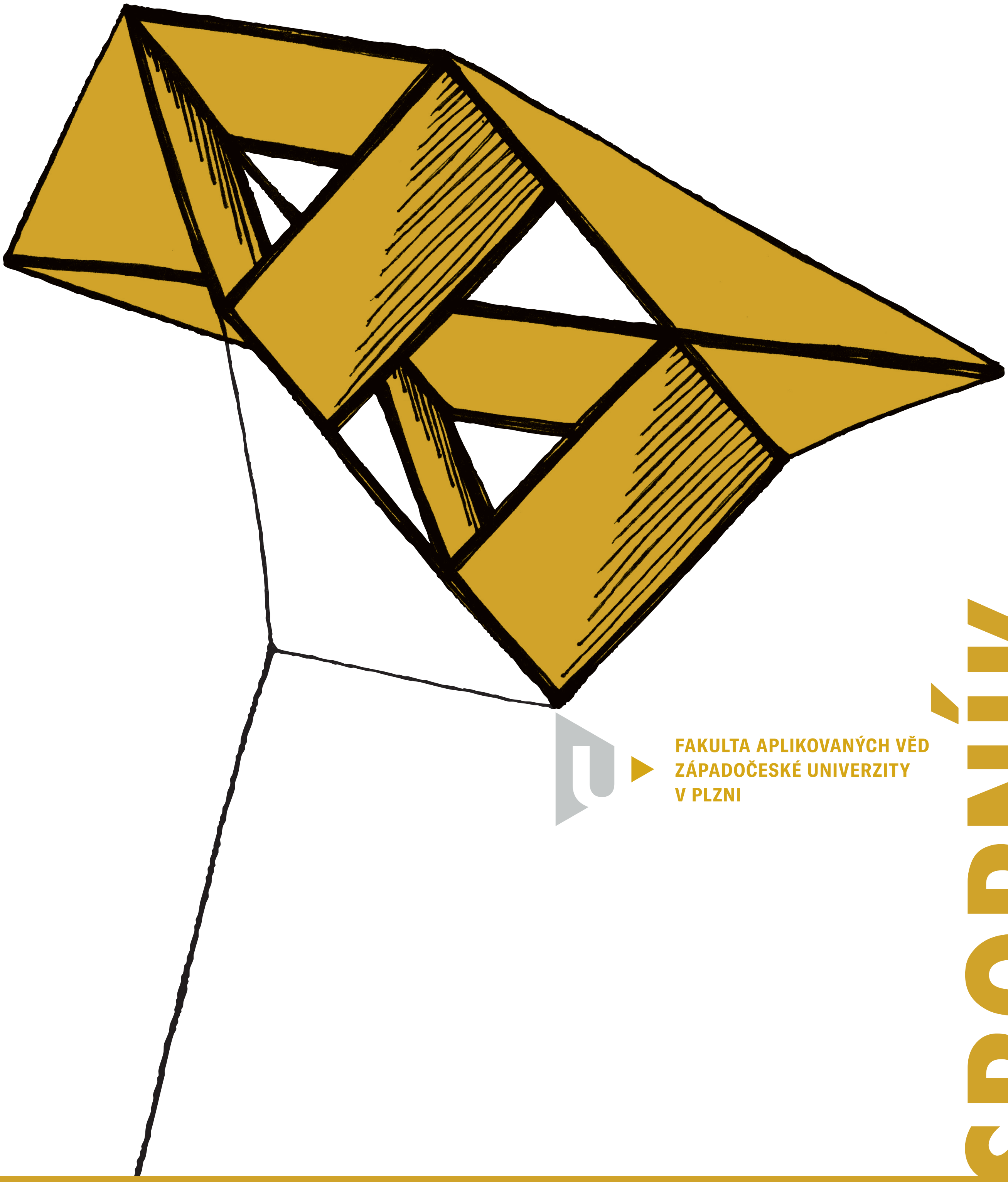
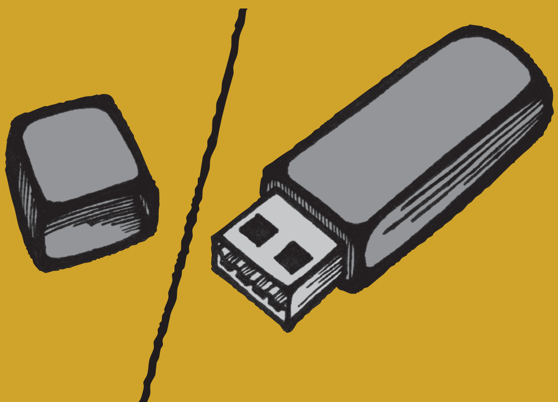


FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

SBORNÍK

STUDENTSKÁ VĚDECKÁ KONFERENCE



MAGISTERSKÉ A DOKTORSKÉ STUDIJNÍ PROGRAMY

4. 6. 2026

FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

STUDENTSKÁ
VĚDECKÁ KONFERENCE 2026

Studentská vědecká konference FAV

Magisterské a doktorské studijní programy

Sborník rozšířených abstraktů

FACULTY OF APPLIED SCIENCES
UNIVERSITY OF WEST BOHEMIA
IN PILSEN

Organizátor konference

Fakulta aplikovaných věd
Západočeská univerzita v Plzni

FACULTY OF APPLIED SCIENCES
UNIVERSITY OF WEST BOHEMIA
IN PILSEN

Studentská vědecká konference je pořádána s podporou prostředků na specifický vysokoškolský výzkum jako projekt SVK1-2026-009.

Organizační výbor

Pavel Novák
Petra Bláhová
Zuzana Majdišová
Tomáš Potužák
Jan Rendl
Jiří Rezek
Vladimír Švígler
Jaroslav Toninger
Lucie Zajícová

Partneři konference

ABB s.r.o.
ČEPS, a.s.
GK Software Czech Republic s.r.o.
Unicorn a.s.
ZF Engineering Plzeň s.r.o.

Název: SVK FAV 2026 – magisterské a doktorské studijní programy
Editor: Jan Rendl
Ilustrace: Jana Trávníčková
Vydavatel: Západočeská univerzita v Plzni, Univerzitní 2732/8, 301 00 Plzeň
Datum vydání: Červen 2026
ISBN 978-80-261-1370-6

Tato publikace neprošla redakční ani jazykovou úpravou.

Partneři konference

ABB

čeps

GK

a Fujitsu company

UNICORN

ZF

Obsah

FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

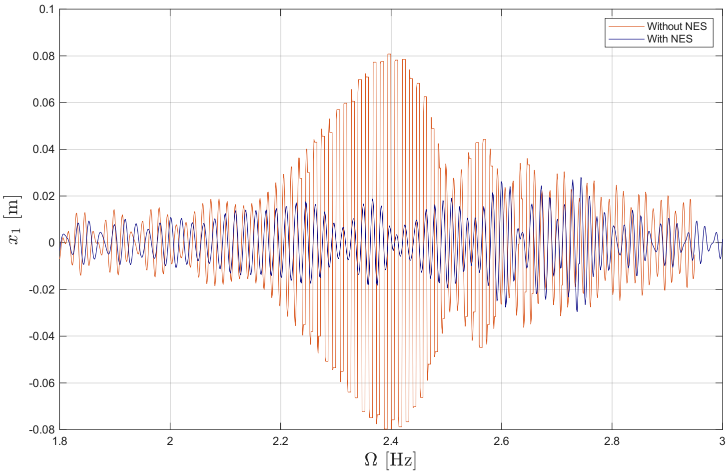
STUDENTSKÁ
VĚDECKÁ KONFERENCE 2026

Matematické modelování a experimentální demonstrace nelineárního pohlcovače energie (NES)

František Bareš[1]

Vibrace jsou fenoménem řešeným v mnoha oblastech strojírenství a stavitelství. Zejména ty s vysokou amplitudou mohou mít zásadní dopad na životnost mechanických součástí nebo stabilitu konstrukcí. Během let se objevilo mnoho řešení vibračních tlumičů a hltičů. Jak popsal Zhang et al. (2017), lze je rozdělit na pasivní, aktivní a semi-aktivní. Tato práce se zabývá těmi pasivními. Jednoduché pasivní hltiče dokáží redukovat vibrace pouze na velmi malé frekvenční oblasti. Nejspíše nejefektivnější způsob, jak dosáhnout více robustního tlumení, je Nelineární pohlcovač energie (anglicky Nonlinear energy sink-NES). Tento koncept je založen na principech nelineární dynamiky, konkrétně na jevu nazvaném Cílený přenos energie (Targeted energy transfer-TET), který popsal Gendelman (2001).

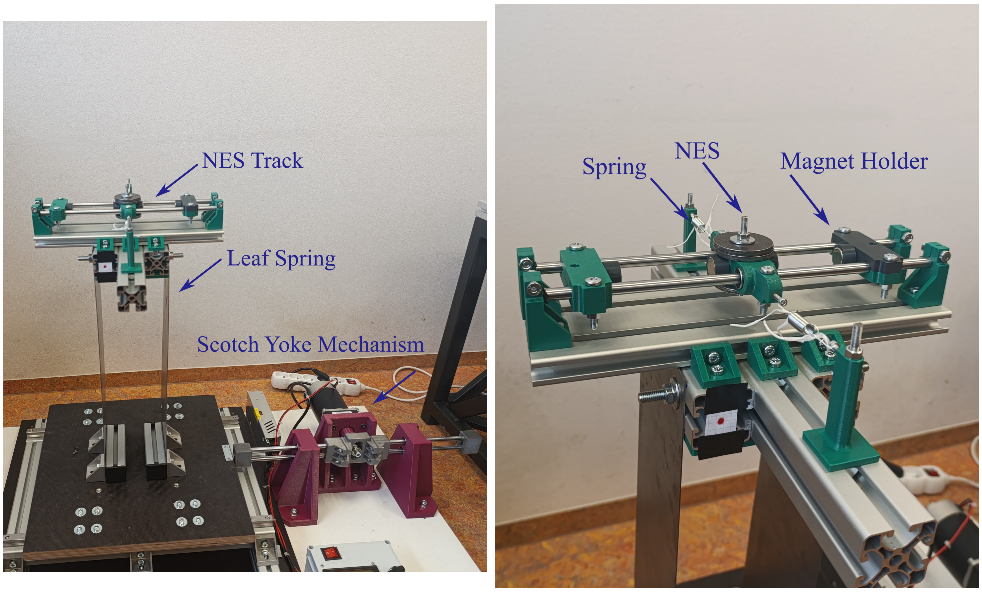
Hlavním cílem této práce je aplikovat NES pro snížení rezonančního vrcholu demonstrátoru s jedním stupněm volnosti buzeného nízkými frekvencemi. Zaměřením této práce je porovnat tradiční NES používající pružiny a NES využívající odpudivé magnetické síly. Nejdříve byl určen matematický model demonstrátoru, následně byla použita Metoda harmonické balance a upravená Metoda délky oblouku. Dále byla provedena kvalitativní analýza parametrů pomocí vyjádření efektivity NESu a následné optimalizace. Nakonec byl zkonstruován fyzický demonstrátor pro experimentální ověření získaných teoretických poznatků. Měření bylo provedeno pomocí vyvinuté metody detekce bodu pomocí videa a následné experimenty ukázaly že pružinový NES dokáže snížit rezonanční vrchol z 0,08 m na 0,02 m. To činí snížení vibrací o 75%. V případě magnetického NESu byly zjištěny nežádoucí vysoké amplitudy zapříčeněny nelinearitou systému. Kombinace obou přístupů však dala uspokojivé výsledky.



Obrázek 1: Naměřený rozběh demonstrátoru s NES a bez něj.

---

[1] student navazujícího studijního programu Aplikovaná mechanika, obor Dynamika konstrukcí a mechatronika, e-mail: bares@students.zcu.cz



Obrázek 2: Konstrukce experimentálního demonstrátoru.

Literatura

Zhang, B.-L., Han, Q.-L., a Zhang, X.-M. (2017) Recent advances in vibration control of offshore platforms. *Nonlinear Dynamics*. Springer, vol. 89, pp. 755–771.

Gendelman, O.V. (2001) Transition of energy to a nonlinear localized mode in a highly asymmetric system of two oscillators. *Nonlinear dynamics*. Springer, vol. 25, no. 1, pp. 237–253.

Global gravitational field modelling for spheroidal planetary bodies: non-singular expressions

Jiří Belinger, Veronika Dohnalová, Martin Pitoňák, Pavel Novák, Michal Šprlák

NTIS – New Technologies for the Information Society, Faculty of Applied Sciences,
University of West Bohemia in Pilsen, Technická 8, 306 14 Plzeň, Czech Republic,
belinger@fav.zcu.cz

1 Introduction

Determination of gravitational fields generated by various planetary bodies (including Earth, Earth's Moon or neighbouring planets) represents a crucial task in modern geodesy. There are various approaches to gravity field calculation. One of the most common methods utilises spherical harmonics and is in principle based on 2D Fourier transform as we combine and scale spherical harmonics to synthesise gravity field quantities (Fig. 1).

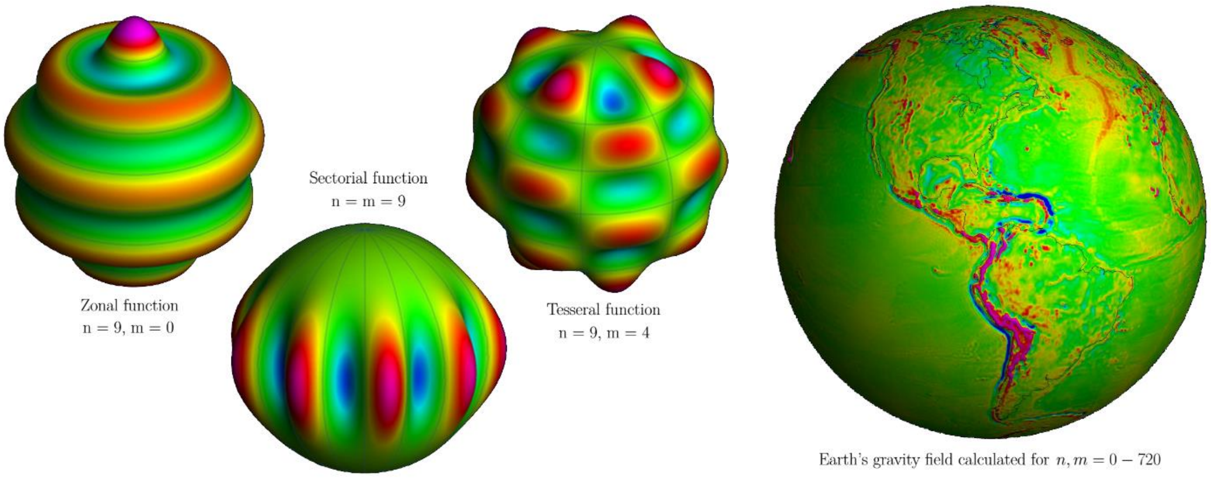


Figure 1: Synthesis of Earth's gravity field from a large set of spherical harmonic functions (spatial resolution rises with degree *n* and order *m*) – Ince et al. (2019).

However, shapes of planetary bodies are often closer to spheroids. Therefore, transition to spheroidal approximation described in Šprlák and Han (2021) is essential as it significantly extends region near the surface of flattened planetary bodies allowing safe computation without any divergence issues described by Hu and Jekeli (2015).

This contribution describes aspects of the gravitational field modelling of planetary bodies using spheroidal harmonic series. We focus on both parts of spheroidal harmonics method, namely spheroidal harmonic analysis and spheroidal harmonic synthesis. We also consider application of outer only and combination of outer and inner expansions allowing for calculation of gravitational field quantities inside the masses of selected body. Results are presented for asteroid Bennu studied by NASA (2026) and include calculation of spheroidal harmonic coefficients (analysis) and 20 functionals of gravitational field (synthesis) including gravitational potential and its derivatives up to the third order.

Acknowledgement

Authors were supported by the project No. 26-21339S of the Czech Science Foundation.

References

Hu X., Jekeli C. (2015) A numerical comparison of spherical, spheroidal and ellipsoidal harmonic gravitational field models for small non-spherical bodies: examples for the Martian moons. *Journal of Geodesy* 89: 159–177. https://doi.org/10.1007/s00190-014-0769-x

Ince E. S., et al. (2016) ICGEM – 15 years of successful collection and distribution of global gravitational models, associated services, and future plans. *Earth System Science Data* 11: 647-674. https://doi.org/10.5194/essd-11-647-2019

National Aeronautics and Space Administration. NASA Images [online]. NASA, [cit. 2026-05-28]. Dostupné z: https://www.nasa.gov/images/

Šprlák M. ,Han S. C. (2021) On the use of spherical harmonic series inside the minimum brillouin sphere: Theoretical review and evaluation by GRAIL and LOLA satellite data. Earth Sci. Rev. 222:103739. https:// doi.org/10.1016/j.earscirev.2021.103739

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

An Analytical Design Method for Fixed-Structure Controllers under Generalized Mixed-Sensitivity Specifications

Michal Brabec[1]

Introduction and Motivation

The design of robust controllers is a key part of a wide range of industrial applications. Although classical $\mathcal{H}_\infty$ synthesis provides a systematic frequency-domain framework for the formulation of complex engineering requirements, the resulting controller is often difficult to use directly in practice. In general, the controller structure cannot be prescribed a priori, and the resulting controller may have a higher order than is desirable for implementation in industrial control systems, where low-order fixed-structure controllers, such as PI, PID, lead–lag, or PR controllers, are still preferred.

Direct design of a fixed-structure controller based on $\mathcal{H}_\infty$ specifications is therefore of considerable practical importance. In general, however, this problem is a non-convex, non-smooth NP-hard optimization problem. For controllers with only two tunable parameters, a parameter-space approach can be used instead. The solution of the design problem can then be represented as a region in the controller parameter plane, whose points correspond to stabilizing controllers that satisfy the prescribed design constraints.

This contribution builds on the analytical design method introduced in Brabec and Schlegel (2023) and extended in Brabec and Schlegel (2025) to systems with a single dead time. The main extension presented here is the inclusion of generalised mixed-sensitivity constraints. This makes it possible to design fixed-structure controllers that satisfy several frequency-domain requirements simultaneously, while preserving the parameter-space interpretation of the feasible set.

Generalised Mixed-Sensitivity Constraints

Classical mixed-sensitivity design is a frequency-domain optimization problem that enables several closed-loop transfer functions to be shaped simultaneously. It is commonly used to express the trade-off between control performance, robustness, and the magnitude of the control effort. In this work, we consider a generalised SISO formulation in which selected weighted sensitivity functions can be constrained simultaneously.

For weighting functions $W_S(s)$, $W_T(s)$, $W_C(s)$, $W_P(s)$, and an upper bound γ, the considered constraints are expressed as sums of squared magnitudes of weighted sensitivity functions. Let $\Omega = \{\omega_1, \dots, \omega_N\}$ denote the selected frequency grid. The S/T mixed-sensitivity trade-off can then be written in the form

$$\max_{\omega_i \in \Omega} \left(|W_S(j\omega_i) S(j\omega_i)|^2 + |W_T(j\omega_i) T(j\omega_i)|^2 \right) \leq \gamma^2.$$

---

[1] student of the doctoral degree program Cybernetics, field of study Robust controllers with constrained structure, e-mail: brabecm@fav.zcu.cz

Other combinations, such as S/S_C, S/S_P, or extended combinations of three weighted sensitivity functions, can be used in the same way. By selecting particular sensitivity functions and weights, one can express requirements on reference tracking, disturbance rejection, control effort limitation, measurement-noise attenuation, and robustness against unmodelled dynamics.

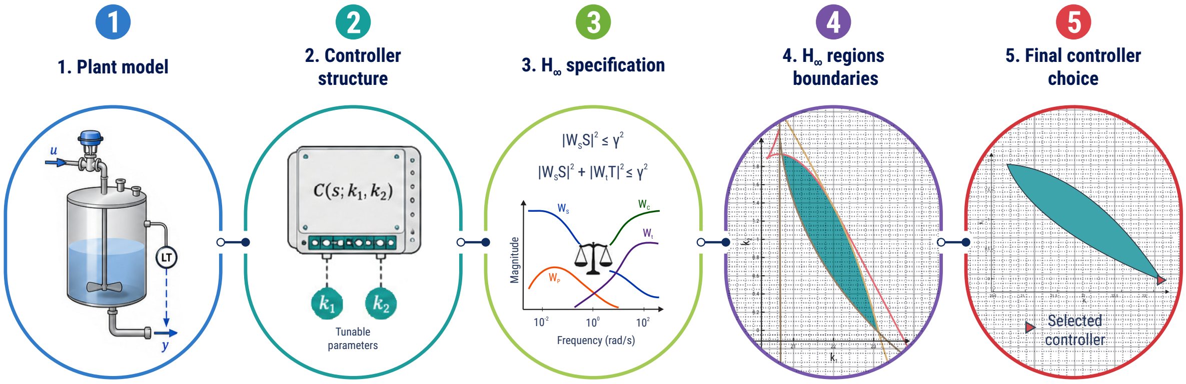


Figure 1: Illustration of the parameter-space design procedure for fixed-structure controllers. The selected $\mathcal{H}_\infty$ specifications are transformed into the boundaries of $\mathcal{H}_\infty$ regions in the controller-parameter plane, from which the final controller is selected.

Analytical Design Method

The design procedure differs from conventional optimization-based approaches. Instead of searching numerically for a single optimal controller, it transforms the specified $\mathcal{H}_\infty$ constraints into boundaries of admissible regions in the controller parameter plane, referred to as $\mathcal{H}_\infty$ regions. These constraints may refer either to individual weighted sensitivity functions or to mixed-sensitivity specifications.

The design procedure is based on the analytical calculation of the boundaries of the corresponding $\mathcal{H}_\infty$ regions. Through the intersection of $\mathcal{H}_\infty$ regions, it is possible to implement both single-model and multiple-model designs and, at the same time, demand fulfilment of several $\mathcal{H}_\infty$ specifications. The result is not a single controller, but the complete set of parameter pairs that guarantee closed-loop stability and fulfilment of the specified constraints. This admissible set can then be used to select the final controller according to additional engineering criteria, such as control effort, damping of resonant modes, bandwidth, robustness, or implementation limitations.

Acknowledgement

This research work was supported by the project SGS-2025-020.

References

Brabec, M. and Schlegel, M., 2023. Analytical Design of a Wide Class of Controllers with Two Tunable Parameters Based on $\mathcal{H}_\infty$ Specifications, 24th International Conference on Process Control (PC), Strbske Pleso, Slovakia, pp. 221-226.

Brabec, M. and Schlegel, M., 2025. Analytical Design of Controllers with Two Tunable Parameters Based on $\mathcal{H}_\infty$ Specifications for Dead-Time Systems, 25th International Conference on Process Control (PC), Strbske Pleso, Slovakia, pp. 1-6.

Autonomous Deep Space Navigation using X-ray Pulsar Timing

Ondřej Dohnal[1]

1 Introduction

Traditional deep space navigation is highly dependent on the Deep Space Network provided by NASA. However, this system is limited by the duration of the communication signal transmission (13 hours at the edge of the solar system), the ground station tracking accuracy of the spacecraft (SC) (lower, if the distance is greater), and the increasing workload of these stations. X-ray Pulsar Navigation (XNAV) offers a promising autonomous alternative. Using the highly stable rotation of millisecond pulsars as ”celestial GPS satellites,” the SC can determine its position, velocity, and time independently on Earth-based support; see [1]. This research focuses on the design and evaluation of an autonomous navigation system based on stable millisecond pulsars.

2 Methodology & Mathematical framework

The basic of XNAV is to transform observed pulse Time of Arrivals (TOAs) into a precise spatial position. Because pulsars are located at interstellar distances (thousands of light years away), the incoming X-ray signal is modeled as a planar wavefront along a unit line-of-sight vectror $\boldsymbol{n}$. The spatial offset between the SC (at barycentric position $\boldsymbol{r}^{ssb}_{sc}$) and the Solar System Barycentr (SSB), see Fig. 1, based on their signal TOA geometric (Romer) delay, can be written as

$$\Delta p = (t_{ssb} - t^{sc}_{ssb}) \cdot c = \boldsymbol{n}^T \cdot \boldsymbol{r}^{ssb}_{sc}. \qquad (1)$$

Here, t_{ssb} and $t^{sc}_{ssb} = t_{sc} - b$ are true barycentric TOA of SSB and SC, c is the speed of light constant, t_{sc} is the TOA observed by the SC clock, and b is the SC's clock bias.

To achieve a higher level of accuracy, the measurement model can account for other geometric or relativistic (such as Shapiro dellay) effects, which are incorporated into the calculation of Δp in (1). In this resaearch the measurement model is defined as pseudorange described by this equation

$$\tilde{\rho}^{ssb}_{i,sc} = (t_{ssb} - t_{sc}) \cdot c + \epsilon = \underbrace{\boldsymbol{n}_i^T \cdot \boldsymbol{r}^{ssb}_{sc}}_{Romer} + \underbrace{2\sum_k \frac{\mu_k}{c^2} \ln\left(\frac{\boldsymbol{n}_i^T \boldsymbol{r}^k_{sc} + ||\boldsymbol{r}^k_{sc}||}{\boldsymbol{n}_i^T \boldsymbol{r}^k_{ssb} + ||\boldsymbol{r}^k_{ssb}||} + 1\right)}_{Shapiro} - c \cdot b + \epsilon, \qquad (2)$$

where $\boldsymbol{n}_i$ is the unit vector from SSB to the i-th pulsar, $\boldsymbol{r}^k_{sc}$ and $\boldsymbol{r}^k_{ssb}$ are the position vectors from k-the included celestial body to SC and SSB at time t_{ssb}, μ_k is the standard gravitational parameter of this body, and ϵ is measurement noise, see [2].

Given the non-linear dynamics and measurements, a recursive estimation approach for estimation of the SC position is mandatory and is currently being developed in MATLAB®:

---

[1] student of the doctoral degree program Applied Sciences, field of study Cybernetics, e-mail: doondra@fav.zcu.cz. This research was funded by the project SGS-2025-020.

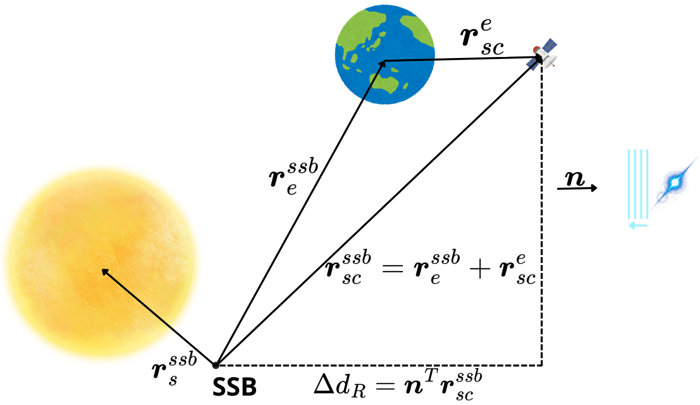


Figure 1: Geometric principle of XNAV coordinate systems.

- **Extended Kalman Filter (EKF)**: Being implemented as the baseline algorithm using first-order Taylor linearization.
- **Unscented/Cubature Kalman Filter (U/CKF)**: Selected to subsequently integrate and improve stability in high non-linear orbital phases by propagating sigma/cubature points instead of linerizing.

This simulation framework utilizes the JPL DE405 planetary ephemerides for celestial body positionong. The pulsar database is driven from the ATNF catalog, focusing on highly stable sources; see [3]. The specific subset selected for the initial implementation of pulsars to be used for SC navigation is summarized in Tab. 1. Studies to date show that position estimates based on the XNAV system are accurate to within a few kilometers; see [1, 2].

Name	Ra[h:m:s]	Dec[°:’:”]	Dist[kpc]
B0531+21	05:34:31.93357	+22:00:52.1927	2
B1937+21	19:39:38.5611889	+21:34:59.119343	3.5
J0437-4715	04:37:16.0494183	-47:15:10.022224	0.157
J1824-2452	18:24:32.009	-24:52:11.10	5.5
J0218+4232	02:18:06.358569	+42:32:17.37515	3.150

Table 1: Parameters of selected navigation pulsars.

3 Conclusion

This work establishes the foundational architecture and relativistic mathematical model for an XNAV simulation framework. Current efforts focus on finalizing the EKF software cod and verifying baseline convergence. Future work will integrate the U/CKF algorithm to perform a rigorous comparative analysis of estimation accuracy and computational robustness.

Resources

[1] Po-Ting Chen. *Pulsar-Based Navigation and Timing: Analysis and Estimation.* Escholarship.org, 2018. URL: `https://escholarship.org/uc/item/8jm6v2x9`.

[2] Shibin Song et al. “Novel X-ray Communication Based XNAV Augmentation Method Using X-ray Detectors”. In: *Sensors* 15.9 (2015), pp. 22325–22342. ISSN: 1424-8220. DOI: `10.3390/s150922325`. URL: `https://www.mdpi.com/1424-8220/15/9/22325`.

[3] R. N. Manchester et al. “The Australia Telescope National Facility Pulsar Catalogue”. Version 2.8.0. In: *The Astronomical Journal* 129.4 (Apr. 2005), pp. 1993–2006. DOI: `10.1086/428488`. URL: `https://www.atnf.csiro.au/research/pulsar/psrcat/`.

Sparse Frequency Response Identification and Robust Controller Design Using Nevanlinna–Pick Interpolation

Lukáš Dostálek[1]

1 Introduction

Frequency-domain identification methods are widely used in control engineering due to their intuitive interpretation and direct applicability to controller synthesis. In many practical applications, however, obtaining a sufficiently dense frequency response function (FRF) is time-consuming and challenging, especially for systems with lightly damped resonances or limited excitation capabilities.

This paper presents a data-driven framework enabling robust controller design from a sparse set of experimentally identified frequency response points. The proposed workflow combines adaptive frequency-domain measurement, Nevanlinna–Pick interpolation-based FRF densification, validation of interpolation assumptions, and robust controller synthesis formulated as a linear programming problem. The entire approach is experimentally verified on a laboratory setup consisting of a cantilever beam with collocated actuation and sensing.

The primary contribution lies in the validation-based interpolation procedure that allows reliable reconstruction of dense frequency-domain data from a limited number of measured samples while preserving consistency with measured system behavior.

2 Sparse Frequency Response Identification

The initial step of the proposed method consists of experimental identification of selected points of the system frequency response fully described in Dostálek and Schlegel (2023) and Dostálek et al. (2025). To reduce measurement time while maintaining sufficient information near resonant regions, an adaptive frequency sweep strategy is employed.

The excitation signal is generated using a frequency sweep with adaptively adjusted sweep rate and excitation amplitude. The sweep rate is reduced or increased depending on the measurement accuracy. This is monitored continuously by pausing the sweep and observing the value to which the estimate converges. Simultaneously, the excitation amplitude is modified to keep the experiment close to the operating point.

The FRF point estimates at which the sweep was stopped are considered to be accurate measurements and serve as the output of the identification method.

3 Nevanlinna–Pick Interpolation and Validation

To reconstruct a dense frequency response from sparse measurements, Nevanlinna–Pick interpolation is employed. This interpolation framework enables estimation of admissible transfer function values while preserving analyticity and stability constraints. The interpolation pro-

---

[1] student of the doctoral degree program Cybernetics, field of study Robust controllers with constrained structure, e-mail: lukasdos@fav.zcu.cz

cedure requires specifying assumptions regarding the permissible location of the poles and the upper bound on the magnitude of the transfer function.

In the proposed model, these assumptions can be reduced to two tuning parameters that define the region of acceptable uncertainty for the interpolated model. To verify the validity of the chosen assumptions, the measured FRF samples are divided into interpolation and validation sets, with the interpolation points used to estimate the permissible FRF values at frequencies corresponding to the validation set. For each validation frequency, the Nevanlinna–Pick framework generates a disk of acceptable transfer function values in the complex plane. The measured validation point is then tested to see if it falls within the predicted uncertainty disk.

If the measured point lies outside the admissible region, the interpolation assumptions are considered inconsistent with the measured data. The interpolation parameters are subsequently adjusted until sufficient validation consistency is achieved across the entire dataset. This validation procedure provides a practical mechanism for tuning interpolation assumptions directly from experimental data without requiring explicit parametric system identification.

4 Robust Controller Design

The densified frequency response data are subsequently used for robust controller synthesis. The controller design procedure is formulated directly in the frequency domain and avoids explicit parametric modeling of the plant. The synthesis problem is expressed as a linear programming (LP) problem derived from frequency-domain robustness and performance specifications. The constraints are formulated geometrically in the complex plane and represent design requirements as values of sensitivity functions, taking into account multiplicative uncertainty.

