

И.С. Фомин, В.Д. Матвеев, А.Е. Архипов

СОПОСТАВЛЕНИЕ СПАЙКОВЫХ ПРЕДСТАВЛЕНИЙ ДЛЯ ЛОКАЛИЗАЦИИ И НАВИГАЦИИ В КАРТЕ КЛЮЧЕВЫХ КАДРОВ

Задача навигации мобильной робототехнической платформы в известном окружении достаточно давно и качественно решается с использованием плоской карты, которая строится с использованием лидара. Тем не менее, регулярно возникают ситуации, когда платформа по тем или иным причинам не оснащена лидаром или иными активными средствами навигации. При этом на платформе обычно установлена камера, предназначенная для визуального контроля обстановки оператором, которая также может быть использована для навигации при перемещении робота в известной среде. Хорошо известны примеры алгоритмов навигации, основанных на использовании последовательности ключевых кадров, например, визуальный SLAM. При этом в качестве ключевых кадров рассматриваются различные варианты видеоизображений (размытые, маскированные и т.д.). В данной работе в качестве базы для навигации рассматривается когнитивная (не-метрическая, непространственная) карта ключевых кадров, представляющих собой спайковое представление наблюдаемых изображений. В работе проанализирована возможность использования разработанных в ЦНИИ РТК нейроморфных информационно-управляющих элементов для выполнения процедуры сопоставления текущего спайкового изображения со спайковыми изображениями ключевой последовательности. Показано, что путем такого сопоставления может быть определен ближайший к текущему ключевой кадр, а также могут быть подобраны параметры сдвига спайковых представлений, что для когнитивной карты является аналогом локализации и навигации. Приведено описание программного средства эмуляции построения карты и перемещения в ней для экспериментальной отработки предложенных алгоритмов. Выполнен сбор данных и экспериментальное исследование качества работы алгоритмов локализации и навигации. Для чего собрано несколько карт ключевых кадров с различным характером перемещения между кадрами. При определении положения кадра в карте получено качество от 70 до 98 %, при определении направления смещения между кадрами точность составила от 94 до 97 %. Полученные результаты оценены как достаточные для решения поставленных перед алгоритмом задач.

Локализация; навигация; спайковое изображение; нейроморфный информационно-управляющий элемент.

I.S. Fomin, V.D. Matveev, A.E. Arkhipov

SPIKING REPRESENTATIONS COMPARISON FOR LOCALIZATION AND NAVIGATION IN THE KEYFRAME MAP

The task of navigating a mobile robotic platform in a known environment has been efficiently solved for a long time and using a flat passability map, which is built using lidar. Nevertheless, situations regularly arise when, for one reason or another, the platform is not equipped with lidar or other active navigation tools. At the same time, a camera is usually installed on the robotic platform, designed for visual monitoring of the situation by the operator, which can also be used for navigation when moving the robot in a known environment. There are well-known examples of navigation algorithms based on the use of sequences of keyframes, for example, visual SLAM. At the same time, various variants of video images (blurred, masked, etc.) are considered as keyframes. In this paper, a cognitive (non-metric, non-spatial) map of keyframes representing a spiking representation of the observed images is considered as a base for navigation. The possibility of using neuromorphic information control elements developed at the RTC to compare the current spiking representation with all spiking representations of a key sequence is analyzed. It is shown that by such a comparison, the keyframe closest to the current one can be determined, and parameters for the shift of spiking representations can also be selected, which is an analog of localization and navigation for a cognitive map. The description of a software tool for emulating the construction of a map and moving in it for experimental testing of the proposed algorithms is given. Data collection and experimental evaluation of the quality of localization and navigation algorithms have been performed. To do this, we have collected several keyframe maps with different patterns of movement between frames. When determining the position of the frame in the map, the quality was from 70 to 98%, when determining the direction of displacement between frames, the accuracy was from 94 to 97%. The results obtained are assessed as sufficient to solve the tasks assigned to the algorithm.

Localization; navigation; spiking representation; neuromorphic information control element.

Введение. Задача навигации мобильного робототехнического средства в заданном окружении является давно известной задачей, к решению которой существует достаточно большое количество различных подходов. В их числе следует отметить, прежде всего, навигацию на основе клеточной карты проходимости, которая строится с использованием плоского LIDAR [1, 2]. Этот подход хорошо зарекомендовал себя при передвижении по достаточно ровной поверхности со статическими и динамическими препятствиями, представленными на высоте, на которой их может определить лазерный луч. Примерами устройств на основе такого решения являются хорошо известные роботы-пылесосы [3] или роботы-газонокосилки [4], успешно осуществляющие перемещение в ограниченной области пространства.

Существуют также алгоритмы интерпретирующей навигации [5], которые основываются на использовании не метрических координат, а некоторой карты ориентиров, по которой принимается решение о местонахождении робота и требуемых действиях для того, чтобы перейти в целевую область пространства. В целом этот подход эффективен, однако, в зависимости от реализации, задает не всегда выполнимые требования к карте ориентиров.

Известны методы, основанные на использовании последовательности ключевых кадров (КК), например, *v*-SLAM, о чем будет более подробно сказано ниже. В данной работе предлагается рассмотреть решение задач локализации и навигации в когнитивной карте ключевых кадров. В общем случае под когнитивной картой понимается некоторый граф, соединяющий переходами области, в которых состав наблюдаемых ориентиров эквивалентен (области информационной эквивалентности, ИЭ). Каждая область ИЭ характеризуется последовательностью уникальных видимых ориентиров. Качество локализации достигается за счет наличия достаточного количества ориентиров с тем, чтобы области ИЭ были малы.

В данной работе под когнитивной картой понимается заранее построенная при прохождении известного маршрута последовательность ключевых кадров, каждый из которых описывается набором не уникальных видимых ориентиров из заданного списка классов. В общем случае карта представляет собой линейный или разветвленный граф, связывающий ключевые кадры. В работе рассматривается элементарный вариант последовательной фиксации таких кадров при преодолении некоторого маршрута (линейный граф).

