

Программные системы и комплексы

УДК 004.032.26

DOI:10.25729/ESI.2026.43.3.014

Разработка гибридного модуля сопровождения воздушных объектов на основе GRU и фильтра Калмана

Фелагерева Евгения Юрьевна, Рудоман Нэлли Радиковна,

Новиков Александр Дмитриевич

Кубанский государственный университет, Россия, Краснодар, *efelagereva@gmail.com*

Аннотация. В статье рассматривается задача устойчивого сопровождения воздушных объектов в видеопотоке, с акцентом на предсказание их траектории при наличии шумов измерений и кратковременных потерь детекции. Целью исследования является разработка гибридного метода трекинга, объединяющего классический фильтр Калмана и рекуррентную нейронную сеть типа GRU для повышения точности и устойчивости прогноза движения объектов. Предлагаемый подход основывается на использовании фильтра Калмана для оценки текущего состояния объекта, заданного вектором координат ограничивающей рамки YOLO и скоростями $[c_x, c_y, w, h, v_x, v_y]$, и нейросетевой модели GRU, которая обучается предсказывать изменение скорости $\Delta v_x, \Delta v_y$ на основе последовательности предыдущих состояний. В отличие от традиционных методов, где модель движения задается априорно, в предлагаемой работе динамика движения частично извлекается из данных. Предсказания GRU интегрируются в модель фильтра Калмана через механизм адаптивного сглаживания с коэффициентом α , что позволяет гибко регулировать вклад нейросетевого прогноза в итоговую оценку состояния. Основное содержание работы включает описание архитектуры гибридной модели, методики подготовки обучающего датасета, а также алгоритма интеграции нейросетевых предсказаний в процесс фильтрации. Особое внимание уделено задаче долгосрочного прогноза – до 30 кадров вперед – что является критически важным для систем активного сопровождения и управления. Научная новизна работы заключается в предложении способа комбинирования GRU и фильтра Калмана через адаптивное сглаживание управляющего воздействия, что позволяет учитывать нелинейные и нестационарные характеристики движения объектов без отказа от интерпретируемости классической модели. В отличие от существующих решений, предлагаемый метод обеспечивает более стабильную работу в условиях шумных измерений и пропусков детекций. В результате проведенных экспериментов показано, что предложенный подход повышает точность предсказания траектории на 16,57% и снижает накопление ошибки при многошаговом прогнозировании по сравнению с классическим фильтром Калмана и базовыми нейросетевыми моделями на 21%. Разработанный метод может быть применен в задачах детекции и сопровождения беспилотных летательных аппаратов, системах видеонаблюдения и автономной навигации.

Ключевые слова: фильтр Калмана, GRU (Gated Recurrent Unit – управляемый рекуррентный блок), БПЛА (беспилотный летательный аппарат), позиционирование

Цитирование: Фелагерева Е.Ю. Разработка гибридного модуля сопровождения воздушных объектов на основе GRU и фильтра Калмана / Е.Ю. Фелагерева, Н.Р. Рудоман, А.Д. Новиков // Информационные и математические технологии в науке и управлении, 2026. – № 3(43). – С. 165-177. – DOI:10.25729/ESI.2026.43.3.014.

Введение. Задача автоматической детекции и сопровождения воздушных объектов является одной из ключевых в современных системах компьютерного зрения и автономного управления. Особую сложность представляет сопровождение малых и быстро движущихся объектов, таких, как беспилотные летательные аппараты (БПЛА), на значительных дистанциях, где наблюдаются низкое разрешение цели, шумы измерений и нестабильность детекции. В таких условиях стандартные методы трекинга часто демонстрируют снижение точности и устойчивости, особенно при наличии пропусков обнаружения и резких изменений траектории движения.

Классическим подходом к решению задачи сопровождения является использование фильтра Калмана, который обеспечивает эффективную оценку состояния объекта на основе линейной динамической модели и зашумленных измерений. Несмотря на широкое распространение, данный метод предполагает априорно заданную модель движения, которая,

как правило, не отражает реальную нелинейную и нестационарную динамику воздушных объектов. Это приводит к накоплению ошибки при длительном прогнозировании и снижению качества сопровождения в сложных сценариях.

В последние годы активно развиваются методы, основанные на применении нейронных сетей, в частности рекуррентных архитектур, таких, как GRU (Gated Recurrent Unit – управляемый рекуррентный блок) и LSTM (Long Short-Term Memory – долгая краткосрочная память), способных извлекать временные зависимости из данных и моделировать сложные динамические процессы. Однако полностью нейросетевые подходы, как правило, требуют значительных объемов данных, обладают меньшей интерпретируемостью и могут демонстрировать нестабильность при работе в условиях шумных измерений.

В связи с этим перспективным направлением является разработка гибридных методов, сочетающих преимущества классических алгоритмов фильтрации и нейросетевых моделей. Такие подходы позволяют сохранить устойчивость и интерпретируемость фильтра Калмана, одновременно обогащая модель движения за счет данных.

В настоящей работе предлагается гибридный метод сопровождения объектов, в котором фильтр Калмана используется для оценки текущего состояния объекта, а рекуррентная нейронная сеть типа GRU – для предсказания изменения скорости на основе последовательности предыдущих состояний. Отличительной особенностью подхода является способ интеграции нейросетевого прогноза: предсказанные изменения скорости рассматриваются, как управляющее воздействие и вводятся в модель фильтра Калмана с использованием механизма адаптивного сглаживания. Это позволяет гибко регулировать влияние нейросети и повышает устойчивость системы к ошибкам предсказания.

Дополнительно в работе рассматривается задача многошагового прогнозирования траектории движения объекта на горизонт до 30 кадров вперед, что имеет практическое значение для систем активного сопровождения и управления. Предлагаемый метод ориентирован на работу в условиях ограниченных вычислительных ресурсов и может быть применен в реальном времени.

Таким образом, основной вклад работы заключается в разработке устойчивого гибридного алгоритма сопровождения, учитывающего сложную динамику движения объектов и обеспечивающего повышение точности прогноза в условиях шумных и неполных данных.

1. Смежные работы. В области сопровождения и прогнозирования траектории движения воздушных объектов активно применяются как классические методы фильтрации, так и современные подходы на основе глубокого обучения. Рассмотрим наиболее релевантные из них.