Since the constraints are affine with respect to the controller parameters, the resulting optimization problem can be solved efficiently using standard linear programming techniques. The proposed workflow therefore enables direct transition from sparse experimental frequency measurements to robust controller synthesis while avoiding intermediate high-order parametric modeling steps.

5 Experimental Verification

The proposed method was experimentally verified using a laboratory setup consisting of a cantilever beam equipped with a voice coil actuator for excitation and an optical laser micrometer sensor for measuring the beam's displacement.

Acknowledgement

This research work was supported by the project SGS-2025-020 .

References

Dostálek, L., and Schlegel, M. (2023, June). The frequency response estimation using swept-sine excitation with observer. In 2023 24th International Conference on Process Control (PC) (pp. 72-77). IEEE.

Dostálek, L., Schlegel, M., and Štětina, M. (2025, June). The Swept-sine-excitation-based Frequency Response Estimation for Nonlinear SIMO Systems. In 2025 25th International Conference on Process Control (PC) (pp. 1-5). IEEE.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Distributed Metasolver

Bc. Jonáš Dufek[1]

1 Introduction

Many modern applications, such as artificial intelligence, machine learning, or medical informatics, are centered around **optimizing** the parameters of specialized computational models. This computationally demanding process is typically handled by solver software, the intent of which is to execute an optimization algorithm and return the optimal parameters.

A novel approach to this concept is using multiple optimization algorithms instead of a single one, solvers designed in this way are called **metasolvers**. Furthermore, the solver presented here is distributed, making it possible to distribute the workload across many heterogeneous computational nodes via the **island-based model**. These two ideas thus complement each other, metasolving has the potential of improving the quality and diversity of solutions, whereas the distributed island model improves the overall execution speed.

This distributed metasolver was implemented as a C++ library, which allows it to be easily integrated into existing applications. Example of such an application is the SmartCGMS framework [2], where a unified solver interface can be used to connect any third party solver, allowing it to be used for the optimization of the framework's various glucose prediction models.

The performance characteristics of this library were evaluated on several standardized optimization problems. The results have demonstrated exceptional performance regarding reliability and distribution of work, although the load-balancing mechanism proved to be less efficient in certain metasolving scenarios.

2 Population-Based Optimization

Traditional optimization algorithms, such as gradient descent, maintain a single candidate solution which is iteratively improved until an optimum of the optimized function is reached. This work utilizes population-based optimization algorithms, which differ from classical methods in the fact that they utilize a "**population**" of candidate solutions of an optimization problem, rather than trying to iteratively improve a single one.

These algorithms are especially beneficial for parallel and distributed implementations, as different portions of the population can be optimized independently on multiple processing elements or nodes.

3 Island Model

The island model is a collection of islands connected by a topology, each island maintains a population of individuals and executes an optimization algorithm.

The topology is represented as a graph whose edges define **migration** channels between

---

[1] Master's programme in Applied Sciences, Distributed Computing Systems, Dept. of Computer Science and Engineering, Faculty of Applied Sciences, UWB Pilsen, `jonasd@students.zcu.cz`

islands.

As highlighted by [3], migration is an important part of this process, because it increases the diversity of individuals in each island and thus allows the algorithms to cooperate.

4 System Architecture

The key dependencies of this software are Pagmo [1] and ZeroMQ [4]. Pagmo provides us with optimization algorithms and the "Archipelago-Island-UDI" model. ZeroMQ provides asynchronous messaging and support for multipart messages.

The most important classes of the implemented library are `distributed_controller`, `distributed_worker` and `distributed_island`, which form the basis of the distributed system architecture.

5 Experimental Results

The library was evaluated in six different scenarios covering distributed single-algorithm solving, distributed metasolving, and local metasolving on standard benchmark problems provided by the Pagmo library (Rastrigin, Rosenbrock, Ackley, Griewank, Schwefel, ZDT1 to ZDT4).

The distributed single-algorithm scenario yielded more than adequate results, achieving up to **115%** faster completion times and **24%** better fitness compared to a local solver, with the load-balancing mechanism successfully absorbing a third low-performance node without degrading the cluster's total computational power.

The metasolving scenarios produced mixed results, when all of the used algorithms had similar computational requirements, the metasolver's performance was comparable to the best individual algorithms. However, in some scenarios such as CMA-ES algorithm on the Rosenbrock function, the load-balancing mechanism was unable to prevent the slower algorithm from becoming a bottleneck.

6 Conclusion

The primary objectives of this work were to create and implement a distributed metasolver, together with assessing its performance.

The distributed controller-worker architecture proved to be efficient and reliable when used as part of the Island-Archipelago model. On the other hand, the metasolving performance turned out to be sensitive to the selection of parameters, algorithms and optimization problems.

Future work could include for example improving the Pagmo library's metasolving capabilities, or implementing more robust load-balancing mechanisms.

References

[1] Francesco Biscani and Dario Izzo. A parallel global multiobjective framework for optimization: pagmo. *Journal of Open Source Software*, 5(53):2338, 2020.

[2] Tomas Koutny and Martin Ubl. Smartcgms as a testbed for a blood-glucose level prediction and/or control challenge with (an fda-accepted) diabetic patient simulation. *Procedia Computer Science*, 177:354–362, 2020.

[3] Fei Peng, Ke Tang, Guoliang Chen, and Xin Yao. Population-based algorithm portfolios for numerical optimization. *IEEE Transactions on Evolutionary Computation*, 14(5):782–805, 2010.

[4] ZeroMQ Contributors. Zeromq: High-performance asynchronous messaging library, 2026.

FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

STUDENTSKÁ
VĚDECKÁ KONFERENCE 2026

Návrh kanálků pouzdra redukčního ventilu s využitím matematického modelování proudění stlačitelné tekutiny

Tomáš Eisenhammer[1]

1 Úvod

Redukční ventily se využívají k regulaci a redukci tlaku páry v náročných provozech, kde je potřeba mařit tlakový spád až 50:1. Perpektivní oblastí vývoje jsou redukční pouzdra tvořená labyrintovými kanálky vyráběná aditivní metodou selektivního laserového tavení (SLM). Struktura kanálků způsobuje opakované náhlé změny směru proudění, čímž vytváří požadovanou tlakovou ztrátu.

V rámci této práce byl vyvinut 0D řešič proudění, který umožňuje velmi rychle navrhnout vhodné parametry jednotlivých úseků kanálku. Navazuje na práci Sladký et al. (2025).

2 0D model proudění

Model proudění v kanálku sestavený v rámci této práce patří mezi 0D modely, kdy je proudová oblast redukována na soustavu diskrétních uzlů. Každý uzel (ohyb či větvení) vytváří tlakovou ztrátu charakterizovanou součinitelem místní ztráty ζ. Výhodou tohoto přístupu je, že numerické řešení výsledného diskretizovaného modelu je velmi rychlé. Pro zajištění postupného efektivního škrcení páry v celém kanálku je požadována konstantní střední rychlost v v celé jeho délce.

Proudění mezi dvěma po sobě umístěnými uzly popisuje obecný integrální tvar Bernoulliho rovnice (1) a rovnice kontinuity (2),

$$\zeta_j \frac{v^2}{2} = -\int_{p_j}^{p_{j+1}} \frac{V_m(p)}{M} \mathrm{d}p, \tag{1}$$

$$\frac{V_{m,j+1}}{V_{mj}} = \frac{A_{j+1}}{A_j}, \tag{2}$$

kde p je tlak, V_m molární objem, M molární hmotnost, A plocha průřezu a indexy j, $j+1$ označují veličiny před a za uvažovaným uzlem. Závislost $V_m(p)$ v integrálu v (1) je implicitně určena Redlichovou-Kwongovou stavovou rovnicí reálného plynu (3) a podmínkou zachování entalpie (4),

$$p = \frac{RT}{V_m - b} - \frac{a}{\sqrt{T} V_m (V_m + b)}, \tag{3}$$

$$\Delta H(V_m, T, V_{mj}, T_j) = 0, \tag{4}$$

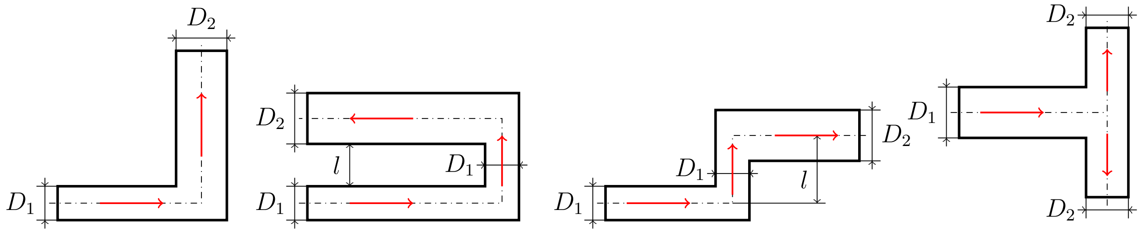
kde R je univerzální plynová konstanta, T teplota, a, b jsou parametry modelu a H je molární entalpie.

---

[1] student navazujícího studijního programu Aplikovaná Mechanika, specializace Dynamika konstrukcí a mechatronika, e-mail: eisentom@students.zcu.cz

3 Určení ztrátových součinitelů CFD simulacemi proudění

Součinitele místní ztráty jednotlivých typů uzlů byly určeny na základě simulací proudění v softwaru ANSYS Fluent. Ztrátový součinitel ζ závisí na geometrických parametrech daného uzlu (viz obr. 1), Reynoldsově čísle a Machově čísle.

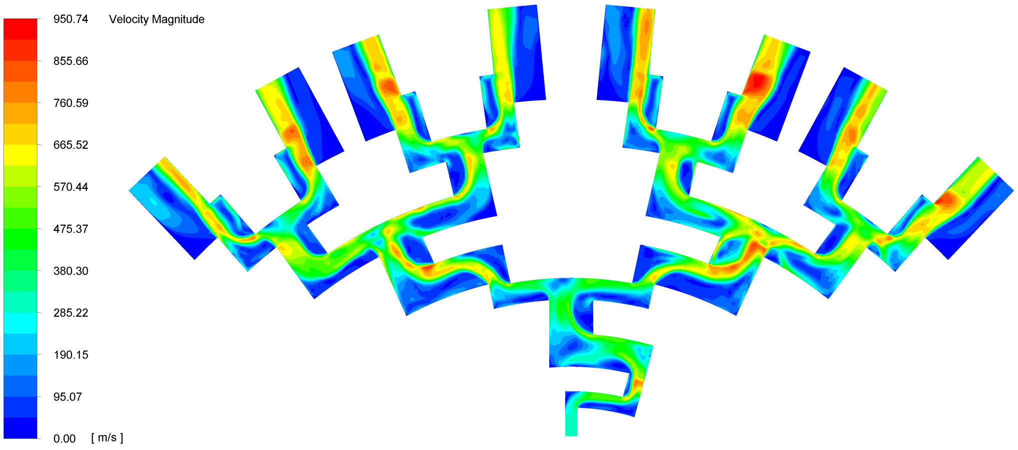


(a) Uzel typu L **(b)** Uzel typu Π **(c)** Uzel typu Z **(d)** Uzel typu T

Obrázek 1: Typy řešených uzlů se zakótovanými geometrickými parametry s vyznačeným směrem proudění

4 Simulace proudění v navrženém kanálku

Přesnost vyvinutého modelu proudění byla otestována porovnáním se simulací proudění ve vybraném navrženém kanálku. Na obr. 2 lze vidět rozložení rychlosti v řezu kanálkem. V predikci průtoku řešeným kanálkem se 0D řešič dobře shoduje se simulací, chyba je 2,5 %.



Obrázek 2: Rozložení rychlosti v podélném řezu středem kanálku

Poděkování

Příspěvek byl podpořen interním grantem SGS-2025-015 ZČU v Plzni. Výpočetní prostředky byly poskytnuty projektem e-INFRA CZ (ID:90254), podporovaného Ministerstvem školství, mládeže a tělovýchovy České republiky.

Literatura

Sladký, J., Blažek, J., Plánička, S., Němeček, S. a Janda, T. Aditivně vyráběná labyrintová pouzdra pro maření tlakových spádů v regulačních armaturách. *Životnost komponent energetických zařízení. Srní, 14.-16. října 2025. Západočeská univerzita v Plzni, 2025, pp. 47-52.*

Modulární GUI pro vizualizační nástroj Vega

Štěpán Faragula[1]

1 Úvod

Vega je mocný nástroj pro tvorbu interaktivních vizualizací, který je využíván především vývojáři a datovými analytiky. Trpí však složitostí vstupního formátu, který vyžaduje ruční psaní JSON specifikací. To může být pro mnoho uživatelů značnou překážkou. Dostupné editory navíc často nabízejí jen omezené možnosti nebo jsou implementovány jako samostatné programy, což komplikuje jejich integraci do vlastních webových aplikací. To v konečném důsledku nutí uživatele neefektivně přepínat mezi různými prostředími.

Hlavním cílem této práce proto bylo navrhnout a implementovat modulární grafické uživatelské rozhraní (GUI) pro nástroj Vega, které umožní vytvářet a upravovat základní typy vizualizací bez nutnosti psaní zdrojového kódu. Výsledkem má být zpřístupnění nástroje Vega širšímu okruhu uživatelů a jeho jednodušší použití při vizualizaci dat.

Návrh uživatelského rozhraní se zaměřuje na snadnou integraci do různých webových aplikací, přičemž jeho stěžejní využití se předpokládá v rámci nástrojové sady Software Process Anti-pattern Detector (SPADe). Tento systém slouží k detekci nevhodných praktik v oblasti procesního a projektového řízení na základě projektových dat. K tomu využívá katalog tzv. anti-vzorů, u nichž byl prokázán negativní dopad na různé oblasti softwarového vývoje, jak je popsáno v práci Brada a Pícha (2019). Vizualizace zde hrají klíčovou roli při prezentaci nalezených anti-vzorů, přičemž nové GUI má umožnit snadnější tvorbu a úpravu těchto grafů.

2 Implementace řešení

Řešení je implementováno jako TypeScript knihovna `@relisa/gui4vega`, která obsahuje celé GUI v podobě jedné React komponenty `VegaEditor`. To umožňuje snadnou integraci do vlastních webových aplikací postavených na tomto frameworku. Projekt je veřejně dostupný na platformě GitHub na adrese `https://github.com/ReliSA/Vega-GUI`. Součástí repozitáře je i uživatelská příručka v souboru *README* a odkaz na programátorskou dokumentaci, která je vygenerována nástrojem TypeDoc z komentářů ve zdrojovém kódu.

Výsledné uživatelské rozhraní je vidět na obrázku 1. Skládá se ze dvou hlavních částí, kde se vlevo nacházejí editační záložky a vpravo náhled vizualizace. Všechny provedené změny se okamžitě promítají do JSON specifikace uložené v paměti, která funguje jako tzv. *Single Source of Truth*. Následně se v reálném čase aktualizuje i samotná vizualizace.

V editační části lze přepínat mezi pěti nástroji. *Wizard* slouží k rychlému generování základních typů grafů podle předdefinovaných šablon. *Data* a *Signals* umožňují manipulovat s datovými sadami a proměnnými specifikace Vega. *Properties* dovoluje upravovat vybrané vlastnosti grafu bez nutnosti zásahu do kódu a *Specification* nabízí textový editor, ve kterém mohou zkušení uživatelé dodefinovat vlastnosti grafu nad rámec toho, co umožňují ostatní záložky.

---

[1] student navazujícího studijního programu Softwarové a informační systémy, e-mail: farag844@students.zcu.cz

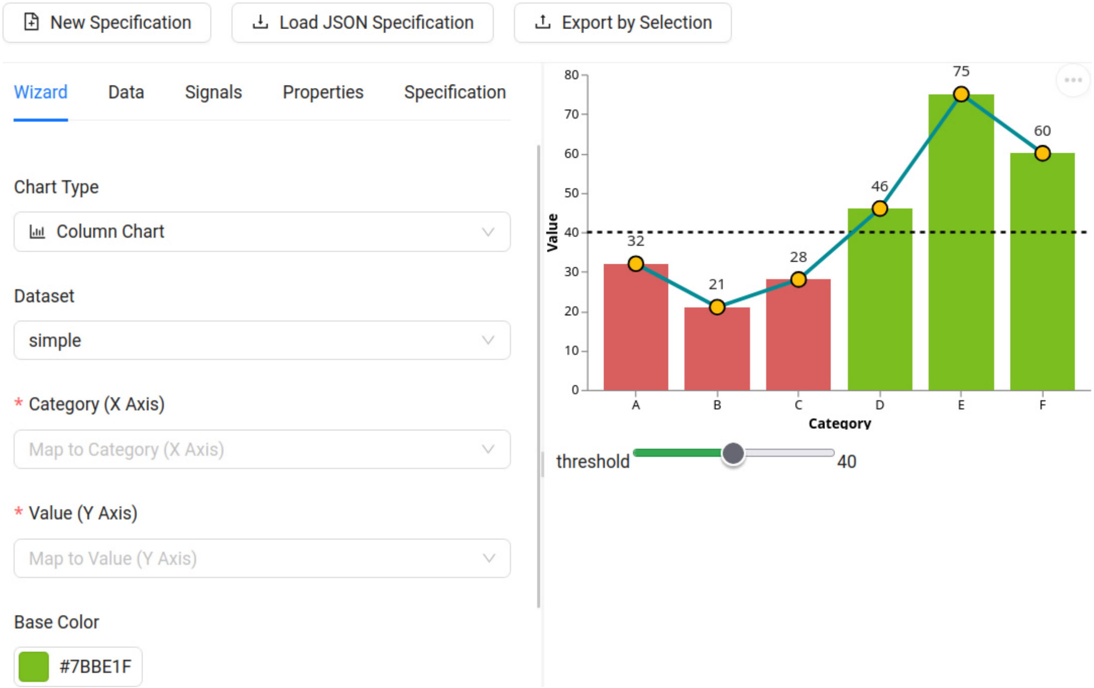


Chart Type

Column Chart

Dataset

simple

* Category (X Axis)

Map to Category (X Axis)

* Value (Y Axis)

Map to Value (Y Axis)

Base Color

🟩 #7BBE1F

threshold 40

Obrázek 1: GUI jako komponenta `VegaEditor` knihovny `@relisa/gui4vega`

3 Dosažené výsledky

Výsledkem práce je modulární a plně funkční grafické uživatelské rozhraní pro nástroj Vega. To lze integrovat do libovolné webové aplikace napsané ve frameworku React na jediném řádku, jak je ukázáno ve zdrojovém kódu 1. Praktická použitelnost řešení byla úspěšně ověřena jeho nasazením do nástrojové sady SPADe, kde se nyní stará o úpravu a vytváření vizualizací anti-vzorů. Kromě toho vznikly i dvě ukázkové aplikace, do kterých je GUI také zakomponováno. Ty jsou součástí repozitáře GitHub v adresářích `demo_antd` a `demo_bootstrap`.

```
import { VegaEditor } from '@relisa/gui4vega';

export default function App() {
    return <VegaEditor height='700px' />;
}
```

Zdrojový kód 1: Zakomponování GUI do vlastního React projektu s minimální konfigurací

Funkčnost a uživatelská přívětivost nového řešení byly ověřeny kombinací jednotkových a uživatelských testů. Jednotkové testy pokrývají logiku GUI a jsou automaticky spouštěny v rámci CI/CD pipeline na platformě GitHub při vydání nové verze knihovny. Uživatelské testování proběhlo se třemi dobrovolníky. Na základě jejich zpětné vazby byly přidány nové funkce a opraveny zjištěné nedostatky. Výsledky ukázaly, že GUI je dostatečně intuitivní a umožňuje vytvářet vizualizace Vega i uživatelům bez hlubší znalosti vstupního formátu tohoto nástroje.

Literatura

Brada, P., a Pícha, P. (2019) Software process anti-patterns catalogue. *Proceedings of the 24th European Conference on Pattern Languages of Programs*. Irsee, Germany: Association for Computing Machinery. Dostupné z DOI: `10.1145/3361149.3361178`.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

W-Doped VO_2 Nanoparticles in SiO_2 Matrix based Thermochromic Coatings via Fast Solid-State Dewetting

Sadoon Farrukh[1], Jaroslav Vlček[2], Elnaz Mohammadi Nia[3], Šárka Zuzjaková[4], Jiri Houška[5],
Radomír Čerstvý[6], Stanislav Haviar[7]

1 Introduction

Global warming and energy crisis drive a focus on energy-saving materials. Buildings have been estimated to produce about 20% of all anthropogenic greenhouse gas emissions and are responsible for up to 40% of the primary energy consumption in the world. Approximately 50% of the total building energy is consumed for compensating the heat gains and losses via windows and glass facades, which are the most energy-inefficient components of buildings. It is evident that energy-saving smart windows with adjustable throughput of solar energy can lower the energy expenditure.

Vanadium dioxide (VO_2) exhibits a reversible phase transition from a low-temperature monoclinic VO_2 (M1) semiconducting phase to a high-temperature tetragonal VO_2 (R) metallic phase at a transition temperature of approximately 68 °C for the bulk material. The automatic response to temperature and the abrupt decrease of infrared transmittance with almost the same luminous transmittance at the transition into the metallic state make VO_2-based coatings a promising candidate for thermochromic smart windows reducing the energy consumption of buildings.

However, practical implementation of VO_2 coatings requires a reduction of the phase transition temperature, enhancement of luminous transmittance, improvement of solar modulation efficiency and long-term environmental stability. Nanostructuring combined with suitable doping represents an effective strategy for overcoming these limitations.

2 Experimental Detail

High-performance thermochromic VO_2-based coatings were prepared using a three-step process consisting of magnetron sputter deposition of SiO_2 films and V-W films without substrate heating followed by fast post-deposition oxidation in pure oxygen. The oxidation process was carried out at temperatures ranging from 400 to 450 °C and resulted in the formation of W-doped VO_2 nanoparticles through a solid-state dewetting mechanism.

The coatings were deposited on standard soda-lime glass (SLG) and quartz glass substrates. Thin V-W precursor layers with thicknesses in the range of 9–11 nm were used for

---

[1] Student of the doctoral degree program Applied Sciences, field of study Plasma physics and physics of thin films, e-mail: sadoon@fav.zcu.cz

[2] Professor of Physics, field Plasma physics and physics of thin films, NTIS, email: vlcek@fav.zcu.cz

[3] Student of the doctoral degree program Applied Sciences, field of study Plasma physics and physics of thin films, elnazmhd@fav.zcu.cz

[4] Researcher at NTIS,field Plasma physics and physics of thin films,email: sazuz@fav.zcu.cz

[5] Professor of Physics, field Plasma physics and physics of thin films, NTIS, email:jhouska@fav.zcu.cz

[6] Researcher at NTIS,field Plasma physics and physics of thin films,email: cerstvy@kfy.zcu.cz

[7] Associate Professor of physics, field Plasma physics and physics of thin films,NTIS ,email: haviar@fav.zcu.cz

the preparation of nanoparticle-based thermochromic layers. The oxidation time was varied between 5 and 20 minutes in order to investigate the influence of processing conditions on the resulting microstructure and functional properties.The final multilayer architecture consisted of two layers of W-doped VO_2 nanoparticles dispersed in a SiO_2 matrix together with a top antireflection SiO_2 layer designed to improve the optical performance of the coating.

3 Conclusion

Thermochromic multilayer coatings containing two layers of W-doped VO_2 nanoparticles dispersed in a SiO_2 matrix were successfully prepared by a fast solid-state dewetting process. The developed fabrication route combines room-temperature magnetron sputtering with fast solid-state dewetting and oxidation using rapid thermal annealing, making it suitable for scalable production on glass substrates.The prepared coatings exhibited a reduced transition temperature of approximately 38 °C together with luminous transmittance exceeding 60% and solar modulation efficiency above 10%, approaching the requirements for practical smart-window applications. The results demonstrate that careful control of the dewetting conditions enables tailoring of nanoparticle morphology, crystallinity and optical properties, leading to enhanced thermochromic performance.The presented approach offers a promising pathway towards cost-effective and energy-efficient smart window coatings based on W-doped VO_2 nanoparticle architecture.

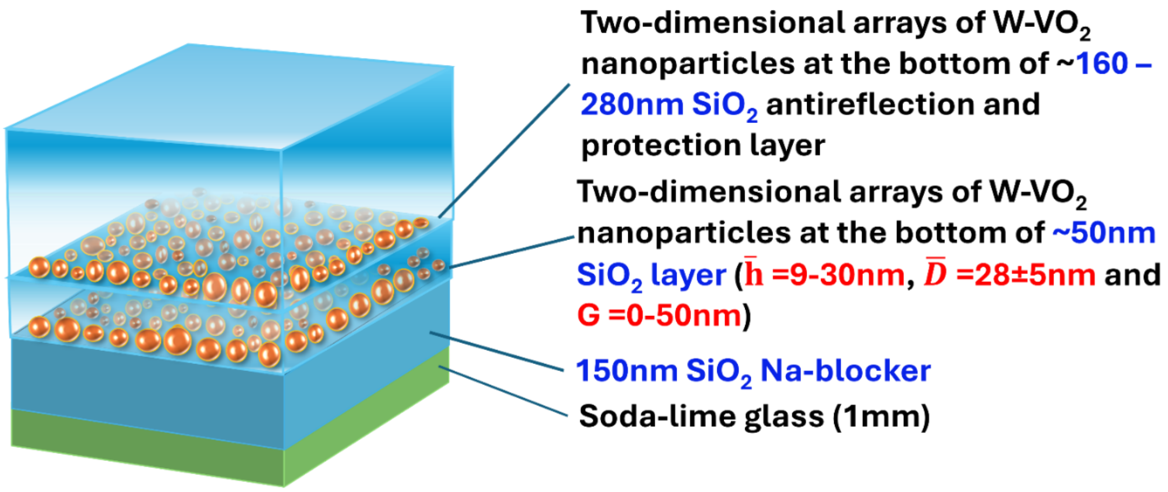


Figure 1*:* Schematic illustration of the thermochromic coating with basic characteristics of the W-doped VO_2 nanoparticles where $\bar{h}$, D and G are their average and diameter, and gap between them, respectively.

4 Acknowledgment

This work was supported by the project Quantum materials for applications in sustainable technologies (QM4ST), funded as project No. CZ.02.01.01/00/22_008/0004572 by Programme Johannes Amos Commenius, call Excellent Research.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Context as a Key: Quantifying Verbatim Data Leakage Across Model Scale, Alignment, and Reasoning Architectures

Ing. Adam Frémund[1]

1 Introduction

Large Language Models (LLMs) introduce severe privacy vulnerabilities through the verbatim memorization of their training data (1). While extraction attacks are well-studied in early base models, the combined impact of scaling, safety alignment (RLHF), and Chain-of-Thought (CoT) reasoning on modern open-weight architectures remains largely unquantified. This paper addresses this gap by presenting a compact empirical evaluation of context-driven data leakage across state-of-the-art models.

2 Methodology

We evaluate verbatim retention using a prefix-suffix context-elicitation attack (1) with prompt prefixes of 50, 150, and 250 tokens, standardized across architectures via the Llama-3 tokenizer baseline. High-entropy sequences from academic and code subsets of The Pile and RefinedWeb are leveraged as proxy pre-training distributions to isolate true memorization from linguistic guessing.

Our testbed evaluates the lifecycle of modern models, including Llama 3.1 (8B/70B) (2), Qwen 2.5 (7B), DeepSeek-R1-Distill-Qwen-7B, Gemma 2 9B, and Mistral Nemo 12B, utilizing deterministic greedy decoding. To robustly capture partial data leaks across varying tokenizer compression ratios, we employ character-level Longest Common Prefix (LCP) as our primary evaluation metric.

3 Experimental Results and Analysis

The empirical results across 5,000 high-entropy probes demonstrate significant architectural variance along four key axes:

3.1 The Context Multiplier Effect

Context acts as an activation threshold for memorized weights. While 50-token prefixes yield shallow memorization, extending prompts to 150 tokens sharply amplifies the leak rate. For Gemma 2, Llama 3.1, and Mistral Nemo, extraction rates saturate at 150 tokens, whereas Qwen 2.5 continues to scale upward at 250 tokens.

3.2 Scale-Memorization Correlation and Alignment Tax

Parameter scaling directly expands memorization capacity, as Llama 3.1 70B Base exhibits a deeper average LCP than its 8B counterpart. Conversely, instruction-tuning imposes

---

[1] student of the doctoral degree program Applied Sciences, field of study Cybernetics, e-mail: afremund@kky.zcu.cz

an alignment tax that significantly suppresses regurgitation depth. However, this defense is partially overridden by the sheer capacity of 70B architectures, which still exhibit substantial baseline leakage.

3.3 Distillation as Suppression

DeepSeek-R1-Distill-Qwen-7B shows a near-total collapse in verbatim recall compared to the standard Qwen 2.5 Instruct. Inference-time ablation studies confirming identical zero-leakage profiles with or without reasoning tokens isolate this privacy resistance to data scrubbing or catastrophic forgetting during the distillation phase, rather than dynamic inference disruption.

3.4 Architectural Outliers

Within the ∼8B class at a 250-token context, Qwen 2.5 7B is the most resilient against extraction. Mistral Nemo and Llama 3.1 8B form a consistent baseline, while Gemma 2 9B acts as a high-risk outlier, with leak frequencies exceeding 7% (see Figure 1).

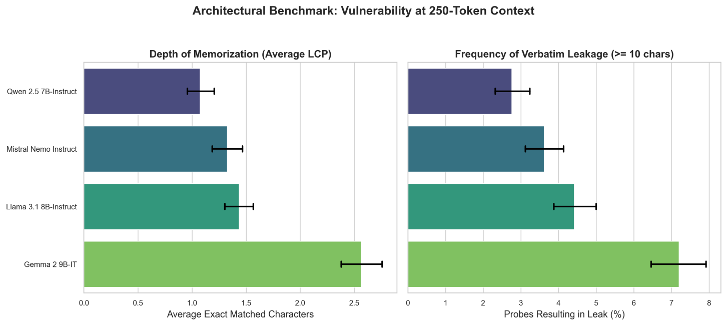


Figure 1: Architectural Vulnerability Benchmark: Ranking modern ∼8B class instruct models by relative extraction risk at a 250-token context window.

4 Conclusion

Our evaluation highlights a fundamental trade-off: features driving top conversational performance—such as parameter scaling and expanded context windows—directly exacerbate data extraction risks. Post-hoc instruction-tuning serves as a superficial friction layer that fails to fully secure large-scale architectures, making structural privacy defenses or intentional distillation filtering a critical imperative for downstream deployment.

Acknowledgement

This work was supported by the University of West Bohemia grant, project No. SGS-2025-011

References

[1] Carlini, N., et al.: Extracting training data from large language models. USENIX Security Symposium (2021).

[2] Dubey, A., et al.: The llama 3 herd of models. arXiv preprint arXiv:2407.21783 (2024).

FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

STUDENTSKÁ
VĚDECKÁ KONFERENCE 2026

Studium kritického chování korelovaných elektronů v jednorozměrných supravodivých nanostrukturách

Jan Gebel[1], Šimon Kos[2], Sunil Wilfred D'Souza[3], Ján Minár[4], Václav Janiš[5]

1 Úvod do problematiky

Supravodivost je makroskopický kvantový jev, který nastává po ochlazení materiálu pod kritickou teplotu T_c a spočívá mimo jiné ve schopnosti materiálu vést elektrický proud s nulovým odporem. Dalším důležitým projevem supravodivosti je tzv. Meissnerův–Ochsenfeldův jev, při němž povrchové stínící proudy vytvářejí magnetické pole opačné orientace vůči vnějšímu poli, čímž dochází k vytlačení vnějšího magnetického pole z objemu supravodiče. Magnetické pole proto proniká pouze do tenké povrchové vrstvy materiálu o charakteristické hloubce λ_L (Londonova penetrační hloubka).

Mikroskopický mechanismus supravodivosti vyplývá z tzv. BCS-teorie supravodivosti (Bardeen, Cooper , Schrieffer). Tato teorie spočívá ve tvorbě tzv. Cooperových párů elektronů, které mají opačný spin a jsou párovány prostřednictvím kmitů atomů krystalové mříže (fononů). Poloměr těchto párů mimo jiné určuje tzv. koherenční délka ξ_N.

Klesající rozměry objemu materiálu jsou doprovázeny změnou mnoha vlastností včetně supravodivosti. Pokud jsou rozměry menší než ξ_N a λ_L, je přechod do supravodivého stavu značně potlačen, případně k němu nemusí vůbec dojít a naopak je posíleno kritické chování. Tento projekt byl věnován teoretickému a numerickému studiu 1d modelu, kde jsou ostatní rozměry zanedbatelně malé, s cílem toto kritické chování, kde se konkrétně uplatňují elektronové korelace, co nejvíce numericky a teoreticky analyzovat.