Под ключевым кадром в работе понимается кадр, полученный при составлении карты, а текущим кадром будем называть кадр, который фиксируется при повторном прохождении маршрута в текущий момент времени. Под задачей локализации в когнитивной карте здесь понимается поиск ключевого кадра, в котором набор и положение ориентиров наилучшим образом соответствуют текущему кадру. Такой КК будем называть целевым кадром. Под навигацией в когнитивной карте будем понимать проверку параметров преобразования текущего кадра к целевому: приблизить-отдалить, сдвинуть влево-вправо.

Для построения когнитивной карты предлагается использовать последовательную фиксацию кадров с фронтальной камеры робота. Получаемые изображения преобразуются в спайковые представления [6] с использованием классических и нейросетевых методов обнаружения и сопровождения объектов, как описано в [7]. Как было показано в [8], такое преобразование сокращает размерность признаков. Принципиально далее возможно аналитическое или численное решение вышеописанных задач локализации и навигации по аналогии с [9]. В данной работе мы провели проверку возможности их решения с использованием энергоэффективной технологии использования элементарных нейроморфных информационно-управляющих элементов (НИУЭ), описанных в [8, 10].

Существующие решения. При решении задачи навигации с использованием камеры классическим подходом является визуальный SLAM (*visual SLAM*, *vSLAM*) [9, 11]. Различные методы *vSLAM* так или иначе выполняют обнаружение точек резкого изменения яркости на кадре (визуальных особенностей) и стереосопоставление для двух и более различных кадров, определяя положение этих точек и камеры в обобщенной системе координат. При навигации используется двухшаговый подход, когда сначала выполняется поиск ближайшего ключевого кадра, а затем определение положения камеры относительно ближайшего ключевого кадра. Задача точного определения положения камеры в данной работе не рассматривается, поскольку когнитивная карта не содержит количественных измерений, только качественную информацию о направлениях на объекты в КК.

В классических алгоритмах vSLAM (2010-е годы), например, [10, 12], при сопоставлении используется сжатие текущего кадра и ключевого кадра [11, 13], что порождает, в частности, грубые ошибки при ложном сопоставлении. К достоинствам данного подхода можно отнести его простоту и небольшие требования к вычислительным ресурсам. Схожий подход использовался в более поздних алгоритмах, например, ORB-SLAM [12–15], где были попытки улучшить работу метода за счет использования более качественных дескрипторов [14, 16]. В работе LSD-SLAM [15–18], рассматривался подход на основе плотного (dense) сопоставления ключевых кадров, но и тот и другой подход очень зависимы от качества определения ориентиров и вообще качества получаемых кадров.

Некоторые методы используют также стереокамеру или RGB-D камеру, чтобы дополнять карту также некоторой метрической информацией, есть попытки построения плотной карты при монокулярном или стереосопоставлении. В силу своих особенностей данные методы являются чувствительными к качеству изображения, получаемого с камеры, к точности определения положения визуальных особенностей и качеству этих особенностей.

Известны также подходы, когда ключевые кадры дополняются ориентирами с некоторыми дескрипторами или даже формируются с их использованием. Поиск определенных зон заметности в качестве естественных ориентиров [17, 19], их сопровождение [18, 20] и ориентация на точки именно из них может существенно улучшать качество работы описанных выше vSLAM методов. Обнаружение и сопровождение искусственных визуальных ориентиров в качестве устойчивых признаков [19, 21] может позволить решить в том числе задачу определения пространственного положения, если с высокой точностью известны их пространственные координаты и ориентиров достаточное количество. Также задача определения положения может быть решена геометрически с разной степенью обусловленности если видимы три и более ориентиров [20, 22], или с помощью расширенного фильтра Калмана [21, 23], если известны закон движения, данные одометра и азимуты линий визирования на ориентиры с некоторой историей. К достоинствам перечисленных методов стоит отнести достаточность перечисленных элементов для определения положения и ориентации мобильного робота. В качестве недостатков необходимо отметить высокие требования к вычислительным ресурсам и необходимость знать положения ориентиров в некоторой заранее определенной системе координат.

Чтобы уйти от требования к наличию уникального идентификатора и координат для каждого объекта можно осуществить переход к более простому спайковому представлению кадра [7]. Для преобразования и сопоставления кадров могут быть использованы различные подходы, включая как аналитические способы, так и прямое вычисление дистанции между спайками. В данной работе были использованы рассмотренные в работе [8, 10] нейроморфные информационно-управляющие элементы, основанные на биоморфной сегментной спайковой модели нейрона [22, 24], отличающейся высокой модульностью и интерпретируемостью и в будущем предполагающей эффективную аппаратную реализацию [23, 25]. Совокупность (сеть) из нескольких связанных между собой НИУЭ позволяет как принять решение о том, какой из кадров карты наилучшим образом соответствует текущему, и в каком направлении сдвинута камера в целевом кадре относительно текущего при условии, что смещение небольшое.

Для обнаружения и сопровождения объектов в процессе исследования местности и сбора данных можно использовать фундаментальные нейронные сети для универсальной сегментации и сопровождения областей SAM [24, 26] и DeAOT [25, 27]. Для определения классов обнаруженных областей применимы нейросетевые классификаторы, например, ResNet [26, 28] или ConvNext [27, 29]. Ранее также рассматривался подход к классификации с использованием комбинации сиамской сверточной и сегментной спайковой сетей [28, 30]. Имея собранные данные, можно обучить нейросетевой детектор, например, Faster R-CNN [29, 31], Mask R-CNN [30, 32]. В случае, когда требуется высокая производительность или перенос на бортовой вычислитель, имеет смысл для распознавания использовать модели семейства YOLO [31, 33], [32, 34], RetinaNet [33, 35] или CenterNet [34, 36].

Механизм построения карты ключевых кадров. Сначала с помощью робота с установленной камерой, которая в будущем будет использоваться при навигации, выполняется запись видеороликов при произвольном движении робота в районе траектории по

которой в будущем планируется движение. Далее, каждый из роликов размечается комбинацией сетей SAM [24, 26] и DeAOT [25, 27]. После этого полученная разметка очищается вручную таким образом, чтобы оставить только те сегменты, которые относятся к объектам, локализованным в пространстве. К примеру, сегменты классов «пол», «потолок», «небо» занимают в кадре значительную площадь и почти не меняют свое положение на изображении при повороте и перемещении камеры и поэтому мало полезны для навигации. Напротив, объекты в помещении такие как «стул», «стол», «шкаф», «компьютер», «монитор» и т.п. не очень велики, хорошо локализованы в трехмерном пространстве помещения и занимают относительно небольшую часть площади кадра, поэтому их координаты на изображении могут использоваться при локализации и навигации.

