Методы обеспечения функциональной безопасности инновационных систем на примере СККСП

Попов П.А., Родиманова Е.С., Раков Д.А.
СПбФ АО «НИИАС»

НИИАС

Актуальность доказательства безопасности систем технического зрения

НИИАС

Кадровый дефицит на сети железных дорог
Угроза стабильности и эффективности перевозок в средне- и долгосрочной перспективе

Автоматизация функций машиниста
Необходима замена с воспроизведением функционала на достаточном уровне безопасности

Сложности верификации систем технического зрения
Низкая степень интерпретируемости и отсутствие формализованных подходов к верификации

Существующие проблемы доказательства безопасности систем технического зрения

НИИАС

1 **Отсутствие четких требований стандартов**
ГОСТ Р МЭК 61508 и ГОСТ Р МЭК 62279 не покрывают системы технического зрения

11 **таблиц требований**

29 **требований обязательны для УПБ2**

63 **требования обязательны для УПБ3**

Нейронные сети как «чёрный ящик»: поведение сложно объяснить и предсказать

Нет прослеживаемости между слоями сети и требованиями

Обучение нейронных сетей стохастично: результаты нестабильны

Качество работы зависит от обучающей выборки

Пример таблицы с методами, которые необходимо применять на этапе детального проектирования

ГОСТ Р МЭК 61508-3-2018, таблица А.4

Метод / средство <1>	Ссылка	УПБ1	УПБ2	УПБ3	УПБ4
1a Методы структурных диаграмм <2>	C.2.1	HR	HR	HR	HR
1a Методы структурных диаграмм <2>	Таблица B.7	R	R	HR	HR
1c Формальные методы проектирования и усовершенствования <2>	B.2.2, C.2.4	-	R	R	HR
2 Средства автоматизированного проектирования	B.3.5	R	R	HR	HR
3 Программирование с защитой	C.2.5	-	R	HR	HR
4 Модульный подход	Таблица B.9	HR	HR	HR	HR
5 Стандарты для проектирования и кодирования	C.2.6, таблица B.1	R	HR	HR	HR
6 Структурное программирование	C.2.7	HR	HR	HR	HR
7 Использование доверительных/проверенных программных модулей и компонентов (по возможности)	C.2.10	R	HR	HR	HR

3

Существующие проблемы доказательства безопасности систем технического зрения

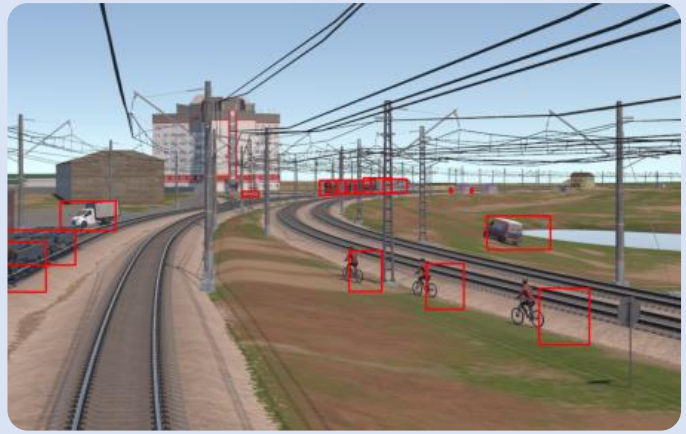
НИИАС

2 **Безграничность возможных сценариев**

Для оценки требуются наборы данных, покрывающие все возможные ситуации

Тестирование в реальных условиях небезопасно, дорого и долго

Тестирование с примирением симулятора: требует подтверждения соответствия реальном условиям и сертификации используемого инструмента



4

Существующие проблемы доказательства безопасности систем технического зрения

НИИАС

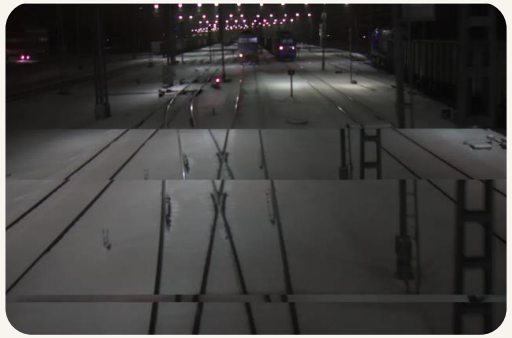
3 **Входные данные влияют на работу**

Даже незначительные отклонения во входной информации могут радикально повлиять на поведение алгоритма