2 Formulace modelu

V rámci projektu byl studován jednopásový Hubbardův model s přitažlivou interakcí pomocí přibližného tzv. Baym-Kadanoffova formalismu, na základě čehož byly odvozeny rovnice pro propagaci elektronů a jejich párové korelace. Tím se problematika studia kritického chování zredukovala na soustavu rovnic, kterou bylo nutno řešit numericky (1).

Klíčovým výstupem numerické analýzy bylo studium veličiny a, která popisuje míru přiblížení ke kritickému bodu. Kvůli rozměrovým omezením však systém není schopen tohoto bodu dosáhnout. Samotný kritický bod by odpovídal nulové hodnotě veličiny a, zatímco supra-

---

[1] Student navazujícího studijního programu Aplikované vědy a informatika, obor Aplikovaná fyzika a fyzikální inženýrství, e-mail: jgebel@students.zcu.cz

[2] Vedoucí diplomové práce a vyučující na Katedře fyziky, e-mail: simonkos@kfy.zcu.cz

[3] Specialista konzultant pro numerickou část práce, Nové technologie – výzkumné centrum, e-mail: dsouza@ntc.zcu.cz

[4] Odborník na kvantové materiály a modelování pomocí metod DFT a DMFT, vedoucí výzkumného týmu Kvantové materiály, Nové technologie – výzkumné centrum, e-mail: jminar@ntc.zcu.cz

[5] Vedoucí projektu, odborník na teoretický popis silně korelovaných materiálů, Fyzikální ústav Akademie věd ČR, e-mail: janis@fzu.cz

vodivé uspořádání by nastalo při dosažení záporných hodnot této veličiny. (1):

$$a = 1 + UY, \tag{1}$$

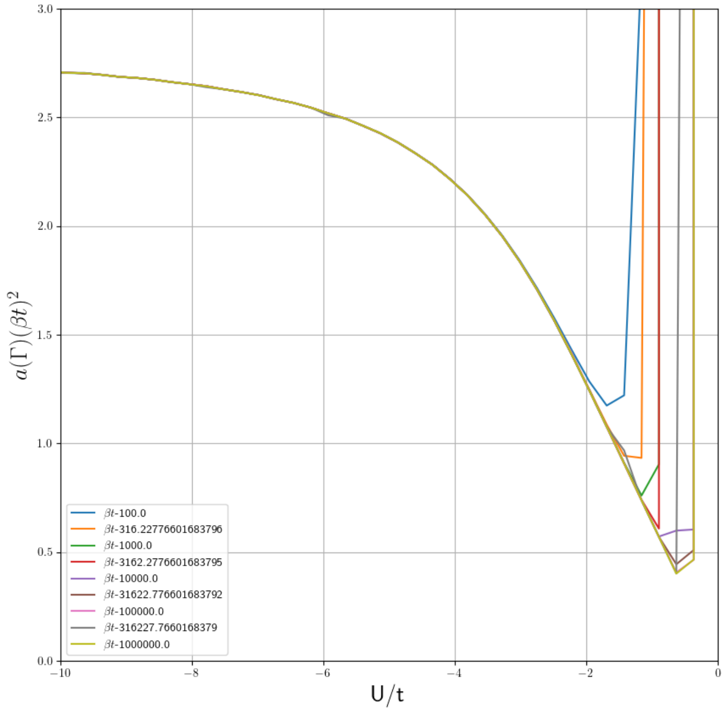
Renormalizovaná interakce korelovaného pohybu (tzv.ireducibilní vrchol) má potom tvar :

$$\Gamma = \frac{U}{\beta\pi}\sqrt{\frac{2}{DU}}\frac{1}{\sqrt{a}}. \tag{2}$$

kde Y popisuje propagaci páru elektronů a D její závislost na hybnosti. Obě veličiny jsou definovány pomocí integrálů, které bylo nutné počítat numericky. U představuje interakci mezi elektrony a β je inverzní teplota (1).

3 Výsledky

Cílem projektu bylo získání veličiny a v závisloti na 2 energetických škálách a to $-U$ a teplotě v jednotkách t, což je amplituda přeskoku v Hubbardově modelu



Hlavním výsledkem bylo popsání přechodu mezi vysokoteplotním nekorelovaným a nízkoteplotním korelovaným režimem. Nízkoteplotní režim vykazuje kritické škálování úměrné druhé mocnině teploty, v jehož důsledku závislosti pro různé teploty kolabují na jednu křivku. Oba režimy zároveň vykazují opačnou závislost na $-U$, což vede ke vzniku minima veličiny a.

Literatura

[1] Kos, Š., D'Souza, S., Gebel, J., Minár, J., a Janiš, V. (2025) Critical behavior of the 1d superconductor in the FLEX approximation. *Prezentace na European COST Action CA21144 "SuperQumap"conference*. Liblice, Czech Republic.

Volume invariant Cosserat rods for position-based (PBD) muscle deformation

Bc. Duc Long Hoang[1]

1 Introduction

When a muscle contracts, its fibers shorten – but muscle tissue is nearly incompressible, so it bulges outward to preserve volume. Current muscle deformation models based on PBD/XPBD (Kohout and Červenka , 2021; Havlíček , 2024) operate only on the muscle surface, lacking internal fiber structure and local volume preservation. This thesis replaces the surface-only approach with bundles of volume-invariant Cosserat rods (Kugelstadt and Schömer , 2016; Angles et al. , 2019) that naturally produce the shortening-and-bulging effect of real muscles.

2 Method

This work builds on Červenka's PBD and Havlíček's XPBD implementations within the Muscle Wrapping 2.0 project. Cosserat rods model 1D elastic structures with bending, twisting, stretching, and shearing. The VIPER extension adds a variable radius, coupling stretch to cross-section – when a segment shortens, its radius increases to preserve volume and vice versa.

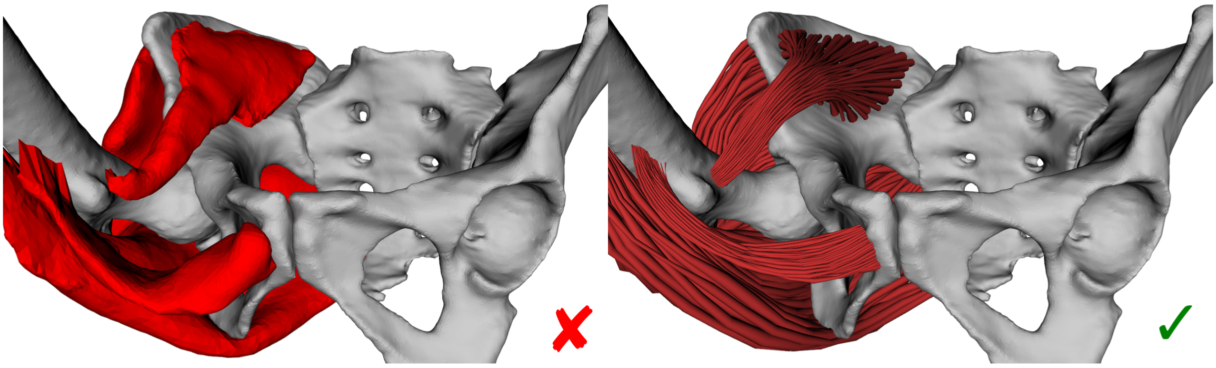


Figure 1: Surface model (left) – muscles sag and lose shape. Cosserat rods (right) – shape and volume are preserved under contraction

Rods are added as a third entity class in the existing XPBD solver alongside muscles and bones. Three constraint types govern the simulation: stretch-shear for rod shape, bend-twist for curvature, and volume preservation for the radius-stretch coupling. Activation signals from OpenSim shorten rod rest lengths; the volume constraint increases radii in response, producing visible bulging (Fig. 1). Per-rod constraints share no degrees of freedom between rods, making them trivially parallelizable.

---

[1] student of the master degree program Software and Information Systems, field of study Computer Science and Engineering, e-mail: duclong@students.zcu.cz

3 Results

Validated on iliacus, gluteus maximus, gluteus medius, and adductor brevis during hip flexion. Volume is perfectly preserved ($V(t)/V(0) = 1.0$) throughout. Moment arms align with published biomechanical data (Blemker and Delp , 2005; Modenese and Kohout , 2020; Kumazaki et al. , 2022) as shown in Fig. 2. The parallel rod-based approach achieves a $24\times$ speedup over the surface-based method.

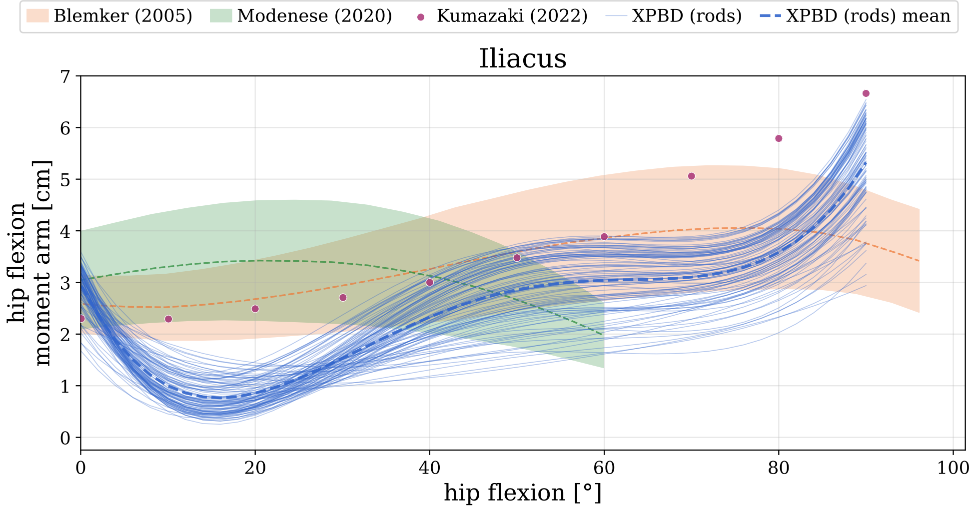


Figure 2: Iliacus hip flexion moment arms compared against Blemker, Modenese, and Kumazaki

4 Conclusion

Rod-based muscle simulation with volume preservation is both feasible and fast, capturing internal fiber behavior that surface-based approaches fundamentally cannot.

Acknowledgement

I would like to thank my thesis advisor doc. Ing. Josef Kohout, Ph.D. for his guidance.

References

Angles, B., et al. (2019) VIPER: Volume Invariant Position-based Elastic Rods. *Proc. ACM Comput. Graph. Interact. Tech.*, 2(2), pp. 1–26.

Blemker, S.S., Delp, S.L. (2005) Three-Dimensional Representation of Complex Muscle Architectures and Geometries. *Ann. Biomed. Eng.*, 33(5), pp. 661–673.

Kohout, J., Červenka, M. (2021) Muscle Deformation Using Position Based Dynamics. In: *Biomedical Engineering Systems and Technologies*. Springer, pp. 486–509.

Havlíček, O. (2024) *Muscle interaction in the context of muscle deformation modelling by a Position Based Dynamics method*. Master's thesis. Pilsen, Univ. of West Bohemia.

Kugelstadt, T., Schömer, E. (2016) Position and Orientation Based Cosserat Rods. *Eurographics/ACM SIGGRAPH Symp. on Computer Animation*.

Kumazaki, T., et al. (2022) Action and Contribution of the Iliopsoas and Rectus Femoris as Hip Flexor Agonists. *Juntendo Medical Journal*, 68(4), pp. 352–362.

Modenese, L., Kohout, J. (2020) Automated Generation of Three-Dimensional Complex Muscle Geometries. *Ann. Biomed. Eng.*, 48(6), pp. 1793–1804.

Automatic Recognition of Parcel Numbers from Historical Maps

Bc. Vladimír Holý[1]

1 Introduction

This master's thesis focuses on automating data extraction from the historical Stable Cadastre using neural networks. The aim of the thesis was the digitization of the Stable Cadastre, specifically the detection of parcel numbers from cadastre maps and text recognition from written records. The primary challenge lies in the materials' age and degraded condition. To address this, two object detection and four text recognition methods were utilized.

2 Methodology & Analysis

As these are historical documents, the primary challenge is their age: the papers are yellowed, stained, and handwritten. Furthermore, the written records suffer from various geometric deformations. Based on an analysis of object detection methods, the EAST and YOLO architectures were selected. For text recognition, both classical architectures based on Convolutional Neural Networks (CNN) and modern Large Vision-Language Models (VLM) were considered. Four architectures were subsequently selected for implementation: Kraken, GOT-OCR 2.0, HunyuanOCR, and QWEN3.

3 Key Aspects of Implementation

A crucial part of the implementation was the development of two annotation applications, one for parcel numbers on maps and another for table cells in written records. Of note is a custom grid-detection algorithm that first uses a small YOLO model to detect line intersections, then filters and clusters the results into rows and columns. This algorithm effectively processes tables with horizontal misalignment and curved rows. A detailed schema of the algorithm is shown in Figure 1, and the final detection is shown in Figure 2. Another phase involved hyperparameter optimization and subsequent training for both detection and text recognition models.

4 Results

Experiments demonstrated that high model reliability is achievable despite the age of the documents. In detection, the EAST model performed best, achieving an impressive Recall rate of 94.0%, meaning it can detect up to 94.0% of parcel numbers on the map. The main limitation is densely built-up areas, where numbers are tightly packed and difficult to detect. Detection results are shown in Table 1.

---

[1] student of the master's degree program Engineering Informatics, field of study Software Engineering, e-mail: holyvl@students.zcu.cz

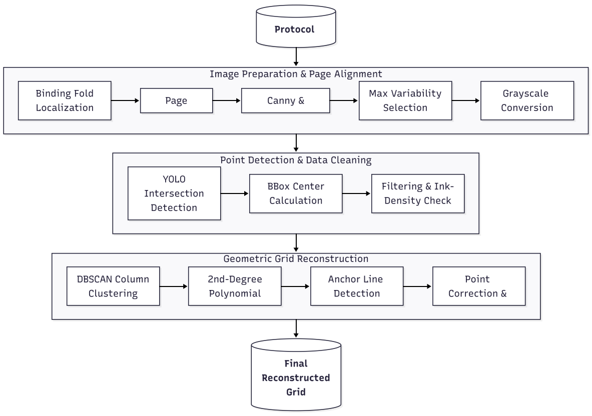


Figure 1: Grid-detection algorithm Schema

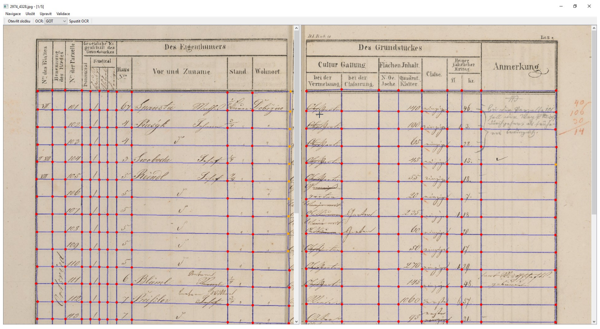


Figure 2: Example of grid detection

Model	Precision	Recall	F1 score	mAP50	IoU
EAST	**91.0**	**94.0**	**92.5**	85.9	75.0
YOLO	86.4	90.0	88.2	**87.3**	**79.5**

Table 1: Detection results [in %].

In both parcel-number and written-record recognition, the Hunyuan model achieved excellent results. In parcel number recognition, the GOT model is also noteworthy; it is up to three times faster than Hunyuan while achieving comparable performance. Surprisingly, the Kraken model achieved results comparable to those of the large VLM model, QWEN. Otherwise, in written-record recognition, the GOT and Kraken models couldn't match the performance of large VLM models. The results are shown in Table 2.

Type of recognition	Metric	Kraken	GOT	Hunyuan	QWEN
Parcel number	CER	20.7	17.0	**16.0**	20.4
	Accuracy	61.4	66.7	**67.4**	64.2
	Time	**0.1s**	0.7s	2.2s	4.3s
Written record	CER	31.8	21.2	**8.6**	11.8
	Accuracy	0.3	0.6	**27.0**	11.4
	Correct Column Count	48.6	81.5	92.6	**95.2**

Table 2: Results of OCR models [in %].

5 Conclusion

Despite the document's age, the models achieved high accuracy and low error rates. The trained models were successfully integrated into the annotation applications, resulting in a functional system capable of automatically processing historical Stable Cadastre maps and written records. This work confirms that modern neural networks can effectively replace manual data extraction from historical documents, facilitating further work with historical data in the Czech Republic.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Learning Where to Look: Adaptive Information Acquisition Under Observation Budget

Tomáš Honzík[1]

1 Introduction

Modern machine learning systems usually assume that the complete input is available before prediction, but this assumption becomes restrictive when inputs are very large, costly to process, or expensive to acquire. Examples include high-resolution video, medical scans, robotic sensing, and scientific measurements requiring experiments or simulations. In such settings, efficiency depends not only on the predictor, but also on which parts of the input are observed.

Naive input reduction, such as downsampling, cropping, frame-rate reduction, or random subsampling, are task-agnostic and may discard informative and uninformative content alike. We instead study whether a model can actively acquire a small number of task-relevant observations before prediction.

As a controlled toy problem, we use MNIST [3] digit classification under a fixed pixel-observation budget. Instead of observing the full 28×28 image, the model may reveal only k pixels before predicting the digit. We compare uniformly random pixel selection with an adaptive policy that chooses each new pixel conditioned on previous observations. Both methods use the same frozen downstream classifier, isolating the effect of the observation policy.

2 Related Work

Related work includes hard visual attention, where models sequentially choose image regions or glimpses to process [6, 1, 7], and active feature acquisition or budgeted prediction, where models select features under an observation or acquisition cost [11, 9, 8, 5, 4, 2, 10]. These approaches are commonly formulated using reinforcement learning, budgeted classifiers, generative models, information-theoretic objectives, or learned acquisition functions. To our knowledge, our work is the first to formulate adaptive information acquisition as supervised imitation of classifier-induced pseudo-oracle observation targets.

3 Method

Each MNIST image is flattened into a vector $x \in [0, 1]^{784}$. At step t, the model observes a pixel-value pair

$$o_t = (i_t, x_{i_t}),$$

where i_t is the selected pixel position and x_{i_t} is the revealed intensity. After observing at most k pixels, the classifier predicts the digit label y. Each observation is encoded by combining a learned embedding of the pixel position with a learned projection of the pixel value. A beginning-of-sequence token represents the state before any real pixel has been observed.

The experiment has two stages. First, we train a causal transformer classifier on randomly sampled pixel sequences. The classifier is trained to predict the digit label from every prefix of the

---

[1] student of the doctoral degree program Applied Sciences, field of study Cybernetics, e-mail: thonzik@students.zcu.cz

sequence, so it can produce predictions after any number of observed pixels. This model defines the random-observation baseline. After training, the classifier is frozen and reused for all evaluations.

Second, we train an adaptive pixel-selection policy. The policy is trained using supervision derived from the frozen classifier rather than by reinforcement learning. For a given partial observation state, several candidate continuations are sampled. The frozen classifier evaluates these candidates by the classification loss obtained after observing them. The candidate that gives the lowest loss is treated as a pseudo-oracle action, and the policy is trained by ordinary gradient descent to imitate this locally preferred next observation. In this way, the classifier supplies a task-dependent signal for where the model should look next.

At inference time, the adaptive policy begins with the beginning-of-sequence token and selects pixels autoregressively. At each step, it scores all pixel positions, reveals the highest-scoring unobserved pixel, appends the corresponding pixel-value pair to the sequence, and repeats until the budget is exhausted. The frozen classifier then predicts from the resulting partial observation sequence. Performance is measured as classification accuracy as a function of the number of observed pixels.

4 Results

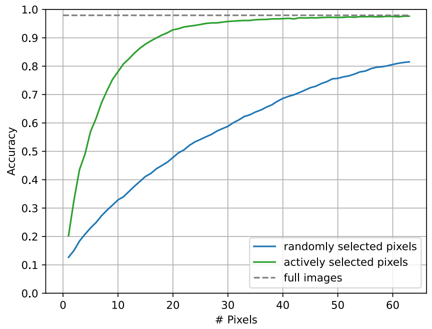


Figure 1: Classification accuracy as a function of the number of observed pixels. The random baseline (blue) selects pixel positions uniformly at random, while the adaptive policy (green) sequentially selects pixels based on the observations seen so far. Adaptive selection achieves substantially higher accuracy at every pixel budget, reaching 95.8% accuracy with 30 observed pixels compared to 58.9% for random selection. At the maximum evaluated budget of 63 pixels, adaptive selection reaches 97.6%, while random selection reaches 81.2%.

Acknowledgement This work was supported by specific university research project **SGS-2025-020**. Computational resources were provided by the e-INFRA CZ project **ID:90254**, supported by the Ministry of Education, Youth and Sports of the Czech Republic.

AI Use Statement Large Language Models were used to assist with rephrasing and structuring the text. All content was reviewed, verified, and finalized by the author.

References

[1] Jimmy Ba, Volodymyr Mnih, and Koray Kavukcuoglu. Multiple object recognition with visual attention. In *International Conference on Learning Representations*, 2015.

[2] Ian Covert, Wei Qiu, Mingyu Lu, Nayoon Kim, Nathan White, and Su-In Lee. Learning to maximize mutual information for dynamic feature selection. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*. PMLR, 2023.

[3] Yann LeCun, Corinna Cortes, and Christopher J. C. Burges. The mnist database of handwritten digits. `http://yann.lecun.com/exdb/mnist/`, 1998. Accessed: 2026-05-19.

[4] Yang Li and Junier B. Oliva. Active feature acquisition with generative surrogate models. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*. PMLR, 2021.

[5] Chao Ma, Sebastian Tschiatschek, Konstantina Palla, José Miguel Hernández-Lobato, Sebastian Nowozin, and Cheng Zhang. EDDI: Efficient dynamic discovery of high-value information with partial VAE. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*. PMLR, 2019.

[6] Volodymyr Mnih, Nicolas Heess, Alex Graves, and Koray Kavukcuoglu. Recurrent models of visual attention. In *Advances in Neural Information Processing Systems 27*, pages 2204–2212, 2014.

[7] Samrudhdhi B. Rangrej, Chetan L. Srinidhi, and James J. Clark. Consistency driven sequential transformers attention model for partially observable scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

[8] Hajin Shim, Sung Ju Hwang, and Eunho Yang. Joint active feature acquisition and classification with variable-size set encoding. In *Advances in Neural Information Processing Systems 31*, 2018.

[9] Kirill Trapeznikov and Venkatesh Saligrama. Supervised sequential classification under budget constraints. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, volume 31 of *Proceedings of Machine Learning Research*. PMLR, 2013.

[10] Michael Valancius, Max Lennon, and Junier B. Oliva. Acquisition conditioned oracle for nongreedy active feature acquisition. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*. PMLR, 2024.

[11] Zhixiang Xu, Kilian Q. Weinberger, and Olivier Chapelle. The greedy miser: Learning under test-time budgets. In *Proceedings of the 29th International Conference on Machine Learning*, 2012.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Use of predictive models in the Czech Salivary Gland Database application

Vojtěch Jelínek[1]

1 Introduction

The Czech Salivary Gland Database (CSGDB) was initially established to collect clinical data, but the clinical potential of this data for predictive modeling led to the formal thesis assignment. The primary objective was to study predictive models in oncology and compare different approaches, including scoring systems, to create a valid and dynamic model capable of updating with new database entries. This work involved analyzing the existing database structure for its suitability in processing clinical and pathological data, designing models for individualized prognostic scores in collaboration with specialists from Motol University Hospital, and integrating these into the database as a clinical decision support tool. Finally, the goal was to deploy the solution at Motol University Hospital and critically evaluate the accuracy of the models using retrospective data.

2 Analysis

A technical and clinical audit of the CSGDB revealed implementation debt, including a lack of database normalization and horizontal redundancy that complicated the extraction of data for machine learning. The existing architecture suffered from tight coupling between the React frontend and the Node.js main process, featuring leaky abstractions that hindered the integration of new analytical features. Furthermore, the rarity of salivary gland carcinomas created a sparse data environment with a low event-per-variable ratio, rendering deep learning architectures like DeepSurv structurally underdetermined and prone to overfitting. Strict security requirements and legal mandates regarding the processing of sensitive oncological data necessitated a solution that could operate entirely offline within the hospital's secure environment. Consequently, the analysis identified a critical need for a decoupled, sidecar-based architecture that could facilitate local survival analysis while maintaining data residency on the physician's workstation.

3 Proposed solution

The redesigned system implements a normalized ERA model that eliminates redundancy by utilizing a unified patient inheritance structure and a flexible TNM staging. To perform complex survival analysis without sacrificing UI responsiveness, a Python-based machine learning engine was integrated via a sidecar pattern (see Fig. 1) communicating with the Electron app through a JSON contract over standard I/O streams. The system utilizes a dual-model approach, leveraging the Random Survival Forest to capture non-linear feature interactions and a regular-

---

[1] student of the master degree program Applied Sciences, field of study Software and Information Systems, e-mail: vjelinek@students.zcu.cz

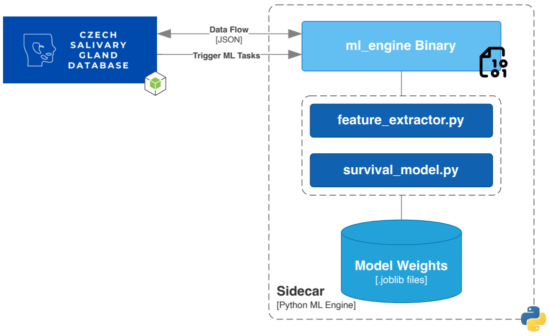
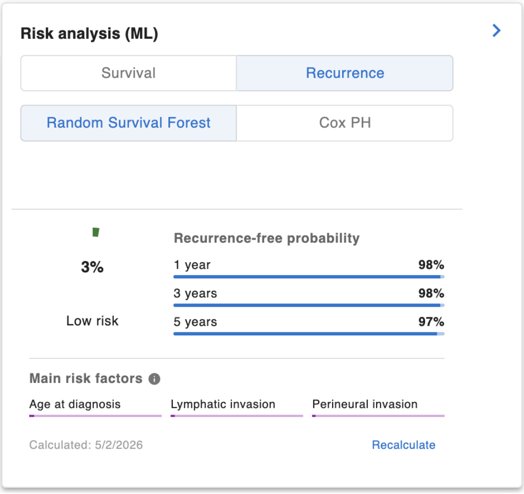


Figure 1: Sidecar pattern

ized Cox Proportional Hazards model to provide hazard ratios compatible with clinical literature.Automated retraining triggers were implemented to ensure the models evolve alongside the local dataset, and a ”cold start” mechanism was introduced to provide pre-trained baseline models from Motol Hospital upon initial deployment. The final user interface features a risk card (see Fig. 2) that presents survival probabilities at 1-, 3-, and 5-year intervals while identifying the top clinical factors driving each prediction to foster clinical trust.



Risk analysis (ML)

Survival | Recurrence

Random Survival Forest | Cox PH

3%

Low risk

Recurrence-free probability

1 year 98%

3 years 98%

5 years 97%

Main risk factors

Age at diagnosis

Lymphatic invasion

Perineural invasion

Calculated: 5/2/2026

Recalculate

Figure 2: Patient risk card

4 Results and Conclusion

Statistical validation on a cohort of 122 patients using the .632 bootstrap estimator has shown that the integrated models meet established performance standards for oncological prognosis. The Random Survival Forest achieved a concordance index of 0.814 for recurrence-free survival, representing a modest improvement over the Cox Proportional Hazards baseline, though the results remain constrained by the limited sample size. Sensitivity analysis justified the use of a 7-feature set, confirming that chronological age and pathological staging are critical for mortality and recurrence estimation. The system has been successfully deployed at Motol University Hospital, where clinical feedback suggests it is a practical tool for prognostic stratification and personalized treatment planning. In conclusion, the project successfully transitioned the database from a data-entry tool into a secure decision support system.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

First Steps Toward a Multimodal System for Spinal Diagnostics: From 3D Skeleton-Based Action Recognition to Clinical Application

Andrea Kadlecová[1]

1 Introduction

Spinal pathologies, such as spondylolisthesis or stenosis of the lumbar spine, are primarily diagnosed and evaluated using X-ray imaging technology. However, these diseases also affect the patient's overall movements, gait, and posture, which is often caused by back pain, tingling in the legs, and other related symptoms.

Relying only on X-ray images limits the ability to objectively assess the severity of the disease and the patient's condition. This could be avoided by creating a multimodal diagnostic system that processes both X-ray images and data obtained using Vicon.

The first step is to focus on the processing and classification of patient's movements. This paper explores the 3D skeleton-based action recognition models. Models CTR-GCN with the G^3CN module and AutoregAd-HGformer are investigated. They were afterwards trained on the UAV-Human dataset to verify their functionality and effectiveness for the given problem. The evaluation of these networks serves as a baseline for future work, which aims to adapt these models for more specific classification of spinal diseases based on complex biomechanical data.

2 3D Skeleton-Based Action Recognition

This section details the models and data configurations utilized for the initial action analysis. The two models mentioned below were created for 3D skeleton-based action recognition, an illustrative example of the skeleton is provided in Figure 1. They were evaluated on the UAV-Human dataset created by Li et al. (2021), which consists of data collected by drones. The authors of the dataset evaluated it using two different cross-subject data splits, referred to as *v1* and *v2*. This work keeps these exact same data splits to ensure a fair comparison between the models.

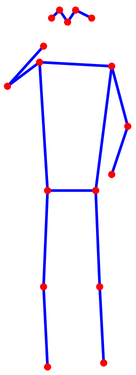


Figure 1: Visualization of the skeleton

The CTR-GCN described by Chen et al. (2021) is utilized as the baseline architecture, enhanced with the G^3CN module proposed by Ren et al. (2025) for advanced spatial feature extraction. In the standard GCN approach, the human skeleton is modeled as a graph, where joints represent nodes and bones represent edges. These connections between joints are subsequently represented by an adjacency matrix. The innovation of the integrated G^3CN module is Gaussian topology, which aims to significantly improve the classification of similar actions.

---

[1] PhD student of Cybernetics, Artificial Intelligence, e-mail: kadlecoa@fav.zcu.cz

In contrast to the CTR-GCN with the G^3CN module, the AutoregAd-HGformer model described by Ray et al. (2025) does not utilize a fixed adjacency matrix. Instead, it processes the 3D skeleton using a hypergraph, which better captures the complex relationships among joints.

3 Results and Conclusion

The obtained results are summarized in Table 1. Among the evaluated networks, the AutoregAd-HGformer achieves the highest overall accuracy. However, the accuracy metrics across both tested models remain relatively comparable.

Given these similar accuracy rates, computational efficiency becomes a crucial factor for a comprehensive evaluation. For this reason, FLOPs are examined to further assess and compare the actual computational cost of each model.

As detailed in Table 1, the complexity analysis indicates a trade-off between model size and computational demands. he CTR-GCN with the G^3CN module is structurally larger, consisting of 6.32 M parameters. In contrast, the AutoregAd-HGformer is significantly more compact, requiring only 3.20 M. However, when evaluating the computational cost during inference, the AutoregAd-HGformer demands nearly twice as many operations compared to the CTR-GCN configuration. Although the AutoregAd-HGformer delivers the highest accuracy and requires fewer parameters, it demands significantly more computational power, which must be considered when deploying the system in practical clinical applications.

Model	*v1* [%]	*v2* [%]	Params $[M]$	FLOPs $[G]$
CTR-GCN w/G^3CN	39.43	65.46	6.32	**8.56**
AutoregAd-HGformer	**39.46**	**68.71**	**3.20**	15.4

Table 1: Results and computational complexity of the evaluated models on the UAV-Human dataset

Acknowledgement

The work was supported by a grant project No. SGS-2025-011 of the University of West Bohemia.

