

Graphical Modelling for Holistic Understanding of IA Disease



Matteo Delucchi
PhD Student, delt@zhaw.ch

Research project
Data-driven decision support for intracranial aneurysms and hospital catering using Bayesian networks

Lead:
Dr. Georg Spinner, Centre of Computational Health

Duration:
March 2020 – December 2021

Partner:
Hôpitaux universitaires de Genève / Département des neurosciences cliniques

Computational Health

Globally, about one-third of all stroke incidents are caused by intracerebral haemorrhages. This Health Research Hub project aims to use data and expert knowledge of intracranial aneurysms (IA) to develop an interpretable disease model. A better understanding of how known risk factors influence the progression of IAs can assist clinical decision making and, in the future, possibly lead to healthier lives.

Stroke and Cerebral Aneurysms

According to the World Stroke Organisation, one in four people over the age of 25 will have a stroke in their lifetime. Symptoms indicating a stroke are sudden facial drooping, numb arms and slurred speech, arising from brain cell death. This results from either an interruption of a cerebral blood vessel (ischemic stroke) or a weakened vessel that ruptures and bleeds into the surrounding brain, known as haemorrhagic stroke, which causes a thunderclap headache. Intracranial aneurysms are bulges filled with blood that form at weak points in cerebral arteries. Although they usually do not cause symptoms, a potential vessel wall rupture would lead to a subarachnoid haemorrhage, a type of haemorrhagic stroke.

Clinical Acceptance of Complex System Models

Over 143 million years of healthy life are lost each year due to subarachnoid haemorrhage-related death and disability. Fortunately, advanced brain imaging means that IAs are increasingly being detected before rupture. Nevertheless, clinicians must still balance treatment risks with the risk of an IA rupture for each patient individually.

While we are currently seeing exponential growth in the collection of health data in hospitals as a result of digitalisation, clinicians' decisions are still often based on subjective experience rather than solid evidence. If, however, this wealth of available data is coupled with advanced statistical modelling, it has the potential to increase understanding of the progression of IAs and ultimately improve clinical decision making. However, such an artificial decision-making system will only receive sufficient clinical acceptance if it is easy to interpret, a factor that often represents a major hurdle for high performing deep learning classifiers.

Graphs Increase Interpretability

Currently, clinicians rely on risk prediction scores from classical statistical regressions, which hardly considers interdependent complex relationships between risk factors. However, this project aims to decipher the interdependencies between IA rupture risk factors using probabilistic graphical models. For example, Bayesian networks (BN) can learn motives and patterns from data combined with prior knowledge. In this way, we can represent variables and their interdependencies as a specific type of graph that is easily accessible – for both computation and interpretation. Furthermore, Bayesian networks allow causal and probabilistic modelling of numerous factors from a complex, hierarchical context.

Can we Improve IA Treatment Outcomes?

Even though BNs are a promising tool for decision support purposes, there is still a long way to go ahead to increase holistic understanding of IA disease. For example, while a healthy lifestyle reduces the risk of a sudden IA rupture, we do not yet fully understand the interplay between risk factors that make people susceptible to IA and that may ultimately lead to a rupture. Meanwhile, we must deal with classic health data problems – unknown, incomplete, imbalanced, heterogeneous, noisy, dirty, erroneous, inaccurate, and missing data sets in arbitrarily high-dimensional spaces – and simultaneously maintain the model's causability¹. ■