Технические сбои

Физические явления

Ошибки алгоритмов обработки



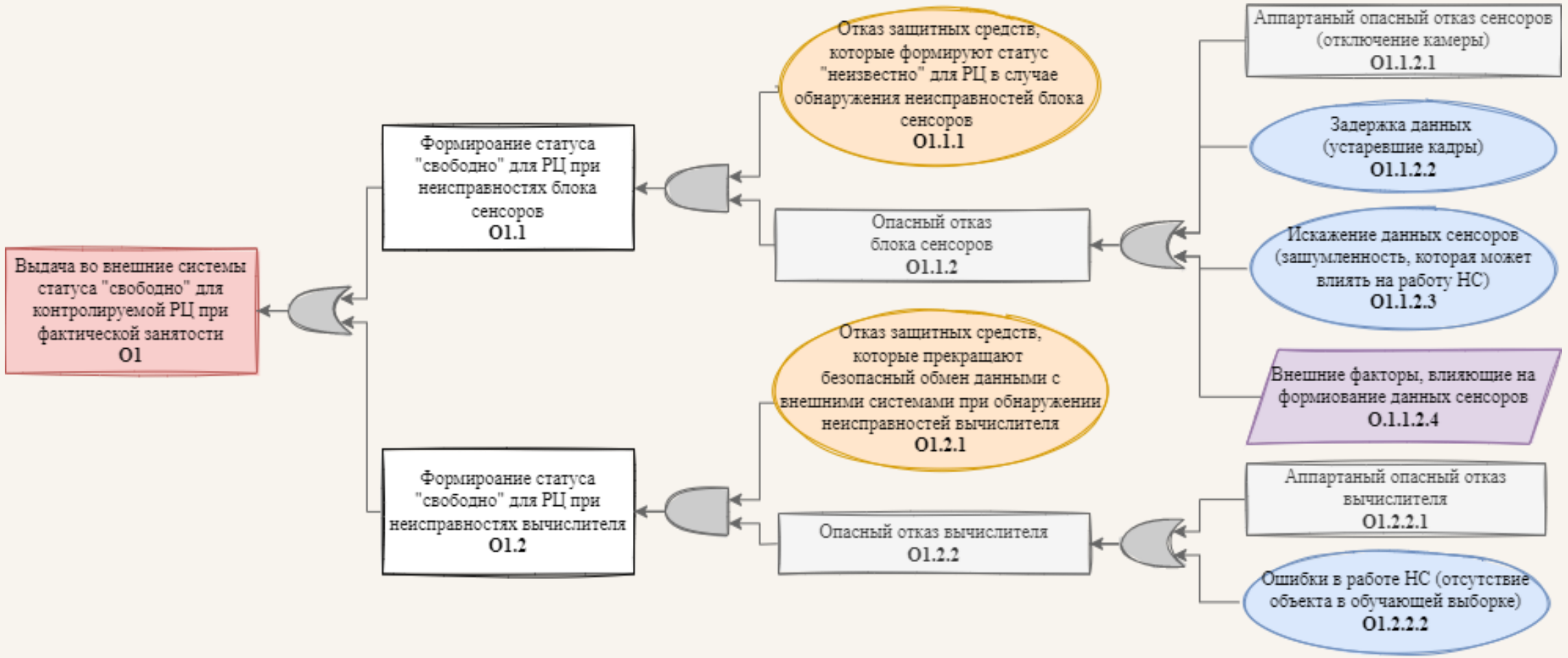
5

Назначение системы

СККСП – стационарный комплекс контроля свободности пути

СККСП предназначен для **обеспечения автоматизации** процессов производства **маневровой работы**, снижения влияния человеческого фактора при **контроле свободности пути при движении вагонами вперед**, исключения «мертвых зон» контроля, контролирует рельсовые цепи (РЦ), которые не централизованы

Возможные отказы системы



Состав системы

Программная часть

- нейронные сети - алгоритмы, поведение которых не явно, полностью задается обучающими данными
- классическое ПО - алгоритмы с четкой логикой, не привязанные к входным данным
- data-drive алгоритмы - алгоритмы с четкой логикой, построенной на базе статистики

реализация контролирующих функций

Аппаратная часть

- блок вычислителя
- блоки сенсоров
  - камеры
  - лидары

Функции системы

Основные функции системы (**функции назначения**):

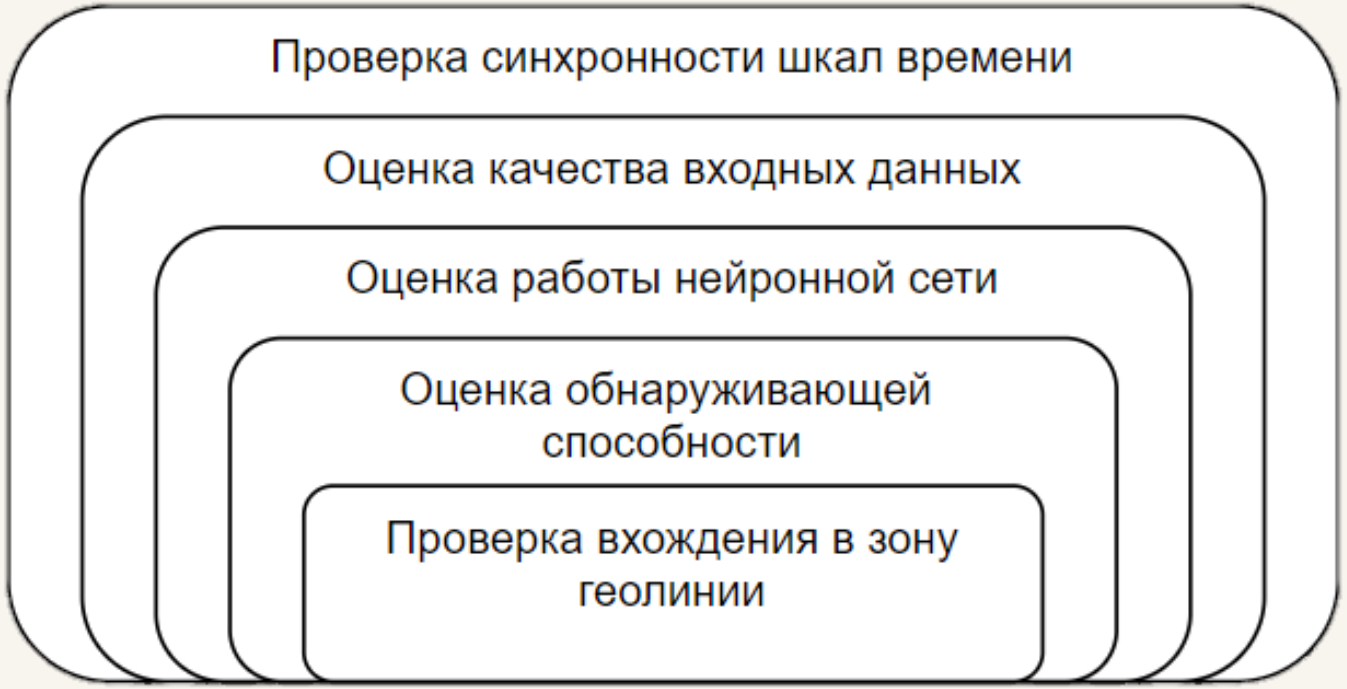
1. Формирование состояния контролируемых зон РЦ (не централизованных)
    - обнаружение объектов (вагон, человек)
    - определение зоны нахождение (красная, желтая)
2. Выдача состояния контролируемых зон РЦ внешним системам (в ЦАМР - центр автоматизации маневровой работы)

Слои безопасности

Работа начинается со слоя проверки синхронизации шкал времени, затем оценка качества входных данных и т.д.