References

Chen, Y., Zhang, Z., Yuan, C., Li, B., Deng, Y., and Hu, W., 2021. Channel-wise Topology Refinement Graph Convolution for Skeleton-Based Action Recognition. *arXiv preprint arXiv:2107.12213*. Available from: https://arxiv.org/abs/2107.12213

Li, T., Liu, J., Zhang, W., Ni, Y., Wang, W., and Li, Z., 2021. UAV-Human: A Large Benchmark for Human Behavior Understanding with Unmanned Aerial Vehicles. *arXiv preprint arXiv:2104.00946*. Available from: https://arxiv.org/abs/2104.00946

Ray, A., Raj, A., and Kolekar, M.H., 2025. Autoregressive Adaptive Hypergraph Transformer for Skeleton-based Activity Recognition. *arXiv preprint arXiv:2411.05692*. Available from: https://arxiv.org/abs/2411.05692

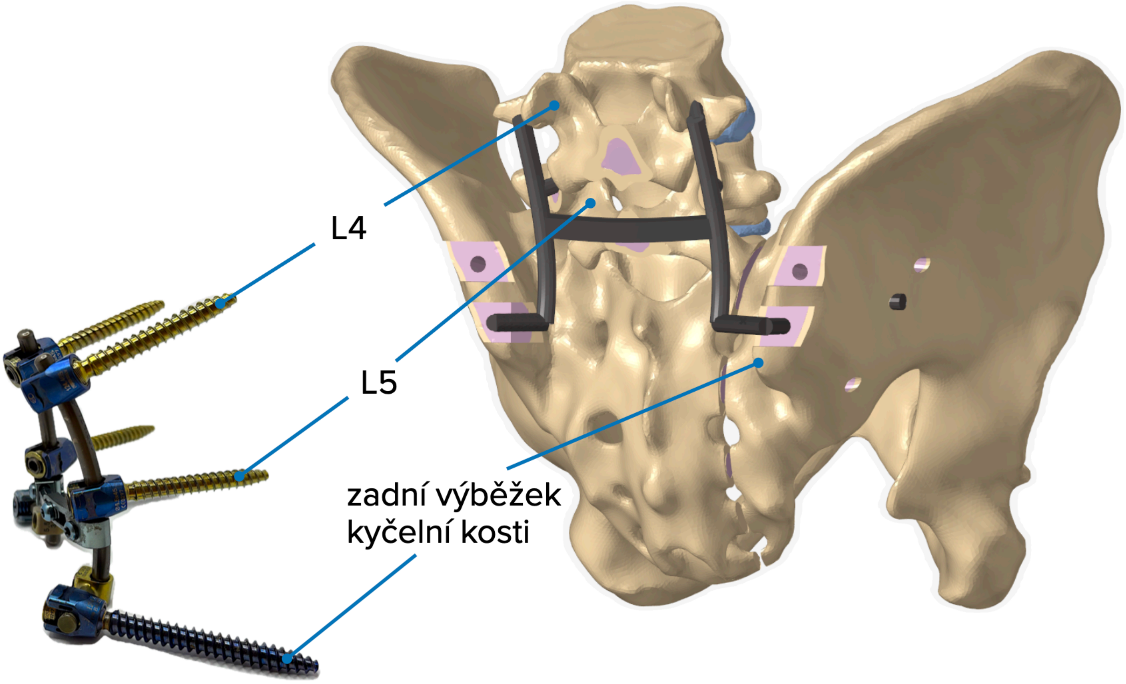
Ren, H., Luo, Z., Fan, H., Yuan, X., Wang, G., and Zhang, L., 2025. G^3CN: Gaussian Topology Refinement Gated Graph Convolutional Network for Skeleton-Based Action Recognition. *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 9661-9668. Available from: https://doi.org/10.1109/IROS60139.2025.11246487

Modelování osteosyntézy zlomeniny křížové kosti pomocí spinopelvických fixací

Anna Karbanová[1], Lukáš Mařík[2], Libor Lobovský[3], Tomáš Mandys[4],
Martin Salášek[5], Jiří Křen[6]

1. Úvod

Tato studie se zaměřuje na osteosyntézu nestabilních zlomenin zadního segmentu lidské pánve pomocí spinopelvických fixačních technik. Tyto mini-invazivní techniky používají ke stabilizaci pánevního pletence kovovou konstrukci, která se zavádí do kyčelních kostí a do oblasti bederní páteře (obr. 1). V rámci této studie je vyhodnocována kvalita stabilizace zlomeniny křížové kosti na základě posuvů zlomených částí kosti při fyziologickém zatížení.



Obrázek 1: Spinopelvická fixace (vlevo) a počítačový model pánve, 4. a 5. bederního obratle včetně meziobratlových plotének a zjednodušené spinopelvické fixace (vpravo)

---

[1] studentka navazujícího studijního programu Aplikovaná mechanika, specializace Biomechanika a mechanika moderních materiálů, e-mail: karbanov@students.zcu.cz

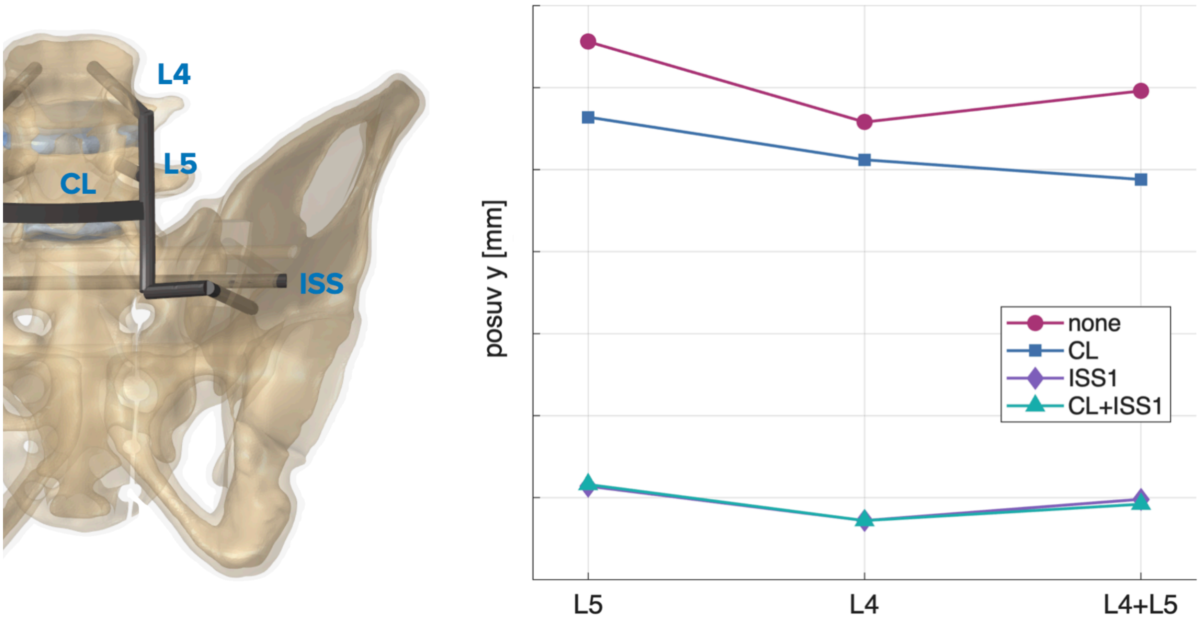
[2] student navazujícího studijního programu Aplikovaná mechanika, specializace Biomechanika a mechanika moderních materiálů, e-mail: mariklu@students.zcu.cz

[3] zaměstnanec KME, e-mail: lobo@fav.zcu.cz

[4] zaměstnanec KME, e-mail: tmandys@fav.zcu.cz

[5] lékař, Klinika ortopedie a traumatologie pohybového ústrojí, Fakultní nemocnice Plzeň

[6] zaměstnanec KME, e-mail: kren@fav.zcu.cz



Obrázek 2: Vyhodnocení posuvů na bázi křížové kosti pro různé kombinace zavedení spinopelvických fixací

2. Výpočtový model

Výpočtový model pánve a posledních dvou bederních obratlů s meziobratlovými ploténkami reálného pacienta vychází z výpočtového modelu pánve, který byl validován na základě experimentů (Karbanová, 2024). Geometrie pánve a posledních dvou bederních obratlů včetně meziobratlových plotének byla zhotovena na základě CT snímků v programu 3DSlicer a následně diskretizována v programu HyperMesh. Numerické simulace pomocí metody konečných prvků byly realizovány v programu Abaqus.

3. Diskuze

Bylo uvažováno zavedení spinopelvické fixace bez crosslink konektoru (NONE) s crosslink konektorem (CL), zavedení spinopelvické fixace v kombinaci s iliosakrálním šroubem (ISS1) a spinopelvická fixace s kombinací crosslink konektoru a iliosakrálního šroubu (CL+ISS1). Na základě výsledků numerických výpočtů bylo prokázáno, že spinopelvická fixace zpevněná iliosakrálním šroubem vykazuje významně větší stabilitu než spinopelvická fixace zpevněná crosslink konektorem (obr. 2). Pro spinopelvickou fixaci zpevněnou kombinací CL+ISS1 nebylo prokázáno výrazné zlepšení.

Poděkování

Příspěvek byl podpořen grantovým projektem Západočeské univerzity v Plzni SGS-2025-015.

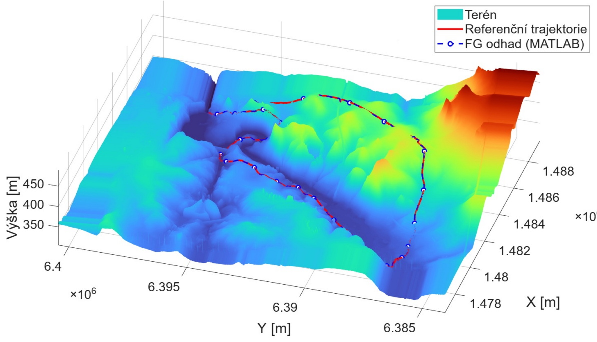
Literatura

Karbanová, A. (2024) *Modelování osteosyntézy tříštivých zlomenin lidské pánve pomocí metody konečných prvků*. Bakalářská práce, Plzeň.

Odhad neznámého stavu pomocí faktorových grafů a bayesovské filtrace

Bc. Filip Kejval[1]

1 Úvod



Obrázek 1: 3D vizualizace odhadu metodou FG.

Spolehlivý odhad vnitřního stavu dynamických stochastických systémů představuje fundamentální problém moderní kybernetiky a navigace. Tradiční rekurzivní algoritmy Bayesovské filtrace (např. EKF, UKF) narážejí u silně nelineárních modelů na své limity, kde jednorázová linearizace vede ke kumulaci chyb a divergenci. Odpovědí je přechod k nelineární optimalizaci pomocí faktorových grafů (Factor Graphs - FG). Tyto pravděpodobnostní modely uchovávají historii stavů i měření a umožňují opakovanou iterativní relinearizaci celého problému Dellaert a Kaess (2017).

2 Metodika

Ve faktorových grafech je odhad stavu formulován jako hledání trajektorie $\boldsymbol{X}$, která maximalizuje aposteriorní pravděpodobnost $p(\boldsymbol{X} \mid \boldsymbol{Z})$, kde $\boldsymbol{Z}$ je množina získaných měření. Za předpokladu gaussovských šumů a nezávislosti měření se problém mění na minimalizaci nelineárních nejmenších čtverců:

$$\hat{\boldsymbol{X}}_{MAP} = \operatorname*{argmin}_{\boldsymbol{X}} \left\{ \|\hat{\boldsymbol{x}}_0 - \boldsymbol{x}_0\|^2_{\boldsymbol{P}_0} + \sum_{k=0}^{K-1} \|f(\boldsymbol{x}_k) - \boldsymbol{x}_{k+1}\|^2_{\boldsymbol{Q}} + \sum_{k=1}^{K} \|h(\boldsymbol{x}_k) - \boldsymbol{z}_k\|^2_{\boldsymbol{R}} \right\}, \qquad (1)$$

kde $f(\cdot)$ a $h(\cdot)$ jsou modely dynamiky a měření, $\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{P}_0$ jsou příslušné kovarianční matice šumů a K představuje počet dat. Faktorový graf tedy řeší úlohu vyhlazování.

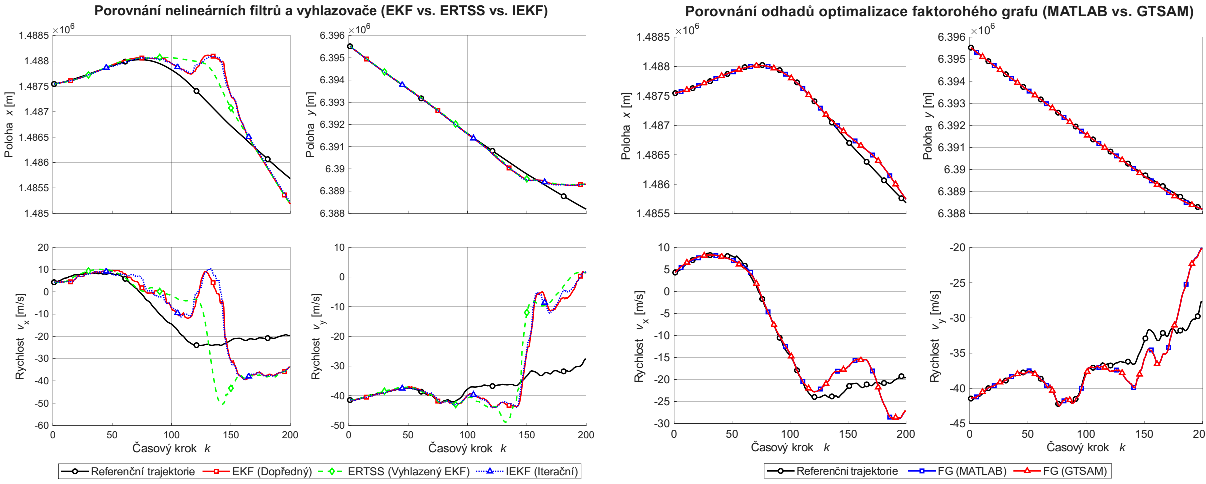
Kvalita estimačních algoritmů byla testována na třech modelech s rostoucí nelinearitou. Nejkomplexnějším experimentem byl silně nekonvexní problém terénní navigace (TAN – Terrain Aided Navigation) Matoušek et al. (2025). Simulován byl také degradovaný model se ztrátou přímého měření směru rychlosti (k dispozici byla pouze výška nad terénem a velikost rychlosti). Hodnocena byla přesnost (RMSE) a statistická konzistence (ANEES).

---

[1] student navazujícího studijního programu Kybernetika a Řídicí technika, specializace Automatické řízení a robotika, e-mail: kejvalf@students.zcu.cz

3 Dosažené výsledky

Kritické rozdíly v robustnosti se projevily u terénní navigace s degradovaným modelem (Obr. 2). Ztráta pozorovatelnosti složek rychlosti vedla na nelineární mapě k rychlé divergenci všech lokálních filtrů (RMSE$_{poloha} = 667.95m$). Oproti tomu obě implementace dávkové optimalizace (MATLAB a C++ GTSAM) trajektorii úspěšně zrekonstruovaly díky opakované relinearizaci celé trajektore (RMSE$_{poloha} = 142.08m$). Monte Carlo analýza metody klouzavého okna na prodloužených trajektoriích navíc odhalila, že pro nekonvexní mapy je optimální volit spíše kratší okna ($N \approx 17$). Dlouhá okna vedla k uvíznutí algoritmu v lokálních minimech, zatímco kratší okna minimalizovala výpočetní čas, dosáhla nižší chyby v porovnání s lokálními filtry a zachovala spolehlivost přes 80 % .



Obrázek 2: Porovnání odhadů trajektorie pro problém terénní navigace. Vlevo: filtry a vyhlazovač (EKF, IEKF, ERTSS). Vpravo: optimalizace faktorových grafů (MATLAB a GTSAM).

4 Závěr

Simulace ukazují, že faktorové grafy představují robustní nástroj pro odhad stavu. U silně nelineárních úloh s omezenou pozorovatelností překonávají tradiční rekurzivní filtry a vyhlazovače. Eliminace fixní linearizace zabraňuje integraci nevratných chyb a umožňuje vyhnout se lokálním extrémům terénní mapy. Vyšší výpočetní nároky lze pro nasazení v reálném čase kompenzovat využitím reprezentace problému ve formě řídkých matic v moderních C++ knihovnách (GTSAM).

Poděkování

Tato diplomová práce byla podpořena projektem Grantové agentury České republiky GA 25-16919J.

Literatura

Dellaert, F., a Kaess, M. (2017) Factor Graphs for Robot Perception. *Foundations and Trends in Robotics*, Volume 6, pp. 1–139.

Matoušek, J., Duník, J., a Straka, O. (2025) Lagrangian Grid-Based Filters With Application to Terrain-Aided Navigation. *IEEE Signal Processing Magazine*, Volume 42(2), pp. 98–104.

Gallium Oxide Thin Films Fabricated by Liquid-Target Reactive DC-Pulsed Magnetron Sputtering

Jan Koloros[1], Petr Novák[2], Sayed Alireza Ataie[3]

Introduction and motivation

Gallium oxide (Ga_2O_3) remains a major focus of research due to its outstanding optoelectronic properties, including an ultra-wide bandgap of approximately 4.8 eV, a high electron saturation velocity, and a high critical breakdown electric field of about 8 MV/cm, making it highly promising for high-power and high-frequency applications. Although high-quality β-Ga_2O_3 films are typically prepared using epitaxial methods such as MBE, MOCVD, or ALD, these techniques suffer from low deposition rates that limit their scalability. In contrast, reactive dc-pulsed magnetron sputtering from a liquid gallium metal target offers high deposition rates, ease of up-scaling, and the ability to tune oxidation conditions via discharge parameters. However, achieving well-ordered crystalline structures with satisfactory electrical properties via sputtering remains a challenge.

Electrical resistivity and crystallinity

Recently, we have demonstrated that optimal electrical performance including a low electrical resistivity of $7 \times 10^3\ \Omega\cdot\text{cm}$ can be achieved for films deposited on sapphire substrates at temperatures in the immediate vicinity of a pronounced crystallization-driven structural transition data are shown in Fig. 1.

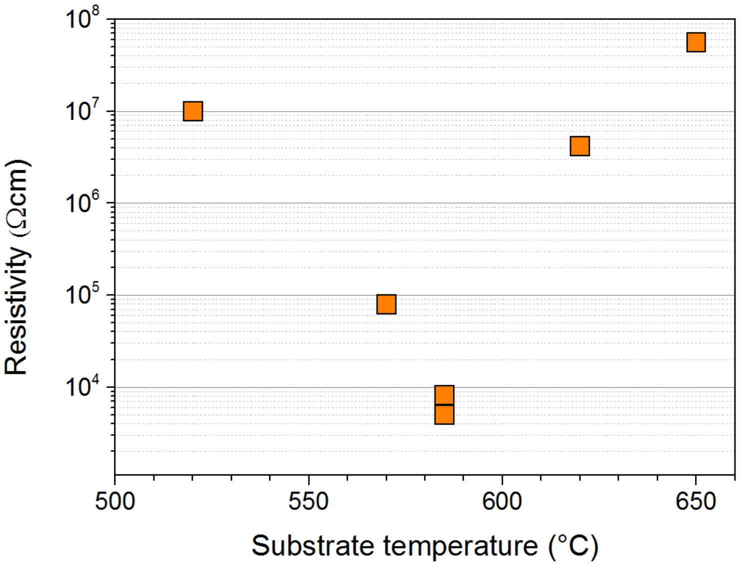


Figure 1: Electrical resistivity

---

[1] student of the doctoral degree program Plasma Physics and Physics of Thin Films, field of study Physics, e-mail: koloros@fav.zcu.cz

[2] scientist at NTC, field of science thin-film materials and magnetron sputtering, e-mail: petrnov@ntc.zcu.cz

[3] scientist at NTIS, field of science thin-film materials and magnetron sputtering, e-mail: sataie@fav.zcu.cz

Within this narrow deposition window, the films remain compact despite incomplete crystallographic ordering, resulting in significantly lower electrical resistivity than that of highly crystallized films. At higher deposition temperatures, further improvements in crystallinity are accompanied by microstructural degradation (ranging from compact solid thin films to wire-like microstructures), which leads to a substantial increase in electrical resistance more details are show in Fig. 2.

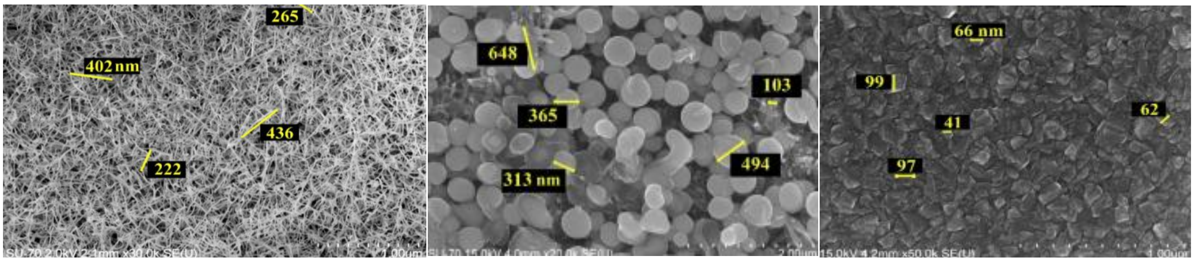


Figure 2: Different microstructural degradation

Futher research and summary

To further enhance the film properties, our current work focuses on optimizing deposition parameters (such as oxygen/argon partial pressures, substrate temperature, and pulse-averaged target power density) in the immediate vicinity of this crystallization threshold, while simultaneously utilizing post-deposition annealing to explore these structural phase transitions. Particular attention is devoted to the temperature region close to the recrystallization threshold, where rapid structural evolution occurs. By systematically examining the changes induced by both process conditions and subsequent annealing using X-ray diffraction (XRD) and scanning electron microscopy (SEM), we aim to clarify the relationship between crystallization, microstructural stability, and the resulting electrical behavior of the films.

These insights provide a practical framework for optimizing thermal treatment conditions to achieve electrical performance comparable to epitaxially deposited Ga_2O_3 films while preserving the scalability advantages of magnetron sputtering for future power electronic applications.

Acknowledgement

This work was supported by the project QM4ST, funded as project No.CZ.02.01.01/00/22_008/0004572 by P JAC, call Excellent Research.

References

Novák, P., Koloros, J., Haviar, S., Rezek, J., a Baroch, P. (2026) Crystalline β-Ga_2O_3 Thin Films Deposited via Reactive Magnetron Sputtering of a Liquid Ga Target. Journal of Vacuum Science & Technology A, Volume 44, Article 033413.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Rigorous Functional Derivatives and Multi-Object Tracking

Jan Krejčí[1]

1 Introduction

Multi-object tracking (MOT) refers to estimating the number and positions of multiple moving objects based on noisy measurements. It is an enabling technology in fields such as public surveillance or autonomous driving. Functional derivatives in MOT are used to derive *elegantly* the analytic expressions that govern how the received measurements are to be ultimately processed during estimation. The use of functional derivatives in MOT is, however, purely formal and heuristic. This contribution reveals how to address this lack of mathematical rigor; see Krejčí et al. (2025) for details.

2 Functional Derivatives in MOT

Let $X = \{\mathbf{x}^1, \ldots, \mathbf{x}^n\} \subset \mathcal{X}$ denote the set of object states (e.g., positions) with $\mathcal{X}$ being the state space. To model both $n \geq 0$ and the individual elements as random, a joint density of the form $p(X) = p(\{\mathbf{x}^1, \ldots, \mathbf{x}^n\})$, $n \geq 0$ is used. Combining $p(X)$ with measurements Z via the Bayes rule results into the posterior $p(X|Z)$, which is the sought entity. The direct derivation of $p(X|Z)$, however, is hindered by the presence of complex expressions resembling convolutions. Similarly to the Laplace transform, the *probability generating functional* (PGFL) defined as

$$G[h] = \sum_{n=1}^{+\infty} \frac{1}{n!} \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} \left(\prod_{i=1}^{n} h(\mathbf{x}^i) \right) p(\{\mathbf{x}^1, \ldots, \mathbf{x}^n\}) \mathrm{d}\mathbf{x}^1 \cdots \mathrm{d}\mathbf{x}^n \,, \tag{1}$$

where h is a function, translates such convolutions into simple products. Once the posterior PGFL $G[h|Z]$ is found, the last step to find $p(X|Z)$ is to take the functional derivative

$$p(X|Z) = \left[\frac{\delta}{\delta X} G[h|X] \right]_{h \equiv 1} . \tag{2}$$

The first-order functional derivative (i.e., with $|X| = 1$) is *formally* defined as a directional derivative in the direction of the Dirac delta "function"

$$\frac{\delta}{\delta\{\mathbf{x}\}} G[h] \triangleq \lim_{\epsilon \to 0} \frac{G[h + \epsilon \delta_{\mathbf{x}}] - G[h]}{\epsilon} \,, \tag{3}$$

and the higher order derivatives (with $|X| > 1$ in "$\frac{\delta}{\delta X}$") are defined iteratively.

---

[1] student of the doctoral degree program Applied Sciences, field of study Cybernetics, e-mail: jkrejci@students.zcu.cz. This research was funded by the project SGS-2025-020.

3 The Rigor Gap

In contemporary MOT, the state space $\mathcal{X}$ is often assumed to be Euclidean while the set of admissible densities $p(X)$ is seldom specified in detail. It should be noted that objects may move on a manifold, such as on a globe, or objects that live in state spaces of different dimensions may coexist, in which cases $\mathcal{X}$ is not Euclidean. Furthermore, the PGFL-based simplification of $P(X|Z)$ derivation is enabled by the availability of *formal* differentiation rules, such as the chain and product rules, see Mahler (2014). It follows from functional analysis, however, that the *safe* use of such rules is subject to specific circumstances, that are clearly missing when the heuristic definition (3) is considered.

The above objections clearly undermine the mathematical rigor in MOT, which may be especially problematic in safety-critical applications such as autonomous driving.

4 The Solution: Utilize Functional Analysis Tools

Krejčí et al. (2025) suggested to fill the rigor gap by utilizing mathematical machinery of measures defined over locally compact and Hausdorff (lcH) topological spaces, which is a well-established theory (Folland , 1999). In particular, assuming $\mathcal{X}$ is any second countable lcH space, the PGFL was redefined as

$$G[\eta] = \sum_{n=1}^{+\infty} \frac{1}{n!} \int_{\mathcal{X}} \cdots \int_{\mathcal{X}} p(\{\mathbf{x}^1, \ldots, \mathbf{x}^n\}) \eta(\mathrm{d}\mathbf{x}^1) \cdots \eta(\mathrm{d}\mathbf{x}^n)\,, \tag{4}$$

so it accommodates any Radon measure η on $\mathcal{X}$ on its input, instead of a “function” h. The set of admissible densities $p(X)$ was clearly specified to be the space of continuous functions vanishing at infinity. Interpreting the Dirac delta $\delta_{\mathbf{x}}$ as a *measure* and showing that it is Radon, it can be legitimately substituted into the functional derivative definition (3); formally replacing h with η. Moreover, the proposed definition was shown to accommodate the safe use of the mentioned differentiation rules by proving it is Fréchet.

5 Conclusion

The proposed definition fills the rigor gap by a large margin, supporting considerably more general spaces $\mathcal{X}$ and admissible densities $p(X)$ than those encountered in the MOT literature. Moreover, the proposed definition is the fist one to provably support the safe use of chain and product differentiation rules used in everyday MOT. Considering the proposed definition thus makes it easier to verify the assumptions under which PGFL-based MOT methods are guaranteed to function as intended.

References

Krejčí, J., Straka, O., Girg, P., and Benedikt, J. (2025) *Revisiting functional derivatives in multi-object tracking.* Available from: `https://arxiv.org/abs/2508.12982` [Accessed 17th April 2026]

Mahler, R. P. S. (2014) *Advances in Statistical Multisource-Multitarget Information Fusion.* Artech House

Folland, G. B. (1999) *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. Wiley

AI-Guided Rewiring of Gene Regulatory Networks in Living Organisms

Lukas Kuhajda[1], Tomas Honzik[2], Jan Svec[3], Daniel Georgiev[4]

1 Introduction

Gene regulation through promoter engineering is a cornerstone of synthetic biology, enabling precise control over transcriptional networks. However, traditional trial-and-error experimental approaches remain highly labor-intensive. Native promoters exhibit intricate, context-dependent behavior heavily influenced by their surrounding sequences and environmental conditions. Rewiring these complex networks by modifying regulatory elements can enable entirely new cellular behaviors. Unfortunately, current synthetic promoter libraries often lack this native context-dependence, and existing computational tools fail to predict functional recombinations between promoters and transcription factor binding sites (TFBS). While artificial neural networks (ANNs) have successfully improved regulatory element predictions, frameworks capable of guiding precise promoter-TFBS integration are still lacking. Our project addresses this gap by developing a two-stage ANN system that identifies optimal TFBS insertion sites and determines the exact promoter rewriting length required to preserve natural sequence syntax.

2 Methods

We extracted promoter sequences from 25 Saccharomycotina species to compile an evolutionarily diverse training and validation dataset of 133,889 sequences, while reserving 6,011 native Saccharomyces cerevisiae promoters strictly for model testing. The developed framework leverages a specialized two-stage ANN architecture. The first stage consists of a self-supervised model called ”Place-Back,” which processes a query sequence and an artificially disrupted base-promoter to output base-pair level insertion probabilities. Two variants of the Place-Back model handle Model-Specific Rewriting Lengths (MSRL) of either 5 bp (simulating exact TFBS insertion) or 40 bp (simulating larger-scale sequence replacements). The second stage, called ”Determiner,” uses an ensemble of 10 diverse CNN or LSTM architectures to resolve prediction inconsistencies, classify the ideal insertion region, and determine the necessary rewrite length. To guarantee experimental reliability, the final system recommendation relies on a multi-model consensus requiring a minimum agreement of 7 out of 10 Determiner models.

3 Results

The multi-model framework achieved an average precision of 48% on held-out validation data, demonstrating massive enrichment over a random baseline insertion precision of

---

[1] PhD student of Applied Sciences and Informatics, field of study Cybernetics, specialization Artificial Intelligence for Synthetic Biology, e-mail: kuhajdal@fav.zcu.cz

[2] University of West Bohemia, Faculty of Applied Sciences, email:thonzik@fav.zcu.cz

[3] University of West Bohemia, Faculty of Applied Sciences, email:honzas@fav.zcu.cz

[4] University of West Bohemia, Faculty of Applied Sciences, email:georgiev@fav.zcu.cz

10.95% for MSRL 40 and 1.26% for MSRL 5. We deployed this system to virtually screen all 6,011 native S. cerevisiae promoters for compatibility with the TetR TFBS (tetO), generating a ranked library that recommended 3,712 high-confidence promoter-TFBS combinations. For in vivo validation, four untested wild-type constitutive promoters of varying expression strengths (pHHF2, pPAB1, pPOP6, pREV1) were engineered using model predictions. Without any prior experimental characterization or manual tuning, the model-designed promoters achieved strong conditional repression rates up to 98.4% (pPOP6_tetO). To demonstrate network rewiring, we introduced a glucose-sensitive Mig1 TFBS (mig1o) into the native promoter of the essential gene PCF11. Luciferase and genomic growth assays verified successful nutrient-dependent regulation, yielding a 57% repression rate in high-glucose media and a 10% lower maximum cell density without causing lethality.

4 Conclusion

The developed two-stage ANN system demonstrates that self-supervised models can successfully capture implicit structural constraints within regulatory DNA without relying on labeled functional datasets. By accurately pinpointing architectural sweet spots, this method streamlines synthetic promoter construction and provides a scalable strategy for reprogramming eukaryotic transcriptional logic. Future work will scale this tool from S. cerevisiae to more complex eukaryotic organisms—including plants and vertebrates—and explicitly train future iterations to handle split-site architectures for enhanced cooperative binding control.

Acknowledgement

This work was supported by the University of West Bohemia grant, project No. SGS-2025-020.

References

Kuhajda, L., Honzik, T., Svec, J., Georgiev, D. (2026) Context-aware synthetic promoter design using neural networks enables rewiring of eukaryotic transcriptional networks. *npj Systems Biology and Applications*.

Advanced Hydrogen Gas Sensing with Porous WO_3 Thin Films Prepared via Glancing Angle Deposition

Akash Kumar[1], Nirmal Kumar[2], Stanislav Haviar[3]

1 Introduction

The emerging hydrogen industry is stimulating efforts in developing new materials for various purposes, including the quest for efficient, sustainable, and low-power hydrogen detectors. Many such devices rely on metal oxide semiconductor materials, which are easily integrable into devices and relatively cheap but suffer from some challenges, such as low sensitivity and selectivity. This study explores the possibility of exploiting a Glancing Angle (GLAD) sputter deposition of WO_3films, targeting the enhancement of active surface area and, therefore, sensor sensitivity improvements.

2 Method

WO_3 films were deposited using conventional reactive DC magnetron sputtering, employing circular tungsten targets in a mixture of argon and oxygen. Films were post-annealed prior to sensing characterization. The Glancing Angle Deposition (GLAD) technique was employed to induce a characteristic columnar nanostructure, thereby increasing the films' porosity and so leading to a desired increase in active surface area. Multiple parameters were tuned to enhance the sensing response, including the angle of deposition (80°, 85°, 88°), thickness (50–300 nm), and reactive sputtering parameters. Quartz and silicon substrates were used for deposition; sensing was evaluated on quartz, and other measurements were performed on silicon.