Традиционно для задачи трекинга используется линейный фильтр Калмана, так как движение объекта между соседними кадрами почти линейно, и использовании нелинейного алгоритма избыточно и не даст нужного прироста качества. Линейный фильтр предполагает заранее заданную модель движения с постоянной скоростью. Такой подход демонстрирует высокую эффективность при краткосрочном прогнозировании, однако при увеличении горизонта предсказания его точность существенно снижается. Это связано с тем, что линейная модель не учитывает ускорения и манёвры объекта, вследствие чего ошибка оценки состояния накапливается, особенно при прогнозировании на десятки кадров вперед.

В работе «VECTOR: Velocity-Enhanced GRU Neural Network for Real-Time 3D UAV Trajectory Prediction» [1] предложен подход к прогнозированию траектории БПЛА с использованием рекуррентной нейронной сети GRU, обученной на последовательностях скоростей. В отличие от моделей, использующих только координаты, авторы показывают, что моделирование изменений скорости позволяет значительно повысить точность предсказания,

особенно при нелинейной динамике движения и резких манёврах. Экспериментальные результаты подтверждают возможность применения модели в реальном времени.

Схожий подход представлен в статье «GRU-Based Deep Learning Framework for Real-Time UAV Trajectory Prediction» [2], где GRU применяется для анализа временных зависимостей в траекториях движения БПЛА. Авторы демонстрируют, что нейросетевые методы превосходят классические модели при моделировании сложных траекторий, однако отмечают снижение устойчивости при наличии шумов и ограниченного объема обучающих данных.

Наряду с полностью нейросетевыми подходами, активно развиваются гибридные методы, объединяющие рекуррентные сети и фильтр Калмана. В работе «Multiple Variational Kalman-GRU for Ship Trajectory Prediction» [3] предложена архитектура, в которой GRU используется для оценки параметров динамической модели, а фильтр Калмана выполняет вероятностную оценку состояния. Показано, что такая комбинация позволяет учитывать нелинейные зависимости и снижает ошибку прогнозирования по сравнению с классическим фильтром.

Аналогичный подход рассмотрен в статье «GRU-UKF Trajectory Prediction with Attention» [4], где выход рекуррентной сети интегрируется в модифицированный фильтр Калмана. Авторы показывают, что объединение нейросетевого предсказания и фильтрации позволяет существенно повысить точность и устойчивость системы в условиях шумных измерений.

Таким образом, существующие исследования показывают, что:

- линейный фильтр Калмана обеспечивает устойчивость, но ограничен простой моделью движения [5];
- рекуррентные нейронные сети эффективно моделируют сложную динамику, но менее устойчивы к шумам [6];
- гибридные подходы позволяют объединить преимущества обоих методов.

В отличие от рассмотренных работ, в настоящей статье предлагается использовать GRU для предсказания изменения скорости, а не абсолютного состояния объекта, с последующей интеграцией этого прогноза в линейный фильтр Калмана через механизм адаптивного сглаживания. Это позволяет компенсировать накопление ошибки при длительном прогнозировании и повысить устойчивость сопровождения без усложнения базовой модели движения.

2. Предлагаемый гибридный метод. Общая схема работы алгоритма включает три основных этапа: получение измерений от детектора [7], оценку состояния с использованием фильтра Калмана и коррекцию скорости объекта с помощью нейросетевого прогноза при отсутствии наблюдений.

В качестве детектора выступает дообученная модель YOLO11s, предсказывающая 3 класса объектов: самолетный и мультироторный типы, птицы. Птицы нужны для того, чтобы при их детекции не сопровождать.

В качестве наблюдений используются выходы YOLO – координаты ограничивающей рамки:

$$z_t = [c_x, c_y, w, h], \quad (1)$$

где c_x – центр рамки по оси x ; c_y – центр рамки по оси y ; w – ширина рамки; h – высота рамки.

Вектор состояния системы задаётся в следующем виде:

$$x_t = [c_x, c_y, w, h, v_x, v_y], \quad (2)$$

где v_x – скорость по оси x ; v_y – скорость по оси y .

2.1 Модель фильтра Калмана. Для оценки состояния используется линейный фильтр Калмана с моделью равномерного движения [8]. Был выбран именно линейный фильтр,

потому что движение объектов между соседними кадрами около линейное, и использование нелинейного алгоритма здесь избыточно. Динамика системы описывается выражением:

$$x_{t+1} = Fx_t + w_t \quad (3)$$

где F – матрица перехода, w_t – шум процесса.

Матрица перехода состояния имеет вид:

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 & dt & 0 \\ 0 & 1 & 0 & 0 & 0 & dt \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

Шум процесса:

$$w_t \sim N(0, Q), \quad (5)$$

где Q – ковариация шума процесса.

Модель наблюдения:

$$z_t = Hx_t + v_t, \quad (6)$$

где H – матрица наблюдения; v_t – шум измерения.

Матрица наблюдения имеет вид:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (7)$$

Шум измерения:

$$v_t \sim N(0, R), \quad (8)$$

где R – ковариация шума измерений.

Фильтр выполняет стандартные этапы прогнозирования и коррекции. При наличии детекции происходит обновление состояния по измерению, что позволяет компенсировать шумы и стабилизировать оценку [9].

2.2 Нейросетевая модель GRU. Для учета сложной динамики движения используется рекуррентная нейронная сеть типа GRU [10], состоящая из последовательных блоков (рис.1). На вход сети подаётся последовательность длины N , состоящая из предыдущих оценок состояния: $\{x_{t-N}, \dots, x_t\}$, где каждый вектор состояния представлен формулой 2.

Выходом сети является вектор приращения скорости:

$$\Delta v_t = \begin{bmatrix} \Delta v_x \\ \Delta v_y \end{bmatrix} \quad (9)$$

Таким образом, GRU моделирует изменение скорости объекта на основе временного контекста, что позволяет учитывать нелинейные траектории и манёвры, но нелинейность здесь скрыта не между соседними кадрами, а в последовательности кадров [11].

2.3 Гибридная схема интеграции. В отличие от классических гибридных подходов, в предлагаемом методе нейросетевой прогноз используется только в условиях отсутствия детекции.

