированных систем на кристалле (SoC) с интегрированными нейронными процессорами (NPU), которые обеспечивают аппаратное ускорение тензорных вычислений при минимальном энергозатратах.

Целью работы является проектирование и программная реализация оптимизированной системы распознавания дорожных знаков на базе платформы Luckfox Pico Max (SoC RV1106). Данная платформа оснащена встроенным NPU с производительностью до 1 TOPS, что позволяет эффективно выполнять задачи компьютерного зрения в условиях ограниченных вычислительных ресурсов. Достижение поставленной цели требует комплексного подхода: от подготовки специализированного набора данных до глубокой аппаратной оптимизации нейросетевой модели под архитектуру выбранного кристалла.

Постановка задачи и обзор набора данных. Система должна локализовать дорожные знаки в видеопотоке (формировать ограничивающую рамку) и классифицировать их тип. Итоговый результат с наложенной разметкой транслируется в сеть посредством реализованного RTSP-сервера, что позволяет удаленно получать видеопоток с результатами распознавания в реальном времени.

Для обучения модели использовался открытый набор данных RTSD (Russian Traffic Sign Dataset). В ходе подготовки датасет был программно модифицирован под требования целевой архитектуры YOLO и уменьшен с 179138 изображений и аннотации к ним до 59188 изображений – удалены лишние изображения, не имеющие дорожные знаки. Ключевым этапом предобработки стало преобразование абсолютных пиксельных координат рамок в нормализованные относительные значения. Пересчет координат выполняется по формуле (1).

$$x_{norm} = \frac{x_{min} + \frac{w}{2}}{W}, y_{norm} = \frac{y_{min} + \frac{h}{2}}{H}, w_{norm} = \frac{w}{W}, h_{norm} = \frac{h}{H}, \quad (1)$$

где x_{norm}, y_{norm} – нормированные координаты центра объекта; w_{norm}, h_{norm} – нормированные ширина и высота объекта относительно размеров всего кадра; x_{min}, y_{min} – координаты левого верхнего угла ограничивающей рамки; w, h – ширина и высота ограничивающей рамки в пикселях; W, H – общая ширина и высота всего входного изображения в пикселях.

Данное преобразование делает вычисления независимыми от разрешения исходного кадра. Например, ограничивающая рамка [730, 445, 16, 18] для изображения размером 1280×720 пикселей преобразуется в вектор [0.576563, 0.630556, 0.012500, 0.025000].

Архитектура и обучение модели. В качестве базовой архитектуры была выбрана нейронная сеть YOLOv5n (nano) – наиболее легковесная версия в семействе YOLO. Выбор обусловлен эффективным балансом между количеством параметров и точностью детекции. Архитектурно модель состоит из трех главных компонентов: магистральной сети (Backbone) на базе CSPDarknet для извлечения признаков, связующей части (Neck) на основе PANet для агрегации признаков на разных масштабах и головной части (Head) для генерации итоговых предсказаний классов и координат.

Процесс обучения модели выполняется в среде фреймворка PyTorch с применением метода переноса обучения (Transfer Learning) на основе предобученных весов, на протяжении 100 эпох. Первоначальное обучение проводилось на стандартном разрешении 640×640 пикселей для извлечения максимально точных пространственных признаков объектов. Однако для обеспечения высокой скорости инференса на платформе RV1106 целевая модель была переобучена и

экспортирована для входного разрешения 320×320 пикселей, чтократно снизило нагрузку на аппаратную часть.

Конвертация модели под аппаратную среду. Запуск модели на NPU платформы Luckfox Pico требует специфический формат нейросетевых графов RKNN. Адаптация обученной модели осуществлялась в несколько этапов. Первоначально веса экспортировались в промежуточный стандарт ONNX (Open Neural Network Exchange), затем, с использованием инструментария RKNN-Toolkit2, модель конвертировалась в формат RKNN. Ключевым этапом является пост-тренировочное квантование, которое переводит веса модели из формата чисел с плавающей точкой (FP32) в 8-битный целочисленный формат (INT8). Данное преобразование описывается формулой линейного квантования (2).

$$q = \text{round} \left(\frac{f}{S} + Z \right), \quad (2)$$

где f – значение в FP32; q – квантованное значение INT8; $S(Scale)$ – коэффициент масштабирования; $Z(Zero - point)$ – смещение.

