

DOI: <https://doi.org/10.17816/gc623520>

Классификация с использованием машинного обучения и кластеризация для анализа формы дендритных шипиков нейронов

П.И. Васильев^{1, 2}, Д.С. Смирнова², В.С. Чуканов^{1, 2}, И.Б. Безпрозванный^{1, 3}, Е.И. Пчицкая^{1*}¹ Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Российская Федерация;² Высшая школа прикладной математики и вычислительной физики, Санкт-Петербургской политехнический университет Петра Великого, Санкт-Петербург, Российская Федерация;³ UT Southwestern Medical Center at Dallas, Даллас, США

АННОТАЦИЯ

Синапс — зона специализированного контакта между двумя нейронами, через который осуществляется передача информации от одной клетки к другой. Формирование синаптических контактов и передача через них сигналов с помощью электрических импульсов — это фундаментальная особенность нейрональных клеток. Со стороны аксона синапс формируется аксональным бутоном, а со стороны дендрита — дендритным шипиком, специализированным выростом дендритной мембраны. Дендритные шипики характеризуются большим разнообразием форм и размеров, которые в сильной степени варьируются между различными отделами мозга, типами клеток и видами животных. Морфология дендритных шипиков изменяется при нейроонтогенетических и нейродегенеративных заболеваниях, а также в ответ на действие внешних стимулов. Считается, что данный процесс обеспечивает синаптическую пластичность, тем не менее требуются дальнейшие исследования для установления связи между формой и функцией шипиков. Для решения данной задачи в современной нейробиологии имеется необходимость разработки методов анализа морфологии синапсов на трёхмерных изображениях нейронов. Нами разработано программное обеспечение с открытым исходным кодом для сегментации дендритных шипиков из трёхмерных изображений дендритов, вычисления 10 наиболее широко используемых морфологических признаков, адаптированных к трёхмерному представлению [1, 2], и выполнения классификации и кластеризации датасета дендритных шипиков для анализа их формы. Помимо численных признаков описания формы шипика было предложено использовать гистограмму длин хорд (chord length distribution histogram, CLDH). В данном подходе строится набор случайных хорд в объёме дендритного шипа, соединяющих его внешние границы, после чего формируется гистограмма. При достижении количества хорд $n=30\ 000$ вероятностные колебания гистограммы становятся незначительно малы. Описанные метрики использовались для кластеризации и классификации датасета.

Классификация по предопределённым морфологическим группам является широко используемым подходом к анализу морфологии дендритных шипиков. В данном подходе шипики делятся на фиксированные категории, такие как тонкие, грибовидные и пеньковые. Классификация обычно выполняется экспериментатором полуавтоматическим способом, что ведёт к значительной ошибке. Мы разработали основанный на алгоритме машинного обучения инструмент классификации, который классифицирует шипики по указанным выше категориям на основе консенсуса, достигнутого путём ручной разметки обучающего датасета 8 различными экспертами. Точность такого метода при классификации с использованием набора классических морфологических признаков сопоставима с экспертной разметкой (>77%). Данный подход позволяет снизить необъективность и трудоёмкость классификации.

Последние исследования, в том числе с использованием прижизненной микроскопии *in vitro* и *in vivo*, свидетельствуют в пользу того, что формы дендритных шипиков представляют собой континуум, а не чётко разделённые классы [3]. В связи с этим существует необходимость в разработке надёжного метода оценки и изучения морфологии дендритных шипиков. Нами разработан инструмент кластеризации, где количество групп и их содержание определяется данными, а не экспериментатором. Инструмент представлен алгоритмами k -средних и DBSCAN. Для определения количества кластеров представлены три различных метода: метод силуэта; метод плеча; а также новый, разработанный нами, метод максимизации критерия расхождения классов. Данный метод основан на предположении, что качество кластеризации лучше, когда кластеры максимально отличаются друг от друга по количеству представленных в них классов грибовидных/тонких/пеньковых шипиков, размеченных экспертами. Преимущество такого подхода заключается в том, что в нём учитывается конкретный тип данных для кластеризации, в то время как без данной информации оценка качества кластеризации затруднена. Кластеризацию с применением метрики CLDH даёт стабильное число кластеров — $n=5$ для всех