Если на текущем кадре измерение отсутствует, выполняется только шаг прогнозирования фильтра Калмана по формуле 3.

При наличии достаточной истории состояний (N шагов) формируется входная последовательность для GRU, на основе которой вычисляется приращение скорости Δv_t .

Коррекция скорости осуществляется по правилу адаптивного сглаживания:

$$v_x = (1 - \alpha)v_x^{KF} + \alpha\Delta v_x, \quad (10)$$

$$v_y = (1 - \alpha)v_y^{KF} + \alpha\Delta v_y \quad (11)$$

где v_x^{KF} , v_y^{KF} – скорость, полученная после шага предсказания фильтра Калмана; Δv_x , Δv_y – изменение скорости, предсказанное GRU; α – коэффициент сглаживания, определяющий вклад нейросетевой коррекции. Структура блока GRU показана на рис. 1.

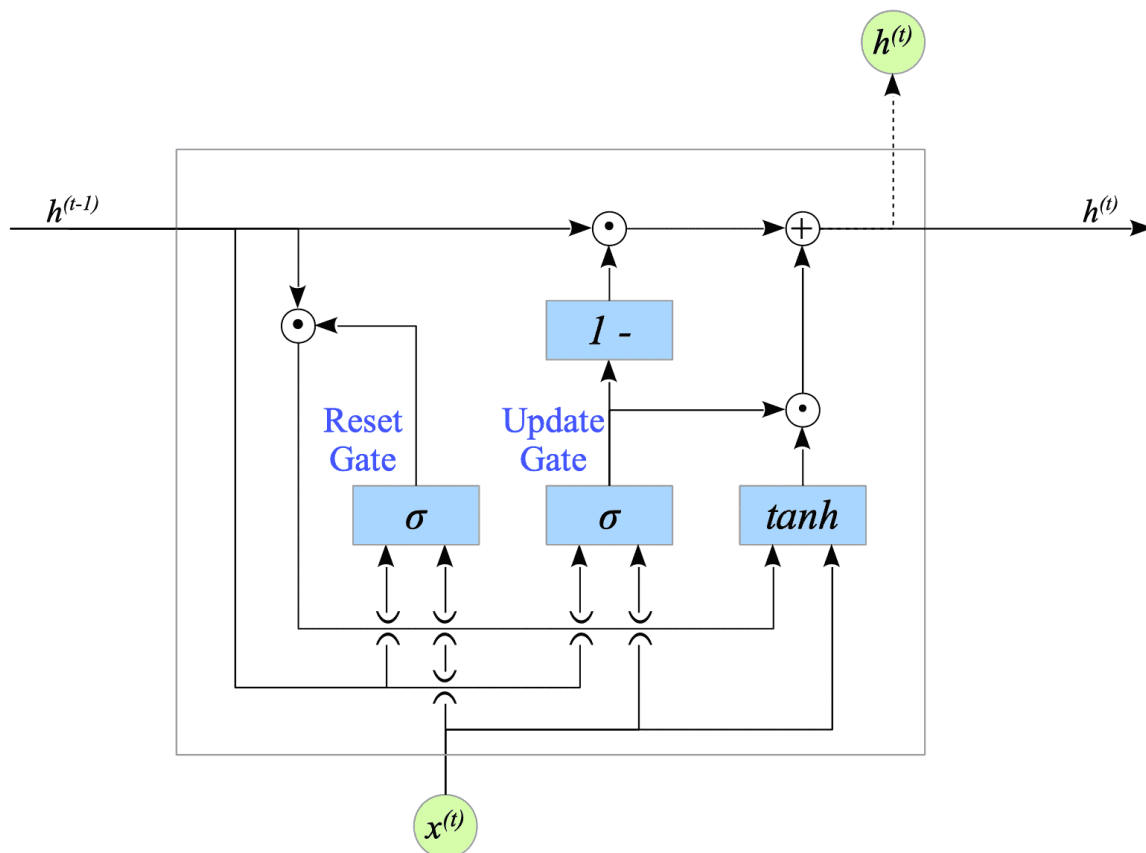


Рис. 1. Блок GRU

Величина α адаптивно изменяется в зависимости от длительности отсутствия детекции:

- при кратковременных пропусках используется большее значение α , что позволяет учитывать динамику движения;
- при увеличении горизонта прогнозирования значение α уменьшается, снижая влияние нейросети и предотвращая накопление ошибки.

Дополнительно вводится ограничение на величину скорости:

$$v_x, v_y \in [-v_{max}, v_{max}], \quad (12)$$

где v_{max} – максимальная скорость, определяемая из распределения скоростей реальных данных.

2.4 Алгоритм гибридного сопровождения. Работа метода на каждом временном шаге включает следующие этапы:

1. Получение измерения z_t от детектора (формула 1).
2. Если измерение доступно:
 - выполняется предсказание и обновление фильтра Калмана.
3. Если измерение отсутствует:
 - выполняется только шаг предсказания;
 - формируется последовательность состояний длины N , состоящая из предыдущих оценок состояния: $\{x_{t-N}, \dots, x_t\}$, где каждый вектор состояния представлен формулой 2;
 - предсказывается Δv_t с использованием GRU;
 - корректируются скорости с использованием адаптивного сглаживания.

4. Обновление состояния объекта.
5. Добавление текущего состояния в буфер истории.

Таким образом, фильтр Калмана обеспечивает устойчивую оценку состояния при наличии наблюдений, а GRU компенсирует ограничения линейной модели в условиях пропусков детекции, улучшая качество прогнозирования траектории.

3. Обучение модели GRU. Для создания датасета использовалась обученная YOLO11s, с помощью которой были обработаны видеопотоки мультироторного дрона. Стоит отметить, что для датасета не имеет разницы, какой класс снимать, важны только координаты детекции модели для формирования итоговых последовательностей. При съемке датасета важны были разнообразные маневры дрона, чтобы нейросеть училась на сложных последовательностях.

