

Step detection via deep learning

А. В. Филиппова¹, Т. Гадаев¹, В. В. Стрижов¹

¹Московский физико-технический институт

В данной работе рассматривается задача предсказания траектории человека по показаниям акселерометра и гироскопа, которые установлены в телефоне. Так как система отсчета, связанная с устройством, постоянно вращается и движется ускоренно относительно мировой системы, поставленная задача не является тривиальной. Существует много различных необучаемых алгоритмов для описания траектории человека. Минус этих алгоритмов в том, что модель не может подстраиваться под конкретную постановку задачи и учитывать детали (пол, возраст, особенность походки объекта). В данной работе предлагается нейросетевой подход для решения задачи, а также описываются полезные эвристики.

Наше исследование можно разделить на три логические части: проектирование устойчивых к ошибкам гироскопов кватернионов, прогнозирование изменения положения на фиксированном периоде с помощью различных нейронных сетей и применение идеи детекции шагов с целью улучшения показаний модели и уменьшения дрейфа.

Ключевые слова: акселерометр и гироскоп; детекция шагов ;кватернионы; дрейфт; нейросетевой подход; предсказание траектории

1 Введение

Задача точного определения положения смартфона в пространстве, и, как следствие, оценка местоположения объекта решается с высокой точностью на открытых площадках с использованием GPS [1]. Современные технологии демонстрируют отличные результаты при отклонении менее чем на несколько метров [2]. Однако у системы есть недостаток: она требует открытого пространства между устройством и спутником для передачи радиосигналов. В реальном мире нас часто окружают деревья, неровности ландшафта, высокие здания. Качество геолокации снижается из-за отражения радиоволн. Например, использование GPS для отслеживания траекторий внутри зданий практически бесполезно [3]. В этом случае используются методы, основанные на данных других датчиков. Наиболее распространенными датчиками IMU смартфона являются гироскоп, магнитометр и акселерометр. Основной проблемой такого подхода является накопление ошибок позиционирования из-за дрейфа, вызванного несовершенствами и шумом в датчиках (тут будет ссылка на статью моего куратора, которая пока не была опубликована). В данной работе предлагается помимо методов, описанных в статье (тут будет ссылка на статью моего куратора, которая пока не была опубликована) использовать детекцию шагов с целью улучшения показаний модели и уменьшения дрейфа.

2 Данные

В эксперименте мы использовали набор данных RuDaCor [4]. Данные состоят из 1200 измерений траекторий для разных положений смартфона (в руке, в сумке, в кармане брюк). Для каждого объекта сняты показания акселерометра, гироскопа, координаты человека в мировой системе отсчета, состояние правой и левой ног (0 - поднята, 1 - на земле). При сборке данных соблюдены следующие требования:

1. Траектории находятся на плоских горизонтальных поверхностях - нет лестниц или значительных изменений высоты ландшафта.

2. Все траектории замкнуты. Это означает, что точка начала равна точке финиша. Участников попросили использовать маркер для отметки положения старта, что позволяет утверждать, что разница в исходном и финишном положениях не более 5 см для каждой ноги.
 3. Участники только ходили, другого движения (прыжков/бега) не было.
 4. Участники не были ограничены в скорости движения.
- Пример показаний акселерометра представлен на изображении 1.