Каждый слой имеет таймер для контроля цикличности поступления данных

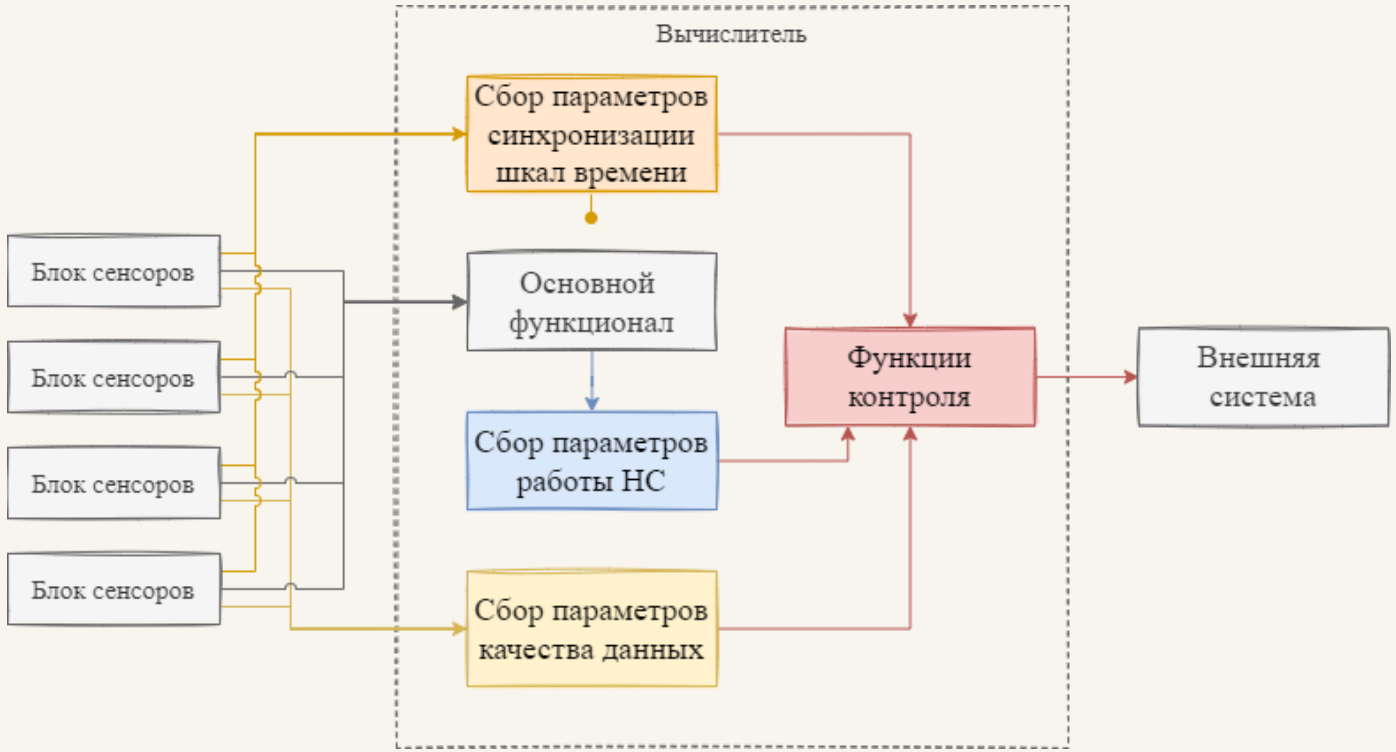
Если не выполнена проверка на текущем слое – переход на следующий слой не осуществляется



Функции системы

Функции безопасности системы (контролирующие функции):

1. Контроль правильности формирования состояния элементов маршрута
2. Прекращение выдачи ответственной информации внешним системам в случае обнаружения неисправностей в работе



Функции безопасности системы

5. Модуль ФБ ПО принимает решение на основании полученных метрик – он соотносит их с заранее заданными граничными значениями

Название метрики	**Что обозначает**	**Диапазон значений**	**Как интерпретировать**	**Какую часть системы контролирует**
data_quality	Оценка качества данных поступающих с сенсора	от 0 до +6 мс	Чем больше значение метрики тем дальше данные находятся от исходного распределения	Часть системы отвечающую за сбор и дальнейшую передачу данных для обработки
ntp_status	Статус NTP синхронизации между сервером обработки и сервером времени	бинарное [0,1]	Если статус синхронизации есть, значит мы можем быть уверены в корректности интерпретации данных внешними подсистемами	Корректность взаимодействия с внешними подсистемами
delay	Размер задержки между меткой времени данных и временем сервера	от -6 мс до +6 мс	Если размер отставания по модулю достигает критической отметки то данные с такой проблемой не могут отражать реальную картину	Часть системы отвечающую за сбор и дальнейшую передачу данных для обработки