Размеченные данные могут быть использованы различным способом для формирования когнитивной карты ключевых кадров. Ранее рассматривались подходы к использованию пространства представлений и распознаванию объектов на основе комбинации сиамской сверточной и сегментной спайковой нейронной сетей [7], что позволяет использовать подход на основе обучения без учителя, когда известно только то, что пары изображений относятся к одному и тому же объекту, а класс при обучении не задан.

Кроме того, если состав классов объектов и варианты внешнего вида заранее определены, то хорошим вариантом является использование вычислительно эффективных нейронных сетей для обнаружения и распознавания объектов. Наиболее известные методы для обнаружения объектов – семейство нейронных сетей R-CNN [29, 31], однопроходные методы на основе построения сетки и перечня допустимых габаритов объектов (SSD, YOLO различных версий [31, 33], [32, 34]), точечные методы оценки (RetinaNet [33, 35], CenterNet [34, 36]). При необходимости в использовании эффективного метода распознавания объектов или в переносе распознавания на непосредственно бортовой вычислитель на основе встраиваемого вычислителя с TPU лучшим решением является использование одной из версий YOLO.

С использованием одного из этих подходов выполняется построение когнитивной карты ключевых кадров. Кадры сохраняются при продольном или угловом смещении камеры на заданную величину. При формировании карты робот с камерой проходит по траектории, которую в будущем требуется воспроизвести. Граф из ключевых кадров формируется последовательно при прохождении траектории. Преобразование изображений в спайковые представления – процесс давно известный, и имеющий некоторые стандартные решения, например, [6, 35, 37]. Здесь для каждого из кадров формируется спайковый образ, в котором задержки напрямую соответствуют горизонтальным координатам объектов, которые близко соответствуют направлениям на объекты относительно оптической оси камеры. Пример спайкового образа одного из кадров представлен на рис. 1.

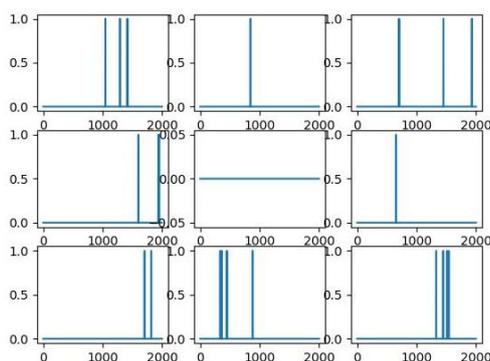


Рис. 1. Спайковый образ ключевого кадра

Алгоритмы локализации и навигации в когнитивной карте ключевых кадров.

Для того, чтобы выполнить локализацию текущего кадра в карте необходимо найти ключевой кадр, который соответствует текущему наилучшим образом. Для сопоставления

спайковых представлений необходимо использовать группу из правильно настроенных НИУЭ [8, 10] для сопоставления последовательностей. Необходимо подать на возбуждающие входы элементов спайки образа ключевого кадра, а на тормозящий вход спайки текущего кадра (см. схему на рис. 2), и, если спайки в последовательностях достаточно близки, спайки на тормозящих входах подавят возбуждение и срабатывания не произойдет, как показано для первых двух пар спайков на рис. 3.

Сравнение последовательностей выполняется попарно для каждого из классов спайкового образа. В конце объединяются срабатывания для всех классов на элементарном НИУЭ типа «ИЛИ» и суммируются на счетчике для каждого из ключевых кадров карты. Кадр с наименьшим числом срабатываний (т.е. отличий) признается за текущее положение в карте. Если для всех кадров число срабатываний слишком велико, карта с большой долей вероятности не соответствует месту, в котором робот с камерой находится в данный момент.

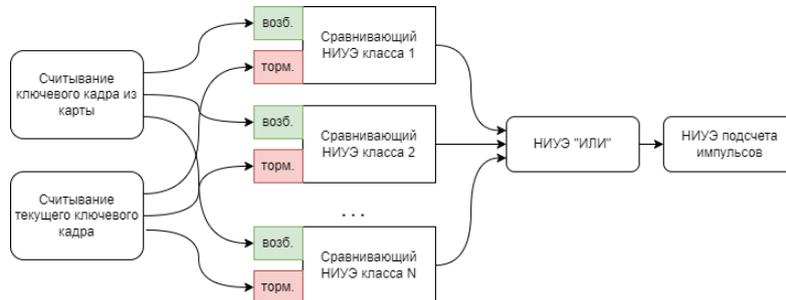


Рис. 2. Принцип сравнения кадров с использованием НИУЭ

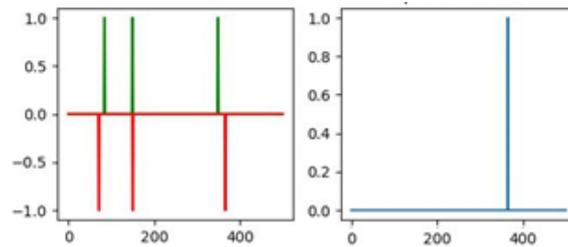


Рис. 3. Пример взаимного гашения импульсов

Кроме задачи локализации, если известны спайковые картины для текущего ключевого кадра и несколько отличающегося от него целевого кадра с помощью НИУЭ может быть решена задача навигации (определение направления смещения). В качестве целевого обычно выбирается следующий ключевой кадр от положения в карте. Для определения того, как именно смещена камера в целевом кадре относительно текущего используются специальные искажающие НИУЭ, а также показанный выше механизм сопоставления ключевых кадров.