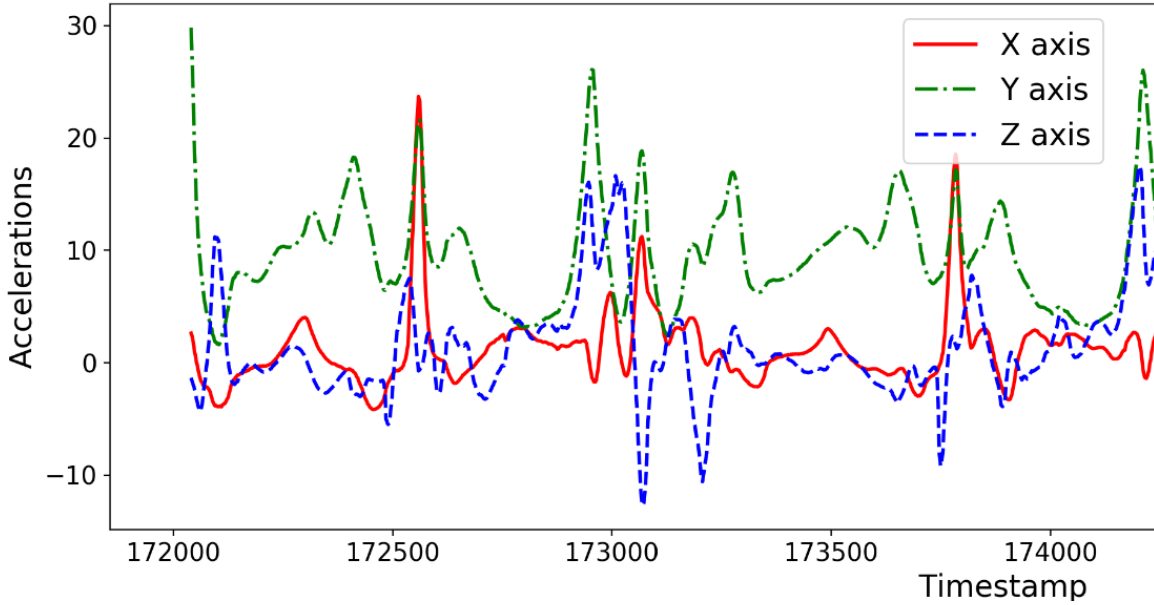


Рис. 1 Пример показаний акселерометра

3 Постановка задачи

Задача состоит в том, чтобы найти суперпозицию функций, которые мы обозначим как F_{tr+st} , которая преобразует данные датчиков в оценку траектории, которая будет близка к истинной, а также дает оценку вероятности совершения шага вдоль траектории в каждый момент времени.

$$\arg \min_{F_{tr+st}} \mathcal{L}(F_{tr+st}(\mathcal{A}, \mathcal{W}), \mathcal{T}, \mathcal{S}) \quad (1)$$

В качестве функции потерь предлагается использовать комбинированную функцию $\mathcal{L}(F_{tr}(\mathcal{A}, \mathcal{W}), \mathcal{T}, \mathcal{S}) = \text{MSE}(F_{tr}(\mathcal{A}, \mathcal{W}), \mathcal{T}) + \text{BCElogloss}(F_{st}(\mathcal{A}, \mathcal{W}), \mathcal{S})$.

Данная функция потерь позволяет обучить модель таким образом, чтобы для вещественных выходов модели решалась задача регрессии, для категориальных - классификации.

3.1 Предобработка данных

Показания акселерометра и гироскопа нарезаются на перекрывающиеся окошки фиксированной длины. Искомое отображение F_{tr+st} по окошку выдается двумерный вектор - смещение за количество временных отсчетов равных длине окна (скорость человека в данном временном интервале).

$(\mathcal{A}, \mathcal{W}) \subset \mathbb{R}^{T \times 6}$, где T - количество отсчетов (временной показатель), $(a_t, w_t)^T \in \mathbb{R}^6$, $t \in [1..T]$ - вектор, составленный из показаний акселерометра и гироскопа в момент вре-