Функции безопасности системы

4. Модуль ФБ ПО принимает решение на основании полученных метрик – он соотносит их с заранее заданными граничными значениями

Название метрики	**Что обозначает**	**Диапазон значений**	**Как интерпретировать**	**Какую часть системы контролирует**
detection_ capability	Оценка обнаруживающей способности модуля детектирования	от 0 до 1	Указывает на наихудшее значение полученное при сравнении заранее заданных координат объектов(статических) с предсказаниями нейронной сети Если нейронная сеть не смогла обнаружить статические объекты на сцене, то такое поведение расценивается как отказ	Блок обнаружения объектов типа “Человек” и “неопознанное препятствие”
semantic_c apability	Оценка обнаруживающей способности модуля сегментации	от 0 до 1		Блок обнаружения объектов типа “Вагон”
nn_quality	Оценка неопределенности нейронной сети в ответов	от 0 до 1	Максимальное значение неопределенности ответа среди текущих ответов нейронной сети Наличие ответов с высокой неопределенностью рассматривается как нетипичное поведение нейронной сети	Блоки обнаружения объектов типа “Человек” ,“неопознанное препятствие” и “Вагон”

Защитное состояние

В случае, если проверка одного из слоев была не пройдена – система переходит в **защитное состояние**

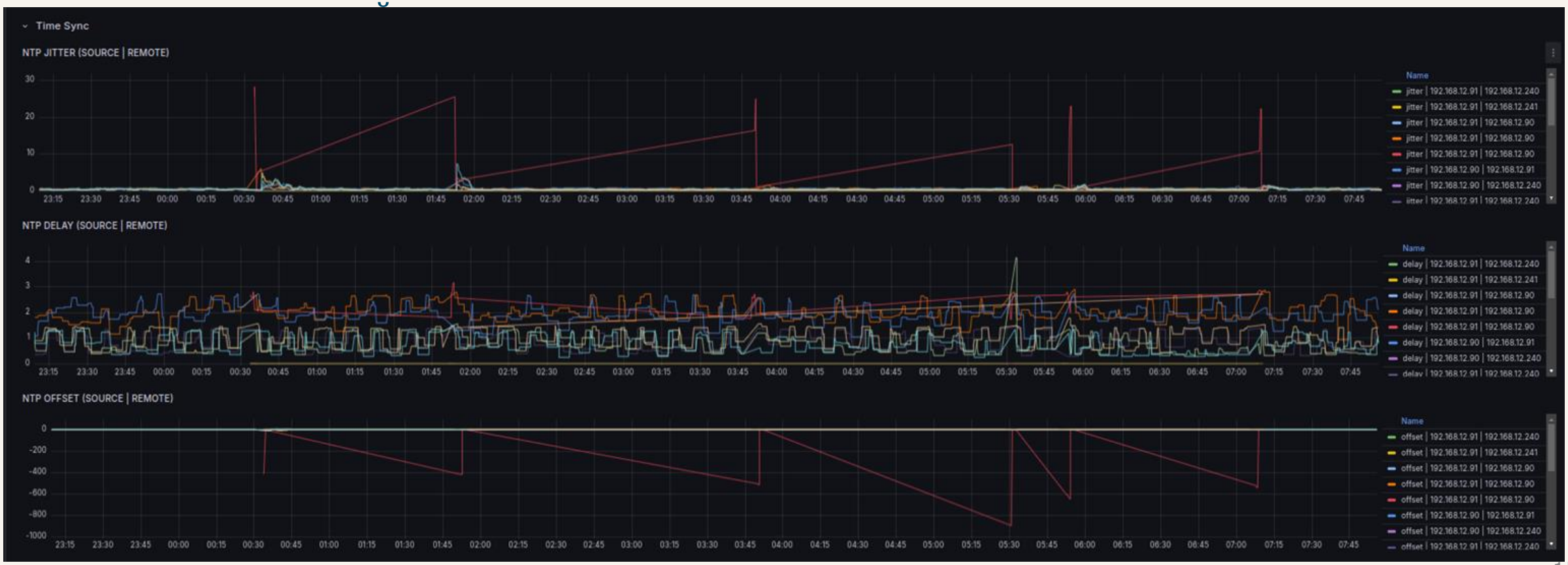
Функция системы	Контроль правильности исполнения	Действия при возникновении отказа
Ф1	ФК1.1 Проверка синхронности шкал времени компонентов системы со шкалой вычислителя	ФД1.1 Формирование статуса "неизвестно"[28] для зоны, которая попадает в область видимости сенсора с выходом параметров синхронизации шкал времени за пределы допустимых
	ФК1.2 Оценка достоверности и целостности входных данных сенсоров	ФД1.2 Формирование статуса "неизвестно" для зоны, которая попадает в область видимости сенсора с выходом параметром оценки корректности входных данных за пределы допустимых
	ФК1.3 Оценка корректности работы нейросетевых модулей для выбранных ракурсов	ФД1.3 Формирование статуса «неизвестно» для зоны, которая попадает в область видимости сенсора, по итогу оценки корректности работы нейросетевых моделей на выбранном ракурсе которого произошел выход параметров за границы допустимых
	ФК1.4 Оценка способности системы обнаруживать требуемые объекты для выбранных ракурсов	ФД1.4 Формирование статуса «неизвестно» для зоны, которая попадает в область видимости сенсора, где отсутствует обнаружение требуемых объектов
	ФК1.5 Проверка вхождения объектов в зоне рельсовой цепи	ФД1.5 Формирование статуса «неизвестно» для зоны, которая попадает в область видимости сенсора, где обнаружено расхождение
Ф2	ФК2.1 Проверка синхронности шкалы времени вычислителя с единым сервером времени	ФД2.1 Прекращение безопасного обмена данными с ЦАМР
	ФК2.2 Проверка средств диагностики, фиксирующих параметры работы аппаратных компонентов вычислителя	ФД2.2 Прекращение безопасного обмена данными с ЦАМР

[28] Здесь и далее по тексту подразумевается, что осуществляется переход в защитное состояние и его удержание до момента, пока не придет команда от машиниста-оператора на выход из защитного состояния и переход в рабочее состояние будет доступен.

