
Методи вирішення задачі розпізнавання людської активності за допомогою генеративних мереж

Денис Каленіченко

Інститут Прикладного Системного Аналізу, КПІ ім. Ігоря Сікорського, Київ, Україна

Валерій Данилов

Інститут Прикладного Системного Аналізу, КПІ ім. Ігоря Сікорського, Київ, Україна

ORCID 0000-0003-3389-3661

Для цитування цієї статті:

Каленіченко Денис, Данилов Валерій. Методи вирішення задачі розпізнавання людської активності за допомогою генеративних мереж. International Science Journal of Engineering & Agriculture. Vol. 3, No.6, 2024, pp. 10-15. doi: 10.46299/j.isjea.20240306.02.

Надійшла до редакції: 09 листопада 2024 р.; **Схвалено:** 30 листопада 2024 р.;

Опубліковано: 01 грудня 2024 р.

Анотація: носимі пристрої, такі як розумні-годинники, здатні передавати значні обсяги даних, зібраних різноманітними сенсорами. У звичайних умовах ці дані є нерозміченими та потребують застосування методів навчання без підкріплення. Генеративні мережі, зокрема варіаційні автокодувальники, мають архітектуру, придатну для дослідження загальної структури даних, отриманих з носимих пристроїв. Автори запропонували систему, що складається з сукупності генеративних мереж для розпізнавання активності людини.

Ключові слова: глибоке навчання, генеративні мережі, автокодувальник, варіаційний автокодувальник, напівкерване навчання.

1. Вступ

У сучасному світі носимі пристрої передають значний обсяг даних на телефон або сервер. Частиною цих даних є вимірювання з акселерометра, гіроскопа, магнетометра тощо.

Найпоширенішими методами вирішення задачі класифікації активності є навчання з підкріпленням:

- Автори [1] використали класифікатор XGBoost [2] і досягли точності 84% у задачі розпізнавання п'яти видів активності.
- Автори [3] застосували мережу LSTM [4] і досягли точності 92,1%, використовуючи датасет WISDM [5].

Проте зазвичай ці дані є нерозміченими, оскільки користувачі не будуть самостійно їх анотувати, а залучення спеціалістів для розмітки наявних даних може бути неможливим через специфіку інформації.

З іншого боку, отримання анонімізованих нерозмічених даних є тривіальною задачею. Це вагомими причинами для глибшого дослідження можливостей мереж, здатних до самостійного навчання.

2. Об'єкт і предмет дослідження

Варіаційний автокодувальник (VAE [6]) — це нейронна мережа, яка навчається зменшувати розмірність вхідних даних, водночас генеруючи нові зразки, схожі на оригінальні.

На відміну від традиційних автокодувальників, VAE формує латентний простір у вигляді ймовірнісного розподілу, що дозволяє моделювати та відтворювати закономірності у вихідних даних. Це особливо корисно для задач розпізнавання активності.

2.1. Ключові аспекти VAE для аналізу людської активності

- Латентний простір як ймовірнісний розподіл: VAE перетворює вхідні дані в латентний простір, де кожна активність (наприклад, ходьба, біг, сидіння) представлена як гаусівський розподіл. Це дозволяє зберігати варіації всередині кожної активності, що є корисним для виявлення закономірностей або аномалій.

- Реконструкція та генерація даних: Під час навчання VAE мінімізує різницю між вхідними та відновленими даними, що дозволяє мережі краще розпізнавати специфічні закономірності активностей. Ця здатність реконструювати вхідні дані допомагає VAE класифікувати активність, оскільки низька похибка реконструкції свідчить про відповідність конкретного класу.

- Виявлення аномалій та нових активностей: Оскільки кожна активність представлена у вигляді ймовірнісного розподілу, VAE здатен виявляти аномалії або нові активності, які не відповідають жодному з уже вивчених розподілів. Це дозволяє VAE гнучко адаптуватися до нових сценаріїв та активностей.