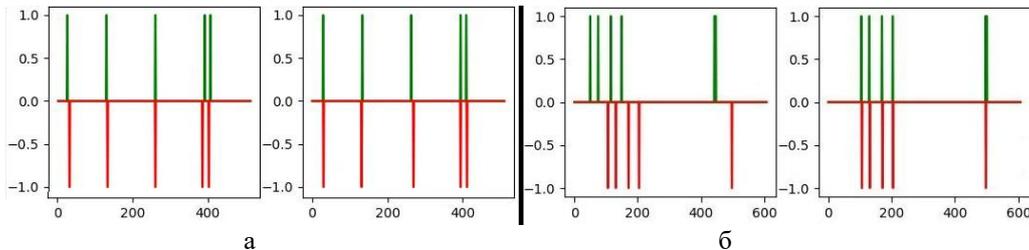


Рис. 4. Пример смещения спайков последовательностей (а – вперед, б – вправо)

Когда целевой кадр сдвинут относительно текущего вперед или назад, объекты на изображении и соответствующие им спайки сдвинуты в стороны относительно середины (см. рис. 4,а, слева). Тогда перед сопоставлением последовательностей текущего и целевого кадра необходимо их исказить с учетом задержки D_1 . Величина задержки зависит от дистанции смещения. Пусть продолжительность обеих последовательностей равна T . Сначала надо добавить задержку D_1 для целевого кадра (зеленые спайки). Затем сдвинуть правую часть ($t \geq T/2$) текущего кадра (красные спайки) на две величины задержки D_1 . Таким образом проверяется, совпадут ли спайковые образы, если текущий кадр сместить вперед на некоторую дистанцию (рис. 4,а, справа). Для движения назад все аналогично, но целевой и текущий кадры меняются местами.

Если целевой ключевой кадр повернут относительно текущего на некоторый угол, то все спайки на целевом кадре будут сдвинуты относительно спайков на текущем на некоторую величину. Если целевой кадр повернут вправо, его спайки придут на вход раньше (рис. 4,б, слева). Для того, чтобы проверить, будут ли совпадать последовательности, если повернуть текущий кадр до целевого, необходимо добавить задержку D_2 для целевой последовательности, тогда картина примет вид как на рис. 4,б, справа. Разные величины D_2 соответствуют повороту на разный угол. При повороте влево аналогичная задержка добавляется для последовательности текущего кадра.

Для всех направлений необходимо исследовать несколько величин задержек D_1 и D_2 , плюс их отсутствие для определения достижения целевого кадра. Общая схема показана на рис. 5. Выбирая вариант с минимальным количеством откликов, блок принятия решений может определить, в каком направлении относительно текущего смещен целевой кадр.

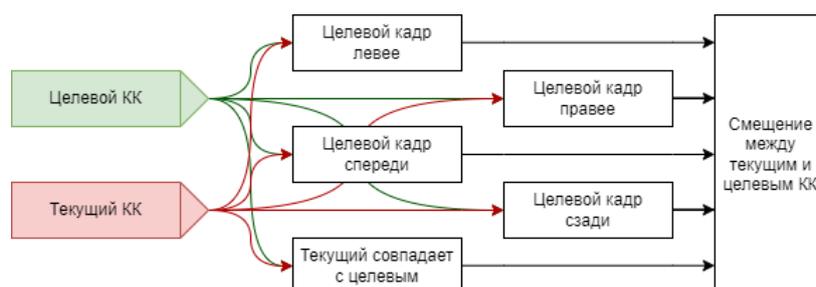


Рис. 5. Схема определения смещения между текущим и целевым КК

Описание эмулятора карты для исследования алгоритмов. Для проведения экспериментов по локализации и навигации в когнитивной карте ключевых кадров были разработаны инструменты для создания плана помещения (вид сверху) и эмуляции формирования карты ключевых кадров по этому плану, локализации и навигации в ней. Оба инструмента разработаны с использованием библиотеки для обработки изображений OpenCV на языке Python.

Инструмент для создания плана позволяет заполнять заданными цветами квадратные области с шагом в 10 см. Каждому цвету (в формате RGB) соответствует один класс объектов, черный соответствует отсутствию объектов. Размер карты, шаг, набор и цвета классов задаются как параметры. В результате получается план, информацию о положении объектов в котором можно получать посредством операций с цветами.

Для симуляции перемещения робота и построения карты был разработан программный эмулятор. Расположение каждого объекта определяется расположением центра масс области, окрашенной на плане соответствующим цветом, форма определяется формой области с точностью до шага ячейки. Робот эмулируется окружностью с допустимым радиусом столкновения и известным направлением. Камера эмулируется заданным углом обзора (в каких пределах по углу относительно направления движения робота камера видит объект) и радиусом восприятия (на каком расстоянии от центра робота видны объ-

екты). Перекрытия объектов друг другом для простоты не учитываются. Задержка спайков в спайковом образе прямо пропорциональна углу направления на объект в створе поля зрения камеры относительно левой его границы (угол и смещение спайка растут по часовой стрелке). В режиме построения когнитивной карты ключевых кадров текущий спайковый образ записывается в карту в качестве нового ключевого кадра, когда угловое или продольное смещение относительно предыдущего превысит заданное значение.

Экспериментальное исследование локализации. В рамках экспериментов по оценке качества работы алгоритма локализации было выполнено попарное сопоставление полученных карт для сформированных наборов данных. Для каждой из пар карт оценивалось число правильно и ошибочно локализованных кадров из первой последовательности во вторую и наоборот.

Для сопоставляемого ключевого кадра выполняется сравнение со всеми ключевыми кадрами карты. Подсчитывается количество спайков. Кадр с минимальным количеством откликов (или один из нескольких кадров, если два и более имеют одинаковое наименьшее число откликов) считается соответствующим сопоставляемому если число откликов меньше установленного порога. Успешность сопоставления определяется углом и дистанцией между текущим кадром и найденным соответствием менее установленного порога. Точность сопоставления – это отношение числа успешно сопоставленных кадров исследуемой последовательности к общему числу. Результаты для 3 пар последовательностей и 6 экспериментов по сопоставлению представлены в табл. 1.

Таблица 1

Результаты экспериментов по локализации

Эксперимент	Правильно	Ошибочно	Точность, %
Левый 1 → Левый 2	65	4	94
Левый 2 → Левый 1	67	1	98
Правый 1 → Правый 2	63	3	95
Правый 2 → Правый 1	65	2	97
Круги 1 → Круги 2	80	28	74
Круги 2 → Круги 1	76	32	70

Полученное по результатам эксперимента качество сопоставления для прямолинейного и кругового движения можно признать удовлетворительным. В рамках исследования алгоритм был поставлен в принципиально более сложные условия, чем те, которые будут в реальной работе, поэтому на практике качество сопоставления будет выше.