Рукопись получена: 15.05.2023

Рукопись одобрена: 26.11.2023

Опубликована online: 20.01.2024

трёх описанных выше методов. В определённых кластерах содержатся схожие по форме дендритные шипики, что было валидировано экспертами. Для набора классических метрик количество кластеров значительно варьировало: от $n=4$ до $n=14$. Подобные данные позволяют сделать предположение, что более сложная метрика CLDH содержит в себе достаточное количество информации о форме синапса для выполнения кластеризации.

Ключевые слова: синапс; дендритный шипик; грибовидный; тонкий; пеньковый; программное обеспечение; сегментация; классификация; кластеризация; гистограмма распределения длин хорд.

Как цитировать:

Васильев П.И., Смирнова Д.С., Чуканов В.С., Безprozванный И.Б., Пчицкая Е.И. Классификация с использованием машинного обучения и кластеризация для анализа формы дендритных шипиков нейронов // Гены и клетки. 2023. Т. 18, № 4. С. 890–893. DOI: <https://doi.org/10.17816/gc623520>

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение исследования и подготовку статьи, прочли и одобрили финальную версию перед публикацией).

Источник финансирования. Исследование профинансировано Министерством науки и высшего образования РФ в рамках Программы стратегического академического лидерства «Приоритет 2030» (соглашение 075-15-2023-380 от 20 февраля 2023 г.).

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

СПИСОК ЛИТЕРАТУРЫ

1. Kashiwagi Y., Higashi T., Obashi K., et al. Computational geometry analysis of dendritic spines by structured illumination microscopy // Nat Commun. 2019. Vol. 10, N 1. P. 1285–1285. doi: 10.1038/s41467-019-09337-0
2. Bokota G., Magnowska M., Kuśmierczyk T., et al. Computational approach to dendritic spine taxonomy and shape transition analysis // Front Comput Neurosci. 2016. Vol. 10. P. 140. doi: 10.3389/fncom.2016.00140
3. Pchitskaya E., Bezprozvanny I. Dendritic spines shape analysis-classification or clusterization? Perspective // Front Synaptic Neurosci. 2020. Vol. 12. N 31. doi: 10.3389/fnsyn.2020.00031

КОНТАКТНАЯ ИНФОРМАЦИЯ

* Е.И. Пчицкая; адрес: Российская Федерация, 195251, Санкт-Петербург, ул. Политехническая, д. 29;
e-mail: pchitskaya.ei@edu.spbstu.ru

DOI: <https://doi.org/10.17816/gc623520>

Classification of dendritic spiking neurons' shape through clustering and machine learning techniques

P.I. Vasiliev^{1,2}, D.S. Smirnova², V.S. Chukanov^{1,2}, I.B. Bezprozvanny^{1,3}, E.I. Pchitskaya^{1*}¹ Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russian Federation;² Department of Applied Mathematics, Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russian Federation;³ UT Southwestern Medical Center at Dallas, Dallas, USA

ABSTRACT

A synapse is a specialized region between two adjacent neurons that allows information to be transmitted from one cell to another. The formation of these connections and transfer of signals via electrical impulses is a critical characteristic of neuronal cells. The synapse is formed by the axonal bouton on the axon side and the dendritic spine, a specialized outgrowth of the dendritic membrane, on the dendrite side. Dendritic spines exhibit diverse shapes and sizes, differing significantly across brain regions, cell categories, and animal species. Changes in dendritic spine morphology occur in neurodevelopmental and neurodegenerative ailments while also responding to external stimuli. Although deemed to facilitate synaptic plasticity, further investigation is essential for establishing the correlation between spine construction and its function. To address the issue in modern neurobiology of characterizing synapse morphology on 3D neuron images, the development of effective analytical methods is necessary. Our team has produced an open-source software solution for precise dendritic spine segmentation using 3D dendrite images. This software calculates the 10 most widely used 3D-adapted morphological features [1, 2] and enables classification and clustering of dendritic spine data sets to determine their shape. In addition to numerical features for describing the shape of a dendritic spine, researchers proposed using a histogram of chord lengths, known as the chord length distribution histogram (CLDH). This involves generating a set of random chords within the dendritic spine's volume, connecting its outer boundaries and forming a histogram. By setting $n=30\ 000$, the probabilistic fluctuations of the histogram become insignificant. The derived metrics were then used for clustering and classifying the dataset.