Чтобы получить итоговый вектор состояния (формула 2), нужно получить скорости, так как на выходе YOLO получаем только вектор из формулы 1. Для этого, а также, чтобы сгладить дрожание рамки, убрать пропуски (несколько секунд) и продлить треки, был применен фильтр Калмана. Как и рассматривалось ранее, была взята линейная модель фильтра, потому что между кадрами объект движется линейно с постоянной скоростью. Фильтр Калмана предсказывал только на 5 кадров вперед, потому что далее из-за линейной модели, рамка начинала «улетать» резко, что далее будет исправлять нейросеть. Итоговая архитектура обработки данных для подачи на вход в модель представлена на рис. 2.

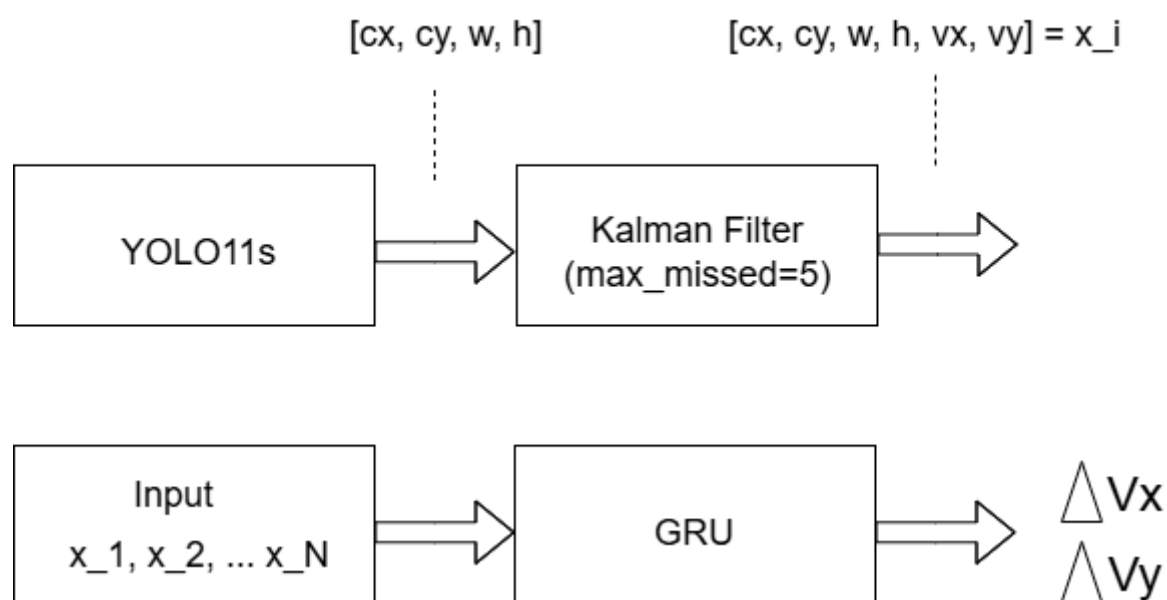


Рис. 2. Архитектура обработки данных для обучения GRU

Для подачи на вход нейросети применялись скользящие окна длиной 30 кадров, на которых сеть обучалась предсказывать приращения скорости на следующие 30 кадров.

Анализ данных показывает интересные закономерности:

- YOLO ограничивающая рамка и скорость: средние координаты объектов примерно по центру кадра ($cx \approx 0,53$, $cy \approx 0,43$), средний размер рамки небольшой ($w \approx 0,15$, $h \approx 0,12$), но наблюдаются редкие крупные значения до 0,97 и 0,77, что отражает сильное изменение масштаба цели. Средняя скорость детектора невелика (0,0128), однако встречаются быстрые движения до 0,835. Это говорит о разнообразии траекторий, чтобы GRU обучалась на сложных данных.
- Дрожание и сглаживание: YOLO имеет заметное дрожание ($\text{mean} \approx 0,025$, $\text{max} \approx 0,91$), фильтр Калмана снижает среднее дрожание до 0,021, демонстрируя эффективность линейного сглаживания.

- Ускорения и предсказания GRU: изменения скорости (Δv) достаточно малы (порядка $\sim 0,001$), из-за плавности большинства траекторий, но встречаются резкие скачки до 0,16 по оси x и 0,12 по оси y . Более того, больше половины значений колеблются около 0, а значит, требуется стандартизировать данные. После нормализации $\text{std} = 1$. Это позволяет сети GRU обучаться на единообразной шкале изменений скорости, учитывая как типичные малые приращения, так и редкие резкие манёвры без того, чтобы выбросы доминировали в процессе обучения. Такой подход улучшает сходимость модели и устойчивость предсказаний на сложных траекториях.
- Автокорреляция: $\Delta v \approx 0,53$ на 1 кадр вперед, что показывает сильную временную зависимость. Эти свойства делают рекуррентные модели верным выбором для предсказания будущей динамики.

Так как в данных обнаружились редкие выбросы, и они сильно влияют на распределение и плохо сказываются на нормализации, то были удалены последовательности, выходящие за пределы 1 и 99 перцентилей.

В совокупности статистика подтверждает, что подготовленный датасет обеспечивает достаточное разнообразие движений и сложность, позволяя GRU эффективно изучать временные зависимости, а фильтр Калмана – стабилизировать данные и компенсировать шум детектора.

Нейросеть имеет один GRU-слой с 64 скрытыми нейронами и последовательную голову из двух линейных слоёв с нормализацией и ELU-активацией. Для оптимизации использовался AdamW с начальной скоростью обучения $3 \cdot 10^{-4}$ и уменьшением LR по Plateau. В качестве функции потерь применялось среднеквадратичное отклонение (MSE), чтобы модель корректно училась на выбросах [12].

Модель обучалась на платформе Kaggle с использованием NVIDIA Tesla T4 с батчами по 128 последовательностей. Для предотвращения переобучения использовался ранний стоп при отсутствии улучшения валидационной ошибки в течение 12 эпох, а градиенты обрезались (clip) для стабильности. График обучения приведен на рис. 3.