мени t . Для того чтобы получить обучающую выборку, используем скользящее окно шириной w и размером шага s_w . Разобьем ряд $(\mathcal{A}, \mathcal{W}) = \mathcal{X}$ на подпоследовательности $Y_j = \{\mathbf{x}_i, \dots, \mathbf{x}_{i+w-1}\}$, где $\mathbf{x}_i = (a_x^{(i)}, a_y^{(i)}, a_z^{(i)}, w_x^{(i)}, w_y^{(i)}, w_z^{(i)})$ - одно измерение. Получим множество $Y = \{Y_j, j \in [1..m]\}$, где $m = \frac{T-w}{s_w}$ - количество подпоследовательностей, $i = w(j-1) + 1$ - номер измерения, которое является первым в подпоследовательности под номером j . Каждой подпоследовательности $Y_j = \{\mathbf{x}_{w(j-1)+1}, \dots, \mathbf{x}_{wj}\}$ длины w поставим в соответствие точку (вектор) $(\mathbf{v}_{wj}, \mathbf{s}_{wj})$, $\mathbf{v}_{wj} \in \mathbb{R}^2$ - скорость в плоскости движения, $\mathbf{s}_{wj} \in \{0, 1\}^2$ - метка положения ноги человека (0 - нога движется, 1 - нога покоится, находится на земле). То есть задачу можно сформулировать следующим образом: необходимо обучить нейронную сеть по подпоследовательности длиной w предсказывать скорость и положение ног.

3.2 Используемые эвристики

Так как система (смартфон) относительно мировой глобальной системы отсчета постоянно движется и вращается, удобно перейти в некую стабилизированную систему отсчета S , в которой вектор гравитации всегда сонаправлен с осью Oz . Авторы статьи [5] используют данную идею с целью уменьшения углов крена и тангажа. Мы считаем, что это хорошая практика и применяем данный подход для решения поставленной задачи. Ориентация системы отсчета представлена на изображении 2.

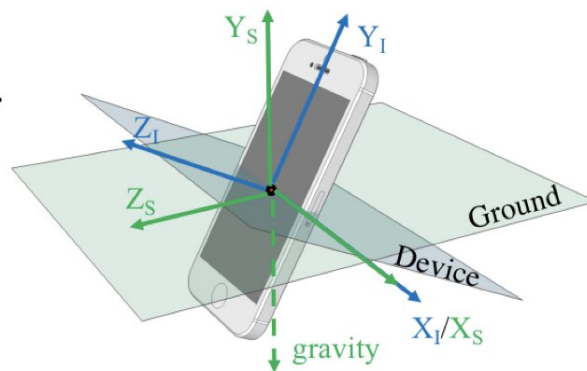


Рис. 2 Ориентация стабилизированной системы отсчета

4 Эксперимент

4.1 Архитектура

Известно, что сверточные сети умеют хорошо выявлять локальную информацию в объектах благодаря обучаемому скрытому представлению (детекция изображений, сегментация изображений). Рекуррентные нейронные сети хорошо подходят для задач регрессии временных рядов, так как имеют способность "хранить" информацию. В данной статье предлагается использовать некоторую модификацию известной архитектуры ResNet [6], состоящей из чередования сверточных слоев, пулингов и residual block - техника позволяющая решать проблему "утекающего градиента". Наша архитектура ResNetLSTM - сетка, в которой последовательно соединены ResNet и двухслойная LSTM. Архитектура сетки представлена на изображении 3.

