

regions that integrate sensory, cognitive and motor functions. Chemogenetic modulation of the activity of various groups and types of neurons changes the connectivity of the brain. In this regard, there is an assumption about the possibility of using chemogenetics to correct behavior and restructure the pathological mechanisms of nervous and mental diseases by influencing connectivity, by increasing or decreasing the activity of one or more groups of interacting neurons. Evaluation of this possibility by chemogenetic parallel unidirectional or multidirectional modulation of the activity of remote neuronal populations of the brain involved in the regulation of psycho-emotional behavior indicates the reality of “control” of behavior based on neural network mechanisms of its regulation.

行為調節的神經網絡方面

Dygalo N.N.

聯邦研究中心細胞學和遺傳學研究所，俄羅斯科學院西伯利亞分院，俄羅斯新西伯利亞

dygalo@bionet.nsc.ru

中樞神經系統活動和行為的基礎是大腦的大型局部神經網絡的功能。磁共振成像研究表明，在外部和/或內部刺激的影響下，大腦活動及其神經集合的功能連接矩陣發生了變化，這被認為會導致某些行為的誘發。連通性基於神經元的活動，是整合感覺、認知和運動功能的大腦區域連接的表現。各種神經元組和類型的活動的化學遺傳學調節改變了大腦的連通性。在這方面，有一種假設認為，通過影響連通性、增加或減少一組或多組相互作用的神經元的活動，可以使用化學遺傳學來糾正行為並重構神經和精神疾病的病理機制。通過化學遺傳學平行單向或多向調製參與心理情緒行為調節的大腦遠程神經元群的活動來評估這種可能性，表明基於其調節的神經網絡機制的行為“控制”的現實。

DOI 10.18699/sblai2023-19

***Применение методов искусственного интеллекта и анализа генных сетей
для интерпретации омиксных данных***

Иванисенко В.А.*, Вензель А.С., Деменков П.С., Иванисенко Т.В., Антропова Е.А.,
Гайслер Е.В., Рогачев А.Д., Покровский А.Г., Колчанов Н.А.

Институт цитологии и генетики СО РАН, Новосибирск, Россия

* salix@bionet.nsc.ru

Работа посвящена когнитивной программно-информационной платформе ANDSystem [1, 2], предназначенной для автоматического извлечения знаний из неструктурированных текстов научных публикаций. ANDSystem позволяет учитывать 12 типов молекулярно-биологических объектов, включая организмы, клетки/ткани, болезни, гены, метаболиты, лекарства и др. Система обеспечивает автоматическую экстракцию из текстов знаний о физических взаимодействиях, каталитических реакциях, регуляторных взаимосвязях, таких как регуляция экспрессии, регуляция активности, регуляция стабильности и т. д. (всего более 20 типов).

Основные модули ANDSystem включают в себя: модуль онтологического описания предметной области, содержащий словари объектов и семантико-лингвистические правила взаимоотношений между ними; модуль распознавания поименованных сущностей; модуль извлечения знаний о взаимодействиях; базу, содержащую экстрагированную информацию, в виде графа знаний; программу ANDVisio, обеспечивающую пользовательский графический интерфейс для доступа к базе знаний, а также реконструкции и анализа генных сетей.

Извлечение знаний осуществляется разработанной системой в четыре этапа. На первом этапе производится первичная экстракция взаимодействий на основе сформированного вручную набора семантико-лингвистических правил и шаблонов. На втором этапе

экстрагированная таким образом информация о взаимодействиях между объектами добавляется в разметку текста для формирования обучающих выборок, а также выполняется автоматическая реконструкция графа знаний. На третьем этапе с помощью сформированных выборок проводится тонкая настройка предобученных генеративных нейронных сетей, реализующих архитектуры GPT-3 и BERT. Также с использованием реконструированного графа знаний и алгоритма GraphSAGE выполняется обучение графовой нейронной сети. Полученные нейронные сети используются для уточнения типов объектов, распознанных по словарю, а также для дополнительного расширения числа взаимодействий, в дополнение к извлеченным с помощью семантико-лингвистическими шаблонами.

ANDSystem нашла успешное применение в решении широкого спектра задач из области биологии и биомедицины, включая исследования молекулярных механизмов заболеваний, интерпретацию омиксных данных и другие. В настоящее время метаболомный анализ является одной из широко применяемых технологий в различных областях биомедицинских исследований. С использованием системы ANDSystem была выполнена интерпретация данных метаболомного анализа образцов крови плазмы пациентов с COVID-19 [3]. В частности, с помощью реконструированных регуляторных генных сетей была выявлена роль неструктурных вирусных белков SARS-CoV-2 в метаболических дисфункциях у пациентов. Реконструированные генные сети играют важную роль в исследовании молекулярных механизмов взаимодействия вирус–хозяин при COVID-19 и служат основой для дальнейших экспериментальных и компьютерных исследований регуляции метаболических путей белками SARS-CoV-2.