Метрики синхронизации шкал времени

Слой, отвечающий за проверку синхронности шкал времени, **на вход получает параметры (метрики) синхронизации вычислителя и сенсоров СККСП относительно единого сервера времени**

В случае **выхода за граничные значения** параметров синхронизации для конкретных

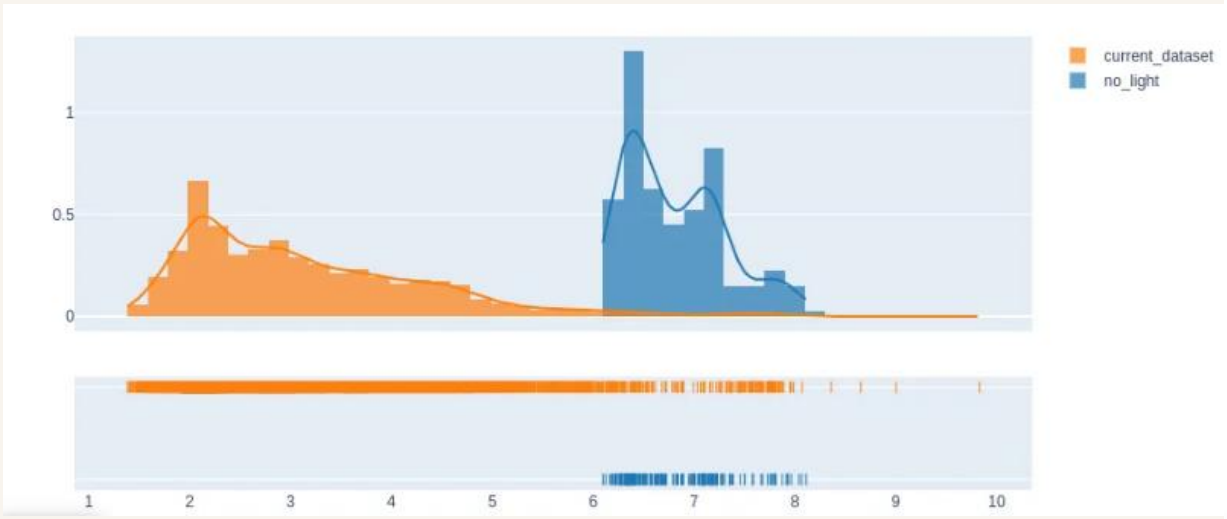


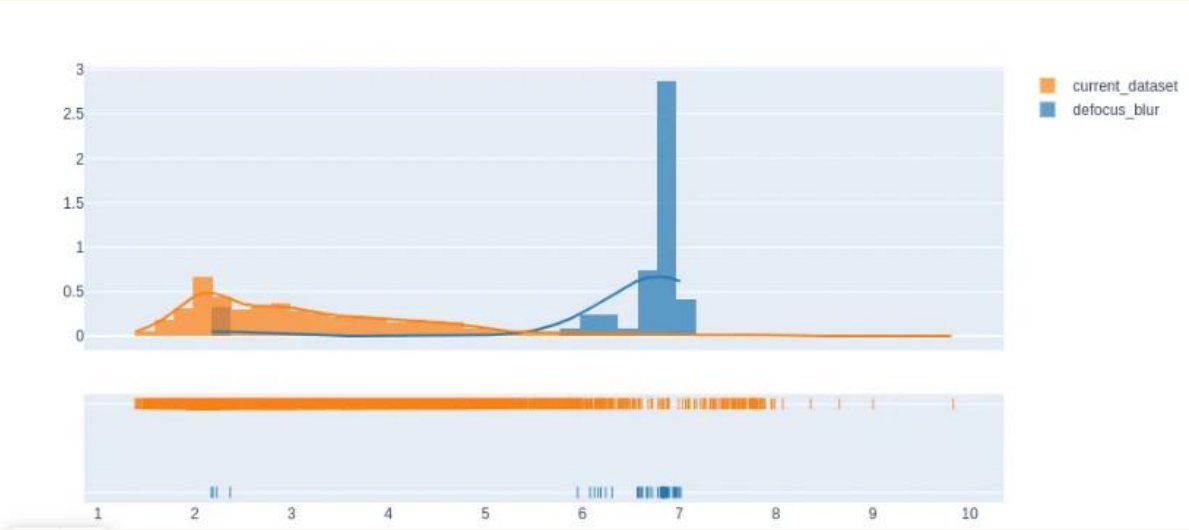
Метрики качества входных данных

Слой, отвечающий за оценку качества данных, **на вход получает показатель шума для данных камер и лидаров**

Показатель содержит бинарное значение, принимающее один из статусов: «корректная работа» в случае, если значение оценки не превышает пороговое, и «не определено» в случае превышения порогового значения

В случае **получения статуса «не определено» для конкретных сенсоров для зоны**, которую они контролируют, **устанавливается статус «неизвестно»**





Метрики работы нейронных сетей

Слой, отвечающий за оценку работы нейронной сети, **на вход получает значение неопределенности нейронной сети в качестве своих предсказаний**

В случае выхода значения за пороговые границы для нейронных сетей для зоны, которая попадает в область видимости конкретного сенсора, **устанавливается статус «неизвестно»**