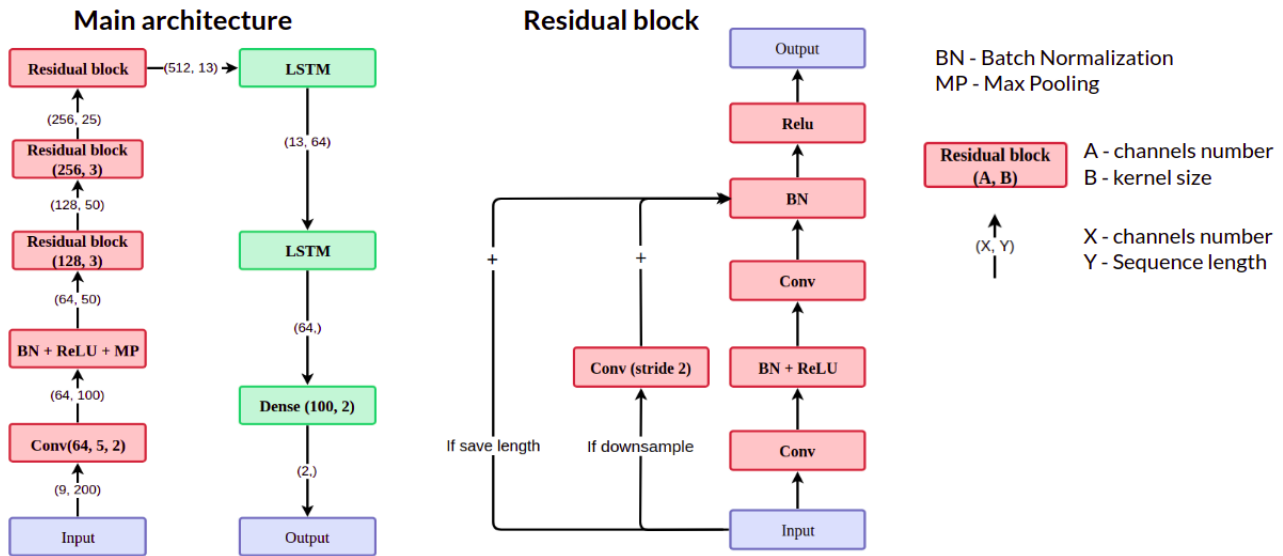


Рис. 3 Схема нейронной сети

4.2 Обучение

Нейронная сеть обучается таким образом, чтобы по окошку определять перемещение человека и положение ноги. Как уже было сказано в разделе 3, для обучения используется комбинированная функция ошибки с целью уменьшения дрейфа.

Также мы попробовали использовать немного иную стратегию и вместо последнего вектора из выхода LSTM в нейронной сети, брать все векторы как последовательные предсказания скорости. Обе стратегии иллюстрированы на изображениях 4(a) и 4(b) соответственно.

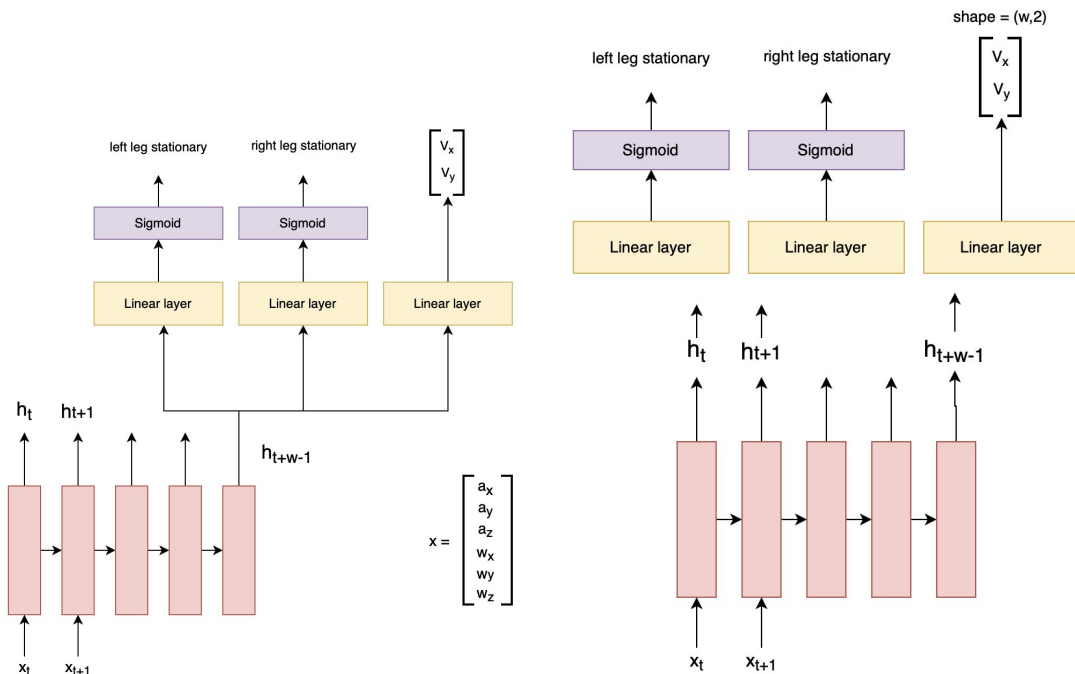


Рис. 4

87 На изображение 3 представлен график зависимости значения функции ошибки во время
 88 обучения и валидации от номера эпохи.