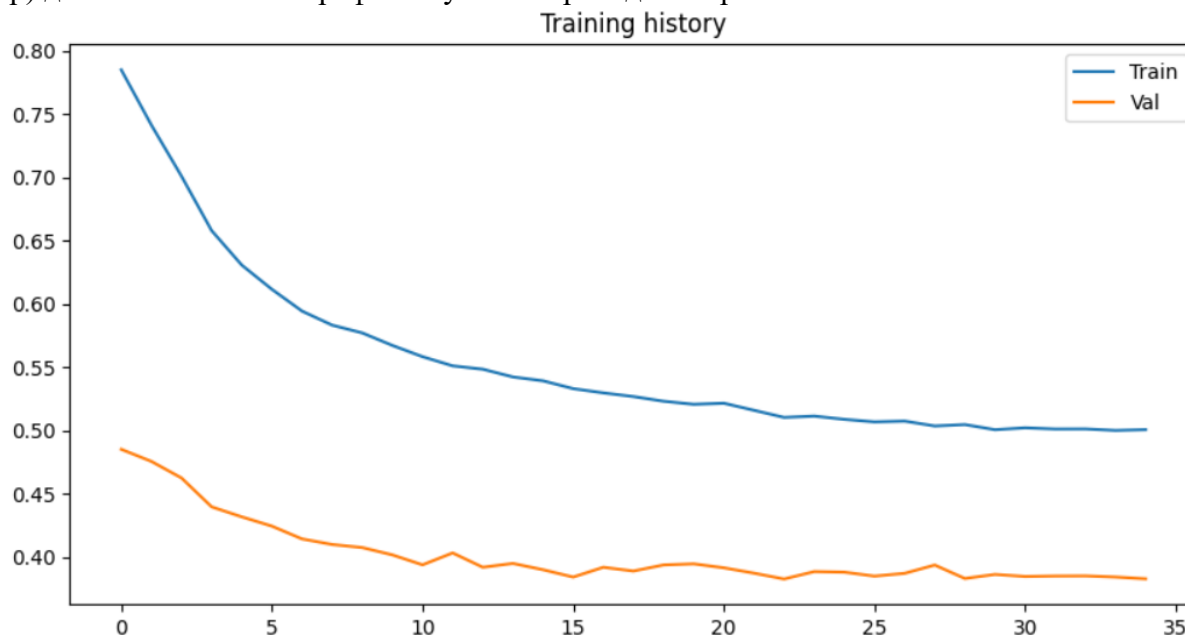


Рис. 3. График обучения GRU

Анализ процесса обучения показал стабильное снижение ошибки как на обучающей, так и на валидационной выборках:

- начальная ошибка (Epoch 1): Train 0,785, Val 0,485;

– наилучшая эпоха (Epoch 35, ранний стоп): Train 0,501, Val 0,383.

Это подтверждает хорошую сходимость модели и способность GRU учить сложные зависимости в последовательностях.

Так как в данных присутствует большое количество значений, близких к нулю, возникла необходимость проверить, не предсказывает ли модель всегда ноль. В противном случае ошибка могла бы оставаться относительно небольшой, но модель не выполняла бы фактическое предсказание скоростей.

Для оценки этого на тестовой выборке были рассчитаны метрики RMSE и MAE для базового сценария, когда выход модели всегда равен нулю, и сравнивались с фактическими предсказаниями модели [13]. Полученные результаты приведены в таблице 1 и демонстрируют, что модель GRU корректно предсказывает изменения скорости, значительно превосходя базовый вариант.

Таблица 1. Сравнение выходов GRU с константным выходом 0

Метрика	Baseline	GRU
MAE	0,396	0,333
RMSE	0,654	0,558

По метрикам видно, что выходы GRU лучше baseline на 19% по MAE и на 17% по RMSE, значит, модель выучила последовательности и эффективнее учитывает редкие резкие манёвры, так как RMSE чувствительна к выбросам.

Таким образом, GRU обеспечивает существенное улучшение точности прогноза по сравнению с наивным подходом, подтверждая целесообразность использования рекуррентной сети в гибридной системе сопровождения.

4. Разработка итогового алгоритма автоматической детекции и сопровождения БПЛА. В рамках работы был реализован итоговый алгоритм сопровождения воздушного объекта, объединяющий результаты детекции, фильтрацию состояния и нейросетевую коррекцию динамики. Основной задачей являлось обеспечение устойчивого трекинга при пропусках детекции и резких изменениях траектории движения.

Алгоритм работает в покадровом режиме и включает два основных состояния: наличие детекции и её отсутствие.

При наличии детекции координаты ограничивающего прямоугольника, полученные от модели YOLO, используются для обновления состояния с помощью фильтра Калмана. Это позволяет компенсировать шумы измерений и стабилизировать положение объекта.

В случае пропуска детекции активируется режим прогнозирования. На первом этапе выполняется предсказание состояния фильтром Калмана. Далее, при наличии достаточной истории (не менее 30 предыдущих состояний), формируется входная последовательность для GRU, которая предсказывает приращение скорости объекта. Полученное значение используется для коррекции скоростей фильтра Калмана с применением адаптивного сглаживания.

Коэффициент влияния GRU зависит от длительности пропуска детекции:

- в первые кадры пропуска используется большее значение (0,7), что позволяет учитывать текущую динамику движения;
- при увеличении горизонта прогнозирования влияние снижается (0,35), предотвращая накопление ошибки.

Дополнительно вводится ограничение на максимальную скорость, что обеспечивает устойчивость алгоритма и исключает выбросы при некорректных предсказаниях.

Таким образом, алгоритм сочетает:

- устойчивость фильтра Калмана при наличии наблюдений;

- способность GRU моделировать сложную динамику при их отсутствии.

Тестирование алгоритма проводилось на видеопоследовательностях с полётами мультироторного дрона в режиме реального времени на графическом процессоре NVIDIA GeForce 4060, что позволило оценить стабильность сопровождения объектов, точность предсказания координат и скоростей, а также поддерживать кадровую частоту видеопотока порядка 45 FPS.

Эксперименты включали сценарии плавного движения, резких манёвров, частичных и длительных пропусков детекции

Визуальный анализ показал, что:

- при наличии детекции алгоритм обеспечивает стабильное сопровождение без заметного дрожания;
- при кратковременных пропусках (до 10 кадров) GRU корректно предсказывает траекторию;
- при длительных пропусках (до 30 кадров) алгоритм сохраняет устойчивость за счёт постепенного снижения влияния нейросети.

Особенно важно, что гибридный подход предотвращает «улёт» ограничивающей рамки, характерный для линейного фильтра при длительном прогнозировании.