Литература

- 1 Ivanisenko T.V., Saik O.V., Demenkov P.S., Ivanisenko N.V., Savostianov A.N., Ivanisenko V.A. ANDDigest: a new web-based module of ANDSystem for the search of knowledge in the scientific literature, BMC Bioinformatics, 2020, vol. 21, P. 228.
2. Ivanisenko V.A., Demenkov P.S., Ivanisenko T.V., Mishchenko E.L., Saik O.V. A new version of the ANDSystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue- specific gene expression, BMC Bioinformatics, 2019, vol. 20.(Suppl 1), P. 34.
3. Ivanisenko V.A., Gaisler E.V., Basov N.V., Rogachev A.D., Cheresiz S.V., Ivanisenko T.V., ... Pokrovsky A.G. Plasma metabolomics and gene regulatory networks analysis reveal the role of nonstructural SARS-CoV-2 viral proteins in metabolic dysregulation in COVID-19 patients, Scientific Reports, 2022, vol. 12, P. 19977.

Application of artificial intelligence methods and gene networks analysis for the interpretation of omics data

Ivanisenko V.A.*, Venzel A.S., Demenkov P.S., Ivanisenko T.V., Antropova E.A.,
Gaisler E.V., Rogachev A.D., Pokrovsky A.G., Kolchanov N.A.
Institute of Cytology and Genetics of SB RAS, Novosibirsk, Russia

* salix@bionet.nsc.ru

The ANDSystem, a cognitive software and information system, has been developed for the automated extraction of knowledge from scientific literature [1, 2]. It encompasses 12 object types, such as organisms, cells/tissues, diseases, genes, metabolites, drugs, and more. The system extracts information on physical interactions, catalytic reactions, and regulatory relationships (over 20 types) from texts.

The core modules of the ANDSystem include: an ontology module with object dictionaries and semantic-linguistic rules; a named entity recognition module; a knowledge extraction module; a knowledge base featuring a knowledge graph; and the ANDVisio program, offering a user-friendly graphical interface for knowledge base access, as well as gene network reconstruction and analysis. Knowledge extraction consists of four stages:

1. Interaction extraction using semantic-linguistic templates.

2. Addition of interaction information between objects to text markup for neural network training, followed by knowledge graph reconstruction.
3. Fine-tuning of pre-trained GPT and BERT transformers with mapped texts, and graph neural network training using the knowledge graph.
4. Identification of new interactions using trained neural networks, supplementing those extracted with semantic-linguistic templates.

The ANDSystem has been employed to address various issues in biomedicine, including the reconstruction of molecular disease mechanisms and the interpretation of omics data. Metabolomic analysis is currently a prevalent omics technology within biomedical research. Utilizing the ANDSystem, we interpreted data from a metabolomic analysis of blood plasma samples from COVID-19 patients [3]. This gene regulatory network reconstruction highlighted the role of nonstructural SARS-CoV-2 viral proteins in metabolic dysregulation among patients. The reconstructed gene networks offer insights into the molecular mechanisms of virus–host interactions in COVID-19 pathology and lay the foundation for subsequent experimental and computational investigations of metabolic pathway regulation by SARS-CoV-2 proteins.

References

1. Ivanisenko T.V., Saik O.V., Demenkov P.S., Ivanisenko N.V., Savostianov A.N., Ivanisenko V.A. ANDDigest: a new web-based module of ANDSystem for the search of knowledge in the scientific literature, BMC Bioinformatics, 2020, vol. 21, P. 228.
2. Ivanisenko V.A., Demenkov P.S., Ivanisenko T.V., Mishchenko E.L., Saik O.V. A new version of the ANDSystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue- specific gene expression, BMC Bioinformatics, 2019, vol. 20.(Suppl 1), P. 34.
3. Ivanisenko V.A., Gaisler E.V., Basov N.V., Rogachev A.D., Cheresiz S.V., Ivanisenko T.V., ... Pokrovsky A.G. Plasma metabolomics and gene regulatory networks analysis reveal the role of nonstructural SARS-CoV-2 viral proteins in metabolic dysregulation in COVID-19 patients, Scientific Reports, 2022, vol. 12, P. 19977.

人工智能方法和基因網絡分析在組學數據解釋中的應用

Ivanisenko V.A.*, Venzel A.S., Demenkov P.S., Ivanisenko T.V., Antropova E.A.,
Gaisler E.V., Rogachev A.D., Pokrovsky A.G., Kolchanov N.A.

