

Адаптация модели гауссовского процесса для решения задачи оценки неопределенности семантической сегментации железнодорожных сцен

Куртеев Д. А.¹

Научный руководитель – канд. техн. наук, доцент Забашта А. С.¹

Научный руководитель – начальник центра систем компьютерного зрения

Раков Д. А.²

¹Университет ИТМО

²АО «НИИАС»

dima.kurteev.work@gmail.com

Введение

В автономном железнодорожном транспорте модуль восприятия программно-аппаратного комплекса должен работать предсказуемо и безопасно, так как его результаты используются для определения возможных маршрутов следования и обнаружения подвижных единиц на пути локомотива. Одной из ключевых задач восприятия является семантическая сегментация изображения с установленной камеры, где каждый пиксель относится к одному из классов: фон, железнодорожная колея, подвижной состав.

В области сертификации аппаратно-программных комплексов для автономного железнодорожного транспорта, связанных с использованием нейронных сетей, сертификационный орган выдвинул требования, соответствующие стандарту ПНСТ 836-2023 [4]. Это означает необходимость внедрения и разработки «функций наблюдения», которые позволят контролировать работу нейронных сетей и, при необходимости, блокировать передачу результатов их работы в модули бизнес-логики системы. Для таких архитектур требуется формирование дополнительного сигнала, который поступает в функцию безопасности и позволяет принять решение: применять ограничивающую логику или использовать результат сегментации без ограничений. В качестве такого сигнала целесообразно использовать оценку неопределенности выходов модели семантической сегментации [1]. Для оценки неопределенности выбран гауссовский процесс, поскольку он дает точечное предсказание и дисперсию, используемую в модуле проверки безопасности. Применение классического гауссовского процесса ограничено из-за высокой вычислительной сложности и требований к памяти, поэтому в работе рассматриваются масштабируемые аппроксимации гауссовского процесса [2].

Техническим заданием работы является реализация системы контроля безопасности выходного сигнала нейронной сети для задачи семантической сегментации железнодорожных сцен за счет исследования и реализации алгоритма оценки неопределенности и функционала модуля проверки безопасности.

Цель работы – улучшить корректность калибровки вероятностных предсказаний выходов модели семантической сегментации железнодорожных сцен при сохранении качества сегментации (mIoU) и обеспечить формирование сигнала для модуля проверки безопасности.

Основная часть

Исходные данные – датасет железнодорожных сцен (изображения) с разметкой по трем классам: фон, железнодорожная колея, подвижной состав; базовая модель

сегментации – SegFormer.

Реализована система контроля безопасности выходов SegFormer, включающая модуль оценки неопределенности на базе гауссовского процесса и модуль проверки безопасности, принимающий решение о применении ограничивающей логики при повышенной неопределенности предсказаний. Предварительная экспериментальная оценка показала улучшение метрики калибровки ECE на 33% и рост качества сегментации на 1 п.п. mIoU по сравнению с базовой моделью. Улучшение калибровки повышает корректность интерпретации неопределенности и, как следствие, надежность работы модуля проверки безопасности. Для достижения указанных результатов были выполнены следующие работы: (1) адаптация и сравнение методов оценки неопределенности для SegFormer на данных железнодорожного домена по метрикам mIoU, ECE и NLL, а также по времени выполнения и потреблению видеопамати; (2) разработка и реализация модуля проверки безопасности, принимающего решение о блокировке/применении ограничивающей логики, с оценкой качества по AUC ROC и FPR@TPR; (3) исследование масштабируемых аппроксимаций гауссовского процесса для $N \geq 100000$, $d \geq 128$ (вычислительная сложность и затраты по времени/памяти, вычисление ковариационной матрицы и решение линейных систем), разработка модифицированного варианта, интегрируемого в GPyTorch [3].

Выводы

Работа направлена на построение системы контроля безопасности результатов нейронной сети семантической сегментации железнодорожных сцен за счет использования оценки неопределенности выходов модели в модуле проверки безопасности. Для повышения корректности калибровки вероятностных предсказаний при сохранении качества семантической сегментации и приемлемых вычислительных затрат рассматривается применение масштабируемых аппроксимаций гауссовского процесса.

Литература

1. Liu J. Z., Padhy S., Ren J. и др. A Simple Approach to Improve Single-Model Deep Uncertainty via Distance-Awareness [Электронный ресурс]. — URL: <https://www.jmlr.org/papers/volume24/22-0479/22-0479.pdf> (дата обращения: 12.02.2026).
2. Rasmussen C.E., Williams C.K.I. Gaussian Processes for Machine Learning [Электронный ресурс]. — 2006. — URL: <https://gaussianprocess.org/gpml/chapters/RW.pdf> (дата обращения: 12.02.2026).
3. GPyTorch Documentation [Электронный ресурс]. — URL: <https://docs.gpytorch.ai/en/stable/> (дата обращения: 12.02.2026).
4. ПНСТ 836-2023. Искусственный интеллект. Функциональная безопасность и системы искусственного интеллекта [Электронный ресурс]. — URL: <https://protect.gost.ru/document1.aspx?control=31&id=256756> (дата обращения: 12.02.2026).