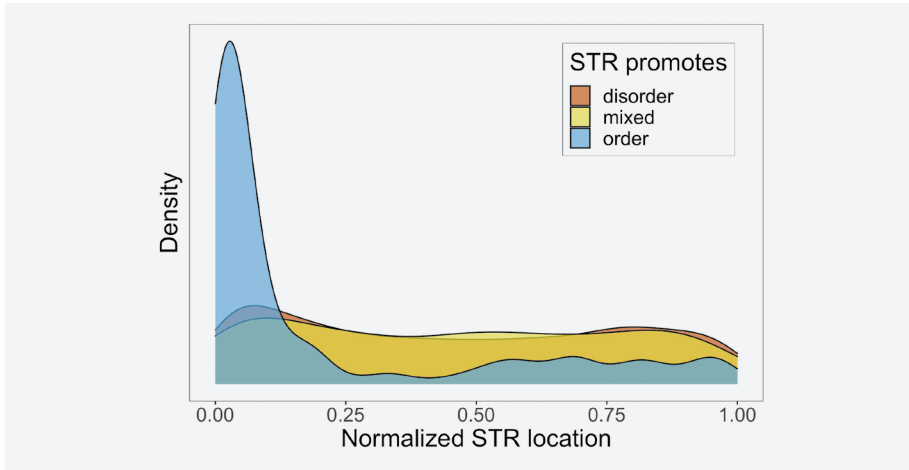
¹ (Holzinger et al., 2019)

Repeated (dis)order

Max Verbiest, PhD student, verb@zhaw.ch

Protein short tandem repeats (STRs) are stretches of consecutive repetitions of short amino acid motifs. In protein-coding sequences, STRs primarily encode disorder-promoting amino acids and are often located in intrinsically disordered regions (IDRs). Max Verbiest, Matteo Delucchi and Maria Anisimova of the Institute of Computational Life Sciences performed a study of STR characteristics across all known human proteins. They were able to confirm that STR-containing proteins are involved in gene regulation and nucleotide binding. Furthermore,

they described for the first time a subset of structured STRs located specifically in N-terminal signal peptides, which target proteins for transport to the cell membrane. This group of STR-containing proteins was also found to be involved in biological functions associated to the cell exterior such as receptor activity and ligand binding. This work was funded by the SNF Sinergia grant CRSII5_193832 and the European Union's Horizon 2020 research and innovation program grant No. 823886. ■



Disordered and mixed protein STRs are distributed across the full length of proteins. A small subset of STRs is ordered and located in N-terminal signal peptides that cause proteins to be moved to the cell membrane and exterior (blue peak on the left). Adapted from Verbiest, Delucchi, Bilgin-Sonay and Anisimova, *Frontiers in Bioinformatics* (2021).

Neue Projekte

Shapescience – AI for morphologically based fruit variety recognition

Dauer: 14.10.2021 – 30.12.2025
Projektpartner: Eidgenössische Technische Hochschule Zürich ETH, Realisation Schmid

Transformernetze für Drohnensignale

Dauer: 28.10.2021 – 29.06.2022
Projektpartner: Bundesamt für Rüstung armasuisse, Kompetenzzentrum Wissenschaft und Technologie

Weitere Projekte

zhaw.ch/icls/projekte

Weiterbildung

11.05.2022
Collecting, Generating & Structuring Data for Simulations

10.06.2022
Digital Ethics

31.08.2022
Validation, Experimentation & Communication Results from Simulations

28.09.2022
Advanced Modelling of Complex Processing

07.10.2022
Digital Ethics

26.10.2022
Natural Language Processing Fundamentals

16.11.2022
The Process Simulation Project

Diverse Kurse und Angebote
zhaw.ch/icls/weiterbildung

ShapeScience – Describing and Identifying Old Fruit Varieties

Anna Wróbel, PhD student, wrob@zhaw.ch

The identification of plant species and varieties is crucial for the resilient cultivation and the diversity of crops. Historically, the identification of fruits was carried out by pomologists or, in recent years, also by DNA analysis due to advances in molecular biology. However, both approaches require time, incur costs and are not free from errors. Machine learning methods offer an alternative for the identification of crops varieties. The ShapeScience project is being carried out by a consortium of ZHAW and the SME Realisation Schmid in collaboration with the Federal Office for Agriculture (BLW) and Agroscope/ETH. The aims of the project are to investigate and find biomarkers for the identification of apple varieties based on their fruits to develop a method, and eventually a cognitive system applicable in the field, for automatic fruit variety identification. Additionally, the PhD thesis based on the project, supported by the Institute of Neuroinformatics (UZH/ETH) embraces the investigation of more fundamental questions pertaining to the practical application of neuromorphic technologies for computer vision and autonomous intelligence. ■



Machine vision system for taking 2D images of apples