3 Result

The synthesized films were characterized by SEM and XRD. SEM micrographs for various film thicknesses deposited at an 85° angle are shown in the figure. Micrograph and sensing response measurements revealed an interesting fact: that neither the surface roughness nor the surface area improves the response to the sensing gas monotonically. In the presented paper, we discuss the geometrical reasons as well as the synthesis parameters that influence the sensing characteristic.

4 Conclusion

GLAD-deposited WO_3 films show a stronger sensing response than non-GLAD films Fig. 1, This likely arises from their larger surface area and anisotropy in resistance, both of which enhance the signal. GLAD films also operate at lower temperatures for H_2 detection

.

---

[1] Student of the Doctoral Degree program Applied Sciences, field of study Physics, e-mail: Akash@fav.zcu.cz

[2] Post Doc., Applied Sciences, field of physics, e-mail: kumarn@fav.zcu.cz

[3] Ass. Prof., Applied Sciences, field of Physics , e-mail: haviar@fav.zcu.cz

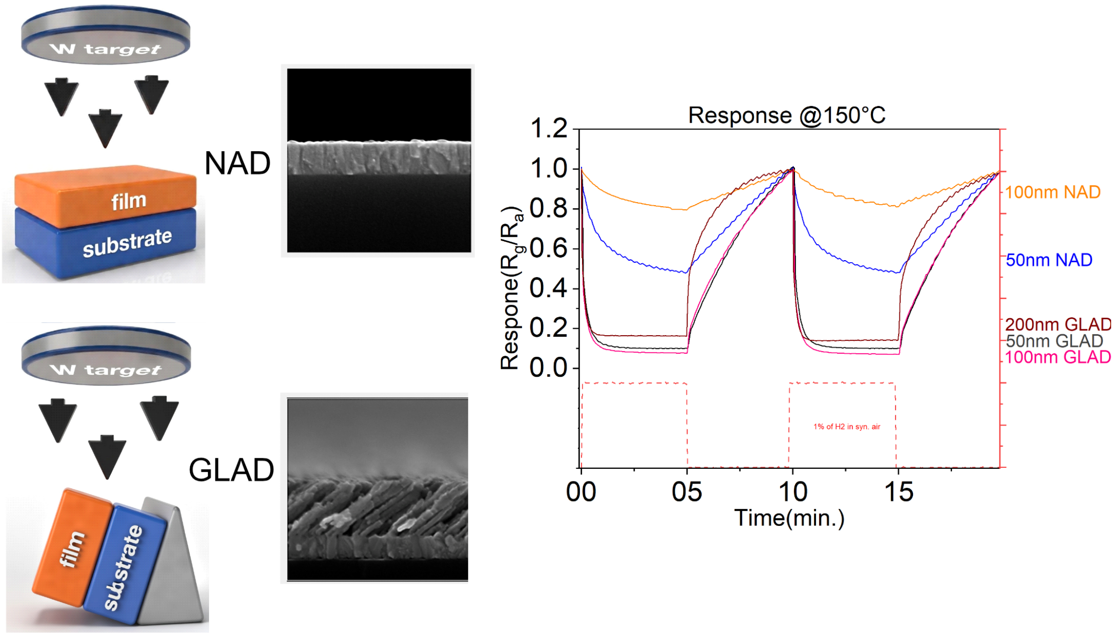


Figure 1: Illustrative Comparison of NAD (Normal Angle Deposition) and GLAD (Glancing Angle Deposition) Techniques

Acknowledgement

This work was supported by the project Quantum materials for applications in sustainable technologies (QM4ST), funded as project No. CZ.02.01.01/00/22_008 /0004572 by Programme Johannes Amos Comenius - Excellent Research. The authors acknowledge professor Jiří Houška for his valuable insights into the growth process discussion, professor Jiří Čapek for his necessary help with the deposition system, and dr. Radomír Čerstvý for XRD measurements and analyses. All acknowledged scientists are from the University of West Bohemia in Pilsen.

References

[1] N. Kumar, S. Haviar, P. Zeman, (2021) Three-layer PdO/CuWO4/CuO system for hydrogen gas sensing with reduced humidity interference, Nanomaterials 11 doi:10.3390/nano11123456

[2] S. Haviar, B. Prifling, T. Kozak, K. Shaji, T. Kosutova, v. Kos, V. Schmidt, J. ˇCapek, (2025) Analysis and 3 D modelling of percolated conductive networks743 in nanoparticle-based thin films, Applied Surface Science Advances 25

[3] N. Kumar, S. Haviar, J. Rezek, P. Baroch, P. Zeman, (2020) Tuning stoichiometry and structure of $Pd-WO_{3-x}$ thin films for hydrogen gas sensing by high-power impulse magnetron sputtering, Materials 13, 1–12.

[4] N. Kumar, A. Kumar, J. ˇCapek, E. Comini, S. Haviar, (2025) $WO_3/CuWO_4$ nanocomposite thin films for humidity resilient acetone gas sensing, Applied Surface Science Advances 30.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Wave-based control of non-uniform systems

Martin Langmajer[1]

1 Introduction

Wave-based control (WBC) is a promising method for position control of flexible systems. Using a single actuator and a single feedback, it performs two seemingly contradictory tasks: rest-to-rest position control, and vibration suppression. Unlike open-loop control, which uses input shaping filters, wave-based control includes feedback. This allows it to react to errors and imperfections while maintaining the same control speed as input shaping filters.

The main principles of wave based control are described in O'Connor (2007). The relationship between WBC and input shaping filters is described in Peled (2013).

Although WBC was originally designed for lumped uniform systems, it is often used successfully for non-uniform systems as well. This has led many authors to make unproven claims about the functionality for any non-uniform systems, e.g. O'Connor (2015). This article presents a rigorously derived refinement of these claims with significant implications.

2 Functionality for non-uniform systems

WBC is primarily used to control systems that can be modelled as mass-spring string (Fig. 1). These include, for example, platoons of autonomous vehicles, but also continua that can be modeled using the FEM method with such a chain.

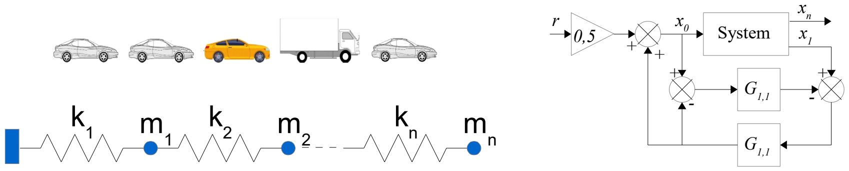


Figure 1: Left: example of a string-like non-uniform system; right: control diagram

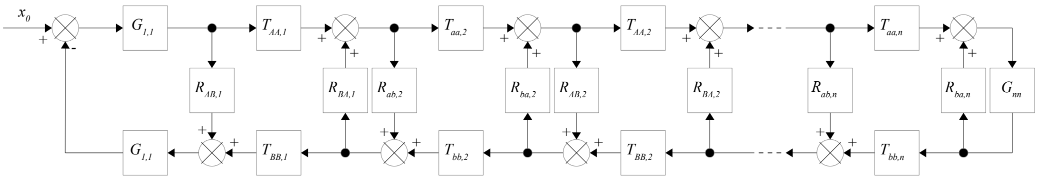


Figure 2: Wave model of a non-uniform system with a free end

Such a system can be modeled using a wave model (Fig. 2), which is an oriented diagram of so-called wave transfer functions (WTF) given by

$$G_{i,j} = 1 + \frac{s^2}{2\omega_{i,j}^2} - \frac{s}{\omega_{i,j}}\sqrt{1 + \frac{s^2}{4\omega_{i,j}^2}}; \quad \omega_{i,j} = \sqrt{\frac{k_i}{m_i}} \tag{1}$$

---

[1] student of the doctoral degree program Applied Sciences, field of study cybernetics, e-mail: gosh@fav.zcu.cz

where s is a complex variable. R is a WTF describing a reflection at the system interface caused by a change in the stiffness k in two consecutive springs or mass m in two consecutive masses. T is WTF for a transmission through this interface. T and R are derived in Martinec (2014).

The control scheme is shown in Fig. 1. The functionality of WBC for non-uniform systems is then determined by the following theorem.

Theorem: *Consider the free-end system in Fig. 1 (left) with a unit DC gain, controlled to a constant value by a WBC of the form in Fig. 1 (right), and consider an ideal actuator with transfer function equal to 1. Next, assume a stable closed loop. WBC then fully suppresses all modes of the controlled system in the frequency range from zero to twice the natural frequency of the first mass-spring pair* $2\omega_{1,1}$*, connected to the actuator, and the controlled system settles to the required value.*

Fig. 3 clearly illustrates the validity of the theorem.

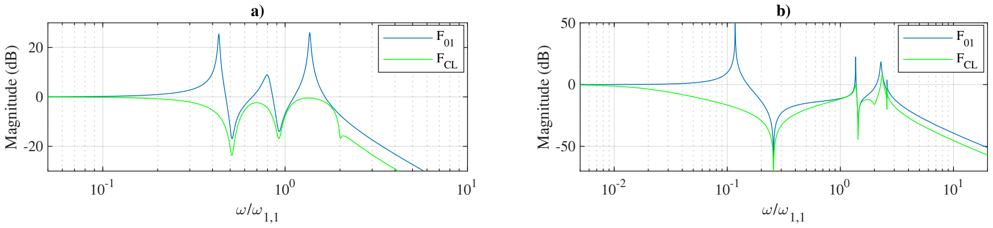


Figure 3: WBC **a)** of a system in which $\omega_{1,1}$ is the highest natural frequency among all mass-spring pairs in the chain, **b)** of a system where a mass-spring pair with a higher natural frequency exists.

3 Conclusion

WBC provides precise and vibration-free control in the range $\langle 0, 2\omega_{1,1}\rangle$. Therefore, if the first mass-spring pair has the highest natural frequency in the system, the control will be stable, accurate, and vibration-free. However, even if this condition is not met, the control can still achieve good results. Modes at higher frequencies are often naturally damped. Furthermore, the dynamics of the fastest mass-spring pair can be artificially created in the actuator control scheme, thereby ensuring functionality across the entire required frequency band.

Acknowledgement

This work was supported by the UWB grant, project No. SGS-2025-020. . . .

References

O'Connor, W.J. (2007) Wave-based analysis and control of lump-modeled flexible robots. *IEEE Transactions on Robotics*, Volume 23(2), pp. 342–352.

Peled, I., O'Connor, W. J., Halevi, Y. (2013) On the relationship between wave based control, absolute vibration suppression and input shaping. *Mechanical Systems and Signal Processing*, Volume 39, pp. 80–90.

O'Connor, W. J., Habibi, H. (2015) Wave-based control of under-actuated flexible structures with strong external disturbing forces. *International Journal of Control*, Volume 88(9), pp. 1818-1829.

Martinec, D., Herman, I., Šebek, M. (2014) The transfer-function approach to travelling waves in path graphs. *ArXiv*, url: https://api.semanticscholar.org/CorpusID:5719109.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Toolkit for Processing and Analysis of Image Description Representation and Semantic Parsing Outputs

Tomáš Lebeda[1]

1 Introduction

Semantic parsing focuses on transforming natural language descriptions into structured semantic representations linked to concrete entities and relations within a reference scene, such as objects in an image. These representations can subsequently be inspected, transformed, evaluated, or processed algorithmically in downstream tasks.

This paper presents the **ref-desc-toolkit**, a modular collection of open-source command-line utilities designed to support workflows involving machine-readable reference scene descriptions and semantic parser outputs. The toolkit was originally developed for a research project focused on semantic analysis of image descriptions for detection of cognitive disorders, where it is used together with the SPGF parser for extraction and evaluation of semantic content derived from speech transcriptions. At the same time, the toolkit was designed as a reusable and extensible infrastructure applicable to other semantic parsing tasks as well.

The primary objective of the toolkit is to enable practical work with semantic representations by providing utilities for extraction, inspection, transformation, evaluation, and visualization of semantic data. By combining these capabilities within a unified modular architecture, the toolkit simplifies development and experimental evaluation of semantic parsing systems.

2 Toolkit Architecture

The toolkit is organized as a set of interoperable CLI utilities targeting different stages of the semantic parsing workflow. At its core is the `ref_desc_lib`, a shared Rust library containing common data structures, parsing logic, serialization support, and utility algorithms. This shared foundation promotes modularity, code reuse, and easier extension of the tooling ecosystem.

2.1 Extractor

The `extractor` enables transformation of free-form tree structures into unified semantic artifacts suitable for further processing and evaluation. The tool processes outputs generated by the SPGF parser (raw parsing trees) and aligns them with a structured reference scene description. Using predefined semantic structures, it identifies and extracts semantic entities corresponding to objects and relations present in the reference scene.

2.2 Inspector

The inspector provides utilities for exploration and debugging of reference scene descriptions. Instead of manually inspecting raw JSON files, users can generate structured summaries containing statistics about objects, attributes, tags, and relations.

---

[1] PhD student of Cybernetics, Artificial Intelligence; email: `lebedat@fav.zcu.cz`

In addition, the tool supports visualization of annotated objects directly on the associated image using bounding boxes and labels. Flexible filtering mechanisms based on object names or semantic tags enable focused analysis of selected subsets of scene entities.

These capabilities are particularly useful during dataset inspection, annotation validation, and parser debugging.

2.3 Evaluator

The evaluator performs quantitative comparison between extracted semantic entities and reference scene descriptions. Given predicted entities, reference annotations, and a predefined loss table, the tool computes a numerical feature vector describing different categories of semantic errors.

The resulting outputs can be exported as CSV files suitable for downstream statistical analysis, benchmarking, or machine learning experiments. This allows reproducible evaluation of grounded semantic parsing systems under consistent scoring rules.

2.4 Visualizer

The visualizer generates interactive browser-based visualizations of parser outputs and derivation structures. Generated outputs consist of HTML, CSS, and text files that display parsing results in a tabular form, where rows correspond to grammar rules and columns correspond to tokens.

To improve readability, the tool supports multiple filtering modes, including displaying only selected rules, limiting derivation depth, or filtering rules using regular expressions. A lightweight console-oriented mode is also available for rapid inspection.

These visualizations simplify qualitative analysis of parser behavior and help identify grammar coverage issues or unexpected derivation structures.

3 Usage in Research Workflows

The toolkit has been developed as supporting infrastructure for research on semantic analysis of image descriptions such as Lebeda et al. (2026) and Šmídl et al. (2026). In the associated workflow, speech transcriptions produced by automatic speech recognition are processed by the SPGF parser, after which the toolkit is used for semantic extraction, normalization, inspection, and evaluation against reference scene descriptions.

The modular architecture enables individual components to be used independently or combined into larger processing pipelines. Particular emphasis was placed on reproducibility, inspectability of intermediate results, and ease of experimentation during iterative parser development.

The toolkit has already been employed in previously published experiments, where extracted semantic information derived from image descriptions was used for subsequent spatial and machine learning analyses.

Acknowledgment

This work was supported by the grant of University of West Bohemia, project No. SGS-2025–011.

References

Lebeda, T., Zajícová, L., Švec, J., Šmídl, L., 2026. Automatic Cognitive Disorder Detection Through Semantic Analysis of Verbal Image Descriptions, in: Text, Speech, And Dialogue. Springer Nature Switzerland, Cham, pp. 177–189.

Šmídl, L., Polák, F., Zajícová, L., Lebeda, T., Švec, J., Tupý, J., Bulín, M., Bartoš, A., 2026. The DigiDiaDem Speech-Cognitive Dataset: Initial Experiments on Detecting Cognitive Impairments From Speech. IEEE Access 14, 22232–22251.. https://doi.org/10.1109/ACCESS.2026.3662045

FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

STUDENTSKÁ
VĚDECKÁ KONFERENCE 2026

Převod stylu řeči do domény vlakových hlášení pomocí modelu StyleTTS2

Karel Müller[1]

Abstrakt. Článek prezentuje přenos prozodických vlastností vlakových hlášení do syntetické řeči při zachování hlasové identity. Jako výchozí model sloužil StyleTTS2 s oddělenou akustickou a prozodickou složkou stylu. Navrženy byly dvě metody přenosu stylu: (1) kombinaci dvou modelů – modelu řečníka a modelu adaptovaného na vlaková hlášení, (2) doladění předtrénovaného modelu s externími akustickými vektory řečníka (při tréninku byla zmražena hlasová složka). Obě metody byly vyhodnoceny pomocí poslechových testů a objektivních metrik. Výsledky ukázaly, že obě metody významně zvýšily stylovou podobnost oproti původnímu modelu a adaptovaná metoda navíc lépe zachovává hlasovou identitu cílového řečníka.

Klíčová slova: syntéza řeči, převod stylu řeči, StyleTTS2, prozodie, adaptace modelu, zachování identity řečníka, vlaková hlášení.

1 Úvod

Řízení řečového stylu a prozodie při zachování identity mluvčího patří k důležitým problémům syntézy řeči. Model StyleTTS2 reprezentuje styl řeči pomocí oddělených akustické a prozodické složky Li et al. (2023), což usnadňuje manipulaci se stylem. Jako cílová doména práce byla vybrána vlaková nádražní hlášení, která se vyznačují pomalejším tempem, specifickým rytmem a výraznou intonací.

2 Metodika

Pro experimenty byl vytvořen dataset 1107 českých vlakových hlášení (ze simulátoru VISL) s textovými i fonetickými přepisy. Data byla rozdělena na trénovací, validační a testovací množiny. Jako výchozí architekturu byl použit model StyleTTS2 s oddělenou akustickou a prozodickou složkou stylu. Pro generování fonémových embeddingů byla využita česká varianta PL-BERT Vladař et al. (2026).

Porovnány byly dva finální přístupy: (1) kombinaci dvou modelů – modelu cílového řečníka a modelu adaptovaného na doménu vlakových hlášení. Při inferenci se spojovala akustická složka stylu z prvního modelu s prozodickou složkou z druhého. Tento přístup umožňuje převod stylu bez dalšího trénování. (2) modifikovanou adaptaci předtrénovaného modelu. Během tréninku nebyla akustická složka stylu extrahována z nahrávek vlakových hlášení, ale byla dodávána externě pomocí předpočítaných akustických vektorů cílového řečníka. Současně byly vybrané komponenty modelu zmrazeny, aby se minimalizovaly změny hlasové identity a adaptace se soustředila na prozodické vlastnosti. Tímto způsobem se omezilo přebírání hlasových vlastností původního řečníka a umožnila se adaptace prozodických vlastností na cílovou doménu.

---

[1] student navazujícího studijního programu Kybernetika a řídicí technika, obor Umělá inteligence a automatizace, e-mail: kmuller@gapps.zcu.cz

3 Experimenty a výsledky

Vyhodnocení proběhlo pomocí tří poslechových testů a objektivních metrik. V testu podobnosti hlasu byl model s upravenou adaptací statisticky významně preferován před kombinovaným modelem ($z = 3{,}31$, $p < 0{,}001$), což naznačuje lepší zachování identity cílového řečníka. V testu stylu mezi oběma metodami nebyl zjištěn významný rozdíl ($z = -1{,}79$, $p = 0{,}073$); obě metody tedy dosáhly podobných výsledků ve stylové podobnosti. V testu stylu vůči referenčním hlášením byl adaptovaný model jednoznačně lepší než původní předtrénovaný model ($z = -20{,}13$, $p < 0{,}001$).

Objektivní metriky (viz Tab. 1) ukázaly výrazné zlepšení obou navržených metod oproti základnímu modelu. Například F0 RMSE kleslo z 62,15 Hz u původního modelu na 41 Hz u adaptovaného a korelace F0 vzrostla z 0,014 na cca 0,61. Adaptovaný model měl navíc nejmenší rozdíl délky projevu vůči referenci a nejvyšší kosinovou podobnost hlasu vůči cílovému řečníkovi (0,792).

Metoda	F0 RMSE [Hz]	F0 korelace	Kosinová podobnost
Kombinovaný model	40,42	0,636	0,781
Adaptovaný model	41,29	0,609	0,792
Původní model	62,15	0,014	—

Tabulka 1: Vybrané objektivní metriky porovnávaných metod.

4 Diskuse a závěr

Výsledky ukazují, že StyleTTS2 je vhodným základem pro řízení prozodických vlastností i v malé stylově specifické doméně vlakových hlášení. Obě finální metody významně zlepšily podobnost generované řeči s cílovým stylem oproti původnímu modelu. Ve stylové podobnosti byly výsledky obou metod obdobné, s rozdíly spíše v detailech. Kombinovaný model dosáhl mírně lepší korelace F0 a umožnil flexibilní přenos stylu bez dalšího trénování, avšak syntéza byla méně stabilní (pravděpodobně kvůli slučování nekompatibilních komponent). Modifikovaná adaptace naopak lépe zachovala hlasovou identitu a poskytla stabilnější výsledky, přestože se i zde vyskytovaly drobné artefakty (např. zhoršená výslovnost některých hlásek). Adaptovaný model vykazoval v tomto ohledu o něco lepší stabilitu. Celkově oba přístupy představují slibné řešení pro řízení stylu řeči ve vysoce specializované doméně.

Literatura

Li, Y.A., Han, C., Raghavan, V., Mischler, G., a Mesgarani, N. (2023) StyleTTS 2: Towards human-level text-to-speech through style diffusion and adversarial training with large speech language models. *Advances in Neural Information Processing Systems*, 36, s. 19594–19621.

Vladař, L., Matoušek, J., Lehečka, J., Řezáčková, M. (2026) Evaluating phoneme-level pretraining in Czech text-to-speech synthesis. In: Ekštein, K., Konopík, M., Pražák, O., Pártl, F. (eds.), *Text, Speech and Dialogue*, s. 158–169. Springer Nature Switzerland, Cham.

Dynamic Network Loading for Short-term Traffic Forecasting

Petr Pernička[1]

1 Introduction

Traffic forecasting ensures that spending and policies are as effective as possible. Traffic models predict the impact of potential projects, such as building city bypass, adding extra lane to a freeway, or introducing a toll on a bridge. Dynamic network loading (DNL) is essentially a macroscopic traffic propagation simulator, simulating how vehicles flow through the traffic network. Metamodels — or models of models — emulate existing models in some way, using lesser computational resources, however often at the cost of lesser accuracy. This work presents a full-fledged DNL and explores its metamodelling using machine learning.

2 Dynamic Network Loading

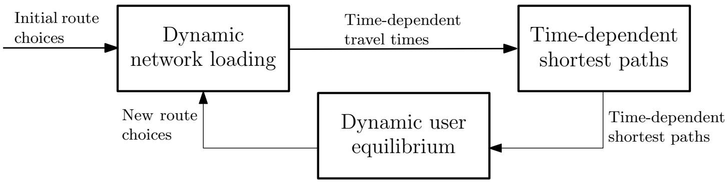


Figure 1: Scheme of dynamic traffic assignment.

As the name suggests, DNL is a dynamic traffic model, meaning it simulates time explicitly. This property makes it more accurate than its static counterparts, but it also makes the theory behind it and its implementation much more difficult. It is also much more computationally demanding, which motivates the usage of metamodels. DNL itself is a submodel of dynamic traffic assignment (fig. 1), which is used for predicting driver behavior in a form of route choices.

DNL divides the model period into short time steps and at each time step, it shifts traffic on roads, followed by its transfer at intersections. The traffic flow at roads is modelled by Lighthill-Whitham-Richards theory (LWR) introduced in Lighill and Whitham (1955) and Richards (1956). The theory represents traffic as continuous flow and operates on three variables: density k, flow q, and speed u. LWR is governed by the following three equations:

$$q = uk, \qquad \frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0, \qquad q = Q(k).$$

One of its numerical solutions, and the one used, is called link transmission model introduced by Yperman et al. (2005). There are plethora of intersection models, each capable of capturing different phenomena such as signal lights or right-of-way. The theory behind them is not as rich but interesting nevertheless. The requirements for realistic intersection model were laid out in Tampère et al. (2011).

---

[1] student of the bachelor master degree program Applied Sciences, field of study Natural Language Processing, e-mail: pperncka@students.zcu.cz

3 Machine Learning Metamodel for DNL

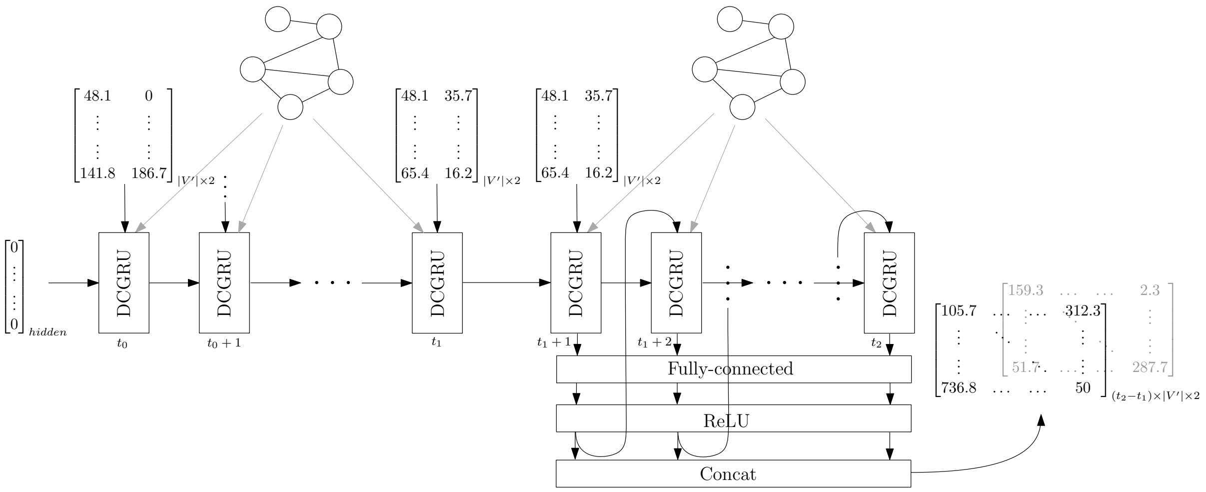


Figure 2: Diffusion convolution recurrent neural network.

Since DNL is very demanding in term of both time and memory, exploring its metamodelling is of interest. Three neural networks were developed and trained on 600 scenarios on two traffic networks. The most promising one is called diffusion convolution recurrent neural network (DCRNN) introduced in Li et al. (2018). It combines a recurrent neural network, specifically GRU, capable of capturing temporal dependencies with diffusion convolution operator, capable of capturing spatial dependencies by propagation information on the traffic network in a manner akin to random processes.

On Chicago network, which has ∼900 nodes and ∼3000 edges, mean absolute error in terms of predicted flows reached 0.27, meaning it is quite accurate. Unfortunately, it is not faster — both DNL and DCRNN produce their results in about 8s on Chicago. However it is expected for the metamodel to beat DNL with larger networks. Developed visualization also revealed that the metamodel is good at predicting zeros, confirmed its inability to conserve vehicles, and showed it has troubles with congested paths.

References

Lighill, M. J., Whitham, G. B. (1955) On Kinematic Waves I. Flood Movement in Long Rivers. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 229, no. 1178, pp. 281–316.

Richards, P. I. (1956) Shock Waves on the Highway. *Operations Research*, vol. 4, no. 1, pp. 42–51.

Yperman, I., Logghe, S., Immers, B. (2005) The link transmission model: An efficient implementation of the kinematic wave theory in traffic networks. *Proceedings of the 10th EWGT Meeting*, vol. 24, pp. 122–127.

Tampère, C. M. J., Corthout, R., Cattrysse, D., Immers, L. H. (2011) A Generic Class of First Order Node Models for Dynamic Macroscopic Simulation of Traffic Flows. *Transportation Research Part B: Methodological*, vol. 45, no. 1, pp. 289–309.

Li, Y., Yu, R., Shahabi, C., Liu, Y. (2018) Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. Available from: `https://arxiv.org/abs/1707.01926`.

Synthetic Tabular Data Generation for Large-Scale Adult Skills Assessments: A Case Study on PIAAC Czech

Duc Thien Pham[1], Roman Mouček[2], František Kalvas[3]

1 Introduction

Synthetic data addresses challenges related to privacy constraints, limited data accessibility, and disclosure risks by generating artificial datasets that preserve the statistical properties and analytical utility of real-world data while reducing the risk of exposing sensitive information. This study presents a comprehensive evaluation of the PIAAC Czech Republic datasets for large-scale adult skills assessment. We benchmark five advanced architectures: Conditional Variational Autoencoder (CVAE), Conditional Tabular Generative Adversarial Network (CTGAN), Transformer-CVAE (T-CVAE), Diffusion, and Transformer-Diffusion (T-Diff), to generate synthetic datasets across varying sample sizes (10,000 and 20,000) and feature densities (100, 300, and 500). Our results identify Diffusion as the most effective approach for high-dimensional tabular synthesis, consistently achieving a strong balance between statistical accuracy and empirical privacy metrics.

2 Method

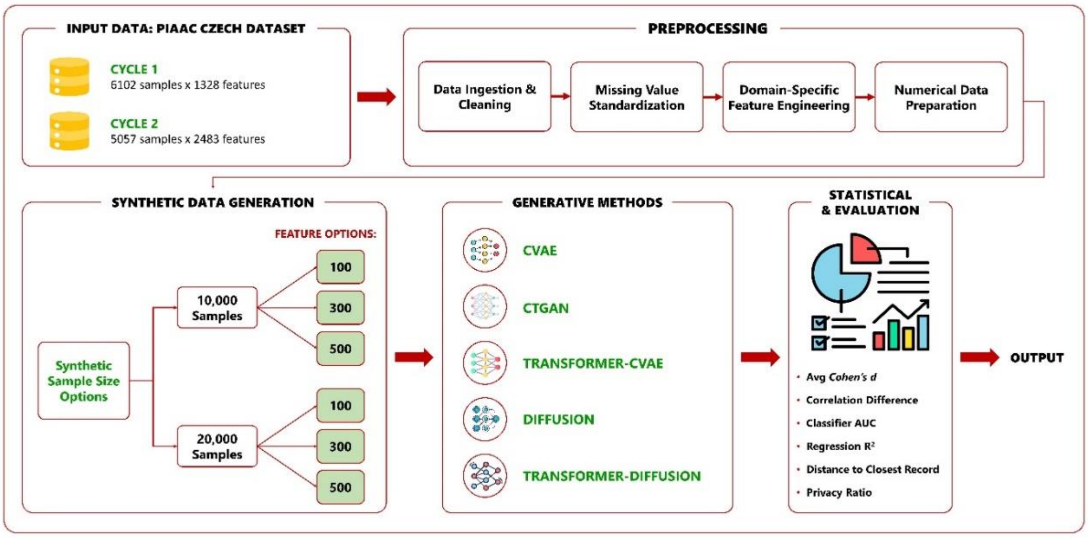


Figure 1: Comprehensive framework for synthetic data generation

3 Results

As shown in Table 1, Diffusion achieves the best overall performance on PIAAC Czech Cycle 1 and 2 by providing the better balance between statistical fidelity, predictive utility,

---

[1] student of the doctoral degree program Computer Science and Engineering, e-mail: ducthien@fav.zcu.cz

[2] Department of Computer Science and Engineering, e-mail: moucek@fav.zcu.cz

[3] Department of Sociology and Social Work, e-mail: kalvas@ff.zcu.cz

privacy, and computational efficiency. Unlike T-Diff, which preserves marginal distributions but fails to maintain multivariate structure (low or negative R^2 and weak AUC), Diffusion consistently preserves correlations and attains the highest predictive scores (up to 0.93 R^2). CVAE and T-CVAE remain competitive but exhibit greater statistical bias, while CTGAN scales poorly in high-dimensional settings.