Экспериментальное исследование навигации. Оценка качества работы предложенного подхода к навигации (т.е. определению направления смещения целевого ключевого кадра относительно текущего) выполнялась следующим образом. Каждая из 6 полученных карт представляет собой последовательность из N ключевых кадров с известным направлением движения. Для каждой из последовательностей сравнивались попарно соседние кадры (т.е. n и $n + 1$ для $n \in (0, N - 1)$) с использованием вышеописанного подхода. Для каждой пары по минимальному числу откликов определялось выбранное направление смещения и прибавлялось к значению в соответствующей графе.

Показатель точности соответствует проценту пар кадров, которые были распознаны правильно. Количество пар, целевые значения и точности представлены в табл. 2. Здесь слева показаны идентификаторы последовательностей, где «Л-П» и «П-Л» – карты ключевых кадров, в процессе получения которых робот в симуляции двигался прямолинейно вперед, слева направо или справа налево на плане, «Пов.» (поворот) – робот между кадрами совершал вращательное движение.

По результатам исследования можно утверждать, что разработанный алгоритм навигации, определения направления смещения между текущим и целевым ключевым кадром продемонстрировал высокое качество сопоставления для всех рассмотренных направлений.

Таблица 2

Результаты экспериментов по навигации

Послед.	Вперед	Назад	Влево	Вправо	Цель	Точность, %
Л-П 1	65	0	2	1	Вперед	95,59
Л-П 2	63	1	0	3	Вперед	94,03
П-Л 1	63	2	0	0	Вперед	96,92
П-Л 2	63	0	1	1	Вперед	96,92
Пов. 1	2	1	1	103	Вправо	96,26
Пов. 2	1	2	2	102	Вправо	95,33

Заключение. В рамках работы были рассмотрены подходы к локализации и навигации в когнитивной карте ключевых кадров с использованием спайковых представлений, построенных по данным изображений, получаемых с камеры. Описан подход к сбору данных и построению когнитивной карты ключевых кадров для работы в новой, ранее не изученной обстановке, приведено описание нескольких вариантов существующих подходов, которые могут быть применены для решения этой задачи.

Предложены подходы для локализации и навигации в когнитивной карте ключевых кадров с использованием групп известных нейроморфных информационно-управляющих элементов для преобразования и сравнения спайковых представлений кадров, построенных основе сегментного спайкового нейрона модели CSNM.

Разработан эмулятор передвижения робота на двумерном плане помещения с заданным набором объектов, соответствующим положению объектов в некотором реальном помещении. С использованием эмулятора собраны наборы данных, на базе которых проведена оценка качества локализации и навигации.

Алгоритм локализации показал очень высокое качество от 94 до 98% в случае решения задачи при прямолинейном движении (основной ожидаемый режим перемещения в процессе патрулирования) и до 74 % при локализации в карте с разными положениями камеры при ее вращении. Сравнительно низкий результат во втором случае объясняется тем, что алгоритм был поставлен в заранее невыгодные условия, которые не должны возникать при плановом использовании.

Алгоритм навигации показал качество от 94 до 97 % при прямолинейном движении и до 96 % при повороте. В целом, полученные показатели качества можно считать достаточно высокими, а разработанный алгоритм – предварительно пригодным для решения задачи навигации в карте ключевых кадров.

В дальнейшем предполагается продолжить исследование разработанных алгоритмов на видеопоследовательностях, полученных с реальной камеры. В частности, для построения образов ключевых кадров требуется выполнить исследование нейронных сетей семейства YOLO, R-CNN, RetinaNet, CenterNet в задаче обнаружения и распознавания заданного набора объектов. В перспективе также планируется рассмотреть вопрос переноса алгоритмов распознавания объектов на встраиваемые вычислители, а алгоритмов локализации и навигации на ПЛИС и/или аналоговые интегральные схемы.

Результаты получены в рамках выполнения государственного задания Минобрнауки России «Исследование методов самоорганизации и семантического взаимодействия в гибридных группах робототехнических агентов повышенной автономности»

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Fox D., Burgard W., Thrun S. The dynamic window approach to collision avoidance // IEEE Robotics & Automation Magazine. – 1997. – Vol. 4, No. 1. – P. 23-33.
2. Brock O., Khatib O. High-speed navigation using the global dynamic window approach // Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C). – IEEE, 1999. – Vol. 1. – P. 341-346.
3. Pan S., Xie Z., Jiang Y. Sweeping robot based on laser SLAM // Procedia Computer Science. – 2022. – Vol. 199. – P. 1205-1212.
4. Hicks II R.W., Hall E.L. Survey of robot lawn mowers // Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision. – SPIE, 2000. – Vol. 4197. – P. 262-269.