3. Аналіз літератури

Для вирішення задачі класифікації людської активності генеративні мережі були застосовані у [8], де автор використав декілька послідовних автоенкодерів для визначення активності та отримав результат $F=0.77$, однак його датасет був дуже малим та містив усього 6 людей.

Варіаційний автокодувальник був використаний у [9], де показав похибку EER [10]= 0.111 для LSTM VAE.

4. Мета та задачі дослідження

Розробка, оптимізація та застосування групи варіаційних автокодувальників для ефективного розпізнавання та класифікації людської активності на основі даних з акселерометра. Метою є досягнення високої точності класифікації активностей та можливість виявлення аномальних закономірностей або нових типів активності в умовах відсутності розмічених даних.

5. Методи досліджень

Для вирішення задачі класифікації активностей за допомогою варіаційного автокодувальника автором було побудовано 4 окремі мережі для визначення кожного окремого виду активності («спуск по сходах», «підйом по сходах», «ходьба», «повільний біг»), які ми викоремили з MotionSense dataset [11].

Архітектура мережі наведена на малюнку 1, (кодувальни) та малюнку 2 (декодувальник).

Layer (type)	Output Shape	Param #
lstm_12 (LSTM)	(32, 100, 128)	69,120
lstm_13 (LSTM)	(32, 64)	49,408
dense_9 (Dense)	(32, 32)	2,080
dense_10 (Dense)	(32, 32)	2,080

Total params: 122,688 (479.25 KB)
Trainable params: 122,688 (479.25 KB)
Non-trainable params: 0 (0.00 B)

Рис 1. Архітектура кодувальника.

Layer (type)	Output Shape	Param #
repeat_vector_3 (RepeatVector)	(32, 100, 32)	0
lstm_14 (LSTM)	(32, 100, 64)	24,832
lstm_15 (LSTM)	(32, 100, 64)	33,024
time_distributed_3 (TimeDistributed)	(32, 100, 6)	390

Total params: 58,246 (227.52 KB)
Trainable params: 58,246 (227.52 KB)
Non-trainable params: 0 (0.00 B)

Рис 2. Архітектура декодувальника.

Навчалась мережа за допомогою алгоритму Adam [12] для градієнтного спуску 100 епох на 32 батчах.

Під час тренування регуляризатор Кульбака-Лейблера [13] став 0, а тренування моделі зупинилося. Як показує дослідження в [14] – це звична проблема для VAE. Авторами було запропоноване рішення – додати константу від 0 до 1, яка зменшить вплив регуляризатора.

6. Результати досліджень

У цьому алгоритмі було побудовано чотири окремі моделі, кожна з яких навчалася на даних, що відповідають певному типу активності. Дані проходять через автоенкодер, після чого обчислюється похибка реконструкції. Найменша похибка серед автоенкодерів вказує на ідентифіковану активність.

Для розпізнавання активності було створено чотири нейронні мережі, де кожна спеціалізується на класифікації окремої активності. Процес тестування відбувався наступним чином:

1. **Обробка вхідних даних:** Кожна модель незалежно пропускає вхідні дані через свій автоенкодер.

2. **Обчислення похибки реконструкції:** Для кожної моделі розраховується середньоквадратична похибка (MSE [14]) між вхідними даними та їх реконструйованими версіями, отриманими на виході автоенкодера. Ця похибка характеризує якість відновлення даних.

3. **Вибір моделі з найменшою похибкою:** Серед усіх отриманих похибок визначається найменша. Модель з найменшою похибкою реконструкції вважається тією, що найбільш відповідно відображає вхідну активність, і, отже, ідентифікує її.

Таким чином, алгоритм використовує сукупність спеціалізованих автоенкодерів та аналізує їх похибки реконструкції для точного розпізнавання типу активності на основі вхідних даних.