Sample	Fea-ture	Method	Time (s)	Cycle 1						Cycle 2					
				Avg	Corr	AUC	R^2	DCR	PR	Avg	Corr	AUC	R^2	DCR	PR
10,000	100	CVAE	**93.1**	0.37	0.09	0.92	0.70	5.54	4.01	0.33	0.09	0.97	0.87	5.71	1.53
		CTGAN	3326.0	0.16	0.08	0.58	0.02	6.32	3.75	0.15	0.10	0.58	-0.05	7.01	1.59
		T-CVAE	250.5	0.27	0.11	0.92	0.77	5.92	4.01	0.33	0.13	0.97	0.90	5.45	1.46
		Diffusion	1422.6	0.12	**0.07**	**0.94**	**0.80**	4.63	2.76	0.12	**0.07**	**0.98**	**0.92**	5.71	1.49
		T-Diff	4117.6	**0.11**	0.09	0.53	-0.05	**12.76**	**9.56**	**0.09**	0.09	0.62	0.03	**12.80**	**3.55**
	300	CVAE	**183.8**	1.07	0.09	0.89	0.72	142.42	17.91	0.25	0.08	0.97	0.85	9.70	0.88
		CTGAN	21874.3	0.93	0.06	0.58	0.03	142.86	**22.56**	0.12	0.07	0.55	-0.11	13.84	1.65
		T-CVAE	532.7	1.08	0.12	0.90	0.77	142.47	22.50	0.24	0.12	0.97	0.88	9.65	1.01
		Diffusion	2757.9	0.90	**0.06**	**0.93**	**0.77**	142.96	21.16	0.09	**0.06**	**0.97**	**0.93**	14.77	1.46
		T-Diff	6737.9	**0.88**	0.07	0.49	-0.02	**144.48**	19.28	**0.06**	0.07	0.54	-0.05	**26.47**	**2.79**
	500	CVAE	**189.7**	0.76	0.09	0.92	0.58	143.15	10.16	0.23	0.09	0.97	0.77	13.55	0.92
		CTGAN	33713.8	0.61	0.05	0.45	-0.05	143.68	**11.06**	0.11	0.05	0.44	-0.04	17.44	1.63
		T-CVAE	549.0	0.80	0.15	0.92	0.71	143.26	9.73	0.26	0.12	0.97	0.80	12.57	0.97
		Diffusion	2751.9	0.55	**0.05**	**0.93**	**0.72**	145.20	10.02	0.06	**0.04**	**0.98**	**0.90**	26.97	1.72
		T-Diff	7025.7	**0.53**	0.05	0.52	0.02	**146.90**	11.05	**0.04**	0.04	0.47	0.00	**35.53**	**2.56**
20,000	100	CVAE	**183.6**	0.38	0.09	0.91	0.75	5.53	4.78	0.32	0.09	0.97	0.87	5.72	1.39
		CTGAN	11501.5	0.17	0.08	0.54	-0.14	5.76	4.47	0.15	0.08	0.72	-0.23	7.09	2.15
		T-CVAE	535.3	0.30	0.12	0.92	0.78	5.80	4.48	0.32	0.11	0.97	0.91	5.60	1.37
		Diffusion	3368.9	0.12	**0.07**	**0.94**	**0.79**	4.74	4.55	0.12	**0.07**	**0.98**	**0.91**	5.69	1.57
		T-Diff	8205.5	**0.11**	0.09	0.55	0.00	**12.82**	**8.39**	**0.09**	0.09	0.43	0.01	**12.79**	**3.02**
	300	CVAE	**192.0**	1.06	0.10	0.90	0.70	142.42	22.03	0.24	0.08	0.97	0.85	9.71	0.97
		CTGAN	22796.7	0.93	0.06	0.57	-0.18	142.87	22.99	0.13	0.07	0.58	-0.23	13.81	1.28
		T-CVAE	543.1	1.08	0.12	0.91	0.73	142.41	**23.03**	0.33	0.14	0.97	0.91	9.62	1.01
		Diffusion	2833.1	0.90	**0.06**	**0.93**	**0.79**	142.97	21.33	0.08	**0.06**	**0.98**	**0.92**	14.82	1.67
		T-Diff	6757.7	**0.87**	0.07	0.51	-0.01	**144.48**	22.74	**0.05**	0.07	0.55	-0.04	**26.49**	**2.74**
	500	CVAE	**196.0**	0.75	0.09	0.92	0.66	143.24	10.31	0.27	0.10	0.97	0.81	13.57	0.96
		CTGAN	33408.5	0.59	0.06	0.53	-0.02	143.71	10.67	0.12	0.04	0.45	-0.16	17.09	1.20
		T-CVAE	543.5	0.71	0.12	0.92	0.48	143.29	9.18	0.27	0.12	0.97	0.84	12.80	0.87
		Diffusion	2744.8	0.56	**0.05**	**0.94**	**0.70**	145.22	10.50	0.06	**0.04**	**0.98**	**0.90**	26.97	1.75
		T-Diff	6999.8	**0.53**	0.05	0.52	0.00	**146.93**	**11.36**	**0.04**	0.04	0.43	-0.00	**35.52**	**2.42**

Table 1: Comprehensive evaluation of generative models on PIAAC Czech Cycle 1 and Cycle 2 datasets

4 Conclusion

The evaluation on the high-dimensional PIAAC Czech dataset shows that Diffusion achieves the most consistent overall performance, preserving correlation structure and predictive utility while maintaining favorable empirical privacy characteristics based on DCR and PR metrics. Future work will focus on improving Transformer-based diffusion architectures, evaluating generalizability across other PIAAC datasets, and optimizing DDIM sampling for more efficient high-dimensional data synthesis.

Acknowledgement

This work was supported by the Ministry of Education, Youth and Sports and the Operational Program Johannes Amos Comenius within the project Open Science II (reg. No. CZ.02.01.01/00/24_030/0015041).

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Digitizing the PICNIR Cognitive Test Using Automatic Speech Recognition

Filip Polák[1], Jan Tupý[2]

1 Introduction

With neurodegenerative diseases rising globally (1), early detection through standardized cognitive tests is a critical public health priority. Effective clinical tools like PICNIR (3) and ALBA (2) identify early-stage cognitive deficits. Specifically, PICNIR evaluates semantic and episodic memory by having patients name 20 drawings and recall as many as possible.

However, these traditional paper-based tests require manual administration and manual evaluation, which focus strictly on linguistic correctness. This entirely discards vital acoustic and temporal nuances, such as response latencies, hesitations, and pause patterns, which are nearly impossible for the human ear to measure accurately.

To address these limitations, we developed **DigiDiaDem**, a multimodal Spoken Dialogue System for automated cognitive screening. It not only replicates standard clinical scoring but also extracts a high-dimensional space of acoustic and temporal biomarkers inherently absent in conventional paper-based assessments.

2 Experiments and Results

To evaluate our ASR-driven pipeline, we utilized a clinical dataset collected via the DigiDiaDem application, comprising 356 participants (247 healthy controls and 109 patients with mild cognitive impairment or dementia). We formulated the task as a binary classification problem (Healthy vs. Impaired) and systematically compared three primary feature configurations:

1. **Paper Features (Baseline):** Only the standard metrics manually scored in the traditional test (number of correctly named pictures and correctly recalled pictures).
2. **Digital Expert Features:** The automated ASR-derived replicas of the paper features, combined with the Total Naming Reaction Time.
3. **All Digital Features:** Approximately 350 ASR-derived digital features, including single-word naming efficiency, semantic confusion indices and self-correction tracking.

We evaluated several machine learning models, reporting the best-performing architecture for each set (Table 1).

The results indicate that Digital Expert features yield outcomes highly comparable to the manual paper test. Crucially, when classifiers utilize the “All Digital” feature space, accuracy drastically increases to 97.0%. Feature importance analysis revealed that item-specific reaction

---

[1] student of the doctoral degree program Applied Sciences, field of study Cybernetics, e-mail: polakf@fav.zcu.cz

[2] student of the doctoral degree program Applied Sciences, field of study Cybernetics, e-mail: jtupy@ntis.zcu.cz

Table 1: Classification results comparing feature sets, with and without sociodemographic data (Age, Sex, Education).

Feature Set	**Best Model**	**Accuracy**	**Macro F1**
Without Sociodemographic Data			
Paper Features	Random Forest	0.890	0.860
Digital Expert Features	Random Forest	0.870	0.847
All Digital Features (82 selected)	Gradient Boosting	**0.970**	**0.962**
With Sociodemographic Data			
Paper	Random Forest	0.920	0.900
Digital Expert	Logistic Regression	0.890	0.874
All Digital (8 selected)	Random Forest	**0.970**	**0.962**

times and exact single-word description counts were among the most discriminative variables. Remarkably, the comprehensive digital feature set achieves state-of-the-art performance completely independently of sociodemographic inputs, preserving user privacy.

3 Conclusion

[1] This study demonstrates that a fully automated, speech-driven pipeline can achieve diagnostic accuracy comparable to, and even exceeding, traditional manual evaluations performed by clinical experts. By digitalizing the PICNIR cognitive test and employing a robust, domain-adapted ASR model, we successfully captured both standard linguistic metrics and novel acoustic biomarkers.

Our experimental results confirm that these digital features yield highly accurate classifications of cognitive impairment, achieving a peak accuracy of 97%. Importantly, the system delivers state-of-the-art results without relying on demographic metadata, and dimensional reduction proved the models are resistant to item-level overfitting. The automated approach provides a fast, mathematically consistent, and objective evaluation, effectively removing the subjective variance inherent in manual human grading. Ultimately, this work paves the way for accessible, scalable, and remote cognitive screening, reducing the logistical burden on both patients and healthcare providers.

References

[1] Alzheimer's Association: 2023 Alzheimer's disease facts and figures. Alzheimer's & Dementia **19**(4), 1598–1695 (2023)

[2] Bartoš, A., Diondet, S.: The sensitive Amnesia Light and Brief Assessment (ALBA) is a valid 3-min test of 4 tasks indicative of mild cognitive deficits. Neurología (English Edition) **40**(6), 586–598 (2025)

[3] Bartoš, A., Diondet, S.: Sensitive written hedgehog PICture Naming and Immediate Recall (PICNIR) as a valid and brief test of semantic and short-term episodic memory for very mild cognitive impairment. Journal of Alzheimer's Disease **102**(2), 396–410 (2024)

---

[1] This research was supported by the SGS-2025-011.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

A Software Tool for PID Controller Tuning in Siemens PLC Systems

Lukáš Slavíček[1]

1 Introduction

With an extension of automation in industrial processes, there is an increasing demand for the development of modern industrial automation systems that meet ever higher requirements for real-time system control. This category includes compact controllers, Programmable Logic Controllers (PLCs), Programmable Automation Controllers (PACs), and Distributed Control Systems (DCSs).

Most industrial automation systems integrate built-in PID control function blocks within their development environments, often with built-in support for automatic tuning. These tools are called autotuners. However, the majority of these tools rely solely on heuristic tuning rules, as described in Åström and T. Hägglund (1995), which provide no formal guarantees of system stability or control performance. Typical implementations of such autotuning tools can be found in systems provided by Siemens, as detailed in Siemens Function Manual (2024).

The operation of an autotuner consists of two phases: system identification and controller parameter computation. System identification is based on exciting the system with an appropriate input signal (such as a step, pulse, or sine wave) and extracting characteristic numbers from its response. These values are then used to compute the controller parameters.

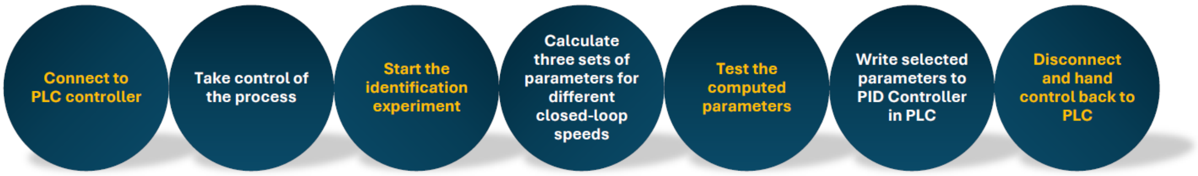


Figure 1: The PID Tuning Procedure of the Tool

2 About the Tool

The tool , whose previous version is described in Slavíček, Balda and Schlegel (2025), is designed to simplify the tuning procedure and improve controller performance for PID_Compact, PID_3Step, and PID_Temp controllers in SIMATIC S7-1200 and S7-1500 industrial automation systems, described in Siemens Function Manual (2024). An industrial controller contains a control scheme consisting of a feedback loop with a PID controller and its associated data structure. The tool's functionality (Fig. 1) is therefore straightforward. Using OPC UA communication, the tool connects to the controller's data structure and temporarily takes over control of the process for the duration of the identification experiment and the subsequent parameter computation.

---

[1] student of the doctoral degree program Cybernetics, field of study Robust controllers with constraind structure, e-mail: slavicel@fav.zcu.cz

For this purpose, the PIDMAX controller described in Function Blocks of REXYGEN (2020) is used to compute three sets of parameters corresponding to different closed-loop speeds and either PI or PID controller structures.

The computed controller parameters can subsequently be tested in the tool on the real system, and after selecting a suitable set, uploaded to the PLC. The tool is then disconnected, and control is handed back to the PLC controller. The PLC controller can operate in both manual and automatic modes.

3 Autotuners Comparison

Figure 2 illustrates the response of closed-loop control systems to a step change in the setpoint followed by a disturbance step for an 8th-order system, whose transfer function is shown in the figure. The results show that, compared to the Siemens PID_Compact controller, PIDMAX achieves improved performance and robustness, including more accurate setpoint tracking and more effective disturbance rejection. With multiple control speed options and direct parameter verification prior to application, the tool enables simpler and more efficient tuning in industrial automation.

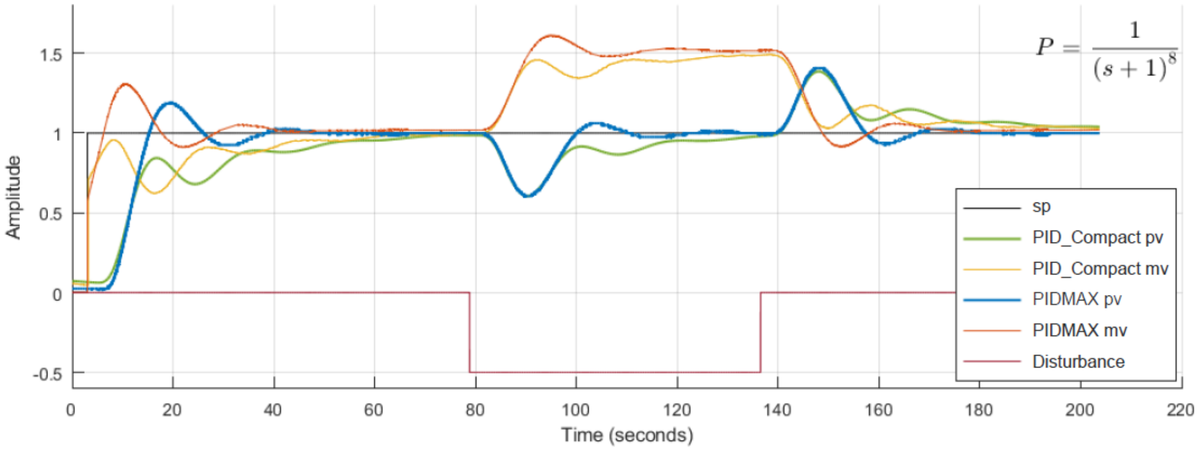


Figure 2: Closed-Loop Performance Comparison

Acknowledgement

This work was supported by project SGS-2025-020.

References

K. J. Åström a T. Hägglund (1995) *PID Controllers: Theory, Design and Tuning*. Research Triangle Park, ISA.

SLAVÍČEK, Lukáš; BALDA, Pavel; SCHLEGEL, Miloš (2025) *RexTuner: A New Software Tool for PID Controller Tuning in Siemens PLC Systems*. In: 25th International Conference on Process Control (PC), 2025. IEEE, pp. 1–6.

Siemens AG (2024) *SIMATIC S7-1200, S7-1500 PID Control, Function Manual*. Entry-ID: 108210036.

REX Controls s.r.o. (2020) *Function Blocks of REXYGEN, Reference Manual*. V2.50.10.

Návrh metody vytváření syntetických dat pro potřeby výzkumu KIR izoforem

David Staníček[*1] , Tomáš Honzík[1] , Kateřina Wolf[1] , Lucie Houdová[1]

1 Úvod

Nepříbuzenská transplantace krvetvorných buněk hraje v současnosti stále širší roli při terapii hematologických onemocnění. I přes neustálý posun ve výběru kritérií za účelem nalezení vhodného dárce je tato procedura spojena s riziem závažných obtíží. Studie od Cibich et al. (2026) uvádí mortalitu přibližně 25 % během 3 let po transplantaci v důsledku následných komplikací.

Tato práce se zabývá návrhem metodiky vytváření syntetických dat reprezentujících izoformy genů ve formě cDNA primárně za účelem testování bioinformatických pipeline určených např. k identifikaci KIR izoforem. Potřeba takovýchto dat je velká, jelikož výzkum KIR genů představuje velmi specifickou oblast, pro kterou v současné době existuje pouze omezené množství veřejně dostupných dat. V rámci práce jsou tyto zdroje prozkoumány v kontextu řešené problematiky. Dále jsou studovány biologické mechanismy vedoucí ke vzniku KIR izoforem, na základě kterých je navržena pipeline tvorby syntetických dat.

2 Biologický kontext

Klíčovou roli v tomto procesu hrají NK buňky, jejichž aktivita je regulována pomocí KIR receptorů. KIR receptory jsou kódovány vysoce polymorfní skupinou genů a tato variabilita sahá od vysokého počtu alel až k úrovni alternativního sestřihu RNA (Hsu et al., 2002). Alternativní sestřih představuje biologický proces vedoucí k různým změnám v exon-intronové struktuře přepisovaného genu, jejichž důsledkem je vznik různých variant výsledného proteinu označovaných jako izoformy (viz obrázek 1). Z klinického hlediska může tato variabilita významně ovlivnit průběh imunitní odpovědi po transplantaci krvetvorných buněk (Houdová et al., 2018).

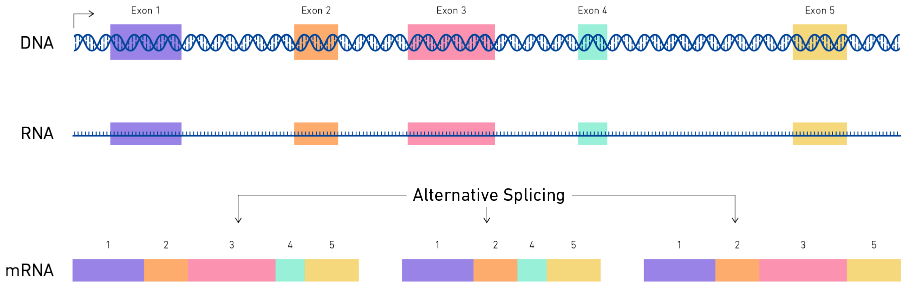
3 Návrh metodiky

Navržená metodika zahrnuje simulaci získávání dat na klinickém pracovišti formou generování syntetických PacBio long-readů, zpracování dat, identifikaci přítomných genů a detekci hranic intronů. K vygenerování izoforem je přistoupeno pomocí skrytého Markovova modelu (HMM), který formou binárních tokenů generuje exon-intronovou strukturu potenciálních izoforem na základě známých sestřihových konfigurací. Součástí práce jsou experimenty za účelem ověření správnosti navrženého přístupu.

---

[*] Student navazujícího studijního programu Kybernetika a řídící technika, specializace Umělá inteligence a automatizace, e-mail: dsta@students.zcu.cz

[1] Fakulta aplikovaných věd, ZČU v Plzni



Obrázek 1: Schéma procesu alternativního sestřihu, převzato (Dornell, 2021)

4 Závěr

Navržené řešení simuluje získávání a zpracování sekvenačních dat od přípravy na klinickém pracovišti až po generování výsledných izoforem pomocí generativního HMM. Navržený přístup představuje částečné zjednodušení biologicky velmi komplexního procesu, nicméně výstupem jsou realistické distribuce známých i nově generovaných izoforem daného genu a vhodným rozšířením metodiky je možné dosáhnout dalšího zpřesnění zejména směrem k vyšší biologické validitě výstupů.

Poděkování

Tato práce byla podpořena z univerzitního projektu SGS-2025-020.

Literatura

Cibich, A., Wechalekar, G., Shanmuganathan, N., Singhal, D., Beligaswatte, A., Wayte, R., Selby, P., Branford, S., Yeung, D., Bardy, P., a Hiwase, D. (2026) Long-Term Survival Gains after Allogeneic Hematopoietic Stem Cell Transplant are Driven by Reductions in Non-Relapse Mortality: A 35-Year Statewide Australian Cohort Study. *Transplantation and Cellular Therapy*. pp. 1–13.

Dornell, J. (2021) Alternative Splicing: Importance and Definition. *Technology Networks*. Dostupné z: `https://www.technologynetworks.com/genomics/articles/alternative-splicing-importance-and-definition-351813` [cit. 2026-05-26].

Houdová, L., Fetter, M., Jindra, P., a Georgiev, D. (2018) Optimal Donor Selection During Verification Process: Which Factors Are Worth Knowing? *Transplantation Proceedings*. 50(10), pp. 3082–3087.

Hsu, C.K., Chida, S., Geraghty, D.E., a Dupont, B. (2002) The killer cell immunoglobulin-like receptor (KIR) genomic region: gene-order, haplotypes and allelic polymorphism. *Immunological Reviews*. 190(1), pp. 40–52.

FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

STUDENTSKÁ
VĚDECKÁ KONFERENCE 2026

Systém automatické eliminace kmitů podvěšené zátěže jeřábu

Daniel Sukovatý[1]

1 Úvod

Manipulace s těžkými břemeny je významnou součástí průmyslových a logistických procesů, u nichž jsou kladeny vysoké požadavky na rychlost, přesnost a bezpečnost. Jedním z nejrozšířenějších zařízení pro tuto činnost jsou jeřáby, které umožňují přesouvání břemen ve více stupních volnosti.

Obr 1. Jeřáb v laboratoři NTISu

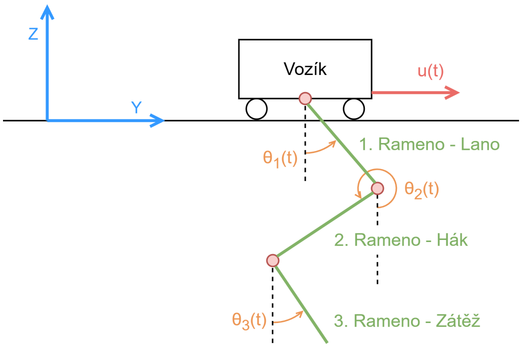
2 Problematika tlumení kmitů

Hlavním problémem jeřábových systémů je vznik nežádoucích kmitů zavěšené zátěže. Ty mohou prodlužovat dobu manipulace, zhoršovat přesnost polohování a v krajním případě ohrozit bezpečnost osob nebo okolních objektů. V praxi jsou kmity často potlačovány manuálně, v současnosti jsou však rozvíjeny také metody automatického tlumení. Ty lze obecně rozdělit na přímovazební metody založené na vhodném tvarování vstupního signálu a zpětnovazební metody využívající měření pohybu zátěže.

Tato práce je zaměřena především na aktivní tlumení kmitů pomocí zpětnovazebního řízení. Pro měření pohybu je využit senzor umístěný na horní části háku, který obsahuje akcelerometry a gyroskopy ve všech třech osách. Pomocí fúze dat v rozšířeném Kalmanově filtru jsou získány veličiny popisující pohyb háku, zejména úhlový náklon a úhlová rychlost.

3 Jeřáb jako dynamický systém

Jeřábový systém se skládá z pohyblivého vozíku, lana, háku a břemene. Vozík lze modelovat jako aperiodický člen s regulací rychlosti, zatímco zbytek systému odpovídá dynamice trojitého kyvadla. Vstupem systému je požadovaná rychlost vozíku a výstupem jsou veličiny popisující pohyb háku. Model trojitého kyvadla obsahuje tři komplexně sdružené dvojice téměř netlumených pólů, které odpovídají třem rezonančním frekvencím. Každá rezonanční frekvence je přitom spojena s jedním konkrétním tvarem kmitu.



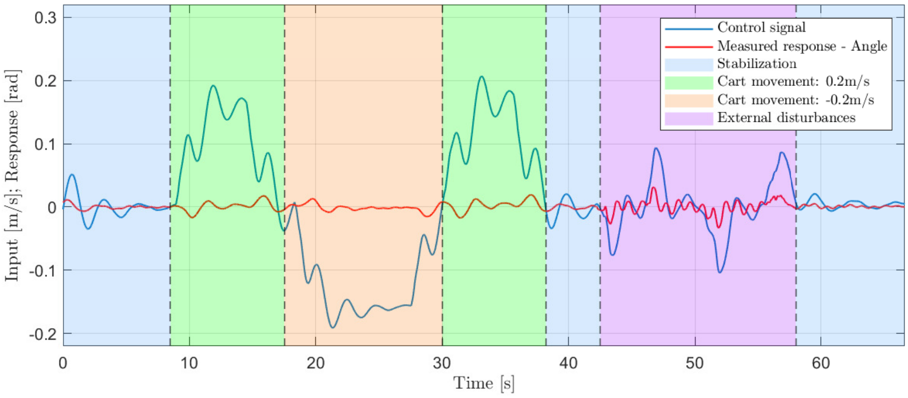
Obr 2. Model trojitého kyvadla s pohyblivým vozíkem

---

[1] student navazujícího studijního programu Kybernetika a řídící technika, specializace Automatické řízení a robotika, a vyzkumný pracovníkk centra NTIS, e-mail: dansuk24@ntis.zcu.cz

4 Experimentální identifikace

Pro odhad rezonančních frekvencí byl navržen algoritmus SCARE (Sequential Crane Active Resonance Estimation). Algoritmus využívá dvoufázový reléový experiment, kterým jsou postupně vybuzeny dominantní módy systému. Z naměřených odezev jsou následně odhadnuty rezonanční frekvence pomocí Hilbertovy transformace, případně v kombinaci s empirickou modální dekompozicí. Na základě těchto odhadů jsou poté navrženy parametry řízení.



Obr 3. Experimentální validace zpětnovazební řízení

5 Adaptace řízení

Při manipulaci s břemenem se často mění délka lana, což vede ke změně dynamiky celého systému. Tuto změnu nelze určit pouze ze znalosti původních rezonančních frekvencí. Proto jsou odhadnuté frekvence nejprve převedeny na neznámé fyzikální parametry zátěže při známých parametrech lana a háku. Při následné změně délky lana je možné z těchto parametrů dopočítat nové rezonanční frekvence a upravit parametry řízení. V tomto případě se tedy nejedná o adaptivní řízení v klasickém smyslu, ale spíše o přístup typu gain-scheduling.

6 Závěr

Navržené identifikační a řídicí algoritmy byly ověřeny experimentálně na reálném jeřábu. Odhad rezonančních frekvencí byl spolehlivý zejména pro první dva módy, zatímco u třetího módu byla pozorována větší odchylka. Při přepočtu na fyzikální parametry zátěže a následné adaptaci řízení může tato odchylka při větších změnách délky lana negativně ovlivnit stabilitu systému. Při menších změnách délky lana však byla stabilita systému zachována a navržené řízení vykazovalo požadovanou účinnost.

Poděkování

Tento příspěvek vznikl za podpory grantového projektu Západočeské univerzity v Plzni, číslo projektu SGS-2025-020.

Literatura

D. Sukovatý, ”Systém automatické eliminace kmitů podvěšené zátěže jeřábu,”diplomová práce, Západočeská univerzita v Plzni, Fakulta aplikovaných věd, Katedra kybernetiky, Plzeň, 2026.

F. R. Kschischang, The Hilbert Transform, The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, 2015.

Introduction to Brain-to-Speech (BTS) Technologies

Lukáš Vladař[1]

1 Introduction

For last decades, researchers around the world have been working on the development of *text-to-speech* (TTS) systems in order to provide an alternative means of communication for people suffering from speech loss. However, these systems are not applicable to completely paralysed patients, who are unable to type the intended utterance on a keyboard. Therefore, the more sophisticated *brain-to-speech* (BTS) technology is necessary to synthesize speech directly based on users' thoughts.

This paper offers a brief introduction to brain activity measurement, it compares basic BCI system categories, and provides an overview of popular BCI models and datasets.

2 Methods for Brain Activity Measurement

There are several ways to measure the brain activity. Some of these methods are *invasive* (i.e., they require the surgical implantation of sensors into the skull), while others are *non-invasive* and thus accessible to anyone. While non-invasive measuring is more user-friendly, the obtained signal is usually burdened with a high level of noise. There are also differences in the temporal and spatial resolution of these approaches (i.e., how accurately the time and location of a particular brain event can be reconstructed from the signal). Additionally, some of these methods require medical supervision or large and expensive equipment, making them unsuitable for non-clinical use. Tab. 1 provides an overall comparison of different brain activity measurements based on the survey of Schultz et al. (2017).

	ECoG	sEEG	Microelectrodes	EEG	MEG	fMRI	fNIRS
Non-Invasive	✗	✗	✗				
Suitable for Non-Clinical Use	✗	✗			✗	✗	
Temporal Resolution	good	good	excellent	good	good	poor	poor
Spatial Resolution	good	good	excellent	poor	good	good	poor
Amount of Noise	low	low	low	high	low	medium	high

Table 1: Comparison of different methods for brain activity measurement.

3 BTS System Categories

BTS systems can be categorized based on different criteria. First of all, these systems differ in output modality – some of them generate audio (i.e., the task of *speech synthesis*) while others output text (i.e., *speech recognition*) and must be followed by a standalone TTS model.

---

[1] student of the doctoral degree program Cybernetics, e-mail: vladarl@fav.zcu.cz

Furthermore, each system is designed for a specific *speech mode*. In terms of applicability, the most useful models are those designed for *imagined speech*, whose input is a brain signal recorded during the mere imagination of the intended utterance. Unfortunately, it is not trivial to train such a model, since it is impossible to record ground-truth audio of imagined speech. Therefore, many researchers focus on the *spoken speech* mode, i.e., recording brain signals while the subject is actually speaking. On the other hand, BTS systems designed for *listened speech* aim to reconstruct the speech that the subject is listening to.

These days, most BTS systems are *user-specific*, but for practical use, a *universal* model that works for anyone without additional training would be very useful. Current BTS models are usually trained using a *small vocabulary* of several words or phrases, however, the ultimate objective is to design a general system working with an *open vocabulary*.

4 Popular BCI Models & Datasets

In recent years, many BCI models have been introduced, some of them have even been published online. For example, Anumanchipalli et al. (2019) created a system synthesizing speech based on ECoG recordings by estimating articulator movements. The NeuroTalk model proposed by Lee et al. (2023) reconstructs speech spectrograms using EEG data. Recently, the first model synthesizing whole sentences from EEG was presented by Rastogi et al. (2025).

Furthermore, several datasets have been published to allow custom BTS system training, such as the Chisco dataset provided by Zhang et al. (2024), the DAIS dataset, EEGIS, etc.

Acknowledgement

The work has been supported by the grant of the University of West Bohemia, project No. SGS-2025-011.

References

Anumanchipalli, G.K., Chartier, J. and Chang, E.F. (2019). Speech synthesis from neural decoding of spoken sentences. *Nature*, 568(7753), pp. 493–498. Available from: https://doi.org/10.1038/s41586-019-1119-1.

Lee, Y.-E., Lee, S.-H., Kim, S.-H. and Lee, S.-W. (2023). Towards Voice Reconstruction from EEG during Imagined Speech. *AAAI Conference on Artificial Intelligence*, 37(5), pp. 60306038. Available from: https://doi.org/10.1609/aaai.v37i5.25745.

Rastogi, S., Dadwal, H., Modi, K., Bedi, J. and Singh, J. (2025). Towards Sentence Level Imagined Speech Generation from EEG signals. *Interspeech 2025*, pp.55585562. Available from: https://doi.org/10.21437/interspeech.2025-2752.

Schultz, Tanja and Wand, Michael and Hueber, Thomas and Krusienski, Dean J. and Herff, Christian and Brumberg, Jonathan S. (2017). Biosignal-Based Spoken Communication: A Survey. *IEEE/ACM transactions on audio, speech, and language processing*, 25(12), pp. 22572271. Available from: https://doi.org/10.1109/taslp.2017.2752365.

Zhang, Z., Ding, X., Bao, Y., Zhao, Y., Liang, X., Qin, B. and Liu, T. (2024). Chisco: An EEG-based BCI dataset for decoding of imagined speech. *Scientific data*, 11(1). Available from: https://doi.org/10.1038/s41597-024-04114-1.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Chronicle Annotator: A Web Platform for Collaborative HTR Annotation of Historical Manuscripts

Bc. Daniel Vodička[1], [prof. Ing. Pavel Král, Ph.D., Doc. Ing. Ladislav Lenc, Ph.D][2]

1 Introduction

Historical chronicles represent an invaluable cultural heritage, yet their handwritten contents remain largely inaccessible to automated search and analysis. The State Regional Archive in Pilsen (Státní oblastní archiv (SOA)[1] v Plzni), which maintains extensive collections of manuscripts, commissioned this work with the goal of developing a software tool that would allow archivists to efficiently produce machine-readable transcriptions of chronicle pages.

Digitising handwritten documents requires Handwritten Text Recognition (HTR) — a task significantly harder than printed-text OCR due to the high variability of individual handwriting styles, archaic vocabulary, and image quality degradation caused by ageing paper and ink. The SOA expressed an interest for an integrated, web-based platform that (i) applies HTR models to scanned pages, (ii) allows domain experts to review and correct the recognised text directly in the browser, and (iii) closes the feedback loop by exporting the corrected annotations so that the underlying models can be continuously improved on the archive's own data.