5. *Кирильченко А.А., Платонов А.К., Соколов С.М.* Теоретические аспекты организации интерпретирующей навигации мобильного робота // Препринт Ин-та прикл. матем. им. МВ Келдыша РАН. – 2002. – №. 5. – С. 40.
6. *Orchard G. et al.* Converting static image datasets to spiking neuromorphic datasets using saccades // *Frontiers in neuroscience.* – 2015. – Vol. 9. – P. 437.
7. *Fomin I., Korsakov A., Arkhipov A.* Comparison of Key Frames of Video Sequences Using the Izhikevich Spiking Neuron // *2023 International Ural Conference on Electrical Power Engineering (UralCon).* – IEEE, 2023. – P. 594-598.
8. *Иванова В.В., Беркман Д.А., Корсаков А.М.* Метод уменьшения размерности пространства признаков при решении задачи классификации // *Перспективные системы и задачи управления.* – 2023. – С. 591-596.
9. *Корсаков А.М., Степанов Д.Н., Смирнова Е.Ю.* Алгоритм определения ориентации транспортного средства на перекрестке по данным системы технического зрения // *Интеллектуальные системы, управление и мехатроника - 2016.* – 2016. – С. 396-399.
10. *Иванова В.В., Демчева А.А., Корсаков А.М.* Нейроморфный механизм управления заданием по результатам анализа ситуации // *Экстремальная робототехника.* – 2024. – Т. 35, № 1. – С. 308-3015.
11. *Taketomi T., Uchiyama H., Ikeda S.* Visual SLAM algorithms: A survey from 2010 to 2016 // *IPSP transactions on computer vision and applications.* – 2017. – Vol. 9. – P. 1-11.
12. *Davison A.J. et al.* MonoSLAM: Real-time single camera SLAM // *IEEE transactions on pattern analysis and machine intelligence.* – 2007. – Vol. 29, No. 6. – P. 1052-1067.
13. *Klein G., Murray D.* Parallel tracking and mapping for small AR workspaces // *2007 6th IEEE and ACM international symposium on mixed and augmented reality.* – IEEE, 2007. – P. 225-234.
14. *Mur-Artal R., Montiel J.M.M., Tardos J.D.* ORB-SLAM: a versatile and accurate monocular SLAM system // *IEEE transactions on robotics.* – 2015. – Vol. 31, No. 5. – P. 1147-1163.
15. *Mur-Artal R., Tardós J.D.* Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras // *IEEE transactions on robotics.* – 2017. – Vol. 33, No. 5. – P. 1255-1262.
16. *Rublee E. et al.* ORB: An efficient alternative to SIFT or SURF // *2011 International conference on computer vision.* – Ieee, 2011. – P. 2564-2571.
17. *Engel J., Schöps T., Cremers D.* LSD-SLAM: Large-scale direct monocular SLAM // *European conference on computer vision.* – Cham: Springer International Publishing, 2014. – P. 834-849.
18. *Engel J., Stückler J., Cremers D.* Large-scale direct SLAM with stereo cameras // *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS).* – IEEE, 2015. – P. 1935-1942.
19. *Smirnova E., Stepanov D., Goryunov V.* A technique of natural visual landmarks detection and description for mobile robot cognitive navigation // *Annals of 26th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM.* – 2015.
20. *Kalal Z., Mikolajczyk K., Matas J.* Tracking-learning-detection // *IEEE transactions on pattern analysis and machine intelligence.* – 2011. – Vol. 34, No. 7. – P. 1409-1422.
21. *Бахшиев А.В., Корбан П.А., Курпаны Н.А.* Программный комплекс определения пространственной ориентации объектов по телевизионному изображению в задаче космической стыковки // *Extreme Robotics.* – 2013. – Т. 1, №. 1. – С. 288-293.
22. *Korsakov A.M. et al.* Determination of an Unmanned Mobile Object Orientation by Natural Landmarks // *AIST (Supplement).* – 2016. – P. 91-101.
23. *Курпаны Н.А., Степанов Д.Н.* Программа моделирования сцены с реперными маркерами и исследования методов определения положения объектов. – 2016 (№ рег. св. 2016612470).
24. *Bakhshiev A., Demcheva A., Stankevich L.* CSNM: the compartmental spiking neuron model for developing neuromorphic information processing systems // *Advances in Neural Computation, Machine Learning, and Cognitive Research V: Selected Papers from the XXIII International Conference on Neuroinformatics, October 18-22, 2021, Moscow, Russia.* – Springer International Publishing, 2022. – P. 327-333.
25. *Boiko A., Bakhshiev A., Korsakov A.* The Hardware Implementation of the Compartmental Spiking Neuron Model (CSNM) Based on Single Supply Operational Amplifiers // *International Conference on Neuroinformatics.* – Cham: Springer Nature Switzerland, 2024. – P. 48-57.
26. *Kirillov A. et al.* Segment anything // *Proceedings of the IEEE/CVF International Conference on Computer Vision.* – 2023. – P. 4015-4026.
27. *Yang Z., Yang Y.* Decoupling features in hierarchical propagation for video object segmentation // *Advances in Neural Information Processing Systems.* – 2022. – Vol. 35. – P. 36324-36336.
28. *He K. et al.* Deep residual learning for image recognition // *Proceedings of the IEEE conference on computer vision and pattern recognition.* – 2016. – P. 770-778.
29. *Liu Z. et al.* A convnet for the 2020s // *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* – 2022. – P. 11976-11986.

30. Fomin I. et al. Investigation of a Spike Segment Neuron in the Offline Multi-Object Tracking Task with Embeddings Constructed by a Convolutional Network // International Conference on Neuroinformatics. – Cham: Springer Nature Switzerland, 2023. – P. 346-354.
31. Ren S. et al. Faster R-CNN: Towards real-time object detection with region proposal networks // IEEE transactions on pattern analysis and machine intelligence. – 2016. – Vol. 39, No. 6. – P. 1137-1149.
32. He K. et al. Mask r-cnn // Proceedings of the IEEE international conference on computer vision. – 2017. – P. 2961-2969.
33. Varghese R., Sambath M. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness // 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). – IEEE, 2024. – P. 1-6.
34. Wang A. et al. Yolov10: Real-time end-to-end object detection // arXiv preprint arXiv:2405.14458. – 2024.
35. Alhasanat M.N. et al. RetinaNet-based Approach for Object Detection and Distance Estimation in an Image // International Journal on Communications Antenna and Propagation (IRECAP). – 2021. – Vol. 11, No. 1. – P. 1-9.
36. Duan K. et al. CenterNet++ for object detection // IEEE transactions on pattern analysis and machine intelligence. – 2023.
37. Chandarana P., Ou J., Zand R. An adaptive sampling and edge detection approach for encoding static images for spiking neural networks // 2021 12th International Green and Sustainable Computing Conference (IGSC). – IEEE, 2021. – P. 1-8.