ICG SB RAS, 新西伯利亞, 俄羅斯

* salix@bionet.nsc.ru

ANDSystem 是一種認知軟件和信息系統，已開發用於從科學文獻中自動提取知識 [1, 2]。它包含 12

種對象類型，例如生物體、細胞/組織、疾病、基因、代謝物、藥物等。

該系統從文本中提取有關物理相互作用、催化反應和調節關係（超過 20 種）的信息。

ANDSystem 的核心模塊包括：帶有對象字典和語義語言規則的本體模塊；命名實體識別模塊；知識抽取模塊；以知識圖譜為特色的知識庫；ANDVisio 程序，為知識網絡庫訪問以及基因重建和分析提供了一個用戶友好的圖形界面。知識抽取包括四個階段：

1. 使用語義-語言模板進行交互提取。
2. 將對象之間的交互信息添加到文本標記中用於神經網絡訓練，然後進行知識圖譜重建。
3. 使用映射文本微調預訓練的 GPT 和 BERT 轉換器，並使用知識圖譜訓練圖神經網絡。
4. 使用經過訓練的神經網絡識別新的交互，補充用語義語言模板提取的交互。

ANDSystem

已被用於解決生物醫學中的各種問題，包括分子疾病機制的重建和組學數據的解釋。

代謝組學分析目前是生物醫學研究中普遍使用的組學技術。利用

ANDSystem，我們解釋了來自 COVID-19 患者血漿樣本的代謝組學分析數據 [3]。

這種基因調控網絡重建突出了非結構性 SARS-CoV-2 病毒蛋白在患者代謝失調中的作用。重建的基因網絡提供了對 COVID-19 病理學中病毒與宿主相互作用的分子機制的見解，並為隨後對 SARS-CoV-2 蛋白調控代謝途徑的實驗和計算研究奠定了基礎。

文學

1. Ivanisenko T.V., Saik O.V., Demenkov P.S., Ivanisenko N.V., Savostianov A.N., Ivanisenko V.A. ANDDigest: a new web-based module of ANDSystem for the search of knowledge in the scientific literature, BMC Bioinformatics, 2020, vol. 21, P. 228.

2. Ivanisenko V.A., Demenkov P.S., Ivanisenko T.V., Mishchenko E.L., Saik O.V. A new version of the ANDSystem tool for automatic extraction of knowledge from scientific publications with expanded functionality for reconstruction of associative gene networks by considering tissue-specific gene expression, BMC Bioinformatics, 2019, vol. 20.(Suppl 1), P. 34.

3. Ivanisenko V.A., Gaisler E.V., Basov N.V., Rogachev A.D., Cheresiz S.V., Ivanisenko T.V., ... Pokrovsky A.G. Plasma metabolomics and gene regulatory networks analysis reveal the role of nonstructural SARS-CoV-2 viral proteins in metabolic dysregulation in COVID-19 patients, Scientific Reports, 2022, vol. 12, P. 19977.

DOI 10.18699/sblai2023-20

Миграционное клеточное обеспечение нормального функционирования иммуно-структурного гомеостаза

Козлов В.А.

Институт фундаментальной и клинической иммунологии СО РАН, Новосибирск, Россия

Тимус, как и костный мозг, является центральным органом иммунной системы. Однако именно тимус является источником Т лимфоцитов, которые мигрируют из тимуса на периферию в виде двух значимых популяций лимфоцитов: наивные Т-клетки и регуляторные Т-клетки (Treg). Функциональная активность этих двух суперпопуляций Т-клеток обеспечивает, наравне с мигрантами из костного мозга, нормальное функционирование иммунной системы в организме. Нарушение нормального функционирования мигрантов из тимуса в первую очередь является фактором формирования в организме иммунопатологических заболеваний, включая аутоиммунные и аллергические заболевания, атеросклероз, онкологию. По-видимому, одним из ведущих условий нормального функционирования иммунной системы являются миграционные процессы клеток между костным мозгом, тимусом и периферическими лимфоидными органами. При этом наряду с популяциями клеток со своими специфическими, органными характеристиками, располагающимися постоянно в различных органах, имеются популяции клеток, недавно мигрировавших из одних органов в другие, которые уже «потеряли» отдельные характеристики органов, из которых они мигрировали, но еще не приобрели характеристики органов, в которые они мигрировали. Любое нарушение процессов миграции клеток между указанными органами иммунной системы несет существенные помехи нормальному функционированию иммунной системы с формированием иммунопатологического состояния. По-видимому, в основном эти клетки, популяционные мигранты, циркулируют по кровеносным и лимфатическим сосудам. Следует учитывать, что именно эти клетки являются источником информации состояния в организме, прежде всего гомеостатических иммунной и кровеносной систем.