This work presents *Chronicle Annotator*, a platform that addresses all three requirements. The system supports both manual and semi-automatic annotation, collaborative work on shared projects, and iterative fine-tuning of two HTR backends — *Kraken* (Kiessling , 2019) and the multimodal vision-language model *HunyuanOCR* (Tencent , 2025).

2 System Architecture

Chronicle Annotator is a web application fully containerised with Docker platform. In production it runs as seven cooperating services; (See Figure 1).

The *Flask* web application (served by *Gunicorn* behind an *Nginx* reverse proxy) handles all user-facing functionality. OCR processing is handled asynchronously by numbered dedicated *Celery* workers communicating through a *Redis* message broker, so that computationally expensive HTR jobs never block the web interface. The `chronicle_worker` container runs *Kraken* for line segmentation and, when *Kraken* HTR is selected, also for text recognition (can run in multiple instances). When more powerful model *HunyuanOCR* is selected, it dispatches per-line tasks to a second worker, `chronicle_hunyan_ocr`.

3 Application Features

Interactive page editor. Figure 2 shows the editor interface. After a page is uploaded under specific project and processed by the OCR pipeline, the scanned image is displayed with coloured polygon overlays marking each detected text line. Users can browse lines, correct the

---

[1] student of the master degree program Applied Sciences, field of study Computer Science and Software Engineering, e-mail: vodickad@students.zcu.cz

[2] e-mail: pkral@fav.zcu.cz, llenc@fav.zcu.cz

[1] https://www.soaplzen.cz/

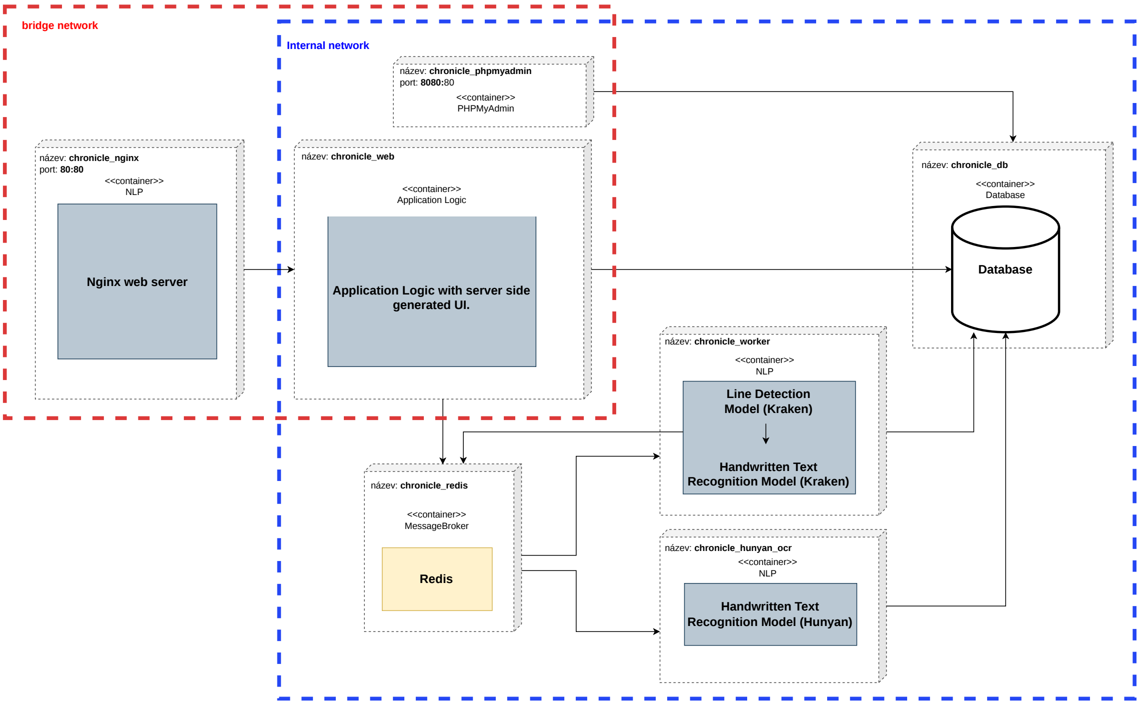
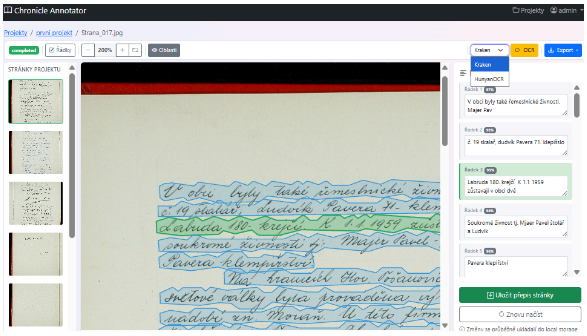


Figure 1: Production Docker container topology. The `backend_chronicle_network` is defined as `internal` (no direct internet access); only Nginx and phpMyAdmin are reachable from the public bridge network.

recognised text inline, add or delete lines, reorder them, and trigger re-processing of any page with a different model.

Export and model fine-tuning. Annotated projects can be exported as ZIP archives in PAGE XML (for *Kraken* fine-tuning via `ketos train`), ALTO XML v4 (for *HunyuanOCR* fine-tuning), or plain TXT. The archive structure matches the input expected by the training scripts directly, requiring no additional pre-processing.



Řádek 1 97%
V obci byly také řemeslnické živnosti. Majer Pav

Řádek 2 95%
č. 19 skalař, dudvik Pavera 71. klepišslo

Řádek 3 95%
Labruda 180. krejčí K 1.1 1959 zůstavají v obci dvě

Řádek 4 94%
Soukromé živnost tj. Mjaer Pavel štolář a Ludvik

Řádek 5 96%
Pavera klepiřství

ⓘ Změny se průběžně ukládají do local storage

Figure 2: Chronicle Annotator page editor.

4 Conclusion

Chronicle Annotator is a web-based platform integrating collaborative editing, asynchronous HTR processing, and model fine-tuning to progressively improve handwriting transcription accuracy. Serving as a pilot project for archives in Pilsen and Cheb, the system is expected to be deployed across numerous other archival institutions in the Czech Republic.

References

Kiessling, B. (2019) Kraken - an Universal Text Recognizer for the Humanities. *Digital Humanities 2019 (DH2019)*, Utrecht, The Netherlands.

Tencent Multimedia Lab (2025) *HunyuanOCR: A Multimodal Vision-Language Model for Handwritten Text Recognition*.

Highly photoactive Cu_2O films with enhanced crystallinity for PEC water splitting

Jan Vosejpka[1], Jiří Čapek[2]

1 Introduction

Cuprous oxide (Cu_2O) is a promising material for photoelectrochemical (PEC) water splitting due to its suitable bandgap and earth-abundant nature. It meets theoretical requirements for efficient solar-driven hydrogen production, making it a key candidate for sustainable energy applications. However, conventional fabrication methods, such as chemical synthesis, thermal oxidation, or DC magnetron sputtering, often fail to produce Cu_2O films with the ideal density, crystallinity, and single-phase composition required for efficient PEC performance. These limitations hinder the material's photocurrent output and stability, necessitating the exploration of advanced deposition techniques and post-deposition strategies.

In our study, we utilise reactive High-Power Impulse Magnetron Sputtering (HiPIMS) to deposit Cu_2O films. The HiPIMS process was optimised by carefully adjusting discharge conditions, primarily the oxygen partial pressure. To further refine the microstructure of the deposited films, we investigated two strategies: increasing the deposition temperature and performing vacuum annealing after deposition. The resulting films were characterised for their structural and photoelectrochemical properties using XRD, SEM, and PEC measurements.

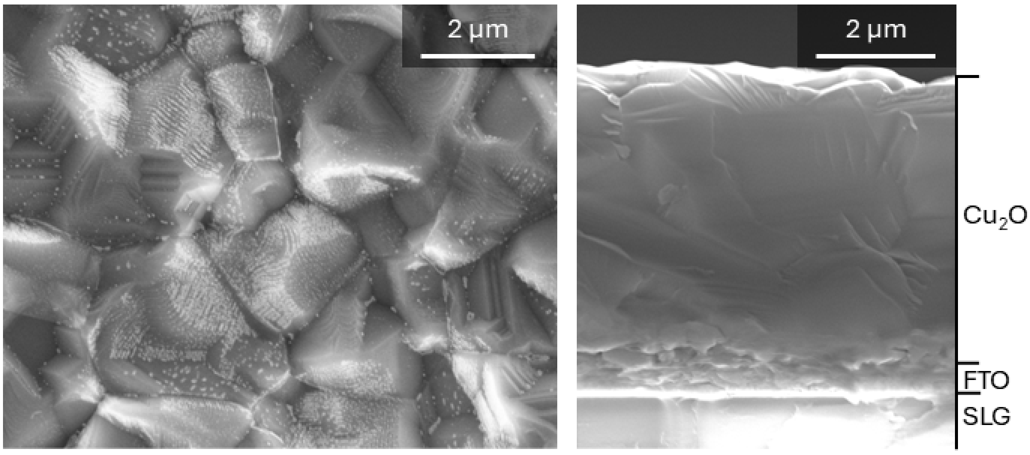


Figure 1: SEM images of 5 μm-thick film of Cu_2O deposited using HiPIMS at partial pressure of oxygen, $p_{O2} = 0.42\,\mathrm{Pa}$, with substrate heating at 500 °C during the deposition and post-annealed at 600 °C in a vacuum: left panel top-view and right panel cross-section of the film.

---

[1] student of the doctoral study program Plasma Physics and Physics of Thin Films, Department of Physics and NTIS – European Centre of Excellence, University of West Bohemia in Pilsen, e-mail: vosejpka@fav.zcu.cz

[2] Department of Physics and NTIS – European Centre of Excellence, University of West Bohemia in Pilsen, e-mail: jcapek@fav.zcu.cz

2 Results

Optimised HiPIMS conditions enabled the deposition of single-phase Cu_2O films with uniform thickness and controlled microstructure. Increased deposition temperature and post-deposition annealing significantly promote crystallisation and recrystallisation of the films, respectively, achieving a highly crystalline and fully densified structure, as shown in Figure 1 and left panel of Figure 2.

Photoelectrochemical measurements demonstrated that precise control of deposition and annealing conditions substantially improves photocurrent performance. The resulting photocurrents, shown in Figure 2, of -3 mA/cm^2 at 0.03 V vs. RHE rank among the highest reported for bare single-phase Cu_2O in the literature. This study demonstrates that reactive HiPIMS, combined with vacuum annealing, is a highly effective method for fabricating single-phase Cu_2O films with superior crystallinity and PEC performance. We also focus on developing multilayered Cu_2O-based photocathodes. This involves tailoring overlayers to improve charge separation and hydrogen evolution reaction (HER) kinetics while minimising degradation of the film, together with underlayers, which help extract charge carriers from the film.

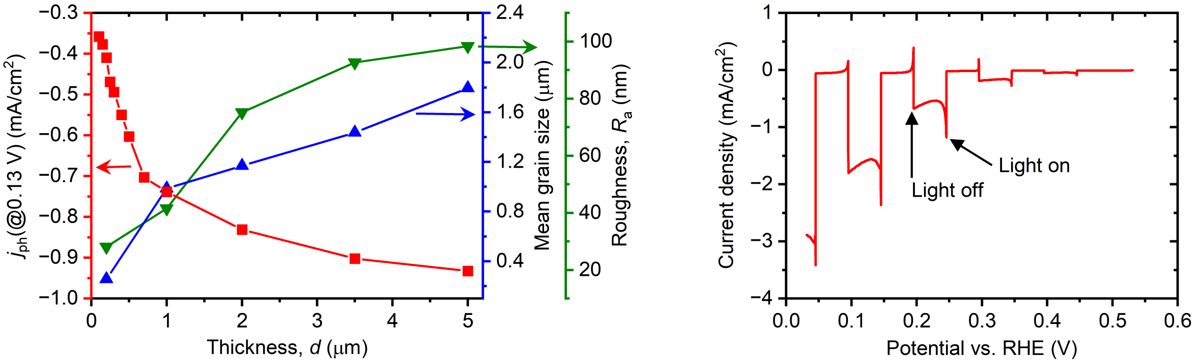


Figure 2: Left panel shows photocurrent density at 0.13 V vs. RHE, $j_{\mathrm{ph}}(@0.13\,\mathrm{V})$ (red line), mean grain size (blue line), and surface roughness R_a (green line) as a function of the film thickness, d. The left panel shows linear sweep voltammetry with a scan rate of 10 mV/s and chopped light. The films were deposited using HiPIMS at a partial pressure of oxygen, $p_{\mathrm{O2}} = 0.42\,\mathrm{Pa}$, with substrate heating during deposition at 500 °C, followed by post-annealing at 600 °C in a vacuum.

Acknowledgement

This work was supported by the project Quantum materials for applications in sustainable technologies (QM4ST), funded as project No. CZ.02.01.01/00/22_008/0004572 by Programme Johannes Amos Comenius, call Excellent Research.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Impact of reduced field-of-view overlap on depth estimation accuracy in modern stereo deep learning architectures

Jan Vrba[1]

1 Introduction

Reliable depth estimation is critical for applications like autonomous driving. To maximize coverage with fewer sensors, industrial multi-camera setups often restrict field-of-view (FoV) overlap. This limits true stereo matching to the central region, creating structural occlusions at the periphery where secondary camera data is missing.

Modern deep learning architectures (e.g., RAFT-Stereo Lipson et al. (2021)) address stereo correspondences via correlation volumes, while advanced models like MonSter++ Cheng et al. (2025) integrate monocular depth priors to improve robustness. This raises a fundamental question: can these networks compensate for hardware-induced overlap loss using single-image context?

This paper presents a quantitative error analysis of deep stereo networks under simulated overlap reduction. Using the KITTI 2015 dataset Menze and Geiger (2015), we artificially restrict the FoV by zero-masking corresponding image boundaries to create occlusion zones of varying widths. By separately evaluating fully stereo (overlap) and occluded regions, we measure accuracy degradation when models predict depth purely from monocular context.

2 Methodology

To simulate reduced stereo FoV without shifting cameras' principal points (which would disrupt correlation volume search ranges), we propose a spatial zero-masking pipeline.

For a rectified image pair of width W, reducing mutual overlap by percentage p requires masking a boundary of $c = \lfloor W \cdot (p/100) \rfloor$ pixels. To preserve epipolar geometry, we zero-mask the input tensors directly: the left image is masked on its right boundary ($x \geq W - c$) and the right image on its left boundary ($x < c$).

Figure 1: Example of masked left input image simulating a reduced mutual overlap.

After network processing, the predicted disparity is cropped to the valid width ($x < W - c$). We evaluate against LiDAR ground truth across two zones: the **Occluded Zone** ($x < c$), relying purely on monocular context, and the **Overlap Zone** ($x \geq c$), containing valid stereo textures. Performance is measured via End-Point Error (EPE), Root Mean Square Error (RMSE), the D1-out outlier metric, and the percentage of pixels exceeding absolute error thresholds of 1, 3, and 5 pixels (> 1px, > 3px, > 5px).

---

[1] PhD student of Cybernetics, Artificial Intelligence, e-mail: janvrb@fav.zcu.cz

3 Results and discussion

We evaluated MonSter++ on KITTI 2015 under varying FoV overlap reductions ($p \in \{0\%, 15\%, 25\%, 33\%\}$). The results are summarized in Table 1.

Table 1: Quantitative evaluation under simulated FoV overlap reduction.

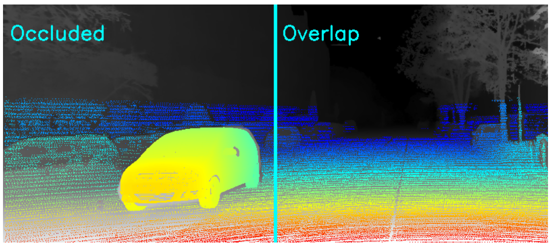
Region	Crop (p)	EPE	RMSE	D1 (%)	> 1px	> 3px	> 5px
All	0%	0.89	2.31	**3.41**	19.83	**3.58**	**2.12**
	15%	**0.86**	**2.15**	4.05	**17.72**	4.20	2.29
	25%	1.07	2.36	7.19	22.78	7.27	3.85
	33%	1.54	2.94	13.88	31.74	13.95	7.62
Occluded	15%	**1.85**	**2.98**	**14.83**	**49.11**	**15.27**	**7.30**
	25%	2.17	3.30	20.01	54.05	20.16	10.25
	33%	2.73	4.02	28.59	60.12	28.71	15.60
Overlap	15%	0.72	1.89	2.57	13.12	2.66	1.63
	25%	0.64	1.65	2.17	10.68	2.19	1.37
	33%	**0.56**	**1.31**	**1.71**	**8.44**	**1.72**	**1.01**

When interpreting these results, it is crucial to consider the geometric proportions. For instance, at a $p = 33\%$ crop, the original image width W is reduced to an evaluation width of $0.67W$. Within this evaluated image, the artificially occluded zone spans exactly $0.33W$. Therefore, the Occluded zone constitutes approximately 50% of the final evaluated image area, leaving the remaining 50% for the true stereo Overlap zone.

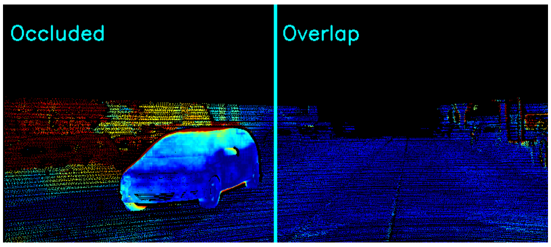
This geometric reality explains the sharp increase in the global error (All zone) at higher crop percentages. While the network maintains exceptional robustness in the Overlap zone (achieving a sub-pixel EPE of 0.56 at 33% crop), its performance collapses in the Occluded zone. Forced to hallucinate depth relying solely on single-image priors, the EPE in the 33% occluded area spikes to 2.73 px, with over 60% of the pixels exceeding a 1-pixel error threshold. This proves that while modern networks leverage monocular context effectively for minor occlusions, they cannot reliably reconstruct depth geometry in wide structural blind spots. Nevertheless, future work must extend this evaluation across diverse datasets and alternative network architectures.

Acknowledgement

This work was supported by project SGS-2025-011.



(a) Disparity prediction with sparse GT overlay



(b) Error map (red indicates high error)

Figure 2: Visual example of prediction and evaluation process

References

Cheng, J., Liao, W., Cai, Z., et al. (2025) *MonSter++: Unified Stereo Matching, Multi-view Stereo, and Real-time Stereo with Monodepth Priors*. Available from: `https://arxiv.org/abs/2501.08643`

Lipson, L., Teed, Z., Deng, J. (2021) *RAFT-Stereo: Multilevel Recurrent Field Transforms for Stereo Matching*. Available from: `https://arxiv.org/abs/2109.07547`

Menze, M., Geiger, A. (2015) *Object scene flow for autonomous vehicles*. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Available from: `https://doi.org/10.1109/CVPR.2015.7298925`

Thermal Engineering and Phase Control of Zirconium Oxynitride Thin Films Synthesized via Reactive HiPIMS for Energy Applications

Minh Thanh Vu[1], David Kolenatý, Jiří Čapek, Stanislav Haviar, Radomír Čerstvý, Petr Zeman

1 Introduction

Transition metal oxynitrides combine the chemical durability of oxides with the narrow, flexible band structures of nitrides, making them excellent candidates for solar water splitting [1, 2]. Zirconium oxynitride Zr–O–N compounds - specifically the Zr_2ON_2 stoichiometry - possess favorable band alignments and visible-light absorption ideal for the oxygen evolution reaction [3, 4]. However, synthesizing phase-pure Zr–O–N with high crystallinity remains challenging due to the narrow thermodynamic window for nitrogen substitution into oxide lattices under conventional physical vapor deposition.

2 Experimental

To address this, Zr–O–N thin films were deposited across a wide compositional range using reactive high-power impulse magnetron sputtering (HiPIMS) from a metallic Zr target. Utilizing the high plasma ionization of the HiPIMS discharge, precise control over oxygen and nitrogen incorporation was achieved on silicon and glass substrates under varied O_2/N_2 gas flow ratios. Depositions were performed at both room temperature and with in-situ heating at 450 C. To drive phase formation, post-deposition vacuum annealing was executed up to 600 C. Film properties were systematically analyzed using WDS, XRD, SEM, four-point probe resistivity, and UV–Vis spectrophotometry.

3 Results and Conclusions

The films demonstrated a clear phase evolution from insulating monoclinic ZrO_2 to semiconducting oxynitrides, and finally metallic ZrN as the O_2/N_2 flow ratio decreased. This structural shift correlates with a systematic narrowing of the optical bandgap (quantified via Tauc analysis) and a multi-order increase in electrical conductivity. Furthermore, photoelectrochemical (PEC) evaluations and electrochemical impedance spectroscopy (EIS) successfully resolved the underlying interfacial charge transfer kinetics and photocurrent stability under simulated solar light. Ultimately, this investigation provides effective processing maps to isolate highly crystalline, visible-light-active zirconium oxynitrides optimized for clean energy conversion.

4 Figures

---

[1] student of the doctoral degree program Applied Sciences, field of study Plasma Physics and Physics of Thin Films, e-mail: thanh@fav.zcu.cz

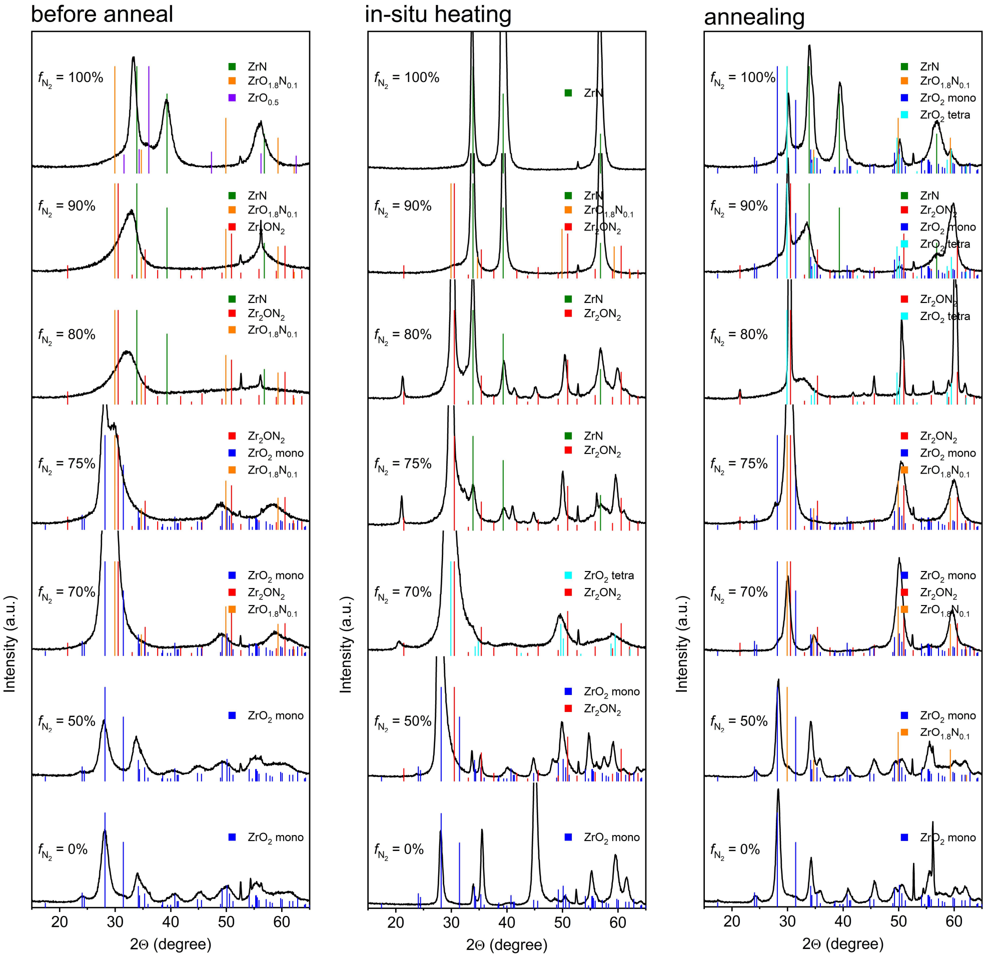


Figure 1. Thermal effect on phase structure of Zr – O – N compound from left (Room temperature) to middle (in-situ 450 C heating) and right (600 C annealing)

5 References

[1] J-H. Huang, Y.-Y. Hu, G.-P. Yu, (2011) *Surf. Coat. Technol*. 205 5093.

[2] Y. Li, M. You, X. Li, B. Yang, Z. Lin, J. Liu, (2022) *IEEE 35th Int. Conf. on Micro Electro-Mechanical Systems*, Tokyo, Japan, pp. 806-809.

[3] V. Streibel, J.L. Schönecker, L.I. Wagner, E. Sirotti, F. Munnik, M. Kuhl, C. Jiang, J. Eichhorn, S. Santra, I.D. Sharp, (2024) *ACS Appl. Energy Mater*. 7 4004.

[4] Y. Wu, P. Lazic, G. Hautier, K. Persson, G. Ceder, (2012) *Energy Environ. Sci*. 6 157.

Acknowledgement

This work was supported by the project Quantum materials for applications in sustainable technologies (QM4ST), funded as project No. CZ.02.01.01/00/22_008/0004572 by Programme Johannes Amos Comenius, call Excellent Research.

Drawing Mining - Title Block Information Extraction

Jiří Vyskočil[1]

1 Introduction

Current workflows for processing technical drawings are still largely manual, time-consuming, and error-prone, which negatively impacts manufacturing planning and cost estimation. Therefore, a project in collaboration with Amitia s.r.o. is being conducted to develop AI-based support for cost estimation and production planning in manufacturing environments.

This paper focuses on a subtask of the project: the extraction of information from title blocks and related global drawing annotations. Besides metadata, such as the author or revision date, these regions may also contain manufacturing-related information, including projection methods, tolerances, surface roughness specifications, and material properties. However, such global information is not strictly confined to the title block and may appear elsewhere in the drawing.

A key challenge arises from the high resolution and visual complexity of technical drawings. While title blocks maintain relatively consistent text size across different drawing formats, their relative spatial extent decreases significantly in large-format drawings. Direct processing of full-resolution drawings using Multimodal Large Language Models (MLLMs) therefore introduces unnecessary visual redundancy and computational overhead. To mitigate this issue, the proposed approach combines a YOLO-based detector by Jocher et al. (2023) with the Qwen3-VL MLLM by Qwen Team (2025), enabling targeted processing of semantically relevant regions only.

2 Methodology

Fig. 1 illustrates the proposed pipeline, consisting of preprocessing via object detection followed by multimodal information extraction.

Table detector. A YOLO-based detector is used to localize relevant regions in technical drawings using resized images with a resolution of 1024×1024 px. The detector identifies title blocks and related metadata regions, as well as other structures such as bills of materials or revision tables, which are not relevant to the task considered in this work. The detected regions are subsequently cropped and forwarded to the MLLM, thereby reducing input complexity and filtering irrelevant content.

Information extractor. The cropped regions are processed using an MLLM; we used Qwen3-VL and GPT-5.4 / GPT-5$_n$. Initial experiments focused on zero-shot prompting and prompt engineering to enforce structured JSON outputs, where ChatGPT additionally supports native response format constraints enforcing a predefined JSON schema. Subsequently, the open-weight Qwen3-VL model was fine-tuned using Low-Rank Adaptation (LoRA) by Mangrulkar et al. (2022). Training was performed on approximately 70 annotated samples, while the initial evaluation was conducted on a test subset consisting of roughly 15 drawings.

---

[1] PhD student, Applied Sciences – Cybernetics program, e-mail: vyskocj@fav.zcu.cz

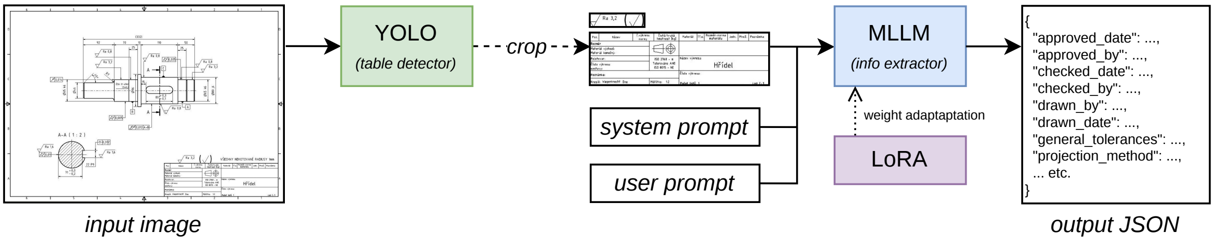


Figure 1: Overview of the proposed pipeline demonstrated on a publicly available technical drawing from www.stovkomat.cz.

Detector	mAP_{50}	mAP_{50-95}
YOLO-11$_x$	**99.5**	99.3
YOLO-26$_x$	**99.5**	**99.5**

Table 1: Object detection performance on localization of tables and text clusters.

MLLM	Fine-tuned	Response format	Exact_m	Exact_μ
GPT-5$_n$	✗	✓	67.0	44.1
GPT-5$_n$	✗	✗	71.3	50.3
GPT-5.4	✗	✓	77.4	62.9
GPT-5.4	✗	✗	78.3	64.3
Qwen3-VL$_{32B}$	✗	–	76.8	61.2
Qwen3-VL$_{8B}$	✓	–	**100.0**	**100.0**

Table 2: Comparison of the title block information extraction for evaluated MLLMs.

3 Results and Conclusion

The proposed pipeline achieves consistent performance in both the preprocessing and multimodal extraction stages. The object detector reliably localizes relevant regions in technical drawings, enabling effective reduction of irrelevant visual content prior to MLLM inference.

On the initial test subset (15 drawings), the fine-tuned Qwen3-VL model achieves 100% exact accuracy, while on an **extended test set** of 48 drawings it achieves **99.6% accuracy**. Here, *exact* denotes the proportion of correctly predicted labels, whereas *macro* and *micro* represent label-wise and frequency-weighted averages, respectively. A notable observation concerns the response format constraint. Although it enforces strict JSON consistency, it occasionally reduced performance by limiting model flexibility. This does not imply that structured response formats are inherently harmful, as they may also be beneficial in other tasks.

Acknowledgement

The work has been supported by the grant of the University of West Bohemia, project No. SGS-2025-011. Computational resources were provided by the e-INFRA CZ project (ID:90254), supported by the Ministry of Education, Youth and Sports of the Czech Republic.

References

Jocher, G., Qiu, J., Chaurasia, A. (2023) *Ultralytics YOLO (Version 8.4.47)* [Software]. Available from: GitHub [Accessed May 2026].

Qwen Team (2025) *Qwen3 Technical Report*. arXiv preprint: 2505.09388.

Mangrulkar, S., Gugger, S., Debut, L., Belkada, Y., Paul, S., Bossan, B., Tietz, M. (2022) *PEFT: State-of-the-art Parameter-Efficient Fine-Tuning Methods*. Available from: GitHub [Accessed May 2026].

Biological data acquisition and its impact on pipeline design

Kateřina Wolf [*1], Tomáš Honzík [1], David Staníček [1], Robin Klieber [2], Pavel Ostašov [3], Zia Ullah [2], Pavel Jindra [3,4], Monika Holubová [2,3] and Lucie Houdová [1]

1 Introduction

Although creating a pipeline for the automatic processing of genetic data is not a new concept, existing pipeline may not be fully applicable or usable due to the specific characteristics of the examined genes or due to new methods of obtaining data from biological material (sequencing methods). Next-generation sequencing (NGS) is currently the most widely used method; however, there is a gradual shift towards long-read sequencing technologies like PacBio and Oxford Nanopore Technologies. NGS, specifically the Illumina sequencer, generates short DNA sequences known as reads, which can typically be around 250 bases long. To analyse these reads, they are typically aligned to a reference sequence. In contrast, PacBio HiFi technology produces longer reads, in our case approximately 1,400 bases. Although this method is more expensive, it can address issues where NGS falls short. This enhanced capability can be demonstrated by examining KIR isoforms.

2 Biological Background

Every year, 700 people in Czechia diagnosed with hemato-oncological diseases receive hematopoietic stem cell transplantation (HSCT) as a vital treatment. Suitable donors are primarily selected based on HLA gene matching, followed by non-genetic factors such as age and gender. Recent studies have also explored the impact of additional genes like the killer-cell immunoglobulin-like receptor (KIR). These genes can play a role in the patient outcome, and they are usually involved in post-transplant conditions like graft-versus-host disease or graft-versus-leukemia. KIR genes are characterised as highly diverse due to KIR gene content, copy number, or allelic polymorphism. Their mutual similarity exists not only at the variant level but also across genes. Additionally, the mechanism of alternative RNA splicing creates a diverse range of isoforms [1], however available public data sources on this topic are currently quite limited, which brings new challenges into data processing.