Пример : Засвет от прожектора и его влияние на оценку неопределенности

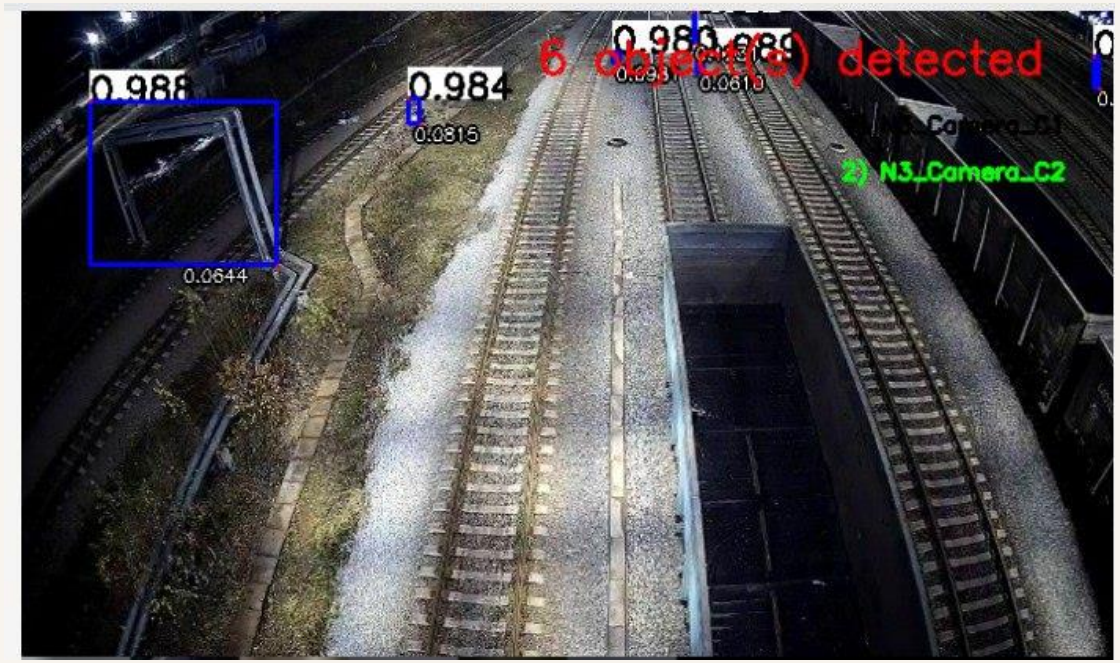
Нейронная сеть не может корректно выделить контур локомотива. Пиксели в этой зоне имеют высокую неопределенность (сложно определить где класс “**Фон**”, а где класс “**Вагон**”)

Чем больше неопределенность тем более синий цвет (колоризация для наглядности)

Проверка обнаруживающей способности

Слой, отвечающий за оценку обнаруживающей способности, **на вход получает метрику обнаруживающей способности железнодорожных путей и метрику неопределенности обнаружения заданных объектов**

В случае выхода метрик за пороговые границы для зоны, которая попадает в область видимости конкретного сенсора, **устанавливается статус «неизвестно»**



Проверка вхождения в красную зону

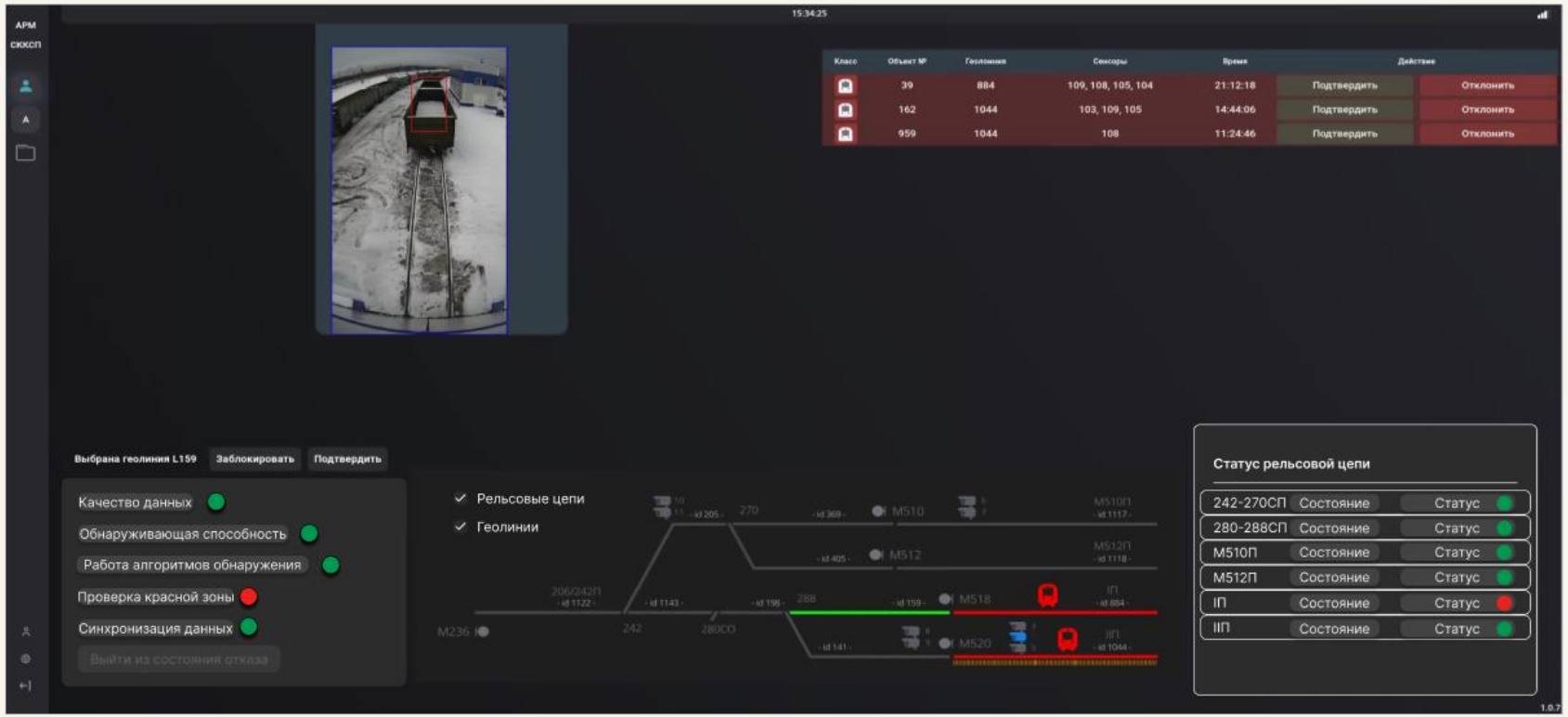
Слой, отвечающий за проверку вхождения в зону зоне рельсовой цепи, содержит реализацию алгоритма, который **повторяет основной алгоритм, но является более простым и классическим** (при этом и более грубым)