При обработке видеопотока с разрешением Full HD (1920×1080) была достигнута производительность:

- детекция (YOLO): ~35–45 FPS;
- GRU + Kalman: >200 FPS (не является узким местом);
- итоговая система: 40–45 FPS;

Таким образом, система работает в режиме, близком к реальному времени, с возможностью обработки видеопотока без существенных задержек.

Следует отметить, что вычислительная нагрузка в основном определяется детектором объектов, тогда как модуль сопровождения (Kalman + GRU) имеет незначительные затраты и может быть эффективно использован даже на менее производительных устройствах.

Для количественной оценки превосходства гибридного модуля трекинга над фильтром Калмана видеопоследовательности были обработаны дообученной моделью YOLO11s с последующим сохранением результатов детекции в формате CSV. Далее формировались фрагменты последовательностей, содержащие не менее 60 последовательных кадров с устойчивой детекцией объекта. Первые 30 кадров использовались для накопления истории наблюдений, необходимой для запуска гибридного алгоритма сопровождения, тогда как детекции на последующих 30 кадрах искусственно маскировались.

Таким образом, оба алгоритма выполняли прогноз положения объекта в условиях отсутствия детекции. Предсказания сравнивались с исходными результатами YOLO, принимаемыми в качестве истинных значений.

Метрики качества рассчитывались по евклидову расстоянию между центрами истинной и предсказанной ограничивающих рамок. Размеры рамок не учитывались, поскольку исследуемая система решает задачу прогнозирования движения объекта, а не восстановления его геометрических параметров. К тому же, ширина и высота приходят из дообученной YOLO11s.

Итоговое сравнение фильтра Калмана и предложенного гибридного метода представлено в таблице 2.

Таблица 2. Сравнение гибридного модуля сопровождения с фильтром Калмана

Метрика	Фильтр Калмана	Гибрид (Калман+GRU)
Mean error	0,094	0,074

RMSE	0,186	0,118
Std	0,161	0,093
Improvement, %	-	16,570

Для оценки качества сопровождения использовался набор метрик, характеризующих точность, устойчивость и стабильность предсказаний при пропадании детекции.

1. Средняя ошибка сопровождения (Mean error).

Метрика отражает среднее отклонение центра предсказанной рамки от истинного положения объекта в кадрах без детекции. Снижение ошибки составляет $\approx 21\%$, это означает, что использование GRU-предсказания движения позволяет значительно уменьшить накопление ошибки положения объекта при отсутствии измерений.

2. Среднеквадратичная ошибка (RMSE).

Метрика RMSE сильнее штрафует большие ошибки и характеризует наличие выбросов и срывов трекинга. Уменьшение RMSE составляет $\approx 36\%$. Полученный результат показывает, что гибридный алгоритм существенно снижает вероятность резкого ухода трека и лучше сохраняет траекторию объекта при длительном прогнозировании.

3. Стандартное отклонение ошибки (Std).

Метрика std характеризует стабильность сопровождения. Снижение дисперсии ошибки составляет $\approx 42\%$. Это свидетельствует о более плавной и устойчивой траектории сопровождения. Гибридный метод демонстрирует значительно меньшие колебания положения объекта относительно истинной траектории.

4. Интегральное улучшение (Improvement).

Показатель отражает относительное снижение средней ошибки сопровождения гибридного подхода относительно фильтра Калмана и составляет 16,57%. Таким образом, предложенный гибридный подход обеспечивает устойчивое повышение точности трекинга по сравнению с классическим фильтром Калмана.

Проведённый эксперимент показывает, что добавление рекуррентной нейросетевой модели движения к классическому фильтру Калмана позволяет:

- снизить среднюю ошибку сопровождения;
- уменьшить вероятность крупных ошибок предсказания;
- повысить устойчивость траектории;
- значительно улучшить качество трекинга при длительной потере детекции.

Полученные результаты подтверждают эффективность гибридного подхода для задач автоматического сопровождения воздушных объектов в видеопотоке и демонстрируют его преимущество над классическими методами фильтрации.

Заключение. В данной работе были рассмотрены классические методы сопровождения объекта на основе фильтра Калмана и выявлено, что они не всегда обеспечивают стабильное отслеживание при пропусках детекции или резких изменениях траектории дрона. Фильтр Калмана эффективно сглаживает траектории и предсказывает координаты при небольших пропусках наблюдений, однако его линейная модель движения ограничена в случаях сложных манёвров, ускорений и внезапных изменений направления. Для повышения точности и устойчивости трекинга был реализован рекуррентный нейросетевой модуль GRU, который корректирует предсказанные скорости объекта на основе истории наблюдений. Архитектура GRU состоит из одного слоя с 64 скрытыми нейронами и последующей линейной головой, предсказывающей приращения скорости по двум координатам. На вход модели подаются скользящие последовательности длиной 30 кадров, что позволяет учитывать динамику движения объекта и предсказывать его траекторию на горизонте до 30 кадров вперёд.

Для подготовки данных перед обучением GRU была использована дообученная модель YOLO11s, которая формировала датасет с мультироторными и самолётными типами дронов, а также птицами для проверки работы на сложных сценах. YOLO позволила выделять ограничивающую рамку объектов и рассчитывать скорости движения, которые затем сглаживались фильтром Калмана и использовались, как целевые приращения скорости для GRU. Такой подход позволил обеспечить высокое качество предсказаний даже при неполной детекции и минимизировать дрожание рамки.

Результаты тестирования алгоритма на видеопоследовательностях с реальными полётами БПЛА показали стабильное сопровождение объектов, точное предсказание координат и скоростей, а также возможность работы в реальном времени на графическом процессоре NVIDIA GeForce 4060. Средняя кадровая частота видеопотока достигла порядка 45 FPS, что обеспечивает достаточную скорость обработки для практического применения. Количественная оценка качества сопровождения выполнялась по евклидову расстоянию между центрами истинной и предсказанной ограничивающих рамок в условиях искусственно смоделированных пропусков детекции. По результатам эксперимента гибридный алгоритм показал снижение средней ошибки сопровождения на 16,57 % по сравнению с классическим фильтром Калмана. Среднеквадратичная ошибка (RMSE) уменьшилась примерно на 36 %, а стандартное отклонение ошибки – более чем на 40 %, что свидетельствует о более стабильной и предсказуемой траектории движения.