3 Discussion

Automating the processing of genetic data requires careful planning. The design of data pipelines must be tailored to the specific data and the insights we aim to gather (such as used primers, sequencing technology and processing tools). It is also important to be aware of the limitations of the results produced by the pipeline.

---

[*] Phd student, Applied Sciences, field of study cybernetics, e-mail: kkratoch@fav.zcu.cz
[1] Faculty of Applied Sciences, University of West Bohemia in Pilsen
[2] Faculty of Medicine in Pilsen, Charles University
[3] Department of Haematology and Oncology, University Hospital Pilsen
[4] Czech National Marrow Donors Registry, Czechia

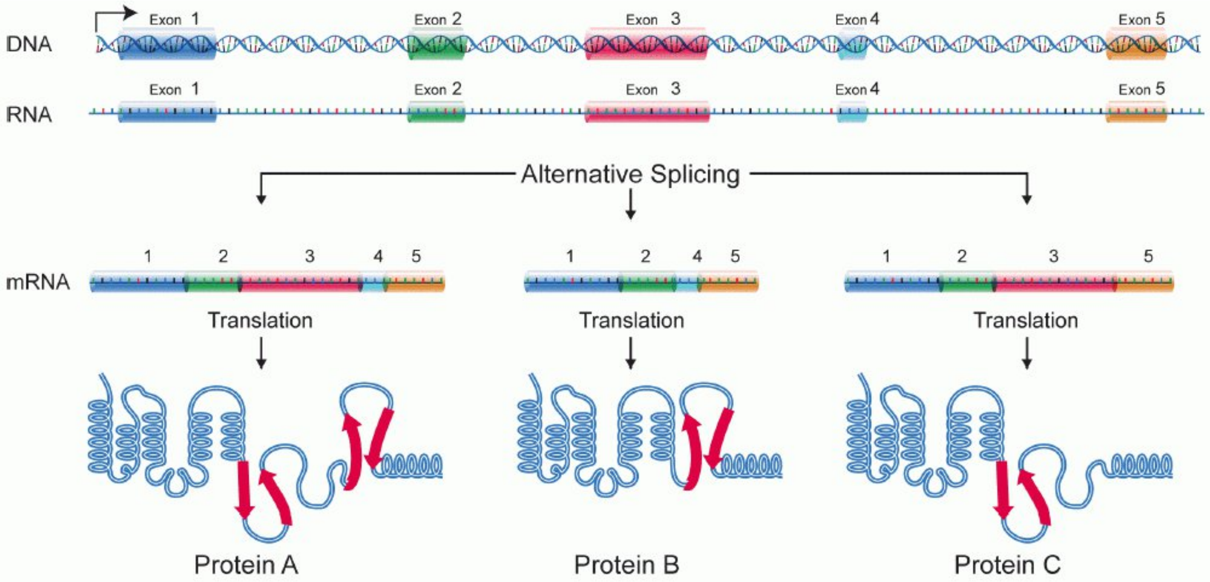


Figure 1: Sequencing data can obtained from various genetic templates, such as DNA, RNA and cDNA. Sequencing data can also be obtained by various technologies, such as: NGS or PacBio. An isoform is a variant of the same gene created through alternative splicing, resulting the production of multiple proteins. Illustrative image from [2].

Acknowledgement

This work was supported by specific university research project SGS-2025-020.

References

1. Bruijnesteijn J, van der Wiel MKH, de Groot N, Otting N, de Vos-Rouweler AJM, Lardy NM, de Groot NG and Bontrop RE (2018) Extensive Alternative Splicing of KIR Transcripts. Front. Immunol. 9:2846. doi: 10.3389/fimmu.2018.02846
2. Alternative splicing; https://en.wikipedia.org/wiki/Alternative_splicing
3. Robinson et al.; IPD – The Immuno Polymorphism Database, Nucleic Acids Research; ; DOI: 10.1093/nar/gks1140
4. Ahmad S, da Costa Gonzales L J, Bowler-Barnett E H, Rice D L, Kim M, Wijerathne S, Luciani A, Kandasaamy S, Luo J, Watkins X, Turner E, Martin M J, UniProt Consortium The UniProt website API: facilitating programmatic access to protein knowledge Nucleic Acids Research, gkaf394 (2025)
5. Chan HW, Kurago ZB, Stewart CA, Wilson MJ, Martin MP, Mace BE, Carrington M, Trowsdale J, Lutz CT. DNA methylation maintains allele-specific KIR gene expression in human natural killer cells. J Exp Med. 2003 Jan 20;197(2):245-55. doi: 10.1084/jem.20021127. PMID: 12538663; PMCID: PMC2193817.
6. Bruijnesteijn J, van der Wiel MKH, de Groot N, Otting N, de Vos-Rouweler AJM, Lardy NM, de Groot NG and Bontrop RE (2018) Extensive Alternative Splicing of KIR Transcripts. Front. Immunol. 9:2846. doi: 10.3389/fimmu.2018.02846

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Modular peripheral system for Raspberry Pi Zero emulator

Matěj Černý[1]

1 Introduction

The development of software for embedded systems is a recurring practical challenge: the target hardware is often unavailable, expensive, or inconvenient to work with directly. Simulation environments resolve this dependency by letting firmware run against a virtual representation of the hardware. One example is ZeroMate developed by Šilhavý (2024), an educational Raspberry Pi Zero emulator targeting the KIV-RTOS operating system (Úbl (2021a)).

Most existing simulators, however, draw a hard line between the emulated processor and everything around it: peripherals are accessories of one fixed emulator, with no facility for arbitrary chips to interact directly. A simulator that lifts this restriction would make it possible to assemble realistic schematics — displays driven by shift registers, sensors reporting over a wire, multiple processors talking over a shared bus.

2 Existing Solutions

Several simulation tools are publicly available, including QEMU, Logisim Evolution, and SimulIDE, but none simultaneously offers a shared wire model with multi-MCU capability and a native C++ plugin SDK. QEMU operates at the memory-mapped register level and stubs key GPIO behavior. Logisim Evolution targets gate-level education with Java-coupled components. SimulIDE comes closest, treating AVR and PIC cores as schematic-level participants, but its MCU integrations are hard-coded and cannot be extended through a stable interface.

3 Proposed Solution – CircuitSim

CircuitSim[2] was developed to fill this gap. It is a tick-accurate digital circuit simulator whose wire model uses six states (strong high and low, floating, weak pull-up and pull-down, and an error state for conflicting drivers), enabling correct simulation of open-drain protocols such as I^2C and of shared-bus designs in general.

Custom chips are authored as native C++ plugins, and a plugin is added to a project by a single configuration line; the same mechanism is used to integrate ZeroMate, whose entire emulator becomes one chip on a schematic with its GPIO pins exposed. Multiple ZeroMate instances, other CPU emulators, or other chips can coexist on the same wire model. The simulator ships with a visual schematic editor showing live wire-state coloring, a casing editor for designing chip visuals, a probe window for inspecting signals during simulation, and a basic built-in chip suite, including a Lua-programmable chip for one-off custom logic.

---

[1] student of the master degree program Computer Science and Engineering, field of study Distributed Systems, e-mail: matcerny@students.zcu.cz

[2] `https://github.com/Cooble/CircuitSim`

4 Results

Testing covers per-chip tests, protocol tests for the I^{2}C and UART tools, and integration tests booting real ARM firmware on the ZeroMate plugin against external chips. The wire-faithful environment uncovered three latent bugs in the original ZeroMate's mini-UART. A side-by-side comparison against a physical Raspberry Pi Zero connected to the KIV-DPP-01 expansion board (Úbl (2021b)), running the same KIV-RTOS-derived kernel image, produced identical OLED text and an identical 8-segment digit driven through an SN74HC595 shift register. With peak compiler optimizations, running ZeroMate as a CircuitSim plugin introduces a throughput overhead of 9–10 % compared to the standalone emulator.

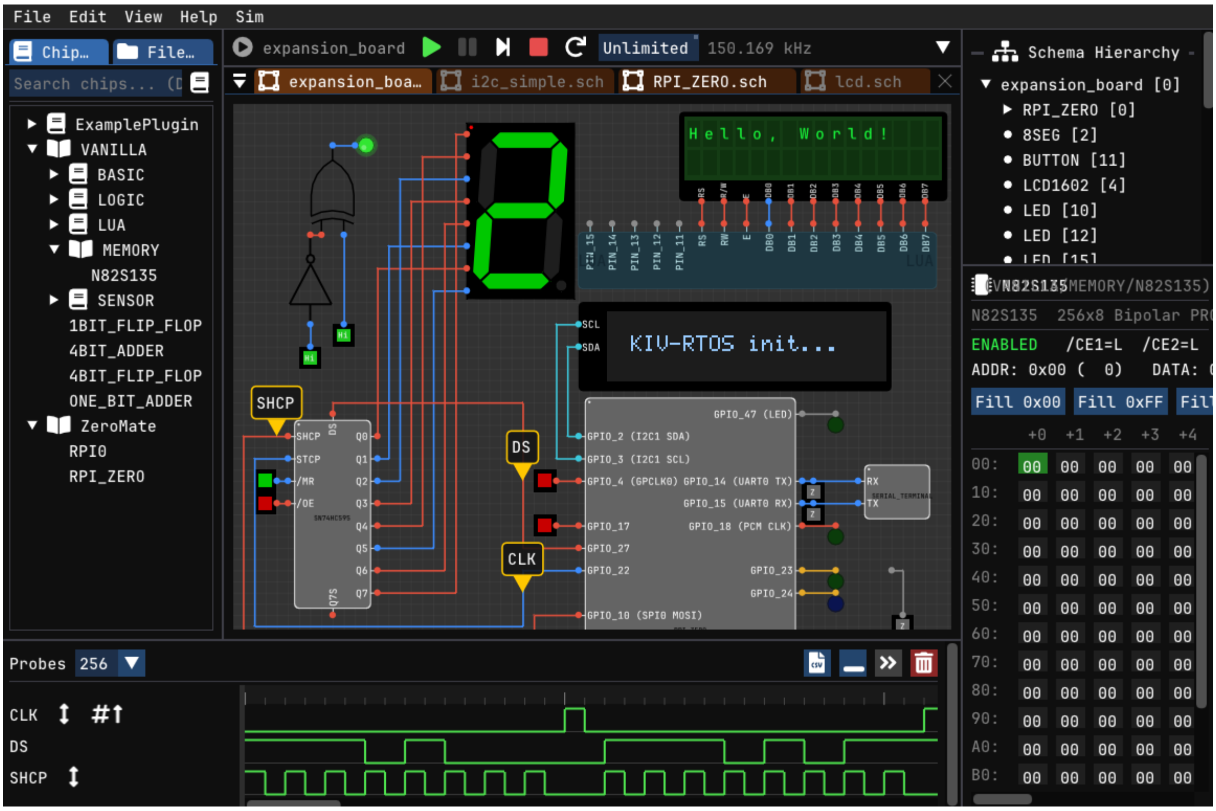


Figure 1: CircuitSim GUI: ZeroMate executing the KIV-RTOS kernel to drive other chips, including an I^{2}C LCD and an 8-segment display via an SN74HC595 shift register. The bottom probe window displays real-time logic waveforms.

References

Šilhavý, J. (2024) *ARMv6 Processor Emulator for Raspberry Pi Environment Emulation*. Master's thesis. Pilsen, University of West Bohemia, Faculty of Applied Sciences.

Úbl, M. (2021a) *KIV-RTOS: Educational Operating System for Raspberry Pi Zero*. Available at: `https://github.com/MartinUbl/KIV-RTOS` [Accessed 6th May 2026].

Úbl, M. (2021b) *KIV-DPP-01 Expansion Board*. University of West Bohemia, Faculty of Applied Sciences. Available at: `https://home.zcu.cz/~ublm/files/os/kiv-dpp-01-en.pdf` [Accessed 6th May 2026].

Speech&Text Deception Detection

Jan Šimek[1]

1 Introduction

Recognizing deception is a difficult task even for humans (Chen et al. 2020), hence automatic detection would be beneficial for interactions where people tend to deceive, for instance interrogations or job interviews. We focus on detection from speech, as it is non-invasive and can even be performed remotely (for example via phone call). Our aim is to compare the detection success rate of various machine learning classifiers and modalities (acoustic and linguistic) to evaluate which representations and classifiers are the most suitable for this task.

2 Training Data

The Columbia SRI Colorado (CSC) corpus (Hirschberg et al. 2005) fits our objective acceptably. The corpus consists of 32 speakers participating in a six areas performance test. Each participant was interviewed and instructed to tell the truth about four of the test areas and deceive about the two remaining. When deceiving, the participant was asked to push a pedal which was hidden from the interviewer. The result is 7 hours of annotated speech with transcriptions – 60% truthful utterances and 40% deceptive ones. For our purposes, the long FLAC files were divided into separate sentences, and these were split into the training (90%) and testing (10%), where a particular speaker appears only in one of these sets.

3 Experiments

First, we extracted acoustic features from the audio files using two different approaches. The predefined ComParE set (Weninger et al. 2013) captures 6,373 features representing various voice characteristics (we selected only the 32 top-ranked ones, as a higher number did not improve results), while the `wav2vec2-base`[1] neural model creates its own embedded representations. These two sets were utilized to train a set of machine learning classifiers to compare which features provide a more suitable representation.

Next, we focused on linguistic features and performed two methods of text embedding: standard TF-IDF and `all-mpnet-base-v2`[2], one of the S-BERT neural models. The goal was to compare the informational richness of the representations. Moreover, we wanted to analyze which would be a better fit for the task: linguistic or acoustic representation.

Finally, we merged the features of both modalities to reveal whether their combination could lead to even better results. To measure the success rate of all the experiments, we trained several machine learning classifiers – decision tree, random forest, SVM, SGD, and XGBoost – and recorded standard evaluation metrics (balanced accuracy: BA, recall: R(lie), precision: P) for every experiment. These results are discussed in the next section.

---

[1] student of the doctoral degree program Cybernetics, e-mail: simio@fav.zcu.cz

[1] `https://huggingface.co/facebook/wav2vec2-base`

[2] `https://huggingface.co/sentence-transformers/all-mpnet-base-v2`

4 Results and Conclusion

Comparing the acoustic representations, the experiments showed that balanced accuracies of most models move a bit above the chance level (0.5) with no apparent dominance of hand-crafted or self-learned features. The best results were achieved with the decision tree classifier (BA = 0.488, R(lie) = 0.438) and XGBoost (BA = 0.520, R(lie) = 0.399) trained with selected ComParE features. Other classifiers ended with worse metrics, or they completely collapsed to generally classify true or false only.

For linguistic features, we can see a marginal superiority of S-BERT embeddings compared to traditional TF-IDF, with best results achieved with SVM with linear kernel (BA = 0.571, R(lie) = 0.368) and decision tree (BA = 0.511, R(lie) = 0.404). However, the highest recall of the "lie" was spotted at the decision tree trained with TF-IDF embedding (0.433). It is hard to state which of these representations performs better, but we can affirm that the decision tree classifier achieves acceptable results, regardless of the input representation, while the performance of other classifiers depends more on the embedding of input data.

From these results, it can be stated that linguistic modality is more advantageous for our approach. The S-BERT embedding exceeds the ComParE features across the majority of classifiers. Moreover the difference between best balanced accuracies from both segments (0.571 and 0.520) is statistically significant according to the McNemar test. (McNemar 1947)

When merging selected ComParE features and S-BERT embedding together, we observed a minor raise of balanced accuracy of SVM with linear kernel (0.579) in contrast to its linguistic-only variant (0.571) and the results of this fusion are overall 1-2% better to the fusion of S-BERT and wav2vec embeddings. Perhaps the low dimension of the selected ComParE features brings auxiliary distilled information, whereas the concatenation of S-BERT and wav2vec embeddings make the high-dimensional data noisy.

In summary, even though our classification results only negligibly surpass the chance level, these numbers are not too different from Hirschberg et al. (2005), introducing the corpus. Moreover, our task is defined as strictly speaker-independent, which is desirable for future work on automated deception detection, and this research can be one of the steps on its journey.

Acknowledgement

This work was supported by project SGS-2025-011.

References

Chen, X., Ita Levitan, S., Levine, M., Mandic, M. & Hirschberg, J. (2020), 'Acoustic-prosodic and lexical cues to deception and trust: Deciphering how people detect lies', *Transactions of the Association for Computational Linguistics* **8**, 199–214.

Hirschberg, J., Benus, S., Brenier, J. M., Enos, F., Friedman, S., Gilman, S., Girand, C., Graciarena, M., Kathol, A., Michaelis, L., Pellom, B. L., Shriberg, E. & Stolcke, A. (2005), Distinguishing deceptive from non-deceptive speech, *in* 'Interspeech 2005', pp. 1833–1836.

McNemar, Q. (1947), 'Note on the sampling error of the difference between correlated proportions or percentages', *Psychometrika* **12**(2), 153–157.

Weninger, F., Eyben, F., Schuller, B. W., Mortillaro, M. & Scherer, K. R. (2013), 'On the acoustics of emotion in audio: What speech, music, and sound have in common', *Frontiers in Psychology* **4**, 292.

Identifikace specifických bodů plochy plasticity tvářených plechů

Jakub Šlegl[1]

V současné době je vývoj konstrukčních dílů v automobilovém průmyslu stále více ovlivňován požadavkem na kombinaci nízké hmotnosti, vysoké pevnosti a dostatečné tvářitelnosti. Tento trend vede k intenzivnímu využití vysokopevnostních ocelí, zejména materiálů typu AHSS, jejichž mechanické chování se však vyznačuje výraznou komplexností. Zatímco zvýšená pevnost přispívá ke zlepšení bezpečnostních parametrů konstrukcí, současně dochází ke snížení tvářitelnosti a zvýšení citlivosti materiálu na lokální nestability, jako je ztenčování, vznik trhlin nebo tzv. „spring back“ efekt. Klíčovým předpokladem pro zvládnutí těchto jevů je detailní znalost plastického chování materiálu. Jedním ze základních nástrojů pro popis plastické odezvy kovových materiálů je plocha plasticity, která reprezentuje hranici mezi elastickým a plastickým stavem v prostoru napětí. Přesná identifikace jejího tvaru zásadně ovlivňuje kvalitu numerických simulací tvářecích procesů, které jsou dnes nedílnou součástí konstrukčního návrhu. Přestože existuje celá řada teoretických modelů, jejich spolehlivá aplikace v praxi je podmíněna kvalitní experimentální kalibrací. Právě tato skutečnost tvoří hlavní motivaci této práce.

Cílem práce je ve spolupráci s firmou COMTES FHT a.s. experimentální určení a numerické ověření bodů na ploše plasticity vysokopevnostní oceli DP1000, a to na základě kombinace několika typů mechanických zkoušek, které umožňují dosažení odlišných stavů napjatosti. Pozornost je věnována třem klíčovým experimentálním metodám – hydraulickému bulge testu vyhodnocený metodou ARGUS (Hallfeldt et al. (2014) a Šlegl (2023)), stack testu (Graser et al. (2019) a Merklein a Kuppertt (2009)) a plane strain testu (Flores et al. (2010) a Min et al. (2022)). Zatímco bulge test a stack test poskytují informace odpovídající přibližně dvojosému tahovému stavu, plane strain test představuje zásadní mezičlánek mezi jednoosým a dvojosým zatěžováním. Právě oblast rovinné deformace je z pohledu materiálového modelování mimořádně citlivá, neboť v ní dochází k výraznému ovlivnění výsledků volbou konstitutivního modelu i experimentální konfigurace. V rámci práce je proto detailně analyzována schopnost různých geometrií zkušebních vzorků dosáhnout stabilního stavu rovinné deformace. Ukazuje se, že dosažení tohoto stavu není samozřejmé a že i relativně malé geometrické nebo technologické odchylky mohou zásadně ovlivnit výsledné chování materiálu. Kombinace optických měřicích metod založených na digitální korelaci obrazu a numerických simulací metodou konečných prvků zde poskytuje silný nástroj pro potvrzení experimentálních předpokladů.

Významným přínosem práce je systematické porovnání experimentálních a numerických výsledků, které umožňuje nejen identifikovat vhodné experimentální postupy, ale také kvantifikovat jejich spolehlivost. Na základě tohoto přístupu jsou vybrány reprezentativní výsledky jednotlivých zkoušek a převedeny do jednotného napěťového prostoru. Takto získané body jsou následně použity pro konstrukci experimentální plochy plasticity a její porovnání s teoretickým popisem podle von Misesova kritéria. Získané výsledky ukazují, že při vhodně zvoleném experimentálním a numerickém přístupu lze dosáhnout velmi dobré shody mezi experimentálně stanovenými body a teoretickou plochou plasticity. Současně však práce

---

[1] student navazujícího studijního programu Aplikovaná mechanika, specializace Výpočty a design konstrukcí, e-mail: slegja@students.zcu.cz

odhaluje limity jednotlivých metod a zdůrazňuje nutnost jejich kritického posuzování, zejména oblasti rovinné deformace. Důležitým závěrem je skutečnost, že správně navržený plane strain test představuje klíčový experiment pro zpřesnění materiálového modelu, avšak jeho realizace vyžaduje pečlivou optimalizaci geometrie i experimentálních podmínek. Popsaný přístup zároveň ukazuje význam provázání experimentu a numerického modelování, které se v tomto případě neomezuje pouze na verifikaci výsledků, ale stává se aktivní součástí jejich interpretace. Práce tak přispívá k lepšímu porozumění chování vysokopevnostních oceli DP1000 v různých stavech napjatosti a poskytuje podklad pro přesnější a spolehlivější predikci odezvy tohoto materiálu v podmínkách reálných tvářecích procesů.

Poděkování

Rád bych poděkoval firmě COMTES FHT a.s. za umožnění provedení veškerých experimentálních testů a zkoušek, které byly využity při vzniku této práce.

Literatura

Flores, P., Tuninetti, V., Gilles, G., Gonry, P., Duchęne, L (2010) *Accurate stress computation in plane strain tensile tests for sheet metal using experimental data*. Journal of Materials Processing Technology.

Graser, M., Lenzen, M., Merklein, M. (2019) *On the inverse identification of Lankford coefficients using geometrical changes under quasi-biaxial loading.* International Journal of Material Forming.

Hallfeldt, T., Hotz, W., Leppin, S., et al (2014) *Sheet bulge testing.* Elsevier.

Merklein, M., Kuppertt, A. (2009) *A method for the layer compression test considering the anisotropic material behavior.* International Journal of Material Forming.

Min, J., Kong, J., Hou, Y., Liu, Z., Lin, J. (2022) *Application of Laser Deposition to Mechanical Characterization of Advanced High Strength Steels Subject to Non-Proportional Loading.* Experimental Mechanics

Šlegl, J. (2023) *Přípravek pro pokročilé testování plechů – Bulge test.* Plzeň.

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

LACA: Improving Cross-lingual Aspect-Based Sentiment Analysis with LLM Data Augmentation

Jakub Šmíd[1]

1 Introduction

Aspect-based sentiment analysis (ABSA) is a natural language processing task that aims to identify sentiment towards specific aspects of a product or service. For example, in the sentence *"Delicious food"*, the aspect term *"food"* expresses positive sentiment. Cross-lingual ABSA transfers knowledge from a source language to a target language. In this work, we focus on the zero-shot cross-lingual setting, where the model is fine-tuned only on source-language data and has no access to labelled target-language examples. To improve adaptation to the target language, we leverage a large language model (LLM) to generate pseudo-labelled target-language data, in contrast to traditional approaches that rely on machine translation.

2 Proposed Solution

We propose the LACA framework (Figure 1), which leverages unlabelled target-language data to improve zero-shot cross-lingual ABSA. The framework first fine-tunes an ABSA model, such as XLM-R, on labelled source-language data. The trained model then predicts labels for unlabelled target-language sentences. Instead of directly using these noisy predictions, we prompt an LLM (e.g. LLaMA-3.1-70B) with the predicted labels to generate new target-language sentences aligned with the predicted sentiment and aspect information. The generated sentences are subsequently paired with the predicted labels to form a pseudo-labelled target-language dataset, which is combined with the original source-language data for final training.

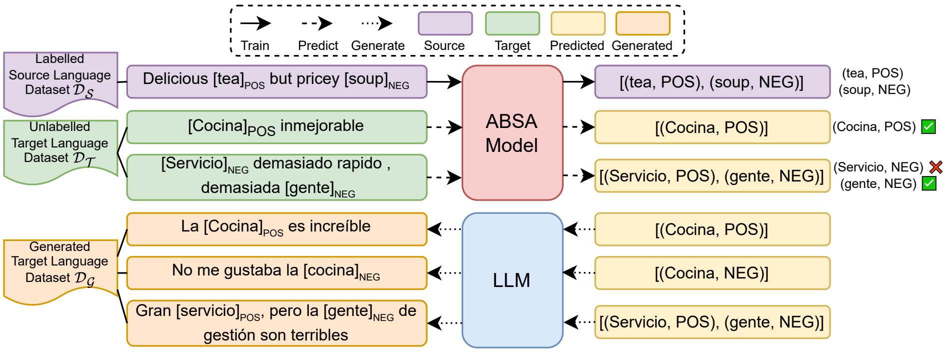


Figure 1: Proposed LACA framework.

Crucially, instead of translating or directly reusing predictions, LACA conditions LLM generation on model-predicted labels, enabling controlled synthesis of target-language training data aligned with the ABSA task structure. In contrast to translation-based approaches that produce semantically similar target-language counterparts of source sentences, our method gener-

---

[1] student of the doctoral degree program Computer Science and Engineering, e-mail: jaksmid@fav.zcu.cz

ates diverse target-language samples conditioned on predicted labels. This improves robustness and generalization by exposing the model to a wider range of expressions and linguistic structures, while also allowing controlled data generation to mitigate label imbalance by synthesizing additional examples for underrepresented sentiment classes.

3 Results

We evaluate our approach on a dataset covering six languages, using English as the source language and the remaining languages as targets. Our primary experiments employ XLM-R as the ABSA backbone and LLaMA-3.1-70B for target-language data generation. Table 1 compares the proposed LACA framework against the MONOLINGUAL baseline (trained directly on labelled target-language data), the ZERO-SHOT baseline (trained only on English data), and the previous state-of-the-art translation-based method (Zhang et al., 2021). The proposed approach outperforms the previous state of the art by more than 2.5% on average across the evaluated target languages, while remaining within 1% of the fully supervised monolingual setting on average. Additional experiments with Turkish and multiple backbone models, including fine-tuned LLMs, demonstrate that the improvements generalize consistently across architectures and can yield even stronger performance gains. We also show that smaller LLMs (e.g. 8B parameters) achieve comparable performance to larger models, indicating robustness and efficiency of the proposed generation strategy.

Method	**Es**	**Fr**	**Nl**	**Ru**	**Avg**
MONOLINGUAL	71.93	67.44	64.28	64.93	67.15
ZERO-SHOT	67.48	58.87	58.95	56.10	60.35
(Zhang et al., 2021)	69.24	59.90	63.74	62.02	63.73
LACA	**71.89**	**64.97**	**65.35**	**63.20**	**66.35**

Table 1: F1 results with XLM-R, with the best cross-lingual scores highlighted in **bold**.

4 Conclusion

We introduce the LACA framework to enhance zero-shot cross-lingual ABSA by leveraging a large language model to generate high-quality pseudo-labelled target-language data conditioned on predictions from an ABSA model. The proposed approach establishes new state-of-the-art results, outperforming translation-based methods, and demonstrates consistent effectiveness across multiple target languages and backbone models, including strong gains with different architectures and LLM-based variants.

Acknowledgement

This work has been supported by the Grant No. SGS-2025-022 – New Data Processing Methods in Current Areas of Computer Science. Computational resources were provided by the e-INFRA CZ project (ID:90254), supported by the Ministry of Education, Youth and Sports of the Czech Republic.

References

Zhang, W., He, R., Peng, H., Bing, L., & Lam, W. (2021, November). Cross-lingual Aspect-based Sentiment Analysis with Aspect Term Code-Switching. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (pp. 9220-9230).

FACULTY OF APPLIED SCIENCES
UNIVERSITY
OF WEST BOHEMIA

STUDENT
SCIENTIFIC CONFERENCE 2026

Towards the Largest Czech Sign Language Translation Dataset in the World

Tomas Zelezny[1]

1 Introduction

Sign languages are rich visual-spatial natural languages used by Deaf communities. They convey information through hand articulation, body pose, facial expressions and non-manual markers. Sign language processing includes tasks such as Isolated Sign Language Recognition (ISLR), Continuous Sign Language Recognition (CSLR), and Sign Language Translation (SLT). Among these, SLT is considered the most challenging because it combines complex visual understanding with natural language translation.

The main limitation of modern SLT research is the lack of large-scale datasets. While high-resource sign languages such as ASL already contain thousands of hours of annotated data, Czech Sign Language (CzSL) remains severely underrepresented. This work presents an ongoing effort towards constructing a large-scale CzSL Translation dataset from multiple heterogeneous data resources.

2 Czech Sign Language Resources

The proposed dataset combines multiple Czech sign language resources with different modalities, annotation quality and task suitability. Table 1 summarizes currently identified datasets and data sources. Highlighted rows denote the primary Czech Sign Language resources newly integrated and analyzed within this work.

Existing Czech datasets are mostly focused on isolated sign recognition or motion-capture acquisition. While such datasets remain valuable for auxiliary tasks and multimodal experiments, they are insufficient for large-scale gloss-free SLT training.

The most promising sources for continuous SLT are interpreted broadcasts, educational materials and online media. However, the collected resources differ substantially in annotation granularity, subtitle quality, synchronization accuracy, visual quality and domain specificity. Some datasets provide sentence-level subtitles, while others contain only paragraph-level annotations requiring further segmentation and alignment.

These differences significantly affect the suitability of individual sources for machine learning training, pretraining and evaluation.

3 Dataset Construction Challenges

Before the collected resources can be used for SLT training, each source must be processed individually. The main challenges include:

- temporal misalignment between signing and subtitles or transcripts,

---

[1] Student of the doctoral degree program Cybernetics, field of study Computer Vision,
e-mail: zeleznyt@fav.zcu.cz

Source	Task	Size	Annotation Quality
UWB-06-SLR-A[†]	ISLR	5685 clips	precise
UWB-07-SLR-P[†]	ISLR	21k clips	precise
Znaky.zcu.cz[†]	ISLR	1000+ clips	precise
SL MoCap Dataset[†]	ISLR/CSLR/SLT	30 min	precise
MC-TRISLAN[†]	SLT	18 h	precise
YouTubeSL-25 Cz	SLT	12 h	weak subtitles
JW Sign Cz	SLT	17.5 h	sentence-level
Dictio	ISLR	30k clips	precise
Hybrid books	SLT	65 h	paragraph-level
PSP broadcasts	SLT	14 h (growing)	weak subtitles
Česká televize	SLT	hundreds of hours	mixed
Minirozhovor ČUN	SLT	1 - 5 h	precise
Národní technické muzeum	SLT	< 1 h	mixed
MTDRadlice	SLT	15 - 30 h	mixed

[†] Product of the University of West Bohemia in Pilsen.

Table 1: Overview of Czech Sign Language resources considered for dataset construction. Highlighted rows denote resources currently integrated within this work.

- paragraph-level annotations that must be split into sentence-level segments,
- different video quality, framing and signer size,
- differences between interpreted, lexical and natural data,
- domain-specific bias, e.g. religious or political language,
- different availability of subtitles, OCR text or audio transcripts,
- licensing and privacy constraints.

These factors significantly affect the usability of individual resources for sign language translation training and evaluation.

4 Conclusion

This work presents an ongoing effort towards construction of a large-scale Czech Sign Language Translation dataset from heterogeneous Czech sign language resources. The collected sources differ substantially in annotation quality, synchronization accuracy, domain, visual properties and linguistic style. As a result, each dataset requires an individual preprocessing and alignment strategy before it can be used for SLT training.

The analysis also suggests that different resources may be suitable for different stages of development. Weakly aligned large-scale corpora may still provide valuable pretraining data, while carefully synchronized sentence-level resources are necessary for fine-tuning and evaluation. Similarly, isolated-sign datasets are not directly suitable for continuous SLT training, but may remain useful for auxiliary tasks or visual encoder pretraining.

Future work will focus on data acquisition, temporal realignment, sentence segmentation and construction of reliable development and test sets for Czech Sign Language Translation.

Acknowledgement

The work has been supported by the grant of the University of West Bohemia, project No. SGS-2025-011.