Квантование позволяет сократить объем занимаемой памяти в 4 раза и задействовать аппаратные блоки умножения-сложения (MAC) нейропроцессора, что критически важно для достижения real-time производительности на чипе RV1106.

Результаты распознавания и оценка метрик. Эффективность разработанной системы оценивалась на тестовой выборке с использованием стандартных метрик компьютерного зрения. Precision (Точность), Recall (Полнота), mAP@50 (Mean Average Precision).

Метрики модели приведены на рисунке 1.

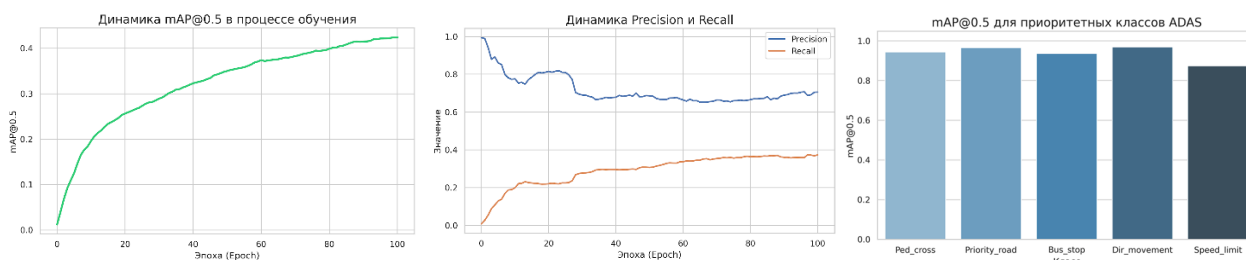


Рисунок 1 – Метрики модели

Результаты распознавания дорожных знаков приведены на рисунке 2.

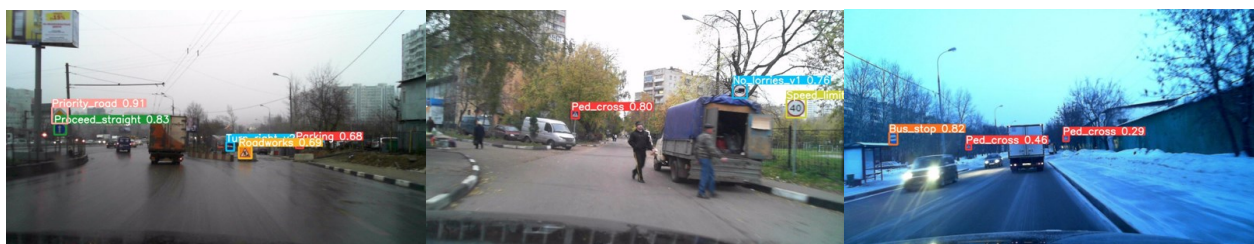


Рисунок 2 – Результаты распознавания дорожных знаков

Заключение. Использование специализированных систем на кристалле с интегрированными нейронными процессорами (NPU) является оптимальным вектором развития автомобильных периферийных вычислений (Edge AI). Применение платформы Luckfox Pico в связке с легковесной архитектурой YOLOv5n обеспечивает высокую энергоэффективность, тепловую стабильность и способность обрабатывать видеопоток в режиме реального времени без привлечения массивных графических процессоров. Ключевым преимуществом данного подхода выступает возможность прямого встраивания нейросетевой аналитики в малогабаритные системы помощи водителю. Дальнейшее повышение качества и надежности распознавания требует сфокусированной работы над обучающей выборкой: в перспективе планируется глубокая доработка используемого датасета за счет синтетической генерации данных, добавления редких классов знаков и включения сцен со сложными погодными условиями.

Список использованных источников:

1. Ultralytics YOLOv5 [Электронный ресурс]. – Режим доступа: <https://github.com/ultralytics/yolov5>. – Дата доступа: 10.04.2026.
2. Luckfox Pico RV1106/RV1103 SDK [Электронный ресурс]. – Режим доступа: <https://wiki.luckfox.com/>. – Дата доступа: 10.04.2026.