REFERENCES

1. Fox D., Burgard W., Thrun S. The dynamic window approach to collision avoidance, *IEEE Robotics & Automation Magazine*, 1997, Vol. 4, No. 1, pp. 23-33.
2. Brock O., Khatib O. High-speed navigation using the global dynamic window approach, *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*. IEEE, 1999, Vol. 1, pp. 341-346.
3. Pan S., Xie Z., Jiang Y. Sweeping robot based on laser SLAM, *Procedia Computer Science*, 2022, Vol. 199, pp. 1205-1212.
4. Hicks II R.W., Hall E.L. Survey of robot lawn mowers, *Intelligent Robots and Computer Vision XIX: Algorithms, Techniques, and Active Vision*. SPIE, 2000, Vol. 4197, pp. 262-269.
5. Kiril'chenko A.A., Platonov A.K., Sokolov S.M. Teoreticheskie aspekty organizatsii interpretiruyushchey navigatsii mobil'nogo robota [Theoretical Aspects of mobile robot interpreting navigation], *Preprint In-ta prikl. matem. im. MV Keldysha RAN* [Preprint of M.V. Keldysh RSA institute of applied math.], 2002, No. 5, pp. 40.
6. Orchard G. et al. Converting static image datasets to spiking neuromorphic datasets using saccades, *Frontiers in neuroscience*, 2015, Vol. 9, pp. 437.
7. Fomin I., Korsakov A., Arkhipov A. Comparison of Key Frames of Video Sequences Using the Izhikevich Spiking Neuron, *2023 International Ural Conference on Electrical Power Engineering (UralCon)*. IEEE, 2023, pp. 594-598.
8. Ivanova V.V., Berkman D.A., Korsakov A.M. Metod umen'sheniya razmernosti prostranstva priznakov pri reshenii zadachi klassifikatsii [A method for reducing the dimension of the feature space when solving a classification problem], *Perspektivnye sistemy i zadachi upravleniya* [Perspective systems and tasks of control], 2023, pp. 591-596.
9. Korsakov A.M., Stepanov D.N., Smirnova E.Yu. Algoritm opredeleniya orientatsii transportnogo sredstva na perekrestke po dannym sistemy tekhnicheskogo zreniya [An algorithm for determining the orientation of a vehicle at an intersection based on data from the vision system], *Intellektual'nye sistemy, upravlenie i mekhatronika – 2016* [Intellectual systems, control and mechatronics-2016], 2016, pp. 396-399.
10. Ivanova V.V., Demcheva A.A., Korsakov A.M. Neyromorfnyy mekhanizm upravleniya zadaniem po rezul'tatam analiza situatsii [Neuromorphic mechanism of task control using situation analysis], *Ekstremal'naya robototekhnika* [Extreme robotics], 2024, Vol. 35, No. 1, pp. 308-3015.
11. Taketomi T., Uchiyama H., Ikeda S. Visual SLAM algorithms: A survey from 2010 to 2016, *IPSN transactions on computer vision and applications*, 2017, Vol. 9, pp. 1-11.
12. Davison A.J. et al. MonoSLAM: Real-time single camera SLAM, *IEEE transactions on pattern analysis and machine intelligence*, 2007, Vol. 29, No. 6, pp. 1052-1067.
13. Klein G., Murray D. Parallel tracking and mapping for small AR workspaces, *2007 6th IEEE and ACM international symposium on mixed and augmented reality*. IEEE, 2007, pp. 225-234.
14. Mur-Artal R., Montiel J.M.M., Tardos J.D. ORB-SLAM: a versatile and accurate monocular SLAM system, *IEEE transactions on robotics*, 2015, Vol. 31, No. 5, pp. 1147-1163.
15. Mur-Artal R., Tardos J.D. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras, *IEEE transactions on robotics*, 2017, Vol. 33, No. 5, pp. 1255-1262.
16. Rublee E. et al. ORB: An efficient alternative to SIFT or SURF, *2011 International conference on computer vision*. Ieee, 2011, pp. 2564-2571.

17. Engel J., Schöps T., Cremers D. LSD-SLAM: Large-scale direct monocular SLAM, *European conference on computer vision*. Cham: Springer International Publishing, 2014, pp. 834-849.
18. Engel J., Stückler J., Cremers D. Large-scale direct SLAM with stereo cameras, *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2015, pp. 1935-1942.
19. Smirnova E., Stepanov D., Goryunov V. A technique of natural visual landmarks detection and description for mobile robot cognitive navigation, *Annals of 26th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM*, 2015.
20. Kalal Z., Mikolajczyk K., Matas J. Tracking-learning-detection, *IEEE transactions on pattern analysis and machine intelligence*, 2011, Vol. 34, No. 7, pp. 1409-1422.
21. Bakhshiev A.V., Korban P.A., Kirpan' N.A. Programmiy kompleks opredeleniya prostranstvennoy orientatsii ob'ektov po televizionnomu izobrazheniyu v zadache kosmicheskoy stykovki [A software package for determining the spatial orientation of objects based on a television image in the task of space docking], *Extreme Robotics [Extreme Robotics]*, 2013, Vol. 1, No. 1, pp. 288-293.
22. Korsakov A.M. et al. Determination of an Unmanned Mobile Object Orientation by Natural Landmarks, *AIST (Supplement)*, 2016, pp. 91-101.
23. Kirpan' N.A., Stepanov D.N. Programma modelirovaniya stseny s repnymi markerami i issledovaniya metodov opredeleniya polozheniya ob'ektov [A program for modeling a scene with reference markers and researching methods for determining the position of objects], 2016 registration No. 2016612470).
24. Bakhshiev A., Demcheva A., Stankevich L. CSNM: the compartmental spiking neuron model for developing neuromorphic information processing systems, *Advances in Neural Computation, Machine Learning, and Cognitive Research V: Selected Papers from the XXIII International Conference on Neuroinformatics, October 18-22, 2021, Moscow, Russia*. Springer International Publishing, 2022, pp. 327-333.
25. Boiko A., Bakhshiev A., Korsakov A. The Hardware Implementation of the Compartmental Spiking Neuron Model (CSNM) Based on Single Supply Operational Amplifiers, *International Conference on Neuroinformatics*. Cham: Springer Nature Switzerland, 2024, pp. 48-57.
26. Kirillov A. et al. Segment anything, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015-4026.
27. Yang Z., Yang Y. Decoupling features in hierarchical propagation for video object segmentation, *Advances in Neural Information Processing Systems*, 2022, Vol. 35, pp. 36324-36336.
28. He K. et al. Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
29. Liu Z. et al. A convnet for the 2020s, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 11976-11986.
30. Fomin I. et al. Investigation of a Spike Segment Neuron in the Offline Multi-Object Tracking Task with Embeddings Constructed by a Convolutional Network, *International Conference on Neuroinformatics*. Cham: Springer Nature Switzerland, 2023, pp. 346-354.
31. Ren S. et al. Faster R-CNN: Towards real-time object detection with region proposal networks, *IEEE transactions on pattern analysis and machine intelligence*, 2016, Vol. 39, No. 6, pp. 1137-1149.
32. He K. et al. Mask r-cnn, *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961-2969.
33. Varghese R., Sambath M. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness, *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*. IEEE, 2024, pp. 1-6.
34. Wang A. et al. Yolov10: Real-time end-to-end object detection, *arXiv preprint arXiv:2405.14458*, 2024.
35. Alhasanat M.N. et al. RetinaNet-based Approach for Object Detection and Distance Estimation in an Image, *International Journal on Communications Antenna and Propagation (IRECAP)*, 2021, Vol. 11, No. 1, pp. 1-9.
36. Duan K. et al. CenterNet++ for object detection, *IEEE transactions on pattern analysis and machine intelligence*. – 2023.
37. Chandarana P., Ou J., Zand R. An adaptive sampling and edge detection approach for encoding static images for spiking neural networks, *2021 12th International Green and Sustainable Computing Conference (IGSC)*. IEEE, 2021, pp. 1-8.