Отримуємо результати, які показані на малюнку 3.

	precision	recall	f1-score	support
Downstairs	0.46	0.85	0.60	2242
Upstairs	0.67	0.74	0.70	2691
Walking	0.93	0.65	0.76	5556
Jogling	0.88	0.68	0.77	2195
accuracy			0.71	12684
macro avg	0.73	0.73	0.71	12684
weighted avg	0.78	0.71	0.72	12684

Рис 3. Пораховані метрики для кожної з активностей.

Порахуємо матрицю помилок, щоб краще зрозуміти природу помилок. Матриця помилок зображена на малюнку 4.

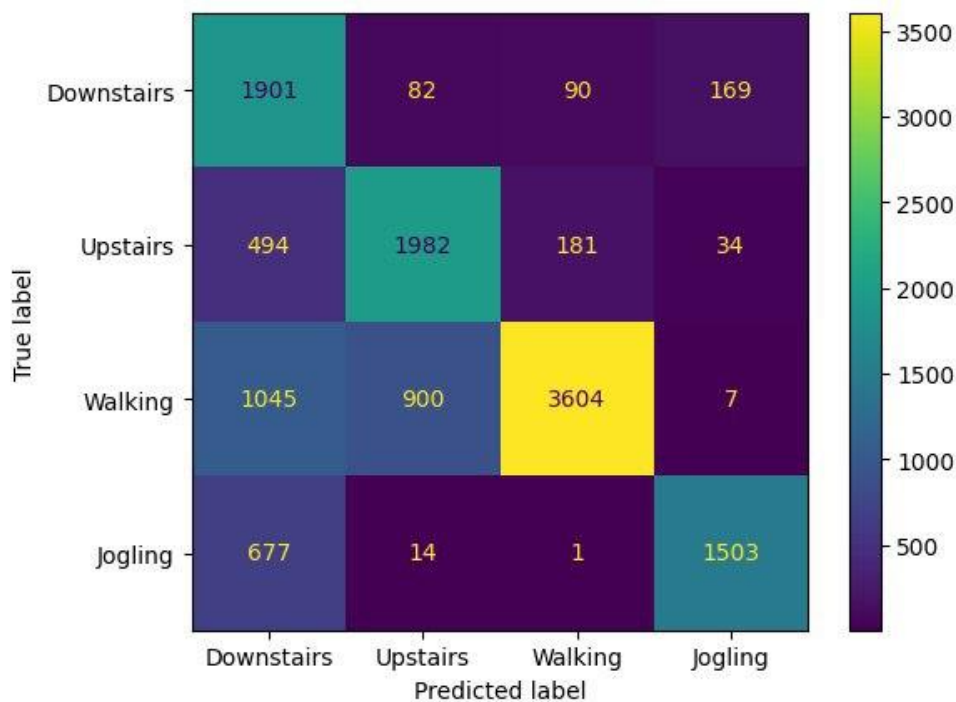


Рис 4. Матриця помилок для кожної з активностей.

Результати показують, що найкраще відпрацювала модель, яка була натренована на даних бігу людей. Враховуючи природу варіаційних автокодувальників їм треба багато даних та довге навчання, щоб найкраще вивчити розподіл вхідних даних. Найменше даних було надано для моделі, яка навчалась на даних «спуску по сходах», що і могло привести до такої точності (f1-score 60%). Найкраще себе показала модель, яка знаходила біг людей – f1-score 77%.

7. Перспективи подальшого розвитку досліджень

Одна з моделей показала f1-score 77%, що показує, що можливо отримати гарні результати, якщо надати достатньо даних. В такому випадку інші моделі потенційно зможуть краще відновлювати дані свого типу, що призведе до зменшення кількості помилок першого типу (класифікація іншої активності як своєї, false positive).

8. Висновки

У цій роботі було досліджено варіаційний автокодувальник та його здатність визначати різні види активностей на основі даних з носимих пристроїв. Використання варіаційного автокодувальника та його модифікацій дозволило досягти певного рівня точності в класифікації.