Данный алгоритм написан на языке C++ с применением требований стандарта кодирования

Проверка считается пройденной, если **результаты** алгоритма внутри машины состояний и основного алгоритма СККСП **сходятся**

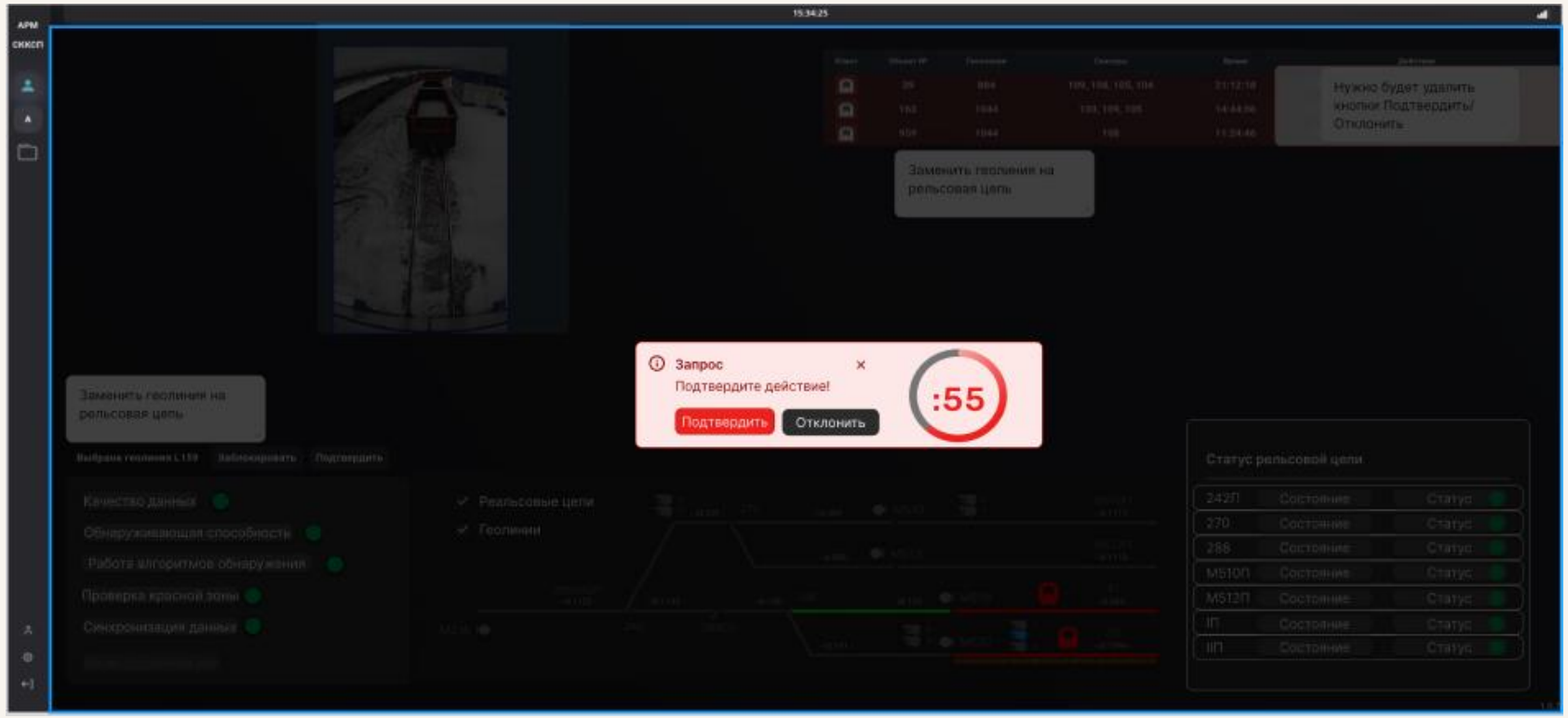
Выход из защитного состояния

Выход из защитного состояния возможен только по внешнему сигналу восстановления, при этом должны выполняться условия: входные данные возобновили поступление, данные удовлетворяют требованиям качества, корректные данные поступают регулярно на протяжении интервала восстановления (5 минут)