Фомин Иван Сергеевич – Государственный научный центр РФ – Федеральное государственное автономное научное учреждение «Центральный научно-исследовательский и опытно-конструкторский институт робототехники и технической кибернетики»; e-mail: i.fomin@rtc.ru; г. Санкт-Петербург, Россия; тел.: +78125523351; научный сотрудник.

Матвеев Виктор Дмитриевич – Государственный научный центр РФ – Федеральное государственное автономное научное учреждение «Центральный научно-исследовательский и опытно-конструкторский институт робототехники и технической кибернетики»; e-mail: v.matveev@rtc.ru; г. Санкт-Петербург, Россия; тел.: +78125523351; инженер.

Архипов Андрей Евгеньевич – Государственный научный центр РФ – Федеральное государственное автономное научное учреждение «Центральный научно-исследовательский и опытно-конструкторский институт робототехники и технической кибернетики»; e-mail: a.arkhipov@rtc.ru; г. Санкт-Петербург, Россия; тел.: +78125523351; м.н.с.

Fomin Ivan Sergeevich – Russian State Scientific Center for Robotics and Technical Cybernetics; e-mail: i.fomin@rtc.ru; Saint-Petersburg, Russia; phone: +78125523351; researcher.

Matveev Victor Dmitrievich – Russian State Scientific Center for Robotics and Technical Cybernetics; e-mail: v.matveev@rtc.ru; Saint-Petersburg, Russia; phone: +78125523351; engineer.

Arkhipov Andrey Evgenievich – Russian State Scientific Center for Robotics and Technical Cybernetics; e-mail: a.arkhipov@rtc.ru; Saint-Petersburg, Russia; phone: +78125523351; junior researcher.

УДК 004.8

DOI 10.18522/2311-3103-2025-4-284-297

С.С. Обаид, В.А. Погонин, И.Б. Кирина**ИНТЕГРАЦИЯ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ПОВЫШЕНИЯ ОТКАЗОУСТОЙЧИВОСТИ МОДЕЛИ ВЛАГОПЕРЕНОСА В СИСТЕМЕ «УМНЫЙ САД»**

Представлено исследование по разработке и интеграции рекуррентной нейронной сети (RNN) для повышения точности и отказоустойчивости модели влагопереноса в системе «Умный сад». Проблема контроля влажности почвы становится особенно актуальной в современных условиях сельского хозяйства и экологического мониторинга, где требуется высокая точность для управления водными ресурсами, прогнозирования урожайности и предотвращения засушливых периодов. Традиционные методы, такие как дистанционное зондирование и модели влагопереноса, имеют существенные ограничения: низкая точность, сложность вычислений, зависимость от точных данных сенсоров и сложность применения в реальных полевых условиях. Для решения этих проблем в исследовании предлагается использование RNN, способной эффективно обрабатывать временные ряды данных и прогнозировать влажность почвы даже при наличии неполных, неточных или искажённых входных данных. В качестве исходных данных использованы глобальный датасет влажности почвы GSSM и погодные данные платформы Meteostat, что позволило учесть климатические особенности регионов с разными типами почвы. Модель включает слой долговременной краткосрочной памяти (LSTM) и полносвязный слой для финального прогнозирования. Особое внимание уделено предобработке данных, включая расчёт среднесуточных, среднемесячных и среднегодовых значений, а также коррекцию данных с учётом характеристик различных типов почв. Проведённое исследование показало, что разработанная модель RNN обладает высокой устойчивостью к сбоям в работе сенсоров, минимальной зависимостью от объёма входных данных и способностью адаптироваться к разным климатическим и почвенным условиям. Предлагаемое решение позволяет улучшить точность мониторинга влажности почвы в системе «Умный сад», оптимизировать использование водных ресурсов и повысить стабильность работы системы в условиях изменяющихся внешних факторов. Таким образом, интеграция RNN открывает новые возможности для развития сельского хозяйства и экологии, обеспечивая более эффективное управление водными ресурсами и повышение производительности агросистем.

Рекуррентная нейронная сеть; влагоперенос в почве; «Умный сад»; отказоустойчивость; прогнозирование.

S.S. Obaid, V.A. Pogonin, I.B. Kirina**INTEGRATION OF A RECURRENT NEURAL NETWORK TO INCREASE THE FAILURE TOLERANCE OF THE MOISTURE TRANSFER MODEL IN THE SMART GARDEN SYSTEM**

The paper presents a study on the development and integration of a recurrent neural network (RNN) to improve the accuracy and fault tolerance of a moisture transfer model in a smart garden system. The problem of soil moisture control is becoming especially relevant in modern agricultural and environmental monitoring, where high accuracy is required to manage water resources, forecast crop yields and prevent drought periods. Traditional methods, such as remote sensing and moisture transfer models, have significant limitations: low accuracy, computational complexity, dependence on accurate sensor data and difficulty in applying in real field conditions. To solve these problems, the study proposes the use of RNN,