Моделі, призначені для визначення активностей «підйом по сходах» та «спуск по сходах», продемонстрували нижчу точність порівняно з іншими класами, що може бути зумовлено недостатньою кількістю даних або схожістю між цими активностями. Зокрема, F1-міра для класу «Downstairs» склала 60%, тоді як для «Upstairs» — 70%.

Модель, відповідальна за розпізнавання активності «біг» («Jogging»), показала високу точність із F1-мірою 77%. Це може бути пов'язано з тим, що біг має більш унікальні характеристики, які відрізняють його від інших активностей, що спрощує процес класифікації.

Можливою проблемою є те, що розподіли даних від одних і тих самих користувачів мають спільні риси незалежно від активності. Це може призвести до нижчої похибки реконструкції для певних випадків, що в свою чергу спричиняє неправильну класифікацію. Цю гіпотезу підтверджує дослідження [9], в якому схожий підхід використовувався для задачі біометричної автентифікації.

Отримані результати свідчать про потенціал використання варіаційних автокодувальників у задачах розпізнавання активностей. Проте необхідні подальші дослідження для покращення точності та надійності моделей, особливо для активностей з меншою кількістю даних або високою схожістю між класами.

Список літератури:

- 1) Zhang, W., Zhao, X., & Li, Z. (2019). A comprehensive study of smartphone-based indoor activity recognition via Xgboost. *IEEE Access*, 7, 80027-80042.
- 2) Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- 3) Chen, Y., Zhong, K., Zhang, J., Sun, Q., & Zhao, X. (2016, January). LSTM networks for mobile human activity recognition. In *2016 International conference on artificial intelligence: technologies and applications* (pp. 50-53). Atlantis Press
- 4) Vennerød, C. B., Kjærran, A., & Bugge, E. S. (2021). Long short-term memory RNN. *arXiv preprint arXiv:2105.06756*.
- 5) Human Activity Recognition - WISDM's activity prediction dataset. Available at https://github.com/coloriz/HAR-WISDM_ar
- 6) Kingma, D. P., & Welling, M. (2019). An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12(4), 307-392.
- 7) Normal distribution. Available at https://en.wikipedia.org/wiki/Normal_distribution
- 8) Munoz-Organero, M., & Ruiz-Blazquez, R. (2017). Time-elastic generative model for acceleration time series in human activity recognition. *Sensors*, 17(2), 319
- 9) Gozhyj, A. (2020). Comparative Analysis of using Recurrent Autoencoders for User Biometric Verification with Wearable Accelerometer.

- 10) Equal error rate. Available at <https://en.wikipedia.org/wiki/Biometrics#Performance>
- 11) MotionSense dataset. Available at <https://github.com/mmalekzadeh/motion-sense>
- 12) Kingma, D. P. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- 13) Kullback–Leibler divergence. Available at https://en.wikipedia.org/wiki/Kullback%E2%80%93Leibler_divergence
- 14) Bowman, S. R., Vilnis, L., Vinyals, O., Dai, A. M., Jozefowicz, R., & Bengio, S. (2015). Generating sentences from a continuous space. arXiv preprint arXiv:1511.06349.
- 15) Mean squared error. Available at https://en.wikipedia.org/wiki/Mean_squared_error

Methods for solving the problem of human activity recognition using generative networks

Denys Kalenichenko

Institute for Applied System Analysis, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

Valeriy Danylov

Institute for Applied System Analysis, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine
ORCID 0000-0003-3389-3661

Abstract: Wearable devices, such as smartwatches, are capable of transmitting significant amounts of data collected by a variety of sensors. Under normal conditions, this data is unlabeled and requires the application of non-reinforcement learning methods. Generative networks, in particular variational autoencoders, have an architecture suitable for investigating the general structure of data obtained from wearable devices. The authors proposed a system consisting of a set of generative networks to recognize human activity.

Keywords: Deep Learning, Generative Networks, Autoencoder, Variational Autoencoder, Semi-Guided Learning