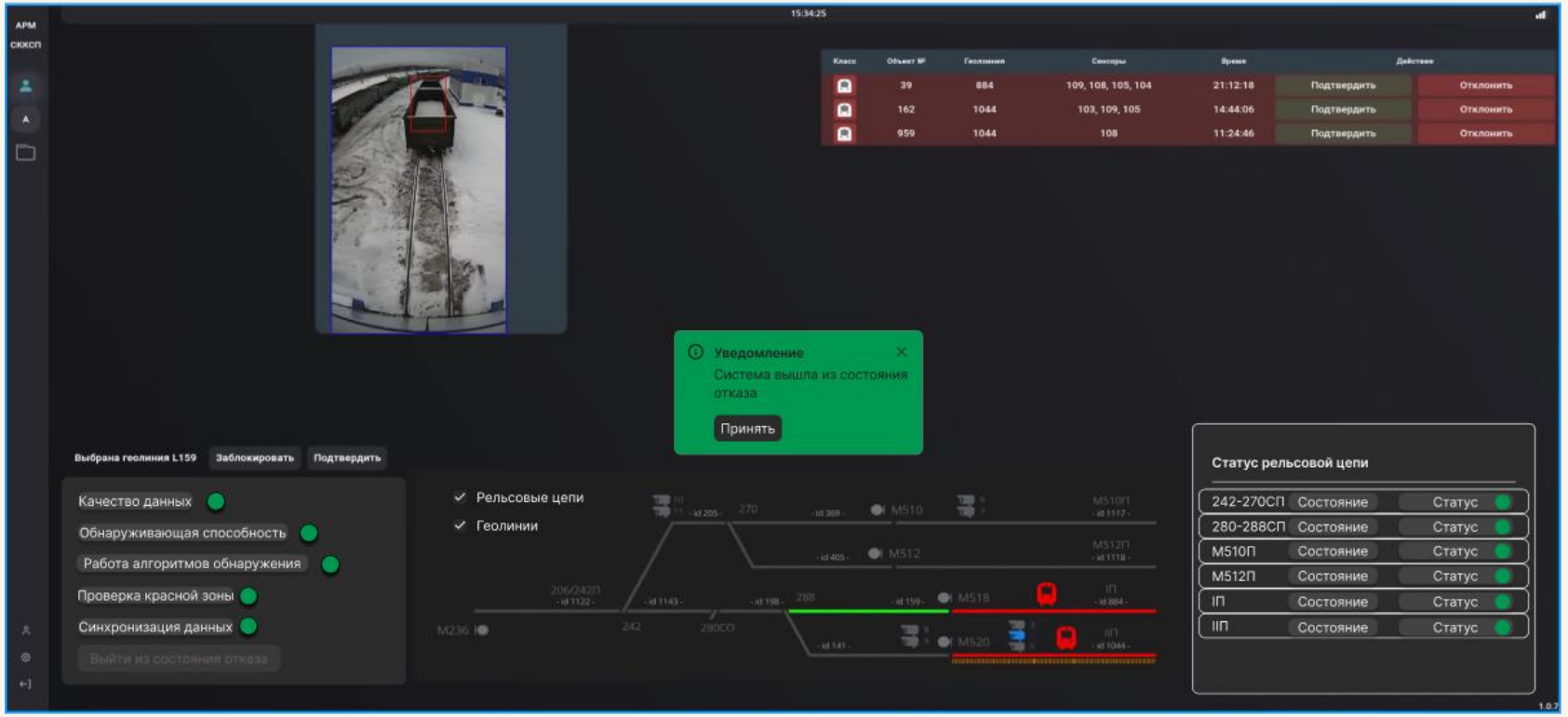
Выход из защитного состояния

Выход из защитного состояния возможен только по внешнему сигналу восстановления, при этом должны выполняться условия: входные данные возобновили поступление, данные удовлетворяют требованиям качества, корректные данные поступают регулярно на протяжении интервала восстановления (5 минут)



Выход из защитного состояния

Выход из защитного состояния возможен только по внешнему сигналу восстановления, при этом должны выполняться условия: входные данные возобновили поступление, данные удовлетворяют требованиям качества, корректные данные поступают регулярно на протяжении интервала восстановления (5 минут)



Испытания

Для проверки ФБ ПО применяются классические методы испытаний

Б.5.2.1.1.1 Подключиться к центральному вычислителю СККСП в режиме администратора в консоли удаленного управления, вызвав на рабочем столе скрипт ./computer_connect.sh.

Б.5.2.1.1.2 Деактивировать сервер времени на центральном вычислителе СККСП, вызвав из командной строки скрипт ./time_server_diactivate.sh.

Б.5.2.1.1.3 Сместить системное время вычислителя на 2 часа назад, вызвав из командной строки скрипт ./time_shift_back_two_ours.sh.

Б.5.2.1.1.4 Убедиться, что на АРМ СККСП индикатор состояния СККСП загорелся «красным», кнопка «Выход из защитного состояния» неактивна. Убедиться, что на АРМ ЦАМР индикатор состояния СККСП загорелся «красным» (рисунок Б.1).

Рисунок Б.1 – Индикация состояния СККСП на АРМ

Б.5.2.1.4.7 Направить прожектор на одну из камер.

Б.5.2.1.4.8 Убедиться, что на АРМ СККСП индикатор состояния СККСП загорелся «красным», кнопка «Выход из защитного состояния» неактивна. Убедиться, что на АРМ ЦАМР индикатор состояния СККСП загорелся «красным» (рисунок Б.1) а выбранная рельсовая цепь недоступна для построения маршрута.

Б.5.2.1.4.9 Убрать прожектор.

Б.5.2.1.4.10 Убедиться, что на АРМ СККСП индикатор состояния СККСП загорелся «желтым», кнопка «Выход из защитного состояния» активна. Убедиться, что на АРМ ЦАМР индикатор состояния СККСП продолжает гореть «красным» (рисунок Б.1), а выбранная рельсовая цепь недоступна для построения маршрута.

СПАСИБО ЗА ВНИМАНИЕ!