Classification based on predetermined morphological groups is a frequently used approach for analyzing the morphology of dendritic spines. This method involves categorizing spines into established groups such as thin, mushroom, and stubby. Experimenters generally perform classification in a semi-automated manner, leading to considerable error. We have created a spine categorization tool using a machine learning algorithm. The tool classifies spines based on a consensus reached by eight experts who manually labeled the training dataset. The accuracy of this method surpasses 77% when using classical morphological features and is comparable to expert labeling. The implementation of this approach reduces classification bias and complexity.

Recent studies, including those using live microscopy *in vitro* and *in vivo*, indicate that dendritic spine shapes exhibit a continuum rather than distinct categories [3]. Therefore, it is essential to establish a dependable methodology for evaluating and examining the morphology of dendritic spines. We have created a clustering tool that establishes both the number of groups and their content based on data, rather than the experimenter's discretion. This tool leverages the k-means and DBSCAN algorithms for its representation. Three clustering methods are presented to determine the number of clusters: the silhouette method, the elbow method, and a novel method, developed by the authors, based on the max class divergence criteria. The authors assume that cluster quality improves as clusters differ significantly in the number of mushroom/thin/stubby spine classes, as marked by experts. The benefit of this approach is that it considers the particular data type for clustering. Without this knowledge, assessing the quality of clustering is challenging. The use of the CLDH metric for clustering yielded a consistent and stable number of clusters ($n=5$) across all three methods described. These clusters contain dendritic spines that share similar shapes and have been validated by experts. In contrast, using classical metrics resulted in a variable cluster count ranging from $n=4$ to $n=14$. These findings suggest that the CLDH metric, with its complexity, provides enough information about synapse shape to enable precise clustering.

Keywords: synapse; dendritic spines; mushroom; thin; stubby; software; segmentation; classification; clustering; chord length distribution histogram.

To cite this article:

Vasiliev PI, Smirnova DS, Chukanov VS, Bezprozvanny IB, Pchitskaya EI. Classification of dendritic spiking neurons' shape through clustering and machine learning techniques. *Genes & cells*. 2023;18(4):890–893. DOI: <https://doi.org/10.17816/gc623520>

Received: 15.05.2023

Accepted: 26.11.2023

Published online: 20.01.2024

ADDITIONAL INFORMATION

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.

Funding sources. The research was funded by the Ministry of Science and Higher Education of the Russian Federation under the strategic academic leadership program "Priority 2030" (Agreement No. 075-15-2023-380 dated 20.02.2023).

Competing interests. The authors declare that they have no competing interests.

REFERENCES

1. Kashiwagi Y, Higashi T, Obashi K, et al. Computational geometry analysis of dendritic spines by structured illumination microscopy. *Nat Commun.* 2019;10(1):1285. doi: 10.1038/s41467-019-09337-0
2. Bokota G, Magnowska M, Kuśmierczyk T, et al. Computational approach to dendritic spine taxonomy and shape transition analysis. *Front Comput Neurosci.* 2016;10:140. doi: 10.3389/fncom.2016.00140
3. Pchitskaya E, Bezprozvanny I. Dendritic spines shape analysis-classification or clusterization? Perspective. *Front Synaptic Neurosci.* 2020;12:31. doi: 10.3389/fnsyn.2020.00031

AUTHORS' CONTACT INFO

* E.I. Pchitskaya; address: 29 Polytechnicheskaya street, 195251 Saint Petersburg, Russian Federation; e-mail: pchitskaya.ei@edu.spbstu.ru