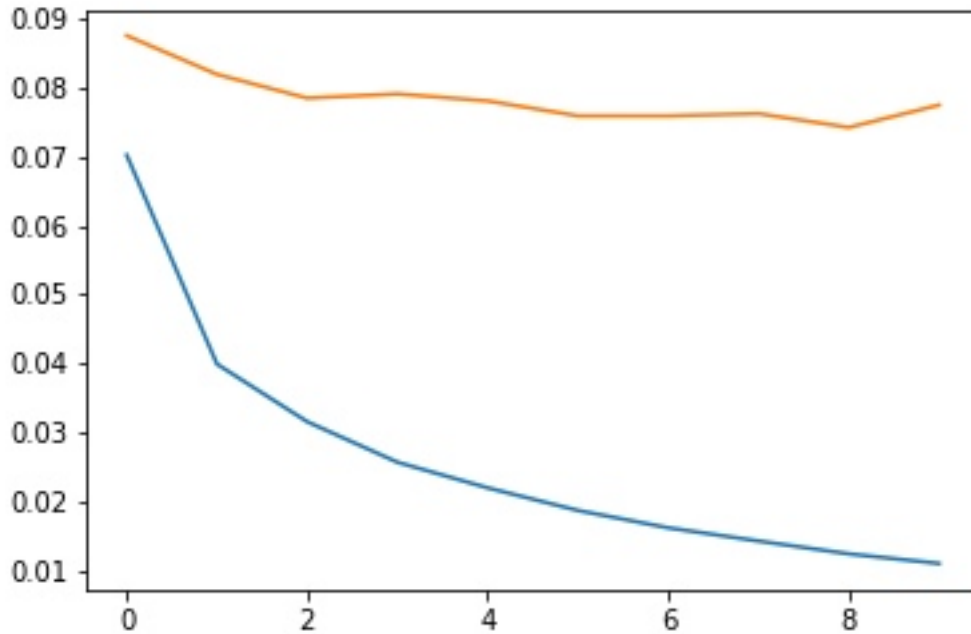


Рис. 5 График зависимости значения функции ошибки во время обучения и валидации.

89 Как видно из графика, оптимизируемая функция достаточно быстро сходится к своему
 90 минимуму, поэтому при обучении достаточно около 8 – 10. Продолжительность каждой
 91 эпохи около 5 минут.