Таким образом, комбинированный подход, сочетающий фильтр Калмана и рекуррентную нейросеть GRU, позволяет реализовать гибридный алгоритм автоматического сопровождения БПЛА, который демонстрирует высокую точность, устойчивость к пропускам и возможность работы в реальном времени с кадровой частотой до 45 FPS. Это делает предложенное решение эффективным для практического применения в задачах трекинга объектов в динамичных условиях полёта.

Список источников

1. Nacar O., Abdelkader M., Ghouti L. et al. VECTOR: velocity-enhanced GRU neural network for real-time 3D UAV trajectory prediction. *Drones*, 2025, vol. 9, no. 1, DOI: 10.3390/drones9010008.
2. Yoon S., Jang D., Yoon H. et al. GRU-based deep learning framework for real-time, accurate, and scalable UAV trajectory prediction. *Drones*, 2025, vol. 9, no. 2, DOI:10.3390/drones9020142.
3. Jia C., Ma J., Kou W.M. Multiple variational Kalman-GRU for ship trajectory prediction with uncertainty. *IEEE Transactions on Aerospace and Electronic Systems*, 2025, vol. 61, no. 2, pp. 3654-3667, DOI: 10.1109/TAES.2024.3491053.
4. Yusheng L., Xinmin T., Xuanming R. GRU-IKF scene taxiing trajectory prediction model based on attention mechanism. *Journal of Applied Artificial Intelligence*, 2024, vol. 2, no. 1, pp. 314–327, DOI:10.59782/aai.v1i2.311.
5. Li Z. et al. GAKP: GRU association and Kalman prediction for multiple object tracking. *arXiv preprint*, 2020, arXiv:2012.14314, DOI:10.48550/arXiv.2012.14314.
6. Li N., Miao Q. LSTM-Kalman filter-based multi-sensor signal fusion for UAV altitude prediction in non-Gaussian environments. *Traitement du Signal*, 2025, vol. 42, no. 2, pp. 1111-1122, DOI:10.18280/ts.420242.
7. Золкин А.Л. Методы идентификации и отслеживания объектов в системах дополненной реальности и компьютерного зрения: учебное пособие для вузов. – СПб.: Лань, 2025. – URL: <https://e.lanbook.com/book/505478> (дата обращения: 22.01.2026).
8. Клетте Р. Компьютерное зрение. Теория и алгоритмы: учебник. – М.: Лань, 2019. – URL: <https://e.lanbook.com/book/131691> (дата обращения: 25.02.2026).
9. Макшанов А.В. Технологии интеллектуального анализа данных: учебное пособие. – СПб.: Лань, 2022. – URL: <https://e.lanbook.com/book/206711> (дата обращения: 12.03.2026).
10. Калин О. Архитектуры глубокого обучения. Математический подход: руководство. – М.: Лань, 2025. – URL: <https://e.lanbook.com/book/514864> (дата обращения: 20.03.2026).
11. Паттерсон Д. Глубокое обучение с точки зрения практика. – М.: Лань, 2018. – URL: <https://e.lanbook.com/book/116122> (дата обращения: 22.03.2026).

12. Парамонов В.А. Методы и модели исследования сложных систем и обработки больших данных / В.А. Парамонов, Н.Е. Смагин, А.Д. Косых, А.Д. Хомоненко. – СПб.: Лань, 2024. – URL: <https://e.lanbook.com/book/425024> (дата обращения: 22.03.2026).
13. Ненашев В.А. Компьютерное зрение. Анализ, обработка и моделирование: учебное пособие. – СПб.: Лань, 2022. – URL: <https://e.lanbook.com/book/341057> (дата обращения: 22.03.2026).

Фелагерева Евгения Юрьевна. Бакалавр в области радиофизики. Основные направления исследования автора: машинное обучение, нейронные сети, радиофизика. AuthorID: 1337017, SPIN: 1499-0400, ORCID: 0009-0001-2406-1225. efelagereva@gmail.com. 350040, г. Краснодар, ул. Ставропольская, 149.

Рудоман Нелли Радиковна. Старший преподаватель кафедры оптоэлектроники ФГБОУ ВО "Кубанский государственный университет". Основные направления исследования автора: физика твердого тела, компьютерное моделирование физических явлений, машинное обучение, искусственный интеллект, квантовые компьютеры, квантовая криптография. AuthorID: 39466, SPIN: 5844-2932, ORCID: 0009-0003-4308-235X. rudnel@ftf.kubsu.ru. 350040, г. Краснодар, ул. Ставропольская, 149.

Новиков Александр Дмитриевич. Аспирант в области оптики. Основные направления исследования: машинное обучение, нейронные сети, радиофизика. ORCID: 0009-0003-1596-8719. nad1100@yandex.ru. 350040, г. Краснодар, ул. Ставропольская, 149.

UDC 004.032.26

DOI:10.25729/ESI.2026.43.3.014

Development of a hybrid module for tracking air objects based on GRU and the Kalman filter

Eugenia Yu. Felagereva, Nelly R. Rudoman, Alexander D. Novikov

Kuban state university, Russia, Krasnodar, efelagereva@gmail.com

Abstract. This paper considers the problem of robust tracking of aerial objects in a video stream, with an emphasis on predicting their trajectories in the presence of measurement noise and short-term detection losses. The aim of the study is to develop a hybrid tracking method combining a classical Kalman filter and a GRU-type recurrent neural network to improve the accuracy and robustness of object motion prediction. The proposed approach is based on the use of a Kalman filter to estimate the current state of an object, specified by the coordinate vector of the YOLO bounding box and the velocities $[c_x, c_y, w, h, v_x, v_y]$, and a GRU-type neural network model, which is trained to predict the change in velocity $\Delta v_x, \Delta v_y$ based on a sequence of previous states. Unlike traditional methods, where the motion model is specified a priori, in this paper the motion dynamics are partially extracted from the data. GRU predictions are integrated into the Kalman filter model through an adaptive smoothing mechanism with a coefficient α , which allows for flexible adjustment of the contribution of the neural network prediction to the final state estimate. The main content of the paper includes a description of the hybrid model architecture, a methodology for preparing a training dataset, and an algorithm for integrating neural network predictions into the filtering process. Particular attention is paid to the problem of long-term forecasting—up to 30 frames ahead—which is critical for active tracking and control systems. The scientific novelty of the paper lies in the proposed method for combining GRU and a Kalman filter through adaptive smoothing of the control action, which allows for the consideration of nonlinear and nonstationary characteristics of object motion without sacrificing the interpretability of the classical model. Unlike existing solutions, the proposed method ensures more stable operation under noisy measurements and missed detections. Experiments demonstrated that the proposed approach improves trajectory prediction accuracy and reduces error accumulation during multi-step forecasting compared to the classical Kalman filter and basic neural network models. As a result of the experiments, it was shown that the proposed approach increases the accuracy of trajectory prediction by 16.57% and reduces the accumulation of errors in multi-step forecasting by 21% compared to the classical Kalman filter and basic neural network models. The developed method can be applied to unmanned aerial vehicle detection and tracking, video surveillance systems, and autonomous navigation.

Keywords: Kalman filter, GRU, UAV, positioning

References

1. Nacar O., Abdelkader M., Ghouti L. et al. VECTOR: velocity-enhanced GRU neural network for real-time 3D UAV trajectory prediction. *Drones*, 2025, vol. 9, no. 1, DOI: 10.3390/drones9010008.
2. Yoon S., Jang D., Yoon H. et al. GRU-based deep learning framework for real-time, accurate, and scalable UAV trajectory prediction. *Drones*, 2025, vol. 9, no. 2, DOI:10.3390/drones9020142.
3. Jia C., Ma J., Kouw W.M. Multiple variational Kalman-GRU for ship trajectory prediction with uncertainty. *IEEE Transactions on Aerospace and Electronic Systems*, 2025, vol. 61, no. 2, pp. 3654-3667, DOI: 10.1109/TAES.2024.3491053.
4. Yusheng L., Xinmin T., Xuanming R. GRU-IKF scene taxiing trajectory prediction model based on attention mechanism. *Journal of Applied Artificial Intelligence*, 2024, vol. 2, no. 1, pp. 314–327, DOI:10.59782/aai.v1i2.311.
5. Li Z. et al. GAKP: GRU association and Kalman prediction for multiple object tracking. *arXiv preprint*, 2020, arXiv:2012.14314, DOI:10.48550/arXiv.2012.14314.
6. Li N., Miao Q. LSTM-Kalman filter-based multi-sensor signal fusion for UAV altitude prediction in non-Gaussian environments. *Traitement du Signal*, 2025, vol. 42, no. 2, pp. 1111-1122, DOI:10.18280/ts.420242.
7. Zolkin A.L. *Metody identifikatsii i otslezhivaniya ob"yektov v sistemakh dopolnennoy real'nosti i komp'yuternogo zreniya: uchebnoye posobiye dlya vuzov* [Methods of identification and tracking of objects in augmented reality and computer vision systems: a textbook for universities]. Saint Petersburg, Lan Publ., 2025. Available at: <https://e.lanbook.com/book/505478> (accessed: 01/22/2026).
8. Klette R. *Komp'yuternoye zreniye. Teoriya i algoritmy: uchebnik* [Computer vision. Theory and algorithms: textbook]. Moscow, Lan Publ., 2019. Available at: <https://e.lanbook.com/book/131691> (accessed: 02.25.2026).
9. Makshanov A.V. *Tekhnologii intellektual'nogo analiza dannykh: uchebnoye posobiye* [Data mining technologies: textbook]. Saint Petersburg, Lan Publ., 2022. Available at: <https://e.lanbook.com/book/206711> (accessed: 03/22/2026).
10. Kalin O. *Arkhitektury glubokogo obucheniya. Matematicheskiy podkhod: rukovodstvo* [Deep learning architectures. A mathematical approach: guide]. Moscow, Lan Publ., 2025. Available at: <https://e.lanbook.com/book/514864> (accessed: 03/22/2026).
11. Patterson D. *Glubokoye obucheniye s tochki zreniya praktika* [Deep learning from a practitioner's perspective]. Moscow, Lan Publ., 2018. Available at: <https://e.lanbook.com/book/116122> (accessed: 03/22/2026).
12. Paramonov V.A., Smagin N.E., Kosykh A.D., Khomonenko A.D. *Metody i modeli issledovaniya slozhnykh sistem i obrabotki bol'shikh dannykh* [Methods and models for research of complex systems and big data processing]. Saint Petersburg, Lan Publ., 2024. Available at: <https://e.lanbook.com/book/425024> (accessed: 03/22/2026).
13. Nenashev V.A. *Komp'yuternoye zreniye. Analiz, obrabotka i modelirovaniye: uchebnoye posobiye* [Computer vision. Analysis, processing and modeling: textbook]. Saint Petersburg, Lan Publ., 2022. Available at: <https://e.lanbook.com/book/341057> (accessed: 03/22/2026).

Felagereva Evgeniya Yurievna. Bachelor's degree in Radiophysics. The author's main research areas are machine learning, neural networks, and radiophysics. AuthorID: 1337017, SPIN: 1499-0400, ORCID: 0009-0001-2406-1225. efelagereva@gmail.com. 350040, Krasnodar, st. Stavropol, 149.

Rudoman Nelly Radikovna. Senior Lecturer at the Department of Optoelectronics, Kuban State University. The author's main research areas are solid state physics, computer modeling of physical phenomena, machine learning, artificial intelligence, quantum computers, and quantum cryptography. AuthorID: 39466, SPIN: 5844-2932, ORCID: 0009-0003-4308-235X. rudnel@ftf.kubsu.ru. 350040, Krasnodar, st. Stavropol, 149.

Novikov Alexander Dmitrievich. Postgraduate student in optics. The author's main research areas are machine learning, neural networks, and radiophysics. ORCID: 0009-0003-1596-8719. nad1100@yandex.ru. 350040, Krasnodar, st. Stavropol, 149.

Статья поступила в редакцию 24.01.2026; одобрена после рецензирования 16.04.2026; принята к публикации 01.08.2026.

The article was submitted 01/24/2026; approved after reviewing 04/16/2026; accepted for publication 08/01/2026.