92 4.3 Результаты

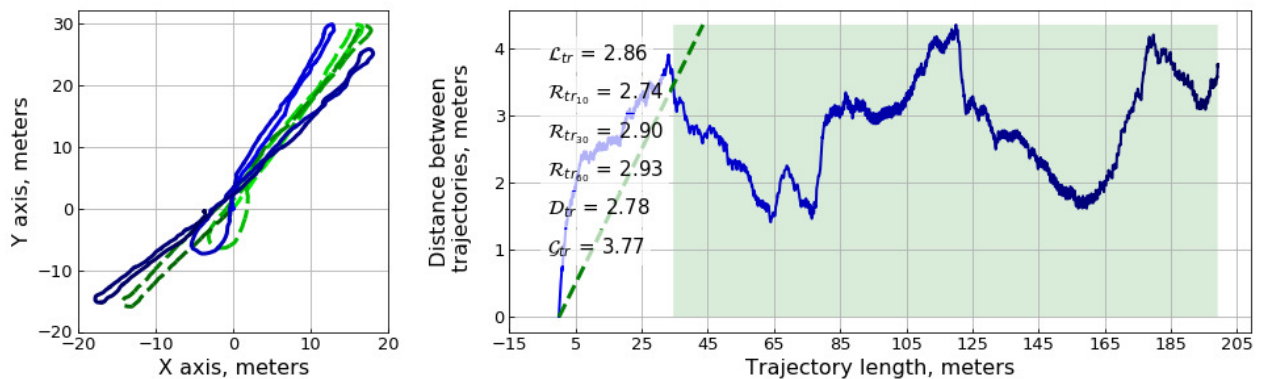


Рис. 6 Пример предсказанной траектории.

Model Steps	GAP	MIE	RMSE	RTE		
	\mathcal{G}_{tr}	\mathcal{D}_{tr}	\mathcal{L}_{tr}	$\mathcal{R}_{tr, w=10}$	$\mathcal{R}_{tr, w=30}$	$\mathcal{R}_{tr, w=60}$
RL	14.54	6.91	8.71	7.09	7.41	7.91
RL _{0.01_BCE}	11.20	6.42	7.71	6.42	6.70	7.14
RL_pc_tc	11.22	6.80	8.07	6.76	7.04	7.45
RL_tc	15.62	8.74	10.39	8.79	9.18	9.77
RL_trend	13.70	7.61	9.09	7.66	7.99	8.55
RL_trend_pc	11.26	6.62	7.90	6.61	6.88	7.32
RL_trend_pc_tc	11.66	6.93	8.24	6.91	7.19	7.63
RL_trend_tc	16.94	9.16	10.97	9.28	9.69	10.38
R	14.41	7.54	9.25	7.65	8.02	8.60
R_pc	11.20	6.40	7.75	6.40	6.68	7.11
R_pc_tc	11.03	6.50	7.81	6.48	6.76	7.18
R_tc	15.50	8.32	10.06	8.41	8.83	9.41
R_trend	13.75	7.15	8.75	7.29	7.63	8.20
R_trend_pc	11.15	6.29	7.65	6.31	6.60	7.04
R_trend_pc_tc	11.96	6.71	8.10	6.74	7.04	7.52
R_trend_tc	15.6	8.21	9.92	8.36	8.76	9.36

Таблица 1 Results of the experiments of various features with different architectures. Here R — ResNet-18, L — LSTM, RL — sequential model of ResNet-18 and LSTM, pc — bias reduction on inference, tc — bias reduction for training set, trend — adding trend of gyroscope signal to the network input. Experiment was conducted using KFold technique with 3 folds. The best result for each metric is bold.

Литература

- [1] AH Mohamed and KP Schwarz. Adaptive kalman filtering for ins/gps. *Journal of geodesy*, 73(4):193–203, 1999.
- [2] Wan Rahiman and Zafariq Zainal. An overview of development gps navigation for autonomous car. In *2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA)*, pages 1112–1118. IEEE, 2013.
- [3] George Dedes and Andrew G Dempster. Indoor gps positioning-challenges and opportunities. In *VTC-2005-Fall. 2005 IEEE 62nd Vehicular Technology Conference, 2005.*, volume 1, pages 412–415. Citeseer, 2005.
- [4] Andrey Bayev, Ivan Chistyakov, Alexey Derevyankin, Ilya Gartseev, Alexey Nikulin, and Mikhail Pikhletsky. Rudacop: The dataset for smartphone-based intellectual pedestrian navigation. In *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8. IEEE, 2019.
- [5] Hang Yan, Sachini Herath, and Yasutaka Furukawa. Ronin: Robust neural inertial navigation in the wild: Benchmark, evaluations, and new methods. *CoRR*, abs/1905.12853, 2019.
- [6] Shaoqing Ren, Kaiming He, Xiangyu Zhang. Deep residual learning for image recognition. 2015.

Поступила в редакцию